

# **Annotation on the tectogrammatical level in the Prague Dependency Treebank**

## **Reference book**



**Marie Mikulová  
Allevtina Bémová  
Jan Hajič  
Eva Hajíčová  
Jiří Havelka  
Veronika Kolářová  
Lucie Kučová  
Markéta Lopatková  
Petr Pajas  
Jarmila Panevová  
Magda Ševčíková  
Petr Sgall  
Jan Štěpánek  
Zdeňka Urešová  
Kateřina Veselá  
Zdeněk Žabokrtský**

**Technical editor: Jiří Semecký**

**Translation: Kateřina Součková**

**Edited by Marie Mikulová**

**Edited from Mikulová et al.: Annotation on the tectogrammatical level in the  
Prague Dependency Treebank. Annotation manual.TR-2006-30, ÚFAL MFF UK,  
Prague, 2006.**

---

## **Annotation on the tectogrammatical level in the Prague Dependency Treebank: Reference book**

by Marie Mikulová, Allevtina Bémová, Jan Hajič, Eva Hajičová, Jiří Havelka, Veronika Kolářová, Lucie Kučová, Markéta Lopatková, Petr Pajas, Jarmila Panevová, Magda Ševčíková, Petr Sgall, Jan Štěpánek, Zdeňka Urešová, Kateřina Veselá, and Zdeněk Žabokrtský

Technical editor: Jiří Semecký

Translation: Kateřina Součková

---

# Table of Contents

Acknowledgements .....	vi
Introduction .....	vii
1. Typographical conventions .....	vii
1. Basic principles for representing sentences at the tectogrammatical level .....	1
2. Relation between the tectogrammatical level and the lower levels .....	4
3. Node types .....	5
4. Tectogrammatical lemma (t-lemma) .....	7
4.1. Basic form of the word .....	7
4.2. T-lemma of derived expressions .....	7
4.2.1. Grammatemes used for representing derivations .....	8
4.2.2. Derivation t-lemma + functor .....	9
4.2.3. Derivation t-lemma + grammatememe .....	10
4.2.4. Derivation t-lemma + functor + grammatememe .....	12
4.3. Multi-word t-lemma .....	16
4.4. Representative t-lemma .....	17
4.5. T-lemma substitute .....	18
5. Complex nodes and grammatememes .....	20
5.1. Semantic parts of speech .....	20
5.1.1. Semantic nouns .....	21
5.1.2. Semantic adjectives .....	24
5.1.3. Semantic adverbs .....	25
5.1.4. Semantic verbs .....	27
5.2. Grammatemes .....	27
5.3. The sentmod attribute .....	32
6. Sentence representation structure .....	34
6.1. Dependency .....	34
6.1.1. Dual dependency .....	34
6.1.1.1. Predicative complement .....	34
6.1.2. Non-dependency edges .....	37
6.1.3. Ambiguous dependency .....	40
6.2. Valency .....	44
6.2.1. Criteria for distinguishing different kinds of modifications .....	44
6.2.2. Argument shifting principle .....	45
6.2.3. Relations between verb meanings and valency frames .....	46
6.2.4. Valency frames and the way they are recorded in the valency lexicon .....	47
6.2.4.1. Surface form of a valency modification .....	47
6.2.4.2. Valency frames of idiomatic expressions and complex predicates .....	49
6.2.4.3. Valency lexicon .....	50
6.2.5. Representing valency in the tectogrammatical trees .....	51
6.3. Clauses (governing, dependent, verbal, non-verbal) .....	51
6.3.1. Verbal and non-verbal clauses .....	51
6.3.1.1. Verbal clauses .....	51
6.3.1.2. Non-verbal clauses .....	53
6.3.2. Dependent and independent clauses (clause connecting) .....	54
6.3.3. Dependent verbal clauses (complex sentences) .....	57
6.3.3.1. Dependent verbal clauses without a finite verb form .....	57
6.3.3.2. Supporting elements .....	60
6.3.3.3. False dependent clauses .....	63
6.4. Parataxis .....	66
6.4.1. Coordination, apposition, mathematical operations and intervals .....	70
6.5. Parenthesis .....	72
6.6. Ellipsis .....	75
6.6.1. Ellipsis .....	76
6.6.1.1. Ellipsis of a governing meaning unit .....	76
6.6.1.2. Ellipsis of a dependent meaning unit .....	80
6.6.1.3. Specific elliptical constructions .....	83
6.6.2. Newly established nodes in more complex syntactic structures .....	91
6.6.3. Nodes representing negation with verbs .....	92
6.6.4. Survey of newly established nodes .....	93
7. Specific phenomena .....	94

7.1. Multi-word lexical units .....	94
7.1.1. Multi-word predicates .....	95
7.1.1.1. Modal predicates .....	95
7.1.1.2. Phase predicates .....	99
7.1.1.3. Quasi-modal and quasi-phase predicates .....	99
7.1.1.4. Complex predicates .....	100
7.1.1.5. Verbonominal predicates (the copula <i>být</i> ) .....	103
7.1.2. Idioms .....	104
7.2. Numbers and numerals .....	105
7.3. Identifying expressions .....	109
7.3.1. Identification structure .....	111
7.4. Foreign-language expressions .....	113
7.5. Direct speech .....	115
7.6. Text segments marked by graphic symbol .....	119
8. <b>Functors and subfunctors</b> .....	121
8.1. Functors for effective roots of independent clauses .....	121
8.2. Argument functors .....	121
8.3. Temporal functors .....	122
8.4. Locative and directional functors .....	124
8.5. Functors for causal relations .....	125
8.6. Functors for expressing manner and its specific variants .....	126
8.7. Functors for rhematizers, sentence, linking and modal adverbials .....	127
8.8. Functors for multi-word lexical units and foreign-language expressions .....	128
8.9. Functors for some specific (new) modifications .....	128
8.10. Specific adnominal functors .....	129
8.11. Functor for the predicative complement (COMPL) .....	130
8.12. Functors expressing the relations between members of paratactic structures .....	130
8.13. Further specification of the meaning of a functor .....	132
8.13.1. Subfunctors .....	132
8.13.2. Attribute for the meaning of “state” .....	138
9. <b>Coreference</b> .....	139
9.1. Representing coreference in the tectogrammatical trees .....	139
9.2. Grammatical coreference .....	140
9.2.1. Control .....	143
9.2.1.1. Types of control constructions .....	146
9.2.2. Textual coreference .....	150
10. <b>Topic-focus articulation</b> .....	152
10.1. Contextual boundness .....	152
10.2. Communicative dynamism .....	154
10.2.1. Basic rules for the nodes ordering in tectogrammatical trees .....	154
10.2.1.1. Paratactic structures and dependent clauses .....	157
10.2.2. Ordering of nodes in verb phrases .....	159
10.2.3. Ordering of nodes in noun phrases .....	160
10.2.4. Projectivity of tectogrammatical trees .....	161
10.3. Topic and focus .....	163
10.4. Rhematizers .....	163
A. <b>Attributes of nodes in tectogrammatical trees</b> .....	167
A.1. Attributes of the technical root .....	167
A.2. Attributes of the other nodes .....	167
A.2.1. a .....	167
A.2.1.1. a/lex.rf .....	167
A.2.1.2. a/aux.rf .....	167
A.2.2. compl.rf .....	168
A.2.3. coref_gram.rf .....	168
A.2.4. coref_special .....	168
A.2.5. coref_text.rf .....	168
A.2.6. deepord .....	168
A.2.7. functor .....	169
A.2.8. gram .....	171
A.2.8.1. gram/aspect .....	171
A.2.8.2. gram/degcmp .....	171
A.2.8.3. gram/deontmod .....	172

A.2.8.4. gram/dispmod .....	172
A.2.8.5. gram/gender .....	172
A.2.8.6. gram/indeftype .....	173
A.2.8.7. gram/iterativeness .....	173
A.2.8.8. gram/negation .....	173
A.2.8.9. gram/number .....	173
A.2.8.10. gram/numertype .....	174
A.2.8.11. gram/person .....	174
A.2.8.12. gram/politeness .....	174
A.2.8.13. gram/sempos .....	174
A.2.8.14. gram/verbmod .....	175
A.2.8.15. gram/resultative .....	175
A.2.8.16. gram/tense .....	175
A.2.9. id .....	176
A.2.10. is_dsp_root .....	176
A.2.11. is_generated .....	176
A.2.12. is_member .....	176
A.2.13. is_name_of_person .....	176
A.2.14. is_parenthesis .....	177
A.2.15. is_state .....	177
A.2.16. nodetype .....	177
A.2.17. quot .....	177
A.2.17.1. quot/set_id .....	177
A.2.17.2. quot/type .....	178
A.2.18. sentmod .....	178
A.2.19. subfunctor .....	178
A.2.20. t_lemma .....	179
A.2.21. tfa .....	180
A.2.22. val_frame.rf .....	180
Subject index .....	181

---

# Acknowledgements

This document was created with the financial support of the grant of the Ministry of Education, Youth and Sport MSM0021620838, LC 536, 1ET100300517, 1ET101120503, 1ET201120505, the Grant Agency of the Czech Republic: GA201/05/H014, GA405/05/0913 and the Grant Agency of the Charles University Prague UK352/2005, UK350/2005, UK375/2005.

---

# Introduction

This manual presents a short description of sentence representation at the tectogrammatical level in the Prague Dependency Treebank (PDT). It is aimed at those PDT users that look for a quick introduction into the used representation.

This reference book is a shortened version of an extensive manual called *Annotation on the tectogrammatical level in the Prague Dependency Treebank. Annotation manual.* (see Prague Dependency Treebank 2.0, CDROM, doc/manuals/en/t-layer/), which contains the complete and detailed description of sentence representation at the tectogrammatical level (it is necessary to consult the “big” manual for exact, detailed and complete information).

The chapters are organized in a way parallel to the way of sentence representation on the tectogrammatical level. The basic principles of the sentence representation are described in Chapter 1, **Basic principles for representing sentences at the tectogrammatical level**, this chapter also includes an explanation of some important notions used further in the manual. In Chapter 2, **Relation between the tectogrammatical level and the lower levels**, the relations between the tectogrammatical nodes and units of the lower levels are described. The following chapter, Chapter 3, **Node types**, divides the tectogrammatical nodes into different types. The next two chapters, Chapter 4, **Tectogrammatical lemma (t-lemma)** and Chapter 5, **Complex nodes and grammatemes**, are devoted to the attributes specifying the lexical unit (represented by a node). A description of the sentence structure follows (Chapter 6, **Sentence representation structure**). Chapter 7, **Specific phenomena** contains a description of certain specific phenomena. A separate chapter is devoted to the functors and sub-functors (Chapter 8, **Functors and subfunctors**). Coreference (Chapter 9, **Coreference**) and topic-focus articulation (Chapter 10, **Topic-focus articulation**) are each assigned one chapter. In the appendix (Appendix A, **Attributes of nodes in tectogrammatical trees**), there is a summary of all attributes.

## 1. Typographical conventions

**Examples.** The manual contains a number of examples illustrating the phenomena in question. The examples have a fixed form. They do not illustrate the structure; they only present the values of the attributes of individual words present or absent in the surface structure of the example sentence.

Example sentences are made-up and usually presented without context. The given annotation corresponds to the most common context in which the sentence can be used.

**NB!** Example sentences necessarily contain only the part that is to be illustrated (i.e. elided expressions do not have to be made visible if they are not the subject of the illustration).

Items represented by a single node in the sentence are underscored. The value of the relevant attribute is given in square brackets in the following form: the name of the attribute = the value of the attribute (if there are more possible values, they are all in the brackets, separated by a semicolon). If the example sentence is supposed to illustrate the values of just one node, the values are presented in square brackets after the example sentence. If there are more nodes whose values are to be illustrated, the values follow (in square brackets) immediately after the last underscored word represented by the given node. An exception to this are the functor values. If the functor values of individual nodes are to be illustrated, they always immediately follow the given word. Functors are not given in square brackets; they are separated from the word by a period. Examples:

*Upadl do nesnází.DIR3 na dlouhou dobu. [is\_state=1]* (=He got into difficulties for a long time.)

*Špičková cena.DENOM [is\_member=1] a.CONJ špičkový výkon.DENOM [is\_member=1]* (=lit. Top price and top performance.)

Words not expressed at the surface level (represented by newly established nodes) are given in curly brackets. The curly brackets always contain the t-lemma of the newly established node, which may but need not be followed by the values of selected attributes. For example:

{#PersPron.ACT} *Přijde.* (=lit. (He) will\_come.)

{#PersPron.ACT [tfa=t]} *Přijde.* (=lit. (He) will\_come.)

If it is necessary to stress that certain words are not represented by a separate node in the tectogrammatical tree, they are given in angle brackets <>. For example:

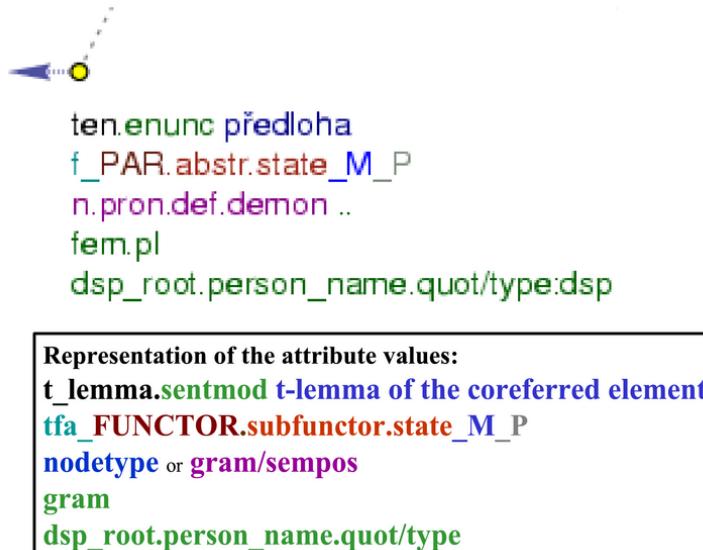
*Přijde jen.RHEM <tehdy>, <když> mu ustoupíš.TWHEN* (=He comes only in case you give in to him.)

**Example tectogrammatical trees.** For a number of example sentences, the corresponding example trees are included as well. Every example tree is equal to a complete analysis of the given sentence.

Tectogrammatical trees in PDT 2.0 make use of two different styles of representation (see Prague Dependency Treebank 2.0, CDROM, doc/tools/tred/PML\_mak.html). Example trees in the manual make use of the following (representation) settings (PML\_T\_Full template).

**Nodes.** The attribute values (below the node) are displayed in the order shown in Fig. 1 (the presented combination can never be found, of course).

**Figure 1. Node**



The attribute values are usually presented directly, without giving the name of the attribute first. Names of the attributes are only provided if the values are not unambiguous. The value of the attribute `quot/type` is always in the form: name of the attribute:its value. As for complex nodes (`nodetype=complex`), the value of the `nodetype` attribute is not specified; the value of the `gram/sempos` attribute is given directly instead.

The notation `state` is included in the list of the attribute values if the value of the `is_state` attribute is 1.

The notation `_M` is included if the value of the `is_member` attribute is 1.

The notation `_P` is included if the value of the `is_parenthesis` attribute is 1.

The notation `person_name` is included if the value of the `is_person_name` attribute is 1.

The notation `dsp_root` is included if the value of the `is_dsp_root` attribute is 1.

Nodes representing words present at the surface level are represented as little circles; newly established nodes (`is_generated=1`) are represented as little squares. (Certain information is carried by the color of the nodes as well: yellow means the node has the `f` value in the `tfa` attribute, green means `c` in the `tfa` attribute, white means `t` in the `tfa` attribute. Nodes with no value assigned in the `tfa` attribute are grey.)

**Edges.** Edges are the connecting lines between nodes. The edge between the technical root node of the tectogrammatical tree and the root node of the represented sentence and the edges between nodes with the PAR, PARTL, VOCAT, RHEM, CM, FPHR and PREC functors and their mother nodes (i.e. edges not representing dependencies; see Section 6.1.2, “Non-dependency edges”) are represented as thin interrupted lines.

The upper half of the edge between a paratactic structure root node and a terminal member of the paratactic structure is represented as a thin (grey) line; the lower half is represented as a thick (grey) line. The upper half of the edge between a paratactic structure root node (that is not a member of another paratactic structure) and its mother node is represented as a thick (grey) line; the lower half is represented as a thin (grey) line. The edge between a paratactic structure root node and the root of a shared modifier is represented as a thin (blue) line. The edge between a paratactic structure root node and a direct member of this structure that is a paratactic structure root node itself (in case of embedded paratactic structures) is represented as a thin (grey) line. (For more on paratactic structures, see Section 6.4, “Parataxis”.)

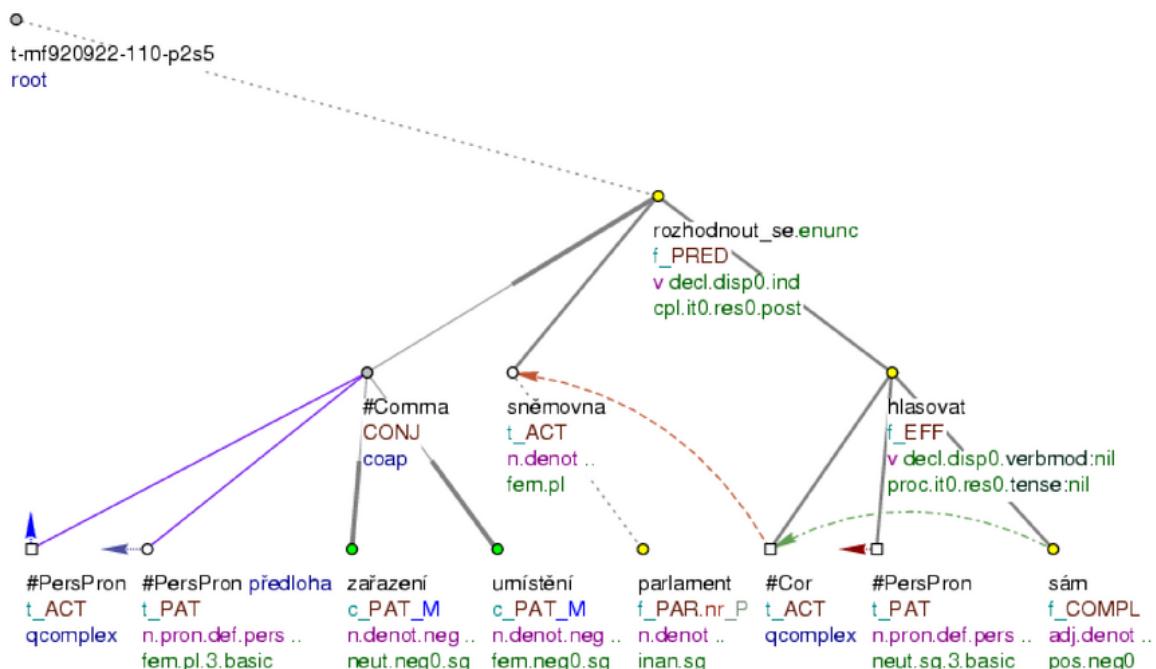
**References.** Attributes of the type reference, marking especially co-referential relations (cf. Chapter 9, *Coreference*), are represented as arrows going from one node to another. Grammatical coreference (the `coref_gram.rf` attribute) is represented by an (orange) interrupted arrow pointing to the co-referred node (starting at the co-referring node). Textual coreference (the `coref_text.rf` attribute) is represented by a (blue) dotted arrow pointing to

the co-referred node (starting at the co-referring node). If the co-referred node is not in the same tree as the co-referring node, the arrow is short and points either to the left or to the right of the node, depending on whether the co-referred node is in the preceding or following tree; next to the co-referring node, the t-lemma of the co-referred node is specified. (**NB!** Textual coreference relations crossing the boundaries of a single tectogrammatical tree are not represented in the example trees at all.) Reference to a segment (`coref_special=segm`) is represented as a short (red) arrow pointing to the left of the node. Exophoric reference (`coref_special=exoph`) is represented as a short (blue) arrow pointing upwards.

The second dependency with predicative complements (cf. Section 6.1.1.1, “Predicative complement”) is represented by a (green) semi-dotted arrow going from the node with the `COMPL` functor to the node representing the governing noun.

An example tree with different types of edges and reference (the illustrated sentence is artificial: *O jejich zařazení a umístění se rozhodl sám hlasovat samy.*) is to be found in Fig. 2.

**Figure 2. Edges and references**



**NB!** In example tectogrammatical trees (just like in the PDT trees; see Section 6.2.5, “Representing valency in the tectogrammatical trees”) the valency of nouns is not represented properly - with the exception of verbal nouns!

---

# Chapter 1. Basic principles for representing sentences at the tectogrammatical level

Natural language is an extraordinarily complex system; therefore, it is useful to decompose its description into several levels. Preceding (lower) levels of PDT were concerned with the morphological annotation (i.e. lemmas, tags, values of morphological categories; where the words are arranged in a linear way, without any structure) and the analytical (surface structure) annotation (dependency structure, analytical functions).

The tectogrammatical structure serves for representing the meaning structure of the sentence. This level is governed by the following principles:

- the basic unit of annotation at the tectogrammatical level is a sentence as a basic means of conveying meaning.
- for every well-formed (Czech) sentence, it is possible to provide its tectogrammatical representation: a tectogrammatical tree structure (tectogrammatical tree in sequel).
- in case of ambiguity, it is in theory possible to assign more tectogrammatical trees to one sentence. However, in PDT only one tree is assigned to every sentence, which corresponds to the reading assigned by the annotator.
- in case of synonymy, on the other hand, different sentences can be assigned a single tectogrammatical tree (however, it has to be a case of strict synonymy, i.e. the truth conditions have to be absolutely identical).

Tectogrammatical trees have these basic properties:

- a. **Tectogrammatical trees** are data structures the basis of which is formed by a rooted tree (in the sense of the theory of graphs): it consists of a set of nodes and a set of edges and one of the nodes is marked as the root of the tree.
- b. **Nodes** of the tectogrammatical tree represent the meaning units of the sentence.

These meaning units usually correspond to the expressions present at the surface in a one to one relation but one meaning unit can also correspond to more surface expressions (this is a case of e.g. prepositional phrases, which are represented by a single node). The meaning unit can also have no counterpart in the surface structure. Then, it is represented by an “artificial”, newly established node (such added nodes are used e.g. for representing omitted obligatory valency modifications). Function words (like prepositions, subordinating conjunctions, auxiliary verbs, supporting expressions) are usually not assigned separate nodes in the tree.

The number of the nodes in the analytical tree generally does not correspond to the number of the nodes in the tectogrammatical tree. For more on the relations between the tectogrammatical and analytical (surface structure) levels, see Chapter 2, *Relation between the tectogrammatical level and the lower levels*

Cf. Fig. 1.1: Examples of nodes representing expressions present in the surface structure of the sentence are: *starý* (=old), *sultán* (=sultan), *nový* (=new), *sultán* (=sultan), *vystřídal se* (=changed places). The prepositional phrase *na trůnu* (=on the throne) is represented by a single node (the preposition *na* is not assigned a separate node). In order to represent the coordination *starý sultán a nový sultán* (=the old and new sultan), the conjunction *a* (=and) is assigned a separate node. An example of a newly established node is the node representing the Patient (functor=PAT) of the verb *vystřídat se* (=exchange, replace) (the newly established node is displayed as a little square).

- c. **Attribute values** Each node is itself a complex unit with inner structure. It is possible to conceive of it as a set of attributes, more precisely as a set of ordered attribute - value pairs. Whether a given attribute is or is not present in a given node follows from the nodetype (see Chapter 3, *Node types*).

Node attributes can be divided into several groups. The basic attributes of a tectogrammatical tree node are the tectogrammatical lemma, grammatemes and functor. The tectogrammatical lemma expresses the lexical meaning of the node (see Chapter 4, *Tectogrammatical lemma (t-lemma)*). The grammatemes correspond mainly to (the meanings of) the morphological categories (see Chapter 5, *Complex nodes and grammatemes*). The functors capture the kind of syntactic dependency between autosemantic expressions, i.e. they correspond to syntactic functions (see Chapter 8, *Functors and subfunctors*). There are also attributes providing the information regarding the coreference (see Chapter 9, *Coreference*), topic - focus articulation and deep word order (see Chapter 10, *Topic-focus articulation*) of the sentence. The remaining attributes concern special properties of the structure and certain syntactic and semantic properties impossible to capture in any other way.

Attribute values are mostly sequences of symbols; the set of the sequences for a given attribute is usually fixed. A special kind of attributes are attributes for representing relations (most often coreferential) between nodes (meaning units), which go “across” the tree or even cross the tree boundaries.

Cf. Fig. 1.1: In the example tree, there is one attribute of the type reference, representing reciprocity (i.e. a grammatical coreference relation) between the Patient and Actor of the predicate *vystřídat se*. It is depicted as a curved (red) interrupted arrow.

For the list of all attributes, see Appendix A, *Attributes of nodes in tectogrammatical trees*.

- d. **Edges** of the tectogrammatical tree primarily capture the dependency relations between the nodes (more precisely between autosemantic expressions). Not every edge, though, represents a linguistic dependency (see Section 6.1.2, “Non-dependency edges”). Edges have no attributes of their own; attributes that actually belong to edges (e.g. the type of dependency) are presented as attributes of the corresponding nodes.

Cf. Fig. 1.1: Edges are represented as straight connecting lines between nodes. Edges representing dependency are marked by a thick grey line.

- e. tectogrammatical tree nodes are **in a linear order**; this linear order represents the deep word order of a sentence (see Chapter 10, *Topic-focus articulation*).

Also the following terms are used when talking about tectogrammatical trees (here explained only informally):

**Technical root node of a tectogrammatical tree.** The root node of a sentence is a node with no linguistic interpretation; it only serves technical purposes (it bears the sentence identifier; for the list of all attributes assigned to the root, see Section A.1, “Attributes of the technical root”). It has always exactly one daughter node. The root of a tree is called **technical root node of a tectogrammatical tree**. When talking about tectogrammatical tree nodes (further in the text), the technical root node is not considered (if not stated otherwise).

Cf. Fig. 1.1: The technical root of a tectogrammatical tree is the highest node in the tree; its only immediate daughter is connected to it by a thin dotted edge; the `nodeType` attribute is assigned the value `root`.

**Mother node.** Node X is the mother of node Y, if there is an edge between X and Y and if X is closer to the technical root node of the tree (i.e. if it is higher in the tree).

Cf. Fig. 1.1: The mother of the node representing the expression *(starý) sultán* is the node for the conjunction *a*.

**Immediate (direct) daughter node.** Node X is an immediate daughter of node Y if Y is the mother of X.

Since tectogrammatical trees make use of linear ordering, there are right and left daughter nodes. **Right (left) immediate daughter** of node M is such an immediate daughter that occurs to the right (left) of node M.

Cf. Fig. 1.1: Immediate daughter nodes of the node representing the verb *vystřídat se* are these three nodes: the node for the conjunction *a*, the newly established node for the Patient and the node for the prepositional phrase *na tránu*. All immediate daughter nodes of the verb *vystřídat se* are its left daughters.

**Governing/dependent node.** If nodes X and Y (or: the expressions represented by them) are in a dependency relation we say that X is the governing (or dependent) node of node Y. The governing node does not have to be the mother node of the dependent node (there can even be more governing nodes for a single node) and the dependent node does not have to be an immediate daughter of its governing node (see also Section 6.1, “Dependency”). (In the technical documentation for PDT, the terms “effective mother node” and “effective daughter node” are used for this type of relation).

Cf. Fig. 1.1: The governing node of the node for *starý* is the node for *sultán* (which is its mother node at the same time). The governing node of the node for *sultán* is the node representing the verb *vystřídat se* (which is not its mother node).

**Sister node.** Node X is a sister of node Y if they have the same mother.

Since tectogrammatical trees make use of linear ordering, there are right and left sister nodes. **Right (left) sister node** of node M is such a sister that occurs to the right (left) of node M.

Cf. Fig. 1.1: The sister nodes of the node for *a* are the newly established node for the Patient of *vystřídat se* and the node representing the prepositional phrase *na tránu*. All the sisters of the node representing the conjunction *a* are its right sisters.

**Path from node M.** For purposes of topic - focus articulation annotation, we also define the term right (left) path from node M and the rightmost (leftmost) path from node M.

**Right (left) path from node M** is such a path in the tree that starts at node M, goes downwards (towards the leaves) and ends in a node that has no right (left) immediate daughters. Node M is not part of the path.

**The rightmost (leftmost) path from node M** is such a right (left) path in the tree for which it holds that no node on the path has a right (left) sister.

Cf. Fig. 1.1: There is no right path going from the node for *vystřídat se*. As for the leftmost path from the node representing *vystřídat se*, it consists of the nodes for *a*, *sultán* and *starý*.

**Subtree of a tectogrammatical tree** is a continuous subgraph of a tectogrammatical tree (a subset of its nodes and edges with a marked root node).

**Root of a subtree** is the node of the subtree whose mother node (if existent) is not part of the subtree.

**Expression**. A linguistically relevant part of a sentence is called expression.

**Root of an expression** is short for the root of the subtree representing a given expression.

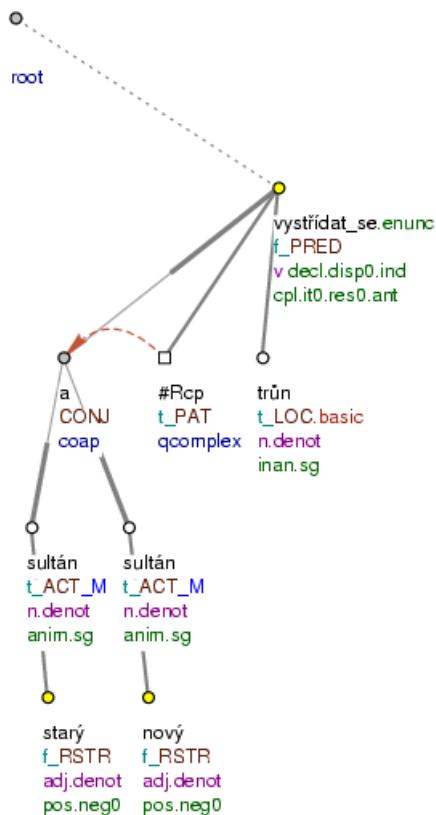
**Root of a sentence** is the root of the subtree corresponding to the whole sentence; i.e. it is the (only) immediate daughter of the technical root node of the tectogrammatical tree.

**Effective root of an expression** is the node that either has no governing node in the given tectogrammatical tree or whose governing node is not part of the subtree representing the expression. The effective root of an expression can be identical to the root of the expression; however, sometimes it is not, e.g. in the case of paratactic structures: the root node (there is only one root) is not identical to the effective root nodes (which are usually more than one).

Cf. Fig. 1.1: The root of the example sentence is the node for *vystřídat se*. This node is also the effective root of the sentence. The coordination *starý sultán a nový sultán* is represented by a subtree of the tectogrammatical tree; the root of the subtree (the root of the coordination) is the node representing the conjunction *a*, the effective root nodes are the two nodes representing the two occurrences of the noun *sultán*.

**Figure 1.1. Tectogrammatical tree**

*Starý sultán a nový sultán se vystřídal na trůnu.* (=lit. Old sultan and new sultan REFL changed on throne)



---

# Chapter 2. Relation between the tectogrammatical level and the lower levels

The relation between the nodes of the tectogrammatical and analytical levels (which is generally of the type M:N, the options 0:N and M:0 included) is captured by the `atree.rf` attribute of the technical root of the tectogrammatical tree and by the `a` attribute (whose value is a structure of a the `lex.rf` and `a/aux.rf` attributes) with other nodes.

**The `atree.rf` attribute.** The `atree.rf` attribute (see Section A.1, “Attributes of the technical root”) refers trivially to the technical root of the analytical tree corresponding to the tectogrammatical tree in question. It contains the identifier of the technical root node of the corresponding analytical tree.

**The `a/lex.rf` and `a/aux.rf` attributes.** The `a/lex.rf` attribute (see Section A.2.1.1, “`a/lex.rf`”) contains the identifier of the node at the analytical level from which the tectogrammatical node got its lexical meaning (or its biggest part). The `a/aux.rf` attribute (see Section A.2.1.2, “`a/aux.rf`”) contains the list of the identifiers of all analytical nodes that influence in some way or other the value of the functor, subfunctor or grammatemes of the tectogrammatical node (these are mostly analytical nodes representing so called function words like prepositions, conjunctions, auxiliaries and supporting expressions).

The values of the attributes `a/lex.rf` and `a/aux.rf` conform to the following rules:

- a. If the tectogrammatical node has no analytical counterpart, both attributes are empty.
- b. If a tectogrammatical node (other than a newly established one with one of the t-lemmas `#Forn`, `#Idph`, `#EmpVerb` or `#EmpNoun`) corresponds to exactly one analytical node, then the `a/lex.rf` attribute contains the reference to the (analytical) node and the `a/aux.rf` is empty.
- c. If a tectogrammatical node (other than a newly established one with one of the t-lemmas: `#Forn`, `#Idph`, `#EmpVerb` or `#EmpNoun`) corresponds to more analytical nodes, then the `a/lex.rf` attribute contains the reference to the node from which the tectogrammatical node got its lexical meaning (or its biggest part) and the `a/aux.rf` attribute contains the list of references to the other analytical nodes, which mostly represent function words (prepositions, subordinating conjunctions, auxiliaries etc.).
- d. With newly established nodes with the t-lemma `#EmpVerb` or `#EmpNoun`, the `a/lex.rf` attribute is always empty since the full verb they represent was not expressed at the surface level. If no function word was expressed at the surface level either, the `a/aux.rf` attribute is also empty; otherwise it contains the list of references to the relevant function words (e.g. auxiliaries that are part of a complex verb form where the full verb is not expressed).
- e. With newly established nodes with the t-lemmas `#Idph` and `#Forn`, which serve for putting parts of identifying and foreign-language expressions together, into a single list (`nodetype=list`), the `a/lex.rf` attribute is always empty. If a foreign-language or identifying expression is syntactically combined with one or more function words at the analytical level, then the references to these function words are contained in the `a/aux.rf` attribute.

**Copied nodes.** The b) and c) groups include also copied nodes (see Section 6.6, “Ellipsis”). The `a/lex.rf` and `a/aux.rf` attributes of the copied nodes contain the identifiers of the analytical nodes for the words present at the surface level that are relevant for the copied node, i.e. that influence its t-lemma, functor and other attributes.

**NB!** A unit of a lower level does not have to have a counterpart at the tectogrammatical level. Thus, various graphic symbols are ignored as well as the reflexive *se* when part of a reflexive passive and constructions with dispositional modality.

---

# Chapter 3. Node types

Tectogrammatical tree nodes are divided into eight groups; these are called **node types**. The node types are defined either on the basis of the t-lemma of the node, or on the basis of its functor, or both. For every node type, essentially the same rules regarding the node's immediate daughters apply.

The node type information is encoded in the value of the `nodetype` attribute (see Section A.2.16, “`nodetype`”). The `nodetype` attribute has eight possible values and applies to every node in a tectogrammatical tree.

For a survey of the node types and their definitions, see Table 3.1, “**Node types**”.

**Table 3.1. Node types**

Node type	<code>nodetype</code>	Definition of the node type	Properties of the immediate daughter nodes
<b>Technical root node of a tectogrammatical tree</b>	<code>root</code>	Artificial node with special attribute values; it includes the identifier of the sentence in the treebank.	Always exactly one immediate daughter, which is either a paratactic structure root node or the effective root node of an independent clause.
<b>Atomic node</b>	<code>atom</code>	Node with the functor: ATT CM INTE MOD PARTL PREC RHEM	As a rule, it has no immediate daughters.
<b>Paratactic structure root node</b>	<code>coap</code>	Node with the functor: ADVS APPS CONJ CONFR CONTRA CSQ DISJ GRAD OPER REAS	An immediate daughter can be: - terminal member of a paratactic structure - shared modifier - rhematizer of a shared modifier ( <code>functor = RHEM</code> ) - conjunction modifier ( <code>functor = CM</code> ) - root node of a(n embedded) paratactic structure ( <code>nodetype = coap</code> )
<b>List structure root node</b>	<code>list</code>	Node with the t-lemma: #Idph #Forn	An immediate daughter can be: - item of the list - modifier of the list  The items of a list under the node with the <code>#Idph</code> t-lemma form a tree structure. The items of a list under the node with the <code>#Forn</code> t-lemma are sisters with respect to each other and their functor is FPHR.
<b>Node representing a foreign-language expression</b>	<code>fphr</code>	Node with the functor: FPHR	It has no immediate daughters.
<b>Node representing the dependent part of an idiom</b>	<code>dphr</code>	Node with the functor: DPHR	As a rule, it has no immediate daughters.

<b>Quasi-complex node</b>	qcomplex	Node with the t-lemma: #Amp #Ast #AsMuch #Cor #EmpVerb #Equal #Gen #Oblfm #Percnt #QCor #Rcp #Some #Total #Unsp  Nodes with the t-lemma (if not of the type coap): #Bracket #Comma #Colon #Dash #Period #Period3 #Slash	Any node except for the types root and fphr.  Quasi-complex nodes are a special type of nodes that occupy the same positions (having the same functor) as complex nodes but they have no grammatemes.
<b>Complex node</b>	complex	All other nodes	Any node except for the types root and fphr.  (Only) complex nodes have grammatemes.

---

# Chapter 4. Tectogrammatical lemma (t-lemma)

**Tectogrammatical lemma**, **t-lemma** in the sequel, is one of the attributes of a tectogrammatical node (the `t_lemma` attribute); it represents the lexical content of the node.

The value of the `t_lemma` attribute can be:

- a. the basic (default) form of the word from which the node got its lexical meaning (see Section 4.1, “Basic form of the word”).

In case of synonymy, the t-lemma is the basic form of one of the variants (the so called representative t-lemma; see Section 4.4, “Representative t-lemma”),

- b. the basic form of the word from which the word represented by the node was derived (see Section 4.2, “T-lemma of derived expressions”),
- c. a so called multi-word t-lemma, consisting of more words (see Section 4.3, “Multi-word t-lemma”),
- d. an “artificial” value, a so called t-lemma substitute (see Section 4.5, “T-lemma substitute”).

## 4.1. Basic form of the word

With nodes representing words present at the surface level and with so called copied nodes (see Section 6.6, “Ellipsis”), the t-lemma usually corresponds to the basic (default) form of the word from which the node got its lexical meaning (i.e. to the basic form of the word at the analytical level the information about which is in the `a/lex.rf` attribute of the tectogrammatical node; see Chapter 2, *Relation between the tectogrammatical level and the lower levels*).

The basic word form is its representative form; i.e. the nominative sg. with nominals, infinitive with verbs, positive with gradable adjectives, positive forms with words that can be negated (as for nouns, only those ending with `-ní`, `-tí` and `-ost` are represented by their non-negated forms); long forms for short forms with adjectives). Morphological meanings are captured by grammateemes (see Section 5.2, “Grammateemes”).

**Exceptions.** Exceptions to this rule:

- a. **Personal and reflexive pronouns and their possessive counterparts** are represented by a node with the t-lemma substitute `#PersPron` (see also Section 4.5, “T-lemma substitute”). For example:

*Bratr prohlásil: Já tam nejdu. [t\_lemma = #PersPron]* (=My brother said: I'm not going there.)

*Pro ni udělám všechno. [t\_lemma = #PersPron]* (=For her I'll do anything.)

*Tvůj názor nesdílím.* [t\_lemma = #PersPron] (=I don't share your view.)

*Své názory ti nesdělím.* [t\_lemma = #PersPron] (=I'm not telling you my opinion.)

- b. **Frozen finite verb forms, frozen transgressives (gerunds) and infinitives** (with adverbial functions; see Section 6.3.3.1, “Dependent verbal clauses without a finite verb form”) are represented by a node whose t-lemma is identical to the surface form of the expression. For example:

*Dnes je, myslím, středa.* [t\_lemma = myslím] (=It's Wednesday, I think.)

*Soudě podle ministra zahraničí, bude to problém.* [t\_lemma = soudě] (=Judging with the Minister, it is going to be a problem.)

- c. **Foreign-language words** (with the FPHR functor; see Section 7.4, “Foreign-language expressions”) are also represented by a node with the t-lemma corresponding to the actual word form.

## 4.2. T-lemma of derived expressions

The main source of differences between the basic word form and the actual t-lemma is the tendency to represent certain kinds of derivation. Then, the t-lemma of the derived expression is the basic form of the base word (the

reference to the word in the analytical tree from which the tectogrammatical node got its lexical meaning is in the `a/lex.rf` attribute; see Chapter 2, *Relation between the tectogrammatical level and the lower levels*).

Word-formation (derivational) relations are represented in PDT 2.0 with deadjectival adverbs, possessive adjectives, pronouns, numerals and pronominal adverbs. The surface forms of these words are derived from the deep-level representation with the help of a set of features characterizing their semantic and syntactic properties (these are the t-lemma, functor and grammatemes). There are the following types of representing derivational relations, depending on the way the basic word form of the derivative is determined:

A. **t-lemma + functor.**

The basic word form of the derivative is inferred from the combination of the t-lemma and the functor assigned to it. For more on type A derivation, see Section 4.2.2, “Derivation t-lemma + functor”.

B. **t-lemma + grammateme(s).**

The basic form of the derivative follows from the combination of the t-lemma and the assigned grammateme(s) (see Section 4.2.1, “Grammatemes used for representing derivations”). For more on type B derivation, see Section 4.2.3, “Derivation t-lemma + grammateme”.

C. **t-lemma + functor + grammateme(s).**

The basic form of the derivative follows from the combination of the t-lemma and the assigned functor and grammateme(s) (see Section 4.2.1, “Grammatemes used for representing derivations”). For more on type C derivation, see Section 4.2.4, “Derivation t-lemma + functor + grammateme”.

## 4.2.1. Grammatemes used for representing derivations

For representing derivations of type B (t-lemma + grammateme) and C (t-lemma + functor + grammateme), the grammatemes `numertype` and `indeftype` are used. The values of these grammatemes reflect the derivational relations at the tectogrammatical level, which makes them different from other grammatemes, which are usually tectogrammatical counterparts of morphological categories.

**The `indeftype` grammateme.** The value of the `indeftype` grammateme (see Section A.2.8.6, “`gram/indeftype`”) captures the semantic feature in which the indefinite, interrogative, negative and totalizing pronouns (and pronominal adverbs) differ from their corresponding relative pronouns, by the t-lemma of which they are represented. We differ between:

- a. **relatives and interrogatives.** Relatives are those pronouns that are in a grammatical coreference relation with another item in the sentence structure (see Section 9.2, “Grammatical coreference”); they usually occur in dependent relative clauses. Interrogatives are those pronouns that do not corefer with anything; they are used in questions and content clauses. Relatives have the `indeftype` grammateme specified as `relat`, interrogatives have the `indeftype` grammateme filled with the value `inter`. For example:

*Muž, kterého jsme dnes potkali, byl můj bratr. [t\_lemma = ktery; indeftype = relat] (=The man (lit. which) we met today was my brother.) (který (=which) corefers with muž (=man))*

*Řekněte, který dům jste si koupili? [t\_lemma = ktery; indeftype = inter] (=Tell me, which house did you buy?)*

- b. **totalizers.** Totalizers are assigned one of the two values: `total1` or `total2`. The value `total1` is assigned to a node for a totalizer referring to a whole (of something); the value `total2` is assigned to a totalizer referring individually to every single item (in the set). For example:

*To je všechno. [t\_lemma = co; indeftype = total1] (=That's all.)*

*každý člověk [t\_lemma = ktery; indeftype = total2] (=every person)*

- c. **negatives.** Negatives are represented by the value `negat`. For example:

*Nikdy už to neudělám. [t\_lemma = kdy; indeftype = negat] (=I'll never do it again.)*

- d. **indefinites** There are several types of indefinites in Czech. The individual types differ from each other by the derivational means they make use of, which also entail subtle meaning differences (see Table 4.1, “**Derivational types of indefinites**”). The individual derivational types of indefinites are assigned one of the values: `indef1` through `indef6`.

**Table 4.1. Derivational types of indefinites**

indeftype	Derivational means	Example
indef1	<i>ně-</i>	<i>Někdo přišel.</i> [ <u>t_lemma</u> = <i>kdo</i> ; <u>indeftype</u> = <i>indef1</i> ] (=Somebody has come.)
indef2	<i>-si, -s</i>	<i>čísi chlapec</i> [ <u>t_lemma</u> = <i>kdo</i> ; <u>indeftype</u> = <i>indef2</i> ] (=somebody's boy)
indef3	<i>-koli; -koliv</i>	<i>jakýkoliv úkol</i> [ <u>t_lemma</u> = <i>jaký</i> ; <u>indeftype</u> = <i>indef3</i> ] (=any boy)
indef4	<i>leda-; lec-; ledas-</i>	<i>Může to být ledaskde.</i> [ <u>t_lemma</u> = <i>kde</i> ; <u>indeftype</u> = <i>indef4</i> ] (=It can be at various places.)
indef5	<i>kde-</i>	<i>Myslí si to kdekdo.</i> [ <u>t_lemma</u> = <i>kdo</i> ; <u>indeftype</u> = <i>indef5</i> ] (=Many people think about it this way.)
indef6	<i>málo-; sotva-; zřídka-; všeli-; nevím-; kdoví-; bůhví-; čertví-</i>	<i>Bůhvíkterý den to bylo.</i> [ <u>t_lemma</u> = <i>který</i> ; <u>indeftype</u> = <i>indef6</i> ] (=lit. God_knows_which day it was.)

**The grammeme numertype.** The value of the numertype grammeme (Section A.2.8.10, “gram/numertype”) expresses the semantic feature in which the given numeral is distinct from the corresponding cardinal numeral (by the t-lemma of which it is represented at the tectogrammatical level).

The values of the numertype grammeme are basic for cardinal numerals *frac* for fractions, *kind* for sort numerals, *ord* for ordinal numerals and *set* for set numerals. For example:

*Přišli jen tři.* [t\_lemma = *tři*; numertype = *basic*] (=Only three of them came.)

*Zdědil polovinu domu.* [t\_lemma = *dva*; numertype = *frac*] (=He inherited one half of the house.)

*Má dvojí občanství.* [t\_lemma = *dva*; numertype = *kind*] (=He has two citizenships.)

*Umístil se na třetím místě.* [t\_lemma = *tři*; numertype = *ord*] (=He came in third.)

*Ztratil už troje klíče.* [t\_lemma = *tři*; numertype = *set*] (=He has already lost three sets of keys.)

## 4.2.2. Derivation t-lemma + functor

The following derivatives are represented with the help of the t-lemma and functor:

- a. **possessive adjectives.** Possessive adjectives (*sempos* = *n. denot*) are thought of as derivatives of their corresponding nouns. They are represented by a node whose t-lemma is the corresponding noun in the nominative singular. For example:

*matčino.APP dítě* [t\_lemma = *matka*] (=mother's child)

*Pavlovo.APP auto* [t\_lemma = *Pavel*] (=Paul's car)

- b. **deadjectival adverbs.** Deadjectival adverbs (*sempos* = *adj. denot*) are represented by a node whose t-lemma is the corresponding adjective in the nominative singular. For example:

*Rychle.MANN odešel.* [t\_lemma = *rychlý*] (=He left quickly.)

*Pěkně.MANN zpívá* [t\_lemma = *pěkný*] (=He sings very well.)

- c. **definite pronominal adverbs with a directional meaning** (derived from the adverbs *tady*, *tam*). Definite pronominal directional adverbs (*sempos* = *adv.pron.def*) are represented by a node whose t-lemma is the corresponding locative adverb. The possible combinations of t-lemmas and functors are to be found in Table 4.2, “**Derivatives of the pronominal adverbs *tady* and *tam***”. For example:

*Odtud.DIR1 je vidět až za hranice.* [t\_lemma = *tady*] (=From here, you can see as far as beyond the border.)

- d. **definite pronominal adverbs with a temporal meaning** (derived from the adverbs *ted'*, *potom* and *tehdy*). Definite pronominal adverbs with various temporal meanings (*sempos = adv. pron. def*) are represented by a node whose t-lemma is the corresponding adverb with the basic temporal meaning (answering the question “*kdy (=when)*”). The possible combinations of t-lemmas and functors are to be found in Table 4.3, “**Derivatives of the pronominal adverbs *ted'*, *potom* and *tehdy***”. For example:

*Doposud.TTILL se nám dařilo dobře [t\_lemma = ted']* (=Until now we did well.)

**Table 4.2. Derivatives of the pronominal adverbs *tady* and *tam***

	t-lemma	
functor	<b>tady</b>	<b>tam</b>
<b>LOC</b>	<i>tady, tu, zde</i>	<i>tam</i>
<b>DIR1</b>	<i>odtud, odsud</i>	<i>odtamtu</i>
<b>DIR2</b>	<i>tudy</i>	<i>tamtudy</i>
<b>DIR3</b>	<i>sem, potud, posud</i>	<i>tam</i>

**Table 4.3. Derivatives of the pronominal adverbs *ted'*, *potom* and *tehdy***

	t-lemma		
functor	<b>ted'</b>	<b>potom</b>	<b>tehdy</b>
<b>TWHEN</b>	<i>ted', tu, nyní</i>	<i>potom, pak, poté</i>	<i>tehdy</i>
<b>TSIN</b>	<i>odted'</i>	-	<i>odtehdy</i>
<b>TTILL</b>	<i>doted', doposud, potud, posud</i>	-	<i>dotehdy</i>

### 4.2.3. Derivation t-lemma + grammateme

The following derivatives are represented with the help of the t-lemma and grammatemes *numertype* or *indef-type*:

- a. **definite non-cardinal numerals**. Definite fraction numerals (*sempos = n. quant. def*) and definite ordinal, sort and set numerals (*sempos = adj. quant. def*) are understood as derivatives of their corresponding cardinal numerals at the tectogrammatical level and they are represented by their t-lemmas. The fact that the given numeral is a fraction, ordinal, sort or set numeral can be inferred from the value of the *numertype* grammateeme. Examples of possible combinations of t-lemmas and functors are to be found in Table 4.4, “**Examples of (non-adverbial) derivatives of definite cardinal numerals**”. For example:

*Ztratil už troje klíče.* [t\_lemma = *tri*; *numertype* = *set*] (=He's already lost three sets of keys.)

- b. **indefinite adjectival numerals** (adjectival derivatives of the numeral *kolik*). Individual types of indefinite adjectival numerals, i.e. various adjectival derivatives of the numeral *kolik* (=how many/much) (*sempos = adj. quant. indef*) are represented by a node with the t-lemma *kolik*. Which numeral is represented by it can be inferred from the value of the grammatemes *numertype* and *indef-type*. Possible combinations of the *numertype* and *indef-type* grammateeme values with nodes whose t-lemma is *kolik* - i.e. all derivatives od this t-lemma - are in Table 4.5, “**Adjectival derivatives of the numeral *kolik***”. For example:

*Člověk může mít několikeré občanství.* [t\_lemma = *kolik*; *numertype* = *kind*; *indef-type* = *indef1*] (=One can have several citizenships.)

- c. **relative, indefinite, interrogative, negative and totalizing pronouns** (derivatives of the pronouns *kdo*, *co*, *který*, *jaky*). Relative, indefinite, interrogative, negative and totalizing pronouns (*sempos = n. pron. indef* or *sempos = adj. pron. indef*) make use of only four t-lemmas: *kdo* (=who), *co* (=what), *který* (=which) and *jaky* (=what). All other pronouns are taken to be their derivatives. Which pronouns are represented by which t-lemmas and which values of the *indef-type* grammateeme they get is summarized in Table 4.6, “**Derivatives of the pronouns *kdo*, *co*, *který* and *jaky***”. For example:

*Nikdo to dělat nebude.* [t\_lemma = *kdo*; *indef-type* = *negat*] (=Nobody is going to do it.)

- d. **relative, indefinite, interrogative, negative and totalizing pronominal adverbs** (derivatives of the adverbs *jak* and *proč*). Relative, indefinite, interrogative, negative and totalizing pronominal adverbs derived from the adverbs *jak* (=how) and *proč* (=why) (*sepos = adv.pron.indef*) are represented by nodes with the t-lemmas *jak* or *proč*. The basic word form of these derivatives follows from the value of the grammateeme *indeftype*. Which adverbs are represented by which t-lemmas and which values of the *indeftype* grammateeme they get is summarized in Table 4.7, “**Derivatives of the pronominal adverbs *jak* and *proč***”. For example:

*Bylo mu všelijak.* [t\_lemma = *jak*; indeftype = *indef6*] (=He felt strange (lit. in\_different\_ways).)

