

Výňatek (obsah a teoretický úvod) z knihy

Markéta Lopatková, Zdeněk Žabokrtský, Václava Kettnerová a kol.:  
Valenční slovník českých sloves. Univerzita Karlova v Praze,  
Nakladatelství Karolinum, Praha, 2008.



# Obsah

<b>Úvod</b>	<b>7</b>
<b>Podoba slovníkového hesla</b>	<b>11</b>
<b>Struktura slovníku VALLEX</b>	<b>13</b>
1 Lexémy, lexikální formy a lexikální jednotky . . . . .	13
2 Lexikální formy a lemmata . . . . .	13
2.1 Reflexivní lemmata . . . . .	14
2.2 Vid a vidové protějšky . . . . .	14
2.3 Varianty lemmatu . . . . .	15
2.4 Homografy (homonyma) . . . . .	16
3 Lexikální jednotky . . . . .	16
4 Valenční rámce . . . . .	18
4.1 Aktanty a volná doplnění . . . . .	18
4.2 Morfematické vyjádření . . . . .	21
4.3 Atribut obligatornosti slovesného doplnění . . . . .	23
4.4 Expanze valenční pozice . . . . .	23
5 Doplnující syntaktické informace . . . . .	24
5.1 Kontrola . . . . .	24
5.2 Reflexivita . . . . .	25
5.3 Reciprocita . . . . .	25
5.4 Syntakticko-sémantické třídy . . . . .	26
5.5 Frazémy a idiomy . . . . .	27
<b>Seznam značek</b>	<b>29</b>
<b>Seznam literatury</b>	<b>31</b>
<b>Slovník</b>	<b>35</b>



# Úvod

Valence patří k jazykovým jevům, které vzbuzují zájem lingvistů, ale i odborníků zabývajících se automatickým zpracováním přirozených jazyků. Její dobrá znalost je totiž nezbytná pro každého, kdo pracuje s jazykem – pro učitele, redaktory, novináře, studenty apod. Valence zároveň hraje klíčovou roli při úlohách automatického zpracování přirozeného jazyka, jakými jsou např. strojový překlad, rozlišování významů slov či vyhledávání informací.

## Co je to tedy valence?

Termínu valence známého z oblasti chemie, kde označuje mocenství atomu, v lingvistickém kontextu poprvé užil uprostřed minulého století francouzský syntaktik Lucien Tesnière. Metaforicky tímto termínem označil schopnost slovesa (podobnou vlastnosti atomů) vázat k sobě určitý počet jazykových elementů.

Pod pojmem valence se rozumí „počet a povaha míst (argumentů), které na sebe dané sloveso (popř. slovo jiného slovního druhu) váže“ (*Encyklopedický slovník češtiny*, Karlík et al., 2002). Je to tedy schopnost slova vázat na sebe určitý počet jiných, syntakticky závislých jazykových jednotek. Tato schopnost se primárně týká významové roviny jazyka, tedy hloubkové větné stavby. Valenční pozice jsou naplňovány valenčními doplněními, jako je aktor (konatel či nositel děje), pacient (zasažený objekt), adresát, původ a výsledek děje, označovanými obvykle jako aktanty, ale i volnými doplněními vyjadřujícími okolnosti děje, jako jsou způsob děje, místo, směr apod.

Jednotlivé valenční pozice jsou různě významově těsné. Neobsazením některých pozic dochází k porušení významové úplnosti, což může vést i k porušení gramatické správnosti věty, srov. např. nepřijatelné věty *\*Petr dává*, *\*Marie nenávidí*, *\*Jan se choval*. Doplnění obsazující takové valenční pozice se nazývají obligatorní. Zvláštní pozornost je v teorii valence věnována případům, kdy obligatorní valenční pozice zůstávají v povrchové podobě věty neobsazeny a posluchač/čtenář si příslušné pozice zaplňuje z kontextu promluvy/textu, např. ve spojeních *obvykle nakupuje v supermarketu* či *děti přišly* není vyjádřen pacient, resp. směr, tyto informace by měly být posluchači/čtenáři zřejmé z kontextu promluvy/textu.

Jiné valenční pozice jsou nepovinné, fakultativní – jsou sice přítomny ve valenčním rámci a mohou být přítomny ve významové (hloubkové) reprezentaci věty, jejich neobsazením však nevzniká významově ani gramaticky narušená věta, např. *Petr se pevně držel (zábradlí)*, *Eva se najedla (ovoce)*, *Dívka píše (mamince) dopis*. Ostatní pozice s velmi volným vztahem ke slovesu se často označují jako nevalenční, např. *Jana se procházela (po lese)*, *Petr se budil (časně)*, *Eva si četla (pro své potěšení)*, i když v teoretickém popisu jsou do valence slovesa (v širším smyslu) zahrnuty.

Valence se promítá i do povrchové realizace věty. Jednotlivé aktanty typicky vlivem slovesa nabývají určité realizace – jejich morfemická podoba je určována rekcí řídicího slovesa. Např. konatel bývá v aktivní větě typicky vyjádřen nominativem, zatímco pacient se obvykle realizuje akuzativem (např. *Petr<sub>konatel</sub> ztratil botu<sub>pacient</sub>*), jiná slovesa vyžadují konatele v dativu (např. *Petrovi<sub>konatel</sub> se ve škole líbí*), opět jiná slovesa mají pacient v dativu (např. *rodiče<sub>konatel</sub> bránili jejich štěstí<sub>pacient</sub>*) či ve formě předložkové skupiny (např. *doufali ve vítězství<sub>pacient</sub>*). Naproti tomu forma valenčních volných doplnění bývá dána významem těchto doplnění (*děti přišly domů/do školy/na hřiště*) a nebývá řízena slovesem.

Valence je teoreticky zkoumána zhruba od poloviny 20. století, jmenujme zde již zmíněného L. Tesnière (viz Tesnière, 1959). Důraz na zkoumání významové složky valence pak přinesly především studie Fillmorovy (Fillmore, 1968, 1969). V českém prostředí se k významným pracím věnovaným valenci řadí zejména studie Danešovy (Daneš, 1971, Daneš – Hlavsa, 1987), práce Panevové (Panevová, 1974, 1980, 1994), Karlíkovy (Karlík, 2000) a Sgallovy (Sgall, 1998, 2006). Další odkazy lze nalézt v seznamu literatury na konci úvodní části.

## Valenční slovník českých sloves *VALLEX*

Je zřejmé, že valenční vlastnosti sloves (i dalších slovních druhů) jsou velmi rozmanité. Nelze je odvodit obecnými pravidly, je třeba je popsat pro jednotlivé lexikální položky, tedy v podobě valenčního slovníku, který obsahuje popis valence jednoho slova po druhém. Z těchto důvodů vzniká v Ústavu formální a aplikované lingvistiky na Matematicko-fyzikální fakultě Univerzity Karlovy v Praze od roku 2001 elektronický Valenční slovník českých sloves *VALLEX*, jehož tištěnou verzi máte nyní v rukou.

Valenční slovník českých sloves (*VALLEX* jako *VALency LEXicon*) je výsledkem snah o formální zachycení valenčních charakteristik českých sloves v rámci teorie Funkčního generativního popisu (FGP), která byla navržena Petrem Sgallem (viz Sgall, 1967) a je dále rozvíjena jím a jeho spolupracovníky (viz zejména Sgall et al., 1986). Teorii valence ve FGP propracovala zejména Jarmila Panevová (Panevová, 1974, 1980, 1994). Slovník *VALLEX* je úzce spojen s Pražským závislostním korpusem (PDT, Prague Dependency Treebank, viz Hajič, 2005, Hajič et al., 2006). Zatímco v PDT je FGP verifikován pomocí komplexní analýzy velkého množství textových dat, ve *VALLEXu* je využita zejména valenční teorie vyvíjená v rámci FGP.

*VALLEX* poskytuje informace o valenční struktuře českých sloves v jejich jednotlivých významech, které charakterizuje pomocí glos a příkladů. Pro jednotlivá valenční doplnění uvádí *VALLEX* možná morfemata vyjádření, pokud jsou jejich formy dány slovesnou rekcí. Kromě těchto základních údajů uvádí i některé další syntaktické, případně syntakticko-sémantické charakteristiky, jako je vlastnost kontroly, možnost recipročního užití či syntakticko-sémantická třída slovesa.

*VALLEX* popisuje chování 2 730 českých lexémů, které zahrnují 6 460 lexikálních jednotek – ‚daných sloves v daném významu‘. Zdůrazněme zde, že *VALLEX* (v návaznosti na teoretické předpoklady FGP) zachycuje valenční chování pravých vidových protějšků v rámci jediného lexému, jednotlivá hesla ve slovníku tedy odpovídají celým lexémům. Pokud bychom počítali dokonavá a nedokonavá slovesa zvlášť, dostali bychom se k počtu 4 250 sloves. Hlavním kritériem výběru sloves ve *VALLEXu* byla jejich frekvence v Českém národním korpusu SYN2000<sup>1</sup> – v prvním kroku bylo vybráno přibližně 2 500 nejčastějších slovesných lemmat, posléze byl tento výběr doplněn tak, aby slovník obsahoval ke každému slovesu i jeho vidové protějšky.

Při navrhování koncepce *VALLEXu* byl kladen důraz na přesnost a lingvistickou adekvátnost popisu valence u velkého množství sloves. Slovníková hesla byla zpracována manuálně s přihlédnutím ke korpusovému i slovníkovému materiálu s následnou rozsáhlou automatickou, poloautomatickou i ruční kontrolou. Od samého začátku se předpokládalo využití *VALLEXu* jak pro člověka jako uživatele jazyka, tak pro počítačové zpracování češtiny a pro další aplikační účely, jako např. strojový překlad, vyhledávání v textech apod. Kromě tištěné verze *VALLEXu* byla připravena též jeho elektronická podoba ve formě webové stránky umožňující pohodlné vyhledávání podle různých kritérií: lemma, syntakticko-sémantická třída, funktor, formy doplnění a další. Tento elektronický slovník i jeho verzi ve formátu XML vhodném pro automatické zpracování (viz zejména Žabokrtský, 2005) lze nalézt na adrese Ústavu formální a aplikované lingvistiky Matematicko-fyzikální fakulty Univerzity Karlovy v Praze.<sup>2</sup>

Důrazem na možnost všestranného využití se *VALLEX* liší od jiných dostupných valenčních slovníků češtiny – jmenujme zejména velmi propracovaný tištěný slovník *Slovesa pro praxi* (Svozilová et al., 1997), který popisuje 767 českých sloves. Ze slovníků pro automatické zpracování jazyka uveďme slovník *BRIEF* (Pala – Ševeček, 1997), který však neobsahuje popis významové struktury, a *Czech syntactic lexicon* (Skoumalová, 2001), který vznikl ze slovníku *BRIEF* automatickým doplněním významové struktury (ani jeden z těchto slovníků není veřejně dostupný). Z dalších významných českých slovníků citujme ještě tištěný *Slovník slovesných, substantivních a adjektivních vazeb a spojení* (Svozilová et al., 2005), elektronický valenční slovník *PDT-VALLEX* (Hajič et al., 2003), zachycující valenci sloves a vybraných

<sup>1</sup> <http://ucnk.ff.cuni.cz>

<sup>2</sup> <http://ufal.mff.cuni.cz/vallex/2.5> (*VALLEX*, verze 2.5)

deverbativních substantiv a adjektiv v PDT (Hajič, 2005, Hajič et al., 2006), a zatím nepublikovaný slovník *Verbalex* (Hlaváčková – Horák, 2006).

Dodejme, že budování valenčních slovníků jak v tištěné, tak v elektronické podobě stojí v centru pozornosti i pro desítky dalších jazyků; odkazy na nejdůležitější slovníky jsou uvedeny v seznamu literatury.

Příprava současné verze slovníku *VALLEX* trvala více než pět let. Jejím primárním cílem bylo vytvořit kvalitní veřejně dostupný valenční slovník, který by vyhovoval lingvistům i odborné veřejnosti na jedné straně a který by na straně druhé mohl sloužit pro počítačové zpracování češtiny. Doufáme, že tohoto cíle bylo dosaženo. *VALLEX* je formálně strukturovaný slovník – jeho tištěnou verzi nyní držíte v ruce a jeho elektronická verze je dostupná na adrese Ústavu formální a aplikované lingvistiky Matematicko-fyzikální fakulty Univerzity Karlovy v Praze. Věříme dále, že snaha o shromáždění syntaktických informací vztahujících se k valenci – stejně jako důraz na konzistentní popis systémových vlastností lexikálních jednotek – činí z *VALLEXu* praktickou pomůcku pro účely didaktické, překladatelské a pro práci s češtinou vůbec, ale i užitečný zdroj dat pro účely aplikační.

