VALLEX 1.0

Valency Lexicon of Czech Verbs

Markéta Lopatková, Zdeněk Žabokrtský,

Karolína Skwarska, Václava Benešová

Contents

Part I

DESCRIPTION

Chapter 1

Introduction

The Valency Lexicon of Czech Verbs, Version 1.0 (VALLEX 1.0) is a collection of linguistically annotated data and documentation, resulting from an attempt at formal description of valency frames of Czech verbs. VALLEX 1.0 was developed at the Center of Computational Linguistics, Faculty of Mathematics and Physics, Charles University, Prague, with the support of the project MSMT LN00A063.

VALLEX 1.0 is closely related to the Prague Dependency Treebank (PDT) project. The Functional Generative Description (FGD), being developed by Petr Sgall and his collaborators since the 1960s, is used as the background theory both in PDT and in VALLEX 1.0. In PDT, FGD is being verified by a complex annotation of large amounts of textual data, whereas in VALLEX 1.0 it is used only for the description of valency frames of selected verbs.

VALLEX 1.0 contains roughly 1400 verbs (counting only perfective and imperfective verbs, but not their iterative counterparts). They were selected as follows: (1) We started with about 1000 most frequent Czech verbs, according to their number of occurrences in a part of the Czech National Corpus (only '*být*' (to be) and some modal verbs were excluded from this set, because of their non-trivial status on the tectogrammatical level of FGD). (2) Then we added their perfective or imperfective aspectual counterparts, if they were missing; in other words, the set of verbs in VALLEX 1.0 is closed under the relation of 'aspectual pair'.

The preparation of the first version of VALLEX has taken more than two years. Although it is still a work in progress requiring further linguistic research, we believe that it will be useful or at least interesting for other researchers in the field. That is why we make it available to them, as well as to anyone else using it for non-commercial purposes, already now.

The whole VALLEX 1.0 can be downloaded from the Internet after filling the on-line registration form at the following address: http://ckl.mff.cuni.cz/zabokrtsky/vallex/1.0/

From the very beginning, VALLEX 1.0 was designed with an emphasis on both human and machine readability. Therefore both linguists and developers of applications within the Natural Language Processing domain can use and critically evaluate its content (of course, any feedback from them will be a valuable source of information to us, as well as a great motivation for further work). In order to satisfy different needs of these different potential users, VALLEX 1.0 contains the data in the following three formats:

- Browsable version. HTML version of the data allows for an easy and fast navigation through the lexicon. Verbs and frames are organized in several ways, following various criteria.

- Printable version. It is identical with Part II of this technical report.

- XML version. Programmers can run sophisticated queries (e.g. based on XPATH query language) on this machine-tractable data, or use it in their applications.

Finally, we would like to thank for an extensive linguistic and also technical advice to our colleagues from CKL and UFAL, especially to professor Jarmila Panevová.

Chapter 2

Logical Structure of the VALLEX Data

Remark on terminology: The terms used here either belong to the broadly accepted linguistic terminology, or come from the Functional Generative Description (FGD), which we have used as the background theory, or are defined somewhere else in this text.

Warning: The primary goal of this text is to explicitly describe the content of VALLEX 1.0 data from the structural point of view. Linguistic issues requiring a long explanation or discussion are mostly left apart – full analysis of all issues related to valency goes far beyond the scope of this text. VALLEX 1.0 is closely related to the Prague Dependency Treebank (PDT) project, basic knowledge about the PDT is useful for full understanding of VALLEX 1.0.