**Table 4.4. Examples of (non-adverbial) derivatives of definite cardinal numerals**

	t-lemma					
numertype	<i>tolik</i>	<i>jeden</i>	<i>dva</i>	<i>tři</i>	<i>čtyři</i>	<i>sto</i>
<b>frac</b>	-	<i>jednina</i>	<i>polovina</i>	<i>třetina</i>	<i>čtvrtina</i>	<i>setina</i>
<b>kind</b>	<i>tolikery</i>	-	<i>dvoji</i>	<i>trojí</i>	<i>čtvverý</i>	<i>stery</i>
<b>ord</b>	<i>tolikátý</i>	<i>první</i>	<i>druhý</i>	<i>třetí</i>	<i>čtvrtý</i>	<i>stý</i>
<b>set</b>	<i>tolikery</i>	-	<i>dvoje</i>	<i>troje</i>	<i>čtvvery</i>	<i>stery</i>

**Table 4.5. Adjectival derivatives of the numeral *kolik***

t_lemma: <b>kolik</b>	numertype			
indeftype	<b>basic</b>	<b>set</b>	<b>kind</b>	<b>ord</b>
<b>relat</b>	<i>kolik (klíčů), kolikery (dveře)</i>	<i>kolikery (klíče)</i>	<i>kolikery</i>	<i>kolikátý</i>
<b>indef1</b>	<i>několik (klíčů), několikery (dveře)</i>	<i>několikery (klíče)</i>	<i>několikery</i>	<i>několikátý</i>
<b>indef2</b>	-	-	-	-
<b>indef3</b>	-	-	-	-
<b>indef4</b>	-	-	-	-
<b>indef5</b>	-	-	-	-
<b>indef6</b>	<i>kdovíkolik (klíčů), kdovíkolikery (dveře), bůhvíkolik, četrvíkolik, nevímkolik, ...</i>	<i>kdovíkolikery (klíče), bůhvíkolikery, četrvíkolikery, nevímkolikery, ...</i>	<i>kdovíkolikery, bůhvíkolikery, četrvíkolikery, nevímkolikery, ...</i>	<i>kdovíkolikátý, bůhvíkolikátý, četrvíkolikátý, nevímkolikátý, ...</i>
<b>inter</b>	<i>kolik (klíčů), kolikery (dveře)</i>	<i>kolikery (klíče)</i>	<i>kolikery</i>	<i>kolikátý</i>
<b>negat</b>	-	-	-	-
<b>total1</b>	-	-	-	-
<b>total2</b>	-	-	-	-

**Table 4.6. Derivatives of the pronouns *kdo*, *co*, *který* and *jaký***

	t-lemma			
indeftype	<b>kdo</b>	<b>co</b>	<b>který</b>	<b>jaký</b>
<b>relat</b>	<i>kdo, kdož</i>	<i>co, což, oč, nač, zač</i>	<i>který, kterýž, jenž</i>	<i>jaky, jakýž</i>
<b>indef1</b>	<i>někdo</i>	<i>něco</i>	<i>některý</i>	<i>nějaký</i>
<b>indef2</b>	<i>kdosi, kdos</i>	<i>cosi, cos</i>	<i>kterýsi</i>	<i>jakýsi</i>
<b>indef3</b>	<i>kdokoli, kdokoliv</i>	<i>cokoli, cokoliv, cožkoli, cožkoliv</i>	<i>kterýkoli, kterýkoliv</i>	<i>jakýkoli, jakýkoliv</i>
<b>indef4</b>	<i>ledakdo, leckdo, ledakdos, ledaskdo</i>	<i>ledaco, lecco, leccos, ledacos, ledasco</i>	<i>leckterý, ledakterý</i>	<i>lecjaký, ledajaký</i>
<b>indef5</b>	<i>kdekdo</i>	<i>kdeco</i>	<i>kdekterý</i>	<i>kdejaký</i>
<b>indef6</b>	<i>málokdo, kdovíkdo, sotvakdo, zřídkakdo, všelikdo, nevímkdo, bůhvíkdo, četrvíkdo, ...</i>	<i>máloco, kdovíco, sotvaco, zřídkaco, všelico, všelicos, nevímco, bůhvíco, četrvíco, ...</i>	<i>málokterý, kdovíkterý, sotvakterý, zřídkakterý, všelikterý, nevímkterý, bůhvíkterý, četrvíkterý, ...</i>	<i>všelijaký, málojaký, kdovíjaký, sotvajaký, zřídkajaký, nevímjaký, bůhvíjaký, četrvíjaký, ...</i>
<b>inter</b>	<i>kdo, kdopak, kdožpak, kdože</i>	<i>co, copak, cožpak, cože, oč, nač, zač</i>	<i>který, kterýpak</i>	<i>jaký, jakýpak</i>
<b>negat</b>	<i>nikdo</i>	<i>nic</i>	<i>žádný</i>	<i>nijaký</i>
<b>total1</b>	<i>všechn, veškerý</i>	<i>všechen, všechno, vše</i>	-	-
<b>total2</b>	-	-	<i>každý</i>	-

**Table 4.7. Derivatives of the pronominal adverbs *jak* and *proč***

	t-lemma	
indeftype	<b>jak</b>	<b>proč</b>
<b>relat</b>	<i>jak</i>	<i>proč</i>
<b>indef1</b>	<i>nějak</i>	-
<b>indef2</b>	<i>jaksi</i>	-
<b>indef3</b>	<i>jakkoli, jakkoliv</i>	-
<b>indef4</b>	<i>lecjak, ledajak, ledasjak</i>	-
<b>indef5</b>	-	-
<b>indef6</b>	<i>všelijak, nevímjak, kdovíjak, bůhvíjak, čertvíjak, ...</i>	<i>kdovíproč, nevímproč, bůhvíproč, čertvíproč, ...</i>
<b>inter</b>	<i>jak, jakpak, jakže</i>	<i>proč, pročpak</i>
<b>negat</b>	<i>nijak</i>	-
<b>total1</b>	-	-
<b>total2</b>	-	-

#### 4.2.4. Derivation t-lemma + functor + grammame

The following derivatives are captured with the help of the t-lemma, the numertype or indeftype grammame and the functor:

- a. **possessive counterparts of relative, indefinite, interrogative, negative and totalizing pronouns** (derivatives of the pronoun *kdo*). Possessive counterparts of relative, indefinite, interrogative, negative and totalizing pronouns (*sempos = n.pron.indef*) make use of a single t-lemma: *kdo* (=who). Which pronouns are represented by which t-lemmas and which values of the indeftype grammame they get is summarized in Table 4.8, “**Possessive derivatives of the pronoun *kdo***”. For example:

*ničí.APP kniha [t\_lemma = kdo; indeftype = negat] (=nobody's book)*

- b. **relative, indefinite, interrogative, negative and totalizing pronominal adverbs with a directional meaning** (derivatives of the adverb *kde*). Relative, indefinite, interrogative, negative and totalizing pronominal adverbs with a directional meaning (*sempos = adv.pron.indef*) are represented by a node with the t-lemma of the corresponding locative adverb (*kde (=where)*). Which adverbs are represented by which t-lemmas and which values of the *indeftype* grammeme they get is summarized in Table 4.9, “**Derivatives of the pronominal adverb *kde***”. For example:

*Nikudy.DIR2 to nejde.* [t\_lemma = *kde*; indeftype = negat] (=There is no way to go; lit. Through\_no\_way it goes.)

- c. **relative, indefinite, interrogative, negative and totalizing pronominal adverbs with a temporal meaning** (derivatives of the adverb *kdy*). Relative, indefinite, interrogative, negative and totalizing pronominal adverbs with various temporal meanings (answering the questions “od kdy (=from when)”, “do kdy (=until when)” etc.; *sempos = adv.pron.indef*) are represented by a node with the t-lemma of the corresponding adverb with the simplest (basic) temporal meaning, i.e. with the t-lemma *kdy*. Which adverbs are represented by which t-lemmas and which values of the *indeftype* grammeme they get is summarized in Table 4.10, “**Derivatives of the pronominal adverb *kdy***”. For example:

*Kdypak.TWHEN přijdete?* [t\_lemma = *kdy*; indeftype = inter] (=When will you come?)

- d. **adverbial derivatives of definite cardinal numerals**. Adverbs with a definite quantitative meaning (*sempos = adj.quant.def*) are represented by a node with the t-lemma of the corresponding cardinal numeral. Which adverbs are represented by which t-lemmas and which values of the *numertype* grammeme they get is summarized in Table 4.11, “**Examples of adverbial derivatives of definite numerals**”. For example:

*Třikrát.THO a dost.* [t\_lemma = *tři*; numertype = basic] (=Three times is enough.)

- e. **adverbial derivatives of the indefinite numeral *kolik*** Adverbial derivatives of the indefinite numeral *kolik* (=how much/many) (*sempos = adj.quant.indef*) are represented by a node with the t-lemma of the corresponding cardinal numeral (i.e. *kolik*). Possible combinations of the t-lemma, functor and grammemes *numertype* and *indeftype* are shown in Table 4.12, “**Adverbial derivatives of the indefinite numeral *kolik***”. For example:

*Pokdovíkoliká té.TWHEN už se to podařilo.* [t\_lemma = *kolik*; numertype = ord; indeftype = indef6] (=I don't know how many times it has turned out well.)

- f. **adverbial derivatives of indefinite gradable numerals**. Adverbial derivatives of indefinite numerals, e.g. *málokrát* (=few times), *hodněkrát* (=many times), *dostkrát* (=lit. enough\_times), *vícekrát* (=lit. more\_times) (*sempos = adj.quant.grad*), are represented by a node with the t-lemma of the corresponding cardinal numeral (i.e. *málo* (=few), *hodně* (=a lot), *dost* (=enough)). The basic word form of these adverbs follows from the combination of the t-lemma, the value of the *numertype* grammeme (basic) and the functor (THO). For example:

*Stalo se to jen málokrát.THO* [t\_lemma = *málo*; numertype = basic] (=It only happened few times.)

**Table 4.8. Possessive derivatives of the pronoun *kdo***

<b>t_lemma = kdo</b>	functor
<b>indeftype:</b>	<b>APP</b> (or possibly AUTH, ACT)
<b>relat</b>	<i>čí</i>
<b>indef1</b>	<i>něčí</i>
<b>indef2</b>	<i>čísi, čís</i>
<b>indef3</b>	<i>číkoli, číkoliv</i>
<b>indef4</b>	<i>lecčí, ledačí</i>
<b>indef5</b>	<i>kdečí</i>
<b>indef6</b>	<i>máločí, kdovíčí</i>
<b>inter</b>	<i>čí, čípak</i>
<b>negat</b>	<i>ničí</i>
<b>total1</b>	-
<b>total2</b>	-

**Table 4.9. Derivatives of the pronominal adverb *kde***

<b>t_lemma = kde</b>	functor			
<b>indeftype</b>	<b>LOC</b>	<b>DIR1</b>	<b>DIR2</b>	<b>DIR3</b>
<b>relat</b>	<i>kde</i>	<i>odkud</i>	<i>kudy</i>	<i>kam</i>
<b>indef1</b>	<i>někde</i>	<i>odněkud</i>	<i>někudy</i>	<i>někam</i>
<b>indef2</b>	<i>kdesi</i>	<i>odkudsi</i>	<i>kudysi</i>	<i>kamsi</i>
<b>indef3</b>	<i>kdekoli, kdekoliv</i>	<i>odkudkoli, odkudkoliv</i>	<i>kudykoli, kudykoliv</i>	<i>kamkoli, kamkoliv</i>
<b>indef4</b>	<i>ledakde, leckde, ledaskde, leckdes</i>	-	-	-
<b>indef5</b>	-	-	-	-
<b>indef6</b>	<i>málokde, sotvakde, zřídkakde, všelikde, nevímkde, bůhvíkde, čertvíkde, ...</i>	-	-	-
<b>inter</b>	<i>kde, kdepak, kdeže</i>	<i>odkud</i>	<i>kudy</i>	<i>kam</i>
<b>negat</b>	<i>nikde</i>	<i>odnikud</i>	<i>nikudy</i>	<i>nikam</i>
<b>total1</b>	<i>všude</i>	<i>odevšad, odevšud</i>	<i>všudy</i>	<i>všude</i>
<b>total2</b>	-	-	-	-

**Table 4.10. Derivatives of the pronominal adverb *kdy***

<b>t_lemma = kdy</b>	functor				
<b>indeftype</b>	<b>TWHEN</b>	<b>TSIN</b>	<b>TTILL</b>	<b>TFHL</b>	<b>THO</b>
<b>relat</b>	<i>kdy</i>	<i>odkdy</i>	<i>dokdy, dokud</i>	-	-
<b>indef1</b>	<i>někdy</i>	-	-	-	-
<b>indef2</b>	<i>kdysi</i>	-	-	-	-
<b>indef3</b>	-	-	-	-	<i>kdykoli, kdykoliv</i>
<b>indef4</b>	<i>ledakdy, leckdy, ledaskdy</i>	-	-	-	-
<b>indef5</b>	-	-	-	-	-
<b>indef6</b>	<i>málokdy, sotvakdy, zřídkakdy, nevímky-dy, kdovíkdy, bůhvíkdy, čertvík-dy, ...</i>	-	-	-	-
<b>inter</b>	<i>kdy, kdypak, kdypakže, kdyže</i>	<i>odkdy</i>	<i>dokdy, dokud</i>	-	-
<b>negat</b>	<i>nikdy</i>	-	-	-	-
<b>total1</b>	<i>vždy, vždycky</i>	-	-	<i>navždy,</i> <i>navždycky</i>	-
<b>total2</b>	-	-	-	-	-

**Table 4.11. Examples of adverbial derivatives of definite numerals**

<b>t_lemma =</b> <b>jeden, tři, sto,</b> <b>tolik ...</b>	functor	
<b>numertype</b>	<b>TWHEN</b>	<b>THO</b>
<b>basic</b>	-	<i>jedenkrát, třikrát, stokrát, tolíkrát, ...</i>
<b>ord</b>	<i>poprvé, potřetí, posté, potolíkáté, ...</i>	-

**Table 4.12. Adverbial derivatives of the indefinite numeral *kolik***

<b>t_lemma:</b> <b>kolik</b>	<b>numertype</b>	
<b>indeftype</b>	<b>basic</b>	<b>ord</b>
<b>relat</b>	<i>kolikrát</i>	<i>pokolikáté</i>
<b>indef1</b>	<i>několikrát</i>	<i>poněkolikáté</i>
<b>indef2</b>	-	-
<b>indef3</b>	-	-
<b>indef4</b>	-	-
<b>indef5</b>	-	-
<b>indef6</b>	<i>kdovíkolikrát, bůhvíkolikrát, čertvíkolikrát, nevímkolikrát, ...</i>	<i>pokdovíkolikáté, pobůhvíkolikáté, počertvíkolikáté, ponevímkolikáté, ...</i>
<b>inter</b>	<i>kolikrát</i>	<i>pokolikáté</i>
<b>negat</b>	-	-
<b>total1</b>	-	-
<b>total2</b>	-	-
<b>functor</b>	<b>THO</b>	<b>TWHEN</b>

## 4.3. Multi-word t-lemma

**Multi-word t-lemma** is assigned to nodes representing some multi-word lexical units (see Section 7.1, “Multi-word lexical units”); it consists of all components of the given lexical unit (relevant at the tectogrammatical level); the components are linked by the underscore mark.

We distinguish two types of multi-word t-lemmas:

- A. the multi-word t-lemma consists of the **basic word forms** of all components of the given (multi-word) lexical unit.

This concerns:

- a. **verbs with *se* or *si*.** The t-lemma of a node representing a verb with *se* or *si* (inherently reflexive or reciprocal verbs), consists of the infinitive of the verb and the reflexive *se* or *si*, connected by the underscore mark. For example:

*Pořád se smáli.* [ $t\_lemma = smát\_se$ ] (=They laughed all the time.)

*Setkali se teprve včera.* [ $t\_lemma = setkat\_se$ ] (=They only met yesterday.)

- b. **multi-word conjunctions and conjunction pairs.** The t-lemma of a node representing a multi-word conjunction or conjunction pair consists of all parts of the given conjunction, which are linked by the underscore marks. The order and form of the individual parts are determined by the chosen representative t-lemma, which serves for various formal and word-order variants (see Section 4.4, “Representative t-lemma”). For example:

*Bud' přijdeš ty, nebo přijdu já.* [ $t\_lemma = bud'_nebo$ ] (=Either you will come or I will.)

*Přišli jak děti, tak dospělí.* [ $t\_lemma = jak_tak$ ] (=Both children and adults came.)

- c. **multi-word operators.** The t-lemma of a node representing a multi-word operator consists of all parts of the given operator, which are linked by the underscore marks. The order and form of the individual parts are determined by the chosen representative t-lemma, which serves for various formal and word-order variants (see Section 4.4, “Representative t-lemma”). For example:

*Trvá to dlouho, od jara přes celé léto až do zimy.* [ $t\_lemma = od_přes_do$ ] (=It takes a long time, from spring to summer and winter.)

- d. **numbers with the function of a “label”.** The t-lemma of a node representing a (compound) number with the function of a “label” (see Section 7.2, “Numbers and numerals”) consists of all parts of the given number, linked by the underscore marks, in the surface word order. For example:

*Zavolej mi na 737 677 228.* [ $t\_lemma = 737_677_228$ ] (=Call number 737 677 228.)

- e. **expressions of the form 'number+adjective'.** The t-lemma of a node representing an expression formed by a number and adjective consists of the number and the basic form of the adjective, which are linked by the underscore mark. For example:

*41 letá žena* [ $t\_lemma = 41\_letý$ ] (=41 year old woman)

*5 hodinová čekací doba* [ $t\_lemma = 5\_hodinový$ ] (=5 hour waiting time)

- f. **surnames containing *van*, *von*, *de* etc.** The t-lemma of a node representing a surname containing the foreign prepositions *van*, *von*, *de* etc. consists of the foreign preposition, which is linked to the surname by the underscore mark. For example:

*skladatel Ludwig van Beethoven* [ $t\_lemma = van_Beethoven$ ] (=the composer L. van B.)

*von Ryanův otec* [ $t\_lemma = von_Ryan$ ] (=von Ryan's father)

- B. the multi-word t-lemma consists of the **actual surface forms** of all words forming the given lexical unit.

This representation is used for **dependent parts of idioms**. The t-lemma of a node representing a multi-word dependent part of an idiom (with the DPHR functor, *node type* = *dphr*; see Section 7.1.2, “Idioms”) consists of the actual surface forms of the dependent parts (including prepositions) as present at the surface. For example:

*Vzal nohy na ramena.* [t\_lemma = nohy\_na\_ramena] (=lit. (He) took legs on shoulders.)

*Měl hluboko do kapsy.* [t\_lemma = hluboko\_do\_kapsy] (=lit. (He) had deep in pocket.)

## 4.4. Representative t-lemma

Synonymy is represented in PDT 2.0 only marginally so far. **Representative t-lemmas** are assigned to nodes that represent various word-order, formal or stylistic variants of some of the following expressions and corresponds to one of the variants:

- a. **pronouns, numerals, pronominal adverbs.** With pronouns, numerals and pronominal adverbs, the synonymy is represented with the help of a set of features attached to the t-lemma. Decisions regarding the stylistic and formal variants of pronominal and numeral expressions are related to the representation of the derivational relations between the expressions: stylistic and formal variants are subsumed under one representative variant; as for the derived (stylistic or formal) variants, they are represented by the basic form of their base word. If in the tables:

Table 4.2, “Derivatives of the pronominal adverbs *tady* and *tam*”,

Table 4.3, “Derivatives of the pronominal adverbs *ted'*, *potom* and *tehdy*”,

Table 4.5, “Adjectival derivatives of the numeral *kolik*”,

Table 4.6, “Derivatives of the pronouns *kdo*, *co*, *který* and *jaky*”,

Table 4.7, “Derivatives of the pronominal adverbs *jak* and *proč*”,

Table 4.8, “Possessive derivatives of the pronoun *kdo*”,

Table 4.9, “Derivatives of the pronominal adverb *kde*”,

Table 4.10, “Derivatives of the pronominal adverb *kdy*”,

Table 4.12, “Adverbial derivatives of the indefinite numeral *kolik*”,

in Section 4.2, “T-lemma of derived expressions” the t-lemma in a given row has more surface forms, they are taken to be synonymous.

- b. **coordinating conjunctions and operators and other expressions.** Synonymy is represented also with certain connectives, operators and other marginal expressions (e.g. abbreviations) - see Table 4.13, “**Representative t-lemmas with conjunctions and operators**” and Table 4.14, “**Representative t-lemma with other expressions**”. For example:

*Mezi smysly patří zrak a sluch a hmat a čich.* [t\_lemma = a] (=Eyesight and hearing and touch belong to the senses.)

*počínaje složitou dopravou na Strahov, přeplněným parkovištěm, až po dlouhé fronty na lístky* [t\_lemma = od\_do] (=from the difficult journey to Strahov to the long queues for the tickets)

*Takové dítě je tzv. nechlibitevné.* [t\_lemma = takzvaný] (=You cannot boast with such a kid, so to speak.)

**Table 4.13. Representative t-lemmas with conjunctions and operators**

Representative t-lemma	Surface variants
ani	ani; ni
bud'_nebo	bud' - nebo; bud' - anebo; bud'to - anebo; bud'to - nebo
od_do	od - do; od - po; od - k; počínaje - po; od - počínaje - po
od_přes_do	od - přes - do; od - přes - k; od - přes - po, počínaje - přes - k
počínaje_konče	počínaje - konče; počínaje - a konče
conjunction (e.g. a)	iterated conjunction (e.g. a - a - a)

**Table 4.14. Representative t-lemma with other expressions**

Representative t-lemma	Surface variants
takzvaný	tzv.; tak zvaný; takzvaný, takzvaně
to_znamená	tn.; to znamená
to_jest	tj.; to jest

## 4.5. T-lemma substitute

The term **t-lemma substitute** is used for artificial t-lemmas beginning with #. T-lemma substitutes are assigned to:

- newly established nodes that are not copies of other nodes** (see Section 6.6, “Ellipsis”).

T-lemma substitutes: #AsMuch, #Benef, #Cor, #EmpNoun, #EmpVerb, #Equal, #Forn, #Gen, #Idph, #Neg, #Ob1fm, #PersPron, #QCor, #Rcp, #Separ, #Some, #Unsp, #Total (see Table 6.8, “**Survey of newly established nodes**” in Section 6.6, “Ellipsis”).

- nodes representing **personal and reflexive pronouns and their possessive counterparts**.

T-lemma substitute: #PersPron.

- nodes representing **punctuation marks and other non-alphabetical/numerical symbols** (if used at the tectogrammatical level).

For their t-lemma substitutes see Table 4.15, “**T-lemma substitutes of non-alphabetical/numerical symbols**”.

**Table 4.15. T-lemma substitutes of non-alphabetical/numerical symbols**

Non-alphabetical/numerical symbol	Description of the symbol	T-lemma substitute
&	ampersand	#Amp
%	per cent	#Percnt
*	asterisk	#Ast
.	period	#Period
...	three periods	#Period3
:	colon	#Colon
,	comma	#Comma
;	semicolon	#Semicolon
-	hyphen	#Dash
—	dash	#Dash
/	slash	#Slash
(	left bracket	#Bracket
)	right bracket	#Bracket

**List of all t-lemma substitutes (in alphabetical order).** In Table 4.16, “**List of all t-lemma substitutes (in alphabetical order)**”, there is a list of all t-lemma substitutes occurring in the tectogrammatical trees. The t-lemmas are in alphabetical order and it is always indicated whether they represent a word/symbol present in the surface structure or whether they correspond to a newly established node (with no counterpart at the surface level). Furthermore, it is specified which node type is usually connected with a particular t-lemma.

**Table 4.16.** List of all t-lemma substitutes (in alphabetical order).

T-lemma	nodetype	is_generated	Node description
#Amp	coap	0	node representing the symbol &
#AsMuch	qcomplex	1	node for an expression introducing a consecutive clause
#Ast	qcomplex	0	node representing the symbol *
#Benef	qcomplex	1	node for the Beneficiary in control constructions
#Bracket	coap	0	node representing a bracket
#Colon	coap qcomplex	0	node representing a colon
#Comma	coap qcomplex	0	node representing a comma
#Cor	qcomplex	1	node representing the controllee in control constructions
#Dash	coap qcomplex	0	node representing a hyphen or dash
#EmpNoun	complex	1	node representing an elided governing expression of syntactic adjectives
#EmpVerb	qcomplex	1	node representing the elided governing predicate of a verbal clause
#Equal	qcomplex	1	node for the elided adjective in positive in comparative constructions
#Forn	list	1	root node of a list structure for foreign-language expressions
#Gen	qcomplex	0	node for a general argument
#Idph	list	1	root node of an identifying structure
#Neg	atom	1	node for syntactic negation
#Oblfm	qcomplex	1	node representing an elided obligatory adjunct
#Percnt	qcomplex	0	node representing the symbol %
#Period	coap	0	node representing a period
#Period3	coap qcomplex	0	node representing three dots
#PersPron	complex	0	node representing personal or possessive pronouns (including reflexives). With newly established nodes, the #PersPron t-lemma signals ellipsis.
#QCor	qcomplex	1	node representing the quasi-controllee in quasi-control constructions
#Rcp	qcomplex	1	node representing a valency modification omitted due to its participation in a reciprocal relation
#Semicolon	coap	0	node representing a semicolon
#Separ	coap	1	auxiliary node for representing parataxis
#Slash	coap	0	node representing a slash
#Some	qcomplex	1	node for the nominal part of a verbonominal predicate in comparative constructions
#Total	qcomplex	1	node for a non-expressed totalizer in constructions with the meaning of a restriction
#Unsp	qcomplex	1	node representing an unspecified valency modification

---

# Chapter 5. Complex nodes and grammatemes

**Complex nodes** (`nodeType = complex`) are nodes representing autosemantic lexical units (nouns, adjectives, verbs, adverbs, numerals and pronouns). To be represented properly, autosemantic lexical units need apart from a t-lemma and a functor (and possibly other attributes) also grammatemes (the tectogrammatical correlates of morphological categories). The fact that complex nodes have grammatemes sets these nodes apart from all other types of nodes.

## 5.1. Semantic parts of speech

Complex nodes can be classified as belonging to one of four semantic parts of speech. **Semantic parts of speech** correspond to the basic onomasiological categories: substances, properties, circumstances and events. They are:

- semantic nouns** (see Section 5.1.1, “Semantic nouns”),
- semantic adjectives** (see Section 5.1.2, “Semantic adjectives”),
- semantic adverbs** (see Section 5.1.3, “Semantic adverbs”),
- semantic verbs** (see Section 5.1.4, “Semantic verbs”).

Semantic nouns, adjectives and adverbs are further classified. The information about the membership of a complex node in a particular (subgroup of a) semantic part of speech is encoded in the `sempos` attribute. The `sempos` attribute (see Section A.2.8.13, “gram/sempos”) has 19 possible values (symbols assigned to individual values are ordered in such a way that the first one is the one referring to the given semantic part of speech, other symbols (referring to the individual characteristics) follow, separated by a period. The characteristics are organized from the more general ones to the more specific ones).

**Traditional autosemantic parts of speech.** Words belonging to traditional autosemantic parts of speech can usually be said to belong to the corresponding semantic parts of speech. Cases when semantic parts of speech do not correspond to the traditional ones occur as a result of representing certain types of derivation (see Section 4.2, “T-lemma of derived expressions”):

- a. **possessive adjectives** (e.g.: *matčin* (=mother's), *Pavlův* (=Pavel's)) correspond to semantic nouns (`sempos = n. denot`).
- b. **deadjectival adverbs** (e.g.: *pěkně* (=nicely), *rychle* (=quickly)) correspond to semantic adjectives (`sempos = adj. denot`).

**Pronouns.** Pronouns are usually understood as either semantic nouns, or adjectives - according to their syntactic function:

- a. **personal and reflexive pronouns and their possessive counterparts** (e.g.: *já* (=I), *ty* (=you), *vy* (=you.pl), *můj* (=my), *jejich* (=their), *se* (=REFL), *svůj* (=self's)) belong to semantic nouns (`sempos = n. pron. def`).
- b. **demonstrative and identifying pronouns** (e.g.: *ten* (=this), *tenhle* (=this), *tentýž* (=the same), *takový* (=such)) are taken to be either semantic nouns (`sempos = n. pron. def. demon`), or semantic adjectives (`sempos = adj. pron. def. demon`, depending on their syntactic function). Cf.:

*Ti už nepřijdou.* (=These will not come again.) → semantic noun

*Ten dům už prodali.* (=They have already sold the house.) → semantic adjective

- c. **relative pronouns *kdo*, *co* and their derivatives**, i.e. indefinite (e.g.: *někdo* (=somebody), *něco* (=something), *ledakdo* (=apprx. anybody), *kdosi* (=someone)), interrogative (*kdo* (=who), *co* (=what)), negative (*nikdo* (=nobody), *nic* (=nothing)) and totalizing pronouns (*všechn* (=all)), possessive counterparts of the pronoun *kdo* (i.e. *čí* (=whose)) and its derivatives (e.g.: *něčí*, *ničí* (=someone's, noone's), *kdovíčí* (=lit. who knows whose))) are taken to be semantic nouns (`sempos = n. pron. indef`).
- d. **relative pronouns *který* (*jenž*) and *jaký* and their derivatives**, i.e. indefinite (e.g.: *některý* (=some), *nějaký* (=some), *všeljaký* (=all kinds of), *ledajaký* (=all kinds of)), interrogative (*který* (=which), *jaký* (=what)), negative (*žádný* (=no), *nijaký* (=no)) and totalizing pronouns (*každý* (=each/every)) are taken to be either semantic nouns (`sempos = n. pron. indef`), or semantic adjectives (`sempos = adj. pron. indef`, depending on their syntactic function). Cf.:

*Šaty, jaké by se hodily na ples, tu neměli.* (=They didn't have the dress that would be suitable for a ball.)  
*Knihu, kterou si přál, nemohla sehnat.* (=She wasn't able to get the book (that) he wished to have.) → semantic nouns

*Nevěděla, jaké šaty se by se na ples hodily.* (=She didn't know what kind of dress would be suitable for a ball.)  
*Kterou knihu si přál? Kup mu nějakou knihu* (=Which book did he wish to have?). *Ta barva je nijaká.* (=The color is insipid.) → semantic adjectives

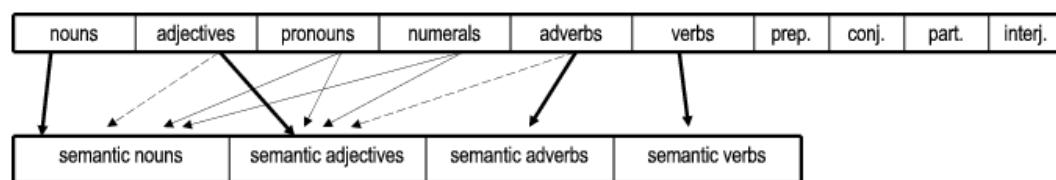
**Numerals.** Numerals are usually taken to be semantic nouns or adjectives, according to their syntactic function:

- a. **definite cardinal numerals** *jedna* through *devětadvadesát* are either semantic nouns (*sempos = n.quant.def*), or adjectives (*sempos = adj.quant.def*, depending on their syntactic function). Cf.:  
*Vybrali tři.* (=They chose three of them.) → semantic noun  
*pět knih* (=five books) → semantic adjective
- b. **definite cardinal numerals** *sto, tisíc, milion* (and other ending with *-ion*), *miliarda* (and other ending with *-iarda*) and **fraction numerals** (e.g.: *třetina* (=one) third), *polovina* (=one) half)) have always the meaning of a “container” and are considered semantic nouns (*sempos = n.quant.def*).
- c. **definite adjectival numerals**, i.e. ordinal, sort and set numerals (e.g.: *druhý* (=the) second), *troje* (=three sets of), *trojí* (=three kinds of)), are considered semantic adjectives (*sempos = adj.quant.def*).
- d. **definite adverbial numerals (multiplicative)** (e.g.: *třikrát* (=three times), *potřetí* (=for the third time), *tisíckrát* (=thousand times), *posté* (=for the hundredth time)) are considered semantic adjectives (*sempos = adj.quant.def*).
- e. **definite cardinal numeral** *tolik* and its **adjectival derivatives** (e.g.: *tolikáty* (=lit. so\_much.ord), *tolikery* (=so\_much.set), *tolikerý* (=so\_much.sort)) as well as its **adverbial derivatives** (*tolikrát* (=so many times), *potolikáty* (=for x-th time)) are considered semantic adjectives (*sempos = adj.quant.def*).
- f. **indefinite gradable cardinal numerals and their adverbial derivatives** (e.g.: *málo, mnoho* (=few/little, many/much), *hodně* (=a lot), *málokrátk* (=few times)) are considered semantic adjectives (*sempos = adj.quant.grad*).
- g. **indefinite numeral** *kolik* and its **adjectival and adverbial derivatives** (e.g.: *několik* (=several), *kolikpak* (=how many/much), *kolikery* (=how many sets), *několikery* (=several sets), *pokolikáty* (=lit. for n-th time), *kolikrát* (=how many times)) are considered semantic adjectives (*sempos = adj.quant.indef*).

**Other traditional parts of speech.** Other parts of speech are represented by other than complex nodes, therefore, no reference to semantic parts of speech is made.

The relations between the semantic and traditional parts of speech are represented in Fig. 5.1. Arrows in boldface indicate that the relation is “prototypical” (nouns prototypically belong to the class of semantic nouns, adjectives to the class of semantic adjectives etc.), simple arrows indicate the distribution of pronouns and numerals between the semantic nouns and adjectives and dotted arrows follow the classification based on derivational relations.

**Figure 5.1. Relations between the semantic and traditional parts of speech**



## 5.1.1. Semantic nouns

**Semantic nouns** correspond to the basic onomasiological category of substance.

The inner structure of semantic nouns is presented in Fig. 5.2. The name of the subgroup is followed by its abbreviation (which is given as the value of the *sempos* attribute). Below the name of the subgroup, there is a list of the grammatemes (in a box) relevant for the subgroup.

For a detailed survey of semantic nouns see Table 5.1, “Survey of semantic nouns”.

Figure 5.2. Inner structure of semantic nouns

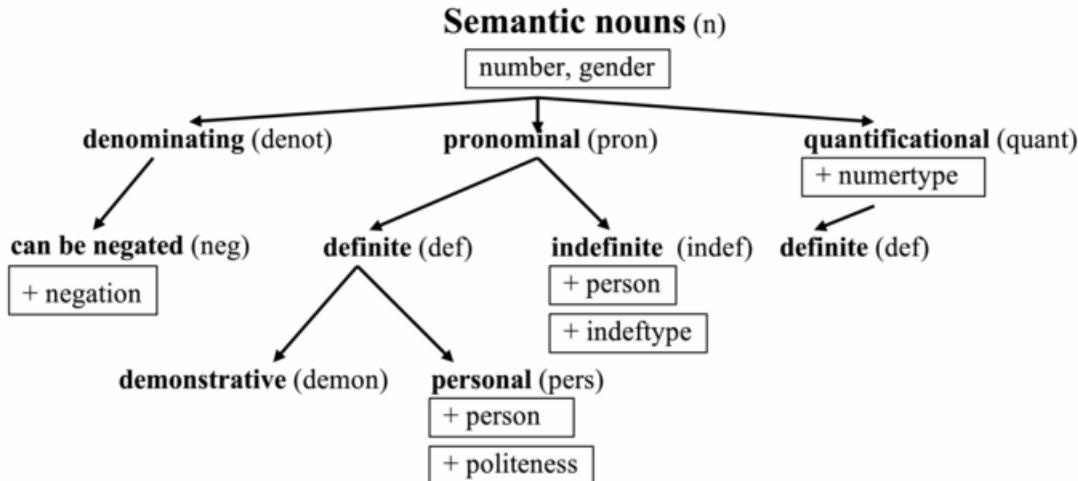


Table 5.1. Survey of semantic nouns

sempos Definition of the subgroup	Grammateemes of the subgroup	Members of the subgroup	T-lemma	Example
<b>n.denot</b> denominating	number gender	traditional nouns (not ending with <i>-ní</i> , <i>-tí</i> or <i>-ost</i> )	basic word form (nom. sg.)	<i>psi</i> (=dogs) [t_lemma = <i>pes</i> ; number = pl; gender = anim]
		possessive adjectives	nom. sg. of the corresponding noun	<i>Martin pokoj</i> (=Marta's room) [t_lemma = <i>Marta</i> ; number = sg; gender = fem]
<b>n.denot.neg</b> denominating that can be negated	number gender negation	traditional nouns ending with <i>-ní</i> , <i>-tí</i> or <i>-ost</i>	basic word form (nom. sg) non-negated form	<i>nezralost dítěte</i> (=the child's immaturity) [t_lemma = <i>zralost</i> ; number = sg; gender = fem; negation = neg1]
<b>n.pron.def.demon</b> pronominal definite demonstrative	number gender	demonstrative pronouns in the position of a syntactic noun	basic word form (nom. sg.)	<i>O tohle mi nejde.</i> (=This is not the point.) [t_lemma = <i>ten</i> ; number = sg; gender = neut]
		nodes with the t-lemma #EmpNoun	#EmpNoun	<i>Mluvil také o anglickém.</i> (=He also talked about the English one.) {#EmpNoun [number = sg; gender = anim inan neut]}

<b>n.pron.def.pers</b> pronominal definite personal	number gender person politeness	personal and reflexive pronouns	#PersPron	<i>Vyjste se už přihlásil.</i> (=You have already registered.)  [t_lemma = #PersPron; number = sg; gender = anim; person = 2; politeness = polite]
		possessive counterparts of personal and reflexive pronouns	#PersPron	<i>Tvůj názor nesdílím.</i> (=I don't share your view.)  [t_lemma = #PersPron; number = sg; gender = anim; person = 2; politeness = basic]
		newly established nodes with the t-lemma #PersPron	#PersPron	<i>Odešel.</i> (=He has left.)  {#PersPron [number = sg; gender = anim; person = 3; politeness = basic]}
<b>n.pron.indef</b> pronominal indefinite	number gender person indeftype	relative pronouns <i>který</i> ( <i>jenž</i> ) and <i>jaký</i> and their derivatives, i.e. indefinite, interrogative, negative and totalizing pronouns in the position of a syntactic noun	<i>který, jaký</i>	<i>Která to je?</i> (=Which one is it?)  [t_lemma = <i>který</i> ; number = sg; gender = fem; person = 3; indeftype = inter]
		relative pronouns <i>kdo</i> , <i>co</i> and their derivatives, i.e. indefinite, interrogative, negative and totalizing pronouns	<i>kdo, co</i>	<i>Někdo to udělat musí.</i> (=Somebody has to do it.)  [t_lemma = <i>kdo</i> ; number = sg; gender = anim; person = 3; indeftype = indef1]
		possessive counterparts of the pronoun <i>kdo</i> and its derivatives	<i>kdo</i>	<i>něčí tužka</i> (=somebody's pen)  [t_lemma = <i>kdo</i> ; number = sg; gender = anim; person = 3; indeftype = indef1]

<b>n.quant.def</b> quantificational definite	number gender numertype	cardinal numerals <i>jedna</i> through <i>devadesát devět</i> in the position of a syntactic noun	basic word form (nom. sg.)	<i>Poznal jen jednu.</i> (=He only recognized one of them.) [t_lemma = <i>jeden</i> ; number = sg; gender = fem; numertype = basic]
		cardinal numerals <i>sto</i> , <i>tisíc</i> , <i>milion</i> (and other ending with <i>-ión</i> ), <i>miliarda</i> (and other ending with <i>-iarda</i> )	basic word form (nom. sg.)	<i>Přišlo sto studentů.</i> (=One hundred students came.) [t_lemma = <i>sto</i> ; number = sg; gender = neut; numertype = basic]
		fraction numerals	nom. sg. of the corresponding cardinal numeral	<i>Koupil dvě třetiny akcií.</i> (=He bought two thirds of the shares.) [t_lemma = <i>tři</i> ; number = pl; gender = fem; numertype = frac]

## 5.1.2. Semantic adjectives

Semantic adjectives correspond to the basic onomasiological category of property/quality.

The inner structure of semantic adjectives is presented in Fig. 5.3 (the structure is parallel to that in Fig. 5.2). For a detailed survey of semantic adjectives see Table 5.2, “Survey of semantic adjectives”.

Figure 5.3. Inner structure of semantic adjectives

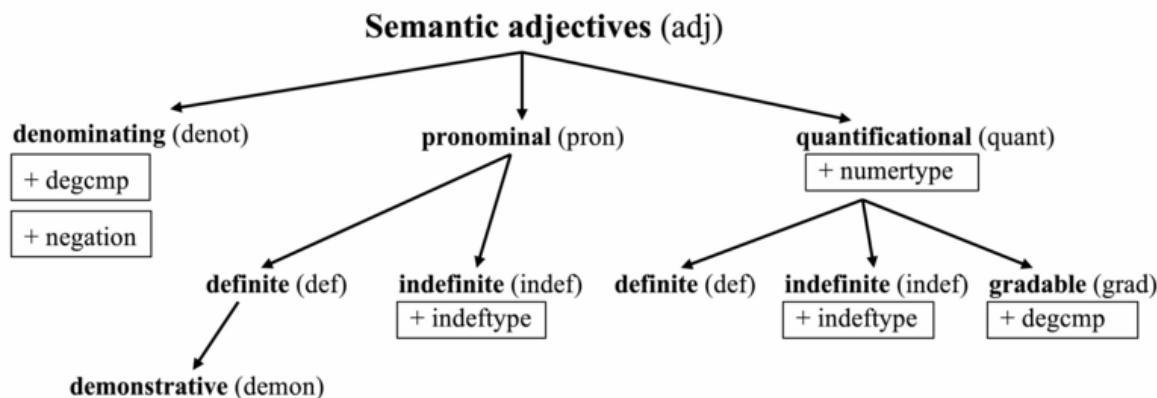


Table 5.2. Survey of semantic adjectives

sempos Definition of the subgroup	Grammateemes of the subgroup	Members of the subgroup	T-lemma	Example
<b>adj . denot</b> denominating	degcmp negation	traditional adjectives	basic word form (nom. sg) non-negated form	<i>příjemnější hudba</i> (=more pleasant music) [t_lemma = <i>příjemný</i> ; degcmp = comp; negation = neg0]
		traditional adverbs (deadjectival)	nom. sg. of the corresponding adjective	<i>Zachoval se nepěkně.</i> (=His behavior wasn't nice.) [t_lemma = <i>pěkny</i> ; degcmp = comp; negation = neg1]

<b>adj . pron . def . demon</b> pronominal, definite, demonstrative	$\emptyset$	demonstrative and identifying pronouns in the position of a syntactic adjective	basic word form (nom.sg)	<i>Ten dům už koupili.</i> (=They have bought the house) [t_lemma = ten]
<b>adj . pron . indef</b> pronominal, indefinite	indeftype	relative pronouns <i>který</i> and <i>jaký</i> and their derivatives, i.e. indefinite, interrogative, negative and totalizing pronouns in the position of a syntactic adjective	<i>který, jaký</i>	<i>Ta barva je nijaká.</i> (=The color is insipid.) [t_lemma = jaký; indeftype = negat]
<b>adj . quant . def</b> quantificational, definite	numertype	definite cardinal numerals ( <i>jedna</i> through <i>devadesát devět</i> ) in the position of a syntactic adjective	basic word form (nom. sg.)	<i>Koupil tři domy.</i> (=He bought three houses.) [t_lemma = tři; numertype = basic]
		definite adjectival numerals, i.e. ordinal, sort and set numerals	nom. sg. of the corresponding cardinal numeral	<i>Umístil se na třetím místě.</i> (=He came in third) [t_lemma = tři; numertype = ord]
		definite adverbial (multiplicative) numerals	nom. sg. of the corresponding cardinal numeral	<i>Volal tam už dvakrát.</i> (=He has already called them twice.) [t_lemma = dva; numertype = basic]
		definite cardinal numeral <i>tolik</i> and its adjectival and adverbial derivatives	<i>tolik</i>	<i>Tolik připomínek nečekal.</i> (=He didn't expect so many comments.) [t_lemma = tolik; numertype = basic]
<b>adj . quant . indef</b> quantificational, indefinite	numertype indeftype	indefinite cardinal numeral <i>kolik</i> and its adjectival and adverbial derivatives	<i>kolik</i>	<i>Vysvětluje to už poněkolikáte.</i> (=It's not the first time he's explaining that.) [t_lemma = kolik; numertype = ord; indeftype = inedf1]
<b>adj . quant . grad</b> quantificational, gradable	numertype degcmp	indefinite cardinal (gradable) numerals and their adverbial derivatives ( <i>málo</i> , <i>málokárat</i> )	nom. sg. of the corresponding cardinal numeral	<i>nejméně chyb</i> (=the fewest mistakes) [t_lemma = málo; degcmp = sup; numertype = basic]

### 5.1.3. Semantic adverbs

Semantic adverbs correspond to the basic onomasiological category of circumstance.

The inner structure of semantic adverbs is presented in Fig. 5.4 (the structure is parallel to that in Fig. 5.2). For a detailed survey of semantic adverbs see Table 5.3, “Survey of semantic adverbs”.

Figure 5.4. Inner structure of semantic adverbs

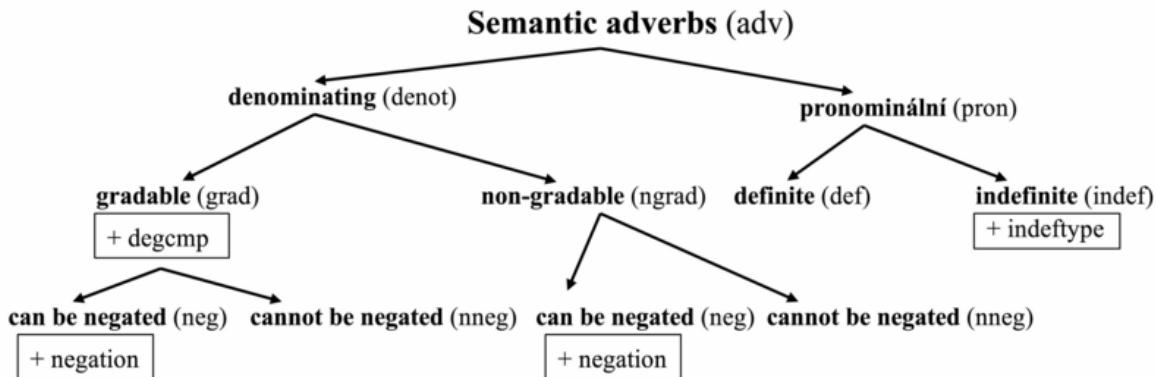


Table 5.3. Survey of semantic adverbs

sempos Definition of the subgroup	Gram-matemes of the subgroup	Members of the subgroup	T-lemma	Example
<b>adv. denot. ngrad. nneg</b> denominating, non-gradable, that cannot be negated	Ø	(non-deadjectival) adverbs that can be neither negated nor are gradable	basic word form	<i>Zůstal dnes doma.</i> (=He has stayed at home today.)
<b>adv. denot. ngrad. neg</b> denominating, non-gradable, that can be negated	negation	(non-deadjectival) adverbs that can be negated but are not gradable	basic word form positive (non-negated) form	<i>Je to nepříliš dobré.</i> (=It's not very good.) [t_lemma = <i>příliš</i> ; negation = neg1]
<b>adv. denot. grad. nneg</b> denominating, gradable, that cannot be negated	degcmp	(non-deadjectival) adverbs that cannot be negated but are gradable	basic word form	<i>Dům ležel nejniže z celé vesnice.</i> (=The house was situated lowest in the village.) [t_lemma = <i>dole</i> ; degcmp = sup]
<b>adv. denot. grad. neg</b> denominating, gradable, that can be negated	negation degcmp	(non-deadjectival) adverbs that can both be negated and are gradable	basic word form positive (non-negated) form	<i>blíže nespecifikovaná trhavina</i> (=further unspecified explosive) [t_lemma = <i>blízko</i> ; degcmp = acomp; negation = neg0]
<b>adv. pron. def</b> pronominal, definite	Ø	definite pronominal locative adverbs and their directional derivatives	<i>tady, tam, taméž</i>	<i>Odtud už to není daleko.</i> (=It's not far from here.) [t_lemma = <i>tady</i> ]
		definite pronominal temporal adverbs and their various (temporal) derivatives	<i>ted', potom, tehdy, onehdy, předtím</i>	<i>Odted' už to nebude mít ríkat.</i> (=From now on, we're not going to say that.) [t_lemma = <i>ted'</i> ]
		other definite pronominal adverbs (demonstrative and identifying)	basic word form	<i>Udělal to právě proto.</i> (=He did it just for this reason.) [t_lemma = <i>proto</i> ]

<b>adv.pron.indef</b> pronominal, indefinite	indeftype	indefinite pronominal adverbs <i>jak</i> and <i>proc</i> and their derivatives, i.e. indefinite, interrogative, negative and totalizing pronominal adverbs	<i>jak, proc</i>	<i>Bylo mu všelijak.</i> (=He felt strange.) [t_lemma = <i>jak</i> ; indeftype = indef6]
		indefinite pronominal adverbs <i>kdy</i> and <i>kde</i> and their derivatives, i.e. indefinite, interrogative, negative, totalizing pronominal adverbs as well as directional or various temporal adverbs	<i>kdy, kde</i>	<i>Nikudy to nejde.</i> (=There is no way to go.) [t_lemma = <i>kde</i> ; indeftype = negat]

## 5.1.4. Semantic verbs

Semantic verbs correspond to the basic onomasiological category of event. Semantic verbs as a group have no further inner structure. For a survey of semantic verbs see Table 5.4, “Semantic verbs”.

Table 5.4. Semantic verbs

sempos	Verbal grammatemes	Definition of the group	T-lemma	Examples
v	verbmod deontmod dispmod aspect tense resultative iterativeness	finite verb form infinitive participle transgressive (gerund)	basic word form (infinitive)	<i>Studenti přišli na schůzi včas.</i> (=The students came in time to the meeting.)  [t_lemma = <i>přijít</i> ; verbmod = ind; deontmod = decl; dispmod = disp0; aspect = cpl; tense = ant; resultative = res0; iterativeness = it0]  <i>Ty musíš přijít.</i> (=You must come.)  [t_lemma = <i>přijít</i> ; verbmod = ind; dentmod = deb; dispmod = disp0; aspect = cpl; tense = sim; resultative = res0; iterativeness = it0]

## 5.2. Grammatemes

Grammatemes are tectogrammatical correlates of morphological categories.

Grammatemes are only assigned to complex nodes (nodetype = complex), they are encoded in the gram attribute, whose value is in fact a structure made of individual attributes-grammatemes. An obligatory attribute of this structure is the sempos attribute, the value of which clearly defines the set of the grammatemes relevant for the given node.

**Grammateme values.** We distinguish basic and special values of grammatemes:

- **basic values.** Every grammeme has at least two basic values (e.g. sg and pl for the number grammeme). In regular cases, the grammeme has always one of the basic values. In case it is possible to assign more basic values (and to exclude at least one of them), all the possible values are listed. For example:

*Viděl jen dva.* [gender = inan|anim] (=He only saw two of them)

- **special values** are described in Table 5.5, “Special values of grammatemes”.

**Table 5.5. Special values of grammatemes**

Special value	Relevant grammatemes	Relevant cases	Examples
<b>nr</b>	all grammatemes	All basic values are possible in the given case; none of them can be excluded.	<i>Výpral si kalhoty.</i> (=He washed his pants.) [number=nr]
<b>nil</b>	dispmod tense verbmod	None of the basic values of the grammeme (i.e. the grammeme as such) is relevant.	<i>Přestaň s tím zlobením!</i> (=Stop that!) [tense=nil]
<b>inher</b>	gender number person politeness	The value of the grammeme follows from the value of the given grammeme of the coreferred node (the value is inherited). The value is assigned to all grammemes of coreferring nodes (reflexives or relatives).	<i>Muži, kteří přišli, jsou naši známi.</i> (=The men who came are our friends.) [gender=inher; number=inher; person=inher] <i>Sestra nemá svůj názor.</i> (=My sister hasn't got her own opinion.) [gender=inher; number=inher; person=inher; politeness=inher]

For a survey of grammatemes of semantic nouns, adjectives and adverbs, see Table 5.6, “Survey of non-verbal grammatemes”. For a survey of grammatemes of semantic verbs see Table 5.7, “Survey of verbal grammatemes”.

**Table 5.6. Survey of non-verbal grammatemes**

Grammateme Values	Relevant subgroups ( <b>sempos</b> )	Definition of the grammeme	Examples
<b>degcmp:</b> pos comp sup acomp nr	adj.denot adj.quant.grad adv.denot.grad.neg adv.denot.grad.nneg (abbreviation: grad and value adj.denot)	tectogrammatical correlate of the morphological category of degree	<i>přijemnější hudba</i> (=more pleasant music) [degcmp = comp] <i>nejméně chyb</i> (=fewest mistakes) [degcmp = sup] <i>Je už pozdě.</i> (=It's already late.) [degcmp = pos] <i>víc</i> [degcmp = comp] <i>příjemně</i> (=more pleasantly) [degcmp = pos] <i>starší žena</i> (=elder woman) [degcmp = accomp]
<b>gender:</b> anim inan fem neut inher nr	n.denot n.denot.neg n.pron.def.demon n.pron.def.pers n.pron.indef n.quant.def (abbreviation: n)	tectogrammatical correlate of the morphological category of gender	<i>děvče</i> (=girl) [gender = neut] <i>Martin pokoj</i> (=Marta's room) [gender = fem] <i>O tohle mi nejde.</i> (=This is not the point) [gender = neut] <i>Někdo to udělat musí.</i> (=Somebody has to do it.) [gender = anim] <i>Byl tam jen jeden.</i> (=Only one of them was there.) [gender = anim   inan] <i>Přišli tři.</i> (=Three of them came.) [gender = anim] <i>Viděl jen tři.</i> (=He only saw three of them.) [gender = nr]

<b>indeftype:</b>	n.pron.indef adj.pron.indef adj.quant.indef adv.pron.indef (abbreviation: indef)	it captures the semantic feature in which the indefinite, interrogative, negative and totalizing pronouns (and their adverbial derivatives) differ from their corresponding relative pronouns, by the t-lemma of which they are represented	<i>Ten, <u>kdo</u> přišel.</i> (=The one who came.) [indeftype = relat] <i>Kdo přišel?</i> (=Who came?) [indeftype = inter] <i>Někdo přišel.</i> (=Somebody has come.) [indeftype = indef1] <i>čísi chlapec</i> (=somebody's boy) [indeftype = indef2] <i>jakýkoliv úkol</i> (=any task) [indeftype = indef3] <i>Může to být ledaskde.</i> (=It can be at various places.) [indeftype = indef4] <i>Myslí si to <u>kdekdo</u>.</i> (=Many people think this.) [indeftype = indef5] <i>Bůhvíkterýden to bylo.</i> (=God knows which day it was.) [indeftype = indef6] <i>To je všechno.</i> (=That's all.) [indeftype = total1] <i>každý den</i> (=every day) [indeftype = total2] <i>Nikdy už to neudělám.</i> (=I'll never do it again.) [indeftype = negat]
-------------------	--	---	--

<b>negation:</b>	n.denot.neg adj.denot adv.denot.grad.neg adv.denot.ngrad.neg (abbreviation: neg and value adj.denot)	it captures the fact whether the semantic noun (ending with <i>-ní</i> , <i>-tí</i> or <i>-ost</i> ), adjective or adverb occurred in its positive or negative form	<i>nezralost dítěte</i> (=the child's immaturity) [negation = neg1] <i>příjemná hudba</i> (=pleasant music) [negation = neg0] <i>Zachoval se k nám nepěkně.</i> (=He treated us not very nicely.) [negation = neg1] <i>Pracuje blízko od domova.</i> (=He works near his home.) [negation = neg0] <i>nekalé úmysly</i> (=bad intentions) [negation = neg0]
------------------	---	---	--

<b>number:</b>	n.denot n.denot.neg n.pron.def.demon n.pron.def.pers n.pron.indef n.quant.def (the value of the sempos attributes contains: n)	tectogrammatical correlate of the morphological category of number	<i>psi</i> (=dogs) [number = pl] <i>otciův názor</i> (=my father's opinion) [number = sg] <i>náš</i> (=our) [number = pl] <i>Přišlo sto studentů.</i> (=One hundred students came.) [number = sg] <i>To, co potřebuješ, tu nemají.</i> (=They don't have what you need.) [number = inher] <i>jedny dveře</i> (=door) [number = sg] <i>Výjste se nepřihlásil?</i> (=You haven't registered?) [number = sg]
----------------	--	--	---

<b>numertype:</b>	n.quant.def adj.quant.def adj.quant.indef adj.quant.grad (abbreviation: quant)	it captures the semantic feature in which the given numeral is distinct from the corresponding cardinal numeral, by the t-lemma of which it is represented at the tectogrammatical level	<i>Koupil dvě z nabízených knih.</i> (=He bought two of the offered books.) [numertype = basic] <i>Natřel troje dveře.</i> (=He has painted three doors.) [numertype = basic] <i>Koupil setinu akcií.</i> (=He bought one hundredth of the shares.) [numertype = frac] <i>Ztratil už troje klíče.</i> (=He has already lost three sets of keys.) [numertype = set] <i>Má dvojí občanství.</i> (=He has two citizenships.) [numertype = kind] <i>Kolikátý pokus jsi provedl?</i> (=How many experiments have you done already?) [numertype = ord] <i>Vysvětluje to už podruhé.</i> (=This is the second time he's explaining that.) [numertype = ord]
-------------------	--	--	--

<b>person:</b>	n.pron.def.pers n.pron.indef (abbreviation: n.pron)	tectogrammatical correlate of the morphological category of person	<i>Já už jdu.</i> (=I am coming.) [person = 1] <i>Tvůj názor nesdílím.</i> (= I don't share your view.) [person = 2] <i>Svůj názor ti neřeknu.</i> (=I'm not telling you my opinion.) [person = inher] <i>Někdo to udělat musí.</i> (=Somebody has to do it.) [person = 3] <i>Zachraň se, kdo můžeš.</i> (=lit. Save REFL who (you) can.) [person = 2]
----------------	---	--	--

<b>politeness:</b>	n.pron.def.pers	it signal the use of the polite form	<i>Já dnes nepřijdu.</i> (=I'm not coming today.) [politeness = basic] <i>Výjste se ještě nepřihlásil.</i> (=You haven't registered.) [politeness = polite] <i>Svého psa jste ještě neviděl.</i> (=You haven't seen your dog yet.) [politeness = inher] <i>Ký tam nepřijdete?</i> (=You are not going there?) [politeness = nr]
--------------------	-----------------	--------------------------------------	--

Table 5.7. Survey of verbal grammatemes

Grammateme Values	Verb forms assigned one of the basic values	Definition of the grammateme	Examples
<b>aspect:</b> proc cpl nr	all verb forms	tectogrammatical correlate of the morpho-lexical category of aspect	<i>Nejraději kupuje nábytek.</i> (=He prefers to buy furniture.) [aspect = proc] <i>Nakoupil už vše potřebné.</i> (=He has bought everything he needed.) [aspect = cpl] <i>Císař tam vždy dobrovolně abdikovali.</i> [aspect = proc] (=The Emperors always abdicated voluntarily) <i>Císař dobrovolně abdikoval.</i> (=The Emperor abdicated voluntarily.) [aspect = cpl]
<b>deontmod:</b> deb hrt vol poss perm fac decl nr	all verb forms	it captures the fact whether the event is understood as necessary, possible, permitted etc.	<i>Musíme zaplatit fakturu včas.</i> (=We have to pay the invoice in time.) [deontmod = deb] <i>Petr ti měl podklady poslat už včera.</i> (=Petr was supposed to send you the documents already yesterday.) [deontmod = hrt] <i>Chtěl odejít, rozloučila se.</i> (=Willing to leave, she said good bye.) [deontmod = vol] <i>Moc tak odejít!</i> (=If only I could leave!) [deontmod = poss] <i>Nesmíš kouřit.</i> (=You mustn't smoke.) [deontmod = perm] <i>Umi se výborně přetvařovat.</i> (=She is very good at pretending.) [deontmod = fac] <i>Přišel na schůzku včas.</i> (=He came to the meetings in time.) [deontmod = decl]
<b>dispmod:</b> disp0 disp1 nil nr	non-imperative finite verb forms	it specifies whether the node is representing the predicate of a clause with dispositional modality	<i>Tato studie se studentům četla dobře.</i> (=lit. This article REFL student.DAT read well; apprx. It was easy for the students to read the article.) [dispmod = disp1] <i>Tato studie se čte dobře.</i> (=lit. This study reads well.) [dispmod = disp1] <i>Spalo se mu tu výborně.</i> (=lit. Slept REFL him.DAT here excellently; apprx. He slept very well here.) [dispmod = disp1] <i>Spí dobré.</i> (=He sleeps well.) [dispmod = disp0]
<b>iterativeness:</b> it0 it1 nr	all verb forms	it specifies whether the event is presented as iterative, multiple	<i>Chodíval k nám často.</i> (=He used to come to see us often.) [iterativeness = it1] <i>Jde zítra plavat.</i> (=He is going swimming tomorrow.) [iterativeness = it0]

<b>resultative:</b> res0 res1 nr	all verb forms	it specifies whether the event is presented as resultative	<u>Uvaril.</u> (=He has cooked (the meal).) [resultative = res0] <u>Měl oběd uvařen.</u> (=lit. (He) had dinner cooked.) [resultative = res1] <u>Má uvařeno.</u> (=lit. (He) has cooked.) [resultative = res1]
<b>tense:</b> sim ant post nil nr	non-imperative finite verb forms transgressive (gerund)	tectogrammatical correlate of the morphological category of tense	<u>Píše dopis.</u> (=He is writing a letter.) [tense = sim] <u>Bude psát dopis.</u> (=He is going to write a letter.) [tense = post] <u>Napiše dopis.</u> [tense = post] (=He will write/will have written a letter) <u>Rád by se díval na tu inscenaci.</u> (=I would like to watch the programme.) [tense = sim] <u>Hlasitě zanaříkavši, odcházela.</u> (=After emitting a loud cry, she was leaving.) [tense = ant] <u>Zitra jedu do Brna.</u> (=Tomorrow I am going to Brno.) [tense = sim]
<b>verbmod:</b> ind imp cdn nil nr	finite verb forms	tectogrammatical correlate of the morphological category of mood	<u>Studenti přišli na schůzi včas.</u> (=The students came to the meeting in time.) [verbmod = ind] <u>Přijdte na schůzi včas!</u> (=Come to the meeting in time!) [verbmod = imp] <u>My bychom přišli určitě včas.</u> (=We would definitely come in time.) [verbmod = cdn] <u>Hlasitě naříkajíc, odcházela.</u> (=She was leaving, crying loudly.) [verbmod = nil]

## 5.3. The sentmod attribute

The `sentmod` attribute (see Section A.2.18, “sentmod”) captures meanings similar to that of grammatemes but it is assigned to the node on the basis of its position in the tree and not according to the value of its `nodetype` and `sempos` attributes. It is assigned to the following nodes:

the root of a sentence,  
the root of a subtree representing direct speech (see Section 7.5, “Direct speech”),  
the root of a subtree representing a (syntactically independent) parenthesis, the effective roots of which are assigned the `PAR` functor (see Section 6.5, “Parenthesis”).