Výsledky uložené v tomto slovníku prezentují valenci jako problém jak syntaktický (kombinatorický), tak lexikografický. Práce s bohatými daty ovšem ukazuje, že slovníkový přístup k zachycení valence přináší stále otevřené teoretické otázky, které vyžadují další podrobné lingvistické zkoumání.

## Poděkování

Při zpracovávání sloves pro slovník *VALLEX* jsme využívali následující slovníky (některé z nich přes webové rozhraní DEBDict)<sup>3</sup> a korpusy českých textů:

- BRIEF (Pala – Ševeček, 1997);
- Slovník spisovné češtiny pro školu a veřejnost (Mejstřík, 2003, dále SSČ);
- Slovník spisovného jazyka českého (Havránek, 1964, dále SSJČ);
- Slovesa pro praxi (Svozilová et al., 1997);
- Slovník slovesných, substantivních a adjektivních vazeb a spojení (Svozilová et al., 2005);
- Český národní korpus (Blatná – Čermák, 2005, dále ČNK)<sup>4</sup>;
- Pražský závislostní korpus (Hajič et al., 2006, dále PDT)<sup>5</sup>.

Rádi bychom zde poděkovali za podporu a za rozsáhlé lingvistické i technologické rady a pomoc kolegům z Ústavu formální a aplikované lingvistiky. Zvláštní dík patří zejména paní profesorce Jarmile Panevové.

Práce na slovníku *VALLEX*, verze 2.5, probíhala v rámci výzkumného záměru Ministerstva školství, mládeže a tělovýchovy ČR č. MSM0021620838 (Moderní metody, struktury a systémy informatiky). Dále byla částečně podporována grantem programu MŠMT ČR Výzkumná centra, č. LC536 (Centrum počítační lingvistiky) a dvěma granty programu Informační společnost číslo 1ET100300517 a 1ET101120503.

<sup>3</sup> <http://nlp.fi.muni.cz/projekty/deb2/debdict/index.html>

<sup>4</sup> <http://ucnk.ff.cuni.cz>

<sup>5</sup> <http://ufal.mff.cuni.cz/pdt2.0/>





## Podoba slovníkového hesla

Zde uvádíme pouze přehled struktury hesla pro usnadnění orientace (viz též obrázek na následující straně), jednotlivé pojmy jsou vysvětleny v následujícím oddíle.

**Lemma** – v záhlaví slovníkového hesla je uvedeno lemma (infinitivní tvar) reprezentující heslové sloveso, příp. seznam lemmat (v pořadí nedokonavé, dokonavé, iterativum), odd. 2.

**Vid** – jako horní index je u každého lemmatu uveden údaj o vidu (značky *impf*, *pf*, *iter*; případně následovány arabskou číslicí, pokud je slovníkové heslo reprezentováno více lemmaty se stejným videm), odd. 2.2.

**Varianty** – pokud má slovesné lemma varianty, jsou všechny varianty (oddělené lomítkem) uvedeny v záhlaví hesla, odd. 2.3.

**Homografy** – jsou rozlišeny římskou číslicí v dolním indexu u lemmatu, odd. 2.4.

**Číslo LU** – pokud má sloveso více významů, tzn. má více lexikálních jednotek (dále LU), je každá z nich označena arabskou číslicí, odd. 3.

**Idiom** – idiomatická užití jsou uvedena značkou *idiom* za číslem LU, odd. 5.5.

**Omezení** – pokud se daná LU vztahuje jen k některým lemmatům ze seznamu uvedeného v záhlaví, jsou za značkou *jen* uvedena všechna lemmata, ke kterým se tato LU vztahuje (s vyznačeným videm a případnými variantami a indexem pro homografy); omezení se neuvádí pro iterativa.

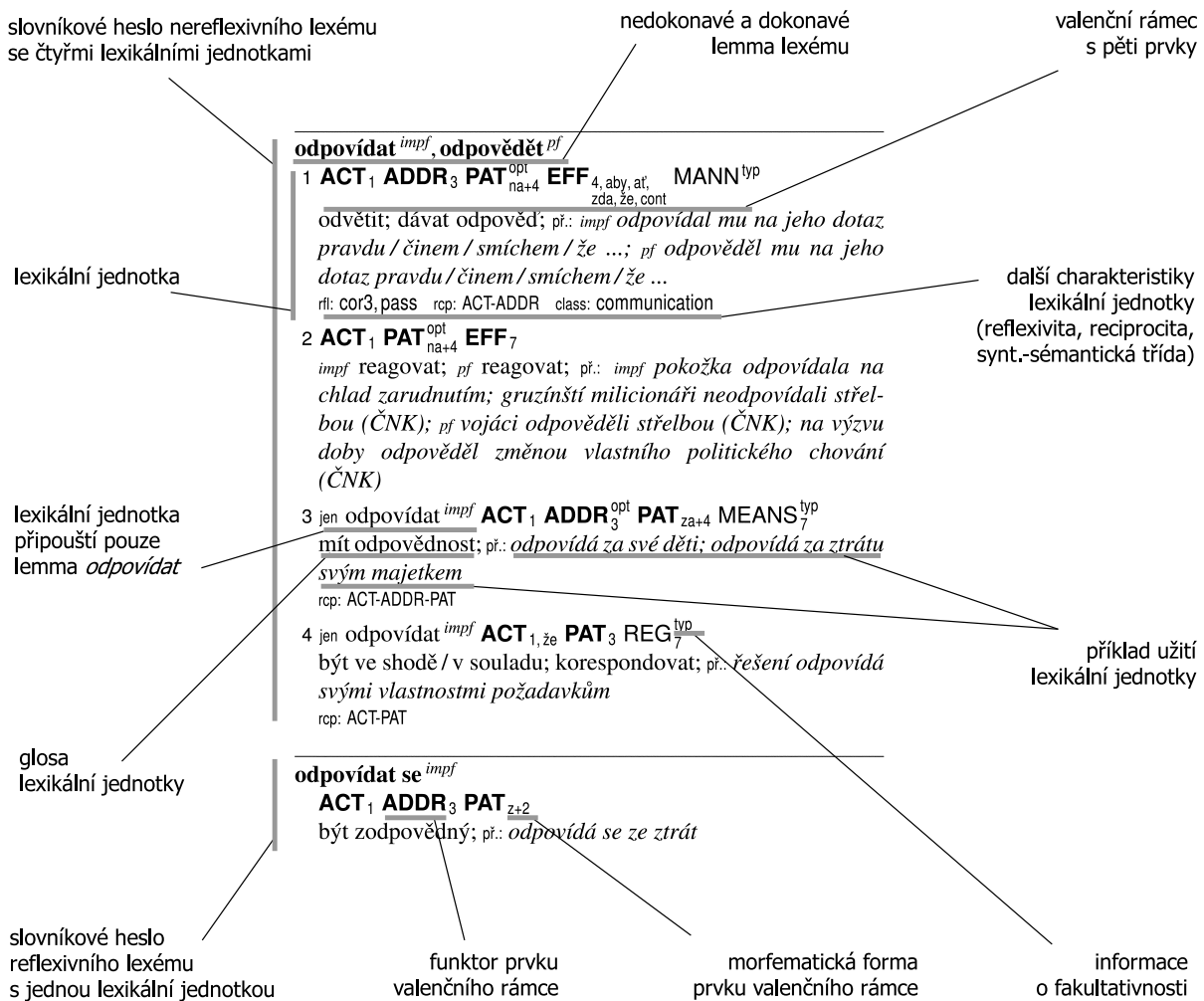
**Valenční rámec** – každá LU je popsána pomocí formálního zápisu rámce, který uvádí počet a typ (tzv. funktor) valenčních doplnění, jejich možná morfemická vyjádření (dolní index) a obligatornost (horní index – pokud není uveden, jde o obligatorní doplnění, značka *opt* zachycuje fakultativní doplnění, značka *typ* doplnění typické), odd. 4.

**Glosa** – každá LU je charakterizována glosou, která je na novém řádku za valenčním rámcem; pokud je heslové sloveso specifikováno několika lemmaty, jsou zde glosy pro všechna nedokonavá a dokonavá lemmata; glosy jsou uvedeny vždy údajem o vidu.

**Příklad** – každá LU obsahuje příklad užití uvedený značkou *př.*; pokud je heslové sloveso specifikováno několika lemmaty, jsou zde příklady pro všechna nedokonavá a dokonavá lemmata; příklady jsou vždy uvedeny údajem o vidu.

### Doplňující syntaktické informace

- **kontrola** – slovesa kontroly mají za značkou *control* uveden funktor členu valenčního rámce tohoto slovesa, který koreferuje se subjektem infinitivu závislého slovesa (kontroluje ho), odd. 5.1;
- **reflexivita** – za značkou *ri* jsou uvedeny možné syntaktické funkce morfému *se/si*, odd. 5.2;
- **reciprocita** – za značkou *rcp* jsou uvedeny dvojice, příp. trojice valenčních doplnění, která mohou vstupovat do vztahu reciprocity, odd. 5.3;
- **syntakticko-sémantická třída** – za značkou *class* je uvedena syntakticko-sémantická třída slovesa v daném významu, odd. 5.4.



### Ukázka dvou slovníkových hesel

# Struktura slovníku VALLEX

Obsah *Valenčního slovníku českých sloves VALLEX* zde popíšeme pouze z hlediska jeho struktury. Lingvistické aspekty vyžadující širší vysvětlení či diskusi jsou většinou ponechány stranou, odkazujeme pouze na základní literaturu týkající se dané problematiky.

Odborná terminologie, kterou v textu užíváme, buď patří k ustálené lingvistické terminologii, nebo vychází z terminologie Funkčního generativního popisu (FGP), který slouží jako podkladová teorie *VALLEXu*. Pokud tomu tak není, jsou potřebné termíny zavedeny v textu.

## 1 Lexémy, lexikální formy a lexikální jednotky

Slovník *VALLEX* je na nejvyšší úrovni tvořen lexémy, kterým odpovídají jednotlivá slovníková hesla. *Lexémem* přitom rozumíme abstraktní jednotku, která v sobě spojuje formální i významovou složku (viz Cruse, 1986), u polysémních/polysémických/víceznačných lexémů též označovanou jako hyperlexém (viz i Filipec – Čermák, 1985, Karlík et al., 2002). Lexém sdružuje množinu všech možných manifestací slovesa v textu/řeči, kterou zde budeme označovat jako množinu všech možných *lexikálních forem*, a množinu *lexikálních jednotek* (LU jako ‚lexical unit‘), které reprezentují jeho významové složky, v české terminologii obvykle označované jako jednotlivé lexie či základní lexikální jednotky, viz terminologickou poznámku níže. Lexikální formy jsou tedy všechny morfematické formy slovesa (celé jeho paradigma), viz odd. 2, zatímco lexikální jednotky zhruba odpovídají lexému v jednom určitém významu a nesou informaci o syntaktických a syntakticko-sémantických rysech slovesa v tomto významu, viz odd. 3.

*Terminologická poznámka:* Zde se odchyľujeme od (ne zcela jednotné) české terminologie, kde se obvykle termíny lexém a lexikální jednotka užívají víceméně synonymně, viz Filipec – Čermák (1985), str. 28: „Tento termín [lexém] je tedy synonymní s termínem LJ [= lexikální jednotka], ale je ještě dále diferencován.“ Každému významu (polysémického) lexému odpovídá jedna základní lexikální jednotka (tamtéž): „Lexikální jednotka jako polysémický lexém je ... útvar zahrnující tolik různých monosémických základních lexikálních jednotek, lexíí (...), kolik má různých významů.“

Ve *VALLEXu* se přikláníme k terminologii, která je běžná v anglicky psané odborné literatuře, kde je termínem ‚lexical unit‘ označován koncept odpovídající základní lexikální jednotce (viz Cruse, 1986, Ruppenhofer et al., 2006).

## 2 Lexikální formy a lemmata

Množiny všech možných lexikálních forem tvořících formální složku jednotlivých lexémů jsou reprezentovány infinitivními tvary slovesa obvykle nazývanými *lemmata*.

Slovníkové heslo je ve *VALLEXu* uvedeno lemmatem, případně seznamem lemmat vztahujících se k danému lexému (včetně příp. morfému *se/si*, viz odd. 2.1). Jednotlivá lemmata jsou doplněna o další informace:

- informace o vidu v horním indexu, viz odd. 2.2;
- římská číslice v dolním indexu rozlišující homografy (viz odd. 2.4).