The description of the logical structure of VALLEX data is divided as follows (the graphical layout is depicted on Fig. 2.1):

- Word Entries – Section 2.1
- Lemmas – Section 2.2
- Lemma Variants – Section 2.3
- Homonyms – Section 2.4
- Frame Entries – Section 2.5
- Valency Frames – Section 2.6
- Functors – Section 2.7
- Morphemic Forms – Section 2.8
- Explicitly Declared Forms – Section 2.9
- Implicitly Declared Forms – Section 2.10
- Types of Complementations – Section 2.11
- Slot Expansion – Section 2.12
- Frame Attributes – Section 2.13
- Control – Section 2.14
- Class – Section 2.15
- Aspect, Aspectual Counterparts – Section 2.16
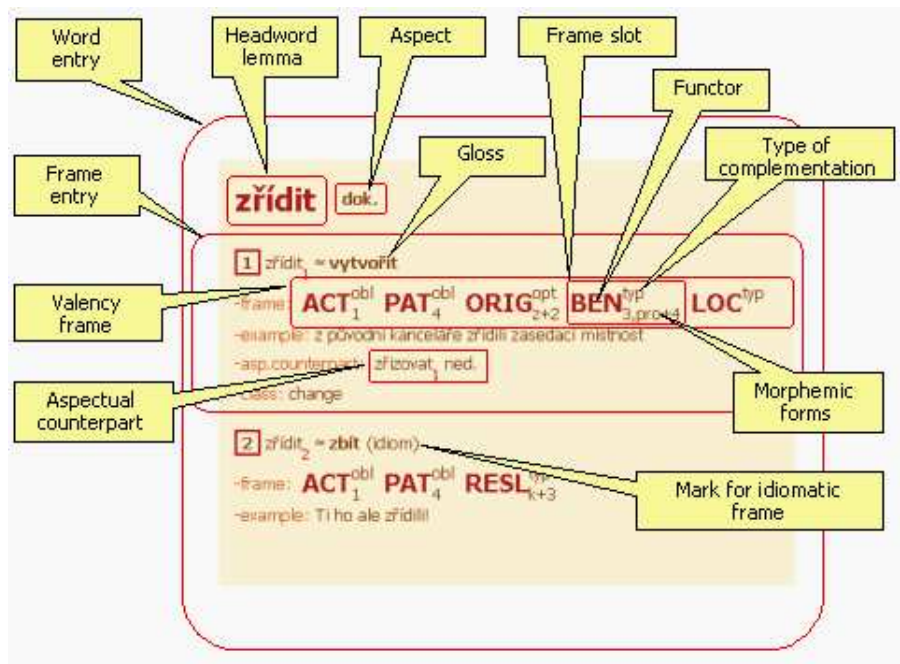- Idiomatic frames – Section 2.17

Figure 2.1: HTML layout of a word entry.

2.1 Word Entries

On the topmost level, VALLEX 1.0 is divided into word entries. Each word entry relates to one or more headword lemmas (Sec. 2.2). The word entry consists of a sequence of frame entries (Sec. 2.5) relevant for the lemma(s) in question (where each frame entry usually corresponds to one of the lemma's meanings). Information about the aspect (Sec. 2.16) of the lemma(s) is assigned to each word entry as a whole.

Most of the word entries correspond to lemmas in a simple one-to-one manner, but the following two non-trivial situations (and even combinations of them) appear as well in VALLEX 1.0:

- lemma variants (Sec. 2.3)
- homonyms (Sec. 2.4)

The content of a word entry roughly corresponds to the traditional term of lexeme.

2.2 Lemmas

Under the term of lemma (of a verb) we understand the infinitive form of the respective verb, in case of homonym (Sec. 2.4) followed by a Roman number in superscript (which is to be considered as an inseparable part of the lemma in VALLEX 1.0!).

Reflexive particles *se* or *si* are parts of the infinitive only if the verb is reflexive tantum, primary (e.g. *bát se*) as well as derived (e.g. *zabít se*, *šířit se*, *vrátit se*).

2.3   Lemma Variants

Lemma variants are groups of two (or more) lemmas that are interchangable in any context without any change of the meaning (e.g. *dovědět se/dozvědět se*). The only difference usually is just a small alternation in the morphological stem, which might be accompanied by a subtle stylistic shift (e.g. *myslet/myslit*, the latter one being bookish). Moreover, although the infinitive forms of the variants differ in spelling, some of their conjugated forms are often identical (*mysli* (imper.sg.) both for *myslet* and *myslit*).

The term 'lemma variants' should not be confused with the term 'synonymy'.