The `sentmod` attribute contains the information regarding the sentence modality. The value `enunc` corresponds to declarative clauses, `excl` to exclamative clauses, `desid` to optative clauses, `imper` to imperative clauses and `inter` to interrogative clauses.

Examples:

Petr nepřišel. [sentmod = enunc] (=Petr didn't come.)

Škola. [sentmod = enunc] (=School.)

Vyhrali jsme! [sentmod = excl] (=We won!)

Kéž by nepřišli! [sentmod = desid] (=I wish they didn't come!)

*Hodně štěstí!* [sentmod = desid] (=Good luck!)

*Přijd'te včas!* [sentmod = imper] (=Come in time!)

*Pozor!* [sentmod = imper] (=Watch out!)

*Zavolali jste už lékaře?* [sentmod = inter] (=Have you called the doctor?)

*Půjdete ven nebo zůstanete tady?* [sentmod = inter] (=Will you go out or will you stay here?)

---

# Chapter 6. Sentence representation structure

## 6.1. Dependency

The tectogrammatical level is based on the dependency conception. The basic idea of the dependency conception is the fact that the use of the **dependent element** is determined by the use of the **governing element**, which stands for the entire combination (the governing part has its syntactic distribution identical to that of the entire combination of the governing and the dependent part). Dependency is reflected in the morphological form of the dependent elements (i.e. by agreement in morphological categories between the dependent and the governing part or in the case of the dependent element). In accordance with some of the new syntactic approaches, the verb is considered the core of the sentence and the subject is taken to be dependent on the verb.

**Representing dependency in tectogrammatical trees.** A dependency relation between two elements in a tectogrammatical tree is primarily indicated by an edge between two nodes that goes from the node representing the governing element (governing node) to the node representing the dependent element (dependent node).

The PDT tectogrammatical trees differ from a dependency tree as defined in the theory, in which each edge represents a dependency between two elements, and each dependency between two elements is represented by an edge, in the following:

- representation of the second dependency with predicative complements, which is expressed by an attribute of the type reference (see Section 6.1.1, “Dual dependency”);
- existence of non-dependency edges (see Section 6.1.2, “Non-dependency edges”),
- cases of ambiguous dependency relations, in which an edge between two nodes does not reflect exact dependency relations in the sentence (see Section 6.1.3, “Ambiguous dependency”).

### 6.1.1. Dual dependency

The term **dual dependency** is used for such cases in which a modification (valency modification, or free modification) participates in a dual semantic dependency relation, i.e. it simultaneously modifies a noun and a verb (which can be nominalized). The dependency on the noun can also be expressed formally (by agreement in grammatical categories). Two cases are distinguished:

- **dual dependency of an argument.** Valency modifications (both prepositional and non-prepositional) with dual dependency are represented as arguments of the governing verb and their functor is usually PAT or EFF (i.e. their dependency on the verb is represented by an edge); their dependency on the noun follows from the meaning of the verb, which is described by its valency frame.
- **dual dependency of an adjunct - predicative complement.** See Section 6.1.1.1, “Predicative complement”.

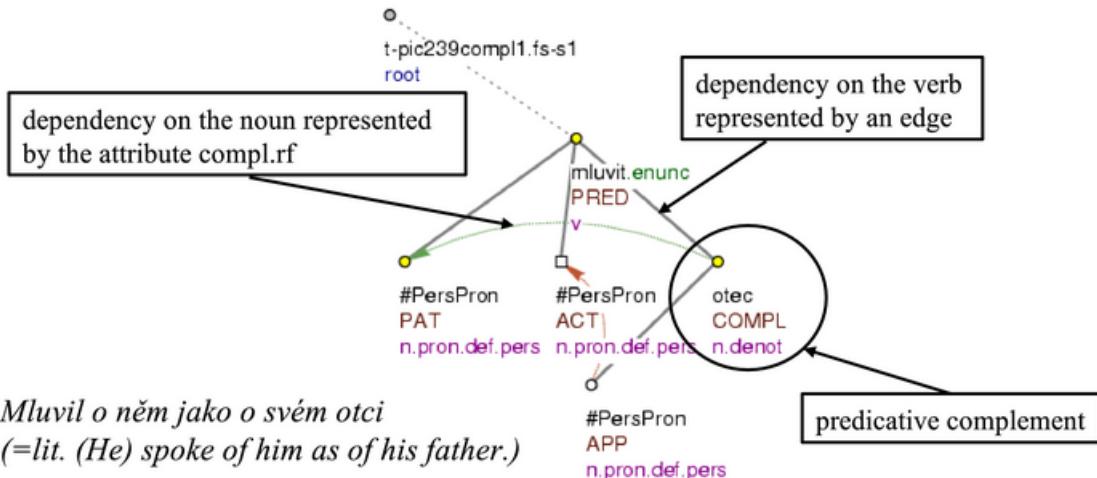
#### 6.1.1.1. Predicative complement

**Predicative complements** are (optional) adjuncts that have two semantic dependency relations, i.e. they simultaneously modify a noun and a verb (which can be nominalized).

**Representing predicative complements in tectogrammatical trees.** Nodes representing predicative complements have always the COMPL functor (see Section 8.11, “Functor for the predicative complement (COMPL)”). The two dependency relations of the predicative complement (functor = COMPL) are represented by the following means:

- the dependency on the verb is represented by an edge,
- the dependency on the (semantic) noun is indicated with the help of the attribute `compl.rf` (see Section A.2.2, “`compl.rf`”), the value of which is the identifier of the modified noun.

Representing predicative complements in tectogrammatical trees is illustrated in Fig. 6.1.

**Figure 6.1. Structure containing a predicative complement**


A predicative complement can be a noun, adjective, numeral, non-finite verb form (participle, transgressive, infinitive) or a dependent clause (introduced by the conjunction *jako*). A predicative complement expressed by a noun (adjective, numeral) can be introduced by the conjunctions *jako*, *jakožto*, *coby* or it can be non-prepositional. Adverbial modifications and prepositional phrases are never considered predicative complements.

Examples:

*Našli kamaráda nemocného.COMPL* (=They found their friend ill.) Fig. 6.2

*Pozvali toho chlapce jako představitele.COMPL hnuti.* (=They invited the boy as a representative of the movement.)

*Kluci přišli tři.COMPL* (=lit. (The) boys came three.)

*Odcházela poražena.COMPL* (=She was leaving defeated.) Fig. 6.3

*Odešel, zpívají si.COMPL* (=He left, singing to himself.)

*Našel Karla ležet.COMPL na posteli.* (=He found Karel lying on his bed.)

*Matka našla dítě, jak spí.COMPL* (=Mother found her child sleeping.) Fig. 6.5

**Construction of the type “seděl hlavu skloněnou”** The group of transgressive predicative complements includes also cases like “seděl hlavu skloněnou (=he was sitting with his head bowed)”. These are constructions in which the verb is followed by a transgressive construction with the governing transgressive of the verb *mít* (=have) elided. The ellipsis is represented by a newly established node for an empty verb (t-lemma #EmpVerb, functor = COMPL).

*Seděla hlavu {#EmpVerb.COMPL} skloněnou.* (=She was sitting with her head bowed) Fig. 6.4

**NB!** Such a newly established node is used also in some cases of direct speech (t\_lemma = #EmpVerb, functor = COMPL, here, this concerns the transgressive form of the verb *říci* (say)). For more details see Section 7.5, “Direct speech”.

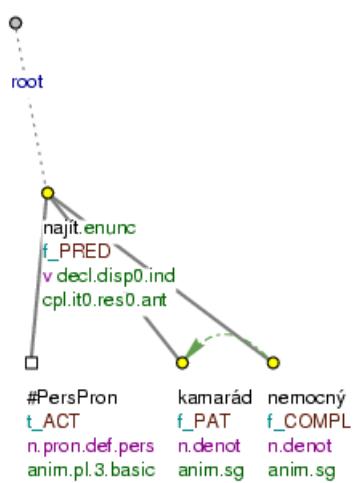
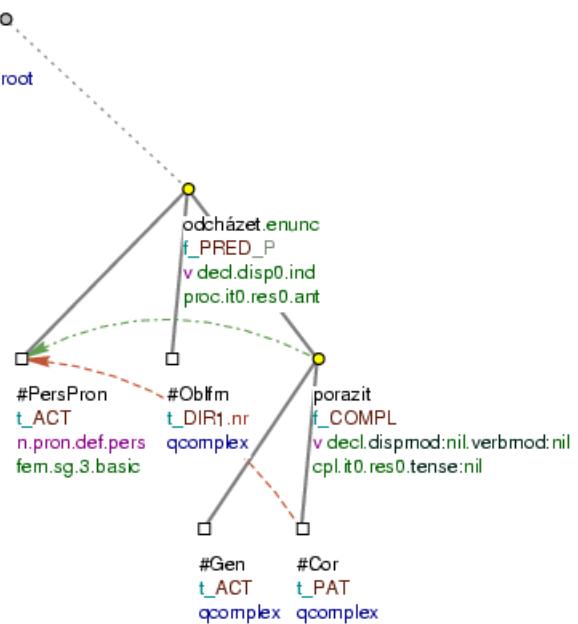
**Nominalization of the governing verb.** Constructions with predicative complements where the verb is nominalized are represented in the same way. For example:

*Poslední volby vyhrál s programem postaveným jako negace.COMPL programu minulého.* (=...conceived of as the negation of the former program.) Fig. 6.6

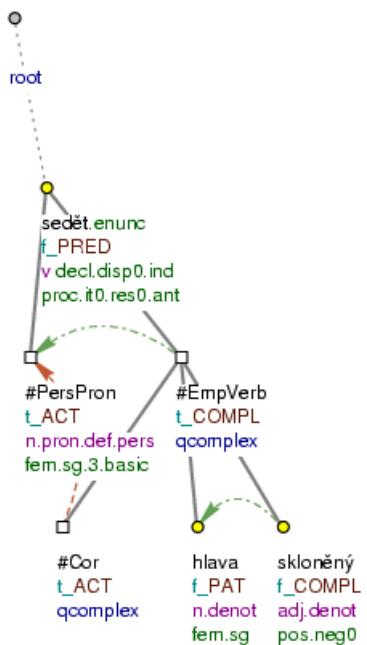
*Předání domu coby záruky.COMPL proběhlo bez problémů.* (=The handing over as a guarantee was without problems.)

**Figure 6.2. Predicative complement**

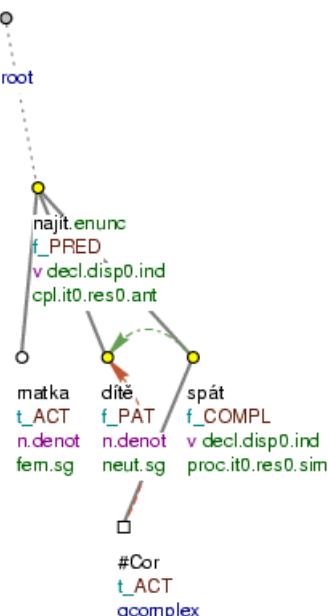
*Našli kamaráda nemocného.* (=lit. (They) found (their) friend ill)


**Figure 6.3. Predicative complement**

**Figure 6.4. Predicative complement**

*Seděla hlavu skloněnou.* (=lit. (She) was\_sitting head bowed)

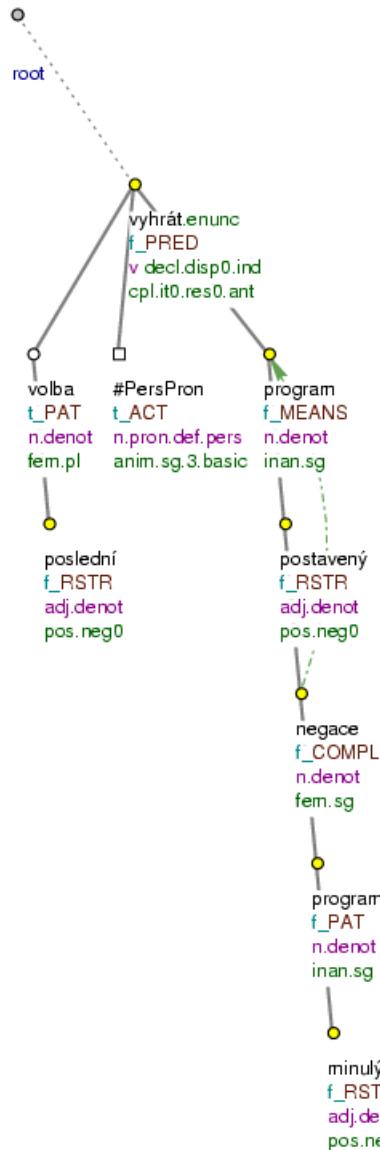

**Figure 6.5. Predicative complement**

*Matka našla dítě, jak spí.* (=lit. Mother found child as (it) sleeps)



**Figure 6.6. Predicative complement**

*Poslední volby vyhrál programem postaveným jako negace programu minulého. (=lit. Last election (he) won with program build as negation (of) program former)*



### 6.1.2. Non-dependency edges

Some edges in a tectogrammatical tree do not represent dependency. We establish such non-dependency edges in a tectogrammatical tree in order to represent parataxis and some other specific syntactic relations.

A non-dependency edge is:

- A. the edge between the root node of the represented sentence and the technical root node of the tectogrammatical tree (`nodeType = root`).

It is an auxiliary edge (of technical nature) without any linguistic interpretation.

- B. the edge between the effective root node of an independent clause and its mother node.**

Functors for effective root nodes of independent clauses (PRED, DENOM, PARTL, VOCAT, PAR; see Section 8.1, “*Functors for effective roots of independent clauses*”) express non-dependency and they determine the clause type. The edge to the mother node only integrates the particular nodes (subtrees) into a tectogrammatical tree.

- ### C. edges in paratactic structures:

- a. the edge between the root of the paratactic structure (`nodeType` = `coap`) and its mother node,
  - b. the edge between a paratactic structure root node and a direct element of the paratactic structure.

- c. the edge between a paratactic structure root node and the root of a shared modifier.

Dependency between modifications within a paratactic structure is always represented by two edges at least. For example, the dependency of a terminal element of a paratactic structure on its governing node is indicated - in a simple, non-embedded paratactic structure - by a combination of type a) and b) edges. In an embedded paratactic structure, the dependency of a terminal element of the paratactic structure on its governing node is indicated by a set of b) type edges and one a) type edge. Dependency of a shared modifier of terminal elements is indicated by a combination of type b) and c) edges. For more on paratactic structures (including definitions of the terms used) see Section 6.4, “Parataxis”.

D. **edges in list structures:**

- a. the edge between the root of a list structure (`nodetype = list`) and its mother node.
- b. the edge between the root of a list structure and an item of the list (`nodetype=fphr`) or the root of an identifying expression (either `functor = ID`, or `nodetype = coap`).
- c. the edge between the root node of the list structure and the root node of the modifier of the list.

Edges between nodes in list structure have various meanings depending on the type of the list structure. In a list structure for foreign-language expressions, type b) edges only collect individual nodes into a list (they do not express dependency) while type a) edges express dependency of the entire list structure (the foreign-language segment as a whole) on its governing node. In a list structure for identifying expressions, dependency is represented by both type a) and b) edges. Type a) edges express dependency of the entire identification structure; type b) edges express a special kind of dependency of the effective root of the identifying expression (the nominative of identity). Type c) edges represent in both cases dependency on all items of the list as a whole.

For more on list structures for foreign-language expressions see Section 7.4, “Foreign-language expressions”. For more on list structures for identifying expressions see Section 7.3.1, “Identification structure”.

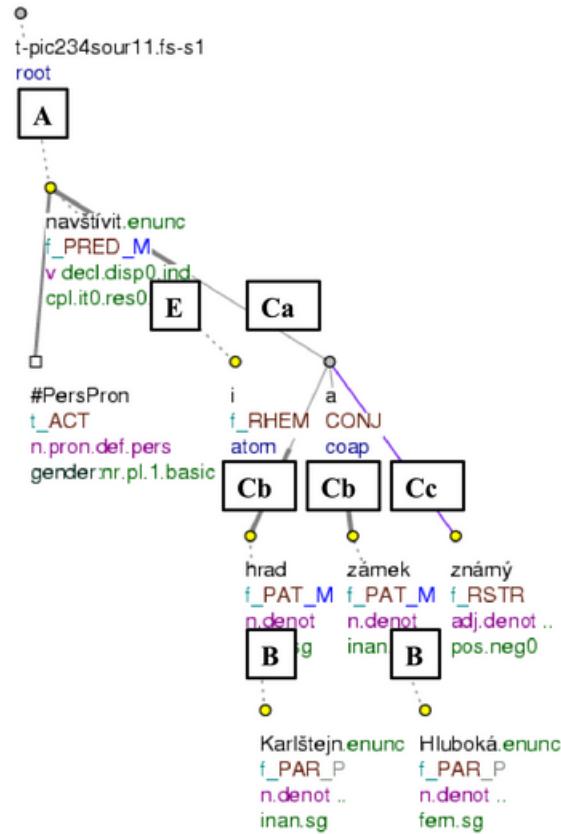
E. **the edge between an atomic node and its mother node.**

Edges above atomic nodes (`nodetype = atom`) integrate these nodes into the tree. Their meaning varies depending on the functor of the particular atomic node. For more on the functors of atomic nodes see Section 8.7, “Functors for rhematizers, sentence, linking and modal adverbials”. For more on rhematizers see Section 10.4, “Rhematizers”.

F. **the edge between a node the functor of which is DPHR or CPHR and its mother node.**

A node the functor of which is DPHR or CPHR expresses the fact that it constitutes a single lexical item together with its mother node (such a lexical item is usually represented by a single node). Therefore, the edge expresses rather the fact that the expressions belong together than dependency. For more details see Section 7.1, “Multi-word lexical units”.

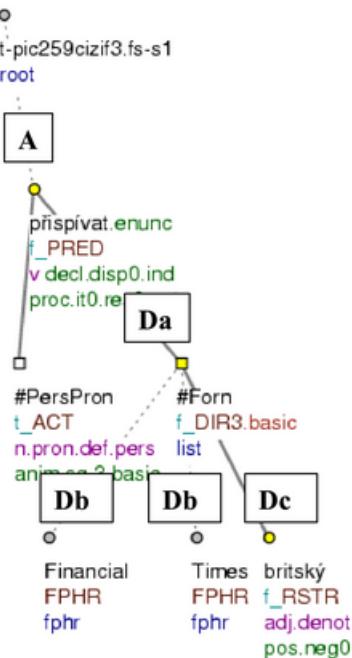
Figure 6.7. Examples of non-dependency edges I



Navštívíme i známý hrad (Karlštejn) a zámek (Hluboká).

(=lit. (We) shall\_visit also famous castle (Karlštejn) and manor (Hluboká).)

Figure 6.8. Examples of non-dependency edges II



Přispíval do britských Financial Times.

(=lit. (He) was\_contributing to British Financial Times.)

### 6.1.3. Ambiguous dependency

It is not always unambiguous what certain adjuncts (expressed by adverbs or prepositional phrases) are dependent on: they do not have to modify only one modification in the sentence but rather they can modify several modifications at the same time. Precise rules for cases of ambiguous dependency are still to be proposed; the following present only a temporary solution.

The basic annotation rule is as follows:

- if a free modification (expressed by a prepositional phrase or adverb) modifies a verb, it is dependent on this verb whether it enters into other semantic relations or not.

The cases of the so-called dual function are the only exception.

Examples:

*Starý muž přišel v otrhaném kabátě.* (=The old man came in a shabby coat.)

*Přecházel po pokoji neklidně.* (=He was pacing the room restlessly.)

**Dual function of a single modification.** In those structures in which a modification has a dual (or multiple) function (i.e. it modifies several modifications at the same time but it is present only once in the surface structure, for stylistic or other reasons), such a modification depends on the node representing the lowest modification modified by it and it is assigned the functor adequate to its real position. There is no explicit indication that the modification has a dual function. Například:

*Koupila jsem si pásek za sedmdesát korun.* (=I bought a belt for seventy crowns.)

= *Koupila jsem si za sedmdesát korun.MEANS pásek za sedmdesát korun.RSTR* (=lit. (I) bought AUX myself for seventy crowns (a) belt for seventy crowns.) Fig. 6.9

*Splatil dluhy pojišťovně.* (=He repaid his debts to the insurance company.)

= *Splatil pojišťovně.ADDR dluhy pojišťovně.ADDR* (=lit. (He) paid\_back (to) (the) insurance company (his) debts (to) (the) insurance\_company.) Fig. 6.10

**Mutual relation of two or more locative/directional or temporal modifications.** Also modifications with the same function (temporal or locative/directional), adjacent in the surface word order, enter into unclear semantic relations. In principle, there are the following three cases:

- a. **apposition of two temporal or locative/directional modifications.** Only those combinations are considered appositions in which the individual modifications are separated by a comma or appositional conjunction; e.g.:

*Zůstal doma, v Krkonoších.* (=He stayed at home, in Krkonoše.)

- b. **one temporal or locative/directional modification dependent on another.** The following cases are considered cases of one modification dependent on another:

- **one modification has the genitive form.** For example:

*Přijel ve čtvrté 5. ledna.APP 1997.* (=He came on Thursday 5th of January.GEN) Fig. 6.11

- cases of the so-called **extent or time-span accusative** further modified by a prepositional phrase. For example:

*Leží to dva kilometry.LOC od řeky.DIR1* (=It is 2 km far from the river.) Fig. 6.12

*Oblékla se půl.TWHEN hodiny před začátkem.TWHEN představení.* (=She got dressed half an hour before the beginning.)

- one modification is required by the valency of the other. Similar to these are cases in which the second modification (expressed by a prepositional phrase) is a **more or less valency modification** with respect to the preceding modification (expressed by an adverb). For example:

*Přijel brzy.TWHEN po Vánocích.TWHEN* (=He came soon after the Christmas.)

*Odehrálo se to daleko.LOC od Moskvy.DIR1* (=It took place far from Moscow.)

- c. **several sister modifications.** If there are two (or more) temporal, locative/directional (or other) modifications present in the construction at the same time and each of the modifications is relatively independent, the entire construction is represented as involving two or more sister modifications dependent on the same governing node (with the same functor): their order can be changed without any change in meaning; any of the modifications can be omitted without any damage to the grammatical structure of the sentence. Both modifications usually relate to the same situational moment and to the same location; one of them gives a general information and the other one is more specific. For example:

*Sejdeme se na Hlavním nádraží.LOC v hale.LOC (=We'll meet at the Main Station in the hall.) Fig. 6.13*

*Přijeli v únoru.TWHEN v roce.TWHEN 1999. (=They came in February 1999.)*

**Dependency relations in noun phrases (concord of two nouns).** Also noun phrases which consist of a sequence of nouns in the same form (not in apposition) present a case of ambiguous semantic relations. Precise rules are available for two-member noun phrases, in which one of the nouns is a proper noun, and for bigger noun phrases denoting persons, in which one of the nouns in the sequence is the name of the person:

- a. **the noun phrase is a name of a person.** If a two-member noun phrase is a name of a single person, the node representing the proper noun is the governing node of the entire phrase. The node representing the common noun depends on the node for the proper noun and has the RSTR functor; e.g.:

*Dej to našemu řediteli.RSTR Novákovi . (=Give it to our director Novák.) Fig. 6.14*

**Bigger noun phrases.** In bigger noun phrases denoting persons in which one of the nouns is the name of the person, all common nouns depend on the proper noun and have the RSTR functor; e.g.:

*o nebožtíku.RSTR panu.RSTR kormidelníkovi.RSTR Janu.RSTR Landgermanovi (=lit. about deceased Mr. steersman Jan Landerman.)*

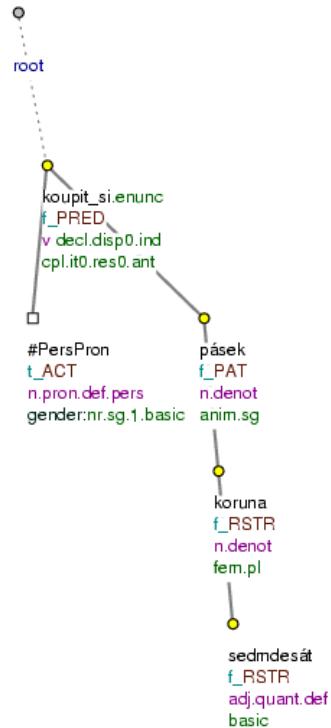
- b. **the noun phrase is not a name of a person.** If the noun phrase is a name of an animal, an inanimate object or another phenomenon, the governing node is a common noun. The node for the proper noun depends on the common noun and has the RSTR functor; e.g.:

*Na řece Vltavě.RSTR jezdí parníky. (=There are steamboats on the Vltava river.) Fig. 6.15*

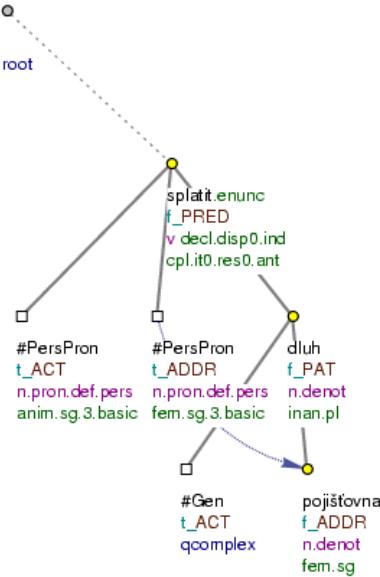
Other types of noun phrases with several nouns in the same form (e.g.: *O nebožtíku panu kormidelníkovi se už nemluvilo.* (=lit. About deceased Mr. steersman REFL any\_more not\_talked)) were analyzed on the basis of the context and the rules above but the decision was up to the annotator. The node representing one of the inflected nouns is chosen to be the effective root of the entire noun phrase. The other nodes depend on this effective root node as its modifications and their functor is RSTR.

**Figure 6.9. Dual function of a single modification**

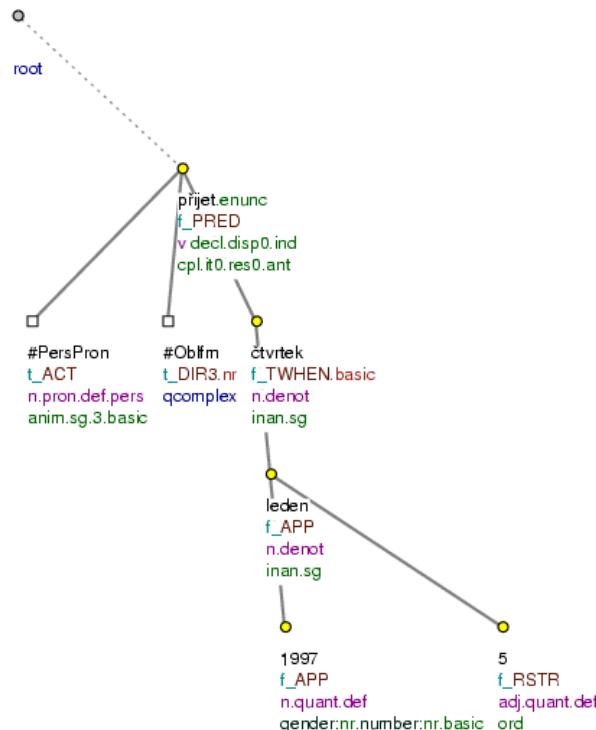
*Koupila jsem si pásek za sedmdesát korun. (=lit. (I) bough AUX REFL belt for seventy crowns.)*


**Figure 6.10. Dual function of a single modification**

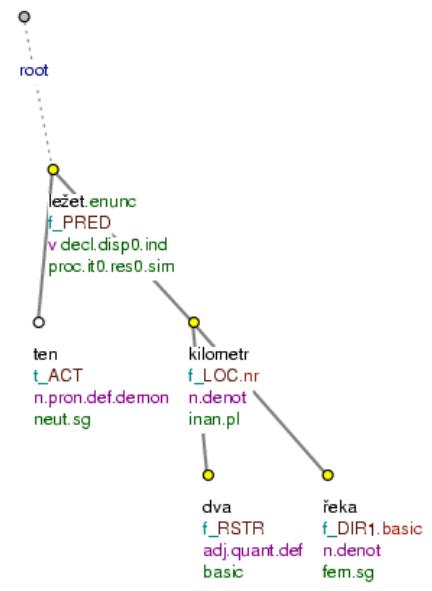
*Splatil dluhy pojišťovně. (=lit. (He) paid\_off (his) debts (to) (the) insurance\_company.)*


**Figure 6.11. Mutual relation of two temporal modifications**

*Přijel ve čtvrtku 5. ledna 1997. (=lit. (He) came on Thursday 5th (of) January 1997.)*

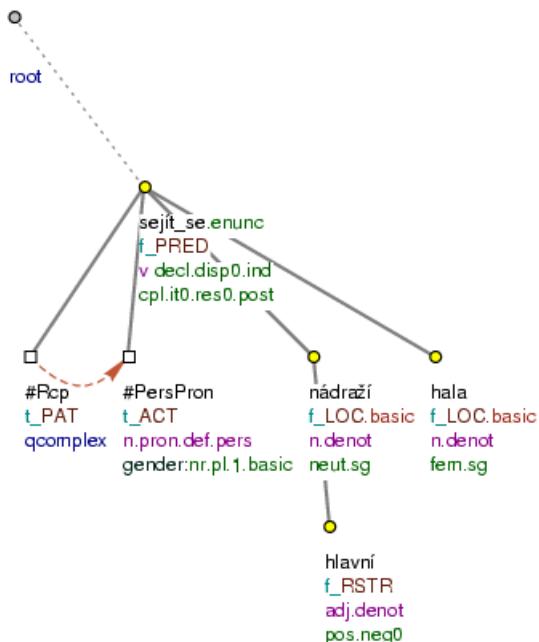

**Figure 6.12. Mutual relation of two locative/directional modifications**

*Leží to dva kilometry od řeky. (=lit. Is\_situated it two kilometers from river.)*

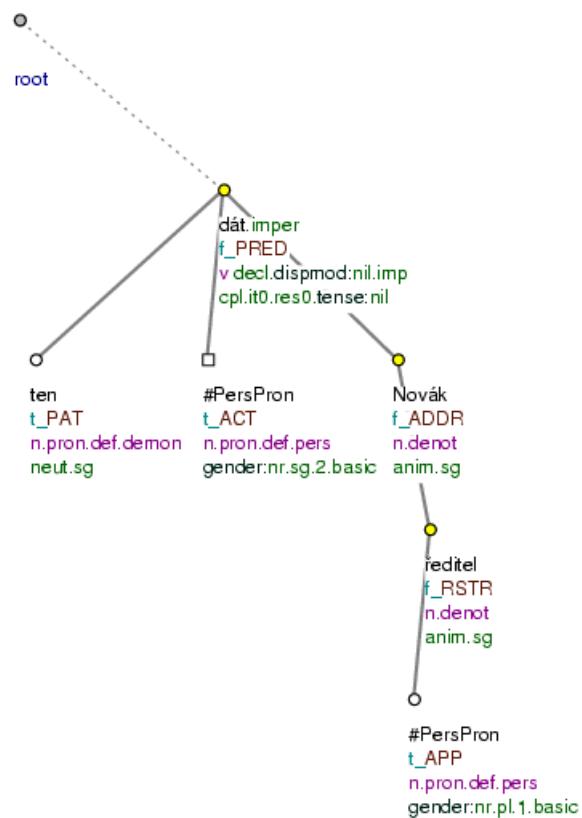


**Figure 6.13. Mutual relation of two locative/directional modifications**

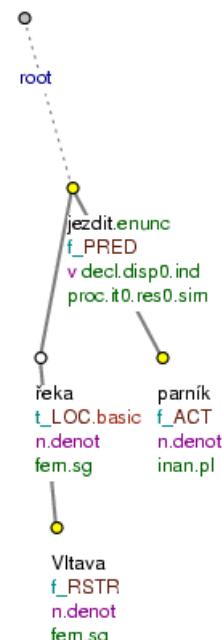
*Sejdeme se na Hlavním nádraží v hale. (=lit. (We) will\_meet REFL at Main Station in hall.)*


**Figure 6.14. Nouns in the same form**

*Dej to našemu řediteli Novákovi. (=lit. Give it (to) our director Novák.)*


**Figure 6.15. Nouns in the same form**

*Na řece Vltavě jezdí parníky. (=lit. On river Vltava go steamboats.)*



## 6.2. Valency

Valency modifications (in the broad sense of the word) include all kinds of elements that can modify a particular lexical unit (or rather a lexical unit in a particular meaning). The term **valency** is, however, used in its narrower sense here, namely: the **valency modifications** of a lexical unit are only its arguments and obligatory adjuncts (for the distinction see Section 6.2.1, “Criteria for distinguishing different kinds of modifications”). These modifications are always specified in the valency frame of the lexical unit (see Section 6.2.4, “Valency frames and the way they are recorded in the valency lexicon”).

### 6.2.1. Criteria for distinguishing different kinds of modifications

**Arguments and adjuncts.** Any modification can be classified as either an inner participant (argument) or free modification (adjunct) - according to the type of dependency:

- **free modifications** are such modifications that can - if it is not excluded for semantic reasons - modify any verb (word) and they can modify a particular verb (word) more than once.
- **inner participants** (arguments in the sequel) are such modifications that can modify a given verb only once (except for the case of coordination) and they only modify a more or less closed class of verbs that can be listed.

It seems that there are the following five (verbal) arguments:

Actor (ACT), Patient (PAT), Addressee (ADDR), Origo (ORIG), Effect (EFF).

With nouns, there is one more argument: MAT. Other types of verbal modifications are considered to be adjuncts, corresponding to temporal, locative/directional, manner and other kinds of adverbials (for the list of recognized functors and their definitions, see Chapter 8, *Functors and subfunctors*).

**Obligatory and optional modifications.** A certain type of modification is either an argument, or an adjunct in all its occurrences. With respect to its governing element, the given modification can be either **obligatory**, i.e. obligatorily present in the deep structure of the sentence, or **optional**, i.e. not necessarily present. The obligatory - optional distinction does not apply to the individual types of modifications directly; it expresses their relation to particular lexical units (their governing verbs/nouns/adjectives..).

For determining whether a given modification is obligatory or optional, the so called dialogue test is used. **The dialogue test** helps to determine which arguments and adjuncts are obligatory and which are optional. It is used whenever a modification is not present at the surface level but when it can be hypothesized that it is in fact (semantically) obligatory. The dialogue test is based on the difference between questions asking about something that is supposed to be known to the speaker - because it follows from the meaning of the verb he/she has used, and questions about something that does not necessarily follow from the meaning of the used verb. Answering a question about a semantically obligatory modification of a particular verb, the speaker - who has used the verb - cannot say: *I don't know*. Cf. the following dialogues.

**Obligatoriness of an argument:**

- a. A: *Když to viděl, koupil to.* (=When he saw it he bought it)  
B: *Kdo?* (=Who?)  
A: \**Nevím.* (=I don't know.)
- b. A: *Když to viděl, koupil to.* (=When he saw it he bought it.)  
B: *Komu?* (=For whom?)  
A: *Nevím.* (=I don't know.)
- c. A: *Když to viděl, koupil to.* (=When he saw it he bought it.)  
B: *Od koho?* (=From whom?)  
A: *Nevím.* (=I don't know.)

The verb *koupit* (=buy) has four arguments: Actor, Patient, Addressee and Origo. With the help of the dialogue test, it can be determined which of these arguments are obligatory and which are optional. In dialogue a), the speaker A cannot answer the question *Kdo?* (=Who?) by saying *Nevím* (=I don't know). The dialogue would not make sense, then. On the contrary, the speaker does not have to know answers to the questions *Komu?* (=For/to whom?) and *Od koho?* (=From whom?) in the dialogues b) and c). These modifications are contained in the meaning of the verb, but not necessarily; they are optional. The Patient is obligatory both at the surface and deep levels.

**Obligatoriness of an adjunct:**

- a. A: *Moji přátelé přijeli.* (=My friends have arrived)  
B: *Kam?* (=Where to?)  
A: \**Nevím.* (=I don't know.)
- b. A: *Moji přátelé přijeli.* (=My friends have arrived)  
B: *Odkud? Proč?* (=Where from? Why?)  
A: *Nevím.* (=I don't know.)

For the verb *přijet* (=come/arrive), the modification answering the question *Kam?* (=Where to?) is obligatory, which can be seen from the impossibility of answering the question by saying *Nevím* (=I don't know). The speaker has used the verb *přijet*, so it would make no sense to answer the question about the goal by saying *Nevím* (=I don't know). On the contrary, the speaker does not have to know answers to the questions *Odkud?* (=Where from?) and *Proč?* (=Why?) in dialogue b). The modification answering the question *Kam?* (=Where to?) is an adjunct; it is, however, obligatory for the verb *přijet*.

**Structure of a valency frame.** By combining the criteria for distinguishing between arguments and adjuncts with the criteria for distinguishing between obligatory and optional modifications, we get four possibilities. **Valency frames**, representing individual meanings of words, contain all arguments of the given word and those adjuncts that are obligatory in the given meaning (cf. the three pluses in Table 6.1, “**Structure of a valency frame**”).

Every verb has at least one valency frame - and often more, with one frame corresponding to one meaning of the verb.

**Table 6.1. Structure of a valency frame**

	Obligatory modifications	Non-obligatory (optional) modifications
Arguments	+	+
Adjuncts	+	-

## 6.2.2. Argument shifting principle

When determining the argument type, with the Actor (ACT) and Patient (PAT) we primarily use syntactic criteria; with the other arguments also semantic criteria. For the discussion of the semantics of the individual arguments (and their definitions), see Section 8.2, “Argument functors”. It holds that:

- A. **the first argument is always the Actor, the second one is the Patient.** From this, it follows that:
  - if a verb has only one argument, it is the Actor (ACT) regardless of its exact semantic relation to the verb.
  - if a verb has two arguments, they are the Actor (ACT) and the Patient (PAT).

**Determining the first and the second argument.** When determining which argument is the first one (i.e. the Actor), the basic rule is that the Actor occupies the subject position, i.e. the structural nominative position. Only if one of the arguments is in the dative case and the other one in the nominative case, the semantics of the arguments comes into account. If the argument in dative refers to the Experiencer (or Agent), we consider the argument the Actor and the argument in the nominative the Patient; cf.:

*Kníha.PAT se mi.ACT libila.* (=I liked the book.)

*Naše výrobky.ACT se vyrovnají cizím výrobcům.PAT* (=Our products are as good as products from other countries; lit. Our products.NOM REFL keep\_pace\_with/are\_a\_match\_for foreign products.DAT)

- B. if a verb has more than two arguments, **the semantics is important for determining the third and any other argument.**

As a consequence of this, the so called argument shifting takes place. **Argument shifting** (cf. Fig. 6.16):

- a. means that if a verb has no argument in its valency frame that bears the cognitive role of an Agent (or another role typical for the first participant - Actor), its position is taken up by the Patient (i.e. what would be assigned the Patient functor under usual circumstances); e.g.

*Kníha.ACT vyšla.* (=The book was published.)

- b. if a verb subcategorizing for two arguments has no argument that bears the cognitive role of a Patient, another argument takes up its position (=is assigned the Patient label/functor). The following rule applies:

- if a verb has a potential Addressee/Origo and a potential Effect but has no Patient-like argument, then the Patient position is taken up by the Effect-like argument. The Addressee and/or Origo-like arguments do not undergo any shifting. For example:

*Petr.ACT vykopal jámu.PAT (=Petr has dug out a hole.)*

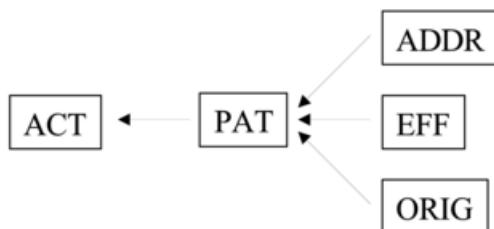
*Jan.ACT vyspěl z jinocha.ORIG v muže.PAT (=Jan grew up into a man - he is not a child any more; lit. Jan grew\_up from adolescent into man.)*

- if a verb has no Effect-like argument, the Patient position is taken up by the cognitive Addressee/Origo (i.e. they shift to the position of the Patient). For example:

*Učitel.ACT vyvolal žáka.PAT (=The teacher asked a pupil to answer a question; lit. Teacher called\_upon pupil.)*

*Z banálního nachlazení.PAT se vyvinulo závažné onemocnění.ACT (=A slight/banal cold developed into a serious illness; lit. From banal cold REFL developed serious illness.)*

Figure 6.16. Argument shifting principle



**NB!** The shifting only concerns the arguments. Adjuncts do not shift to argument positions. An adjunct that is obligatory for a given verb (according to the criteria in Section 6.2.1, “Criteria for distinguishing different kinds of modifications”) is always assigned an adjunct-like functor; e.g.:

*Petr.ACT přijel do Prahy.DIR3 (=Petr came to Praha.)*

The argument shifting applies to valency frames of all verbs, with the exception of complex predicates (for the discussion see Section 6.2.4.2, “Valency frames of idiomatic expressions and complex predicates”). Argument shifting also does not apply to valency frames of nouns and adjectives not referring to events and to valency frames of adverbs.

### 6.2.3. Relations between verb meanings and valency frames

Verbs usually have more than one meaning: **each meaning is assigned a separate valency frame**. This principle is violated in cases of. **competing valency modifications**. These are cases in which one of the valency positions may be occupied by modifications of different functors while the meaning of the verb is preserved. The potential competition arises either between an argument and adjunct or among different types of adjuncts. There are two ways to deal with these cases:

- the basic way is to introduce the concept of modification alternatives. So far, this is the solution adopted only for the cases of different types of manner adjuncts competing for the same position.

For example, the valency frame for one of the meanings of the verb *chovat se* (=behave):

ACT(1) MANN(\*)|CRIT(\*)|ACMP(\*)|BEN(\*)|CPR(\*)

*chová se laskavě.MANN (=his behavior is kind); ch. se podle pravidel.CRIT (=he behaves according to the rules); ch. se otrocky.CPR (=his behavior is slavish); ch. se bezchybně.ACMP (=his behavior is flawless); ch. se ku prospěchu věci.BEN (=his behavior is for the good)*

- in other cases of competition:

competition of the Addressee and directional modification (*odebrat děti rodičům.ADDR - odebrat děti od rodičů.DIR1* (=take away the children from their parents)),

competition between locative and directional modifications (*umístit obrázek na nástěnu.DIR3 - umístit obrázek na nástěnce.LOC* (=place the picture on the notice board))

the given meaning of the verb is assigned as **many valency frames as there are competing modifications**.

The basic principle: one meaning - one valency frame is violated here.

Cf. the three valency frames for one of the meanings of the verb *podat* (=submit):

ACT(.1) PAT(.4) ADDR(.3)

*podali své listiny úřadu* (=they submitted their documents at the office)

ACT(.1) PAT(.4) DIR3(\*)

*podali své listiny na úřad*

ACT(.1) PAT(.4) LOC(\*)

*podali své listiny na úřadě*

## 6.2.4. Valency frames and the way they are recorded in the valency lexicon

A valency frame record in the valency lexicon (PDT-VALLEX) is a sequence of records of the individual valency modifications (types of dependents), separated by spaces. Valency modification alternatives (see Section 6.2.3, “Relations between verb meanings and valency frames”) are separated by the | mark. The lexical meaning linked to a given valency frame is specified by examples; often, synonyms and antonyms are provided, too, or aspectual counterparts, if possible. In the example part of a valency frame record, one can also occasionally find so called typical adjuncts, i.e. those modifications that are not required (semantically obligatory) but which are typical for the given verb (noun, adjective) in the given meaning.

Valency modifications (in a valency frame) are presented in the following order: ACT, CPHR, DPHR, PAT, ADDR, ORIG, EFF, BEN, LOC, DIR1, DIR2, DIR3, TWHEN, TFRWH, TTILL, TOWH, TSIN, TFHL, MANN, MEANS, ACMP, EXT, INTT, MAT, APP, CRIT, REG. A valency modification record contains the information regarding the functor and surface form of the given modification (see Section 6.2.4.1, “Surface form of a valency modification”). The question mark preceding a functor specification indicates optionality; if the question mark is not present, the modification is obligatory.

Cf. the valency frame for one of the meanings of the verb *zmenšovat* (=shrink/decrease):

ACT(.1) PAT(.4) ?ORIG(z+2) ?EFF(na+4)

*zmenšovat nájem z 8 na 6 tisíc* (=to reduce the rent from 8 to 6 thousand)

*z. objem odpadu o přijatelné procento.DIFF* (=to reduce the waste volume by a reasonable amount/per cent)

**Empty valency frames.** Valency frames may also be empty, i.e. they may contain no valency positions. Such a valency frame is recorded as EMPTY.

### 6.2.4.1. Surface form of a valency modification

The surface form of a valency modification is the form as found at the analytical level. The surface form specification involves:

- A. **indication of the syntactic dependency.** To indicate dependency, square brackets ( [ ] ) are used; sister nodes are separated by a comma ( , ). The notation is, then:

governing-node[dependent-node1, dependent-node2]

- B. **indication of the part-of-speech and morphemic properties.** The requirements regarding the word class (part of speech) and morphemics of individual nodes are encoded in an abbreviated form (using one symbol for each category), introduced after a period or colon, in the following order: part of speech, gender, number, case, degree. If a surface-level category is not specified, it means that the given valency modification may get any value of the category.

In some cases, surface-form specifications include also information regarding the analytical forms (lemmas) of certain dependent (analytical) nodes; prepositions, subordinating conjunctions and also dependent parts of

idiomatic expressions. For the sake of simplicity, when specifying what kind of prepositional phrase is required by a given verb, an abbreviated form is used. For example, *na+4* is short for: *na-1[.4]*.

Examples of surface-form specifications:

accusative:.4

adjective in the instrumental: .a7

possessive pronoun or adjective: .u

numeral: .m

pronoun: .p

infinitive: .f

adverb: .d

interjection: .i

direct speech: .s

masculine: .M

feminine: .F

neuter: .N

singular: .S

plural .P

(asyndetic) content clause (a subordinate clause beginning with a relative pronoun/adverb): .c

dependent clause, with any kind of conjunction: j[.v]

dependent clause with the conjunction *že*: že[.v]

preposition *o* and a noun in the locative: o+6

the multi-word preposition *na rozdíl od* plus a noun in the genitive: od[na,rozdíl,.2]

A surface form of an obligatory adjunct is usually not specified, which means that usual forms can be used. This is indicated by the star symbol ( \* ), which is used instead of the explicit specification of the surface form. With arguments, the surface forms are always specified.

**Regular changes in the surface form (not indicated in the valency frame).** A surface-form specification contains all variants found in the analyzed data. A number of surface forms are the result of a productive (syntactic) process, however. These derived forms are not specified in the valency frame. This concerns especially the following cases:

- passivization.** Valency frames only specify those surface forms that occur in active sentences. For example:

*Stavební firma*.ACT *postavila dům*.PAT (=The building company has built a house.)

Passive: *Dům*.PAT *byl postaven stavební firmou*.ACT (=The house.NOM was built by a building company.IN-STR)

The valency frame of the verb *postavit* (=build):ACT(.1) PAT(.4) ?ORIG(z+2)

- resultative.** The surface form variants that are the result of a verb occurring in a resultative construction are not specified in the verb's valency frame. For example:

*Otec*.ACT *pronajal auto sousedovi*.ADDR (=Father rented out a car to a neighbour.)

Resultative: *Soused*.ADDR *má auto pronajato od otce/otcem*.ACT (=lit. Neighbour:NOM has car rented from/by Father.)

The valency frame of the verb *pronajmout* (=rent out): ACT(.1) PAT(.4) ADDR(.3)

- dispositional modality.** The surface form variants that are the result of a verb occurring in a construction with the dispositional modality meaning are not specified in the verb's valency frame. For example:

*Žáci*.ACT *počítají příklady*.PAT (=The pupils are doing exercises.)

Dispositional modality construction: *Příklady*.PAT *se žákům*.ACT *počítají dobře*.MANN (=lit. Examples.NOM REFL pupils.DAT count/do well.)

The valency frame of the verb *počítat* (=count): ACT(.1) PAT(.4,že[.v],zda[.v],jestli[.v],.v[kolik])

- forms used for expressing subtle shifts in the meaning of arguments.** The basic form of an argument (e.g. nominative for the Actor or accusative for the Patient) may be replaced by another form if a slightly different/more specific meaning is to be expressed. These forms are used for a given meaning regularly, therefore they are not mentioned as possible forms of individual valency modifications. For example:

*Deset knih*.ACT *leží na stole*. (=Ten books are lying on the table.)

Okolo deseti knih.ACT leží na stole. (=Around ten books are lying on the table.)

Kolem deseti knih.ACT leží na stole. (=Around ten books are lying on the table.)

Přes deset knih.ACT leží na stole. (=More than ten books are lying on the table.)

K deseti knihám.ACT leží na stole. (=Something like ten books are lying on the table.)

Po deseti knihách.ACT leží na stole. (=Ten books are lying on each table.)

The valency frame of the verb *ležet* (=lie): ACT(.1) LOC(\*)

The presented meanings (partitivity, distributivity, approximation) are going to be represented by subfunctors (assigned to the arguments) in a future version of PDT.

- e. **reciprocity.** The fact that there is a reciprocal meaning in the sentence is signalled by the presence of *se* (*mezi sebou, k sobě* (=lit. among themselves, to themselves; meaning: with/to/... each other)); a typical form used for expressing reciprocity is the form *mezi+7*. These expressions are understood as formal means of expressing reciprocity; they are not recorded in valency frames (i.e. in their surface-form specification part). For example:

*jednání premiéra s prezidentem* (=lit. negotiation (of) Prime minister with President) → *jednání mezi premiérem a prezidentem* (=lit. negotiation between Prime minister and President)

The valency frame of the noun *jednání* (=negotiation): ACT(.2,.u) PAT(o+6) ADDR(s+7)

- f. **numeral+noun constructions.** Certain numeral+noun constructions (see Section 7.2, “Numbers and numerals”) are analyzed in such a way that the formally dependent noun (in the genitive) is understood as the governing node of the construction whereas the formally governing numeral is taken to be the dependent node. The genitive form is not included in the surface form specification. For example:

*Dívky*.ACT koupily dětem čokoládu.PAT (=The girls bought the children some chocolate.ACC)

*Pět* dívek.ACT koupilo dětem hodně čokolády.PAT (=Five girls bought.GEN the children a lot of chocolate.GEN)

The valency frame of the verb *koupit* (=buy): ACT(.1) PAT(.4) ?ADDR(.3,pro+4) ?ORIG(od+2)

### 6.2.4.2. Valency frames of idiomatic expressions and complex predicates

Idiomatic expressions (see Section 7.1.2, “Idioms”) and complex predicates (see Section 7.1.1.4, “Complex predicates”) represent more complex cases; their dependent part is included in the valency frames as one of the valency modifications (functor = CPHR or DPHR).