Slovník *VALLEX* (v návaznosti na teorii FGP) zachycuje valenční chování vidových protějšků v rámci jediného lexému. Proto je slovníkové heslo typicky uvedeno dvěma (příp. více) slovesnými lemmaty, nedokonavým a dokonavým infinitivním tvarem slovesa, viz odd. 2.2. Dalším důvodem, proč se v záhlaví slovníkového hesla může vyskytnout více lemmat, jsou pravopisné varianty slovesa, viz odd. 2.3.

## 2.1 Reflexivní lemmata

Z hlediska zachycení reflexivity ve slovníku jsou ve *VALLEXu* rozlišovány dva základní typy reflexivních konstrukcí (viz též Karlík et al., 2002):

- **Reflexivní lexémy.** Jako reflexivní lexémy jsou označována inherentně reflexivní slovesa, *reflexiva tantum*, u nichž je morfém *se/si* považován za součást lemmatu.

Řadí se k nim:

- primární reflexiva tantum (v Karlík et al., 2002 označovaná jako inherentně reflexivní slovesa), tedy slovesa, která se v nereflexivní formě vůbec nevyskytují (např. *bát se*, *smát se*), či slovesa, která nereflexivní formy sice mají, lexikální jednotky odpovídající reflexivním a nereflexivním formám si jsou však natolik významově vzdáleny, že jsou obvykle vyčleňovány do dvou lexémů (např. *chovat se* vs. *chovat*);
- tzv. odvozená/sekundární reflexiva (v Karlík et al., 2002 též inherentně reflexivní varianta slovesa), tedy slovesa, kde *se/si* je slovotvorně motivovaným morfémem, reflexivní forma je tudíž v nějakém významovém vztahu k nereflexivní formě, např. vyjadřuje samovolnou či bezděčnou činnost (např. *šířit se*, *vrátit se*).

Reflexivní lexémy jsou ve *VALLEXu* zachyceny v samostatných slovníkových heslech.

- **Reflexivní užití nereflexivních lexémů.** Pokud reflexivní morfém *se/si* nese syntaktickou funkci, jsou reflexivní formy sloves zachyceny v rámci nereflexivního lexému, kde je též specifikována jejich syntaktická funkce (viz oddíly 5.2 a 5.3):

- *se* může být součástí tvaru tzv. reflexivního pasiva (např. *pátrá se po zloději*);
- *se/si* může označovat doplnění obsazující valenční pozici řídicího slovesa u tzv. vlastního reflexiva (kde *se/si* lze nahradit silnou podobou zájmena *sebe/sobě*, např. *vidět se* (= sebe), *darovat si* (= sobě) *dort*, kde *se/si* je PAT (patient), resp. ADDR (adresát) koreferující s ACT (aktorem) řídicího slovesa);
- *se/si* může mít reciproční funkci (např. *kopat se v kopou se vzájemně do nohou*).

*Poznámka:* Některé lexikální jednotky mají reflexivní i nereflexivní podobu lemmatu beze změny významu, např. *myslím (si)*, *že to tak není* (někdy se označuje se jako volné *se/si*). Tento typ *se/si* bývá zaznamenán pouze jako jeden z příkladů u nereflexivní lexikální jednotky.

## 2.2 Vid a vidové protějšky

V češtině se pro kategorii vidu rozlišují dvě základní hodnoty, nedokonavost a dokonavost. Vedle toho se vyčleňují též iterativa jako specifická podtřída nedokonavých sloves a slovesa obouvidová (slovesa, která se v určitých kontextech chovají jako dokonavá, v jiných kontextech jako nedokonavá).

Ve *VALLEXu* jsou do jediného lexému spojeny tzv. *pravé vidové dvojice* tvořené sufixálně – nedokonavé sloveso je utvořeno od dokonavé formy slovesa např. příponou *-va(t)* (např. *ochutnat* → *ochutnávat*), příponou *-ova(t)* (např. *dokončit* → *dokončovat*), příponou *-a(t)* (např. *vyrůst* → *vyrůstat*) či příponou *-ě/e(t)* (např. *otočit* → *otáčet*). Dále jsou v jednom lexému zachyceny i supletivní páry (např. *vzít* – *brát*, *najít* – *nacházet*).<sup>6</sup> Pokud existuje i běžně užívaná iterativní podoba, je rovněž zahrnuta do příslušného lexému (např. slovesa *nasedat<sup>impf</sup>*, *nasednout<sup>pf</sup>* a *nasedávat<sup>iter</sup>* jsou popsány v jednom slovníkovém hesle odpovídajícím jednomu lexému).

<sup>6</sup> Dokonavé protějšky nedokonavých sloves tvořené prefixálně v rámci stejného lexému zachyceny nejsou. K tomuto rozhodnutí vedly praktické důvody, neboť ne vždy je zcela jednoznačné, které z více možných prefigovaných lemmat považovat za vidový protějšek.

Ve *VALLEXu* se informace o vidu zachycuje u každého lemmatu jako horní index, který může nabývat následujících hodnot:

- *impf* pro nedokonavá slovesa;
- *pf* pro dokonavá slovesa;
- *iter* pro iterativa (násobená slovesa);
- *biasp* pro obouvidová slovesa.

V rámci jediného slovníkového hesla je v některých případech zachyceno více podob jednoho ze členů vidové dvojice (aniž by šlo o varianty, viz odd. 2.3), a to v případech, kdy jednomu lemmatu jedné vidové hodnoty odpovídají různá lemmata s druhou hodnotou (např. dokonavá slovesa *dohonit<sup>pf</sup>* i *dohnat<sup>pf</sup>* mají nedokonavý protějšek *dohánět<sup>impf</sup>*, naopak dokonavé sloveso *odvinout<sup>pf</sup>* má dva nedokonavé protějšky *odvinovat<sup>impf</sup>* a *odvíjet<sup>impf</sup>* – každá tato trojice lemmat reprezentuje jediný lexém, odpovídá jí tedy jediné slovníkové heslo). Při stanovování vidových protějšků se *VALLEX* přidržuje vztahů stanovených ve slovníku SSJČ.

Typicky taková lemmata sdílejí (alespoň jednu) lexikální jednotku, viz odd. 3, i když u nich může docházet k modifikaci významu (spočívající zejména ve změně ‚způsobu slovesného děje‘, např. u některých sloves pohybu, u sloves distributivních či u sloves momentálních s příponou *-nou(t)*). Např. lemmata *odříznout<sup>pf</sup>* a *odřezat<sup>pf</sup>* jsou zahrnuta spolu s lemmatem *odřezávat<sup>impf</sup>* do jediného lexému.

*Poznámka k notaci:* V seznamu lemmat reprezentujících lexém se tedy může vyskytnout více lemmat se stejnou vidovou charakteristikou. V takovém případě je tato charakteristika doplněna arabskou číslicí tak, aby index s vidovou charakteristikou mohl sloužit jako jednoznačný identifikátor lemmatu ve slovníkovém hesle. Tento identifikátor uvádí glosy a příklady, případně další údaje, které se vztahují pouze k některým z lemmat reprezentujících lexém. Ve *VALLEXu* se to týká např. sloves *dohánět<sup>impf1</sup>*, *dohnat<sup>pf1</sup>* *dohonit<sup>pf2</sup>*; *odvinovat<sup>impf1</sup>*, *odvíjet<sup>impf2</sup>*, *odvinout<sup>pf</sup>*; *odřezávat<sup>impf</sup>*, *odříznout<sup>pf1</sup>*, *odřezat<sup>pf2</sup>*.

Lexikální jednotka se typicky vztahuje ke všem lemmatům daného lexému, která jsou uvedena v záhlaví slovníkového hesla. Toto obecné pravidlo má však řadu výjimek, konkrétní lexikální jednotky (viz odd. 3) se mohou vztahovat jen k některým z uvedených lemmat. Například sloveso *odpovědět<sup>pf</sup>* je dokonavým protějškem slovesa *odpovídat<sup>impf</sup>* ve smyslu ‚dát odpověď‘, ale již ne ve smyslu ‚reagovat‘, ‚mít odpovědnost‘ či ‚být ve shodě/v souladu; korespondovat‘ (v těchto významech jde o imperfektum tantum). V takovém případě jsou za arabskou číslicí uvádějící příslušnou lexikální jednotku uvedena za značkou jen všechna lemmata, ke kterým se tato lexikální jednotka vztahuje (s vyznačeným videm a případnými variantami a indexem pro homografy, viz níže); toto omezení se neuvádí pro iterativa.

Ve třídě nedokonavých sloves se dále vyčleňuje skupina sloves násobených (iterativní slovesa, iterativa) označujících opakovaný děj. V češtině je tvoření iterativ velmi produktivní a do značné míry pravidelné, tvoří se od nedokonavých sloves příponou *-va(t)* s kvantitativní či kvalitativní změnou vokálu před příponou (např. *volat* → *volávat*, *křičet* → *křičívat*, *být* → *bývat*), příp. příponou *-a(t)* (např. *jíst* → *jídat*). Vzhledem k vysoké produktivitě při tvorbě iterativ nejsou ve *VALLEXu* iterativa zachycena vyčerpávajícím způsobem: iterativa s rozšířenou variantou přípony *-váva(t)* (např. *chodívat*) nejsou uváděna vůbec, z ostatních iterativ jsou uváděna pouze ta, která se ve sledovaných textech a zdrojových slovnících vyskytovala pravidelněji.

Lemmata iterativních sloves jsou součástí záhlaví slovníkového hesla, neuvádějí se však pro ně glosy a příklady ani se nezaznamenává omezení u lexikálních jednotek, pro které iterativum nelze užít.

### 2.3 Varianty lemmatu

Varianty lemmatu (často ortografické alternativy) jsou chápány jako skupina dvou nebo více lemmat reprezentujících daný lexém, která jsou zaměnitelná v jakémkoliv kontextu beze změny významu, vztahují se k nim tedy stejné lexikální jednotky (např. *dozvědět/dovědět se*, *dýchnout/dechnout*). Obvykle se varianty

liší jen alternací v morfematickém kmenu slovesa, která je případně doprovázena stylistickým posunem (např. *myslet/myslit*), navíc někdy mohou mít obě lemmata společné některé tvary paradigmatu (např. *mysli* (imperativ) je společným tvarem pro *myslet* i *myslit*).

Všechny varianty lemmatu jsou popsány společně v jednom lexému. Varianty jsou uvedeny v záhlaví hesla, jsou oddělené lomítkem, např. *dověděť/dozvěděť se, dýchnout/dechnout, myslet/myslit*.

Přes toto základní vymezení variant *VALLEX* obsahuje řídke výjimky, kdy lze užít pouze jedno z lemmat. Např. lemmata *plavat* a *plovat* jsou tradičně považována za varianty (viz SSJČ), přestože v některých kontextech lze užít jen *plavat*, např. *plavat při zkoušce* vs. *\*plovat při zkoušce*. V případě, že lze v daném významu užít pouze jednu z variant, jsou za arabskou číslicí uvádějí příslušnou lexikální jednotku za značkou jen uvedena ta lemmata a jejich varianty, ke kterým se tato lexikální jednotka vztahuje (s vyznačeným videm a indexem pro homografy, viz níže).

## 2.4 Homografy (homonyma)

Jako *homografy* jsou ve *VALLEXu* označována lemmata s identickou grafickou podobou, ale bez zřejmého sémantického vztahu. Jde tedy o různé lexémy, jejichž lemmata jsou reprezentována stejnou kombinací morfémů, tj. mají stejné grafematické vyjádření.

Často se též liší jejich etymologie (např. *nakupovat<sub>I</sub>* jako *nakupovat dětem oblečení* vs. *nakupovat<sub>II</sub>* jako *nakupovat kolem sebe hromady věcí*), vid (např. *stačit<sub>I</sub><sup>impf</sup>* jako *stačí mu to ke štěstí* či *Petr stačí Pavlovi v běhu* vs. *stačit<sub>II</sub><sup>pf</sup>* jako *stačí dorazit do školy včas*) nebo některé tvary paradigmatu (např. infinitiv *žít* je homograf, který má pro 3. osobu singuláru minulého času formu *žil* pro význam ‚být naživu; trávit čas‘, např. *Jan žil v Praze*, ale formu *žal* pro význam ‚kosit; sekat‘, např. *Jan žal trávu*).

Ve *VALLEXu* jsou homografy rozlišovány římskou číslicí v dolním indexu (v případě reflexivních lemmat před morfémem *se/si*, např. *dít<sub>I</sub>*, *dít<sub>II</sub> se*).