2.4   Homonyms

There are pairs of word entries in VALLEX 1.0, the lemmas of which have the same spelling, but considerably differ in their meanings (there is no obvious semantic relation between them). They also might differ as to their etymology (e.g. *nakupovat^I* - to buy vs. *nakupovat^{II}* - to heap), aspect (Sec. 2.16) (e.g. *stačit^I* pf. - to be enough vs. *stačit^{II}* impf. - to catch up with), or conjugated forms (*žilo* (past.sg.fem) for *žít^I* - to live vs. *žalo*(past.sg.fem) *žít^{II}* - to mow). Such lemmas (homonyms)[1] are distinguished by Roman numbering in superscript. These numbers should be understood as an inseparable part of lemma (Sec. 2.2) in VALLEX 1.0.

2.5   Frame Entries

Each word entry (Sec. 2.1) consists of a non-empty sequence of frame entries, typically corresponding to the individual meanings (senses) of the headword lemma(s) (from this point of view, VALLEX 1.0 can be classified as a Sense Enumerated Lexicon).

The frame entries are numbered within each word entry; in the VALLEX 1.0 notation, the frame numbers are attached to the lemmas as subscripts.

The ordering of frames is not completely random, but it is not perfectly systematic either. So far it is based only on the following weak intuition: primary and/or the most frequent meanings should go first, whereas rare and/or idiomatic meanings should go last. (We do not guarantee that the ordering of meanings in this version of VALLEX 1.0 exactly matches their frequency of the occurrences in contemporary language.)

Each frame entry[2] contains a description of the valency frame itself (Sec. 2.6) and of the frame attributes (Sec. 2.13).

---

[1]Note on terminology: we have adopted the term 'homonyms' from Czech linguistic literature, where it traditionally stands for what was stated above (words identical in the spelling but considerably different in the meaning); in English literature the term 'homographs' is sometimes used to express the same notion.

[2]Note on terminology: The content of 'frame entry' roughly corresponds to the term of lexical unit ('lexie' in Czech terminology).

2.6   Valency Frames

In VALLEX 1.0, a valency frame is modeled as a sequence of frame slots. Each frame slot corresponds to one (either required or specifically permitted) complementation[3] of the given verb.

The following attributes are assigned to each slot:

- functor (Sec. 2.7)
- list of possible morphemic forms (realizations) (Sec. 2.8)
- type of complementation (Sec. 2.11)

Some slots tend to systematically occur together. In order to capture this type of regularity, we introduced the mechanism of slot expansion (Sec. 2.12) (full valency frame will be obtained after performing these expansions).

2.7   Functors

In VALLEX 1.0, functors (labels of 'deep roles'; similar to theta-roles) are used for expressing types of relations between verbs and their complementations. According to FGD, functors are divided into inner participants (*actants*) and free modifications (this division roughly corresponds to the argument/adjunct dichotomy). In VALLEX 1.0, we also distinguish an additional group of quasi-valency complementations.

Functors which occur in VALLEX 1.0 are listed in the following tables (for Czech sample sentences see [3], page 43):

Inner participants:

- ACT (actor): *Peter read a letter.*
- ADDR (addressee): *Peter gave Mary a book.*
- PAT (patient): *I saw him.*
- EFF (effect): *We made her the secretary.*
- ORIG (origin): *She made a cake from apples.*

Quasi-valency complementations:

- DIFF (difference): *The number has swollen by 200.*
- OBST(obstacle): *The boy stumbled over a stumb.*
- INTT (intent): *He came there to look for Jane.*

Free modifications:

- ACMP (accompaniement): *Mother came with her children.*

---

[3]Note on terminology: in this text, the term 'complementation' (dependent item) is used in its broad sense, not related to the traditional argument/adjunct (complement/modifier) dichotomy (or, if you want, covering both ends of the dichotomy).