**Valency frames of idiomatic expressions.** When specifying the surface form of the dependent part of an idiomatic expression, it is necessary to capture the following facts: how many parts (words) the dependent part has, what are their morphological categories and often also the precise lexical content of these parts. There is a convention adopted for encoding these requirements.

Examples of valency frames for idioms:

The valency frame for the verbal idiom: *lapat po dechu* (=gasp for breath):

ACT(.1) DPHR(po-1[dech.S6])

*lapat po dechu*

The valency frame for the verbal idiom: *běhat mráz po zádech* (=approx.: give sb the creeps, the experiencer is in the dative, the source is a PP):

ACT(.3) DPHR(mráz.S1,po-1[záda:P6])

*mráz mi běhal po zádech* (=it was giving me the creeps)

**Valency frames of complex predicates.** For the establishment of valency frames for the verbal component of a complex predicate two basic rules are to be followed:

- the nominal component of the complex predicate (with the functor CPHR) is recorded as a member of the valency frame, as is its valency modification.

All complex predicates that have the same verb in their verbal part and the nominal part of which may be formed by various synonyms and antonyms are assigned the same valency frame. In the surface-form specification of the nominal part of a complex predicate (with the CPHR functor), first, the set of synonyms and antonyms is given in curly brackets and only after this enumeration the representation of the forms follows. The list of the synonyms and antonyms (their lemmas) ends with three dots, which indicates that the list is not exhaustive; it only comprises the cases collected so far.

- in the valency frame of the complex predicate there is no argument shifting.

The valency frame of a verb which is part of a complex predicate is always considered in relation to the valency frame for the unmarked use of the verb. A new implementation of the shifting principle (actually, its doubling) would blur the relationships between equivalent valency positions in the two valency frames: With complex predicates, one valency position (Actor or Patient) of the unmarked valency frame becomes the nominal part of the predicate and it is assigned the functor CPHR. Other valency positions are, in the majority of cases, taken from the valency frame for the unmarked (semantically non-empty) use without change. They may, however, undergo certain modifications: when part of a complex predicate, the verb may acquire another valency modification or lose one of those it has in the unmarked use.

Examples:

Vedoucí.ACT dal podřízenému.ADDR výplatu.PAT (=The manager gave his subordinate his salary.)

The valency frame for one of the meanings of the predicate dát (=to give): ACT(.1) PAT(.4) ADDR(.3).

Vedoucí.ACT dal podřízenému.ADDR příkaz.CPHR přijít. (=The manager gave his subordinate the order to come.)

The valency frame for the complex predicate dát příkaz (=to give an order): ACT(.1) CPHR({pověření, podpora, souhlas, zpráva, impuls, odpověď, možnost, příkaz, naděje, popud, příčina, právo, příležitost, signál, šance,...}.4) ADDR(.3).

Udělal tuto část.PAT diplomové práce. (=He did this part of the thesis)

The valency frame of the predicate in the unmarked use udělat (=do): ACT(.1) PAT(.4).

Udělal na mě.ADDR dojem.CPHR (=He made an impression on me)

The valency frame of the complex predicate udělat dojem (=to make an impression): ACT(.1) CPHR({dojem,...}.4) ADDR(na+4).

No special valency frame is assumed for the noun that is part of a complex predicate. The nominal part carries the meaning of the complex predicate; the noun enters the complex predicate with its “full” meaning (unlike the verb, which becomes semantically empty), and thus it also has an “unimpoverished” valency frame.

### 6.2.4.3. Valency lexicon

The valency lexicon contains valency frames of semantic verbs, nouns, adjectives and adverbs. Individual valency frames are clustered on the basis of what t-lemma they are related to (for the discussion of t-lemmas see Chapter 4, *Tectogrammatical lemma (t-lemma)*).

The valency lexicon was being constituted during the annotation; therefore, only those verbs, nouns, adjectives and adverbs - i.e. those of their meanings - are included which occurred in the analyzed data. For example, if a verb has two different valency frames in the lexicon, it means that these two meanings of the verb were found in the analyzed data; however, the given verb may have other meanings (i.e. other valency frames), too.

The current version of the valency lexicon contains:

- valency frames of all semantic verbs (and verbal idioms) found in the analyzed data.
- valency frames of those semantic nouns which constitute the nominal part of complex predicates (i.e. those with the CPHR functor), found in the analyzed data.
- valency frames of those semantic nouns, adjectives and adverbs that have at least one argument as their daughter node, i.e. a node with one of the following functors: ACT, PAT, ADDR, EFF, ORIG.
- valency frames for non-verbal idioms if the governing node is either a semantic adverb or a semantic noun.
- valency frames for non-verbal idioms if the governing node is a semantic verbal noun (a noun ending with *-ní* or *-ti*). Other nouns that function as the governing nodes of idiomatic expressions are included in the valency lexicon only selectively.

**NB!** Only complex nodes' t-lemmas are included in the valency lexicon (`nodetype = complex`). T-lemmas of traditional verbs, nouns, adjectives and adverbs the `nodetype` attribute of which has a value other than `complex`

(according to the rules in Chapter 3, *Node types*) are not included in the valency lexicon (even if they have arguments).

## 6.2.5. Representing valency in the tectogrammatical trees

The valency of a node is represented in a tectogrammatical tree in the following way:

- by **assigning the node an adequate valency frame** (from the valency lexicon). The `val_frame.rf` attribute (see Section A.2.22, “`val_frame.rf`”) contains the identifier assigned to the valency frame corresponding to the given meaning of the given word.
- by **filling in the valency frame** in the tectogrammatical tree. Filling in the valency frame in a tectogrammatical tree means assigning functors to the dependent valency modifications (according to the assigned valency frame) and generating new nodes for those obligatory modifications that are not present at the surface level of the sentence. The rules for adding new nodes (for obligatory modifications) into a tectogrammatical tree are described mainly in Section 6.6.1.2, “Ellipsis of a dependent meaning unit”.

**The present state of valency representation in PDT.** Not every case of valency requirements is represented properly in the tree structures. Valency is represented properly with the following groups of complex nodes (checked nodes):

all semantic verbs,

all semantic verbal nouns (ending with *-ni* and *-ti*) that are included in the valency lexicon.

all semantic nouns that represent the nominal part of a complex predicate (i.e. for all semantic nouns with the CPHR functor),

all semantic adverbs included in the valency lexicon.

For the checked nodes it can be guaranteed that:

- they are assigned the appropriate valency frame. The value in the `val_frame.rf` attribute is valid (which might not be the case with unchecked nodes).
- their valency frames are filled in, i.e. the dependent modifications are assigned the appropriate functors and new nodes are generated if necessary (i.e. nodes for non-expressed obligatory modifications).

## 6.3. Clauses (governing, dependent, verbal, non-verbal)

Sentences (represented by tectogrammatical trees) are formed by one or more clauses. Annotation of clauses differs depending on whether the clause is verbal or non-verbal (see Section 6.3.1, “Verbal and non-verbal clauses”), and on the kind of dependency (see Section 6.3.2, “Dependent and independent clauses (clause connecting)”).

### 6.3.1. Verbal and non-verbal clauses

The following types of clauses are distinguished, on the basis of their governing node (effective root):

verbal clauses (see Section 6.3.1.1, “Verbal clauses”),

non-verbal clauses (see Section 6.3.1.2, “Non-verbal clauses”).

#### 6.3.1.1. Verbal clauses

**Verbal clause** is such a clause the governing node of which is a finite verb form or other forms with the function of a verbal predicate. There are both dependent and independent verbal clauses.

The effective root node of an independent verbal clause has the `PRED` functor. If an independent verbal clause is a parenthesis (see Section 6.5, “Parenthesis”), its effective root has the `PAR` functor (see Section 8.1, “Functors for effective roots of independent clauses”). Effective roots of dependent verbal clauses are assigned functors on the basis of their relation to their governing node.

The governing node (predicate) of a verbal clause can be:

- **finite verb form.**

*Otec spí.PRED (=Father is sleeping.)*

- **non-finite verb form.**

*Sparta poražena.PRED ! (=Sparta defeated!)*

*Nevíme, kam jít.PAT (=We don't know where to go.)*

- **contextual ellipsis of a predicate** (see Section 6.6.1.1, “Ellipsis of a governing meaning unit”).

*Otec spí, matka taky. = Otec spí, matka taky { spí.PRED} (=Father is sleeping, mother too.)*

- **grammatical ellipsis of a predicate.** Also those clauses are considered verbal in which there is no verb but which contain morphologically dependent forms of words. These are called constructions with an empty verb. Their governing node is a newly established node for an empty verb, i.e. a node with the t-lemma substitute #EmpVerb (see Section 6.6.1.1, “Ellipsis of a governing meaning unit”).

{#EmpVerb.PRED} *Vodu!* (=Water!)

{#EmpVerb.PRED} *V Praze, v pět hodin.* (=In Praha, at five o'clock.) Fig. 6.18

- **punctuation mark.** A punctuation mark is the governing node of a verbal clause in those cases in which it occupies the place of the missing verb and has its function. It is always possible to insert a simple verb (like *být*) into the clause and it is also possible to determine the function of the individual lexical units with respect to the missing verb.

*Doprava: vlastní. [#Colon.PRED] (=lit. Transport: own.)* Fig. 6.17

- **three dots.** Three dots are the governing element of a verbal clause if they signal that the clause is not finished and the predicate is not expressed.

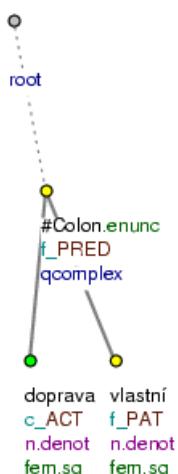
*Jenže... {#Period3.PRED} (=But...)*

- **interjection.** An interjection is the governing node of a verbal clause only if it plays the role of a verbal predicate (otherwise it is the governing node of an interjectional clause Section 6.3.1.2, “Non-verbal clauses”).

*Zajíc hop.PRED do jámy.* (=lit. Hare jump.interjection to hole.)

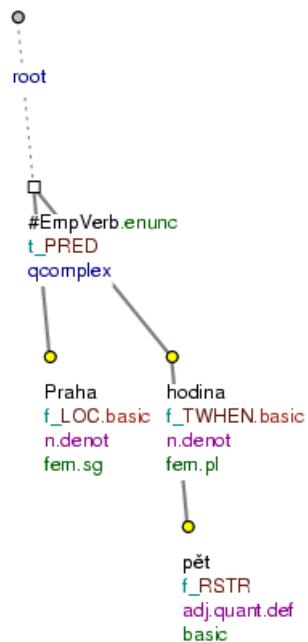
**Figure 6.17. Verbal clauses with a punctuation mark in the role of the predicate**

*Doprava: vlastní. (=lit. Transport: own.)*



**Figure 6.18. Verbal clauses with ellipsis of the predicate**

*V Praze, v pět hodin.* (=lit. In Praha, at five o'clock.)



### 6.3.1.2. Non-verbal clauses

**Non-verbal clause** is a clause whose governing node is not a verb. Non-verbal clauses are usually independent. They are dependent only in specific cases. Non-verbal clauses:

- a. **nominative clauses.** The governing node of a nominative clause is a noun in the nominative (and other forms with the same function).

If a nominative clause is independent, its effective root has the DENOM functor. If an independent nominative clause is a parenthesis (see Section 6.5, “Parenthesis”), its effective root has the PAR functor (see Section 8.1, “Functors for effective roots of independent clauses”). In those specific cases in which the nominative clause is dependent (direct speech, nominative of identity), its effective root has the functor according to its relation to the governing node.

Examples of nominative clauses:

*Důležitá událost.DENOM* (=An important event.) Fig. 6.19

*Vltavská.DENOM*

*1989.DENOM*

*10 let.DENOM* (=10 years)

*čtk.DENOM*

- b. **vocative clauses.** The governing node of a vocative clause is a noun in the vocative.

The effective root of a vocative clause has always the VOCAT functor (see Section 8.1, “Functors for effective roots of independent clauses”); the only exception are cases in which the vocative clause stands in the position of the nominative of identity (see Section 7.3, “Identifying expressions”).

Examples:

*Jirko.VOCAT* (= Jirka! VOC) Fig. 6.20

*Zeptali se: Občane.VOCAT , chceš dýchat čistý vzduch a mít také teplo?* (=They asked: Citizen, do you want to breathe fresh air..?)

nápis Občane.ID (=lit. *Inscription: Citizen.VOC*)

- c. **interjectional clauses.** The governing node of an interjectional clause is an interjection or a yes-no particle.

The effective root of an interjectional clause has always the VOCAT functor (see Section 8.1, “Functors for effective roots of independent clauses”); the only exception are cases in which the vocative clause stands in the position of the nominative of identity (see Section 7.3, “Identifying expressions”).

Examples:

Pardon.PARTL (=Pardon me.) Fig. 6.21

Ano.PARTL (=Yes)

nápis Aha.ID (=lit. *inscription: I\_see*)

**Figure 6.19. Nominative clauses**

Důležitá událost. (=lit. *Important event.*)



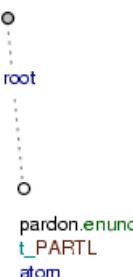
**Figure 6.20. Vocative clauses**

Jirko! (=lit. *Jirka! VOC*)



**Figure 6.21. Interjectional clauses**

Pardon.



## 6.3.2. Dependent and independent clauses (clause connecting)

Verbal and non-verbal clauses can be combined in two ways: either by a dependency relation or in a non-dependency connection. There are:

- **independent clauses**, i.e. clauses the governing node of which is not dependent on any other node (of any clause). Independent clauses are both verbal and non-verbal.
- **dependent clauses**, i.e. clauses the governing node of which is dependent on another node (of another clause). Dependent clauses are mostly verbal clauses. In specific case they can also be non-verbal.

**Non-dependency.** Several types of non-dependency connections of verbal and non-verbal clauses are distinguished:

- a. **paratactic connection.** The following combinations of clauses are considered paratactic (i.e. coordination or apposition; for the annotation rules see Section 6.4, “Parataxis”):

verbal clause + verbal clause	<i>Kočka je.PRED [is_member = 1] savec, ale savcem je.PRED [is_member = 1] i velryba.</i> (=The cat is a mammal but the whale is a mammal too.)
nominative clause + nominative clause	<i>Jan Novák.DENOM [is_member = 1], Brno.DENOM [is_member = 1]</i> Fig. 6.22

vocative clause + vocative clause	<i>Milý Jirko.VOCAT [is_member = 1], milý Petře.VOCAT [is_member = 1]! (=Dear Jirka, dear Petr!)</i>
interjectional clause + interjectional clause	<i>Cha.PARTL [is_member = 1], cha.PARTL [is_member = 1]</i>
verbal clause + nominative clause	<i>Recenze.DENOM [is_member = 1] knihy: Novou knihou jsou.PRED [is_member = 1] Rozbité obrazy. (=Review: Rozbité obrazy is a new book.) Fig. 6.23</i>
vocative clause + interjectional clause	<i>Ach.PARTL [is_member = 1], Jirko.VOCAT [is_member = 1]! (=Oh, Jirka!)</i>

- b. **specific non-dependency relations.** The combination of a verbal or nominative clause and an interjectional or vocative clause is not considered a paratactic connection but it is a specific non-dependency relation. The effective root of the interjectional or vocative clause is represented as dependent on the effective root of the verbal or nominative clause. The fact that this is not a dependency relation follows from the functor of the effective root of the interjectional or vocative clause, which is always PARTL or VOCAT. The following combinations are represented this way:

verbal clause + interjectional clause	<i>Ejhle.PARTL, to byla.PRED právě ta kapička. (=Oh, this was the droplet.) Fig. 6.24</i>
verbal clause + vocative clause	<i>Pane.VOCAT, nehodlám.PRED tu zůstat déle. (=Sir, I'm not staying any longer.) Fig. 6.25</i>
nominative clause + interjectional clause	<i>Ach.PARTL, ta prožluklá jména.DENOM (=Oh, those bloody names!)</i>
nominative clause + vocative clause	<i>Zavolal: Jirko.VOCAT, voda.PAT! (=He called: Jirka, water!)</i>

- c. **parenthesis.** Another case of non-dependency relations is the case of syntactically non-incorporated parenthesis. For the annotation rules see Section 6.5, “Parenthesis”.

**Dependency.** The basic type of combining two clauses in a dependency relation is:

- a. **a complex sentence.** Complex sentences are combinations of two or more verbal clauses in a dependency relation. This means:

governing verbal clause + dependent verbal clause	<i>Nevím.PRED, proč odešel.PAT (=I don't know why he left.)</i>
---	---

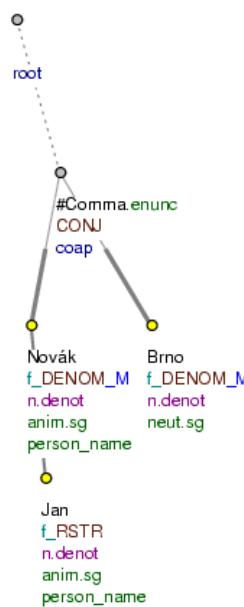
The effective root of the dependent verbal clause has the functor corresponding to the type of dependency. For more rules regarding dependent verbal clauses see Section 6.3.3, “Dependent verbal clauses (complex sentences)”.

Specific cases of dependency relations:

- **dependent direct speech.** All types of clauses can be used in the position of direct speech. For the annotation rules see Section 7.5, “Direct speech”.
- **nominative of identity.** All types of clauses can occur in the position of the nominative of identity. The effective root of any clause in the position of the nominative of identity has the ID functor. For detailed rules see Section 7.3, “Identifying expressions”.

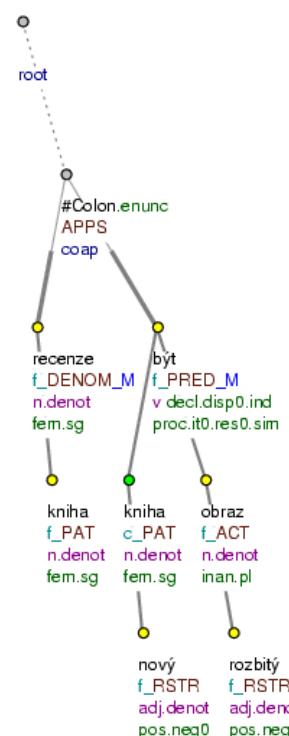
**Figure 6.22. Paratactic connection of two nominative clauses**

Jan Novák, Brno.



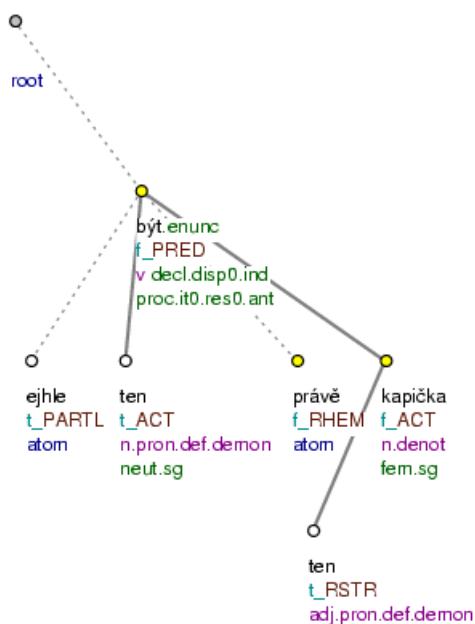
**Figure 6.23. Paratactic connection of a nominative clause and a verbal clause**

Recenze knihy: Novou knihou jsou Rozbité obrazy. (=lit. Review: New book is Broken pictures.)



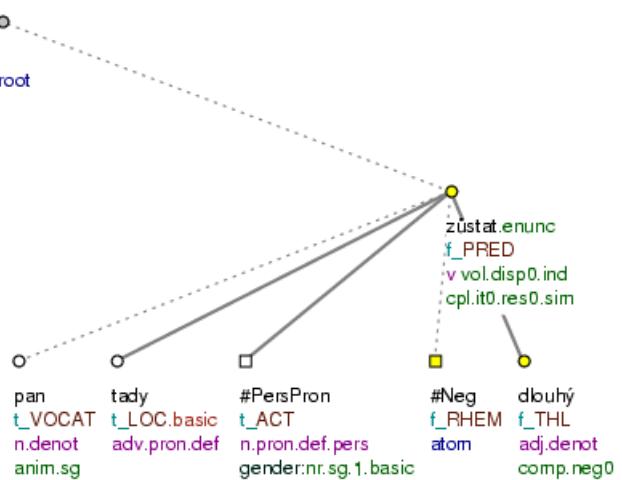
**Figure 6.24. Connection of an interjectional clause and a verbal clause**

Ejhle, to byla právě ta kapička. (=lit. Lo\_and\_behold, this was just the droplet.)



**Figure 6.25. Connection of a vocative clause and a verbal clause**

Pane, nehozlám tu zůstat déle. (=lit. Sir, I\_am\_not\_willing here to\_stay any\_longer.)



**Table 6.2. Types of dependent verbal clauses**

Dependent clause	Definition	Connective	Examples
<b>Content clause</b>	It stands for an argument of a word (verb, event noun..) in the governing clause. The effective root has an argument functor.	subordinating conjunction relative expression Relative elements introducing content clauses have no coreferred elements.	<i>Řekl, &lt;že&gt; přijde.EFF</i> (=He said he would come.) <i>Zeptal se, kdopřijde.PAT</i> (=He asked who was coming.)
<b>Relative clause</b>	It further specifies, modifies a noun phrase in the governing clause The effective root of the dependent clause has the RSTR functor.	relative expression connective <i>co</i> Relative elements introducing relative clauses corefer with the modified noun (see Section 9.2, “Grammatical coreference”).	<i>Otázka, která nebyla zodpovězena.RSTR, si žádá odpověď.</i> (=The question that was not answered...) <i>Ten kluk, &lt;co&gt; ho.PAT</i> <i>Jirka potkal.RSTR, bydlí v naší ulici.</i> (=The boy Jirka met lives in our street.)
<b>Adverbial clause</b>	It is a temporal, locative/directional, manner or other modification of an element in the governing clause. The effective root of a dependent clause has an adjunct functor.	subordinating conjunction relative expression	<i>&lt;Když&gt; bude.COND hezky, půjdeme ven.</i> (=If it is nice we'll go out.) <i>Šel, kam ho nohy nesly.DIR3</i> (=He went where his feet took him.)

### 6.3.3. Dependent verbal clauses (complex sentences)

In a complex sentence, the following is distinguished:

- **governing clause**, i.e. the clause (a part of) which is modified by another clause.
- **dependent clause**, i.e. a clause that modifies another clause or its part.

The effective root of a dependent clause always depends on the effective root of the modified element. If a dependent clause modifies the content of the whole governing clause, its effective root node depends on the effective root of the governing clause. Three types of dependent verbal clauses are to be distinguished in the annotation (see Table 6.2, “Types of dependent verbal clauses”).

#### 6.3.3.1. Dependent verbal clauses without a finite verb form

Dependent verbal clauses with no finite verb form include:

dependent infinitival constructions,  
dependent participial constructions,  
transgressival constructions.

If there is a non-finite verb form in a dependent clause (instead of the finite verb form; i.e. the infinitive, participle, transgressive) this non-finite form is the effective root of the dependent clause. A dependent clause without a finite verb form can be:

- a **argument**. It is the predicative-complement-like position (i.e. there is dual dependency involved, see Section 6.1.1, “Dual dependency”).
- a **predicative complement** (for the annotation rules see Section 6.1.1.1, “Predicative complement”).
- a **adverbial clause**. In some exceptional cases, a dependent verbal clause without a finite verb form can also have an adverbial meaning, especially when introduced by a subordinating conjunction.

Transgressival constructions are always analyzed as predicative complements.

Examples:

*Profesor, inspirován.COMPL článkem, přednášel o nových problémech.* (=The professor, inspired by the article, held a lecture on the new issues.) Fig. 6.28

*Zůstává inspirován.PAT článkem.* (=He stays inspired by the article.)

*Dům, ač zadalžen.CNCS, byl prodán velmi rychle.* (=The house, although indebted, was sold very quickly.)

*Máš dvě možnosti, jak získat.PAT peníze.* (=You have got two possibilities how to get money.) Fig. 6.26

*Nebýt.COND vás, nebyl bych tady.* (=If it were not for you I wouldn't be here.)

*Odešel, maje.COMPL vztek na celý svět.* (=He left, being mad at the whole world.)

**Frozen infinitival and transgressival constructions.** The verb form in frozen infinitival and transgressival constructions (the subject of the transgressive does not agree with the subject of the governing clause) is considered an adverb *sempos* = *adv*, which has kept part of its verbal valency. Frozen verbal constructions are often interpreted as non-verbal idioms (see Section 7.1.2, “Idioms”). The t-lemma of a node representing a frozen verb form is the actual frozen form. The node for the frozen verb form is assigned a functor according to its position in the sentence. For example:

*Soudě.COND podle ministra zahraničí, je to špatný výkon.* (=Judging with the Minister, it is a bad performance.) Fig. 6.29

*Přijdu,co.DPHR nevidět.TWHEN* (=I'll come very soon, lit. what not\_to\_see.)

**NB!** Certain transgressival constructions have petrified to such an extent that they are considered secondary prepositions, e.g.:

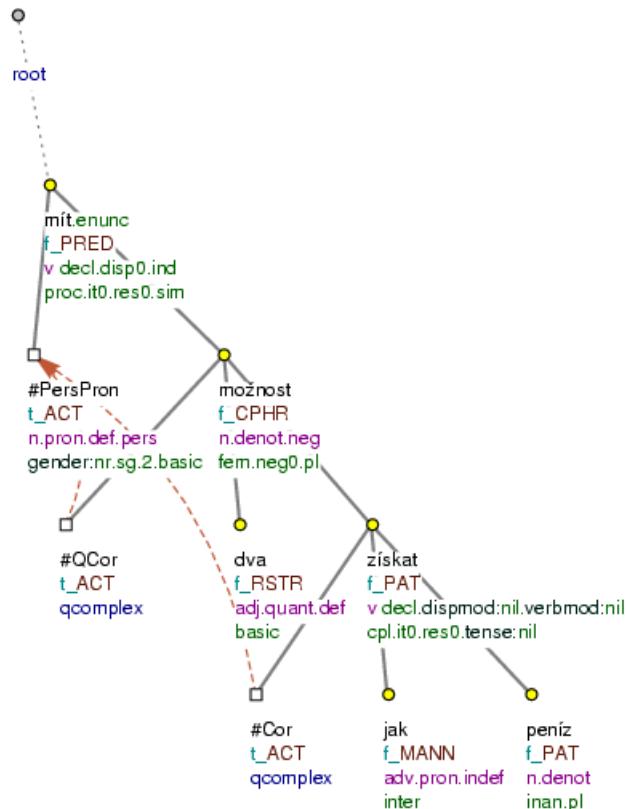
*Pozvali všechny příbuzné <vyjma> jeho bratra.RESTR* (=They invited all the relatives except for his brother.)

**Non-agreeing participial constructions.** Non-agreeing participial constructions are analyzed as conditional clauses (for syntactically incorporated parentheses, see Section 6.5, “Parenthesis”). For example:

*Upřímně řečeno.COND, vybrala si špatného partnera. [is\_parenthesis = 1]* (=Frankly speaking, she has chosen a wrong partner.) Fig. 6.27

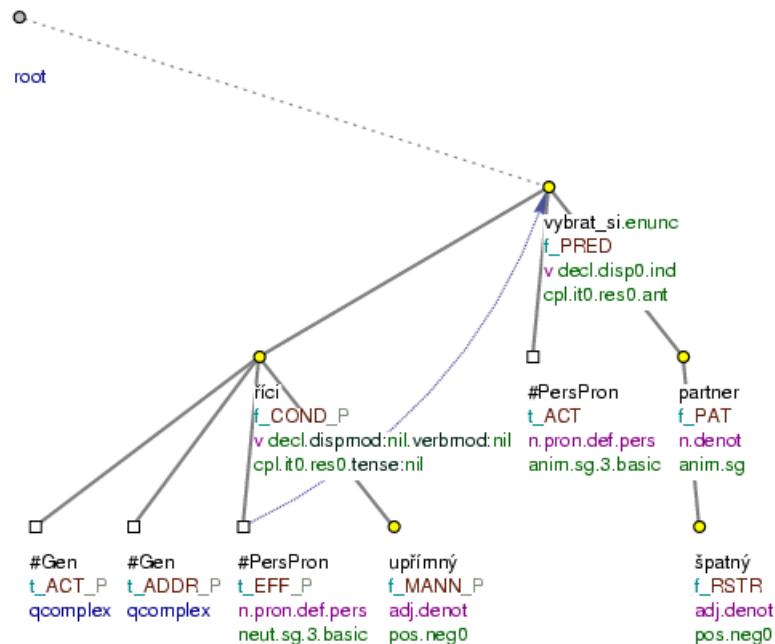
**Figure 6.26. Dependent infinitival constructions**

*Máš dvě možnosti, jak získat peníze.* (=lit. (You) have two possibilities how to\_get money.)

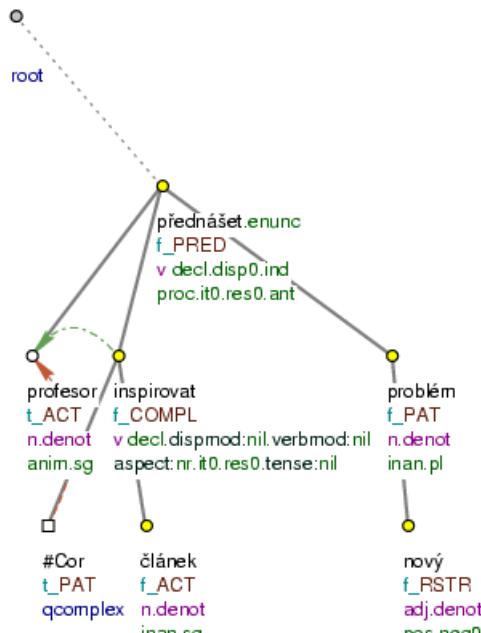


**Figure 6.27. Non-agreeing participial constructions**

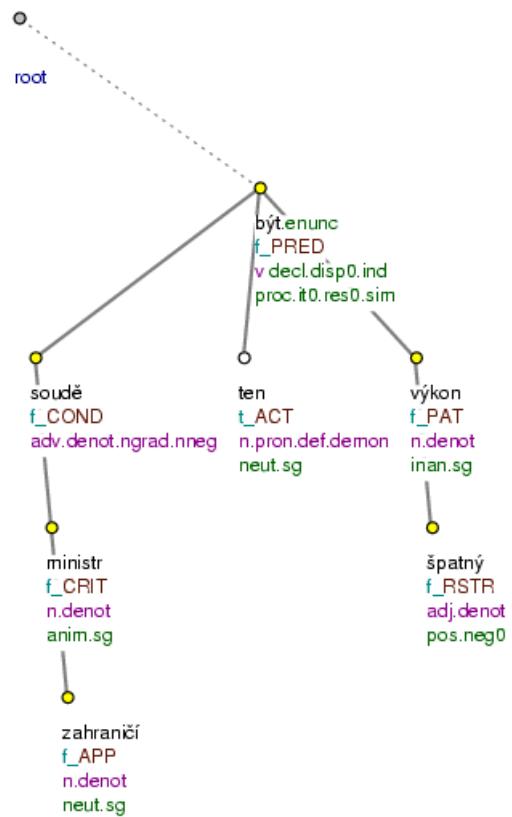
Upřímně řečeno, vybrala si špatného partnera. (=Frankly said, (she) has\_chosen REFL wrong partner.)


**Figure 6.28. Dependent participial constructions**

Profesor, inspirován článkem, přednášel o nových problémech. (=lit. Professor, inspired by\_the\_article, lectured on new issues.)


**Figure 6.29. Frozen transgressival constructions**

Soudě podle ministra zahraničí, je to špatný výkon. (=Judging according\_to Minister (of) Foreign\_affairs, is it bad performance.)



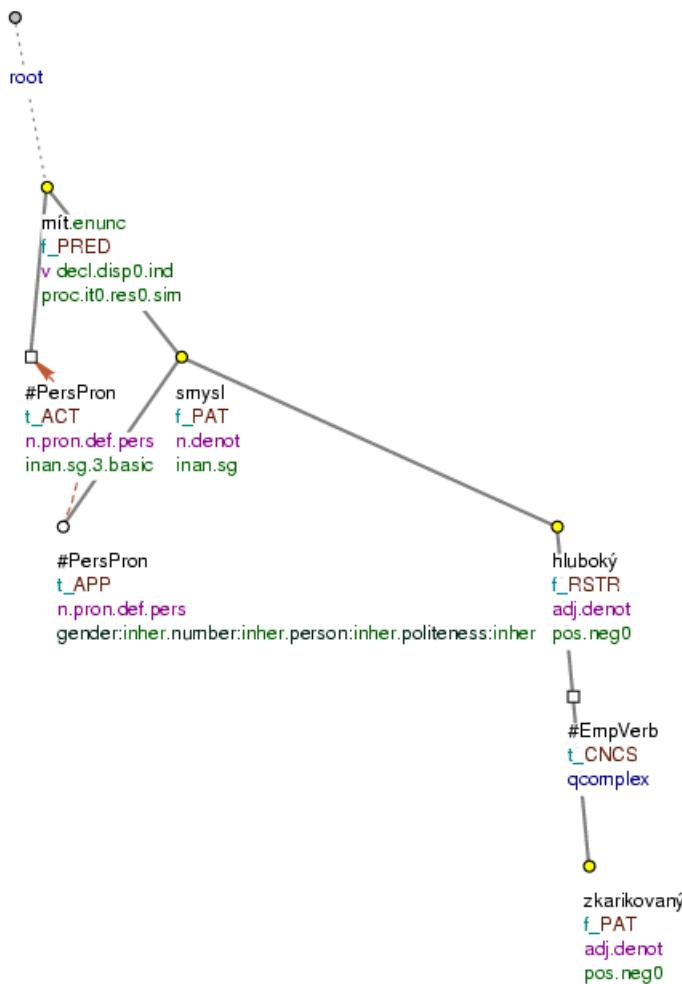
**Constructions with adjectives introduced by subordinating conjunctions.** If an adjective modifying some modification is introduced by a subordinating conjunction, this construction is analyzed as a dependent verbal clause in which the predicate is omitted. A new node for the missing predicate is added to the tree (#EmpVerb) with the functor corresponding to the meaning of the conjunction. The node for the adjective depends on the node for the empty verb as its Patient. The dependent clause modifies either another adjective or a(n entire) noun phrase. For example:

Má svůj hluboký, přestože {#EmpVerb.CNCS} zkarikovaný.PAT smysl. (=It has its deep, though twisted sense.)  
Fig. 6.30

Měřit něco platným, byť {#EmpVerb.CNCS} spletěným.PAT zákonem. (=To judge something with respect to a valid, though confusing law.)

Figure 6.30. Adjective introduced by a subordinating conjunction

Má svůj hluboký, přestože zkarikovaný smysl. (=lit. (It) has its deep, though twisted sense.)



### 6.3.3.2. Supporting elements

**Supporting elements** are pronominal expressions (pronouns, pronominal adverbs) in the governing clause that refer to the dependent clause and signal its function (by overt morphology).

Supporting elements do not have a node of their own. They are referred to in the `a/aux.rf` attribute of the effective root of the dependent clause. The effective root of the dependent clause is assigned a functor according to the meaning of the supporting element and the connective and depends on the effective root node of the governing clause.

**Correlative pairs.** Supporting elements form correlative pairs with the connectives in the dependent clause. These correlative pairs are divided into two groups according to whether the connective in the dependent clause is a subordinating conjunction or a relative expression (a pronoun or an adverb):

- A. **supporting element + conjunction.** When a supporting element and a conjunction constitute a correlative pair, the entire correlative pair is hidden in the tectogrammatical tree. the `a/aux.rf` attribute of the effective root node of the dependent clause contains a reference both to the supporting element and the conjunction.
- B. **supporting element + relative expression.** If the connective is a relative expression, only the supporting element is hidden in the tree: the `a/aux.rf` attribute of the effective root node of the dependent clause contains a reference to it. Relative expressions are always represented by a separate node.

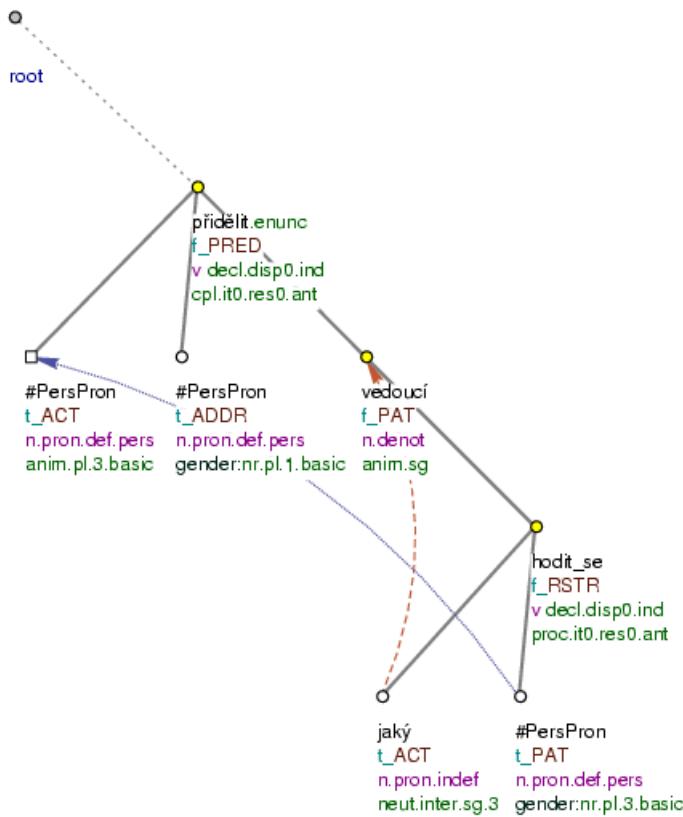
The basic types of correlative pairs are described in Table 6.3, “**Types of supporting elements**”.

**Table 6.3. Types of supporting elements**

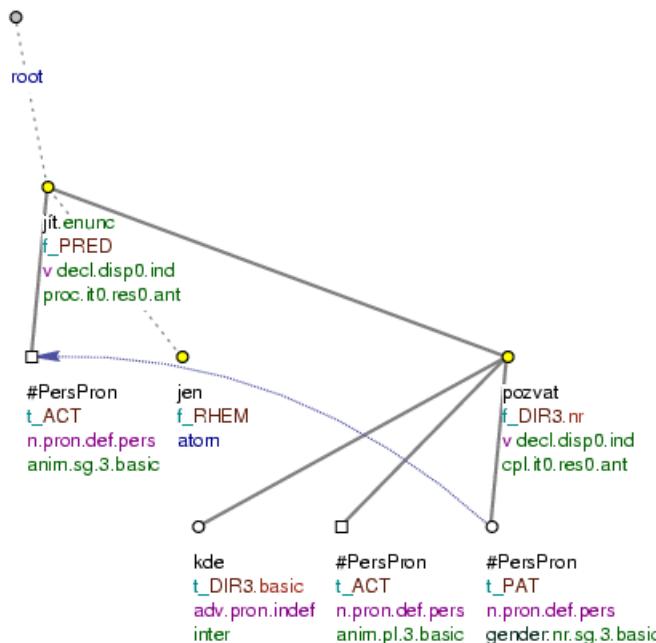
Correlative pair	Definition	Examples
<b>pronoun “ten”</b> + connective	The pronoun <i>ten</i> is considered a supporting element only when preceding a content or adverbial clause (not a relative clause).	<i>Neuměla vysvětlit &lt;to&gt;, co.PATudělala.PAT</i> (=lit. (She) couldn't explain that what (she) did.) Fig. 6.33  <i>Znepokojil se &lt;tím&gt;, &lt;že&gt; nepřišla.MEANS</i> (=He was worried by the fact (lit. that) that she hadn't come.)  <i>Čím.DIFFje.DIFF víno starší, &lt;tím&gt; je lepší.</i> (=The older the wine, the better it is; lit. By what is wine older by that is better.) Fig. 6.34
<b>pronoun “takový”</b> + connective	The pronoun <i>takový</i> can be a supporting element when preceding a predicate clause or a relative clause, which are introduced by the relative pronoun <i>jaky</i> .	<i>Přidělili nám vedoucího &lt;takového&gt;, jaký.ACT se jim hodil.RSTR</i> (=They assigned us a boss that (lit. such which) was convenient for them.) Fig. 6.31
<b>pronominal adverb</b> + connective	Pronominal adverbs ( <i>tam, odtud, tudy, tehdy, proto</i> ) are considered supporting elements if their function in the governing clause is the same as the function of the relative adverb ( <i>kde, kam, odkud, kdy</i> ) in the dependent clause or if the dependent clause is introduced by a conjunction.	<i>Šel jen &lt;tam&gt;, kam.DIR3 ho pozvali.DIR3</i> (=lit. (He) went only there where (they) him invited.) Fig. 6.32  <i>Bydleli tam &lt;odtehdy&gt;, odkdy.TSIN jim to bylo dovoleno.TSIN</i> (=They lived there from the time (lit. from then from when) they were allowed to.)  <i>Udělal něco &lt;proto&gt;, &lt;aby&gt; přišla.AIM</i> (=He did something for her to come.)  <i>Dostal &lt;tolik&gt;, kolik.PAT chtěl.PAT</i> (=He got as much as he wanted.)

**Figure 6.31. Supporting element “takový”**

Přidělili nám vedoucího takového, jaký se jim hodil. (=lit. (They) assigned us boss such which REFL them suited.)

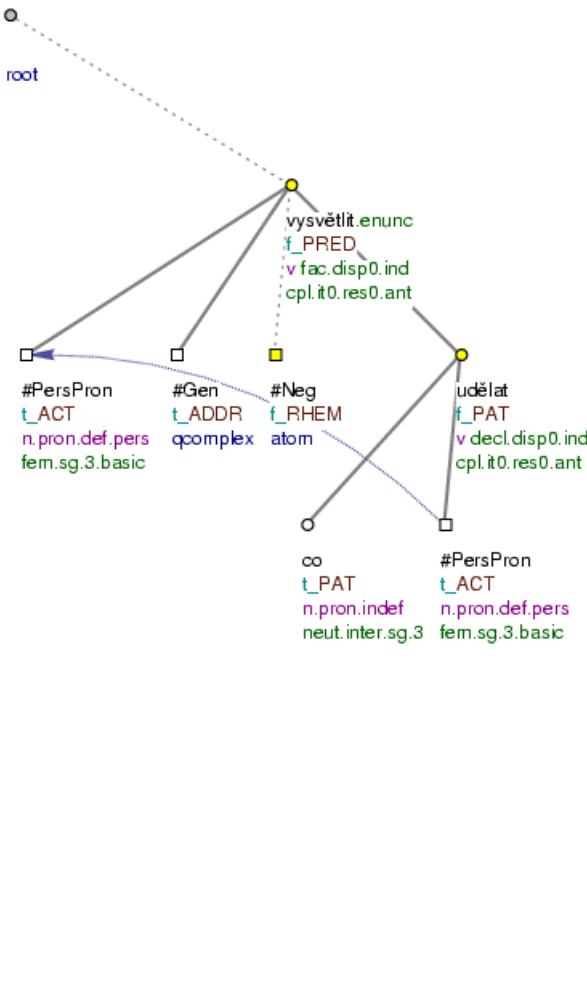

**Figure 6.32. Pronominal adverb as a supporting element**

Šel jen tam, kam ho pozvali. (=lit. (He) went only there where (they) him invited.)

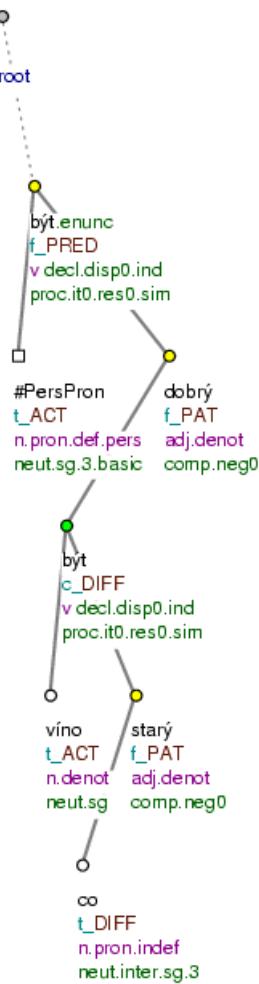


**Figure 6.33. Supporting element “ten”**

*Neuměla vysvětlit to, co udělala. (=lit. (She) couldn't explain that what (she) did.)*


**Figure 6.34. Supporting element “ten”**

*Čím je víno starší, tím je lepší. (=lit. By\_what is wine older by\_that is better.)*



### 6.3.3.3. False dependent clauses

**False dependent clauses** are such clauses that have the form of a dependent clause but their semantic relation to the other (governing) clause is rather that of coordination. False dependent clauses are either relative or conjunctival clauses. By the use of a subordinating conjunction the speaker introduces a new meaning (purpose, condition) into the sentence, which is in fact not present between the clause contents. When analyzing constructions with false dependent clauses, the form and not the content is the criterion. The effective root of the dependent clause is assigned a functor according to the meaning of the connective and depends on the effective root node of the governing clause.

Examples:

*Spadl pod vlak, který ho přejel.RSTR (=He fell under a train, which ran over him.)*

*Odešel, <aby> se už nevrátil.AIM (=He left not to come back again.)*

<Jestliže> Sparta v první třetině vyhrávala.COND, nakonec prohrála. (=Sparta was leading in the first third only to lose in the end.)

**NB!** If the connective could be considered a coordinating conjunction, the construction is analyzed as paratactic. The following constructions are considered paratactic:

- constructions with the connectives *což*, *přičemž*, *načež*, *pročež*, *začež*, *aniz*.** The connectives *přičemž*, *načež*, *pročež*, *začež*, *aniz* are taken to be coordinating conjunctions. The connective *což* (also in oblique cases: *bez*

*čehož, čemuž, za což*) is often a valency modification in the clause it introduces. The root of the paratactic structure, then, is the node for the present punctuation mark. The effective root of the attached clause has the same functor as the other paratactically connected element. For example:

*Nedohodli se, pročež.CSQ nastal nový boj. [nodetype=coap] (=They didn't find a solution, which is why new fights started.)* Fig. 6.36

*Při reklamaci došlo k chybě, za což.PAT se vám omlouváme. [#Comma.CONJ] (=There was an error in processing the complaint, for which we apologize.)* Fig. 6.35

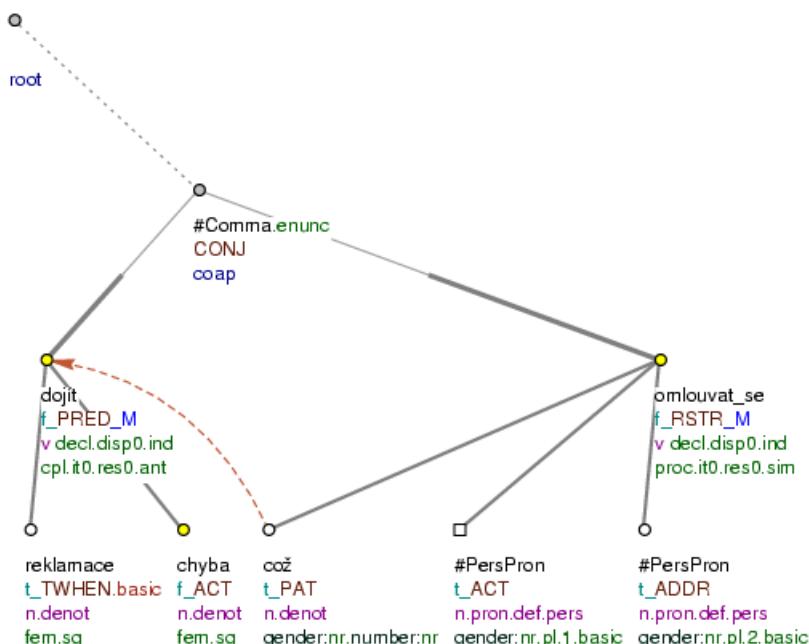
- b. **constructions with the connectives *kdežto* and *takže*.** The conjunctions *kdežto* and *takže* are considered co-ordinating conjunctions and the clauses they connect are paratactically connected. For example:

*Udělalo se hezky, takže.CSQ jsme mohli jít ven. [nodetype=coap] (=The weather turned better so we could go out.)* Fig. 6.37

*Svobodní mládenci mívají nepořádek kolem sebe, kdežto.CONFR ženatí mívají nepořádek v duši. [node-type=coap] (=Bachelors often have a mess all around them whereas married men have a mess in their souls.)*

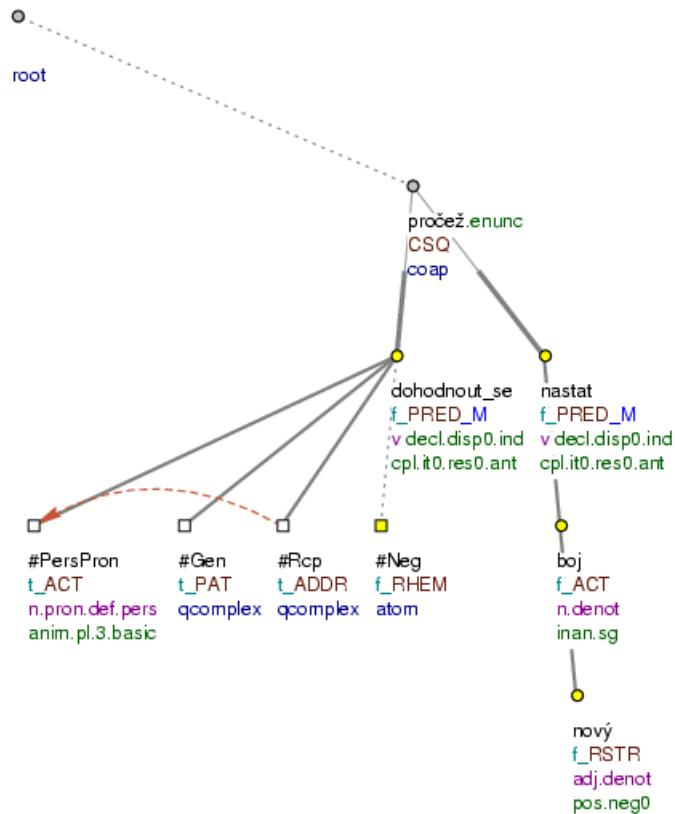
**Figure 6.35. Constructions with the connective “což”**

*Při reklamaci došlo k chybě, za což se vám omlouváme. (=lit. With complaint occurred - error; for what REFL (we) you apologize.)*

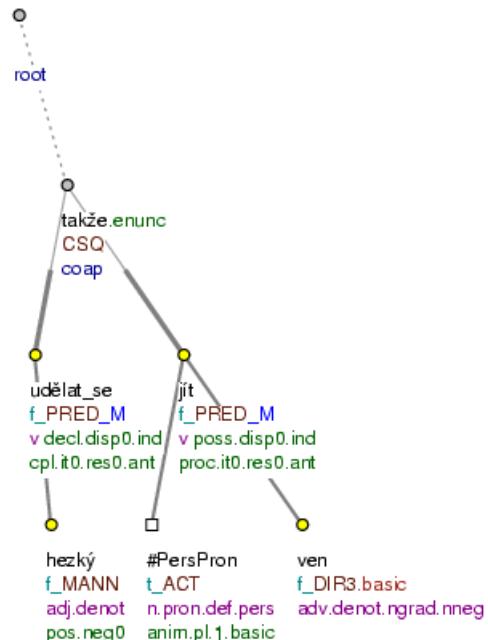


**Figure 6.36. Constructions with the connective “pročež”**

Nedohodli se, pročež nastal nový boj. (=lit. (They) did\_not\_find\_a\_solution REFL, which\_is\_why started new fight.)


**Figure 6.37. Constructions with the connective “takže”**

Udělalo se hezky, takže jsme mohli jít ven. (=lit. (It) became REFL nice, so\_that (we) AUX could go out.)



## 6.4. Parataxis

**Parataxis** is a non-dependency connection of two or more elements (modifications or clauses) that are on the same level and that depend on the same governing element (in the same way).

The tectogrammatical trees are two-dimensional and we do not introduce a third dimension for paratactic structures (which leads to the violation of the dependency principle; see also Section 6.1.2, “Non-dependency edges”).

**Representing parataxis in the tectogrammatical trees.** Paratactic connections are represented by a **paratactic structure** (see Fig. 6.38 and Fig. 6.39). **Paratactic structure root node** is a node for the coordinating connective or operator. In those rare cases in which there is no coordinating connective nor punctuation mark present in the surface structure, the root node of the paratactic structure is a newly established node with the t-lemma #Separ. Paratactic structure root nodes are assigned the value `coap` (see Chapter 3, *Node types*) in the `nodetype` attribute.

A paratactic structure root node is an immediate daughter of the governing node of the effective roots of the paratactically connected elements (i.e. terminal members of the paratactic structure).

The root nodes of the paratactically connected elements are immediate daughters of the paratactic structure root node and the value of their `is_member` attribute (see Section A.2.12, “`is_member`”) is 1. Paratactically connected elements are thus distinguished from nodes for shared modifiers. The root node of a shared modifier is also an immediate daughter of the paratactic structure root node but the value in its `is_member` attribute is not 1.

A paratactically connected element can also be represented by an **embedded paratactic structure**. Further, direct and terminal members of paratactic structures are distinguished. **Terminal members** of a paratactic structure are the effective root nodes of paratactically connected elements. **Direct members** of a paratactic structure are all immediate daughters of the paratactic structure root node whose value in the `is_member` attribute is 1. A direct member of a paratactic structure can also be a terminal member but a direct member can also be represented by the root node of an embedded paratactic structure; the root node of a paratactic structure is never a terminal member.

An immediate daughter of the root of a paratactic structure (`nodetype = coap`) can be:

- a. the effective root of a paratactically connected element (i.e. the terminal member of the paratactic structure), whose value of the `is_member` attribute is 1.
- b. the root of a(n embedded) paratactic structure, whose value in the `is_member` attribute is 1.
- c. the root of a shared modifier, whose value in the `is_member` attribute is 0.

**NB!** A shared modifier can also be instantiated by a paratactic structure. The root of the shared modifier is, then, a paratactic structure root node; its `is_member` attribute is however assigned the value 0.

- d. a node for a rhematizer of a shared modifier, i.e. a node with the RHEM functor. The value in its `is_member` attribute is then 0.
- e. a node for a conjunction modifier, i.e. a node with the CM functor. The value in its `is_member` attribute is then 0.