*Terminologická poznámka:* Zde se držíme terminologie běžné v anglicky psané literatuře, která rozlišuje homografy jako jednotky se stejnou psanou podobou (bez ohledu na podobu zvukovou) a homofony jako jednotky se zvukově stejnou podobou. Termín *homograf* užívaný ve *VALLEXu* tedy zahrnuje termíny homonymum (jednotka se stejnou psanou i zvukovou podobou) i homograf (jednotka se stejnou psanou, ale odlišnou zvukovou podobou), jak je užívá česky psaná literatura.

## 3 Lexikální jednotky

V koncepci slovníku *VALLEX* reprezentují lexikální jednotky významovou složku, která spolu se složkou formální (s možnými lexikálními formami) vytváří lexém. Každý lexém je tedy tvořen množinou lexikálních jednotek (LU), kterým jsou přiřazeny příslušné lexikální formy (reprezentované lemmaty). V souladu s D. A. Crusem (viz Cruse, 1986) považujeme lexikální jednotky za „komplexní jednotky s (relativně) stálými, diskrétními sémantickými vlastnostmi.“<sup>7</sup> Stručně řečeno, jde zhruba o ‚dané slovo v daném významu‘.

*Poznámka k vyčleňování jednotlivých LU:* Pro vyčleňování jednotlivých významů daného lexému neexistují všeobecně přijatá testovatelná kritéria, přechod od jednoho významu k druhému je v řadě případů pozvolný. Ve *VALLEXu* je při rozlišování jednotlivých LU kladen důraz na syntaktická kritéria, zejména na podobu valenčního rámce, včetně povrchové realizace jednotlivých valenčních doplnění (viz odd. 4). Přitom se ovšem přihlíží též k sémantice.

- Změny ve valenčním rámci (s výjimkou morfematických variant) vedou k vyčlenění více LU, i když je význam těchto LU blízký (např. následující užití slovesa *poslat* bude popsáno dvěma LU: *poslat peníze do banky/na účet.DIR3* vs. *poslat peníze dětem.ADDR*).

<sup>7</sup> Cruse (1986): “form-meaning complexes with (relatively) stable and discrete semantic properties”.

- Podobně vede k vyčlenění různých LU i různá syntaktická strukturace (např. *naložit vůz.PAT senem.EFF* vs. *naložit seno.PAT na vůz.DIR3*; *žnout louku.PAT* vs. *žnout trávu.PAT na louce.LOC*). Systematické provázání těchto blízkých LU pomocí tzv. alternačního modelu navrženého ve studiích Žabokrtský (2005) a Lopatková et al. (2006) není v současné verzi slovníku uplatněno.
- Jednotlivé valenční členy jsou specifikovány syntakticko-sémantickým vztahem k řídicímu slovesu, odlišnost tohoto vztahu opět vede k různým LU (např. tři LU pro lexém *pocházet*: *rukopis pochází ze 14. stol.TFRWH* vs. *pochází z venkova.DIR1* vs. *všechno zlo pochází z bídy a neznalosti.PAT*).
- Pokud má sloveso dva (či více) zřetelně odlišné významy, jsou tyto významy popsány různými LU i v případech, kdy se valenční rámec neliší (např. dvě LU pro sloveso *chovat*: *chovat dítě.PAT v náručí.LOC* vs. *chovat prasata.PAT na farmě.LOC*).

*Terminologická poznámka:* Lexikální jednotka spolu se svými lexikálními formami, jak je chápána ve VALLEXu, odpovídá jednotce v české tradici označované jako monosémní/monosémický lexém, lexie nebo též základní lexikální jednotka (viz Filipec – Čermák, 1985, Karlík et al., 2002), zde též odd. 1.

Lexikální jednotky jsou ve VALLEXu číslovány arabskými číslicemi – pokud má lexém více významů popsaných několika lexikálními jednotkami, je každá lexikální jednotka uvozena touto číslicí.

Pořadí lexikálních jednotek není zcela arbitrární, není ale přísně systematické. V této podobě slovníku je dáno intuicí autorů (s přihlédnutím ke vzorku korpusového materiálu) – primární a/nebo velmi frekvencované významy jsou uváděny na prvních místech, zatímco řídké a idiomatické významy jsou řazeny na konec slovníkového hesla.

Pokud není specifikováno jinak, vztahuje se lexikální jednotka ke všem lemmatům reprezentujícím daný lexém (uvedeným v záhlaví slovníkového hesla). V případech, kdy se daná lexikální jednotka vztahuje jen k některým lemmatům ze seznamu v záhlaví, jsou za značkou jen uvedena všechna lemmata, ke kterým se tato lexikální jednotka vztahuje (s vyznačeným videm a případnými variantami a indexem pro homografy).<sup>8</sup>

Lexikální jednotky zhruba odpovídají lexému v určitém významu a nesou informaci o syntaktických a sémantických rysech slovesa v daném významu. Příslušné informace jsou ve VALLEXu zachyceny jako povinné a nepovinné atributy lexikální jednotky. Povinné atributy musí být vyplněny pro každou lexikální jednotku. Nepovinné atributy mohou být nevyplněny, buď protože se u dané lexikální jednotky nevyskytují (např. kontrola se uvádí jen u sloves s touto vlastností, viz odd. 5.1), nebo protože dané informace nejsou v současné podobě dostupné (např. určení syntakticko-sémantické třídy slovesa, viz odd. 5.4).

Povinné atributy lexikální jednotky:

- valenční rámec, viz odd. 4;
- glosa – sloveso nebo parafráze charakterizující daný význam slovesa; glosy nelze pokládat za synonyma nebo dokonce za lexikografické definice, slouží pouze pro orientaci ve slovníkovém hesle;
- příklad – věty nebo fragmenty vět obsahujících dané sloveso v daném významu, případně s označením zdroje příkladu, např. ČNK, SSJČ apod.

Nepovinné atributy lexikální jednotky:

- kontrola, viz odd. 5.1;
- možný typ reflexivních konstrukcí, viz odd. 5.2;
- možný typ recipročních konstrukcí, viz odd. 5.3;
- příslušnost k syntakticko-sémantické třídě, viz odd. 5.4;
- označení idiomu, viz odd. 5.5.

<sup>8</sup> Pokud se daná lexikální jednotka vztahuje ke všem lemmatům uvedeným v záhlaví slovníkového hesla s výjimkou iterativa, žádné omezení se neuvádí.

## 4 Valenční rámce

Nejdůležitější sémanticko-syntaktická charakteristika slovesa je zachycena ve formě *valenčního rámce*. Valenční rámec (v užším smyslu) ve FGP sestává z aktantů (obligatorních i fakultativních) a z obligatorních volných doplnění, v novějších studiích je pak obohacen o tzv. kvazivalenční doplnění. Ve *VALLEXu* se kromě členů takto pojímaného valenčního rámce uvádí i nevelké množství fakultativních volných doplnění (dále typická doplnění). S daným slovesem se mohou vyskytovat též ostatní volná doplnění, ta však nejsou ve valenčním rámci uváděna, neboť jejich výskyt není podle FGP podmíněn syntakticky. Klasifikaci valenčních doplnění tvořících obohacený valenční rámec ve *VALLEXu* je zde věnován oddíl 4.1.

Ve *VALLEXu* jsou valenční rámce modelovány jako posloupnosti valenčních a nevalenčních pozic, kde každá pozice odpovídá jednomu valenčnímu, příp. typickému doplnění daného slovesa. Každá pozice je charakterizována:

- funktorem, viz odd. 4.1;
- seznamem možných morfematických forem, viz odd. 4.2;
- informací o obligatornosti, viz odd. 4.3.

Jistá volná doplnění se systematicky objevují společně. Tato pravidelnost je zachycena pomocí mechanismu expanze valenční či nevalenční pozice, viz odd. 4.4; plný valenční rámec se získá expanzí pozice uváděné ve slovníku.

### 4.1 Aktanty a volná doplnění

Ve valenční teorii FGP se slovesná doplnění dělí na *aktanty* (vnitřní doplnění ve všech svých výskytech) a *volná doplnění* (viz zejména Panevová, 1974, 1980, 1994). Zkratky pro jednotlivé aktanty a volná doplnění se dále souhrnně označují jako *funktory*. Jednotlivé funktoři ve *VALLEXu* tedy označují typ sémanticko-syntaktického vztahu mezi slovesem a jeho doplněním.

*Aktanty* jsou určovány převážně na základě syntaktických pravidel:

- počet pozic pro aktanty je charakteristický pro každé sloveso a pro každé sloveso tedy musí být vymezen ve slovníku;
- jako rozvíti nějakého konkrétního slovesa se daný aktant vyskytuje nejvýše jednou (vyjma případů souřadnosti a apozice).

Doplňme, že aktanty jsou doplnění typicky reálná.

Empiricky bylo stanoveno pět aktantů: aktor (ACT), pacient (PAT), výsledek děje (EFF), adresát (ADDR) a původ (ORIG). Zásady pro určování jednotlivých aktantů lze nalézt např. v Panevová – Skoumalová (1992), nově v Mikulová et al. (2005), kde je i jejich (zatím nejpodrobnější) charakteristika; zde je krátká charakteristika aktantů uvedena níže.

*Volná doplnění* jsou na rozdíl od aktantů sémanticky distinktivní. Charakterizují je následující vlastnosti:

- omezení na slučitelnost slovesa s volnými doplněními nemají podle FGP z velké části syntaktický charakter;
- dané sloveso může být rozvíjeno jedním typem volného doplnění i více než jedenkrát.

Volná doplnění (včetně typických forem) jsou popsána v Mikulová et al. (2005), níže uvádíme typické příklady.

Dichotomie aktant – volné doplnění byla v novějších studiích obohacena o třetí typ tzv. *kvazivalenčních doplnění* (viz Panevová, 2003, Lopatková – Panevová, 2006), která jsou na hranici mezi aktanty



a volnými doplněními. Jde o doplnění rozvíjející relativně uzavřenou (sémanticky homogenní) třídu sloves, jsou to doplnění rekční a dané doplnění nelze u jednoho řídicího slovesa opakovat. Podobně jako volná doplnění jsou však sémanticky distinktivní a typicky se nejedná o doplnění obligatorní. Jde např. o záměr INTT u sloves pohybu (třídy motion a transport, odd. 5.4, např. *Petr jel nakoupit, Maruška šla na jahody*) či o překážku OBST (třída contact, např. *zakopl o kořen, zachytil šálou o hřebík*).

Dále se ve *VALLEXu* uvádí nevelké množství fakultativních volných doplnění, která obvykle nespecifikují význam slovesa, ale typicky se vztahují k celé syntakticko-sémantické třídě sloves. Pro některá doplnění mají prototypickou formu (např. instrumentál pro způsob, *psal perem, jel vlakem*, či předložková skupina pro+4 pro benefaktiv, *dělal to pro děti*), jindy je jejich forma dána sémantikou příslušného doplnění (např. směrová doplnění DIR1, DIR2 a DIR3 u sloves pohybu, více viz odd. 4.2). Takováto doplnění se obvykle chápou jako doplnění nevalenční, ve *VALLEXu* se však uvádějí, protože tato informace může být s úspěchem využita při automatické analýze češtiny.

*Terminologická poznámka:* Aktanty a obligatorní volná doplnění ve FGP víceméně odpovídají konstitutivním větným členům (obligatorním i potenciálním), typická a kvazivalenční doplnění odpovídají větným členům nekonstitutivním (viz Daneš, 1971, Daneš – Hlavsa, 1987, Grepl – Karlík, 1998). V termínech Mluvnice češtiny 3 (Daneš et al., 1987) odpovídají aktanty a obligatorní volná doplnění participantům intenčního pole.

Dělení na aktanty a volná doplnění se také víceméně shoduje s dělením na argumenty a adjunktory podle Grepl – Karlík (1998), Karlík et al. (2002).

Funktory označující typ sémanticko-syntaktického vztahu jsou blízké tzv. hloubkovým pádům/rolím C. J. Fillmore (např. Fillmore, 1969) či theta rolím podle N. Chomského (viz např. Chomsky, 1981, Jackendoff, 1990).