- AIM (aim): *John came to a bakery for a piece of bread.*
- BEN (benefactive): *She made this for her children.*
- CAUS (cause): *She did so since they wanted it.*
- COMPL (complement): *They painted the wall blue.*
- DIR1 (direction-from): *He went from the forest to the village.*
- DIR2 (direction-through): *He went through the forest to the village.*
- DIR3 (direction-to): *He went from the forest to the village.*
- DPHR (dependent part of a phraseme): *Peter talked horse again.*
- EXT (extent): *The temperatures reached an all time high.*
- HER (heritage): *He named the new villa after his wife.*
- LOC (locative): *He was born in Italy.*
- MANN (manner): *They did it quickly.*
- MEANS (means): *He wrote it by hand.*
- NORM (norm): *Peter has to do it exactly according to directions.*
- RCMP (recompense): *She bought a new shirt for 25 $.*
- REG (regard): *With regard to George she asked his teacher for advice.*
- RESL (result): *Mother protects her children from any danger.*
- SUBS (substitution): *He went to the theatre instead of his ill sister.*
- TFHL (temporal-for-how-long): *They interrupted their studies for a year.*
- TFRWH (temporal-from-when): *His bad reminiscences came from this period.*
- THL (temporal-how-long): *We were there for three weeks.*
- TOWH (temporal-to when): *He put it over to next Tuesday.*
- TSIN (temporal-since-when): *I have not heard about him since that time.*
- TWHEN (temporal-when): *His son was born last year.*

Note 1: Besides the functors listed in the tables above, also value DIR occurs in the VALLEX 1.0 data. It is used only as a special symbol for slot expansion (Sec. 2.12).

Note 2: The set of functors as introduced in FGD is richer than that shown above, moreover, it is still being elaborated within the Prague Dependency Treebank. We do not use its full (current) set in VALLEX 1.0 due to several reasons. Some functors do not occur with a verb at all (e.g. APP - appuertenace, '*my.*APP *dog*'), some other functors can occur there, but represent other than dependency relation (e.g. coordination, '*Jim or.*CONJ *Jack*'). And still others can occur with verbs as well, but their behaviour is absolutely independent of the head verb, thus they have nothing to do with valency frames (e.g. ATT - attitude, '*He did it willingly.*ATT').

2.8 Morphemic Forms

In a sentence, each frame slot can be expressed by a limited set of morphemic means, which we call forms. In VALLEX 1.0, the set of possible forms is defined either explicitly (Sec. 2.9), or implicitly

(Sec. 2.10). In the former case, the forms are enumerated in a list attached to the given slot. In the latter case, no such list is specified, because the set of possible forms is implied by the functor of the respective slot (in other words, all forms possibly expressing the given functor may appear).

2.9 Explicitly Declared Forms

The list of forms attached to a frame slot may contain values of the following types:

- **Pure (prepositionless) case.** There are seven morphological cases in Czech. In the VALLEX 1.0 notation, we use their traditional numbering: 1 - nominative, 2 - genitive, 3 - dative, 4 - accusative, 5 - vocative, 6 - locative, and 7 - instrumental.

- **Prepositional case.** Lemma of the preposition (i.e., preposition without vocalization) and the number of the required morphological case are specified (e.g., *z+2*, *na+4*, *o+6...*). The prepositions occurring in VALLEX 1.0 are the following: *bez, do, jako, k, kolem, kvůli, mezi, místo, na, nad, na úkor, o, od, ohledně, okolo, oproti, po, pod, podle, pro, proti, před, přes, při, s, u, v, ve prospěch, vůči, v zájmu, z, za*. ('*jako*' is traditionally considered as a conjunction, but it is included in this list, as it requires a particular morphological case in some valency frames).

- **Subordinating conjunction.** Lemma of the conjunction is specified. The following subordinating conjunctions occur in VALLEX 1.0: *aby, at', až, jak, zda,*[4] *že*.

- **Infinitive construction.** The abbreviation 'inf' stands for infinitive verbal complementation. 'inf' can appear together with a preposition (e.g. '*než+inf*'), but it happens very rarely in Czech.

- **Construction with adjectives.** Abbreviation 'adj-digit' stands for an adjective complementation in the given case, e.g. adj-1 (*Cítím se slabý* - I feel weak).

- **Constructions with '*být*'** . Infinitive of verb '*být*' (to be) may combine with some of the types above, e.g. *být+adj-1* (e.g. *zdá se to být dostatečné* - it seems to be sufficient).