**Figure 6.38. Paratactic structure I**

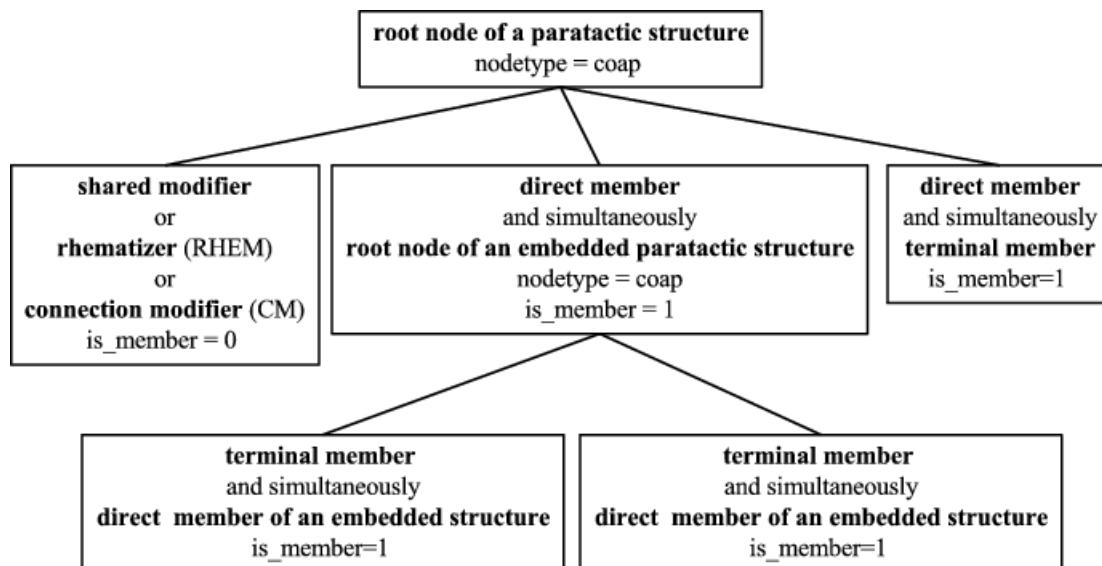
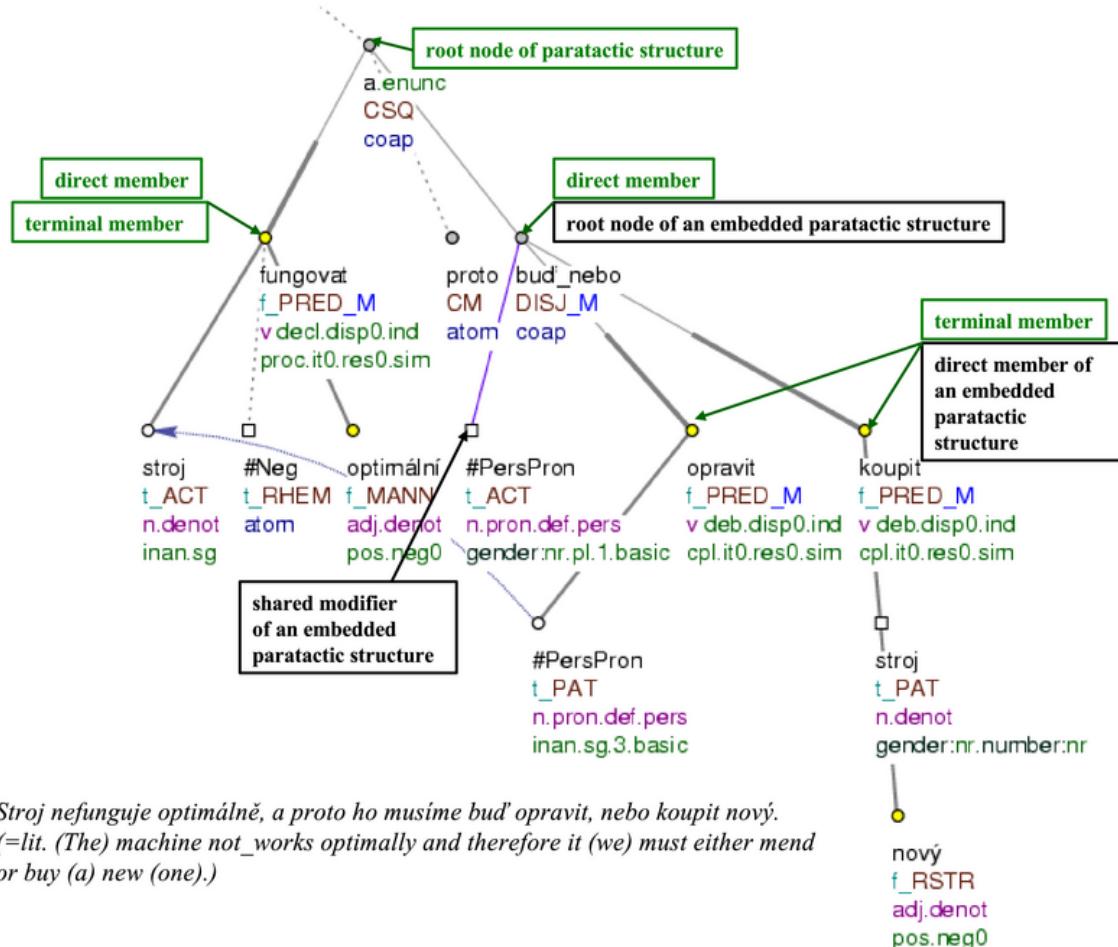


Figure 6.39. Paratactic structure II



**Shared modifiers.** A shared modifier is such a modification that relates to every paratactically connected terminal element and that is expressed only once at the surface level. Any kind of modification (i.e. both arguments and adjuncts) can be a shared modifier. Non-obligatory modifications are analyzed as shared modifiers only in unambiguous cases.

The root node of a shared modifier is represented as an immediate daughter of the root of that paratactic structure the terminal elements of which it modifies. It is distinguished from the paratactically connected elements by the value of the `is_member` attribute, which is 0. For example:

*Marii jsem viděl a slyšel zpívat.* (=I saw and heard Marie sing.) Fig. 6.40

*Petr celý den pracoval na své disertaci a připravoval se na zkoušku z angličtiny, ale večer už nedělal nic.* (=Petr worked the whole day on his dissertation.. but in the evening he did nothing.) Fig. 6.42

**NB!** If a potential shared modifier requires a different value in any attribute with respect to any of the terminal members (e.g. the functor or the value in the `tfa` attribute), it is not possible to represent the modification as a shared modifier but it has to be represented separately for every terminal member (with the help of newly established nodes). For example:

*Přišel Jirka a posadil se.* (=Jirka came and sat down.)

=*Přišel Jirka.ACT [tfa=f] a {#PersPron.ACT [tfa=t]}* posadil se.

**Principle of the simplest structure.** Generally, we represent paratactic structures as deep as possible in the tree structure and we make use of the possibility of shared modification. Therefore, it is usually not necessary to add new nodes into the tree for the elided modifications. The simplest possible structure is chosen, which means the parataxis of sentence parts is preferred over clausal parataxis.

Nevertheless, it is not always possible to represent the construction as parataxis of sentence parts. All cases which do not fulfil the conditions on the parataxis of sentence parts (agreement in form and function), are represented as clausal parataxis, i.e. new nodes for the governing predicates of the clauses are added to the tree (see Section 6.6.1.1, “Ellipsis of a governing meaning unit”). For example:

*Přišel Petr a asi i Pavel.* (=Petr came and Pavel probably as well.)

= *Přišel Petr a asi přišel i Pavel.* (=Petr came and Pavel probably came as well.) The Actors *Petr* and *Pavel* cannot be captured as being in constituent coordination; the expression *asi* modifies the absent predicate. The construction is therefore represented as clausal coordination. Fig. 6.41.

**Functors for terminal members of paratactic structures.** Paratactic structures are usually formed by elements with the same functor. The functors of the terminal members of paratactic structures can also differ, but it holds that:

- a. the functors of all operands for expressing mathematical operations and intervals are always identical.
- b. the functors of the terminal members in the case of clausal parataxis are always identical.
- c. in the case of parataxis of sentence parts, the terminal members can only have differing functors if it is co-ordination or apposition of non-obligatory adjuncts. For example:

*pracovní doba osmihodinová.RSTR [is\_member = 1] a.CONJ bez přestávky.ACMP [is\_member = 1]*  
*(=lit. eight-hour working hours and without break)*

*Udělali to s úžasem.ACMP [is\_member = 1] , tedy dobře.MANN [is\_member = 1] [#Comma.APPS]*  
*(=They did it with astonishment, that is well.)*

- d. if the paratactic connection is mixed, the non-clausal modification is assigned a functor depending on its relation to the governing node. The clausal (verbal) terminal member has the same functor as the non-clausal terminal member; e.g.:

*O zajímavých místech.PAT [is\_member = 1] , jako.APPS je.PAT [is\_member = 1] třeba Litomyšl,  
 Kutná Hora nebo Český Krumlov, zahraniční turisti většinou nevěděj.* (=lit. About interesting places, as is e.g.  
*L., K.H. or C.K....)*

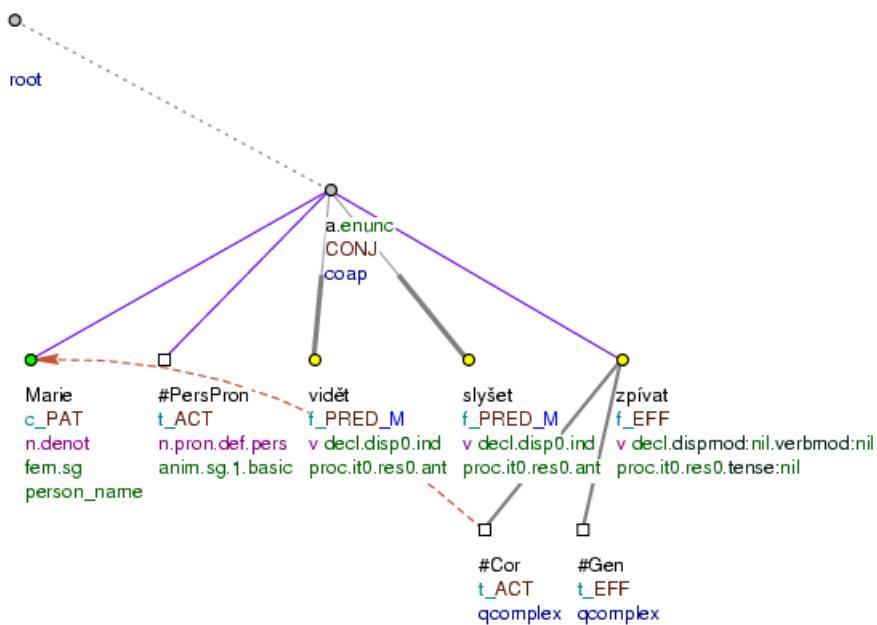
If the paratactic connection is a connection of a verbal clause and a nominative clause, the functor assignment follows the rules in Section 6.3.1, “Verbal and non-verbal clauses”. For example:

*Téma.DENOM [is\_member = 1] : Co právě dělám.PRED [is\_member = 1] [#Colon.APPS]* (=The topic: What I am doing at the moment.)

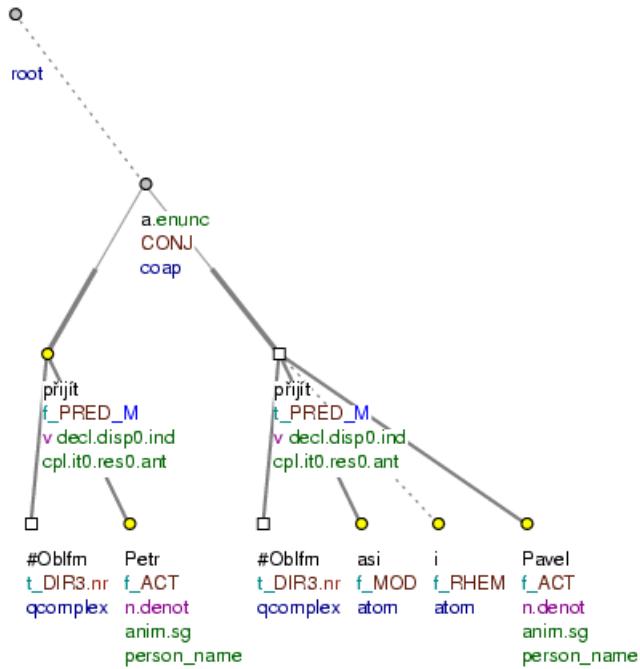
**Semantic types of paratactic connections.** The functor assigned to the root of a paratactic structure expresses the semantic relation between the connected elements. All functors (and their definitions) for paratactic structure root nodes are in Section 8.12, “Functors expressing the relations between members of paratactic structures”.

**Figure 6.40. Paratactic structure**

*Marii jsem viděl a slyšel zpívat. (=lit. Marie (I) AUX saw and heard sing.)*

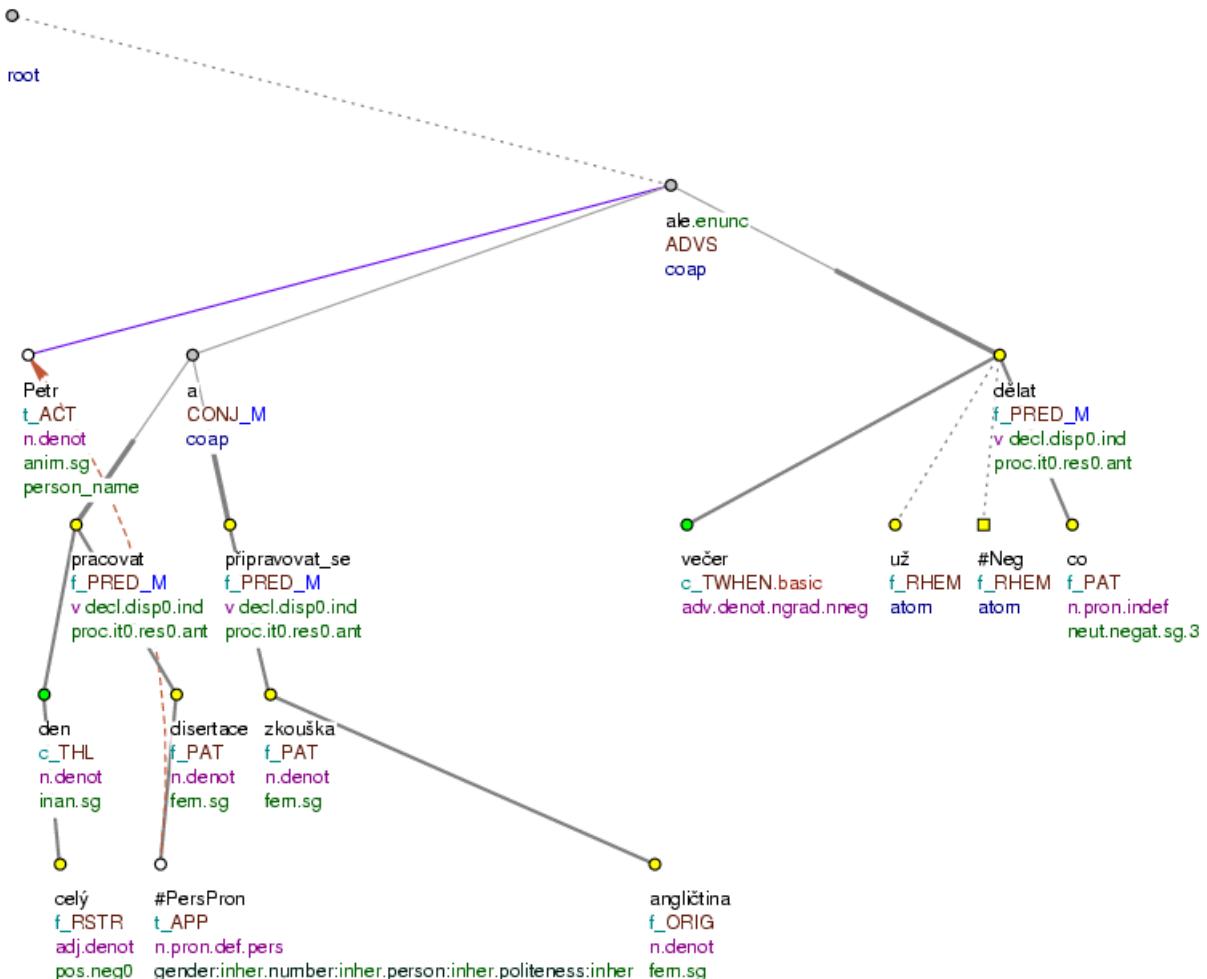

**Figure 6.41. Paratactic structure**

*Přišel Petr a asi i Pavel. (=lit. Came Peter and maybe also Paul.)*



**Figure 6.42. Paratactic structure**

*Petr celý den pracoval na své disertaci a připravoval se na zkoušku z angličtiny, ale večer už nedělal nic.* (=lit. Petr whole day worked on his dissertation and prepared himself for exam from English but in\_the\_evening (he) already not\_did nothing.)



## 6.4.1. Coordination, apposition, mathematical operations and intervals

The following cases are represented as paratactic structures:

coordination or apposition,  
mathematical operations and intervals.

**Coordination or apposition.** Only those combinations of two or more elements are considered coordination or apposition which are connected by a coordinating connective.

To be analyzed as apposition, the two elements have to be separated by a comma (e.g.: *český král, Karel* (=lit. Czech king, Karel); *hlavní město, Praha* (=lit. capital city, Praha)). If there is no comma (e.g.: *český král Karel* (=lit Czech king Karel); *hlavní město Praha* (=capital city Praha); *stalo se to v Praze na Vyšehradě* (=It happened in Praha at Vyšehrad); *v únoru v roce 1999* (=lit. in February i year 1999)) the connection is not analyzed as apposition (see especially Section 6.1.3, “Ambiguous dependency”).

Also some specific constructions are represented as coordination or apposition: constructions with the abbreviations *atd.*, *apod.*, *aj.* (=etc.); constructions in which a modification follows expressions like *a to, a since* (=i.e.). Also constructions with *což* (=which) (see also Section 6.3.3.3, “False dependent clauses”) and some other specific constructions like addresses etc. are analyzed as paratactic structures.

**Mathematical operations and intervals.** Constructions expressing mathematical operations and intervals are represented as paratactic structures even if the elements are connected by hypotactic means:

- A. **mathematical operations** (addition, subtraction, multiplication, division) are analyzed exclusively by means of the OPER functor. For example:

Zápas skončil 5:0. (=lit. match ended 5:0.) Fig. 6.43

Prodáváme byt 4+1. (=lit. We\_are\_selling flat 4+1.)

10 minus 2 je 8. (=Ten minus two is eight.)

- B. **intervals** (temporal, spatial and other) are represented in two ways:

- a. with the help of appropriate temporal and locative/directional functors. For example:

Sněžilo od Vánoc.TSIN až do Velikonoc.TTILL (=It snowed from Christmas to Easter.)

Stalo se to mezi pondělkem.TWHEN [subfunctor=betw] a středou.TWHEN [subfunctor = betw]  
 (=It happened between Monday and Wednesday.)

Znám to od Aše.DIR1 přes Prahu.DIR2 až po Brno.DIR3 (=I know it here from Aš to Praha and Brno.)

- b. those temporal and spatial intervals in which the interval meaning would get lost in the annotation by means of temporal or locative/directional functors, and all other intervals (that have no temporal or spatial meaning) are represented as a paratactic structure, with the functor OPER at the root of the structure. For example:

V období 1995 až 1999 jsem studoval na gymnáziu. (=In the years 1995 - 1999...) Fig. 6.44

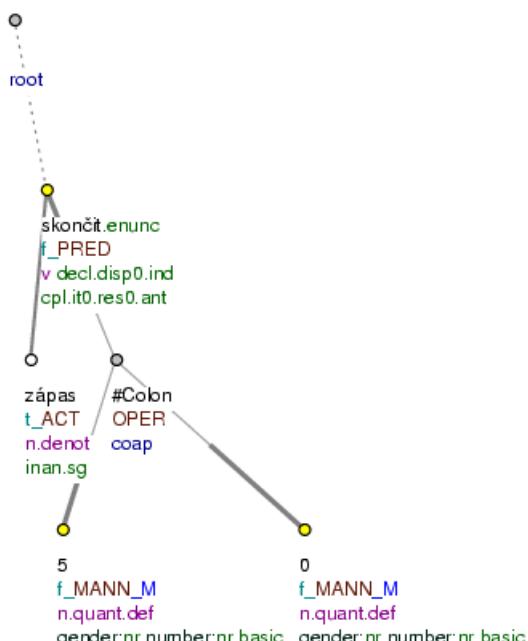
Sledovali to všichni, od dětí přes mládež až po dospělé. (=Everybody watched it, from children to young people and adults.) Fig. 6.45

trest od tří do pěti let (=three to five year sentence)

Na trase Praha - Brno došlo k nehodě. (=On the road Praha -Brno...)

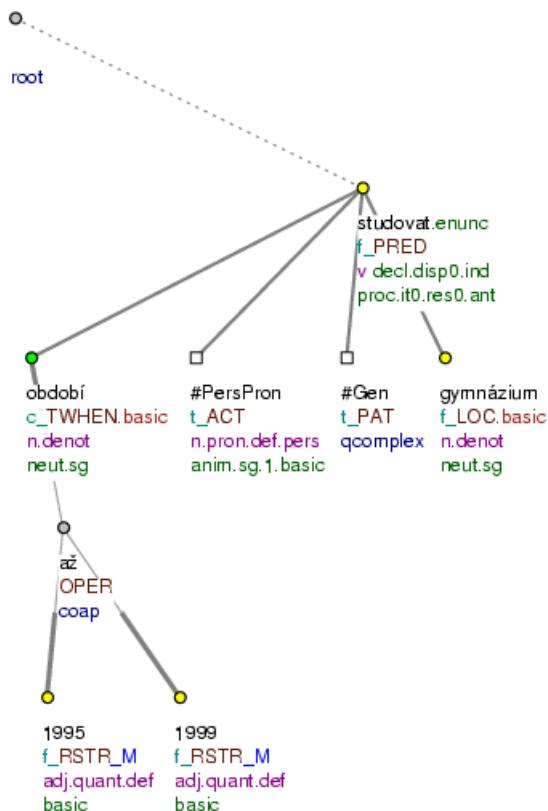
**Figure 6.43. Mathematical operations**

Zápas skončil 5 : 0. (=lit. Match ended 5:0.)

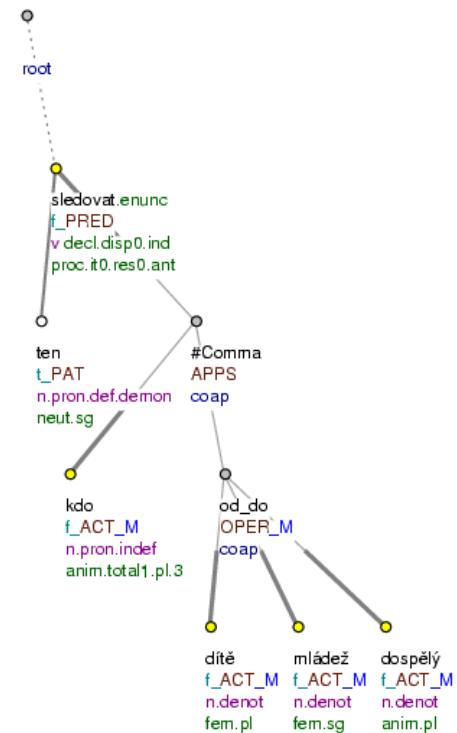


**Figure 6.44. Interval as a paratactic structure**

*V období 1995 až 1999 jsem studoval na gymnáziu.* (=lit. *Sledovali to všichni, od dětí přes mládež až po dospělé.*  
*In period 1995 to 1999 (I) AUX studied at secondary\_grammar\_school.)*


**Figure 6.45. Interval as a paratactic structure**

*V období 1995 až 1999 jsem studoval na gymnáziu.* (=lit. *Sledovali to všichni, od dětí přes mládež až po dospělé.*  
*(=lit. Watched it everybody, from children through youth up to adults.)*



## 6.5. Parenthesis

**Parenthesis** is defined as those parts of the text that do not belong to its basic level but that rather interrupt it by inserting additional information, an ex-post explanation, evaluating comments etc. Parenthesis is usually signalled by graphic means (dashes or brackets).

**Representing parenthesis in the tectogrammatical trees.** Parenthesis is represented with the help of the attribute `is_parenthesis` (see Section A.2.14, “`is_parenthesis`”). All nodes that are part of parenthesis have the value 1 in this attribute. The reason for this are cases of the so called “discontinuous” parenthesis, e.g.:

*Přišel tam Petr (a Pavel). (=Petr (and Pavel) came there.)* Fig. 6.46

The following two types are distinguished:

parenthesis proper,  
lexicalized parenthesis

**Parenthesis proper** is a parenthesis used uniquely in a given situation. There are the following subtypes:

- a. **syntactically incorporated parenthesis.** If a parenthesis is syntactically incorporated in the sentence (as one of its modifications) its effective root node is assigned a functor corresponding the type of dependency. For example:

*Podmětem (jestliže vyjadřuje.COND činnost), může být i infinitiv.* (=The subject can be - if it expresses an activity - also the infinitive.) Fig. 6.47

*Vidím náš dům (a naši zahradu.PAT).* (=I can see our house and our garden.)

- b. **syntactically non-incorporated parenthesis.** Syntactically non-incorporated parenthesis corresponds - both semantically and in its form - to some kind of independent clause (see Section 6.3, “Clauses (governing, dependent, verbal, non-verbal)”). The effective root of a parenthesis is assigned a functor on the basis of what kind of independent clause it is, see Table 6.4, “**Functors for syntactically non-incorporated parenthesis**”. The root of such a parenthesis is a direct daughter of the node it most closely relates to. For example:

*Mužstvo skončilo až třetí (loni bylo PAR první). (=The team was on the third place (last year it was the first place).)* Fig. 6.48

*Pavel Novák (Praha.PAR).*

*Zase nesehnal práci (ach.PARTL).* (=Oh, he didn't get a job again.)

**Table 6.4. Functors for syntactically non-incorporated parenthesis**

Type of independent clause	Functor of the root of the parenthesis
verbal clause	PAR
nominative clause	PAR
vocative clause	VOCAT
interjectional clause	PARTL

**Lexicalized parenthesis** is a parenthesis fixed to such an extent that it has become a particle. The lexicalized parenthesis is formed by a frozen finite verb form, which can keep part of the original valency.

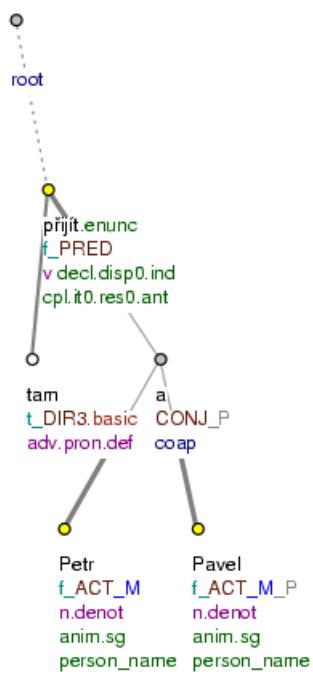
The effective root of a lexicalized parenthesis is assigned the ATT functor (node\_type = atom). A multi-word lexicalized parenthesis is usually considered a non-verbal idiom (see Section 7.1.2, “Idioms”). The effective root of a lexicalized parenthesis is an immediate daughter of the effective root of the clause in which the parenthesis is inserted. For example:

*Dnes je, myslím, středa.* (=It's Wednesday today, I think.) Fig. 6.49

*To se, nedej bůh, snad nestane.* (=This will, hopefully, not happen.)

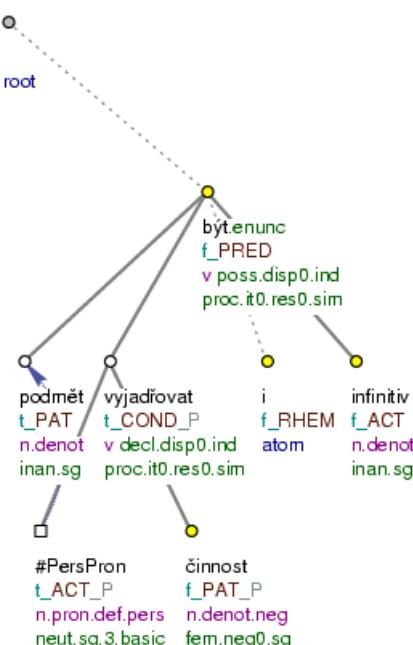
**Figure 6.46. Discontinuous parenthesis**

*Přišel tam Petr (a Pavel).* (=lit. Came there Petr (and Pavel).)



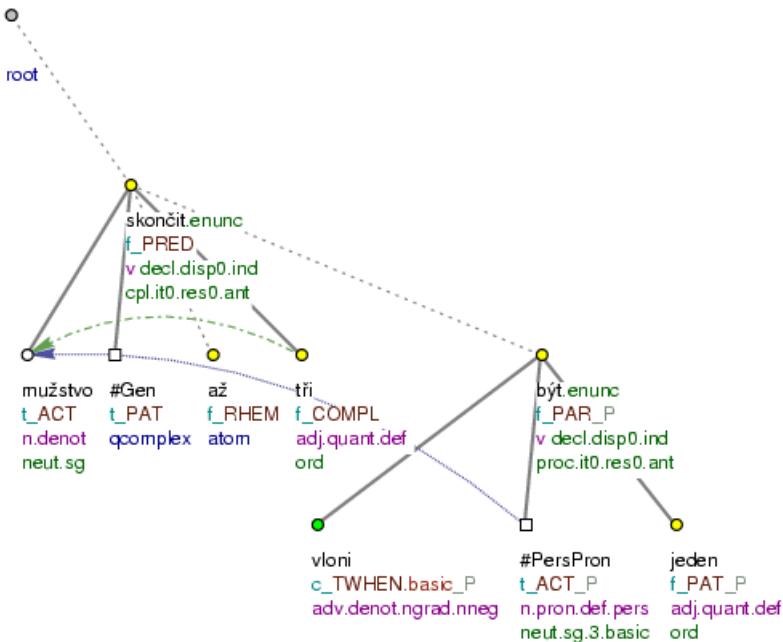
**Figure 6.47. Syntactically incorporated parenthesis proper**

*Podmětem (jestliže vyjadřuje činnost), může být i infinitiv.* (=lit. Subject.INSTR (if (if) expresses activity) can be also infinitive.)

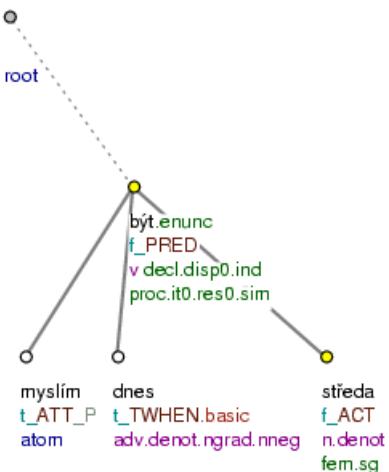


**Figure 6.48. Syntactically non-incorporated parenthesis proper**

*Mužstvo skončilo až třetí (loni bylo první). (=lit. Team ended only third (last\_year (it) was first).)*


**Figure 6.49. Lexicalized parenthesis**

*Dnes je, myslím, středa. (=lit. Today is, I\_think, Wednesday.)*



**Speaker's comments such as “aby bylo jasno”** Clauses introduced by the conjunction *aby* in constructions such as *Aby bylo jasno, já jsem tu pánem* (=Just to make things clear: I'm the boss here.) are in fact parentheses as they are comments of the speaker. It is assumed that such comments are fixed, lexicalized and therefore the dependent clause is represented as a syntactically non-incorporated parenthesis. For example:

*<Aby> bylo.PAR jasno, já jsem tu pánem. (=Just to make things clear: I'm the boss here.)*

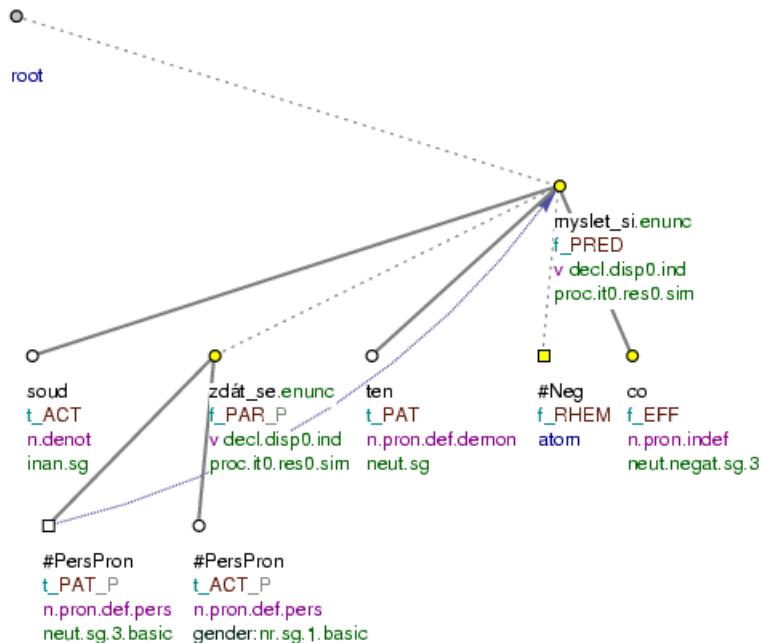
**Clauses with a reversed syntactic relation.** Construction in which the syntactic relation between the clauses is reversed is such a construction in which the inserted clause (*Soud, zdá se, nemyslí si o tom nic.* (=The court, it seems, has no opinion on the subject)), or a clause introduced by the connective *jak* (*Soud, jak se zdá, nemyslí si o tom nic.* (=The court, as it seems, has no opinion on the subject)) is in fact (originally) the governing clause. One argument of the verb in the original governing clause is usually missing as it is expressed by the original content clause.

The original governing clause is represented as a syntactically non-incorporated parenthesis. The missing argument of the verb in the original governing clause is represented by a newly established node with the t-lemma substitute #PersPron and the appropriate functor. There is a textual coreference relation between the newly established node and the effective root of the original content clause. The connective *jak*, if present, is assigned no node in the tree. For example:

*Soud, jak se mi zdá.PAR, nemyslí si o tom nic.* / *Soud, zdá se.PAR mi, nemyslí si o tom nic.* Fig. 6.50

**Figure 6.50. Clauses with a reversed syntactic relation**

*Soud, jak se mi zdá, nemyslí si o tom nic.* (=lit. Court, as REFL to \_me seems, does \_not\_think REFL about it nothing.)



## 6.6. Ellipsis

Newly established nodes are necessary for a complete representation of the meaning of the sentence.

Every newly established node is assigned the value 1 in the attribute `is_generated` (see Section A.2.11, “`is_generated`”).

### Types of newly established nodes:

- nodes for omitted meaning units (see Section 6.6.1, “Ellipsis”),
- auxiliary nodes for representing more complex syntactic structures (see Section 6.6.2, “Newly established nodes in more complex syntactic structures”),
- nodes representing negation with verbs (see Section 6.6.3, “Nodes representing negation with verbs”).

Newly established nodes are added to the structure in two ways. There are:

- copied nodes**, i.e. newly established nodes that have the values of certain attributes the same as some other node. A node is copied as a lexical unit, which is represented by its t-lemma, some grammatemes and a valency frame. The following attributes remain the same when a node is copied:

```

t_lemma
a/lex.rf
val_frame.rf
is_name_of_person
gram/gender
gram/aspect
gram/iterativeness
    
```

gram/negation  
gram/indeftype  
gram/numertype

A copied node does not have to be present in the same tree as the original one. Copied nodes are used for representing ellipsis of governing elements (see Section 6.6.1.1, “Ellipsis of a governing meaning unit”).

- B. **newly established node with a t-lemma substitute**, i.e. a newly established node with one of the following t-lemma substitutes:

<b>#EmpNoun</b>	<b>Grammatical ellipsis of governing meaning units</b> (see Section 6.6.1.1, “Ellipsis of a governing meaning unit”)
<b>#EmpVerb</b>	<b>Ellipsis of governing meaning units in specific constructions</b> (see Section 6.6.1.3, “Specific elliptical constructions”)
<b>#AsMuch</b> <b>#Equal</b> <b>#Total</b>	
<b>#Benef</b> <b>#Cor</b> <b>#Gen</b> <b>#Oblfm</b> <b>#PersPron</b> <b>#QCor</b> <b>#Rcp</b> <b>#Some</b> <b>#Unsp</b>	<b>Ellipsis of dependent meaning units</b> (see Section 6.6.1.2, “Ellipsis of a dependent meaning unit”)
<b>#Forn</b> <b>#Idph</b> <b>#Separ</b>	<b>Auxiliary nodes for representing more complex syntactic structures</b> (see Section 6.6.2, “Newly established nodes in more complex syntactic structures”)
<b>#Neg</b>	<b>Verbal negation by means of the morpheme <i>ne</i>-</b> (see Section 6.6.3, “Nodes representing negation with verbs”)

## 6.6.1. Ellipsis

**Ellipsis** is omission of a meaning unit at the surface level. Such a meaning unit is however necessary for the semantic interpretation of the sentence. Depending on the dependency relations, we distinguish:

ellipsis of a governing meaning unit (see Section 6.6.1.1, “Ellipsis of a governing meaning unit”),  
ellipsis of a dependent meaning unit (see Section 6.6.1.2, “Ellipsis of a dependent meaning unit”).

### 6.6.1.1. Ellipsis of a governing meaning unit

**Ellipsis of a governing meaning unit** is such ellipsis when the surface form of the sentence does not contain a meaning unit governing the present modifications or clauses. There are the following subtypes:

- ellipsis of the governing predicate**, i.e. omission of the governing predicate in verbal clauses.
- ellipsis of the governing noun**, i.e. omission of the governing noun with adjectival modifications.
- ellipsis of the governing clause**, i.e. omission of the governing clause in cases where there is a dependent clause.
- specific types of ellipsis of a governing meaning unit**, i.e. omission of a meaning unit that has to be represented in the deep structure in order to represent properly the meaning of some more complex constructions. These are:

ellipsis of the lexical unit expressing the degree of similarity/agreement/disagreement, which is the governing node for a dependent comparative clause,  
ellipsis of the totalizer governing a dependent restrictive clause,  
ellipsis of the lexical unit expressing the extent, which is the governing node for consecutive clauses.

The following types are distinguished:

- contextual ellipsis**, tj. such ellipsis when the lexical content of the omitted element is clear from the context and easily recoverable. The element was elided since it is not necessary to repeat it for full interpretation.

- b. **grammatical ellipsis**, i.e. such ellipsis when the elided element cannot be recovered from the context but when it is necessary for grammatical and semantic reasons.

**Representing ellipsis of a governing meaning unit.** For a survey of individual types of ellipsis of a governing meaning unit, see Table 6.5, “**Representing ellipsis of a governing meaning unit**”.

**Table 6.5. Representing ellipsis of a governing meaning unit**

Types of ellipsis		Contextual ellipsis	Grammatical ellipsis
<b>Ellipsis of the predicate</b>		copied node	#EmpVerb
<b>Ellipsis of the governing noun</b>		copied node	#EmpNoun
<b>Ellipsis of the governing clause</b>	for a dependent content or adverbial clause	copied node	#EmpVerb
	for a dependent relative clause	copied node	#EmpNoun
<b>Specific types of ellipsis</b>	Ellipsis of the element governing a dependent comparative clause	#Equal	
	Ellipsis of a totalizer governing a dependent restrictive clause	#Total	
	Ellipsis of the element expressing the extent, governing a dependent consecutive clause	#AsMuch	

Examples of ellipsis of the predicate:

(Jirka navštívil Marii.) Honza Jiřinu. = (Jirka visited Marie.) Honza Jiřina.) = (Jirka navštívil Marii.) Honza navštívil Jiřinu. (= (Jirka visited Marie.) Honza visited Jiřina.) Fig. 6.51

Nac̄ {#EmpVerb.PRED} ten spěch? (=Why the rush?) Fig. 6.52

{#EmpVerb.PRED} Samozřejmě. (=Of course.)

Examples of ellipsis of the governing noun:

Modré pantofle jsou maminky, zelené jsou bratrovy. (=The blue slippers are my mother's, the green ones my brother's.) = Modré pantofle jsou maminky, zelené pantofle jsou bratrovy. (=The blue slippers are my mother's, the green ones are my brother's.) Fig. 6.53

Přišli jen {#EmpNoun.ACT} mladší. (=Only the younger ones came.) Fig. 6.54

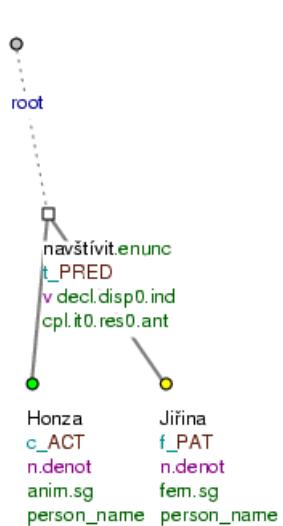
Examples of ellipsis of the governing clause:

(Neodešla.) Protože by to nestihla. (=She didn't leave.) Because she wouldn't be there in time.) Fig. 6.55

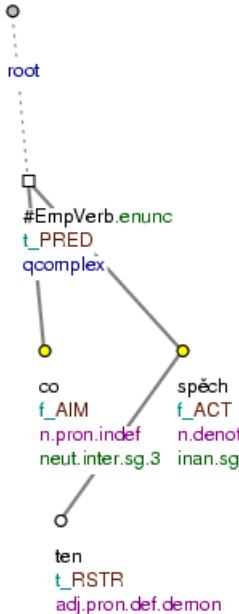
(Čtu všechny knihy). Které jsou dobrodružné. (=I read all books) Which are full of adventure.) Fig. 6.56

**Figure 6.51. Contextual ellipsis of the governing predicate**

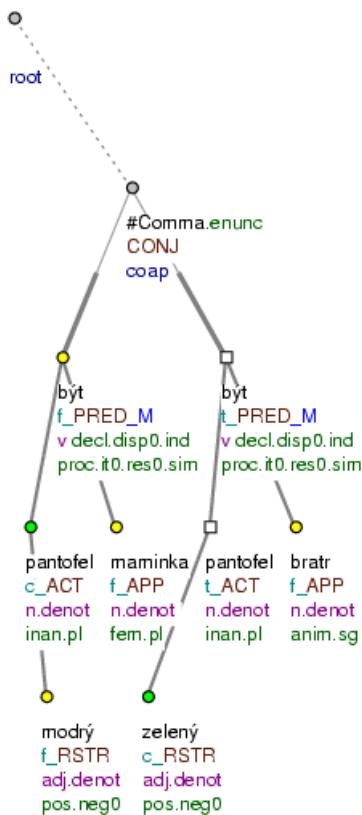
(Jirka navštívil Marii.) Honza Jiřinu. (=lit. (Jirka visited Marie) Honza Jiřina.)



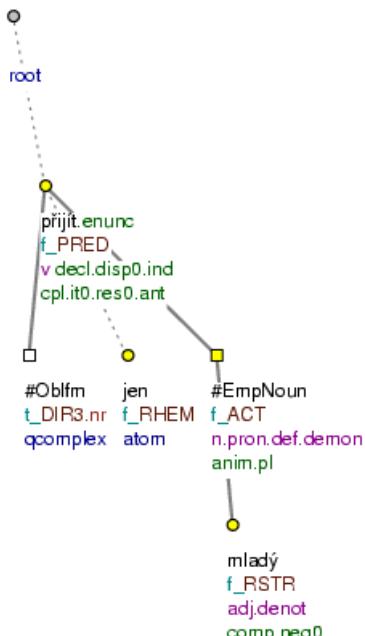
**Figure 6.52. Grammatical ellipsis of the governing predicate**



**Figure 6.53. Contextual ellipsis of the governing noun**  
Modré pantofle jsou maminky, zelené bratrov. (=lit.  
Blue slippers are mother's, green brother's.)

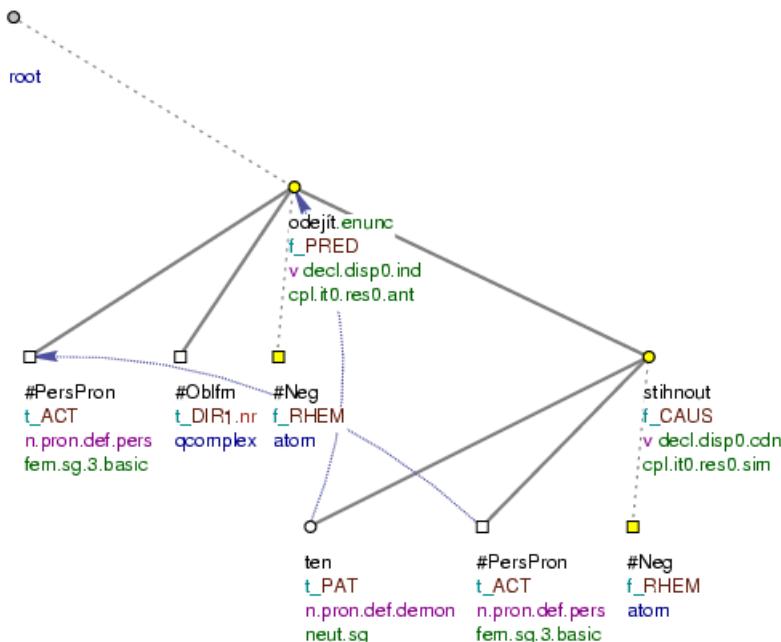


**Figure 6.54. Grammatical ellipsis of the governing noun**  
Přišli jen mladší. (=lit. Came only younger.)

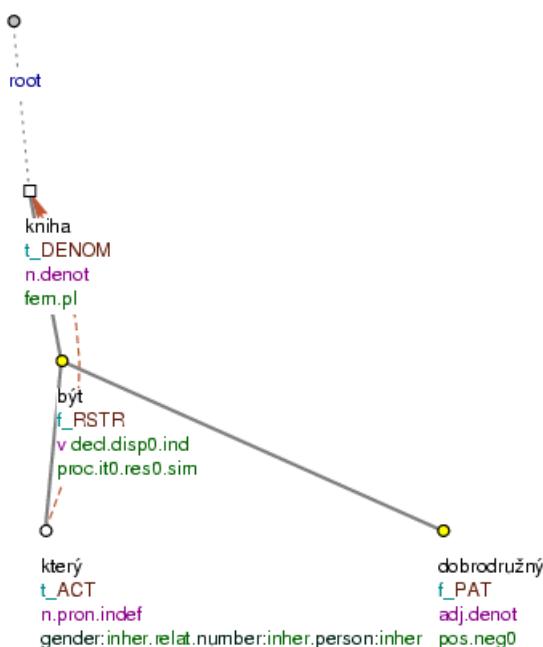


**Figure 6.55. Ellipsis of the clause governing an adverbial clause**

(Neodešla.) Protože by to nestihla. (=lit. ((She) did\_not\_leave.) Because (she) would it not\_make.)


**Figure 6.56. Ellipsis of the clause governing a relative clause**

(Čtu všechny knihy). Které jsou dobrodružné. (=lit. ((I) read all books) Which are adventurous.)



**Contextual ellipsis of a multi-word predicate.** In the case a multi-word predicate is elided, both its parts are copied. For example:

(Jirkovi běhá mráz po zádech.) Honzovi také. (=It's giving Jirka the creeps) Honza too.) = (Jirkovi běhá mráz po zádech.) Honzovi také běhá mráz po zádech.

(Jirka má zájem studovat.) Honza také. =(Jirka is interested in studying) Honza too.) = (Jirka má zájem studovat.) Honza má také zájem studovat.

**Contextual ellipsis of a full verb (following a modal or auxiliary verb).** Cases in which a full verb following a modal or auxiliary verb is elided also belong to cases of elided governing predicates. The whole predicate is represented by a single copied node, the meaning of the modal or auxiliary verb is captured by appropriate grammatical features. For example:

(*Budeš se učit?*) *Budu.* (=Will you be learning?) I will.) = (*Budeš se učit?*) *Budu se učit.*

(*Musiš už jít?*) *Musím.* (=Do you have to go?) I do.) = (*Musiš už jít?*) *Musím jít.*

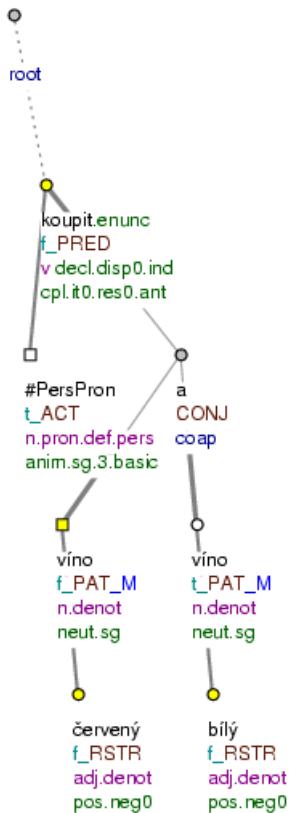
**Contextual ellipsis of the governing noun in constituent coordination.** Ellipsis in constituent coordination is also a case of ellipsis of the governing noun. These are those cases of constituent coordination (parataxis of sentence parts) in which the governing noun is the same in both (all) coordinates and therefore not repeated. It has to be clear that there are really two or more different entities and not just one entity described by several modifiers. For example:

*Koupil červené a bílé víno.* (=He bought red and white wines.) = *Koupil červené víno a bílé víno.* Fig. 6.57

*ministerstvo vnitra a zdravotnictví* (=Ministries of the Interior and Health) = *ministerstvo vnitra a ministerstvo zdravotnictví*

**Figure 6.57. Contextual ellipsis of the governing noun in constituent coordination**

*Koupil červené a bílé víno.* (=lit. (He) bought red and white wine.)



### 6.6.1.2. Ellipsis of a dependent meaning unit

**Ellipsis of a dependent meaning unit** is such ellipsis when a dependent meaning unit is omitted at the surface level but it is present in the meaning of the sentence.

**Representing ellipsis of a dependent meaning unit.** Ellipsis of a dependent meaning unit is represented with the help of newly established nodes with t-lemma substitutes. Various types of ellipsis of a dependent meaning unit are distinguished by the use of different t-lemma substitutes. A survey of individual types of ellipsis of a dependent meaning unit and the relevant t-lemma substitutes is to be found in Table 6.6, “**Representing ellipsis of a dependent meaning unit**”.

**Table 6.6. Representing ellipsis of a dependent meaning unit**

Types of ellipsis		T-lemma of the newly established node
<b>Ellipsis of an obligatory modification</b>	Contextual ellipsis of an argument	#PersPron
	Controlled argument	#Cor
	Quasi-controlled argument	#QCor
	Valency modification missing due to the presence of reciprocity	#Rcp
	General argument	#Gen
	Unspecified Actor	#Unsp
	Obligatory adjunct	#Ob1fm
<b>Ellipsis of a non-obligatory modification</b>	Non-expressed Beneficiary in control constructions	#Benef
<b>Specific type</b>	Non-verbal part of an elided verbonominal predicate (in a dependent comparative clause), which cannot be represented by a copied node due to semantic reasons	#Some

For more on control see Section 9.2.1, “Control”. For more on quasi-control see Section 7.1.1.4.1, “Quasi-control with complex predicates” For more on comparative constructions see Section 6.6.1.3.1, “Comparative constructions: comparison of two events”.

**Contextual ellipsis of an argument** is such a case of ellipsis in which the lexical content of the omitted argument is clear from the context and easily recoverable. There is a textual coreference involved (see Section 9.3, “Textual coreference”). Examples:

(Zabalil prodavač už tu knihu?) Zabalil {#PersPron.ACT} {#PersPron.PAT} (=Has the shop assistant wrapped the book?) He has.) Fig. 6.58

(Firma měla doručit zboží zákazníkovi.) Doručení {#PersPron.ACT} {#PersPron.PAT} {#PersPron.ADDR} se však neuskutečnilo. (=The company was supposed to deliver the goods to the customer) The delivery did not take place, however.)

{#PersPron.ACT} Jsi zlý. (=You are mean.)

**Ellipsis of an obligatory adjunct:** Example:

Vedoucí podniku odcestoval {#Ob1fm.DIR1} (=The boss has left.) Fig. 6.59

**General argument (#Gen)** is a term used for such non-expressed arguments that refer to a type of modification usual in a given position, rather than a particular lexical unit (as is the case with contextual ellipsis). The lexical content of an argument has to be the usual one for the given verb (noun, adjective) in the given position in order for the argument to become general. A general argument refers to the “usual”, “typical entity”. The t-lemma substitute for general arguments is #Gen. For example:

Domy se stavějí z cihel. {#Gen.ACT} (=Houses are built from bricks.) Fig. 6.60

V téhle troubě se mi dobře peče. {#Gen.PAT} (=This oven is good for baking.)

**Unspecified Actor (#Unsp).** Apart from cases of contextual ellipsis on the one hand and general arguments on the other, there is also a transitory case: the so called unspecified Actor. This involves cases when a modification absent at the surface level denotes an entity more or less known from the context which is, however, not explicitly referred to. The entity with the same reference as the non-expressed Actor cannot be really specified; the absent Actor refers rather to the preceding context than to a particular lexical unit; nevertheless, the choice of the referent is not free and it is in general possible to narrow down the possibilities. The t-lemma substitute for unspecified Actors is #Unsp. For example:

Hlásili to v rozhlas. {#Unsp.ACT} (=They announced it on the radio.) Fig. 6.61

**Reciprocity (#Rcp).** The term reciprocity is used for the syntactic operation on valency frames that puts two different valency modifications (arguments or obligatory adjuncts) in a symmetric relation, which can be expressed in the following way:

*Jan a Marie se setkali.* (=Jan and Marie met.) = *Jan se setkal s Marií a (zároveň) Marie se setkala s Janem.* (=Jan met Marie and (simultaneously) Marie met Jan.)

As a result of the presence of a reciprocal relation in the sentence, one of the obligatory modification positions is lost at the surface level. The other position involved is occupied by both modifications standing in the reciprocal relation at the same time (coordination, plural). Semantically, however, they correspond to two different valency positions. The fact that there is a reciprocal meaning in the sentence is usually signalled by the presence of *se* (*k sobě, mezi sebou*), nevertheless it is not a necessary condition for the construction to be interpreted as involving reciprocity.

**Representing reciprocity in the tectogrammatical trees.** Reciprocity is represented by means of inserting a newly established node with the #Rcp t-lemma into the structure, in the position of the valency modification that was left out. The newly established node has the functor corresponding to the unoccupied valency position. The relation between the newly established node and the valency modification containing both members of the relation is indicated in the tree as a case of grammatical coreference (see Section 9.2, “Grammatical coreference”). *Se*, possibly present at the surface level, is not represented by a separate node at the tectogrammatical level; the reference to the relevant analytical node(s) is included in the *a* attribute of the newly established node with the #Rcp t-lemma.

Examples:

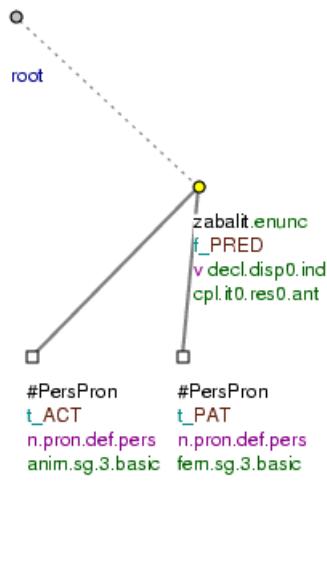
*Jan.ACT a Marie.ACT <se> líbali. {#Rcp.PAT}* (=Jan and Marie kissed.) Fig. 6.62

*Státy.ACT Evropské unie <mezi sebou> obchodují. {#Rcp.ADDR}* (=EU states trade with each other.)

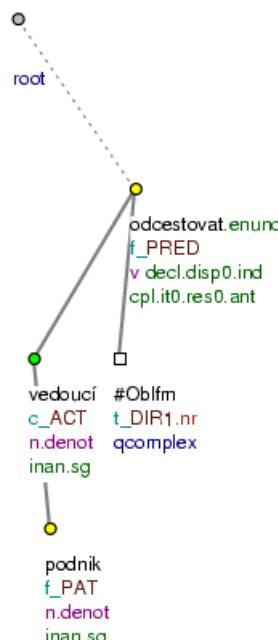
*Porovnávali Německo.PAT a Koreu.PAT {#Rcp.EFF}* (=They compared Germany and Korea.)