### Charakteristika aktantů

- **Aktor/konatel (funktor ACT).** Valenční doplnění aktor je (levovalenční) aktant, který je vymezen jako první aktant slovesa – označuje doplnění zaplňující první syntaktickou pozici slovesa (např. *maminka.ACT upekla koláč, voda.ACT naplnila jámu, kniha.ACT vyšla*). V zásadě je to doplnění v pozici syntaktického subjektu u aktivní konstrukce (v případě pasivní konstrukce se jedná o doplnění se stejným sémantickým vztahem ke slovesu, např. *nakladatelství.ACT Odeon vydalo knihu i kniha byla vydána nakladatelstvím.ACT Odeon*). Jde o rozšířené pojetí konatele děje zahrnující jak činitele, tak i nositele stavu/děje a příbuzné sémantické role. Je-li jeden z aktantů vyjádřen dativní formou (a druhý nominativní formou), přihlíží se též k sémantice aktantu. Vyjadřuje-li aktant s dativní formou proživatele, hodnotí se tento aktant jako aktor (a aktant v nominativu jako patient) (např. *knihy se mi.ACT líbila*).
- **Patient (funktor PAT).** Valenční doplnění patient je (pravovalenční) aktant, který je vymezen jako druhý aktant slovesa – označuje doplnění zaplňující druhou syntaktickou pozici slovesa (např. *Marie postavila vázu.PAT na stůl, maminka upekla koláč.PAT, kniha patří Janovi.PAT, obraz.PAT se mi nelíbil, vzdal se odměny.PAT, učil se zahradníkem.PAT, vyprávěl nám o dovolené.PAT*). V zásadě je to doplnění v pozici přímého (syntaktického) objektu u aktivní konstrukce (v případě pasivní konstrukce se jedná o doplnění se stejným sémantickým vztahem, např. *nakladatelství Odeon vydalo knihu.PAT, kniha.PAT byla vydána nakladatelstvím Odeon*). Jde o rozšířené pojetí předmětu zasaženého dějem.
- **Výsledek děje, efekt (funktor EFF).** Valenční doplnění efekt je (pravovalenční) aktant, který se uplatňuje u sloves se třemi (a více) syntaktickými pozicemi. Funktor EFF je přiřazován zejména doplněním obsazujícím třetí syntaktickou pozici u sloves, která odpovídá jednak tzv. doplňku u sloves neplnovýznamových (viz též Šmilauer, 1966),<sup>9</sup> jednak jde o druhý věcný předmět sémanticky se blížící výsledku děje. Obecně vyjadřuje vlastnost nebo stav, které má doplnění s funktoem PAT za

<sup>9</sup> Doplněk doplňovací v prvním vydání Novočeské skladby (Šmilauer, 1947).

jistého děje nebo které se mu jistým dějem přisuzují (např. *považoval Pavla za odborníka.EFF, jmenovali ho ředitelem.EFF, my tomu říkáme efekt.EFF sněhové koule, Petr přeložil knihu do češtiny.EFF, svazovali kmeny do voru.EFF*).

- **Adresát (funktor ADDR).** Valenční doplnění adresát je (pravovalenční) aktant, který je vymezen jako aktant slovesa typicky vyjadřující roli příjemce děje (např. *dal dceři.ADDR k narozeninám knížku, řekl synovi.ADDR pravdu, bratrovi.ADDR nezaplátili dohodnutou mzdu, celé dětství soupeřil o matčinu přízeň s bratrem.ADDR*). Funktor ADDR se uplatňuje u sloves se třemi (a více) syntaktickými pozicemi. Jeho typickým rysem je životnost. Prototypicky jde o doplnění v pozici nepřímého objektu (např. *předal knihu Janovi.ADDR, kniha byla předána Janovi.ADDR*).
- **Původ (funktor ORIG).** Valenční doplnění původu je (pravovalenční) aktant, který je vymezen jako aktant slovesa vyjadřující roli původu (např. *vyrábějí ze dřeva.ORIG stoly i židle, slyšel o neštěstí od sousedů.ORIG, nevzal od něj.ORIG za práci peníze*). Funktor ORIG se uplatňuje u sloves se třemi (a více) syntaktickými pozicemi.

*Poznámka:* Podle valenční teorie FGP, viz zejména Panevová (1974, 1980), platí pro určování funktořů následující princip, který je označován jako *princip posouvání* (shifting): pokud má sloveso jediný aktant, jde o aktor, sloveso se dvěma aktanty má vždy aktor a patient; teprve u sloves se třemi a více aktanty přistupují při výběru funktořů sémantická kritéria.

**Funktory ve VALLEXu.** V následujícím výčtu jsou shrnuty funktoř, které se v tomto slovníku vyskytují. Pro úplnost zde uvádíme všechny funktoř pro volná doplnění s alespoň jedním výskytem ve VALLEXu bez ohledu na to, zda jsou v konkrétních příkladových větách příslušná doplnění valenční, nebo zda jde pouze o doplnění typická, tedy nevalenční.

Aktanty:

- ACT (aktor): *Petr čte dopis.*
- PAT (patient): *Potkal jsem bratra.*
- EFF (výsledek děje, efekt): *Jmenovali ho ředitelem.*
- ADDR (adresát): *Petr dal Marii knihu.*
- ORIG (původ): *Upekla z jablek koláč.*

Kvazivalenční doplnění:

- DIFF (rozdíl): *Hodnota akcií stoupla o 100 %.*
- INTT (záměr): *Přišel navštívit Janu.*
- OBST (překážka): *Chlapec zakopl o kořen.*

Volná doplnění (abecedně):

- ACMP (doprovod): *Matka tam šla s dětmi.*
- AIM (účel): *Jan šel do pekárny pro chléb.*
- BEN (benefaktiv): *Připravila snídani pro děti.*
- CAUS (příčina): *Petr pro nemoc končí s prací.*
- COMPL (doplněk): *Pracoval jako učitel.*
- CRIT (kritérium): *Třídili diamanty podle velikosti.*
- DIR1 (směr – odkud): *Přišel z lesa promočený.*
- DIR2 (směr – kudy): *Vydal se do sousední vesnice přes les.*

- DIR3 (směr – kam): *Vydal se do sousední vesnice přes les.*
- DPHR (závislá část frazému): *Novináři ho neustále chytali za slovo.*
- EXT (míra): *Tatínek měřil 2 metry.*
- HER (dědictví): *Pojmenovali nejstaršího syna po otci.*
- LOC (místo): *Narodil se v Itálii.*
- MANN (způsob): *Choval se k ní laskavě.*
- MEANS (prostředek): *Napsal dopis rukou.*
- RCMP (náhrada): *Koupila si nové tričko za 350 Kč.*
- REG (zřetel): *Situace se v tomto ohledu výrazně zlepšila.*
- SUBS (substituce): *Startoval za Slávii.*
- TFHL (čas – na jak dlouho): *Přerušil studium na rok.*
- TFRWH (čas – ze kdy): *Jeho špatné vzpomínky pocházejí právě z tohoto období.*
- THL (čas – jak dlouho): *Strávili jsme tam tři týdny.*
- TOWH (čas – na kdy): *Odložili zkoušku z pondělka na úterý.*
- TSIN (čas – od kdy): *Lhůtu počítáme od okamžiku dodání.*
- TTIL (čas – do kdy): *Potrvá to do večera.*
- TWHEN (čas – kdy): *Babička přijde zítra.*

*Poznámka:* Kromě těchto funktorů se ve VALLEXu vyskytuje ještě hodnota DIR. Ta je však užívána jen v souvislosti s expanzí pozice valenčního rámce, viz odd. 4.4.

Množina funktorů, se kterými se pracuje ve FGP a která je využita v PDT, je bohatší, viz např. Mikulová et al. (2005). Některé z těchto funktorů se však nevyskytují u slovesných doplnění (např. MAT – partitiv, jako ve spojení *sklenice piva.MAT*), jiné funktory specifikují vztahy, které nejsou závislostní (např. koordinaci, *Petr nebo.DISJ Marie*). Další funktory reprezentují závislostní vztahy u sloves, nemají však nikdy valenční povahu (např. ATT – postoj, *udělal to dobrovolně.ATT*).

## 4.2 Morfematické vyjádření

Každá valenční i nevalenční pozice může být ve větě vyjádřena omezenou množinou výrazových prostředků, morfematických forem. Ve VALLEXu je množina možných forem specifikována buď explicitně, nebo implicitně.

U explicitně zachycených forem jsou možné morfematické formy dány výčtem u dané pozice valenčního rámce (dolní index u příslušného funktoru). U aktantů a kvazivalenčních doplnění je tento seznam forem úplný (udávají se formy pro užití slovesa v aktivním tvaru) – jiné prostředky nelze pro vyjádření těchto valenčních doplnění užít.<sup>10</sup> V případě volných doplnění jsou explicitně uvedené formy pro dané sloveso pouze typické, lze užít i další formy dané sémantikou doplnění.

U implicitně zachycených forem se předpokládá, že množina možných forem je dána sémantikou doplnění, tedy vyplývá z příslušného funktoru. Jinými slovy, doplnění se může realizovat jakoukoliv formou vyjadřující daný typ doplnění; její výběr je ovšem závislý na lexikálním obsazení a kontextových podmínkách, např. *bydlí na kopci vs. ve vesnici, napsal dopis rukou vs. na počítači*.

**Explicitně zachycené formy.** Seznamy morfematických forem, které se mohou vyskytnout u jednotlivých valenčních pozic, sestávají z následujících typů hodnot:

<sup>10</sup> Zcela stranou jsou však ponechány formy, které jsou dány gramatickými pravidly, např. pasivizací nebo reciproční konstrukcí, a dále formy pro partitiv (*dodat sůl – dodat soli*), distributivnost (*rozdal jim jablíčka – rozdál jim po jablíčku*) či méně přesnou kvantifikaci (*přišlo padesát lidí – přišlo na padesát lidí, přišlo okolo padesáti lidí*).

- **Bezpředložkové pády.** Jednotlivé pády jsou označeny příslušnými číslicemi: 1 – nominativ, 2 – genitiv, 3 – dativ, 4 – akuzativ, 5 – vokativ, 7 – instrumentál.
- **Předložkové skupiny.** Jsou určeny lemmatem předložky (v její nevokalizované podobě) a číslem pádu (např. z+2, na+4, o+6, ...). Ve *VALLEXu* se vyskytují následující předložky: *bez, do, jako*,<sup>11</sup> *k, kolem, mezi, místo, na, nad, o, od, po, pod, podle, pro, proti, před, přes, při, s, u, v, z, za*.
- **Infinitivní konstrukce.** Značka inf reprezentuje valenční doplnění ve formě infinitivu slovesa (ve vzácných případech též se spojkou než+inf).
- **Závislé věty.** Závislé věty obsahové uvozené podřadicí spojkou jsou reprezentovány lemmatem této spojky; ve *VALLEXu* se vyskytují následující spojky: *aby, ať, až, jak, zda*,<sup>12</sup> *že*. Závislé věty obsahové, které nejsou uvozeny spojkami (např. nepřímé otázky uvozené tázacím zájmenem nebo adverbem), jsou reprezentovány zkratkou cont.
- **Konstrukce s adjektivy.** Zkratka adj-číslice specifikuje doplnění ve formě přídavného jména v příslušném pádu (např. adj-1 pro *cítím se slabý*).
- **Konstrukce s být.** Infinitiv slovesa *být* se může vyskytnout v konstrukci s adjektivem či v bezpředložkovém pádu (např. být+adj-1 pro *zdá se to být dostatečné*).
- **Část frazému.** U frazeologických jednotek platí, že pokud je množina lexikálních forem, které naplňují určitou valenční pozici, omezená (často jednočlenná), jsou ve *VALLEXu* uvedeny přímo tyto lexikální formy (např. *napospas* pro frazém *ponechat napospas*).

**Implicitně zachycené formy.** Pokud není pro valenční pozici explicitně určena možná forma doplnění, potom množina možných forem vyplývá z funktoru pro toto doplnění. Následující výčet udává formy obvyklé pro dané funktoři (seznam vychází z nejčastějších forem pro jednotlivé funktoři v PDT).