- **Part of phraseme.** If the set of the possible lexical values of the given complementation is very small (often one-element), we list these values directly (e.g. '*napospas*' for phraseme '*ponechat napospas*' - to expose).

2.10 Implicitly Declared Forms

If no forms are listed explicitly for a frame slot, then the list of possible forms implicitly results from the functor of the slot according to the following (yet incomplete) lists:

- LOC: adverb, na+6, v+6, u+2, před+7, za+7, nad+7, pod+7, okolo+2, kolem+2, při+6, vedle+2, mezi+7, mimo+4, naproti+3, podél+2 . . .

- MANN: adverb, 7, na+4, . . .

- DIR3: adverb, na+4, v+4, do+2, před+4, za+4, nad+4, pod+4, vedle+2, mezi+4, po+4, okolo+2, kolem+2, k+3, mimo+4, naproti+3 . . .

- DIR1: adverb, z+2, od+2, zpod+2, zpoza+2, zpřed+2 . . .

---

[4]Note: form '*zda*' is in fact an abbreviation for couple of conjunctions '*zda*' and '*jestli*'.

- DIR2: adverb, 7, přes+4, podél+2, mezi+7, ...
- TWHEN: adverb, 2, 4, 7, před+7, za+4, po+6, při+6, za+2, o+6, k+3, mezi+7, v+4, na+4, na+6, kolem+2, okolo+2, ...
- THL: adverb, 4, 7, po+4, za+4, ...
- EXT: adverb, 4, na+4, kolem+2, okolo+2, ...
- REG: adverb, 7, na+6, v+6, k+3, při+6, ohledně+2, nad+7, na+4, s+7, u+2, ...
- TFRWH: z+2, od+2, ...
- AIM: k+3, na+4, do+2, pro+4, proti+3, aby, ať, že, ...
- TOWH: na+4 ...
- TSIN: od+2 ...
- TFHL: na+4, pro+4, ...
- NORM: podle+2, v duchu+2, po+6, ...
- MEANS: 7, v+6,na+6,po+6, z+2, že, s+7, na+4, za+4, pod+7, do+2, ...
- CAUS: 7, za+4, z+2, kvůli+2, pro+4, k+3, na+4, že, ...

2.11 Types of Complementations

Within the FGD framework, valency frames (in a narrow sense) consist only of inner participants (both obligatory[5] and optional, 'obl' and 'opt' for short) and obligatory free modifications; the dialogue test was introduced by Panevová as a criterium for obligatoriness. In VALLEX 1.0, valency frames are enriched with quasi-valency complementations. Moreover, a few non-obligatory free modifications occur in valency frames too, since they are typically ('typ') related to some verbs (or even to whole classes of them) and not to others. (The other free modifications can occur with the given verb too, but are not contained in the valency frame, as it was mentioned above (Sec. 2.7))

The attribute 'type' is attached to each frame slot and can have one of the following values: 'obl' or 'opt' for inner participants and quasi-valency complementations, and 'obl' or 'typ' for free modifications.

2.12 Slot Expansion

Some slots tend systematically to occur together. For instance, verbs of motion can be often modified with direction-to and/or direction-through and/or direction-from modifier. We decided to capture this type of regularity by introducing the abbreviation flag for a slot. If this flag is set (in the VALLEX 1.0 notation it is marked with an upward arrow), the full valency frame will be obtained after slot expansion.

If one of the frame slots is marked with the upward arrow (in the XML data, attribute 'abbrev' is set to 1), then the full valency frame will be obtained after substituting this slot with a sequence of slots as follows:

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

---

[5]It should be emphasized that in this context the term obligatoriness is related to the presence of the given complementation in the deep (tectogrammatical) structure, and not to its (surface) deletability in a sentence (moreover, the relation between deep obligatoriness and surface deletability is not at all straightforward in Czech).