*jednání Petra.ACT a Pavla.ACT o prodeji domu trvalo několik hodin. {#Rcp.ADDR}* (=negotiations between Petr and Pavel...)

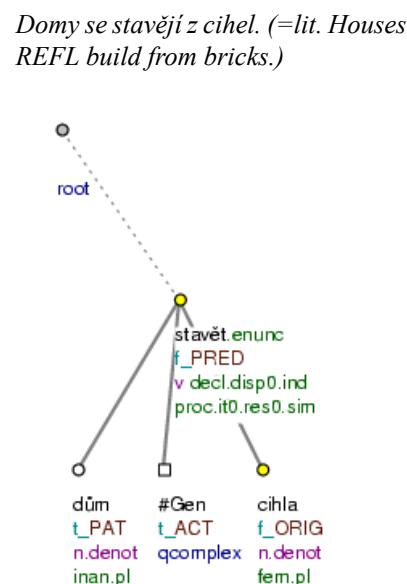
**Figure 6.58. Contextual ellipsis of an obligatory argument**



**Figure 6.59. Ellipsis of an obligatory adjunct**

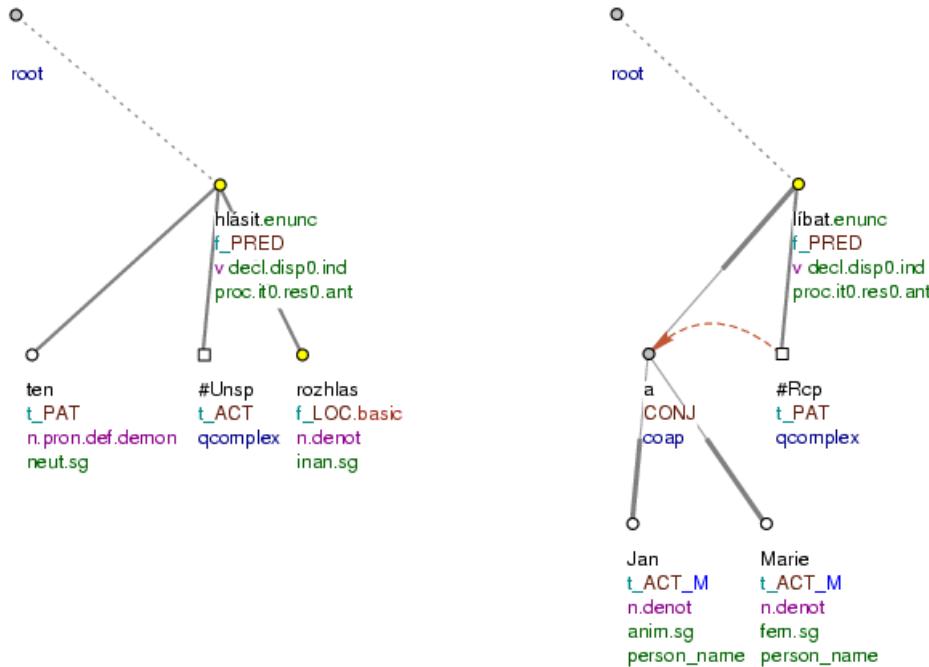


**Figure 6.60. General argument**



**Figure 6.61. Unspecified Actor**

*Hlásili to v rozhlas. (=lit. (They) announced it on radio.) Jan a Marie se libali. (=Jan and Marie REFL kissed.)*



### 6.6.1.3. Specific elliptical constructions

Special cases of ellipsis of a governing meaning unit are:

- comparative constructions (comparison of two events, see Section 6.6.1.3.1, “Comparative constructions: comparison of two events”),
- constructions with the meaning of a restriction (see Section 6.6.1.3.2, “Constructions with the meaning of a restriction”),
- constructions with consecutive clauses (see Section 6.6.1.3.3, “Constructions with consecutive clauses”).

#### 6.6.1.3.1. Comparative constructions: comparison of two events

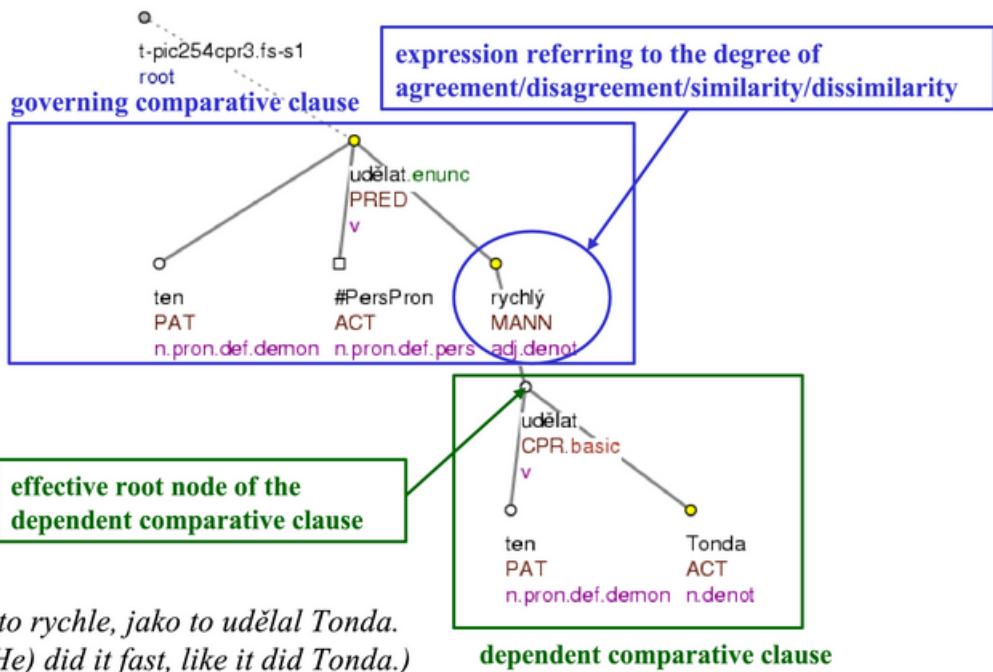
**Constructions with the meaning of comparison of two events** are constructions in which two events or states are compared. There is always a property or degree of agreement/disagreement/similarity with respect to which the events are compared.

Two basic types are distinguished, depending on the form of comparison:

- a. comparison based on similarity or identity, by means of the conjunction *jako* (*Pavel běhá stejně rychle, jako běhá Honza.* (=Pavel runs as fast as Honza.)),
- b. comparison on the basis of dissimilarity, by means of the conjunction *než* (*Dorazil dříve, než dorazil Jirka.* (=He arrived earlier than Jirka.)).

**Representing constructions with the meaning of comparison of two events.** There is a governing comparative clause and a dependent comparative clause. The governing comparative clause contains an element expressing the degree of agreement/disagreement/similarity/dissimilarity. The effective root of the dependent clause is assigned the CPR functor. In regular cases, it depends on the expression referring to the degree of agreement/disagreement/similarity/dissimilarity in the governing clause. For an illustration of the way comparative constructions are analyzed, see Fig. 6.63.

Figure 6.63. Comparative constructions: comparison of two events



**Ellipsis of the expression referring to the degree of agreement / disagreement / similarity / dissimilarity.** If the expression referring to the degree of agreement/disagreement/similarity/dissimilarity is omitted at the surface level (constructions with *jako*), a new node is added into the structure, with the t-lemma substitute `#Equal`. For example:

*Udělal to {#Equal.RSTR}, jako to udělal Tonda.* (=He did it like Tonda.)

= *Udělal to stejně/stejným způsobem, jako to udělal Tonda.* (=He did it the same way Tonda did it.)

**Ellipsis of the predicate in the dependent clause.** The predicate of the dependent comparative clause is often omitted in the surface form of the sentence. This ellipsis results from the identity of the verbs in the governing and the dependent clauses: their lexical values are identical (see Section 6.6.1.1, “Ellipsis of a governing meaning unit”). If the governing verb of the dependent clause is not present in the surface structure of the sentence, a new node with the functor CPR is inserted into the dependent clause on the position of the effective root of the clause (following the rules in Section 6.6.1.1, “Ellipsis of a governing meaning unit”). For example:

*Udělal to jako Tonda.* (=He did it like Tonda.)

= *Udělal to stejně/stejným způsobem, jako to udělal Tonda.* (=He did it the same way Tonda did it.) Fig. 6.64

*Dorazil dříve než Jirka.* (=He arrived earlier than Jirka.)

= *Dorazil dříve, než dorazil Jirka.* (=He arrived earlier than Jirka did.)

*Je zdravý jako ryba.* (=He is fit as a fiddle; lit. as fish.)

= *Je zdravý, jako je zdravá ryba.* (=lit. (he) is fit as is fit fish.) Fig. 6.65

**Node with the t-lemma substitute #Some.** A node with the t-lemma substitute `#Some` is added to the structure on the position of the non-verbal part of the elided verbonominal predicate in the dependent comparative clause if it is impossible (due to semantic reasons) to copy the corresponding node from the governing clause (where this can be e.g. *stejný* (=the same), *podobný* (=similar), *jiný* (=other), *více* (=more), *méně* (=less), *jiný* (=different)). For example:

*Je stejný jako já.* (=He's the same as me.)

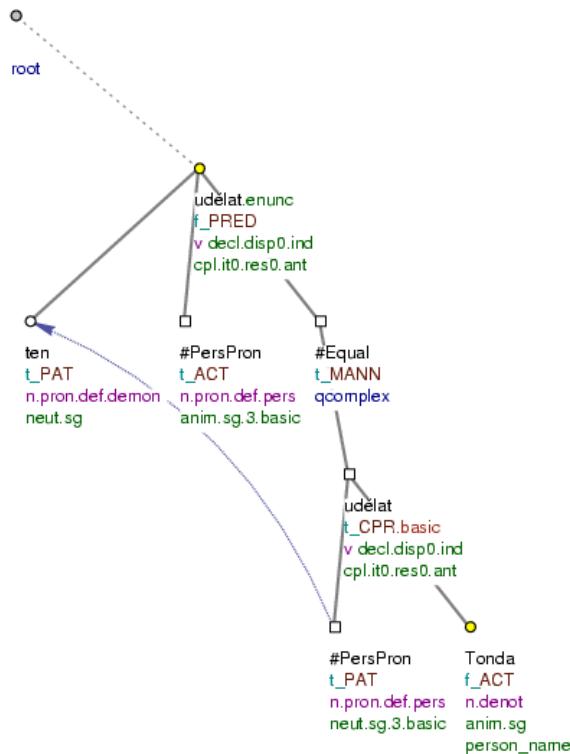
= *Je stejný, jako já jsem "nějaký".* (=He's the same as I am such-and-such.) Fig. 6.66

*Situace v armádě je jiná než na ministerstvu.* (=The situation in the army is different than the one at the Ministry.)

= Situace v armádě je jiná, než je situace na ministerstvu "nějaká". (=The situation in the army is different than the situation at the Ministry is such-and-such.)

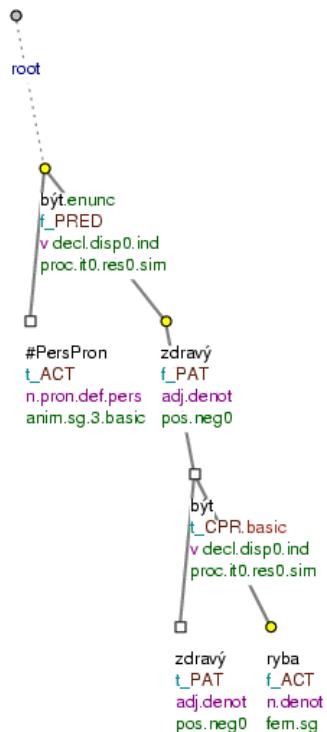
**Figure 6.64. Comparative constructions: comparison of two events**

*Udělal to jako Tonda.* (=lit. (He) did it like Tonda.)



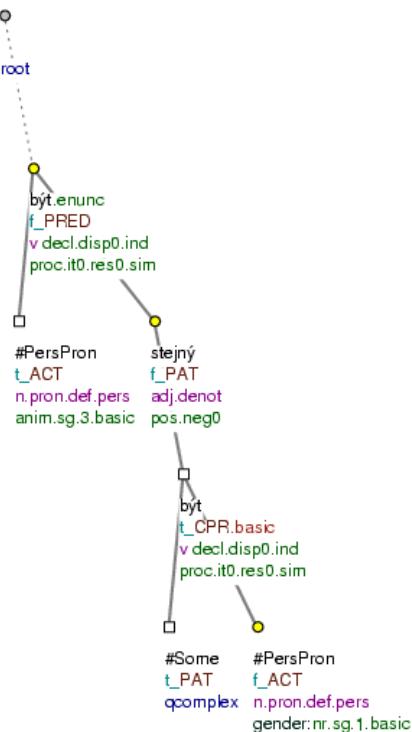
**Figure 6.65. Comparative constructions: comparison of two events**

*Je zdravý jako ryba.* (=lit. (He) is healthy as fish.)



**Figure 6.66. Comparative constructions: comparison of two events**

*Je stejný jako já.* (=lit. (He) is same as me.)



### 6.6.1.3.2. Constructions with the meaning of a restriction

**Constructions with the meaning of a restriction** (restrictive constructions) are such constructions that restrict the validity of a totalizer like *každý* (=every), *celý* (=whole), *všechn* (=all), *nic* (=nothing), *nikam* (=nowhere) etc., or introduce an exception to a regular state.

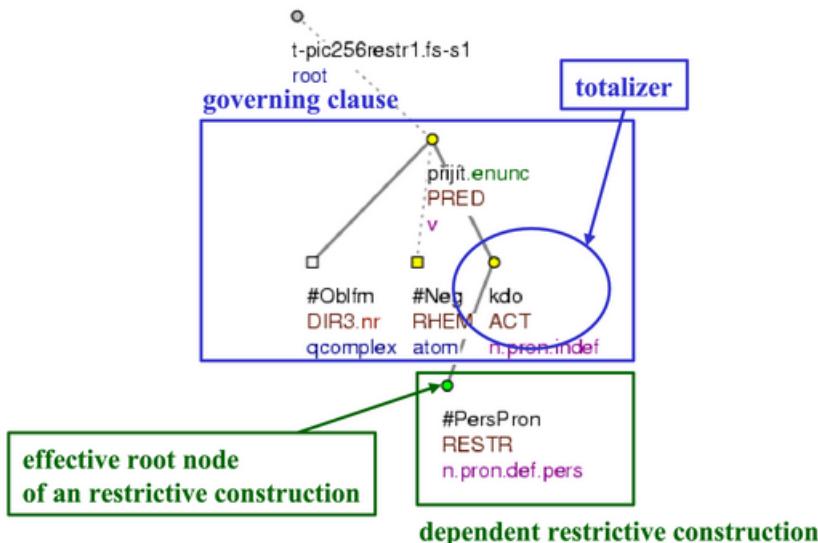
**Representing constructions with the meaning of a restriction.** There is a governing clause and a dependent clause in restrictive constructions. The governing clause contains a totalizer or an expression referring to a regular state etc. The effective root of the dependent restrictive clause is assigned the RESTR functor and depends on the node representing the totalizer. For an illustration of the way restrictive constructions are analyzed, see Fig. 6.67.

**Ellipsis of the totalizer.** If the totalizer is not present at the surface level, a new node with the t-lemma substitute #Total is added to the tectogrammatical tree. The node with the t-lemma substitute #Total stands for any absent positive totaliser (*všichni* (=all), *všechno* (=everything), *každý* (=each), *všude* (=everywhere), *vždycky* (=always) or a negative totalizer (*nic* (=nothing), *nikdo* (=nobody), *žádný* (=no/none), *nikam* (=nowhere), *nikdy* (=never)). The added node for the totaliser has a functor corresponding to the position in which it was added. For example:

*Mimo datum.RESTR se piší {#Total.RSTR} řadové číslice slovy.* (=Except for dates, the ordinals are written in words.)

= *Mimo datum se piší všechny řadové číslice slovy* (=Except for dates, all ordinals are written in words.) Fig. 6.68

Figure 6.67. Constructions with the meaning of a restriction



*Mimo tebe nepřišel nikdo.*  
 (=lit. Apart\_from you, not\_came nobody.)

**Effective root of a restrictive construction.** The effective root of a restrictive clause (with the RESTR functor) is determined depending on what means are used for expressing the restriction:

- A. **restrictive construction introduced by a preposition.** As for constructions introduced by prepositions (*až na*, *vyjma*, *kromě*, *mimo* (=except), *vedle* (=beside)), there are several types:
  - a. **direct restriction:** the following noun is in the required case. The effective root node of the restrictive construction is the node representing the governing (syntactic) noun in the prepositional phrase. For example:

*Kromě tohoto týdne.RESTR budu učit pravidelně.* (=Except for this week, I'll be teaching regularly.)

- b. **restriction with ellipsis of the verb.** the noun following the preposition is in a case different from that required by the preposition or the preposition is followed by another prepositional phrase. The effective root of the restrictive clause is a newly established node for a verb (see the rules in Section 6.6.1.1, "Ellipsis of a governing meaning unit"). For example:

*Kromě do Prahy chtěli jet všude.* (=Except for Prague, they wanted to go everywhere.)

= *Kromě toho, že nechtěli jet do Prahy, chtěli jet všude.* (= Apart from the fact that they did not want to go to Prague, they wanted to go everywhere.) Fig. 6.69

- c. **restrictive construction introduced by “kromě / vedle / mimo toho, že (=except for/apart from the fact that)”.** The effective root of the restrictive constructions is the effective root of the dependent clause. *Toho* is a supporting element. For example:

<Kromě toho, že> krásně zpívá.RESTR , neumí nic.. (= Apart from the fact that he/she sings beautifully, he/she is not capable of anything.) Fig. 6.70

- B. **restrictive construction introduced by a connective.** Restrictive constructions introduced by connectives (*než, ledaže*) are always interpreted as dependent verbal clauses. If there is no governing verb present at the surface level in the restrictive construction, a new node is added to the tree in place of this absent predicate (following the rules in Section 6.6.1.1, “Ellipsis of a governing meaning unit”). For example:

*Nepůjdu nikam než do Prahy.* (=I won't go anywhere but to Praha.)

= *Nepůjdu nikam, než půjdu do Prahy.* Fig. 6.71

*Nezbývá mi nic než doufat*.RESTR (=I can only hope.)

**Constructions with the meaning of exceptional conjoining.** Constructions with the meaning of exceptional conjoining are different from standard restrictive constructions (with the prepositions *kromě, mimo, vedle*); they do not express a simple restriction nor simple conjoining. Both conjoined constituents are semantically more closely defined: one has the meaning of the “commonplace, self-evident”, the other conjoined constituent, by contrast, has the meaning of the “unusual, exceptional”.

Annotation rules are similar to those for restrictive constructions (the effective root of these constructions is also assigned the RESTR functor). The difference is that the effective root of the construction with the meaning of exceptional conjoining depends on the node for the predicate of the governing clause. For example:

*Kromě do Říma chtěli jet i do Benátek.* (= Except for Rome, they also wanted to go to Venice.)

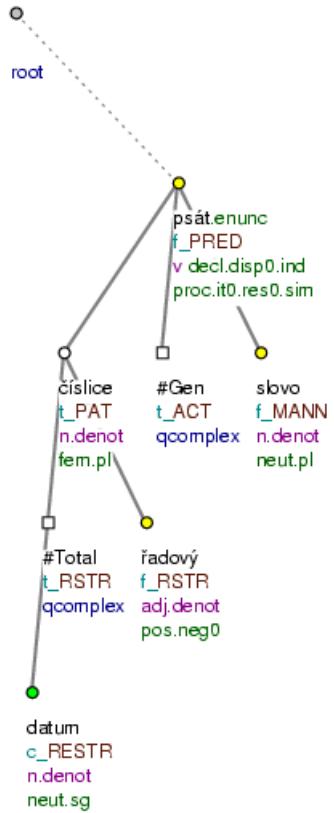
= *Kromě toho, že chtěli jet do Říma, chtěli jet i do Benátek.* Fig. 6.72

*Kromě ní*.RESTR tam byl i Mirek. (= Except for her, also Mirek was there.)

<Kromě toho, že> byl.RESTR učitelem, byl i dobrým člověkem. (= Except for the fact that he was a teacher he was a good person too.)

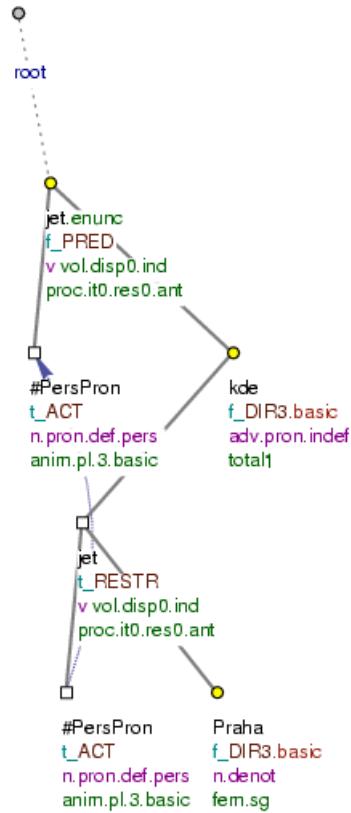
**Figure 6.68. Constructions with the meaning of a restriction**

*Mimo datum se píší řadové číslice slovy (=lit. Except date REFL write ordinal numerals in \_words.)*



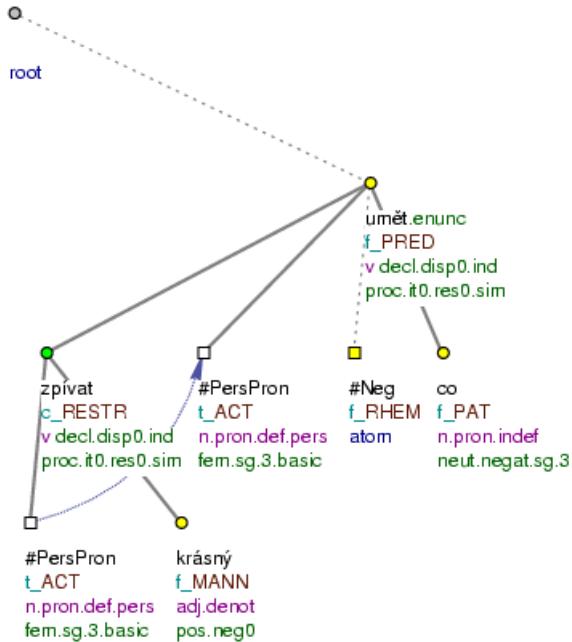
**Figure 6.69. Constructions with the meaning of a restriction**

*Kromě do Prahy chtěli jet všude.* (=lit. Except to Prague (they) wanted to go everywhere.)



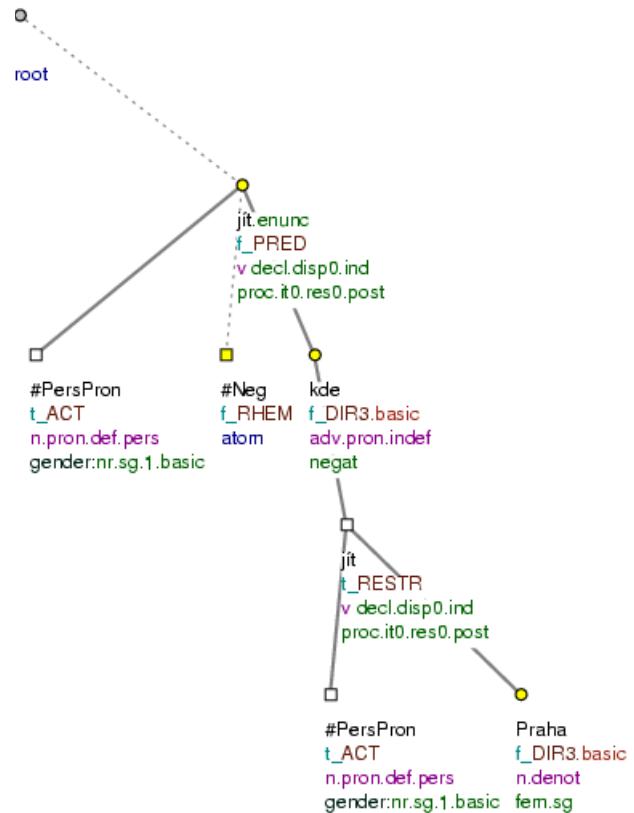
**Figure 6.70. Constructions with the meaning of a restriction**

*Kromě toho, že krásně zpívá, neumí nic.* (=lit. Except that that (he) beautifully sings, (he) cannot nothing.)



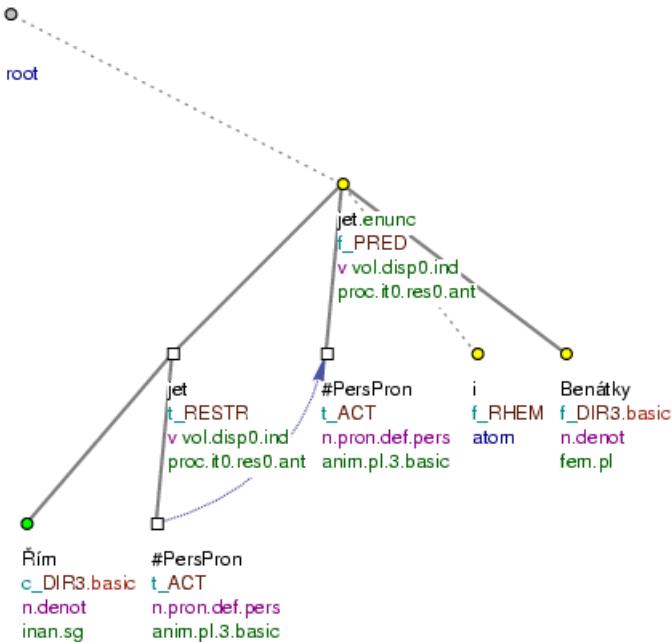
**Figure 6.71. Constructions with the meaning of a restriction**

*Nepůjdu nikam než do Prahy.* (=lit. (I) will\_not\_go nowhere but to Praha.)



**Figure 6.72. Constructions with the meaning of exceptional conjoining**

*Kromě do Říma chtěli jet i do Benátek.* (=lit. Except to Rome (they) wanted to \_go also to Venice.)

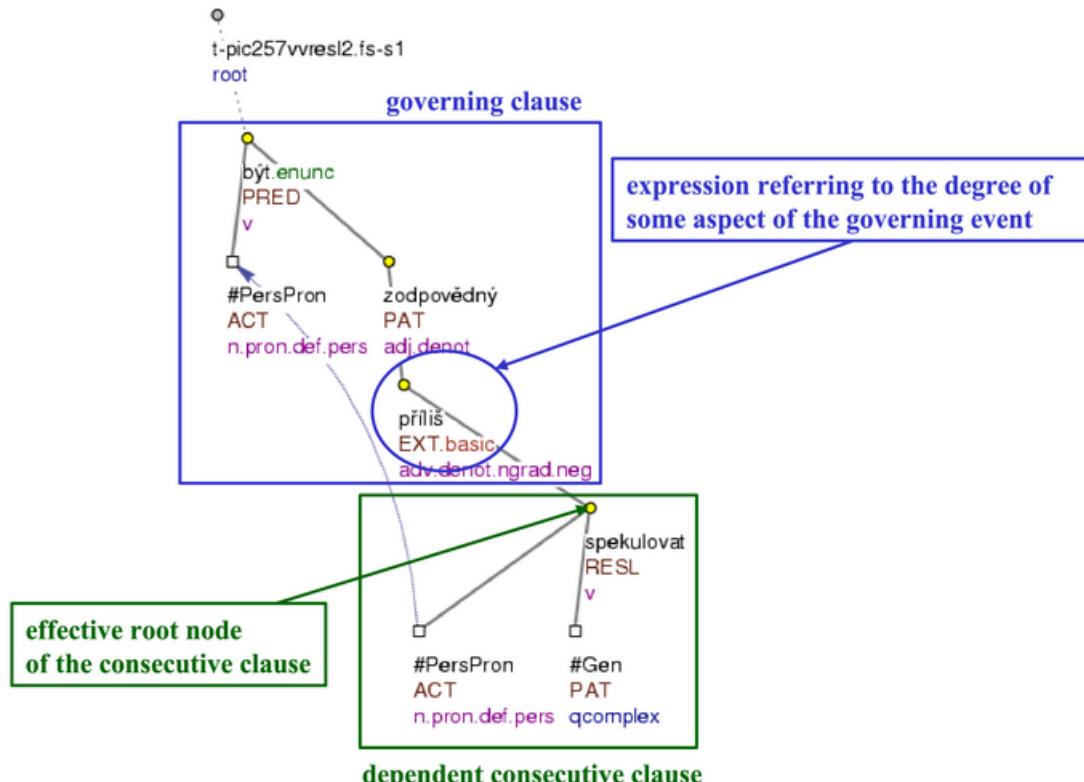


### 6.6.1.3.3. Constructions with consecutive clauses

**Dependent consecutive clause** is such a dependent clause that expresses the effect that comes about as a result of the extent of some aspect of the governing event. The governing clause contains an expression like *tolik* (=so much), *natolik* (=enough), *příliš* (=too much), *tak* (=so), *takový* (=such), *dost* (=enough), *dostatečně* (=enough), *do té míry* (=to such a degree) that introduces the consecutive clause.

**Representing constructions with consecutive clauses.** The node for the expression referring to the degree of some aspect of the governing event is assigned a functor according to its position; if it is a non-valency modification, it is usually assigned the EXT functor. The effective root of the consecutive clause has the RESL functor and depends on the expression referring to the degree of some aspect of the governing event. An illustration of the way consecutive clauses are analyzed can be found in Fig. 6.73.

Figure 6.73. Representing constructions with consecutive clauses



*Byl příliš zodpovědný, než aby spekuloval.  
(=lit. (He) was too responsible than so that (he) speculated.)*

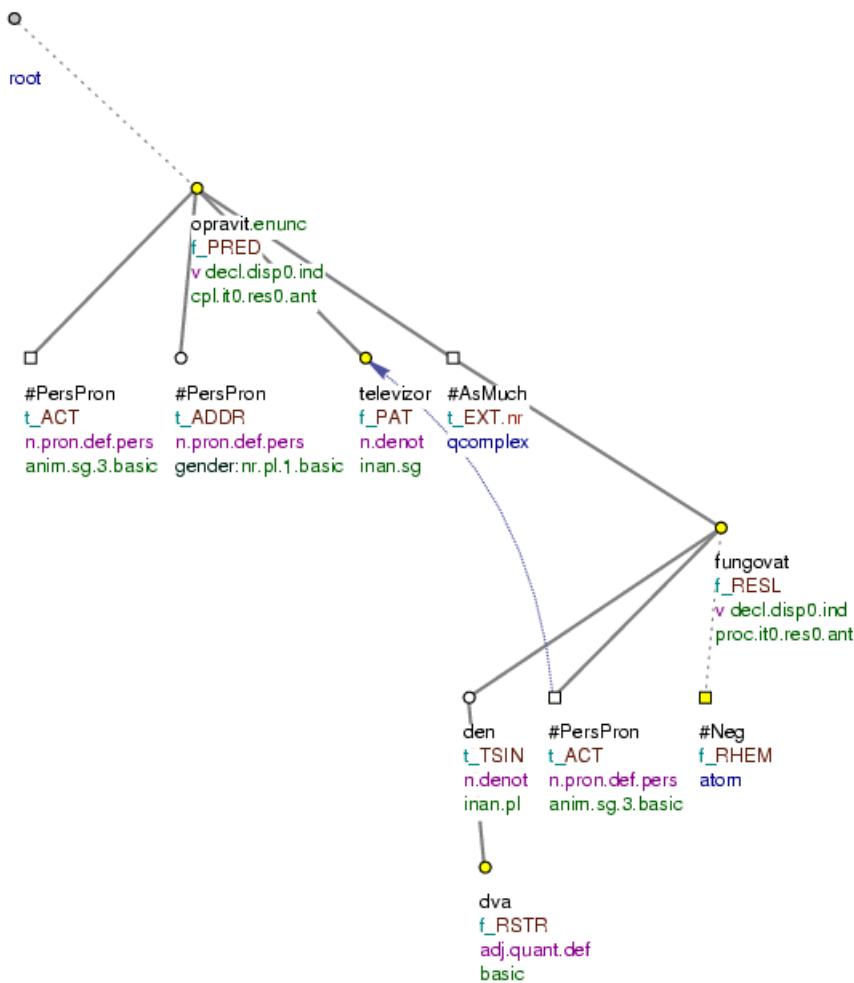
**Ellipsis of the expression referring to the degree of some aspect of the governing event.** The expression referring to the degree of some aspect of the governing event can be omitted in the surface structure; then, a new node with the t-lemma substitute #AsMuch and an appropriate functor (usually EXT) is inserted in its position. The node with the t-lemma #AsMuch stands for both a large and small degree of an aspect of the event (e.g.: *tak málo* (=so little), *tak špatně* (=so badly), *tak dobře* (=so well), *tak hodně* (=so much), *tak moc* (=so much)). For example:

*Opravil nám televizor, že za dva dny nefungoval.RESL* (=He repaired the TV set so badly that it didn't work in two days again.)

= *Opravil nám televizor tak špatně, že za dva dny nefungoval.* Fig. 6.74

**Figure 6.74.** Dependent consecutive clause

Opravil nám televizor, že za dva dny nefungoval. (=lit. (He) repaired us TV\_set that in two days (it) did\_not\_work.)



## 6.6.2. Newly established nodes in more complex syntactic structures

In order to represent properly certain more complex structures, some auxiliary nodes are needed. If there is no surface-level expression suitable for the purpose, a new node is added to the tree. The newly established node has always a t-lemma substitute. Different t-lemma substitutes distinguish between various subtypes of auxiliary nodes (see Table 6.7, “Newly established nodes in more complex syntactic structures”).

**Table 6.7.** Newly established nodes in more complex syntactic structures

Type of newly established node	T-lemma of the newly established node
paratactic structure root node	#Separ
root node of a list structure for foreign-language expressions	#Forn
root node of an identification structure	#Idph

For more on paratactic structures, see Section 6.4, “Parataxis”. For more on list structures for foreign-language expressions see Section 7.4, “Foreign-language expressions”. For more on identification structures, see Section 7.3.1, “Identification structure”.

### 6.6.3. Nodes representing negation with verbs

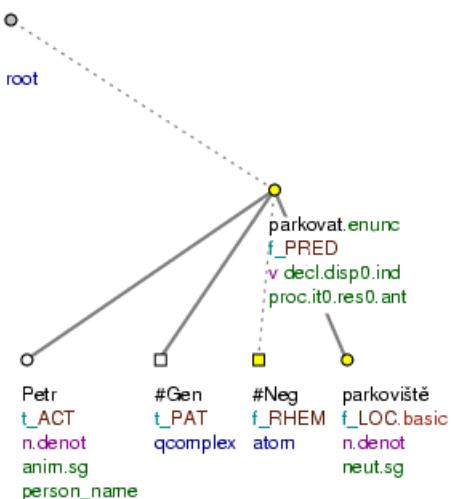
A specific type of newly established node is a node for negation expressed by the verbal morpheme *ne-*. The morpheme *ne-* with negated verbs is represented by a newly established node with the t-lemma substitute #Neg (the node is either a rhematizer (functor = RHEM; see Section 10.4, “Rhematizers”), or conjunction modifier (functor = CM; see Section 7.1, “Multi-word lexical units”)). The t-lemma of the node for the verb is in its non-negated form. The node with the t-lemma substitute #Neg is added also under copied nodes if necessary.

Example:

*Petr neparkoval na parkovišti. {#Neg}* (=Petr didn't park in the parking lot.) Fig. 6.75

**Figure 6.75. Nodes representing negation with verbs**

*Petr neparkoval na parkovišti. (=lit. Petr not\_parked in parking\_lot.)*



## 6.6.4. Survey of newly established nodes

The following table (Table 6.8, “Survey of newly established nodes”) offers a survey of newly established nodes (they are listed in the alphabetical order according to their t-lemmas).

**Table 6.8. Survey of newly established nodes**

T-lemma	Type of newly established node	Example
#AsMuch	ellipsis of the expression referring to the degree of an aspect of the governing event, which governs a consecutive clause	<i>Opravil nám televizor {#AsMuch.EXT}, že za dva dny nefungoval.RESL</i> (=He repaired the TV set in such a way that it didn't work in two days again.)
#Benef	ellipsis of the Beneficiary in control constructions	<i>Je vhodné {#Benef.BEN} odejít.</i> (=It's good to leave.)
#Cor	controlled element	<i>Rozhodl se {#Cor.ACT}odejít</i> (=He decided to leave)
#EmpNoun	grammatical ellipsis of the governing noun	<i>Přišli jen {#EmpNoun.ACT} mladší.</i> (=Only the younger ones came.)
#EmpVerb	grammatical ellipsis of the predicate	<i>{#EmpVerb.PRED} Nač ten spěch?</i> (=Why the rush?)
#Equal	ellipsis of the expression referring to the degree of similarity/agreement/disagreement, which is the governing node for a dependent comparative clause	<i>Udělal to {#Equal.MANN}, jako to udělal.CPR Tonda.</i> (=He did it just like Tonda.)
#Forn	root node of a list structure for foreign-language expressions	<i>Nenahraditelný je pro sledování {#Forn.PAT} cash flow.</i> (=Irreplaceable for cash flow monitoring.)
#Gen	general argument	<i>V téhle troubě se mi dobře peče {#Gen.PAT}</i> (=apprx. This oven is good for baking.)
#Idph	root node of an identification structure	<i>Čtu {#Idph.PAT} Proti všem.</i> (=I'm reading Proti všem.)
#Neg	verbal negation (the morpheme <i>ne-</i> )	<i>{#Neg.RHEM} Neodejdu.</i> (=I'm not leaving.)
#PersPron	contextual ellipsis of an obligatory argument	<i>{#PersPron.ACT} Rozhodl se odejít.</i> (=He decided to leave.)
#QCor	quasi-controlled element	<i>Mám {#QCor.ACT} plán studovat.</i> (=He is planning to study.)
#Rcp	obligatory modification missing due to its participation in a reciprocal relation	<i>Poslanci jednali o novém zákoně {#Rcp.ADDR}</i> (=The members of the Parliament discussed the new law.)
#Separ	paratactic structure root node	<i>{#Separ.CONJ} střed nákup prodej</i> (=middle sell buy)
#Some	non-verbal part of an elided verbonominal predicate, which cannot be represented by a copied node due to semantic reasons	<i>Je stejný jako {být} {#Some.PAT}, já.</i> (=He is just like me.)
#Total	ellipsis of a totalizer governing a dependent restrictive clause	<i>Mimo datum.RESTR se пиší {#Total.RSTR} řadové číslovky slovy.</i> (=Except for dates, the ordinals are written in words.)
#Unsp	unspecified Actor	<i>{#Unsp.ACT} Hlásili to v rozhlasu.</i> (=They announced it on the radio.)
copied node (verb)	grammatical ellipsis of the predicate	<i>(Pavel navštívil Janu.) David {navštívit.PRED} Jiřinu.</i> (=P. visited J.) David Jiřina.)
copied node (noun)	contextual ellipsis of the governing noun	<i>Má rád červené {víno.PAT} i bílé víno.</i> (=He likes both red and white wine.)

---

# Chapter 7. Specific phenomena

## 7.1. Multi-word lexical units

**Multi-word lexical units** are such collocations of two (or more) words that have a single lexical meaning. Multi-word lexical units are represented in several ways:

A. **single node + multi-word t-lemma**: the multi-word lexical unit is represented by a single node and all its parts are contained in the t-lemma. Such a t-lemma is called multi-word t-lemma. Multi-word lexical units represented in this way are listed in Section 4.3, “Multi-word t-lemma”.

B. **single node + grammateeme**: the multi-word lexical unit is represented by a single node whose t-lemma corresponds to one of the components of the lexical unit. The information regarding other components of the unit is encoded in the values of various grammateemes. This holds for:

**modal predicates** (the `deontmod` grammateeme; see Section 7.1.1.1, “Modal predicates”)

C. **more than one node+ special functor**: the multi-word lexical unit is represented by several nodes and the fact that these form a single unit is captured by using a special functor. This holds for:

**complex predicates** (the `CPHR` functor; see Section 7.1.1.4, “Complex predicates”)

**idioms** (the `DPHR` functor; see Section 7.1.2, “Idioms”)

**complex (coordinating) connectives** (the `CM` functor)

**Complex co-ordinating connectives.** When analyzing coordinating connectives, the following two are distinguished: **coordinating conjunction** and **conjunction modifier**. A complex co-ordinating connective is represented by at least two nodes: a node for the conjunction (`nodetype = coap`; see also Section 6.4, “Parataxis”) and a node for the conjunction modifier. The node for the conjunction modifier is an immediate daughter of the root of a paratactic structure and has the `CM` functor. For example:

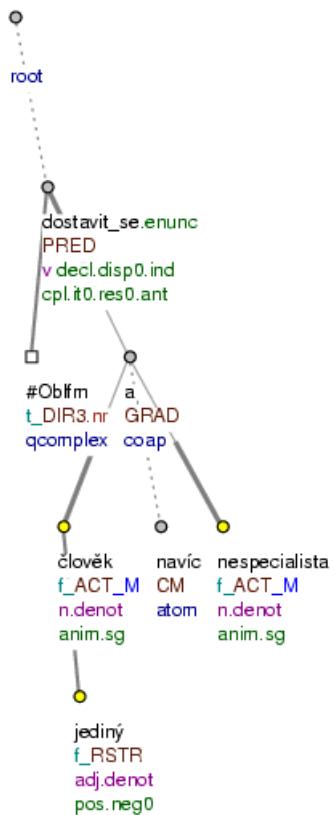
*Dostavil se jediný člověk, a navíc.CM nespecialista.* (=Only one person turned up, a layman to boot.) Fig. 7.1

D. **more than one node+ special structure**: the multi-word lexical unit is represented by several nodes and the fact that they form a unit is expressed by assigning them a special structure. This holds for:

**non-inflected titles** (see Section 7.3.1, “Identification structure”)

**Figure 7.1. Multi-word coordinating connective**

*Dostavil se jediný člověk, a navíc nespecialista.* (=lit. Arrived REFL only\_one person, and moreover layman.)



## 7.1.1. Multi-word predicates

**Multi-word predicates** are defined as cases where the predicate as a lexical unit is represented at the surface level of the sentence not only by a finite verb form but where the predicate additionally incorporates the meaning of other words - verbs, nouns, adverbs. Multi-word predicates include:

- modal predicates (see Section 7.1.1.1, “Modal predicates”),
- phase predicates (see Section 7.1.1.2, “Phase predicates”),
- quasi-modal and quasi-phase predicates (see Section 7.1.1.3, “Quasi-modal and quasi-phase predicates”),
- complex predicates (see Section 7.1.1.4, “Complex predicates”),
- verbonominal predicates (see Section 7.1.1.5, “Verbonominal predicates (the copula *byt*)”),
- verbal idioms (see Section 7.1.2, “Idioms”).

### 7.1.1.1. Modal predicates

**Modal predicates** are defined as multi-word predicates comprising a modal verb which expresses (in addition to the finite verbal meanings) the modal meaning of the predicate, and the infinitive of a full verb, carrying the main lexical meaning of the expression as a whole.

The following are distinguished:

- **modal verb:** *dát se* (=be possible); *dovést* (=be able); *hodlat* (=be willing); *chtít* (=want); *mít* (=have an obligation); *moci / moct* (=can); *muset* (=must); *smět* (=may); *umět* (can).
- **infinitive of a full verb.**

**The basic annotation of modal predicates.** A modal predicate is represented by a single node with the t-lemma of the full verb. Information on the modality of this predicate expressed by the modal verb is contained in the value of the deontic modality grammeme (*deontmod*; see Section A.2.8.3, “gram/deontmod”). The following combinations are represented in the basic way (by copying the nodes) :

(negated) modal verb + non-negated infinitive of a full verb.	<i>Petr &lt;chce&gt; přijít na koncert. [deontmod = vol] (=Petr wants to come to the concert.) Fig. 7.2</i>
(negated) modal verb + paratactic connection of non-negated infinitives of full verbs	<i>Petr &lt;nechce&gt; přijít [deontmod = vol] a zůstat [deontmod = vol]. (=Petr doesn't want to come and stay.) Fig. 7.3</i>
paratactic connection of (negated) modal verbs + non-negated infinitive of a full verb.	<i>Petr nemohl a nemůže přijít na koncert. (=Petr could not, and cannot, come to the concert.) = Petr &lt;nemohl&gt; {přijít} [deontmod = poss] a &lt;nemůže&gt; přijít [deontmod = poss] na koncert. Fig. 7.4</i>
paratactic connection of (negated) modal verbs + paratactic connection of non-negated infinitives of full verbs	<i>Takže to nemohli a nemohou potvrdit ani vyvrátit. (=So they could not and cannot confirm or deny it.) = Takže to &lt;nemohli&gt; {potvrdit} [deontmod = poss] a &lt;nemohou&gt; potvrdit [deontmod = poss] ani vyvrátit [deontmod = poss] a vyvrátit [deontmod = poss].</i>

**The basic annotation cannot be used for** modal predicates with a negated infinitive of the full verb and cases of more modal meanings within one modal predicate (layering of modal meanings). This concerns the following combinations:

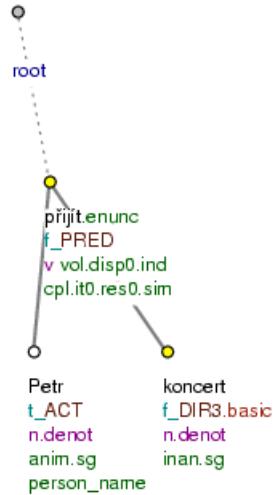
(negated) modal verb + negated infinitive of a full verb	<i>Karel může nepřijít na koncert. (=Karel may not come to the concert.)</i> <i>Karel nemůže nepřijít na koncert. (=Karel cannot not come to the concert.) Fig. 7.5</i> <i>Petr mohl a může nepřijít na koncert. (=Petr could and can not come to the concert.) Fig. 7.6</i>
modal verb + modal verb + infinitive of a full verb	<i>Petr může [deontmod = decl] &lt;chtít&gt; přijít.PAT [deontmod = vol] na koncert. (=Petr may want to come to the concert.) Fig. 7.7</i> <i>Petr nemůžechtí.PAT nepřijít.PAT na koncert. (=Petr cannot want not to come to the concert.)</i>

**Modal predicate with a negated infinitive of a full verb.** In cases in which the infinitive of the full verb is negated both the modal verb and the infinitive of the full verb are represented by separate nodes. The grammateeme deont-mod is assigned the value decl at both nodes. The node representing the infinitive of the full verb has the functor PAT and is dependent on the node representing the modal verb.

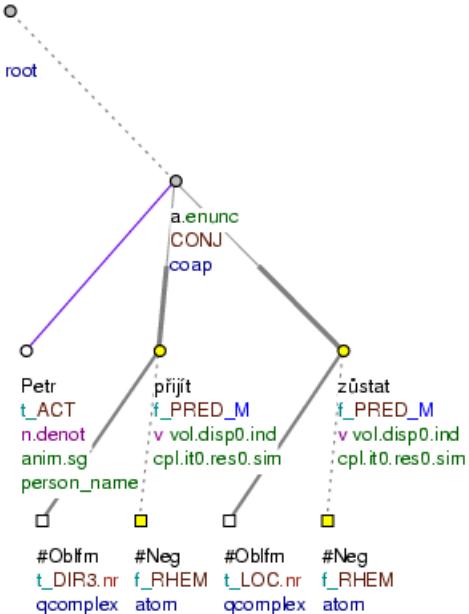
**Layering of modal meanings within one modal predicate.** In cases of layering of modal meanings, the first modal verb in the sequence is always represented by a separate node with the t-lemma corresponding to the infinitive of this modal verb. In the deontic modality grammateeme the value decl is entered at that node (the modal meaning of the first modal verb is not represented by a grammateeme, but by the lexical meaning of the node of the given modal verb). The second modal verb in the sequence and the infinitive of the full verb are represented by a single node (unless the infinitive is negated in which case they would be represented by separate nodes).

**Figure 7.2. Modal predicates**

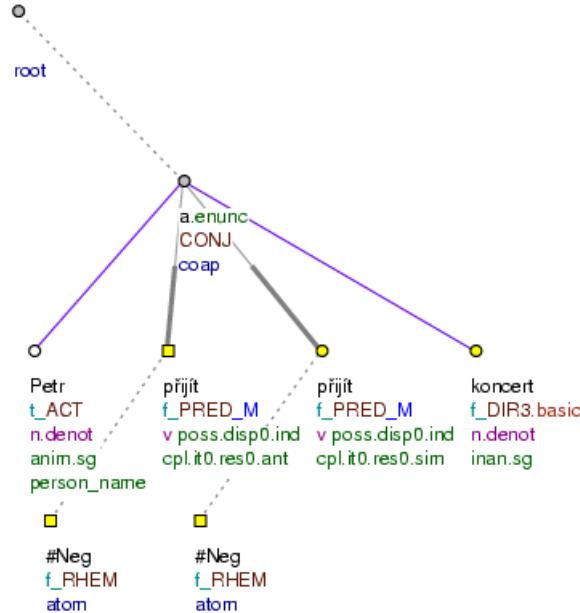
Petr chce přijít na koncert. (=lit. Petr wants to\_come to (the) concert.)


**Figure 7.3. Modal predicates**

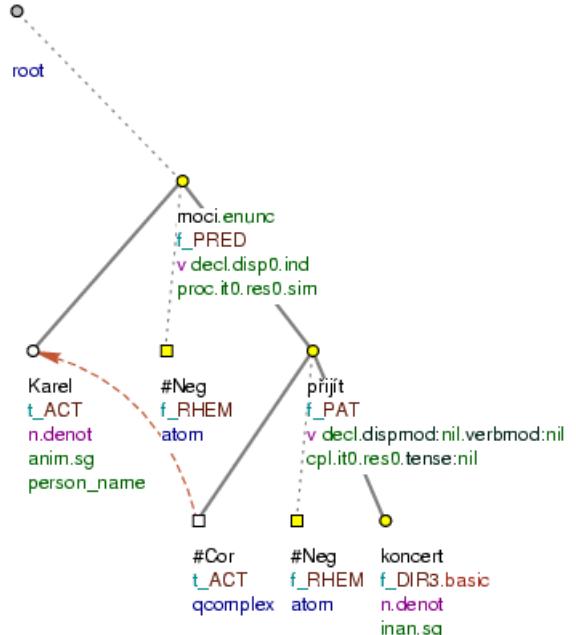
Petr nechce přijít a zůstat. (=lit. Petr does\_not\_want to\_come and stay.)


**Figure 7.4. Modal predicates**

Petr nemohl a nemůže přijít na koncert (=lit. Petr could\_not Karel nemůže nepřijít na koncert. (=lit. Karel cannot not\_come to (the) concert.)

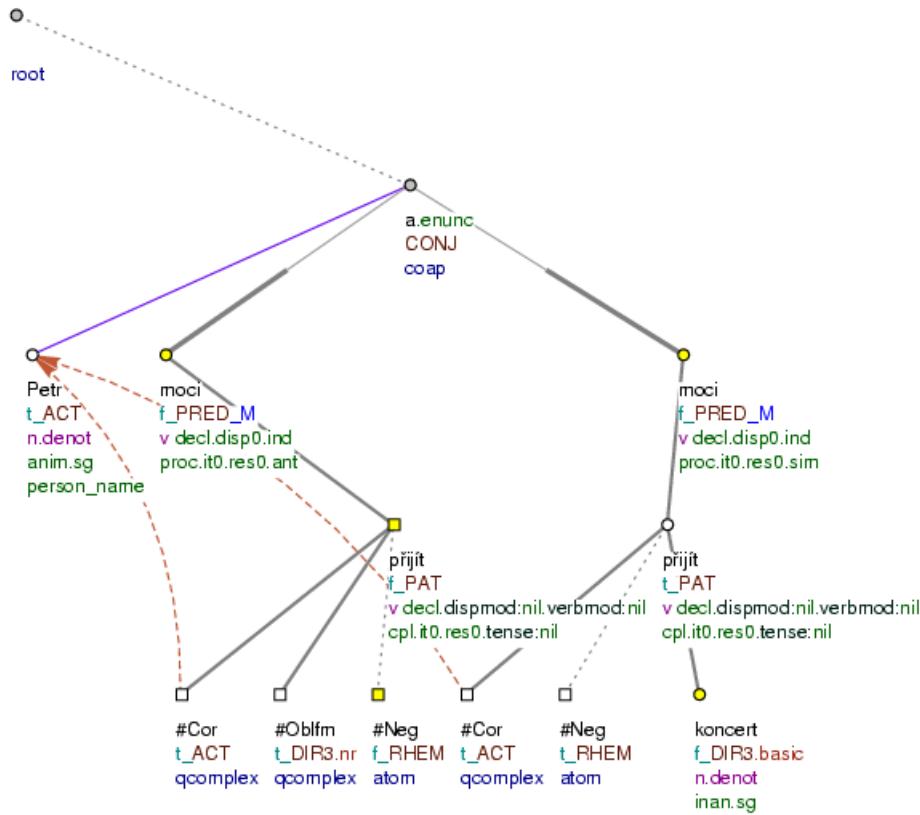

**Figure 7.5. Modal predicates**

Karel nemůže nepřijít na koncert (=lit. Karel cannot not\_come to (the) concert)

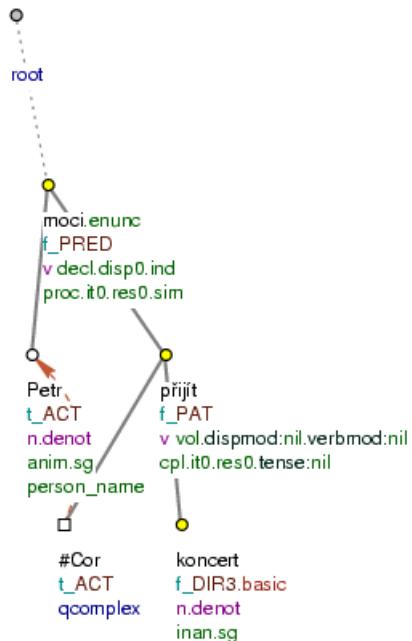


**Figure 7.6. Modal predicates**

Petr mohl a může nepřijít na koncert. (=lit. Petr could and can not\_to\_come to (the) concert.)


**Figure 7.7. Layering of modal meanings**

Petr může chít přijít na koncert. (=lit. Petr may want\_to\_come to (the) concert.)



### 7.1.1.2. Phase predicates

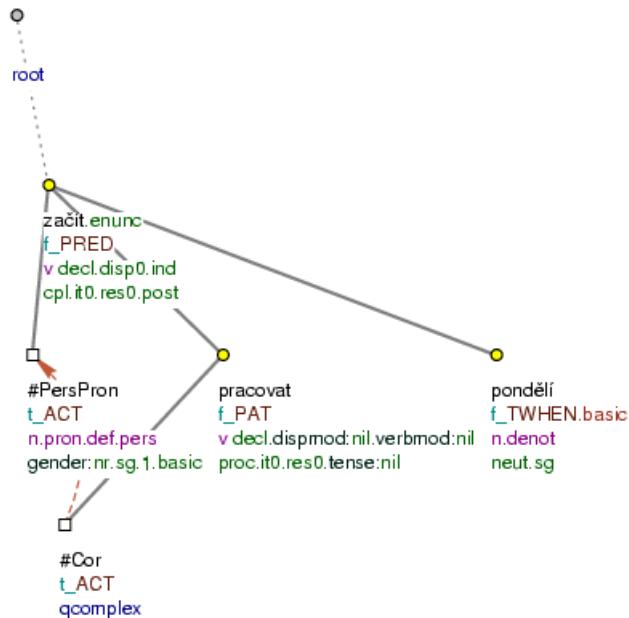
**Phase predicate** is a multi-word predicate consisting of a phase verb (*začít* (=begin), *zahájit* (=start), *přestat* (=stop), *skončit* (=finish) etc.), which, beside the grammatical meanings, refers to the phase of the event, and the infinitive of a full verb, which carries the main lexical meaning of the predicate as a whole.