- ACMP: bez+2, s+7, společně s+7, spolu s+7, v čele s+7, v souvislosti s+7, ve spojení s+7, včetně+2, ...;
- AIM: aby, ať, do+2, k+3, na+4, o+4, pro+4, pro případ+2, proti+3, v zájmu+2, za+4, za+7, že, ...;
- BEN: 3, na+4, na účet+2, na úkor+2, na vrub+2, pro+4, proti+3, v+4, ve prospěch+2, v rozporu s+7, v zájmu+2, ...;
- CAUS: 7, aby, adverb, díky+3, jelikož, ježto, kvůli+3, na+4, na+6, na základě+2, nad+7, následkem+2, od+2, pod+7, pod nápor+2, pod tíhou+2, pod vahou+2, poněvadž, pro+4, protože, v+6, v důsledku+2, v souvislosti s+7, vinou+2, vlivem+2, vzhledem k+3, z+2, z důvodu+2, za+4, za+7, zásluhou+2, že, ...;
- CRIT: 2, 7, dle+2, na+6, na základě+2, po vzoru+2, podle+2, přiměřeně+3, v+6, v duchu+2, v rozporu s+7, v souladu s+7, v souhlase s+7, v závislosti na+6, ve shodě s+7, ve smyslu+2, ve světle+2, z titulu+2, ...;
- DIR1: adverb, od+2, s+2, z+2, ze strany+2, zpod+2, zpoza+2, zpřed+2, ...;
- DIR2: 7, adverb, cestou+2, kolem+2, mezi+7, napříč+7, po+6, podél+2, přes+4, skrz+4, v+6, ...;
- DIR3: 7, adverb, do+2, do čela+2, k+3, kolem+2, mezi+4, mimo+4, na+4, na+6, nad+4, naproti+3, okolo+2, po+4, po+6, pod+4, proti+3, před+4, přes+4, směrem do+2, směrem k+3, směrem na+4, v+4, vedle+2, za+4, za+7, ...;
- EXT: 2, 4, 7, adverb, do+2, k+3, kolem+2, na+4, na+6, nad+4, okolo+2, po+6, pod+7, přes+4, v+4, z+2, za+4, ...;
- LOC: adverb, blízko+2, blízko+3, daleko+2, do+2, kolem+2, mezi+7, mimo+4, na+4, na+6, na úroveň+2, nad+7, naproti+3, nedaleko+2, okolo+2, po+6, poblíž+2, pod+7, podél+2, proti+3, před+7, přes+4, při+6, stranou+2, u+2, uprostřed+2, uvnitř+2, v+6, v čele+2, v oblasti+2, v rámci+2, v řadě+2, vedle+2, za+4, za+7, ...;
- MANN: 7, adverb, do+2, formou+2, na+4, na+6, nad+4, o+4, po+6, pod+7, proti+3, před+7, při+6, přes+4, s+7, v+4, v+6, v podobě+2, ve formě+2, vedle+2, z+2, za+4, za+7, jak, že, ...;
- MEANS: 7, adverb, cestou+2, díky+3, do+2, na+4, na+6, o+6, po+6, pod+7, pomocí+2, prostřednictvím+2, přes+4, s+7, s pomocí+2, skrz+2, v+6, z+2, za+4, za pomoci+2, že, ...;

<sup>11</sup> Slovo *jako* je sice tradičně považováno za spojku, zde je však zahrnuto mezi předložkami, neboť konkrétní valenční doplnění uvozené touto spojkou vyžaduje vždy určitý pád substantiva.

<sup>12</sup> Spojka *zda* reprezentuje též spojkou *jestli*.

- REG: 7, adverb, bez ohledu na+4, bez zřetele k+3, k+3, kolem+2, na+4, na+6, na téma+2, nad+7, nezávisle na+6, o+6, ohledně+2, po+6, pro+4, před+7, při+6, s+7, s ohledem na+4, se zřetelem k+3, se zřetelem na+4, u+2, v+6, v otázce+2, v případě+2, v rámci+2, v souvislosti s+7, ve věci+2, ve vztahu k+3, vůči+3, vzhledem k+3, z+2, z hlediska+2, za+4, ...;
- SUBS: jménem+2, místo+2, namísto+2, výměnou za+4, za+4, ...;
- TFHL: adverb, do+2, na+4, po+2, pro+4, ...;
- TFRWH: od+2, z+2, ...;
- THL: 2, 4, 7, adverb, až, dokud, do+2, na+4, po+4, po dobu+2, přes+4, v+2, za+4, ...;
- TOWH: adverb, do+2, k+3, na+4, pro+4, ...;
- TSIN: adverb, od+2, počínaje+7, z+2, ...;
- TTILL: adverb, do+2, dokud, k+3, než, po+4, ...;
- TWHEN: 2, 4, 7, adverb, až, do+2, jakmile, k+3, když, kolem+2, koncem+2, mezi+7, na+4, na+6, na závěr+2, než, o+6, okolo+2, po+6, počátkem+2, postupem+2, poté co, před+7, předtím než, při+6, s+7, u příležitosti+2, v+4, v+6, v době+2, v období+2, v průběhu+2, v závěru+2, z+2, za+2, za+4, začátkem+2, ...

### 4.3 Atribut obligatornosti slovesného doplnění

Ve *VALLEXu* se v souladu s valenční teorií FGP slovesná doplnění dělí na obligatorní a fakultativní. *Obligatorností* se rozumí povinná přítomnost daného doplnění v hloubkové (tektogramatické) struktuře, a to bez ohledu na jeho možnou povrchovou vypustitelnost ve větě, viz poznámku níže. Jako kritérium obligatornosti byl stanoven *dialogový test* (viz Panevová, 1974, Sgall et al., 1986). Tento test slouží pro určení obligatornosti doplnění, je-li zkoumaný člen v povrchové větě vypuštěn – např. test obligatornosti doplnění směru-kam (funktor DIR3) pro sloveso *přijít* simuluje dialog mluvčího A a B:

A: *Přátelé už přišli.*

B: *Kam?*

A: *\*Nevím.*

Odpověď mluvčího A činí dialog deviantní (A musí vědět, o jakém místě mluví) – sloveso *přijít* má tedy obligatorní doplnění DIR3.

Opozice obligatornosti a fakultativnosti se týká aktantů, kvazivalenčních doplnění i volných doplnění.

*Poznámka:* Některá doplnění obligatorní na rovině významové stavby mohou být vypuštěna (elidována) v povrchové realizaci věty, aniž dojde k porušení gramatičnosti věty (lze říci, že dané doplnění má nulovou lexikální realizaci). K takové elipse dochází tehdy, je-li možné příslušný aktant či volné doplnění snadno doplnit z kontextu, např. *Děti už přišly* (= na místo dané kontextem/sem.DIR3) *a jsou celé promrzlé* (= děti.ACT), případně pokud je daná pozice realizována nějakým typem všeobecného aktantu, např. *Do této buchy se dává sůl*, *Psali to v novinách* (= všeobecný aktor), viz Daneš (1971), Panevová – Řezníčková (2001).

U každé pozice valenčního rámce je ve *VALLEXu* kódována informace o obligatornosti či fakultativnosti daného doplnění. Obligatorní doplnění (aktanty, kvazivalenční i volná doplnění) jsou tištěna zvýrazněným písmem. Stejným písmem jsou tištěny i fakultativní aktanty a kvazivalenční doplnění (patří též do úzce chápaného valenčního rámce), ty jsou navíc odlišeny značkou *opt* v horním indexu. Typická volná doplnění, která rozšiřují tradiční valenční rámec, jsou tištěna obyčejným písmem a označena horním indexem *typ*.

### 4.4 Expanze valenční pozice

Jistá volná doplnění se systematicky objevují společně. Např. slovesa pohybu lze často rozvíjet všemi typy směrových doplnění, tedy DIR1 (směr-odkud), DIR2 (směr-kudy) a DIR3 (směr-kam). Tato pravidelnost je zachycena pomocí mechanismu expanze pozice valenčního rámce. Pokud je u některé pozice uveden symbol pro expanzi ↑ před funktorem, je plný valenční rámec získán expanzí dané pozice rámce.

Ve *VALLEXu* se symbol pro expanzi  $\uparrow$  vyskytuje u funktorů DIR, DIR1, DIR2, DIR3 a THL, expanze je popsána následujícími pravidly:

- $\uparrow\text{DIR}^{\text{typ}} \rightarrow \text{DIR1}^{\text{typ}} \text{DIR2}^{\text{typ}} \text{DIR3}^{\text{typ}}$

Typické doplnění  $\uparrow\text{DIR}$  expanduje ve tři typická doplnění DIR1, DIR2 a DIR3;

např. rámec pro sloveso *jít* vznikne následující expanzí:

$\text{ACT}_1 \text{INTT}_{k+3,na+4,inf}^{\text{opt}} \text{MANN}^{\text{typ}} \text{MEANS}^{\text{typ}} \uparrow\text{DIR}^{\text{typ}} \rightarrow$

$\rightarrow \text{ACT}_1 \text{INTT}_{k+3,na+4,inf}^{\text{opt}} \text{MANN}^{\text{typ}} \text{MEANS}^{\text{typ}} \text{DIR1}^{\text{typ}} \text{DIR2}^{\text{typ}} \text{DIR3}^{\text{typ}}$

(*Petr.ACT jel nakoupit.INTT autem.MEANS z domova.DIR1 přes celou Prahu.DIR2 do Makra.DIR3*)

Obdobně i pro další typy expanze:

- $\uparrow\text{DIR1} \rightarrow \text{DIR1} \text{DIR2}^{\text{typ}} \text{DIR3}^{\text{typ}}$

Doplnění  $\uparrow\text{DIR1}$  expanduje v obligatorní doplnění DIR1 a typická DIR2 a DIR3.

- $\uparrow\text{DIR2} \rightarrow \text{DIR2} \text{DIR1}^{\text{typ}} \text{DIR3}^{\text{typ}}$

Doplnění  $\uparrow\text{DIR2}$  expanduje v obligatorní doplnění DIR2 a typická DIR1 a DIR3.

- $\uparrow\text{DIR3} \rightarrow \text{DIR3} \text{DIR1}^{\text{typ}} \text{DIR2}^{\text{typ}}$

Doplnění  $\uparrow\text{DIR3}$  expanduje v obligatorní doplnění DIR3 a typická DIR1 a DIR2.

- $\uparrow\text{THL} \rightarrow \text{TSIN}^{\text{typ}} \text{THL} \text{TTIL}^{\text{typ}}$

Doplnění  $\uparrow\text{THL}$  expanduje ve tři typická doplnění TSIN, THL a TTILL;

např. rámec pro sloveso *trvat* vznikne následující expanzí:

$\text{ACT}_1 \text{PAT}_3^{\text{opt}} \uparrow\text{THL} \rightarrow \text{ACT}_1 \text{PAT}_3^{\text{opt}} \text{THL} \text{TSIN}^{\text{typ}} \text{TTIL}^{\text{typ}}$

(*Práce na novém obraze.ACT mu.PAT trvala půl roku.TH od jara.TSIN až do konce října.TTILL*)

## 5 Doplnující syntaktické informace

Jednotlivé LU mohou být obohaceny o nepovinné, doplňující syntaktické, případně syntakticko-sémantické informace, které s valencí souvisejí jen volně. Ve *VALLEXu* je zachycena kontrola (odd. 5.1), reflexivita (odd. 5.2) a reciprocita (odd. 5.3) – jde o gramatické jevy, které přímo ovlivňují povrchové projevy valence. Dále se u vybraných LU uvádí jejich zařazení do syntakticko-sémantické třídy (odd. 5.4), které umožňuje zkoumat, jak se sémantická blízkost sloves odráží v jejich valenčních vlastnostech, a příznak pro idiom (odd. 5.5), neboť frazémy a idiomy často vykazují specifické valenční chování.

### 5.1 Kontrola

Termínem *kontrola* (značka *control*) se v tomto kontextu rozumí vlastnost některých sloves (tzv. sloves kontroly) vyžadovat koreferenci mezi svým valenčním doplněním („controller“) a valenčním doplněním podřízeného slovesa („controllee“), viz Panevová (1996). Ve *VALLEXu* je tento vztah zaznamenán pouze pro slovesa, která mohou mít doplnění ve formě infinitivu (bez ohledu na jeho funktor). Za kontrolovaný člen (controllee) je pak považován subjekt tohoto infinitivu (který se v povrchové podobě věty v češtině nevyjadřuje), kontrolující člen (controller) je výraz s ním koreferenční, typicky člen valenčního rámce řídicího slovesa kontroly. Ve *VALLEXu* je kontrola zachycena v atributu *control* následujícím způsobem:

- koreferenční vztah mezi (nevyjádřeným) subjektem infinitivu a jedním z členů valenčního rámce řídicího slovesa kontroly – atribut *control* má hodnotu funktoru tohoto valenčního doplnění;
- ostatní případy (tj. pokud takový člen valenčního rámce řídicího slovesa neexistuje) – atribut *control* má hodnotu *ex*.

Příklady:

- *pokusit se* (např. *Jiří.ACT se pokusí přijít*) – *control*: ACT;

- *slyšet* (např. *děti.ACT slyší někoho.PAT přicházet.EFF*) – control: PAT;
- *doporučit* (např. *doporučili mu.ADDR jít.PAT k lékaři*) – control: ADDR;
- *jít* (např. *jde to udělat*, ve smyslu *lze to udělat*) – control: ex.