- $\uparrow\text{DIR1}^{obl} \rightarrow \text{DIR1}^{obl} \text{ DIR2}^{typ} \text{ DIR3}^{typ}$
- $\uparrow\text{DIR2}^{obl} \rightarrow \text{DIR1}^{typ} \text{ DIR2}^{obl} \text{ DIR3}^{typ}$
- $\uparrow\text{DIR3}^{obl} \rightarrow \text{DIR1}^{typ} \text{ DIR2}^{typ} \text{ DIR3}^{obl}$
- $\uparrow\text{TSIN}^{obl} \rightarrow \text{TSIN}^{obl} \text{ THL}^{typ} \text{ TTIL}^{typ}$
- $\uparrow\text{THL}^{typ} \rightarrow \text{TSIN}^{typ} \text{ THL}^{typ} \text{ TTIL}^{typ}$

2.13   Frame Attributes

In VALLEX 1.0, frame attributes (more exactly, attribute-value pairs) are either obligatory or optional. The former ones have to be filled in every frame. The latter ones might be empty, either because they are not applicable (e.g. some verbs have no aspectual counterparts), or because the annotation was not finished (e.g. attribute class (Sec. 2.15) is filled only in roughly one third of frames).

Obligatory frame attributes:

- gloss – verb or paraphrase roughly synonymous with the given frame/meaning; this attribute is not supposed to serve as a source of synonyms or even of genuine lexicographic definition – it should be used just as a clue for fast orientation within the word entry!
- example – sentence(s) or sentence fragment(s) containing the given verb used with the given valency frame.

Optional frame attributes:

- control (Sec. 2.14)
- class (Sec. 2.15)
- aspectual counterparts (Sec. 2.16)
- idiom flag (Sec. 2.17)

2.14   Control

The term 'control' relates in this context to a certain type of predicates (verbs of control)[6] and two correferential expressions, a 'controller' and a 'controllee'. In VALLEX 1.0, control is captured in the data only in the situation where a verb has an infinitive modifier (regardless of its functor). Then the controllee is an element that would be a 'subject' of the infinitive (which is structurally excluded on the surface), and controller is the co-indexed expression. In VALLEX 1.0, the type of control is stored in the frame attribute 'control' as follows:

- if there is a coreferential relation between the (unexpressed) subject ('controllee') of the infinitive verb and one of the frame slots of the head verb, then the attribute is filled with the functor of this slot ('controller');
- otherwise (i.e., if there is no such co-reference) value 'ex.' is used.

---

[6]Note on terminology: in English literature the terms 'equi verbs' and 'raising verbs' are used in a similar context.

Examples:

- *pokusit se* (to try) - control: ACT
- *slyšet* (to hear), e.g. 'slyšet někoho přicházet' (to hear somebody come) - control: PAT
- *jít*, in the sense 'jde to udělat' (it is possible to do it) - control: ex

2.15   Class

Some frames are assigned semantic classes like 'motion', 'exchange', 'communication', 'perception', etc. However, we admit that this classification is tentative and should be understood merely as an intuitive grouping of frames, rather than a properly defined ontology.

The motivation for introducing such semantic classification in VALLEX 1.0 was the fact that it simplifies systematic checking of consistency and allows for making more general observations about the data.

2.16   Aspect, Aspectual Counterparts

Perfective verbs (in VALLEX 1.0 marked as 'pf.' for short) and imperfective verbs (marked as 'impf.') are distinguished between in Czech; this characteristic is called aspect. In VALLEX 1.0, the value of aspect is attached to each word entry as a whole (i.e., it is the same for all its frames and it is shared by the lemma variants, if any).

Some verbs (i.e. *informovat*, *charakterizovat*) can be used in different contexts either as perfective or as imperfective (obouvidová slovesa, 'biasp.' for short).

Within imperfective verbs, there is a subclass of of iterative verbs (iter.). Czech iterative verbs are derived more or less in a regular way by affixes such as *-va-* or *-iva-*, and express extended and repetitive actions (e.g. *čítávat*, *chodívat*). In VALLEX 1.0, iterative verbs containing double affix *-va-* (e.g. *chodívávat*) are completely disregarded, whereas the remaining iterative verbs occur as aspectual counterparts in frame entries of the corresponding non-iterative verbs (but have no own word entries, still).