No grammatemes have been established to represent the phase of an event. Phase predicates are therefore always represented by two nodes: a node representing the phase verb and a node representing the infinitive of the full verb. The node for the full verb has the functor PAT and is dependent on the node for the phase verb. For example:

*Začnu.PRED pracovat.PAT v pondělí.* (=I will start work on Monday.) Fig. 7.8

**Figure 7.8. Phase predicate**

*Začnu pracovat v pondělí.* (=lit. (I) will\_start work on Monday.)



### 7.1.1.3. Quasi-modal and quasi-phase predicates

**Quasi-modal predicates** and **quasi-phase predicates** are defined as combinations of a quasi-modal or quasi-phase verb (expressing the grammatical meanings and modal or phase meaning) with the infinitive of a full verb, carrying the main lexical meaning of the predicate. A quasi-modal or quasi-phase predicate is formed by:

- a **quasi-modal or quasi-phase verb**, i.e. by a multi-word synonym of a modal verb (semantically empty verb + noun or adverb with a modal meaning). In addition to the modality, also a phasal meaning can be present with quasi-modal verbs.

Depending on the means of expression of the verbal and non-verbal components, three groups of quasi-modal and quasi-phase verbs are identified, distinguished in the annotation rather due to the gradual development of the annotation rules than to satisfy a need to distinguish the respective groups; see Table 7.1, “**Representing quasi-modal and quasi-phase verbs**”.

- the **infinitive of a full verb**. The infinitive of a full verb may also be nominalised and the main lexical meaning of the predicate is then expressed by a noun (frequently in a prepositional case). Cf:

*začít pracovat* (=to start working) → *dostat chut' pracovat* (=to get an appetite to work) → *dostat chut' k práci* (=to get an appetite for work)

The infinitive of the full verb (or its nominalization) is usually a valency modification of the non-verbal part of the quasi-modal or quasi-phase verb. The fact that the three components form a single predicate is not represented in any way. The infinitive is not assigned any deontic modality grammeme.

**Table 7.1. Representing quasi-modal and quasi-phase verbs**

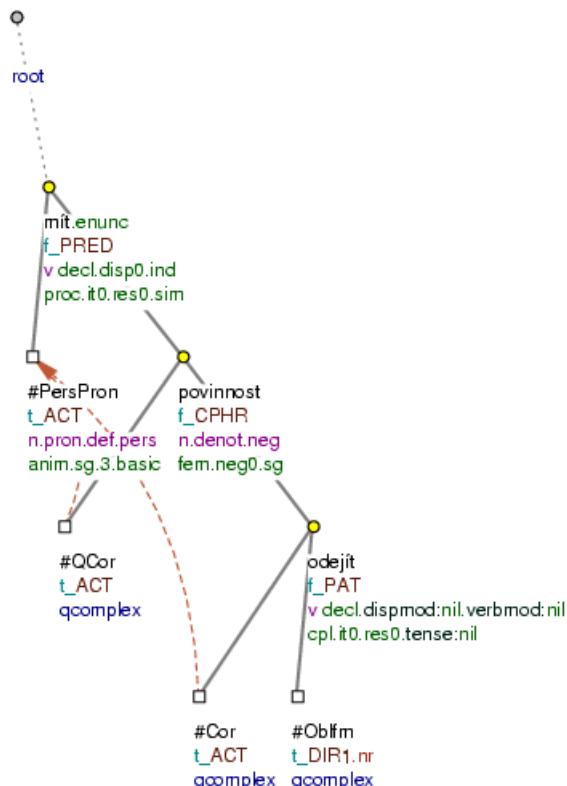
Definition of the group	Representation	Examples
verb (not <i>být</i> ) + noun	following the same rules as in the case of complex predicates (see Section 7.1.1.4, “Complex predicates”)	<i>mít plán</i> (=have a plan) <i>dát se do práce</i> (=start work) <i>dostat nápad</i> (=get an idea)
<i>být</i> + adjective or noun	following the same rules as in the case of verbonominal predicates (see Section 7.1.1.5, “Verbonominal predicates (the copula <i>být</i> )”)	<i>být schopen</i> (=be able) <i>být nutné</i> (=be necessary) <i>být povinností</i> (=be obliged)
<i>být</i> + predicative adverb	the CPHR functor is used (it is assigned to the predicative adverb)	<i>být nutno</i> (=be necessary) <i>být třeba</i> (=be necessary)

Example:

*Má povinnost.CPHR odejít.* (=He has an obligation to leave.) Fig. 7.9

**Figure 7.9. Quasi-modal predicate**

*Má povinnost odejít.* (=lit. (He) has (an) obligation to leave.)



#### 7.1.1.4. Complex predicates

**Complex predicates** are multi-word predicates consisting of a semantically empty verb which expresses the grammatical meanings in the sentence, and a noun (frequently denoting an event or a state), which carries the main lexical meaning of the entire predicate. A complex predicate forms a single multi-word lexical unit for which an appropriate synonymous expression can usually be found in the form of a one-word predicate. Cf.:

- hovořit* (=talk) → *vést rozhovor* (=have a talk/conversation)
- plánovat* (=to plan) → *mít plán* (=to have a plan)
- nárokovat si* (=to claim) → *činit si nárok* (=to make a claim)
- připravovat se* (=to prepare (oneself)) → *dělat přípravy* (=to make preparations)
- omezit* (=to limit) → *provést omezení* (=to impose a limitation)
- zajímat se* (=to be interested) → *projevit zájem* (=to show interest)

The existence of an adequate synonymous one-word expression is not however necessary for considering a certain collocation of a semantically empty verb and a meaning-bearing noun as a complex predicate.

The following parts of complex predicates are distinguished:

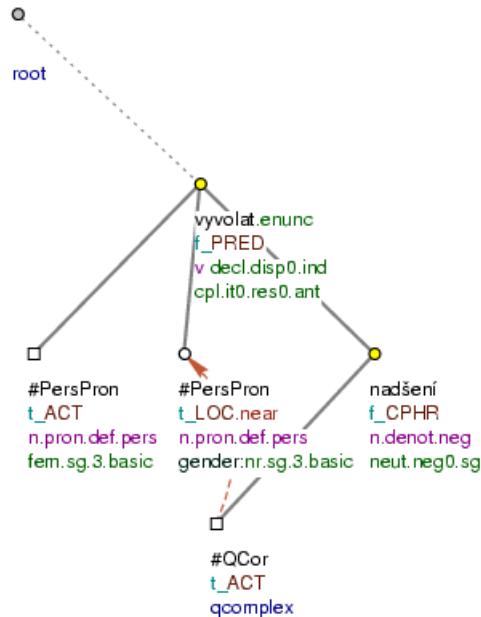
- **the verbal part of a complex predicate.**
- **the nominal part of a complex predicate**, i.e. the dependent noun, which carries the main lexical meaning of the entire predicate.

**Basic annotation rules for complex predicates.** A complex predicate is represented in the tectogrammatical tree by two nodes: by a node representing the verbal component of the complex predicate and by a node representing the nominal component of the complex predicate. The node representing the verbal part is assigned a functor according to the function of the entire complex predicate in the sentence structure. The nominal part is assigned the CPHR functor, which signals that it is a part of a multi-word predicate. The node is represented as an immediate daughter of the node for the verbal component. For example:

*Vyvolala u něho nadšení.CPHR (=She aroused his enthusiasm.)* Fig. 7.10

**Figure 7.10. Complex predicate**

*Vyvolala u něho nadšení. (=lit. (She) aroused at him enthusiasm.)*



#### 7.1.1.4.1. Quasi-control with complex predicates

The fact that a complex predicate is semantically a single unit has the consequence that certain valency modifications of the verbal and nominal parts are identical in reference. In other words, the nominal and verbal parts of complex predicates share certain valency modifications. This is called **quasi-control**, a specific type of grammatical coreference (see Section 9.2, “Grammatical coreference”).

**Representing quasi-control with complex predicates.** This referentially identical (shared) valency modification is expressed only once at the surface level. The annotators must first of all determine whether the given valency modification belongs to the verbal or the nominal part of the complex predicate. If the valency modification occurs in the valency frames of both parts of the complex predicate but the form corresponds only to one of them, it is represented by a node dependent on the part that determines the form. However, those cases are problematic where the expressed valency modification occurs in the same form in the valency frames of both parts of the complex predicate. A simple convention has been accepted for these cases: the shared modification is represented by a node dependent on the verbal part.

In place of the other (shared) valency modification that is omitted at the surface level (as a rule it is a valency modification of the nominal part of the complex predicate), a new node is added to the tectogrammatical tree, with the t-lemma substitute #QCor. In addition to the special t-lemma, referential identity is also indicated by the

grammatical co-reference relation leading from this added node to the node for the second shared valency modification. If the shared valency modification is not expressed at the surface level at all, it is represented in the nominal component of the complex predicate by a newly established node with the t-lemma #QCor; in the verbal component, the newly established node for this modification has a t-lemma substitute depending on the type of elision (for the rules see Section 6.6.1.2, “Ellipsis of a dependent meaning unit”); thus: #Gen, #PersPron, possibly #Unsp.

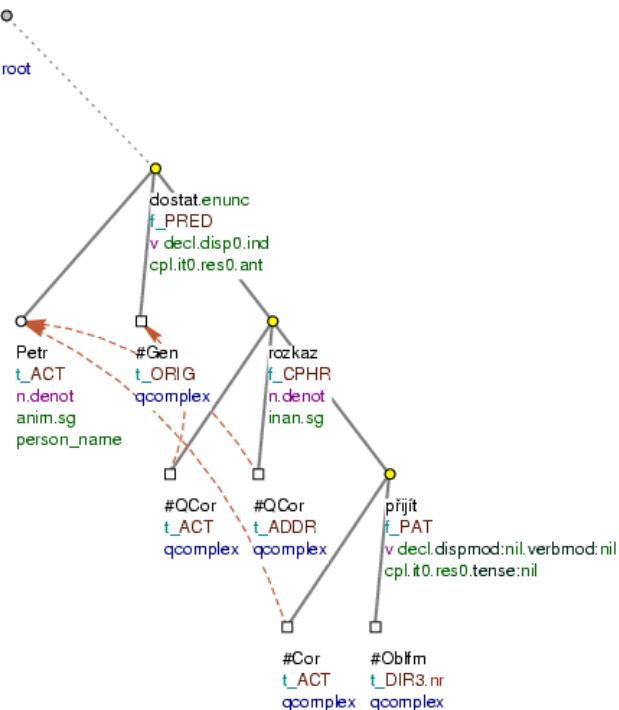
Examples:

*Petr dostał {#Gen.ORIG} {#QCor.ACT} rozkaz pójść. (=Petr got the order to come.)* Fig. 7.11

*Pavel.ACT dal Petrovi.ADDR {#QCor.ACT} {#QCor.ADDR} radu.CPHR (=Pavel gave Petr a piece of advice.)*  
Fig. 7.12

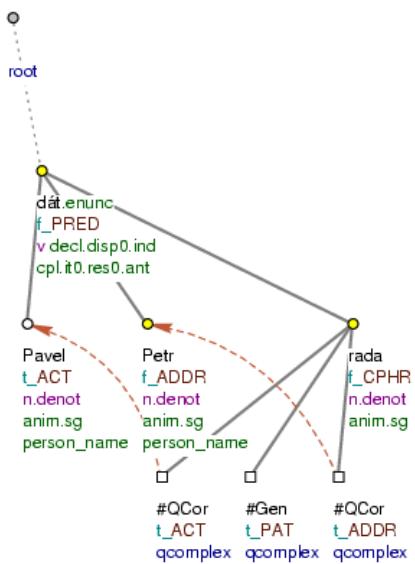
**Figure 7.11. Quasi-control with complex predicates**

*Petr dostał rozkaz pójść. (=lit. Petr got order to\_come.)*



**Figure 7.12. Quasi-control with complex predicates**

*Pavel dal Petrovi radu.* (=lit. Pavel gave Petr advice.)



### 7.1.1.5. Verbonominal predicates (the copula *být*)

**Verbonominal predicate** is a collocation of the verb *být* (=be) (whose function is primarily to carry the grammatical meanings) and another word, carrying the main lexical meaning of the predicate. Verbonominal predicates can express a wide range of meanings: identity of the Actor and Patient, description, classification or quantification. With verbonominal predicates, we distinguish:

- **the verbal part**, i.e. the copula *být*.
- **the non-verbal part**, i.e. the semantic adjective or noun in the nominative or instrumental. The non-verbal part can also be expressed by a noun in the genitive, infinitive, dependent clause, adverb or interjection.

**Representing verbonominal predicates in the tectogrammatical trees.** Verbonominal predicates are represented by two nodes: one for the verbal and one for the non-verbal part. The node representing the verbal part is assigned a functor depending on the function of the whole predicate in the sentence structure. The node for the non-verbal part is assigned the PAT functor (which stands for a wide range of meanings) and is an immediate daughter of the node for the verbal part. The fact that the two parts form a single predicate is not indicated in any way (except for the valency frame of the verb *být*). For example:

*Kočka je savec.PAT* (=The cat is a mammal.) Fig. 7.13

*Jist je obřad.PAT* (=Eating is a ritual.)

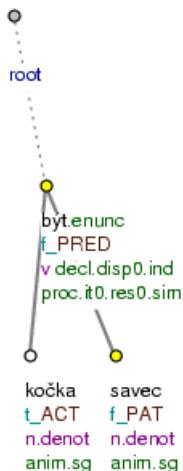
*Jirka je hodný.PAT* (=Jirka is kind.)

*Dětí je pět.PAT* (=There are five children.)

*To je fuk.PAT* (=It doesn't matter.)

**Figure 7.13. Verbonominal predicate**

*Kočka je savec.* (=lit. Cat is mammal.)



## 7.1.2. Idioms

**Idiom (idiomatic expression)** is a collocation of two or more words with a fixed lexical content that form a single lexical unit, which has a metaphorical meaning (as a whole; the meaning cannot be computed from its parts). Idiomatic expressions have two parts:

- **the governing element**, i.e. the syntactically governing part of the idiom.
- **the dependent part**, i.e. all other expressions that are part of the idiom.

Further, we distinguish:

- a. **non-verbal idioms**, i.e. idioms whose governing node is not a verb with a regular paradigm. The governing node of a non-verbal idiom can be a verb, but this verb never has a complete paradigm, it is always a more or less fixed verb form (e.g.: *stúj co stúj* (=at any cost); *chtě nechte* (=willy nilly)). If it is not possible to determine unequivocally the governing element of the idiom, it is the first expression in the sequence.
- b. **verbal idioms**, i.e. idioms the governing node of which is a verb with a full paradigm.

**Representing idioms in the tectogrammatical trees.** Idiomatic expressions are represented by two nodes: a mother node and its immediate daughter. The mother node is the governing node of the idiom and has a functor depending on the position of the whole idiom in the structure. The dependent part is represented by a single node with the functor DPHR (node type = dphr), which indicates that the node forms an idiom with its mother node. The t-lemma with the functor DPHR consists of all the dependent parts of the idiom (incl. prepositions), linked by the underscore character, the order being identical to the surface word order (see also Section 4.3, “Multi-word t-lemma”).

Examples of non-verbal idioms:

*Hledá investici šitou.COMPL na míru.DPHR* (=..made to measure) Fig. 7.14

*Zavřeli mě pro nic.CAUS za nic.DPHR* (=They arrested me for no reason.)

*Široko.LOC daleko.DPHR nebylo vidět žádnou policii.* (=For miles around, there was no police to be seen.)

Examples of verbal idioms:

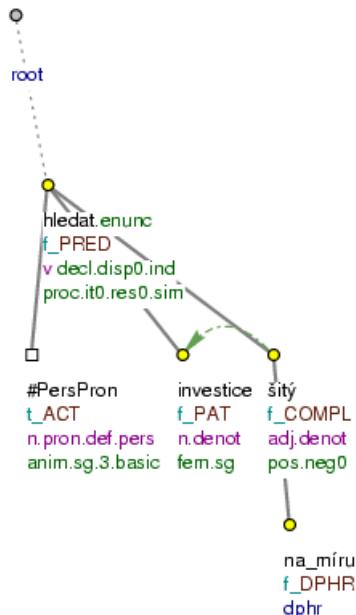
*Házeli.PRED nám klacky pod nohy.DPHR* (=They put obstacles in his way.) Fig. 7.15

*Dával.PRED mi neustále najevlo.DPHR svou převahu.* (=He made me feel his superiority.)

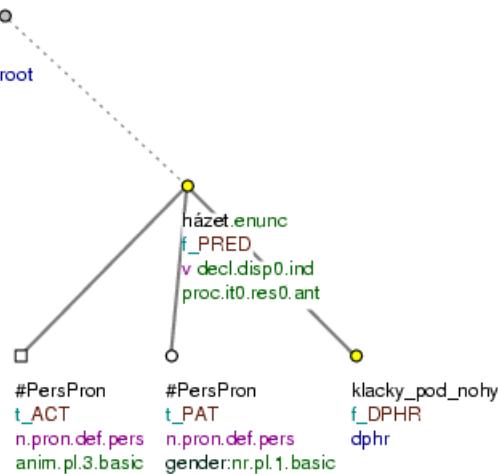
*Běhal.PRED mu mráz po zádech.DPHR* (=It was giving him the creeps.)

**Figure 7.14. Non-verbal idiom**

*Hledá investici šitou na míru. (=lit. (He) is \_looking\_for Házeli nám klacky pod nohy. (=lit. (They) were \_throwing us bows under legs.)*


**Figure 7.15. Verbal idiom**

*Házel nám klacky pod nohy. (=lit. (They) were \_throwing us bows under legs.)*



## 7.2. Numbers and numerals

There are several groups of **numerals** and they are distinguished in the annotation:

- numerals in the role of an attribute (RSTR):** numerals in the combination with the counted object (except for the numerals with the function of a container). The governing node is the node for the counted noun; the numeral depends on it. The node for the numeral is assigned the RSTR functor. For example:

*Mám pět.RSTR domů a tři.RSTR auta. (=I've got five houses and three cars.) Fig. 7.16*

*náš druhý.RSTR nejlepší hráč (=our second best player)*

- cardinal numerals without the counted object.** The numeral is considered a syntactic noun. The node for the numeral can get different functors. For example:

*Zvolili tři.PAT z pěti místopředsedů. (=They elected three out of five vice-chairs.) Fig. 7.20*

*Kolik.PAT mi dáš? (=How much will you give me?)*

- numerals with the function of a “container”:** numeral expressions: *milion* (=million), (and other ending with -ion), *miliarda* (=billion), *polovina* (=half), *polovice* (=half), *půl(e)* (=half), *třetina* (=one third), *čtvrt* (=quarter), *čtvrtina* (=quarter), *tisícina* (=one thousandth), *tucet* (=dozen), *veletucet* (=twelve dozen), *kopa* (=heap), *řada* (=row), *spousta* (=a lot), *hromada* (=pile), *zástrup* (=crowd), *dav* (=crowd), *dvojice* (=couple), *trojice* (=trio), *sto* (=hundred), *tisíc* (=thousand), *trocha/u* (=a bit) with the counted object. The governing node is the node for the numeral-container, the node for the counted object depends on it. The node for the counted object is assigned the functor MAT. For example:

*Žije tu jeden.RSTR milion lidí.MAT (=One million people live here.) Fig. 7.17*

- numerals used as “labels”:** definite cardinal numerals (mostly written in digits) that are used to number/label objects (e.g. phone/fax numbers, house numbers, postal codes, product codes, numbers in product names). The node for the numeral is usually assigned the RSTR functor (in addresses). For example:

*Nový Golf 500.RSTR už je na trhu.* (=New Golf 500 is already on the market.) Fig. 7.18

- e. **numerals with adverbial meanings:** numerals: *pětkrát* (=five times), *několikrát* (=several times), *jednou* (=once), *podruhé* (=for the second time) etc. The node for the numeral is assigned an adjunct functor corresponding to the position in the structure. For example:

*Výhráli jsme jen dvakrát.THO* (=We only won twice.) Fig. 7.19

**NB!** Adjectival numerals and numeral expressions *hodně* (=a lot), *více* (=more), *méně* (=less), *mnoho* (=much/many), *málo* (=little/few), *stejně* (=the same), *plno* (=plenty), *dost* (=enough) used separately, without a counted object, are considered syntactic adjectives (attribute function). A new node for the governing noun is added to the structure, following the rules in Section 6.6.1.1, “Ellipsis of a governing meaning unit”. For example:

*Třetí.RSTR {#EmpNoun.PAT} už jsme nestihli.* (=We missed the third one.) Fig. 7.21

*Má hodně.RSTR {#EmpNoun.PAT}* (=He's got a lot.)

**Numerals written in digits.** For the annotation of numerals written in digits essentially the same rules apply. If a complex numeral expression is written in digits, it is represented by a single node; e.g.:

*Mám 38 234.RSTR korun.PAT* (=I've got 38 234 crowns.)

**Numeral expression of the type “sto čtyřicet tisíc lidí”.** Complex numeral expressions are not assigned any inner structure

- a. if the numeral expression contains one or more numerals with the function of a container, the one with the highest value is the governing node of the whole expression. All other parts of the numeral expression are dependent on this node (as sisters with respect to each other) and they have the functor RSTR. For example:

*sto.RSTR čtyřicet.RSTR tisíc židů.MAT* (=one hundred and forty thousand Jews.) Fig. 7.22

- b. if there is no numeral with the function of a container present, the governing node is the node for the counted object and all parts of the numeral expression depend on it (as sisters) and have the RSTR functor. For example:

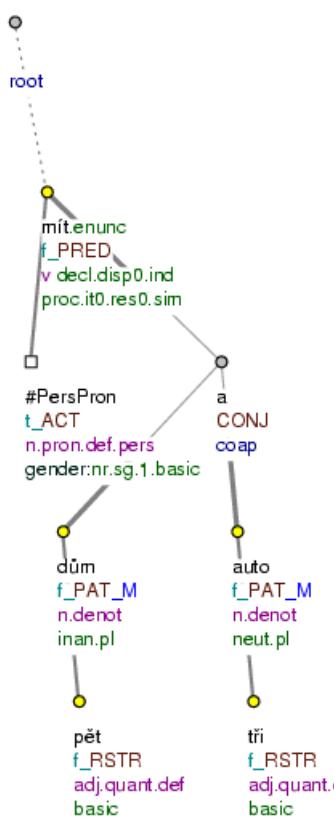
*třicet.RSTR osm.RSTR žáků* (=thirty-eight pupils)

**Numeral expression of the type “tyč dlouhá 2 m 10 cm 4 mm”.** Combinations of numerals and measure units are not assigned any inner structure. The measure unit with the highest value is considered the governing node of the whole collocation. The nodes for the other measure units depend on this node (as sisters) and are assigned the RSTR functor; the nodes for the numerals depend on their respective measure units and also have the RSTR functor. For example:

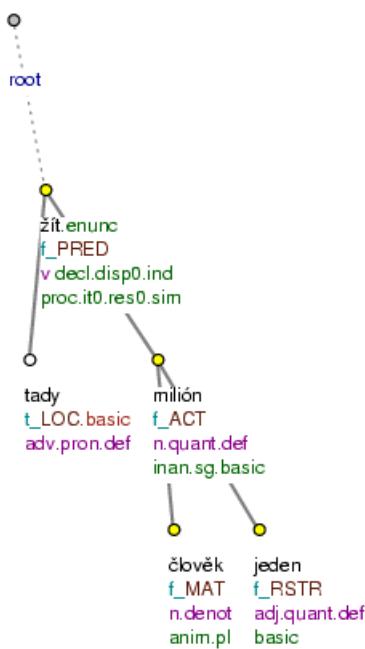
*Tyč je dlouhá 2 m.EXT 10 cm.RSTR 4 mm.RSTR* (=The pole is 2m 10cm and 4mm long.) Fig. 7.23

**Figure 7.16. Numeral in the attribute position (RSTR)**

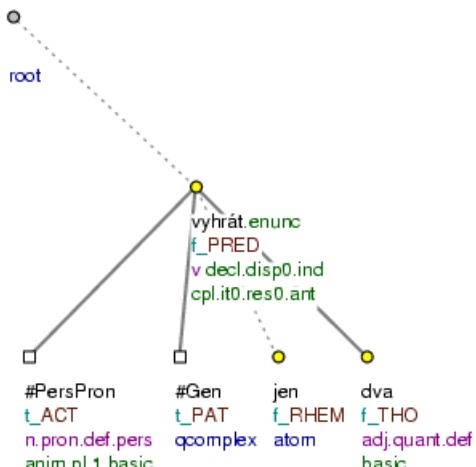
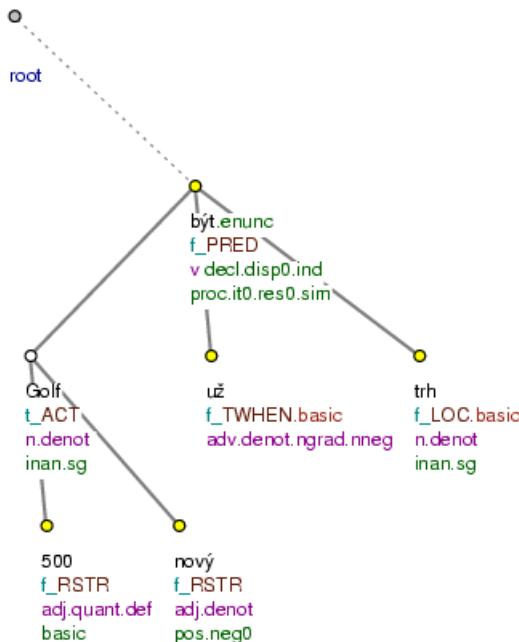
*Mám pět domů a tři auta.* (=lit. *(I) have five houses and three cars.*)


**Figure 7.17. Numeral with the function of a “container”**

*Žije tu jeden milión lidí.* (=lit. *Lives here one million (of) people.*)

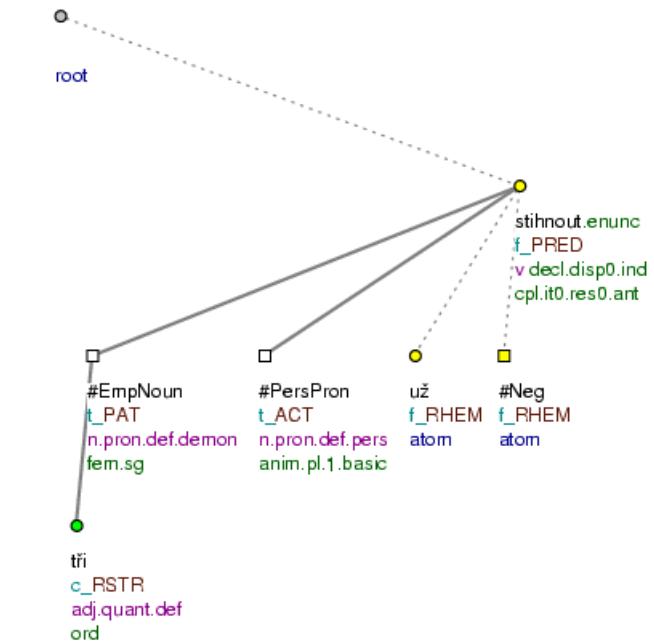
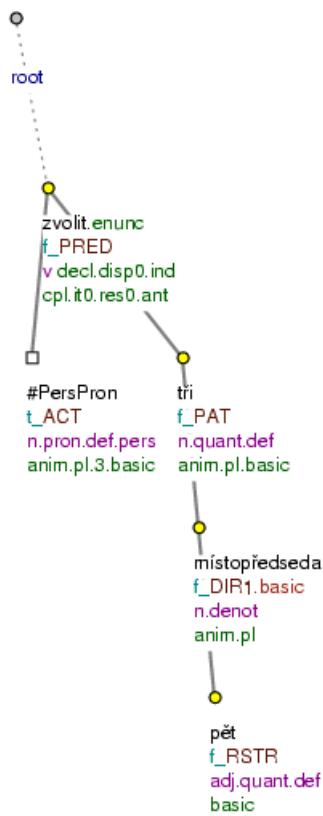

**Figure 7.18. Numeral with the function of a “label”**

*Nový Golf 500 už je na trhu.* (=lit. *New Golf 500 already is on market.*)



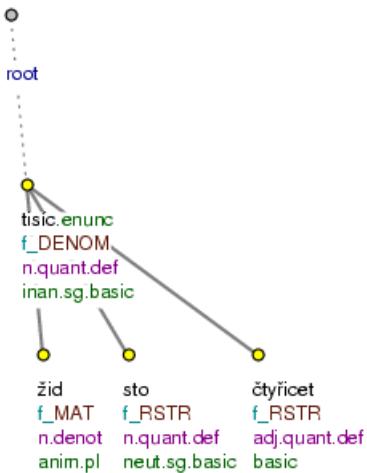
**Figure 7.20. Cardinal numerals without the counted object**

Zvolili tři z pěti místopředsedů. (=lit. (They) elected three out\_of five vice-chairs.)



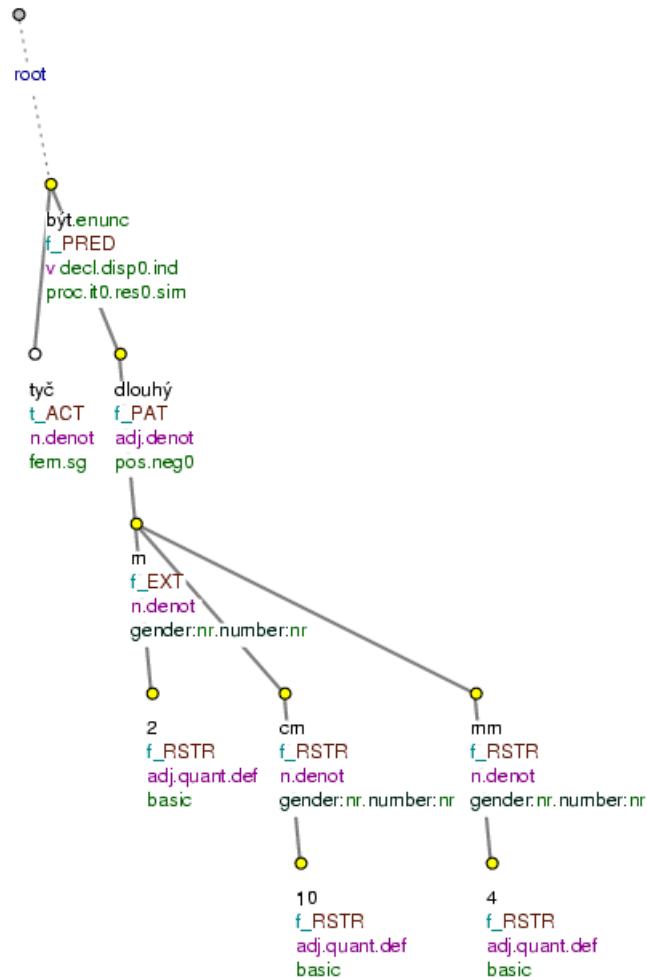
**Figure 7.22. Complex numeral expressions**

sto čtyřicet tisíc židů (=lit. hundred forty thousand Jews.)



**Figure 7.23. Complex numeral expressions**

Tyč je dlouhá 2 m 10 cm 4 mm. (=lit. Pole is long 2 m 10 cm 4 mm.)



## 7.3. Identifying expressions

**Identifying expressions** are expressions used for identification (proper names, and titles, meta-language use).

Identifying expressions are divided into two basic groups and these are distinguished in the annotation

- identifying expressions with the governing element inflected regularly.** The effective root of an identifying expression (the inflected noun) has a functor according its position in the structure. Nodes dependent on the effective root are analyzed according to the standard annotation rules. The root of the modifier of an identifying expression is always an immediate daughter node of the root of the identification structure. For example:

Čtu Babičku.EFF (=I'm reading Babička.)

Organizace.DENOM Spojených.RSTR národů.APP pro výchovu.BEN , vědu.BEN a.CONJ kulturu.BEN (=The United Nations Organization for Education, Science and Culture) Fig. 7.24

Jeho.APP Máj.ACT je otrhaný. (=His Máj is all torn.) Fig. 7.25

- other identifying expressions:** identifying expressions that do not conform to the criteria for being included in the first group (their governing node is not a regularly inflected noun). These are represented as identification structures (see Section 7.3.1, “Identification structure”). For example:

kniha Obsluhoval jsem.ID anglického.RSTR krále.PAT (=the book Obsluhoval jsem anglického krále)

**Nominative of identity:** The nominative of identity is a modification of a (common) noun introducing a proper noun, title, an expression used metalinguistically or an expression quoted word for word. Both groups of identifying expressions can occupy the position of the nominative of identity. When in the position of the nominative of identity, the identifying expression is always represented as an identification structure (see Section 7.3.1, “Identification structure”). For example:

*hory Krkonoše.ID (=the Krkonoše mountains)*

*symfonická báseň Z českých.RSTR luhů.ID a hájů.ID (=the symphonic poem Z českých luhů a hájů)*

**Explicative genitive:** an expression in the genitive modifying a common noun while the following transformation is applicable: *trest smrti* (=death penalty, lit. penalty of death) → *smrt je (druhem) trest(u)* (=death is a kind of punishment). This specific type of identifying expressions is analyzed with the help of an identification structure (see Section 7.3.1, “Identification structure”). For example:

*pojem času.ID (=the concept of time)*

*otázka laickosti.ID (=the laicism issue)*

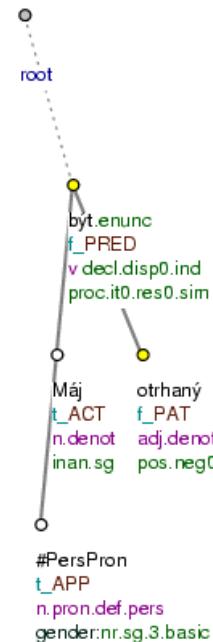
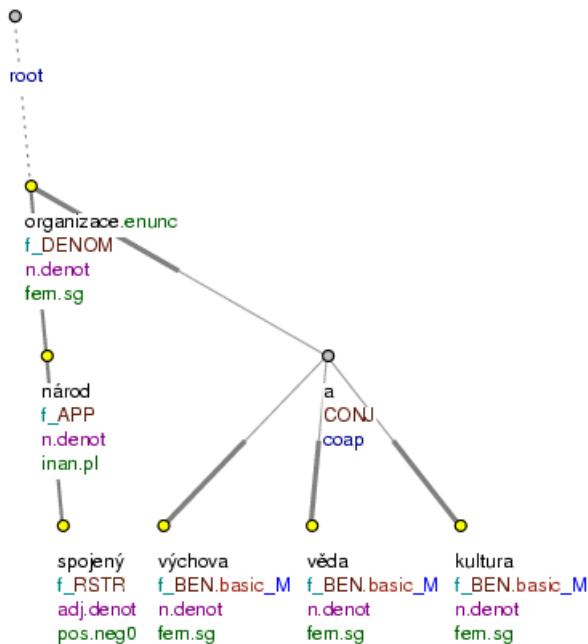
*osoba V.Klause.ID (=lit. person (of) VK)*

**Proper names of people (the `is_name_of_person` attribute).** At all nodes representing expressions which are constituents of proper names of people (nodes representing first name or surname) the value 1 is entered in the attribute `is_name_of_person` (see Section A.2.13, “`is_name_of_person`”).

**Figure 7.24. Identifying expression with the inflected governing noun**

**Figure 7.25. Identifying expression with the inflected governing noun**

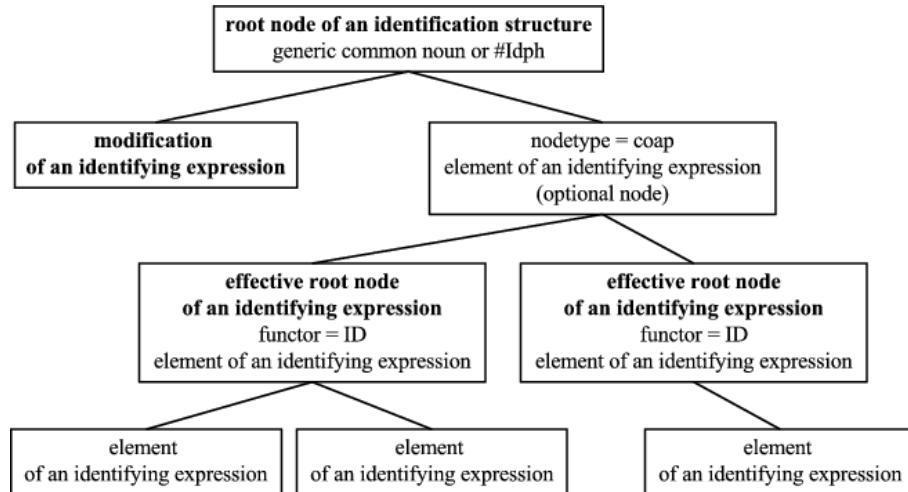
*Organizace Spojených národů pro výchovu, vědu a kulturu.* (=lit. Organization (of) United Nations for Education, Science and Culture)



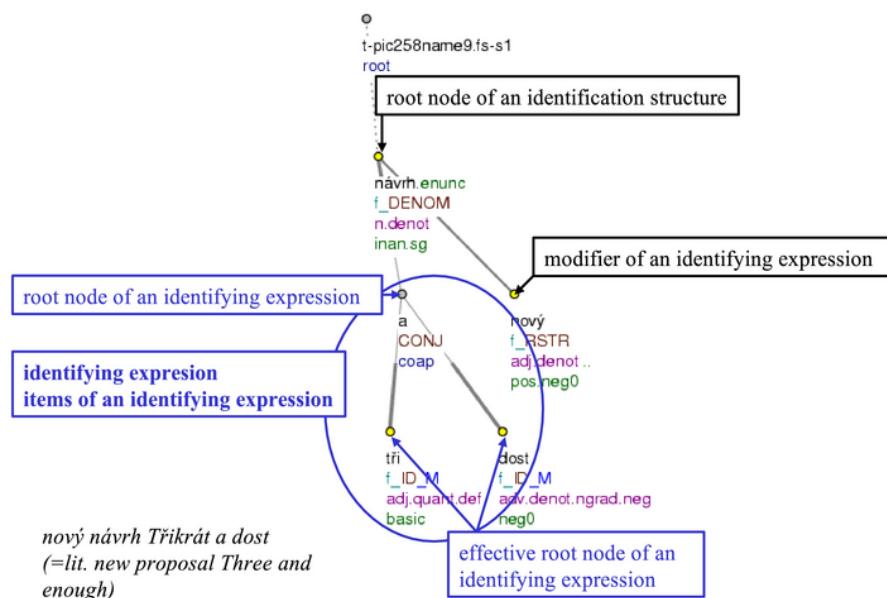
### 7.3.1. Identification structure

Identification structure is illustrated in Fig. 7.26 and Fig. 7.27.

**Figure 7.26. Identification structure I**



**Figure 7.27. Identification structure II**



**Root node of an identification structure.** The root of an identification structure is the node representing the common noun. If the identifying expression occurs at the surface level without its common noun, a new node is added to the tectogrammatical tree in the position of the root of the identification structure, with the t-lemma substitute #Idph. The root node of an identification structure is assigned a functor depending on its position in the sentence structure.

**Effective root node of an identifying expression.** The effective root nodes of the identifying expression, which all have the functor ID are dependent on the root of the identification structure. The effective roots of an identifying expression are nodes representing expressed governing nodes of the identifying expression not dependent on any other node. This means that a possible case of ellipsis of the governing element is not represented with identifying expressions.

The effective root of an identifying expression is usually identical to **the root of an identifying expression**. The effective root nodes of an identifying expression are not identical with the root of the identifying expression only in cases where the identifying expression has more governing nodes, which are paratactically connected. Identifying expressions can also have more roots: in cases in which the identifying expression has more governing elements (effective roots), which are not paratactically connected.

**Items of an identifying expression.** All nodes representing the individual words of an identifying expression are the items of the identifying expression. The structure of identifying expressions undergoes further analysis. Unless

stated otherwise, the rule is that the elements dependent on the effective root nodes of an identifying expression are annotated according to the standard annotation rules and their functor is assigned according to the nature of their dependency.

**Modifiers of identification structures.** Identifying expressions can be further modified (as a whole). The root of the modifier is always an immediate daughter node of the root of the identification structure.

Examples:

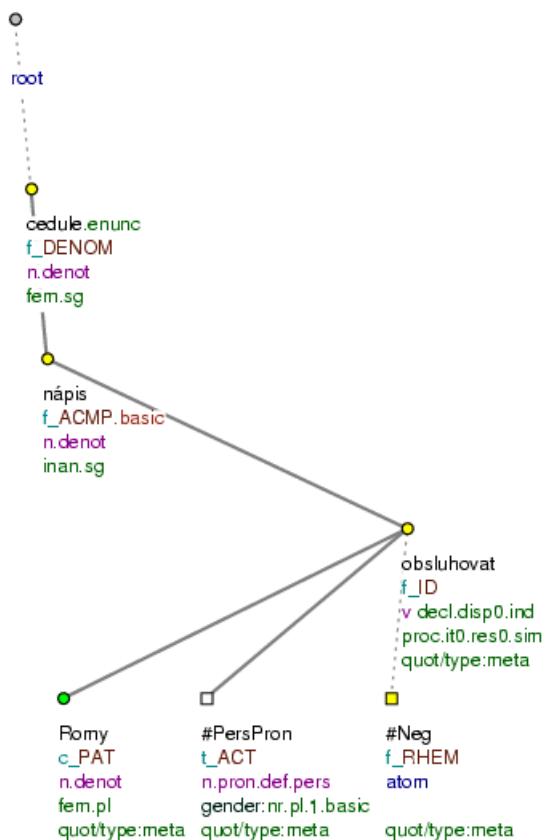
*cedule s nápisem.ACMP „Romy neobsluhujeme.ID .” (=the notice We do not serve Romanies) Fig. 7.28*

*{#Idph.ACT} V sobotu.ID v poledne.ID je hezký film. (=V sobotu v poledne is a good film.) Fig. 7.29*

*Jiráskovo {#Idph.DENOM} Proti všem.ID (=Jirásek's Proti všem) Fig. 7.30*

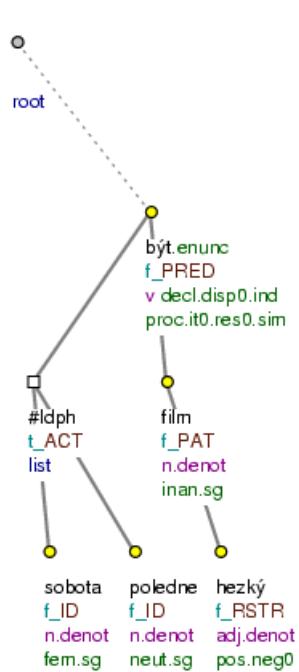
**Figure 7.28. Identification structure**

*cedule s nápisem „Romy neobsluhujeme.” (=lit. board with inscription Romanies (we) do\_not\_serve)*

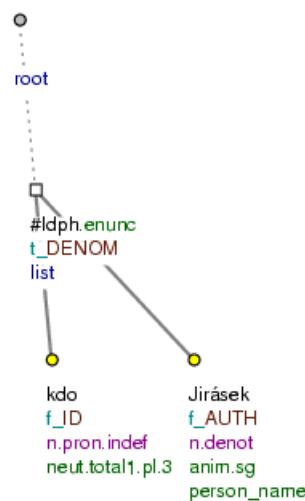


**Figure 7.29. Identification structure**

*V sobotu v poledne je hezký film.* (=lit. *On Saturday at noon is nice film.*)


**Figure 7.30. Identification structure**

*Jiráskovo Proti všem.* (=lit. *Jirásek's Proti všem*)



## 7.4. Foreign-language expressions

**Foreign-language expressions** those text segments that are written in a language other than Czech.

**List structure for foreign-language expressions.** Foreign-language expressions are represented as a list structure (see Fig. 7.31 and Fig. 7.32). **The root of a list structure for foreign-language expressions** is a newly established node with the t-lemma substitute #Forn (nodetype = list). The root of a list structure is assigned a functor corresponding to the function of the foreign-language expression (as a whole) in the sentence structure. All expressions of the foreign-language text segment as well as all punctuation marks and other symbols are represented by separate nodes in the tree as immediate daughters of the root of the list structure. Thus, they are represented as sister nodes, in the surface word order. These nodes (representing the **items of the list**) are assigned the FPHR functor (nodetype = fphr). The t-lemmas of these nodes are the unchanged forms of the foreign-language words.

**Modifiers of a list:** Foreign-language expressions can be further modified (as a whole) by a Czech expression. The root of the modifier is always an immediate daughter node of the root of the list structure.

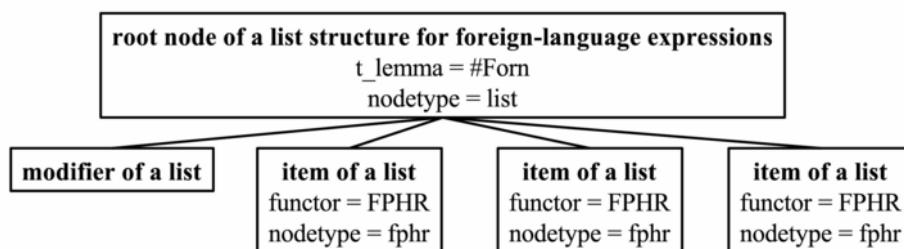
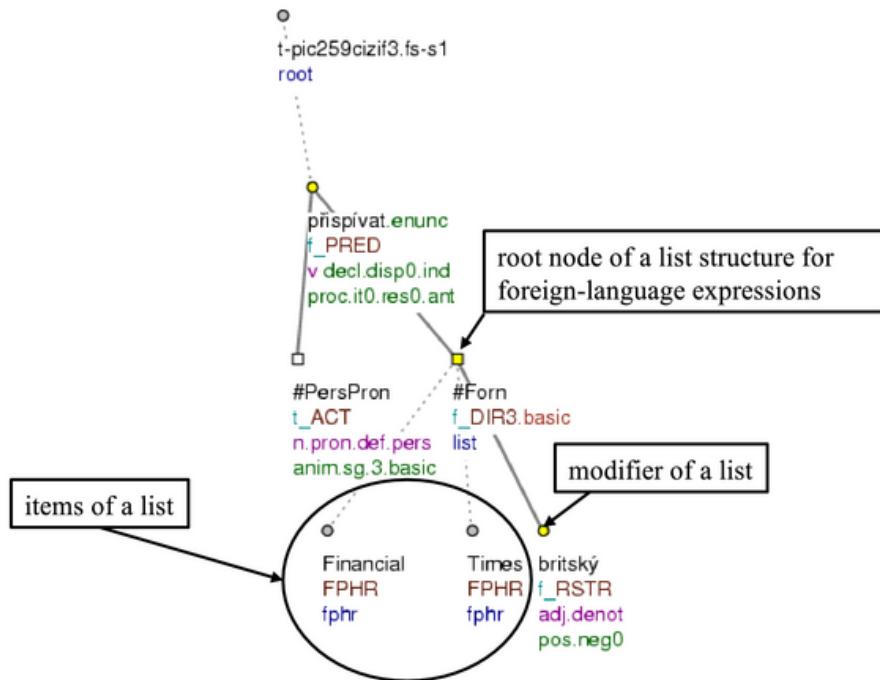
**Figure 7.31. List structure for foreign-language expressions I**


Figure 7.32. List structure for foreign-language expressions II



*Přispíval do britských Financial Times.  
 (=lit. (He) was\_contributing to British Financial Times.)*

Examples:

*firma {#Forn.ID} Eagle.FPHR Group.FPHR V.FPHR, [#Period.FPHR] A.FPHR, [#Period.FPHR] (=the EGVA company)* Fig. 7.33

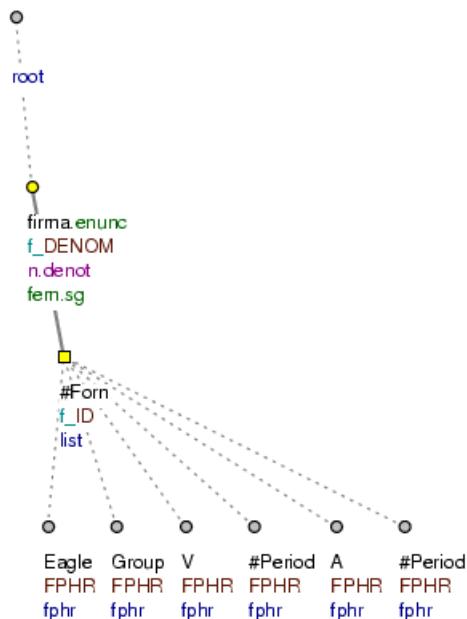
*Nenahraditelný je pro sledování {#Forn.PAT} cash.FPHR flow.FPHR (=important for cash flow monitoring)*

*v deniku {#Forn.ID} Financial.FPFR Times.FPFR (=in the FT newspaper)*

**NB!**Common loan words (even if indeclinable) and foreign-language words with Czech morphology are not represented by means of a list structure (*Vyprávěl o včerejší extra.RSTR show.PAT* (=He was telling us about the yesterday extra show)); the same holds for European proper names and one-word foreign-language titles in the nominative of identity (*časopis Times.ID* (=the Times magazine)).

**Figure 7.33. Foreign-language expression**

firma Eagle Group V. A. (=lit. company Eagle Group V. A.)



## 7.5. Direct speech

By **direct speech** we mean quoted spoken or written speech that is not formally integrated into the structure. The entire original utterance is quoted, including the original grammatical tenses and persons.

**Representing direct speech.** At the root node of the sub-tree representing direct speech, the value 1, is entered in the attribute `is_dsp_root` (see Section A.2.10, “`is_dsp_root`”), even if the direct speech is not graphically marked. Nodes for expressions that are part of graphically marked direct speech have the `quot/type` attribute filled with the value `dsp` (see Section 7.6, “Text segments marked by graphic symbol”).

Annotation rules for direct speech vary according to whether the direct speech is independent or introduced by a reporting clause:

- A. **independent direct speech.** If the direct speech is independent (i.e. not introduced by a reporting clause), the direct speech is annotated according to the rules described in Section 6.3, “Clauses (governing, dependent, verbal, non-verbal)”; i.e. it is represented as a verbal or non-verbal clause. For example:

„Máme.PRED několik set členů.“ [`is_dsp_root=1`] (=“We have several hundred members.”)

„Pardon.PARTL“ [`is_dsp_root=1`]

„Rozchod.DENOM!“ [`is_dsp_root=1`] (=“Dismiss!”)

- B. **direct speech introduced by a reporting clause.** The basic ways of direct speech representation are:

- a. **the direct speech is an argument of a verb or adjective.** If the direct speech is an argument of a verb (or adjective) in the reporting clause and unless the given valency position is already occupied by another modification, the effective root node of the direct speech has an appropriate argument functor and depends on the verb. For example:

*Jeho odpověď* byla: „Přijdu.PAT“ (=His answer was: I will come.)