## 5.2 Reflexivita

Nepovinný atribut *reflexivity* (značka  $\text{R}$ ) udává možnou syntaktickou funkci reflexivního morfému *se/si*, který je v češtině (kromě jiného, viz poznámku níže) formálním prostředkem pro vyjádření následujících syntaktických konstrukcí:

- **sekundární diateze:** částice *se* je součástí reflexivní formy slovesné (viz Karlík et al., 2002), a tedy součástí tvaru tzv. reflexivního pasiva:
  - pro tranzitivní slovesa (slovesa s akuzativní vazbou) (např. *připravovat – plány se připravují*, *bojovat – bojovala se těžká bitva*); atribut  $\text{R}$  má hodnotu *pass*;
  - pro intranzitivní slovesa (např. *pátrat – pátrá se po zloději*, *chodit – v neděli se chodí do kostela*, *bojovat – s nepřáteli se nakonec nebojovalo*); atribut  $\text{R}$  má hodnotu *pass0*;
- **gramatická koreference:** zájmeno *se/si* zaujímá pozici valenčního doplnění, které je koreferenční se jménem v subjektu a vyjadřuje, že subjekt vykonává děj sám na sobě; jde o tzv. vlastní reflexiva:
  - je-li příslušná valenční pozice zaplňovaná doplněním s akuzativní formou (a jde tedy o formu zájmena *se*), má atribut  $\text{R}$  hodnotu *cor4* (např. *mýt se (= sebe)*, *vidět se (= sebe)*, *darovat se (= sebe)*, kde *se* je patient (PAT) koreferující s aktorem (ACT) řídicího slovesa *mýt*, *vidět* a *darovat*);
  - pro valenční doplnění s dativní formou (a tedy formou zájmena *si*) má atribut  $\text{R}$  hodnotu *cor3* (např. *darovat si (dort) (= sám sobě)*, kde *si* je adresát (ADDR) řídicího slovesa *darovat* koreferující s aktorem (ACT) tohoto slovesa).

*VALLEX* se omezuje na zachycení případů, kdy zájmeno *se/si* zaplňuje pozici aktantu s akuzativní nebo dativní formou.<sup>13</sup>

*Poznámka:* Atribut reflexivity se netýká případů, kdy je morfém *se/si* součástí slovesného lemmatu (tyto případy jsou popsány v odd. 2.1) nebo kdy je *se/si* příznakem reciprocit (těm je věnován následující odd. 5.3).

## 5.3 Reciprocita

*Reciprocitou* se rozumí možnost vyjádření vztahu vzájemnosti mezi dvěma (či více) valenčními doplněními, přičemž vztah mezi těmito doplněními je symetrický (doplnění přitom splňují jisté sémantické podmínky), viz Karlík et al. (2002).

Pokud je do vztahu reciprocit zapojen aktor (ACT), užívá se reflexivní (zvrtné) sloveso, reciproční doplnění se potom vyjadřují jako koordinované členy podmětu (*Petr a Marie se hádali*) nebo podmět plurálový (*přátelé se navštěvují*); reciprocita může být zdůrazněna příslovci *spolu*, *navzájem* apod.

Pokud do vztahu reciprocit není zapojen aktor (ACT), reciproční vztah typicky vyplývá z koordinace či plurálové formy doplnění (např. *seznámil je, seznámil Jana a Marii*), konstrukce může být opět zdůrazněna příslovci *spolu*, *navzájem* apod.

Možnost recipročního užití je ve *VALLEXu* vyznačena v atributu reciprocit (značka  $\text{rcp}$ ), jehož hodnotou jsou dvojice, příp. trojice funktořů identifikující doplnění, která mohou vstupovat do vztahu reciprocit

<sup>13</sup> *VALLEX* tedy nepokrývá případy, kdy se zájmeno *se/si* může vyskytovat v jiném pádu (např. *praštil sebou o postel*) či v předložkové skupině (např. *ode dneška děláme na sebe.PAT, nechali si to u sebe.LOC*).

(např. ACT-ADDR pro *hádat se – neustále se spolu hádali*, ACT-ADDR-PAT pro *mluvit – mluví spolu o sobě (navzájem)*).

V případech odvozených reflexiv (viz odd. 2.1), která je možno klasifikovat jako inherentně reciproční varianty slovesa, typicky s obligatorním doplněním s formou s+7 (viz Panevová, 2007; Panevová – Mikuřová, 2007), je ve VALLEXu uváděna reciprocita u nereflexivního i reflexivního lexému.

VALLEX se omezuje na zachycení případů reciprocity, do které vstupují aktanty a obligatorní volná doplnění.

#### 5.4 Syntakticko-sémantické třídy

Část lexikálních jednotek (2 903 z celkového počtu 6 460, tedy přibližně 45 % všech lexikálních jednotek) má určenu *syntakticko-sémantickou třídu* (značka *class*). Tyto třídy byly budovány striktně ‚zdola nahoru‘ – seskupováním lexikálních jednotek s podobnými syntaktickými vlastnostmi, přičemž se přihlíželo k jejich sémantice. Zdůrazněme zde, že syntakticko-sémantické třídy jsou tvořeny jednotlivými lexikálními jednotkami, nikoliv celými lexémy – víceznačný lexém se tedy může vyskytovat v několika třídách.

Bylo vytvořeno následujících 22 syntakticko-sémantických tříd:

- appoint verb (23 LU), např. *nominovat, určovat* (ve smyslu *určovala své zástupce*), *ustanovovat*, ...;
- cause motion (43 LU), např. *hýbat* (*hýbat pravou rukou*), *mávat, vrhat*, ...;
- combining (96 LU), např. *míchat* (*míchat žlutky s moukou v těsto*), *přidávat, spojovat*, ...;
- communication (364 LU), např. *číst, hovořit, nařizovat*, ...;
- contact (115 LU), např. *dotýkat se, narážet, tisknout*, ...;
- emission (22 LU), např. *pouštět* (ve smyslu *tričko pouštělo barvu*), *vysílat* (ve významu *vysílat signály*), ...;
- exchange (177 LU), např. *dávat, dostávat, měnit, platit, pronajímat*, ...;
- expansion (19 LU), např. *pronikat, šířit*, ...;
- extent (20 LU), např. *činit* (ve smyslu *činí to 30 Kč*), *dosahovat, vycházet* (ve smyslu *boty vycházejí na tisíc korun*), ...;
- change (318 LU), např. *budovat, klesat* (ve smyslu *teplota prudce klesala*), *proměňovat, růst*, ...;
- intervention (10 LU), např. *zasahovat, mluvit* (*do toho nemůžu mluvit*), ...;
- location (399 LU), např. *doplňovat* (*doplňovat zboží do regálu*), *nacházet, shromažďovat*, ...;
- mental action (304 LU), např. *cítit se* (ve smyslu *cítit se dobře*), *jásat, mrzet*, ...;
- modal verb<sup>14</sup> (15 LU), např. *dovést* (ve smyslu *dovede plavat*), *chtít, moci, smět*, ...;
- motion (309 LU), např. *běžet, dorážet, hýbat se* (*Nehýbej se!*), ...;
- perception (104 LU), např. *hledět, pamatovat, všímat si*, ...;
- phase of action (80 LU), např. *končit* (*přednáška končí v 5 hodin*), *vrcholit, vznikat*, ...;
- phase verb (76 LU), např. *iniciovat, končit* (*končit školu*), *najet* (ve smyslu *najeli aspoň 500 mil*), ...;
- providing (51 LU), např. *naplnit* (ve smyslu *naplnit vanu vodou*), *vybavovat*, ...;
- psych verb (83 LU), např. *klamat, potěšit*, (ve smyslu *potěšila ho dárkem, dárek ho potěšil*), ...;

<sup>14</sup> Ve zpracování modálních sloves, která jsou na pomezí gramatiky a lexika, se VALLEX odchyľuje od teorie FGP. Ve FGP jsou modální slovesa *muset, mít, chtít, hodlat, moci, dát se, smět, dovést* a *umět* zachycena pomocí gramatémů u významových sloves (nemají tedy valenční rámec). Naproti tomu ve VALLEXu jsou kvůli úplnosti a lexikální proměnlivosti pro modální význam těchto sloves vyčleněny LU, nejsou však zachyceny všechny jejich syntaktické zvláštnosti. Protože některá z těchto sloves jsou víceznačná (např. *mít* je modální v užití *Jan má připravit večeři*, ale plnovýznamové v užití *Jan má spoustu peněz*), mohou být popsána v několika LU.



- social interaction (86 LU), např. *potkávat se* (*potkává se s přáteli v baru*), *spojovat* (*spojím se s ním co nejdříve*), *souhlasit*, ...;
- transport (189 LU), např. *donášet*, *přemíst'ovat/přemíst'ovat*, *shrnovat*, ...

Upozorňujeme, že toto rozdělení lexikálních jednotek do syntakticko-sémantických tříd je pouze pracovní a nelze je považovat za klasifikaci splňující požadavky dobře definované ontologie. Je zřejmé, že takto hrubé rozdělení není syntakticky ani sémanticky homogenní, jde o základní vymezení skupin sloves, které je potřeba dále podrobně studovat. Motivací pro tuto předběžnou klasifikaci lexikálních jednotek byla skutečnost, že i takovéto pracovní třídění zachycuje vztahy mezi slovesy a díky tomu usnadňuje kontroly konzistence slovníku a dovoluje formulovat obecnější pozorování týkající se slovníkových dat.

## 5.5 Frazémy a idiomy

Při vytváření slovníku *VALLEX* byl kladen důraz především na úplné pokrytí primárních a obvyklých významů sloves. Zároveň bylo zpracováno mnoho lexikálních jednotek popisujících okrajová a idiomatická užití sloves; jejich pokrytí však není (a nemůže být) úplné. Takové lexikální jednotky jsou odlišeny značkou *idiom* za číslem lexikální jednotky.

Idiomatická užití sloves jsou taková ustálená užití, která jsou pracovní charakterizována buď podstatným posunem ve významu (vzhledem k primárnímu významu, např. *přišel o hodinky*), omezenou, obvykle velmi malou množinou možných lexikálních hodnot, kterých můžou jejich doplnění nabývat (např. *brát roha*, *mráz mi z toho běhal po zádech*), nebo jinými nepravidelnostmi a anomáliemi.

*Poznámka:* Metaforické užití slovesa – pokud nedošlo k jeho výrazné lexikalizaci – je obvykle pokryto lexikální jednotkou pro primární význam slovesa (například *po městě šla řeč, že se budeš stěhovat* je řazeno do lexikální jednotky slovesa *jít* popisující význam ‚pohybovat se po vlastních nohou; přemísťovat se chůzí‘).



## Seznam značek

<i>biasp</i>	obouvidové sloveso (viz odd. 2.2)
<i>class</i>	syntakticko-sémantická třída (viz odd. 5.4)
<i>cont</i>	závislá věta obsahová (viz odd. 2)
<i>control</i>	atribut kontroly (viz odd. 5.1)
<i>cor3</i>	gramatická koreference s dativní formou (viz. odd. 5.2)
<i>cor4</i>	gramatická koreference s akuzativní formou (viz. odd. 5.2)
ČNK	příklad pochází z Českého národního korpusu (Blatná – Čermák, 2005)
<i>ex</i>	hodnota atributu kontrola (viz odd. 5.1)
<i>idiom</i>	idiomatická lexikální jednotka (viz odd. 5.5)
<i>impf</i>	nedokonavé sloveso (viz odd. 2.2)
<i>inf</i>	infinitiv (viz odd. 2)
<i>iter</i>	iterativní sloveso (viz odd. 2.2)
<i>jen</i>	omezení lemmat, která reprezentují lexikální jednotku (viz odd. 2.2, 2.3)
<i>opt</i>	fakultativní valenční nebo kvazivalenční doplnění (viz odd. 4.3)
<i>pass</i>	sekundární diateze u tranzitivního slovesa (viz. odd. 5.2)
<i>pass0</i>	sekundární diateze u intransitivního slovesa (viz. odd. 5.2)
PDT	příklad pochází z Pražského závislostního korpusu (Hajič et al., 2006)
<i>pf</i>	dokonavé sloveso (viz 2.2)
př.	příklad užití lexikální jednotky
<i>rcp</i>	možné typy recipročních konstrukcí (viz 5.3)
<i>rfl</i>	možné typy reflexivních konstrukcí (viz 5.2)
SSČ	příklad pochází ze Slovníku spisovné češtiny pro školu a veřejnost (Mejstřík, 2003)
SSJČ	příklad pochází ze Slovníku spisovného jazyka českého (Havránek, 1964)
<i>typ</i>	typické volné doplnění (viz odd. 4.3)

*Poznámka:* Neopakujeme zde znovu výčty hodnot funktorů (uvedeno v odd. 4.1) a syntakticko-sémantických tříd (uvedeno v odd. 5.4).