A verb in its particular meaning can have aspectual counterpart(s) - a verb the meaning of which is almost the same except for the difference in aspect (that is why the counterparts constitute a single lexical unit on the tectogrammatical level of FGD; however, each of them has its own word entry in VALLEX 1.0, because they have different morphemic forms). The aspectual counterpart(s) need not be the same for all the meanings of the given verb, e.g., *odpovědět* is a counterpart of *odpovídat* - to answer, but not of *odpovídat* - to correspond. Therefore the aspectual counterparts (if any) are listed in frame attribute 'asp. counterparts' in VALLEX 1.0. Moreover, for perfective or imperfective counterparts, not only the lemmas are specified within the list, but (more specifically) also the frame numbers of the counterpart frames (which is of course not the case for the iterative counterparts, for they have no word entries of their own as stated above).

One frame might have more than one counterpart because of two reasons. Either there are two counterparts with the same aspect (impf. *působit* and impf. *způsobovat* for pf. *způsobit*), or there are two counterparts with different aspects (impf. *scházet*, pf. *sejít*, iter. *scházívat*).

2.17   Idiomatic frames

When building VALLEX 1.0, we focused mainly on primary or usual meanings of verbs. We also noted many frames corresponding to peripheral usages of verbs, however their coverage in VALLEX is not exhaustive. We call such frames idiomatic and mark them with label 'idiom'. An idiomatic frame is tentatively characterized either by a substantial shift in meaning (with respect to the primary sense), or by a small and strictly limited set of possible lexical values in one of its complementations, or by occurence of another types of irregularity or anomaly.

References[7]

[1] Eva Hajičová, Jarmila Panevová, Petr Sgall. *A Manual for Tectogrammatical Tagging of the Prague Dependency Treebank.* ÚFAL/CKL Technical Report TR-2000-09, 2000.

[2] Miroslav Grepl and Petr Karlík. *Skladba češtiny.* Votobia, Olomouc, 1998.

[3] Markéta Lopatková, Zdeněk Žabokrtský, Karolína Skwarska, Vendula Benešová: *Tektogramaticky anotovaný valenční slovník českých sloves.* Technical Report ÚFAL/CKL TR-2002-15. 2002.

[4] Markéta Lopatková: *Valency in the Prague Dependency Treebank: Building the Valency Lexicon.* PBML 79-80, 2003.

[5] Markéta Straňáková-Lopatková, Zdeněk Žabokrtský: *Valency Dictionary of Czech Verbs: Complex Tectogrammatical Annotation.* In: LREC 2002, Proceedings, vol. III, eds. M.G.Rodríguez and C.P.S.Araujo, ELRA, pp. 949-956, 2002.

[6] *Mluvnice češtiny II.* Academia, Praha, 1986.

[7] *Mluvnice češtiny III.* Academia, Praha, 1987.

[8] Karel Pala and Pavel Ševeček. Valence českých sloves. In *Sborník prací FFBU*, pages 41–54, Brno, 1997.

[9] Jarmila Panevová. *On verbal frames in functional generative description.* In The Prague Bulletin of Mathematical Linguistics 22, pages 3–40. 1974.

[10] Jarmila Panevová. *Valency Frames and the Meaning of the Sentence.* In: Ph.L.Luelsdoft (ed.): The Prague School of Structured and Functional Linguistics, Amsterdam-Philadephia, John Benjamins, pp. 223-243. 1994.

[11] Petr Sgall, Eva Hajičová, and Jarmila Panevová. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects.* D. Reidel Publishing Company, Dordrecht, 1986.

[12] Hana Skoumalová. *Verb frames extracted from dictionaries.* The Prague Bulletin of Mathematical Linguistics 77, 2002.

[13] Naďa Svozilová, Hana Prouzová, and Anna Jirsová. *Slovesa pro praxi.* Academia, Praha, 1997.

[14] *Slovník spisovné češtiny pro školu a veřejnost.* Academia, Praha, 1978.

[15] Vladimír Šmilauer. *Novočeská skladba.* SPN, 1966.

---

[7]Only a few references are given here. For more references, see [3].

Part II

DATA