## Seznam literatury

- APRESJAN, J. D. *Eksperimental'noje issledovanije semantiki russkogo glagola*. Moskva, Nauka, 1967.
- BABKO-MALAYA, O. et al. Proposition Bank II: Delving Deeper. In MEYERS, A. (ed.) *HLT-NAACL 2004 Workshop: Frontiers in Corpus Annotation*, s. 17–23, Boston, USA, 2004.
- BLATNÁ, R. – ČERMÁK, F. (ed.). *Jak využívat Český národní korpus*. Praha, Nakladatelství Lidové noviny, 2005.
- BOGUSLAVSKY, I. – IOMDIN, L. – SIZOV, V. Multilinguality in ETAP-3: Reuse of Lexical Resources. In *Proceedings of PostCOLING Workshop on Multilingual Linguistic Resources*, 2004.
- BOND, F. – SHIRAI, S. Practical and Efficient Organization of a Large Valency Dictionary. In *Proceedings of the 4th Natural Language Processing Pacific*, Phuket, Thailand, 1997.
- CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht, Foris, 1981.
- CINKOVÁ, S. From PropBank to EngValLex. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, s. 2170–2175. ELRA, 2006.
- CRUSE, D. A. *Lexical Semantics*. Cambridge, Cambridge University Press, 1986.
- DANEŠ, F. Větné členy obligatorní, potenciální a fakultativní. *Miscellanea Linguistica*. 1971, s. 131–138.
- DANEŠ, F. *Věta a text*. Praha, Academia, 1985.
- DANEŠ, F. The Sentence-Pattern Model of Syntax. In LUELSDORFF, P. A. (ed.) *The Prague School of Structural and Functional Linguistics*, s. 197–221. Philadelphia, John Benjamins Publishing Company, 1994.
- DANEŠ, F. – HLAVSA, Z. *Větné vzorce v češtině*. Praha, Academia, 1987.
- DANEŠ, F. – GREPL, M. – HLAVSA, Z. (ed.). *Mluvnice češtiny 3*. Praha, Academia, 1987.
- DORR, B. J. et al. LCS Verb Database, Online Software Database of Lexical Conceptual Structures and Documentation. Technical report, University of Maryland, 2001.
- DOWTY, D. *Word meaning and Montague grammar. The semantics of verbs and times in Generative Semantics and in Montague's PTQ: Synthese Language Library*. Dordrecht, Reidel, 1979.
- ELLSWORTH, M. et al. PropBank, SALSA, and FrameNet: How Design Determines Product. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004), Workshop on Building Lexical Resources from Semantically Annotated Corpora*, Lisbon, 2004.
- ERK, K. et al. Towards a Resource for Lexical Semantics: A Large German Corpus with Extensive Semantic Annotation. In *Proceedings of ACL-03*, Sapporo, Japan, 2003.
- FILIPEC, J. – ČERMÁK, F. *Česká lexikologie*. Praha, Academia, 1985.
- FILLMORE, C. J. The Case for Case. In BACH, E. – HARMS, R. T. (ed.) *Universals in Linguistic Theory*, s. 1–88. New York, Holt, Rinehart and Winston, 1968.
- FILLMORE, C. J. Types of lexical information. In KIEFER, F. (ed.) *Studies in syntax and semantics*, s. 109–137. New York, Kluwer Academic Publishers, 1969.
- FILLMORE, C. J. FrameNet and the Linking between Semantic and Syntactic Relations. In TSENG, S.-C. (ed.) *Proceedings of COLING 2002*, s. xxviii–xxxvi. Howard International House, 2002.
- FILLMORE, C. J. – BAKER, C. – SATO, H. Seeing Arguments through Transparent Structures. In RODRÍGUEZ, M. G. – ARAUJO, C. P. S. (ed.) *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002)*, s. 787–791. ELRA, 2002.
- GREPL, M. – KARLÍK, P. *Skladba češtiny*. Olomouc, Votobia, 1998.
- HAIJČ, J. Complex Corpus Annotation: The Prague Dependency Treebank. In ŠIMKOVÁ, M. (ed.) *Insight into Slovak and Czech Corpus Linguistics*, s. 54–73. Bratislava, Veda, 2005.
- HAIJČ, J. et al. PDT-VALLEX: Creating a Large-coverage Valency Lexicon for Treebank Annotation. In *Proceedings of The Second Workshop on Treebanks and Linguistic Theories*, 9, s. 57–68, 2003.
- HAIJČ, J. et al. *Prague Dependency Treebank 2.0*. Philadelphia, PA, USA, Linguistic Data Consortium, 2006.
- HAVRÁNEK, B. (ed.). *Slovník spisovného jazyka českého*. Praha, Academia, 1964.
- HELBIG, G. – SCHENKEL, W. *Wörterbuch zur Valenz und Distribution deutscher Verben*. Leipzig, Bibliographisches Institut, 1969.

- HLAVÁČKOVÁ, D. – HORÁK, A. VerbaLex – New Comprehensive Lexicon of Verb Valencies for Czech. In *Computer Treatment of Slavic and East European Languages*, s. 107–115, Bratislava, 2006.
- JACKENDOFF, R. S. *Semantic Structures*. Cambridge, MIT Press, 1990.
- KARLÍK, P. Hypotéza modifikované valenční teorie. *Slovo a slovesnost*. 2000, 61, s. 170–189.
- KARLÍK, P. – NEKULA, M. – PLESKALOVÁ, J. (ed.). *Encyklopedický slovník češtiny*. Praha, Nakladatelství Lidové noviny, 2002.
- KARLÍK, P. – NEKULA, M. – RUSÍNOVÁ, Z. (ed.). *Příruční mluvnice češtiny*. Praha, Nakladatelství Lidové noviny, 1996.
- KIPPER, K. – DANG, H. T. – PALMER, M. Class-Based Construction of a Verb Lexicon. In *AAAI-2000 Seventeenth National Conference on Artificial Intelligence*, Austin, TX, 2000.
- KIPPER-SCHULER, K. *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis, Computer and Information Science Dept., University of Pennsylvania, Philadelphia, PA, 2005.
- KIRCHMEIER-ANDERSEN, S. *Lexicon, Valency and the Pronominal Approach. An Application of the Pronominal Approach to Danish Verbs and Nouns*. PhD thesis, Odense Universitet, 1997.
- KOMÁREK, M. et al. (ed.). *Mluvnice češtiny 2*. Praha, Academia, 1986.
- LEVIN, B. C. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago and London, The University of Chicago Press, 1993.
- LOPATKOVÁ, M. – PANEVOVÁ, J. Recent developments in the theory of valency in the light of the Prague Dependency Treebank. In ŠIMKOVÁ, M. (ed.) *Insight into Slovak and Czech Corpus Linguistics*, s. 83–92. Bratislava, Veda, 2006.
- LOPATKOVÁ, M. – ŽABOKRTSKÝ, Z. – SKWARSKA, K. Valency Lexicon of Czech Verbs: Alternation-Based Model. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, 3, s. 1728–1733. ELRA, 2006.
- MEJSTRÍK, V. (ed.). *Slovník spisovné češtiny pro školu a veřejnost*. Praha, Academia, 2003.
- MEL'ČUK, I. A. *Dependency Syntax: Theory and Practice*. Albany, State University of New York Press, 1988.
- MEL'ČUK, I. A. Actants in semantics and syntax I: actants in semantics. *Linguistics*. 2004a, 42 (1), s. 1–66.
- MEL'ČUK, I. A. Actants in semantics and syntax II: actants in syntax. *Linguistics*. 2004b, 42 (2), s. 247–291.
- MEL'ČUK, I. A. – ZHOLKOVSKY, A. K. *Explanatory Combinatorial Dictionary of Modern Russian*. Vienna, Wiener Slawistischer Almanach, 1984.
- MIKULOVÁ, M. et al. Anotace na tektogramatické rovině Pražského závislostního korpusu. Anotátorská příručka. Technical Report TR-2005-28, ÚFAL MFF UK, Praha, 2005.
- NIŽNÍKOVÁ, J. – SOKOLOVÁ, M. *Valenčný slovník slovenských slovies*. Prešov, Filozofická fakulta Prešovskej univerzity, 1998.
- PALA, K. – ŠEVEČEK, P. Valence českých sloves. In *Sborník prací FFBU*, s. 41–54, Brno, 1997.
- PANEVOVÁ, J. On Verbal Frames in Functional Generative Description. *The Prague Bulletin of Mathematical Linguistics*. 1974, 22, s. 3–40.
- PANEVOVÁ, J. *Formy a funkce ve stavbě české věty*. Praha, Academia, 1980.
- PANEVOVÁ, J. More Remarks on Control. *Prague Linguistic Circle Papers*. 1996, 2, s. 101–120.
- PANEVOVÁ, J. Ještě k teorii valence. *Slovo a slovesnost*. 1998, 59, s. 1–13.
- PANEVOVÁ, J. Znovu o recipitě. *Slovo a slovesnost*. 2007, 68, s. 91–100.
- PANEVOVÁ, J. Some Issues of Syntax and Semantics of Verbal Modifications. In *Proceedings of the First International Conference on Meaning-Text Theory (MTT 2003)*, s. 139–146, Paris, 2003. Ecole Normale Supérieure.
- PANEVOVÁ, J. Valency Frames and the Meaning of the Sentence. In LUELSENDORFF, P. A. (ed.) *The Prague School of Structural and Functional Linguistics*, s. 223–243. Amsterdam, Philadelphia, John Benjamins Publishing Company, 1994.
- PANEVOVÁ, J. – MIKULOVÁ, M. On Reciprocity. *The Prague Bulletin of Mathematical Linguistics*. 2007, 87, s. 27–40.
- PANEVOVÁ, J. – ŘEZNÍČKOVÁ, V. K možnému pojetí všeobecnosti aktantu. In HLADKÁ, Z. – KARLÍK, P. (ed.) *Čeština - univerzália a specifika 3*, s. 139–146, 2001.

- PANEVOVÁ, J. – SKOUMALOVÁ, H. Surface and Deep Cases. In *Proceedings of COLING 1992*, s. 885–889, Nantes, France, 1992.
- PAULINY, E. *Štruktúra slovenského slovesa*. Bratislava, Slovenská akadémia vied a umení, 1943.
- PETR, J. et al. (ed.). *Mluvnice češtiny I*. Praha, Academia, 1986.
- POLAŃSKI, K. (ed.). *Słownik syntaktyczno-generatywny czasowników polskich*. Wrocław, Wydawnictwo Polskiej Akademii Nauk, 1980–1992.
- POPOVA, M. *Kraták valenten rečnik na glagolite v sávremennia bălgarski knižoven ezik*. Sofia, Bulgarian Academy of Sciences Publishing House, 1987.
- PUSTEJOVSKY, J. *The Generative Lexicon*. Cambridge, MIT Press, 1995.
- RUPPENHOFER, J. et al. *FrameNet II: Extended Theory and Practice* (<http://framenet.icsi.berkeley.edu/>), 2006.
- SGALL, P. Valence jako jádro jazykového systému. *Slovo a slovesnost*. 2006, 67, s. 163–178.
- SGALL, P. *Generativní popis jazyka a česká deklinace*. Praha, Academia, 1967.
- SGALL, P. Teorie valence a její formální zpracování. *Slovo a slovesnost*. 1998, 59, s. 15–29.
- SGALL, P. – HAJIČOVÁ, E. – PANEVOVÁ, J. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Dordrecht, Reidel, 1986.
- SILNICKIJ, G. *Korreljacionnaja tipologija glagolnych sistem indoevropskich i inostrukturych jazykov*. Smolensk, Russia, Rossijskaja akademija nauk, Institut lingvističeskich issledovanij, 1999.
- SKOUMALOVÁ, H. *Czech syntactic lexicon*. PhD thesis, Univerzita Karlova, Filozofická fakulta, 2001.
- ŠMILAUER, V. *Novočeská skladba*. Praha, Nakladatel Ing. Mikuta, 1947.
- ŠMILAUER, V. *Novočeská skladba*. Praha, SPN, 1966.
- SVOZILOVÁ, N. – PROUZOVÁ, H. – JIRSOVÁ, A. *Slovesa pro praxi*. Praha, Academia, 1997.
- SVOZILOVÁ, N. – PROUZOVÁ, H. – JIRSOVÁ, A. *Slovník slovesných, substantivních a adjektivních vazeb a spojení*. Praha, Academia, 2005.
- TESNIÈRE, L. *Eléments de syntaxe structurale*. Paris, Librairie C. Klincksieck, 1959.
- EYNDE, K. – MERTENS, P. La valence: l'approche pronominale et son application au lexique verbal. *French Language Studies*. 2003, s. 63–104.
- ŽABOKRTSKÝ, Z. *Valency Lexicon of Czech Verbs*. PhD thesis, Charles University, Prague, 2005.