*Řekl, že nepřijde.EFF* [`is_member=1`] : „*Určitě se nedostavím*.EFF [`is_member=1`] .“ (=He said he wouldn't come: I will certainly not arrive.) Fig. 7.34

*Trenér Sparty: „Nehráli jsme.PAT špatně.“* [`#Colon.PRED`] (=The coach: We didn't play too badly.) Fig. 7.35

- b. **direct speech is a modification of a noun** Direct speech may be an argument of a noun, or, if the noun introducing the direct speech has no valency, it is represented as a non-valency modification with the functor RSTR. For example:

*Nesdílím názor trenéra: „Hráč pro zranění nenastoupí.PAT .“ (=I don't share the view of the coach: The player will not play.)*

*Jeho slova „Závidím.RSTR mu.“ nás překvapila. (=His words I envy him surprised us.) Fig. 7.38*

If the direct speech cannot be interpreted as a modification of a word in the reporting clause, one of the following representations is used:

- c. if it is possible to add a verb of saying as an argument of the verb in the reporting clause (e.g. říkat), **a new node for this verb of saying is inserted into the tree**, a node with the t-lemma substitute #EmpVerb and an appropriate functor. The effective root node of the direct speech is dependent on the newly added node and it has the functor EFF. For example:

*Posadil se a začal{#EmpVerb.PAT} : „Nejdříve mi vysvětlete.EFF , co se stalo.“ (=He sat down and started: First of all tell me what happened.)*

- d. if it is possible to attach the direct speech to the reporting clause with the help of a transgressive (gerund) of a verb of saying (the subject of the verb in the reporting clause and the subject of the added transgressive are identical), then **a new node for the transgressive is added to the tree**, with the t-lemma substitute #EmpVerb and the functor COMPL (the node stands for expressions like řka (=saying) etc. The effective root node of the direct speech is dependent on the newly established node and it has the functor EFF. For example:

*Nepřesvědčivý výkon vysvětoval trenér {#EmpVerb.COMPL} : „Hráli jsme.EFF jen napůl.“ (=The coach was explaining the failure: We didn't play well enough.) Fig. 7.36*

*Vtrhl do dveří {#EmpVerb.COMPL} : „Kdy bude.EFF večeře?“ (=He rushed into the door: When will the dinner be ready?)*

- e. if the direct speech is not a modification of the reporting clause, we represent it as **paratactically connected to the reporting clause**. The root node of the paratactic structure is usually a node representing a colon. For example:

*V jednotce se lepší.PRED [is\_member = 1] nálada: „Porazíme.PRED [is\_member = 1] je.“ (=The mood is getting better: We will beat them.) Fig. 7.37*

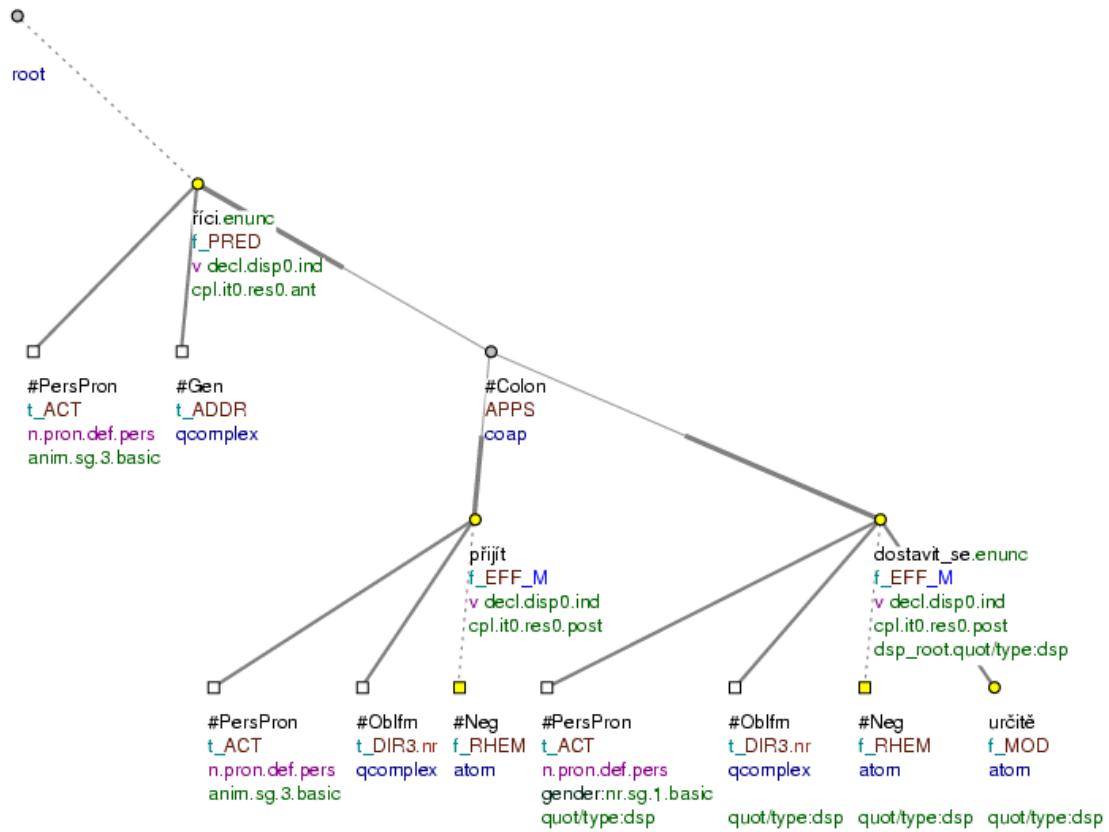
**NB!** In case the direct speech is a verbal or nominative clause, the argument (or another) functor is assigned directly to the effective root node of this nominative or verbal clause. If the direct speech is an (independent) vocative or interjectional clause, the effective root node of the direct speech is a newly established node for an empty verb (#EmpVerb) and the effective root node of the interjectional or vocative clause is represented as dependent on this newly established node. For example:

*Zavelel: „Rozchod.PAT !“ [is\_dsp\_root=1] (=He issued a command: Dismiss!)*

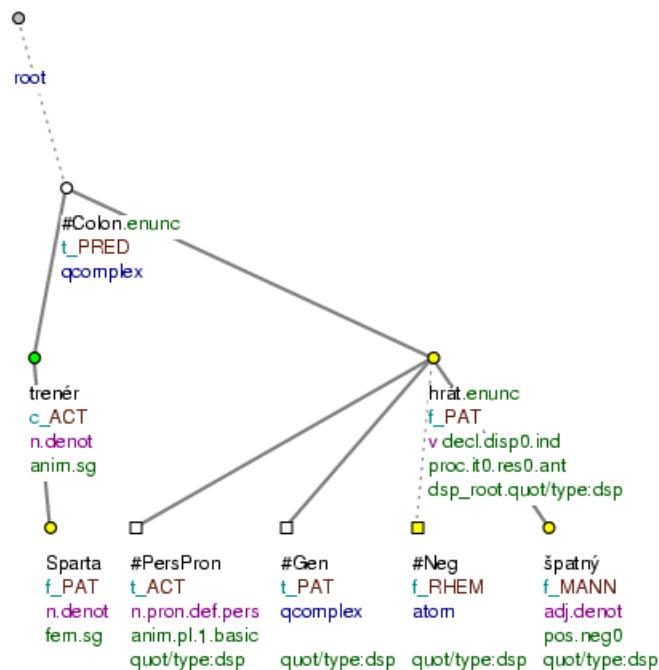
*Řekl: „ {#EmpVerb.EFF [is\_dsp\_root=1]} Pardon.PARTL “ (=He said: Pardon.)*

**Figure 7.34. Direct speech as an argument of the reporting verb**

*Řekl, že nepřijde: „Určitě se nedostavím.“ (=lit. (He) said that (he) will\_not\_come: Certainly (I) REFL will\_not\_arrive.)*

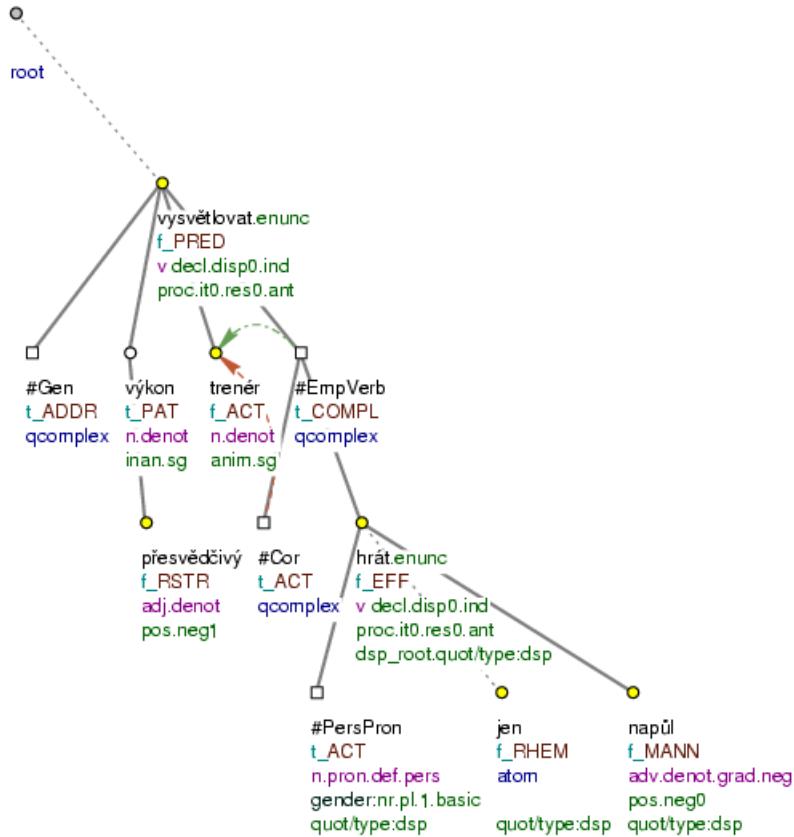

**Figure 7.35. Direct speech as an argument of the reporting verb**

*Trenér Sparty: „Nehráli jsme špatně.“ (=lit. Coach (of) Sparta: (We) did\_not\_play AUX badly.)*

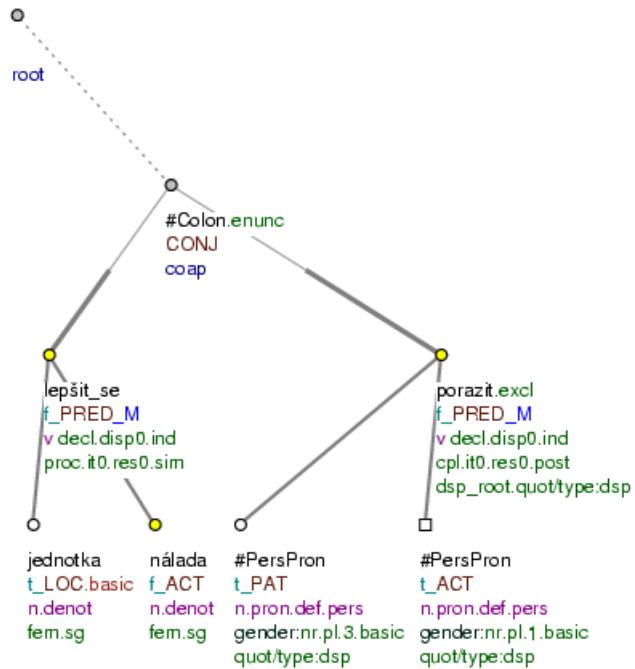


**Figure 7.36. Direct speech as an argument of the reporting verb**

*Nepřesvědčivý výkon vysvětloval trenér: „Hráli jsme jen napůl.“ (=Unconvincing performance.ACC explained coach.NOM: (We) played AUX only halfway.)*

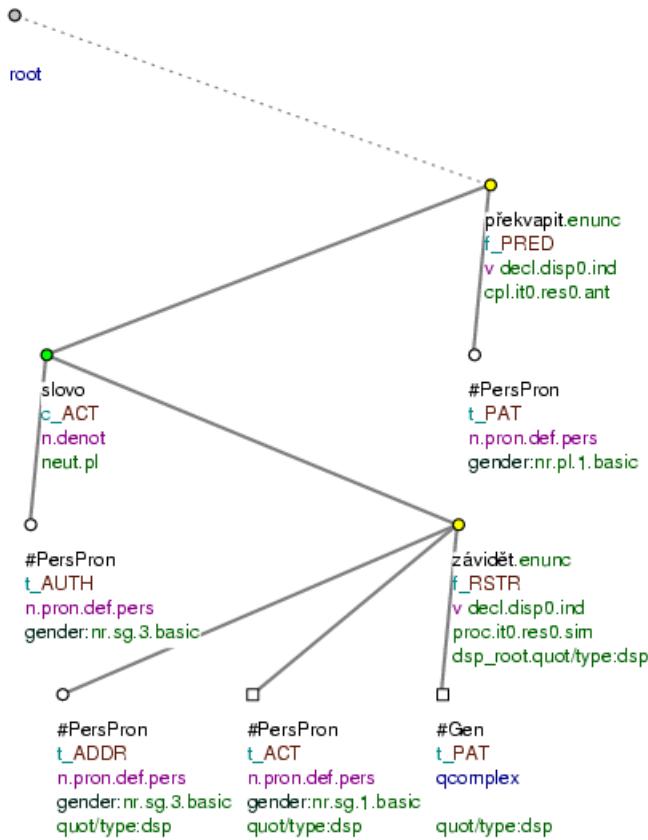

**Figure 7.37. Direct speech is not a modification of the reporting clause**

*V jednotce se lepší nálada: „Porazíme je.“ (=lit. In unit REFL is\_getting\_better mood: (We) will\_beat them.)*



**Figure 7.38. Direct speech as a modification of a noun**

*Jeho slova „Závidím mu.“ nás překvapila. (=lit. His words „(I) envy him“ us surprised.)*



## 7.6. Text segments marked by graphic symbol

Text segments marked by graphical symbols are text segments in quotation marks, brackets or between dashes.

**Quotation marks.** When analyzing text segments in quotation marks (double or simple) (“text segment in quotation marks”) two kinds of information are specified (in the `quot` attribute (see Section A.2.17, “`quot`”), which is a list every item of which is a structure of attributes `quot/type` and `quot/set_id`):

- a. **scope of the quotation marks**, i.e. which part of the tree represents the expressions within the quotation marks.

For each text segment in quotation marks a unique identifier is selected. For all nodes representing expressions in the given text segment within quotation marks this unique identifier is encoded in the attribute `quot/set_id`. A node can be a member of one or more sets of such marked nodes (embedded quotation), or of none.

- b. **type of use**, i.e. in which function the quotation marks have been used (direct speech, title, quotation).

Information on the type of use of the quotation marks is encoded in the `quot/type` attribute. Every node representing an expression that is part of a text segment in quotation marks is, apart from the `quot/set_id` attribute, which gathers the nodes of the text segment together, assigned also a value of the attribute `quot/type`. The type of use of the quotation marks is specified for the whole text segment. Therefore, it holds that nodes with the same identifier in the `quot/set_id` attribute have also the same value in the `quot/type` attribute. The `quot/type` attribute distinguishes four types of quotation mark use; see Table 7.2, “Types of quotation mark use”.

Table 7.2. Types of quotation mark use

Type	quot/type	Definition	Examples
<b>quotation</b>	citation	a text segment representing a quoted speech, which is formally integrated in the construction. The quoted segment is not complete; usually just a part of utterance is quoted (even one word).	<i>Dodal, že SRN se nechce s Japonskem „<u>tlacít</u>“, nýbrž „<u>podporovat</u>“. (=He added that the FRG did not want to "pressurise" Japan but to "be supportive")</i> <i>Řekl, že „<u>to není nutné</u>.“ (=He said that "it is not necessary")</i>
<b>direct speech</b>	dsp	direct speech (see Section 7.5, "Direct speech"); the quoted speech is a coherent (complete) unit	<i>„<u>Máme několik set členů</u>.“ (=We have several hundreds of members)</i> <i>Vhodnost investic obhajuje slovy: „<u>Velká část budov je v zanedbaném stavu</u>.“ (=He defends the investments: "A big number of the buildings are in a bad condition")</i>
<b>meta-language</b>	meta	expressions used metalinguistically, i.e. expressions in which the words are not used in the usual way; the words themselves are the topic, i.e. their meaning or form	<i>cedule s nápisem „<u>Romy neobsluhujeme</u>“ (=the notice "We do not serve Romanies")</i> <i>Germanismus <u>klika</u> se užívá ve významu „<u>štěstí</u>“. (=The word <u>klika</u> is used in the meaning "luck")</i> <i>Potetovaná kůže znamená „<u>Já jsem tady kápo</u>“. (=Skin full of tattoos means: "I am the boss here")</i>
<b>title</b>	title	proper name, title (identifying expression)	<i>Lépe než „<u>Otvíráme</u>“ měl pan ředitel svůj text nazvat „<u>otevřená dlaň</u>“. (=The text should have been named "open palm" rather than "we are opening")</i> <i>Do města přijeli vyzvání motem „<u>Kdo nebyl v Poděbradech, nemá rád tenis</u>“. (=They came, inspired by the motto "Who wasn't in Podebrady, doesn't like tennis")</i>
<b>other cases</b>	other	other types of use (irony, metaphor etc.)	<i>Firma respektuje vlivy „<u>výšší moci</u>“. (=The company respects the "higher forces")</i> <i>Jedna městská čtvrť „<u>váží</u>“ sto kilogramů. (=One town quarter "weighs" 100 kilograms)</i>

**Text segments in brackets and between dashes** are not marked in any special way. **Brackets** are interpreted as signalling parenthesis. Text segments in brackets are annotated according to the rules in Section 6.5, "Parenthesis". The function of **dashes** delimiting a text segment can vary. In a number of cases text segments between dashes are treated as parenthesis (for the rules, see Section 6.5, "Parenthesis"); in other cases the dashes are treated as co-ordinating connectives.

# Chapter 8. Functors and subfunctors

**Functors** are semantic values of syntactic dependency relations. The information regarding the functor of each node is contained in the attribute `functor` (see Section A.2.7, “functor”).

Subfunctors are described in Section 8.13, “Further specification of the meaning of a functor”.

## 8.1. Functors for effective roots of independent clauses

**Functors for effective roots of independent clauses** are functors that indicate the clause type and signal that the clause is independent. There are different functors for verbal, nominative, interjectional and vocative clauses, as well as for parenthesis. For a survey of the functors for effective roots of independent clauses and their definitions see Table 8.1, “**Functors for effective roots of independent clauses**”.

**Table 8.1. Functors for effective roots of independent clauses**

Functor	Definition	Example
<b>DENOM</b>	effective root node of an independent nominative clause (which is not parenthetical)	<i>Názory čtenářů.</i> (=Our readers' opinions.)
<b>PAR</b>	effective root node of an independent verbal or nominative clause (parenthesis)	<i>Přijedu 13. prosince (pátek).</i> (=I'll arrive on December 13th (Friday).)
<b>PARTL</b>	effective root node of an independent interjectional clause	<i>Pozor!</i> (=Watch out!)
<b>PRED</b>	effective root node of an independent verbal clause (which is not parenthetical)	<i>Pavel dal kytku Marii.</i> (=P. gave a flower to M.)
<b>VOCAT</b>	effective root node of an independent vocative clause	<i>Pane majore, hodlám zavést nové prvky.</i> (=Major; I intend to introduce some new elements.)

## 8.2. Argument functors

The basic definitions of the arguments are in Section 6.2, “Valency”. For a brief survey of argument functors and their definitions see Table 8.2, “**Argument functors**”.

**NB!** The modification with the `MAT` functor is also an argument; it is described in Section 8.10, “Specific adnominal functors”.

**Table 8.2. Argument functors**

Functor	Definition	Examples
<b>ACT</b>	functor for the first argument In those cases when there is no argument shifting, the modification with the <code>ACT</code> functor refers to the human or non-human originator of the event, the bearer of the event or a quality/property or the experiencer.	<i>Ten román mě oslovil.</i> (=The novel appealed to me.) <i>Je mi smutno.</i> (=I feel sad.) <i>zločiny mafie na Sicílii</i> (=mafia's crimes on Sicily) <i>jeho výkon</i> (=his performance) <i>K nehodě došlo v noci na včerejšek.</i> (=The accident took place the night before the last night.)

<b>ADDR</b>	functor used for arguments with the cognitive role of the recipient of the event. In those cases when the argument shifting does apply, the modification with the ADDR functor is assigned if the verb (noun, adjective) has at least three arguments.	<i>Dal dítěti hračku.</i> (=He gave the child a toy.) <i>Vzal dítěti hračku.</i> (=He took the toy from the child.) <i>Obrátil se na soud s problémem.</i> (=He turned to the court with a problem.) <i>Albertyn vede před Puzarem.</i> (=A. is leading before P.) <i>projev necitlivosti vůči mrtvým</i> (=insensitivity towards the dead)
<b>EFF</b>	functor used for arguments with the cognitive role of the effect/result of the event. In those cases when the argument shifting does apply, the modification with the EFF functor is assigned if the verb (noun, adjective) has at least three arguments.	<i>Považoval Pavla za odborníka.</i> (=He considered P. a professional.) <i>Cinili si život snesitelným.</i> (=They made each other's life bearable.) <i>Změnila účes z kudrn na rovné vlasy.</i> (=She changed her hairstyle from curly hair to straight hair.) <i>Slyšet hodiny tikat.</i> (=to hear the clock tick) <i>Řekl, že nepřijde.</i> (=He said he wouldn't come.)
<b>ORIG</b>	functor used for arguments with the cognitive role of the source of the event. In those cases when the argument shifting does apply, the modification with the ORIG functor is assigned if the verb (noun, adjective) has at least three arguments.	<i>Vyráběli ze dřeva nábytek.</i> (=They produced furniture made out of wood.) <i>Na malých kroužcích založili novou organizaci.</i> (=They build a new organization on small groups.) <i>Slyšel o neštěstí od učitele.</i> (=He heard about the accident from the teacher.) <i>Zdražili vstupenky z 500 na 550 Kč.</i> (=The price of the tickets rose from 500 to 550Kc.) <i>kaluž z tajícího sněhu</i> (=a puddle from melting snow)
<b>PAT</b>	functor for the second argument. In those cases when there is no argument shifting, the modification with the PAT functor refers to the affected object (in the broad sense of the word).	<i>Postavili stany.</i> (=They pitched the tents.) <i>Žasl jsem nad kvalitou sýrů.</i> (=I was astonished at the quality of the cheeses.) <i>Zapomněli jsme dýchat.</i> (=We forgot to breathe.) <i>Doporučují, aby stejnou studii udělali pro ostatní závody.</i> (=They recommend to organize a competition.) <i>kniha o dinosaurech</i> (=a book on dinosaurs) <i>lidé odpovědní za deportace Židů</i> (=the people responsible for the deportations of Jews.)

## 8.3. Temporal functors

**Temporal functors** represent a set of semantically differentiated functors for adjuncts that denote various time points or intervals to which the content of the governing node is related.

The individual temporal functors are differentiated according to which question about time they answer. For a survey of the temporal functors and their definitions see Table 8.3, “**Temporal functors**”.

**Table 8.3. Temporal functors**

Functor	Definition	Examples
<b>TFHL</b>	for how long?	<i>Přijel <u>na</u> dva dny.</i> (=He came for two days.) <i>Navždy vám budu vděčný.</i> (=I'll be forever grateful.) <i>kontejnery určené k <u>dlouhodobému</u> uskladňování vyhořelého paliva</i> (=containers used for long-time storage of burn-out fuel)
<b>TFRWH</b>	from when?	<i>cukroví <u>od</u> Vánoc</i> (=sweets from Christmas) <i>Z <u>dětství</u> si nic nepamatuji.</i> (=I don't remember anything from my childhood.) <i>Zápis y jsou <u>odvčera</u>.</i> (=The registration has been in progress since yesterday.)
<b>THL</b>	how long? in what time?	<i>Dělal to <u>po</u> tři léta.</i> (=He was doing it for three years.) <i>Napsal to <u>za</u> dvě hodiny.</i> (=He wrote it in two hours.) <i>Psal ten úkol dvě hodiny.</i> (=He was writing the homework for two hours.) <i>Dlouho se neviděli.</i> (=They haven't seen each other for a long time.) <i>Pracoval, <u>dokud</u> měl síly.</i> (=He was working until he felt tired.) <i>tři čtvrtě hodinové vystoupení</i> (=forty-five minute performance.)
<b>THO</b>	how often? how many times?	<i>Rakoustí přepravci musejí stavět <u>po</u> každých šesti hodinách.</i> (=In Austria, the drivers have to stop every six hours.) <i>Každé čtyři hodiny si musím vzít prášek.</i> (=Every four hours I have to take a pill.) <i>Scházeli se <u>denně</u>.</i> (=They met every day.) <i>každoroční vyúčtování nákladů</i> (=annual accounts.)
<b>TOWH</b>	to when?	<i>Svolal schůzi <u>na</u> šestou hodinu.</i> (=He called the meeting for six o'clock.) <i>Pro nejbližší období plánuje ODA setkání se slovenskými poslanci.</i> (=In the next few months, ODA is planning...) <i>Odložíme to <u>napříště</u>.</i> (=We'll put it off until next time.)
<b>TPAR</b>	during what time? in parallel with what?	<i>Během naší dovolené ani jednou nepršelo.</i> (=During our holiday it didn't rain once.) <i>Souběžně s přednáškou probíhaly semináře.</i> (=Simultaneously with the lecture, there were some seminars.) <i>Cestou do Norimberku jsme třikrát stavěli.</i> (=On our way to Nuernberg, we stopped three times.) <i>Zatímco spala, přemýšlel jsem.</i> (=While she was sleeping, I was thinking.)
<b>TSIN</b>	since when?	<i>Od soboty nepršelo.</i> (=It didn't rain since Saturday.) <i>Počínaje snídaní nic nejedl.</i> (=He didn't eat anything since breakfast.) <i>Odjakživa jsem neměl rád tohle město.</i> (=I never liked this town.) <i>Co odešli, je tu klid.</i> (=Since they left, it is calm here.)
<b>TTILL</b>	till when?	<i>Do večera budu v Praze.</i> (=I'll be in Praha by the evening.) <i>K tomuto termínu bylo podáno 2173 žádostí.</i> (=By this date, 2173 applications were submitted.) <i>Dodnes nevím, kde je.</i> (=Until today I don't know where it is.) <i>Udělej to, než se vrátím.</i> (=Do it before I come back.)

<b>TWHEN</b>	when?	<p>Přijdu <u>k večeru</u>. (=I'll come in the evening.)</p> <p>Uvidíme se <u>koncem roku</u>. (=We'll see each other at the end of the year.)</p> <p><u>V období deštů</u> sloni zůstávají v horách. (=In the rainy period, the elephants stay in the mountains.)</p> <p><u>Před týdnem</u> přestala dojíždět za prací. (=A week ago, she stopped commuting to work.)</p> <p>Sejdeme se 2. <u>února</u>. (=We'll meet on February 2nd.)</p> <p>Dočkáme se <u>brzy</u> rychlé dopravy. (=There will soon be a faster means of transport.)</p> <p><u>Sotva se naučil</u> jeden jazyk, už se učí další. (=As soon as he learned one language, he started another one.)</p> <p><u>budoucí</u> zvyšování výroby (=the future production increase)</p>
--------------	-------	--

## 8.4. Locative and directional functors

**Locative and directional functors** represent a set of semantically differentiated functors for adjuncts that - in a broader sense - denote places to which the content of the governing node is related.

The individual functors are differentiated according to which question about location or direction they answer. For a survey of the locative and directional functors and their definitions see Table 8.4, “**Locative and directional functors**”.

Table 8.4. Locative and directional functors

Functor	Definition	Examples
<b>DIR1</b>	where from?	<p><u>Ustoupil od stěny</u>. (=He made a step from the wall.)</p> <p><u>Směrem od Prahy</u> jede víc aut. (=In the direction from Prague, there are more cars.)</p> <p><u>jeden z chlapců</u> (=one of the boys)</p> <p><u>Zevnitř</u> se ozval hluk. (=There was some noise coming from inside.)</p>
<b>DIR2</b>	which way?	<p><u>Prošli blízko toho domu</u>. (=They passed near the house.)</p> <p><u>Souběžně s železnicí</u> vede silnice. (=The road is parallel with the railway.)</p> <p><u>Využila pohledu skrz okno</u>. (=...through the window)</p> <p><u>Maršál Rommel ustupoval údolím řeky Vardaru</u>. (=...through the valley...)</p> <p><u>Tudy cesta nevede</u>. (=We can't go this way.)</p>
<b>DIR3</b>	where to?	<p><u>Lidé to mají do sběren daleko</u>. (=It is far to the collecting places.)</p> <p><u>Přiměla ho k návratu za moře</u>. (=She made him to come back over the sea.)</p> <p><u>Hleděl tváří v tvář problému</u>. (=He faced the problem.)</p> <p><u>Nikam se nepůjde</u>. (=We are not going anywhere.)</p> <p><u>zpětné získání celého podílu</u> (=getting the share back)</p>
<b>LOC</b>	where?	<p><u>Nebezpečí může být všude kolem nás</u>. (=Danger can be everywhere around us.)</p> <p><u>V oblasti vzdělávání</u> máme velké mezery. (=In the area of education...)</p> <p><u>Za vyřešenými problémy</u> se objevily další. (=In place of the solved problems, there some new again.)</p> <p><u>Studentku přepadl poblíž stadionu</u> neznámý mladík. (=...near the stadium)</p> <p><u>Zůstaň doma</u>. (=Stay at home.)</p> <p><u>zahraniční utkání</u> (=a match in a different country.)</p>

## 8.5. Functors for causal relations

**Functors for causal relations** represent a set of semantically differentiated functors that are assigned to adjuncts referring to different causal relations among events or states. Modifications with these functors refer to something with respect to what the governing event is assessed, explained etc.

Functors for causal relations are differentiated depending on the type of relation between the two events, states etc.; i.e. depending on whether it is a cause, purpose, condition etc. For a survey of the functors for causal relations and their definitions see Table 8.5, “**Functors for causal relations**”.

**Table 8.5. Functors for causal relations**

Functor	Definition	Examples
<b>AIM</b>	purpose, aim	<p><i>Jsem tu, abych vám pomohl.</i> (=I am here to help you.)</p> <p><i>Domek byl zastaven už dříve ke krytí půjčky.</i> (=... in order to cover the loan)</p> <p><i>V zájmu zkvalitnění legislativního procesu by měl každý zákon projít oběma komorami.</i> (=For the improvement of the process... it is necessary...)</p> <p><i>Pracoval jen za účelem výdělku.</i> (=in order to make some money)</p>
<b>CAUS</b>	cause proper	<p><i>Nepovím vám to, protože byste mi stejně nevěřili.</i> (=I am not going to tell you because you wouldn't believe me anyway.)</p> <p><i>Kvůli tobě mám teď po dobré náladě.</i> (=Because of you I'm in a bad mood.)</p> <p><i>Z důvodu nemoci zavřeno.</i> (=Closed for illness.)</p> <p><i>Radoval se nad dárkem jako malé dítě.</i> (=He was delighted because of the present.)</p> <p><i>Zemřel hladem.</i> (=He died with hunger.)</p>
<b>CNCS</b>	concession	<p><i>Vyhral, přestože nepatřil mezi favority.</i> (=Although nobody expected it, he won.)</p> <p><i>Navzdory velkým studijním úspěchům se v praxi neuplatnil.</i> (=Although he wasn't very successful...)</p> <p><i>Přes své dobré výchování se nezachoval nejlépe.</i> (=Although his manners are normally good...)</p>
<b>COND</b>	condition	<p><i>Jestliže nepůjde dobrovolně, použijeme násilí.</i> (=If he fights, we'll use the force.)</p> <p><i>V případě, že se nedostaví, schůzi rozpustíme.</i> (=If he doesn't arrive there will be no meeting.)</p> <p><i>Nebýt vás, nebyl bych tady.</i> (=If it were not for you I wouldn't be here.)</p>
<b>INTT</b>	intention	<p><i>Návštěvnici sem přijíždějí lyžovat.</i> (=The people are coming here for skiing.)</p> <p><i>Spěchal zmáčknout spoušť svého nikonu.</i> (=He rushed to make the picture.)</p> <p><i>Pozvali ho k dialogu.</i> (=They invited him to the dialogue.)</p> <p><i>Izraelský premiér je v Praze na návštěvě.</i> (=the Israeli Prime Minister on the visit.)</p>

## 8.6. Functors for expressing manner and its specific variants

**Functors for expressing manner and its specific variants** are a rather heterogeneous group of adjunct functors that express all kinds of inner characteristics of events, i.e. the manner in which the event (state) is carried out.

Manner can be expressed in different ways - by comparison, by specifying the result or instrument used for accomplishing the event, by expressing quantity etc.; these different ways of expressing manner correspond to different (manner) functors. For a survey of the functors for expressing manner and its specific variants and their definitions see Table 8.6, "Functors for expressing manner and its specific variants".

**Table 8.6. Functors for expressing manner and its specific variants**

Functor	Definition	Examples
<b>ACMP</b>	accompaniment	<i>Pracuje bez brýlí. (=He works without his glasses.)</i> <i>tatínek s maminkou (=Father with Mother)</i> <i>V souvislosti s růstem mezd se zvýšila poptávka. (=In the connection to the growth of the wages...)</i> <i>Koupili dvě sady lega s tím, že dají každému synovi jednu. (=...with the idea that they'll give each to one of the sons.)</i>
<b>CPR</b>	comparison	<i>Musíme udělat nepochyběně menší a snazší manévr, než byl ten minulý. (=...smaller and easier than the last one.)</i> <i>Podmínky jsou až nebezpečně příznivé, podobně jako tomu bylo v Americe. (=..just like it was in America.)</i> <i>V porovnání s tebou budu vždycky lepší. (=When compared to you I'll always be better.)</i>
<b>CRIT</b>	criterion	<i>Posuzuje lidi dle obleku. (=He is judging people on the basis of their clothes.)</i> <i>Po vzoru Německa bude firma pořádat fotbalové turnaje. (=Following Germany...)</i> <i>Ve světle těchto faktů se hroutí i tato představa. (=In the light of these facts...)</i> <i>Tento svazek nebyl v rozporu s platnými předpisy registrován. (=This was not registered which is not in accordance with the rules.)</i>
<b>DIFF</b>	difference	<i>Tuzemský výrobce dodal hlavy o čtyři dny později. (=...four days later.)</i> <i>Čím je víno starší, tím je lepší. (=The older the wine the better it is.)</i>
<b>EXT</b>	extent	<i>Cesta je dlouhá přes dvacet kilometrů. (=...more than twenty km.)</i> <i>Obyvatelstvo proudí po tisících. (=Thousands of people are coming.)</i> <i>Náklady na palivo činily téměř miliardu korun. (=The costs were almost one billion crowns.)</i> <i>Dnes je docela hezky. (=The weather is quite nice today.)</i>
<b>MANN</b>	manner proper	<i>Pracuje pomalu. (=He works slowly.)</i> <i>Hráči si vychutnali radost z vítězství o samotě. (=They enjoyed the victory alone.)</i> <i>Dochovala se v podobě iluminovaného rukopisu. (=.. in the form of an illuminated manuscript.)</i> <i>Jednal, jak mu ukládá zákon. (=He acted as the law required.)</i> <i>elektrické ovládání (=electric control)</i>

<b>MEANS</b>	means	<p><i>Výstražnými výstřely zahnala pobřežní hlídka rybářské lodě.</i> (=...using the warning shots.)</p> <p><i>Zloděj odjel na bicyklu.</i> (=The thief left on the bike.)</p> <p><i>Investor může prodat akcie prostřednictvím makléřské firmy.</i> (=...through a broking house.)</p> <p><i>Časopisecky jsem povídky představil již v roce 1965.</i> (=In magazines, I published the stories already in 1965.)</p>
<b>REG</b>	regard	<p><i>Stanoviště je třeba upravit se zřetelem k množství účastníků.</i> (=It is necessary to prepare the spot with regard to the number of participants.)</p> <p><i>U celního zákona však s hlasováním po jménech nikdo nepřišel.</i> (=As for this law, nobody proposed voting...)</p> <p><i>školy pro sluchově postižené</i> (=schools for people with a hearing defect; lit. aurally handicapped.)</p> <p><i>Co se týče odhadu budoucího vývoje, ten je již usměrněn ekonomickými faktory.</i> (=As for the future development...)</p> <p><i>penzijní připojištění</i> (=a contributory pension scheme.)</p>
<b>RESL</b>	result	<p><i>maso upečené do zlatova</i> (=roasted until golden brown)</p> <p><i>Ukradenými auty zaútočili na vůz tak, že havaroval</i> v příkopu. (=...they pushed the car in such a way that it ended up in a ditch.)</p> <p><i>Autor se snažil napsat povídku tak, aby si ji mohl přečíst každý.</i> (=The author is trying to write it in such a way that anybody can read it.)</p> <p><i>deficitní hospodaření</i> (=deficit economy)</p>
<b>RESTR</b>	exception/restriction	<p><i>Kromě dosavadních úkolů bude úřad vydávat licence pro podnikání na dráze.</i> (=Apart from its current job, the office is going to issue...)</p> <p><i>Veřejnost není zlá až na některé jedince.</i> (=The society is not evil, except for certain individuals.)</p> <p><i>S tím nelze než souhlasit.</i> (=One can only agree.)</p>

## 8.7. Functors for rhematizers, sentence, linking and modal adverbials

**Functors for rhematizers, sentence, linking and modal adverbials** are functors of atomic nodes representing adjuncts that have the rhematizing function, that link the sentence to the preceding text or express various modal and attitudinal characteristics. For a survey of the functors for rhematizers, sentence, linking and modal adverbials and their definitions see Table 8.7, “**Functors for rhematizers, sentence, linking and modal adverbials**”.

**Table 8.7. Functors for rhematizers, sentence, linking and modal adverbials**

Functor	Definition	Examples
<b>ATT</b>	evaluating or emotional attitude of the speaker to the content of the utterance	<p><i>Je to bohudíky za námi.</i> (=...fortunately..)</p> <p><i>Taková je bez nadsázky Honda Prelude 3.2.</i> (=...without exaggeration...)</p> <p><i>Když ale vy to nevidíte dobře.</i> (=But you can't see it clearly.)</p> <p><i>takzvané podnikání</i> (=so called business)</p>

<b>INTF</b>	expression intensifying the subject (so called “false subject”)	<i>On Jirka ještě spí.</i> (=lit. Well, Jirka still is_asleep.) <i>To Karel nepřišel.</i> (=lit. That Karel did not come.) <i>Ono táhne.</i> (=lit. There is_(a)_draught.) <i>To hučí letadla.</i> (=lit. That are_buzzing (the) planes.)
<b>MOD</b>	modal characteristics of the content of the utterance	<i>Pravděpodobně přijdeme.</i> (=We'll probably come.) <i>Vedlo by to zřejmě ke zničujícímu rozkolu.</i> (=Apparently, this would lead to...) <i>Takoví pracovníci v zásadě neexistují.</i> (=In principle, such workers do not exist.) <i>možné selhání</i> (=possible failure)
<b>PREC</b>	expression linking the clause to the preceding text	<i>Jsem tedy šťasten.</i> (=Hence, I'm happy.) <i>Ale to zatím není náš případ.</i> (=But this is not our case.) <i>Izolovaný výzkum však nemůže přinášet výsledky.</i> (=However; isolated research cannot have good results.)
<b>RHEM</b>	rhematizer	<i>Jen on o tom nevěděl nic.</i> (=Only he didn't know anything.) <i>Teprve před týdnem přestala za prací do Púchova dojízdět.</i> (=Just a week ago...)

## 8.8. Functors for multi-word lexical units and foreign-language expressions

Functors for multi-word lexical units and foreign-language expressions are functors assigned to nodes representing a foreign-language expression (see Section 7.4, “Foreign-language expressions”), or the dependent part of a multi-word predicate (see Section 7.1.1, “Multi-word predicates”). For a survey of the functors for multi-word lexical units and foreign-language expressions and their definitions see Table 8.8, “**Functors for multi-word lexical units and foreign-language expressions**”.

Table 8.8. Functors for multi-word lexical units and foreign-language expressions

Functor	Definition	Examples
<b>CPHR</b>	nominal part of complex predicates or a pre-dicative adverb with quasi-modal verbs	<i>Dostali rozkaz nevycházet ze stanů.</i> (=They got the order not to leave the tents.) <i>Je třeba odejít.</i> (=It is necessary to leave.)
<b>DPHR</b>	the dependent part of an idiomatic expression	<i>Jde mi na nervy.</i> (=He's getting on my nerves.) <i>křížem krážem</i> (=criss cross)
<b>FPHR</b>	foreign-language expression	<i>Cizinec zvolal: "This is not true."</i> (=The foreigner called:...)

## 8.9. Functors for some specific (new) modifications

The present section is devoted to the description of functors that are traditionally not included in grammar books. These functors are not classified as being members of different semantic groups yet. For a survey of the functors for specific (new) modifications and their definitions see Table 8.9, “**Functors for some specific (new) modifications**”.

**Table 8.9. Functors for some specific (new) modifications**

Functor	Definition	Examples
<b>BEN</b>	adjunct expressing to whose advantage or disadvantage something happens	<i>Peníze zůstanou klientovi dále na účtu.</i> (=The money will stay on the client's account.) <i>Bylo by to <u>ku prospěchu</u> věci.</i> (=This would be for the good.) <i>šance pro movité nájemníky</i> (=a chance for rich tenants) <i>To je další argument proti tomu, aby mzdová regulace existovala.</i> (=This is another argument against the existence...)
<b>CONTRD</b>	adjunct expressing a contrasting background (confrontation) for the governing event	<i>Zatímco dříve se dotovaly byty, které ovšem zdaleka nedostávali ti nejpotřebnější, napříště by měly peníze směřovat přímo k lidem, tedy k nájemníkům.</i> (=While in the past..., in the future ...)
<b>HER</b>	adjunct referring to a person (group of people, institution, time) which is the source of an inherited object	<i>Po Juliu Fučíkovi pobírala 320 korun vdovské penze.</i> (=She got a pension after JF.) <i>Jmenovala se Barbora podle patronky horníků.</i> (=Her name was B., after the patron of miners.)
<b>SUBS</b>	adjunct expressing that something was substituted for something else	<i>Libuše Šroubková jménem firmy Inreka předala základní škole deset nových tříd.</i> (=...in the name of the company...) <i>Do učeben zasednou otcové místo svých synů.</i> (=...fathers instead of their sons.) <i>Za otce jednal strýc.</i> (=In place of my father...) <i>Svobodná inteligence musí spolupracovat a pomáhat, místo aby se posmívala.</i> (=...they should help instead of laughing at us.)

## 8.10. Specific adnominal functors

**Specific adnominal functors** are functors for modifications exclusively modifying (semantic) nouns. These adnominal modifications have specific functions that cannot be found with adverbial modifications. For a survey of the specific adnominal functors and their definitions see Table 8.10, “**Specific adnominal functors**”.

**Table 8.10. Specific adnominal functors**

Functor	Definition	Examples
<b>APP</b>	adjunct referring to the person or thing something or someone belongs to	<i>organizace <u>neslyšících</u></i> (=organization of the hearing-impaired) <i>můj klobouk</i> (=my hat)
<b>AUTH</b>	author, creator, originator of artefacts	<i>deset básnických sbírek mladých básníků</i> (=young poets' poetry books) <i>Svěrákův Akumulátor I</i> (=Svěrák's Akumulátor I) <i>skladba Svatý Václav od Marty Jiráčkové</i> (=MJ's Svatý Václav)

<b>ID</b>	effective root of an identifying expression represented by an identification structure	<i>v případu <u>Kott</u> (=in the Kott case) pojem <u>času</u> (= notion of time) v pražské Galerii <u>U Řečických</u> (=the gallery U Řečických)</i>
<b>MAT</b>	adnominal argument referring to the content (material etc.) of something	<i>balík <u>papíru</u> (=a box of paper) polovina <u>dortu</u> (=one half of the cake)</i>
<b>RSTR</b>	adnominal adjunct more closely specifying its governing noun	<i>drsné <u>počasí</u> (=harsh weather) <u>pět</u> <u>dětí</u> (=five children) <u>rozhodčí</u> <u>Severýn</u> (=referee Severýn) <u>lod'</u> <u>o výtlaku</u> 9700 tun (=a vessel with a displacement of 9700 tons) <i>Udeřil i člověka, který si to <u>nezasloužil.</u> (=He hit a man who didn't deserve it.)</i></i>

## 8.11. Functor for the predicative complement (COMPL)

Due to the specific properties of predicative complements, which cannot be attributed to any other adjuncts, we take this functor to be a special functor that does not belong to any group of functors. The functor for the predicative complement has the value COMPL (see Table 8.11, “**Functor for the predicative complement**”).

**Table 8.11. Functor for the predicative complement**

Functor	Definition	Examples
<b>COMPL</b>	predicative complement	<i>Našli kamaráda <u>nemocného</u>. (=They found their friend ill.) Hráči odcházeli ze hřiště <u>nepřemoženi</u>. (=The players were leaving the field undefeated.) Sledoval ho, jak <u>se chová</u> k mladším spolužákům. (=He watched him how he behaved to his younger classmates.)</i>

## 8.12. Functors expressing the relations between members of paratactic structures

**Functors expressing the relations between members of paratactic structures** present a special group of functors. These functors do not express a kind of dependency; they rather capture the relation between members of paratactic structures (either clauses or modifications). They are functors that are assigned to the root nodes of paratactic structures (see Section 6.4, “Parataxis”). For a survey of the functor for relations between members of paratactic structures see Table 8.12, “**Functors for coordination**”, Table 8.13, “**Functor for apposition**” and Table 8.14, “**Functor for mathematical operations and intervals**”.

Apart from the functors assigned directly to the roots, there is also a specific functor CM, which is assigned to nodes representing conjunction modifiers - see Table 8.15, “**Functor for conjunction modifiers**”.

**Table 8.12. Functors for coordination**

Functor	Relation between the coordinated elements	Examples
<b>ADVS</b>	adversative	<i>Vypadalo to, že bude hezky, ale začalo pršet.</i> (=The weather seemed nice but it started raining.) <i>V dalších řádcích nebude odkaz na žádný zákon, a přesto půjde o záležitost aktuální.</i> (=There will be no reference to a law, and still it will be highly relevant..) <i>Přijel do Prahy, nikoli do Brna.</i> (=He came to Praha, not to Brno.)
<b>CONFR</b>	confrontational	<i>Bristol je v Anglii, kdežto Glasgow je ve Skotsku.</i> (=B. is in England, while G. is in Scotland.) <i>Svobodní mládenci mírají nepořádek kolem sebe, ale ženatí naopak mírají nepořádek v duši.</i> (=Bachelors often have a mess all around them but married men, on the other hand, have a mess in their souls.)
<b>CONJ</b>	simple conjoining	<i>Mezi smysly patří zrak a sluch a hmat.</i> (=Eyesight and hearing and touch belong to the senses.) <i>Vyniká jak svědomitostí, tak houževnatostí.</i> (=She stands out due to her thoroughness, as well as her stringiness.)
<b>CONTRA</b>	coordination of two fighting, competing subjects	<i>akademie věd kontra vysoké školství</i> (=Academy of Science contra universities) <i>Bukač versus Hlinka očima Jana Horáka</i> (=Bukač versus Hlinka in JH's eyes) <i>utkání Rusko - Švédsko</i> (=Russia - Sweden)
<b>CSQ</b>	consecutive	<i>Pracoval nezodpovědně, a proto dostal výpověď.</i> (=He was irresponsible, therefore he was fired.) <i>Byl nemocný, a tudíž nepřišel.</i> (=He was sick so that's why he didn't come.) <i>Potřebujete jen 1/4 vody, takže sprcha vás stojí 0,46 Kč.</i> (=You only need 1/4 of the water so a shower costs 0.46 Kc.)
<b>DISJ</b>	disjunctive	<i>Vysloví se bud' pro, nebo proti návrhu.</i> (=They will be either for or against the proposal.) <i>Měl dvě možnosti – nechat se předstihnout, nebo způsobit havárii.</i> (=He had two possibilities - to be outrun, or to cause a crash.) <i>Mají, či nemají pravdu?</i> (=Are they right, or not?) <i>Na konzultaci musí přijít alespoň jeden z rodičů, nebo alespoň někdo z rodiny.</i> (=At least one parent has to come, or at least a member of the family.)
<b>GRAD</b>	gradation	<i>Stát neposkytne na nákup bytů žádné slevy, ani žádnou jinou finanční pomoc.</i> (=The state will not offer a discount, not even any other financial help.) <i>Nestavíme jen domy, nýbrž budujeme i dětská hřiště.</i> (=Not only do we build houses, we also build playgrounds.) <i>Závod plán splnil, ba dokonce jej překročil.</i> (=The factory fulfilled the plan; it even overfulfilled it.)
<b>REAS</b>	causal	<i>Tento krok je nebezpečný, neboť se do ekonomiky занáší nesmírná džungle.</i> (=This step is dangerous since it lets the jungle into our economy.) <i>Úkol splníme, vždyť není obtížný.</i> (=We'll fulfil the task, for it is not difficult.)

**Table 8.13. Functor for apposition**

Functor	Relation between the appositioned elements	Examples
<b>APPS</b>	apposition	<i>Božena Němcová, autorka Babičky (=BN, the author of Babička)</i> <i>jeden rok, neboli dva semestry (=one year, or two semesters)</i> <i>půdní režim, to znamená půdní vláha (=soil environment, i.e. soil humidity)</i>

**Table 8.14. Functor for mathematical operations and intervals**

Functor	Definition	Examples
<b>OPER</b>	paratactic connection of operands of mathematical operations or intervals	<i>poměr deset <u>ku</u> jedné (=10:1)</i> <i>ve věku 34 - 44 let (=at the age of 34-44)</i> <i>od hlavních bodů <u>přes</u> příklady <u>do</u> nejmenších detailů (=from the main points to the tiniest details)</i>

**Table 8.15. Functor for conjunction modifiers**

Functor	Definition	Examples
<b>CM</b>	expressions modifying coordinating connectives	<i>Rozpočet <u>nejenže</u> není přebytkový, ale <u>dokonce</u> je skrytě deficitní. (=The budget not only is not surplus, but it is even covertly deficit.)</i>

## 8.13. Further specification of the meaning of a functor

Two attributes are used to specify the meaning of certain modifications in a more detailed way:

the `subfunctor` attribute (see Section 8.13.1, “Subfunctors”);

the `is_state` attribute (see Section 8.13.2, “Attribute for the meaning of “state””).

### 8.13.1. Subfunctors

**Subfunctors** describe semantic variations within a particular functor, they provide closer specification of the meaning of the functor, and they specify more precisely the semantic relation between the modification and the governing word. These various differences within one functor are expressed by various (prepositional) cases, or by conjunctions.

The information on the subfunctor is carried by the attribute `subfunctor` (see Section A.2.19, “subfunctor”).

Subfunctors are not assigned to all PDT functors that would need a more detailed specification of their meaning but rather to a limited group of functors. Subfunctors are assigned to nodes with the following functors:

ACMP (see Table 8.16, “Subfunctors for the functor ACMP”)

BEN (see Table 8.17, “Subfunctors for the functor BEN”)

CPR (see Table 8.18, “Subfunctors for the functor CPR”)

DIR1 ( see Table 8.19, “Subfunctors for the functor DIR1”)

DIR2 (see Table 8.20, “Subfunctors for the functor DIR2”)

DIR3 (see Table 8.21, “Subfunctors for the functor DIR3”)

EXT (see Table 8.24, “Subfunctors for the functor EXT”)

LOC (see Table 8.22, “Subfunctors for the functor LOC”)

TWHEN (see Table 8.23, “Subfunctors for the functor TWHEN”)

Subfunctors were assigned to the functors in PDT on the basis of the surface form of the modification with the given functor.

For every functor (out of the selected ones), there is a subfunctor with the values `basic` and `nr`:

- **basic.** Each functor has its set of forms that express the basic meaning of the functor. A modification realized by one of these forms has the attribute `subfunctor` filled with the value `basic`.

- **nr.** Modifications in a surface form that has not been assigned to any subfunctor have the attribute `subfunctor` filled with the value `nr` (not recognised).

**Table 8.16. Subfunctors for the functor ACMP**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	positive accompaniment	<i>s+7</i> <i>spolu s+7</i> <i>společně s+7</i> <i>v čele s+7</i> <i>ve spojení s+7</i> <i>zároveň s+7</i> <i>rukou v ruce s+7</i>	<i>tatínek s maminkou</i> (=Father with Mother) <i>celá rodina v čele s otcem</i> (=the whole family with the Father in the lead) <i>Zároveň s hovorem dostane potřebné informace o volajícím zákazníkovi.</i> (=Together with the call...)
<b>circ</b>	vague circumstance	<i>bez toho, aby+vfin</i> <i>s tím+vfin</i> <i>s tím, že+vfin</i> <i>v souvislosti s+7</i> <i>v souvislosti s tím, že+vfin</i>	<i>Tenis může být podívanou i bez toho, aby po kurtě chodily polonahé děvky.</i> (=Tennis can be a good show even without half-naked girls walking on the tennis court...) <i>Vyšetřovatelé si v souvislosti s korupčním skandálem přišli vyslechnout i dalšího svědka.</i> (...in the connection with the scandal...)
<b>incl</b>	inclusion	<i>včetně+2</i>	<i>Ročně by tedy zaplatila na pojistném včetně úrazového připojištění 4104 korun.</i> (=...accident insurance including.)
<b>wout</b>	negative accompaniment	<i>bez+2</i>	<i>pokoj bez vlastního příslušenství</i> (=a room without conveniences)

**Table 8.17. Subfunctors for the functor BEN**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	benefit	<i>ku prospěchu+2</i> <i>pro+4</i> <i>ve prospěch+2</i> <i>v zájmu+2</i>	<i>Profit připravuje pro své čtenáře poradnu.</i> (...for the readers...) <i>Zřikají se kariéry ve prospěch dětí.</i> (=.for the benefit of their children)
<b>agst</b>	disadvantage/detrimen-	<i>na úkor+2</i> <i>proti+3</i> <i>proti tomu, aby+vfin</i> <i>v neprospěch+2</i>	<i>Hlavní překážkou státu fungovat na úkor jednotlivce je demokratický systém.</i> (=.at the expense of an individual..) <i>válečný vývoj v neprospěch Německa</i> (=..to the disadvantage of Germany)

**Table 8.18. Subfunctors for the functor C<sub>P</sub>R**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	similarity	jako jak	<i>Sloni jsou jako Angličani.</i> (=Elephants are like Englishmen.) <i>Jenže stejně rychle, jak naděje svitla, tak rychle pohasla.</i> (=The hope was gone as fast as it came.)
<b>than</b>	difference	genitive než nežli	<i>To byla otázka pro 982 respondentů starších 14 let.</i> (=...respondents older than 14 years) <i>Výrobili více než 495 milionů metrů krychlových pitně vody.</i> (=They produced more than 495 million...)
<b>wrt</b>	“the entity something is compared to”	<i>naproti+3</i> <i>na rozdíl od+2</i> <i>proti+3</i> <i>oproti+3</i> <i>ve srovnání s+7</i> <i>ve srovnání s tím+vfin</i> <i>v porovnání k+3</i> <i>v porovnání s+7</i>	<i>Proti dřívějšku se však zase objevili noví zájemci.</i> (=In comparison to the past...) <i>ČR je v tomto ohledu výjimečná ve srovnání s okolními zeměmi střední Evropy.</i> (=When compared to the other countries of Central Europe..) <i>Na rozdíl od hotelů nemohly lázně využívat zisku tak dobré.</i> (=Unlike the hotels, the spas couldn't...)

**Table 8.19. Subfunctors for the functor DIR1**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	“where from”	adverb <i>od+2</i> <i>z+2</i>	<i>Zezadu se prodral ke stříbru domácí miláček.</i> (=From behind..) <i>Ze zahraničí k nám začali jezdit chudší turisté.</i> (From abroad...)

**Table 8.20. Subfunctors for the functor DIR2**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	“which way”	adverb instrumental	<i>Musí ovládat umění nabízet své výrobky přímo ze dvora třeba turistovi, který projde okolo.</i> (=...a tourist who is going past him.) <i>přenos dat <u>vzduchem</u></i> (=air transmission)
<b>across</b>	“across”	<i>napříč+7</i> <i>přes+4</i>	<i>Přibližování satelitů k Atlantické alianci spouští novou železnou oponu <u>napříč kontinentem</u>.</i> (=...across the continent) <i>přeprava transportů <u>přes Českou republiku</u></i> (=via the Czech republic)
<b>along</b>	“along”	<i>po+6</i> <i>podél+2</i>	<i>V roce 1997 pravděpodobně projedou první vozidla <u>po dálnici Praha -Plzeň</u>.</i> (=..on the highway Praha -Plzeň.) <i>Jel jsem <u>podél něj</u>.</i> (=I went along it.)
<b>around</b>	“around”	<i>kolem+2</i> <i>okolo+2</i>	<i>Kolem právní <u>ochrany</u> software se u nás chodí v širokých kruzích.</i> (=lit. Around legal protection...)
<b>betw</b>	“between”	<i>mezi+7</i>	<i>Plechové krabičky aut se plazí po výmolech hliněných cest <u>mezi nevzrušenou zvěří</u>.</i> (=..between the uninterested animals.)
<b>near</b>	“near”	<i>vedle+2</i>	<i>A tak se naše věda stala popelkou, jen nesměle kráčející <u>vedle civilizovaného světa</u>.</i> (=...shyly walking beside the civilized world.)

**Table 8.21. Subfunctors for the functor DIR3**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	“where to”	adverb <i>do+2</i> <i>do čela+2</i> <i>na+4</i> <i>směrem na+4</i> <i>směrem do+2</i> <i>směrem k+3</i>	<i>Soused v tichosti vyčkal mezi větvemi a pak se odplížil domů.</i> (=...he crawled home.) <i>Máme zaměstnance, které občas vysíláme na služební cestu.</i> (=...sending the employess to business trips.) <i>Včera odpoledne byl na Nuselském mostě pruh <u>směrem do centra</u> normálně průjezdny.</i> (=...to the center...)
<b>above</b>	“above”	<i>nad+4</i>	<i>Už odpoledne se vrátila zpět <u>nad hranici</u> 1100 lir za marku.</i> (=It is back above the 1100 lires for a mark level.)
<b>behind</b>	“behind”	<i>za+4</i>	<i>Že se <u>za tento plášt'</u> schová leccos dalšího, neřeba připomínat.</i> (=...can be hidden under this coat...)
<b>below</b>	“below”	<i>pod+4</i>	<i>Dostal se <u>pod auto</u>.</i> (=He got under the car.)
<b>betw</b>	“between”	<i>mezi+4</i>	<i>Dal to <u>mezi ty drobnosti</u>.</i> (=He put it among those little things.)
<b>elsew</b>	“outside something”	<i>mimo+4</i>	<i>Postavil se <u>mimo hrací plochu</u>.</i> (=He stood outside the playing field.)
<b>ext</b>	“extent”	<i>po+4</i>	<i>Čs. vojsko nikdy neobsadilo celé Těšínsko až <u>po Bílsko</u>.</i> (=The army never occupied the whole region up to Bílsko.)
<b>front</b>	“in front of”	<i>před+4</i>	<i>Tím mi ulehčili práci a já mohl předstoupit <u>před lidí</u>.</i> (=...so I could face the people; lit. come in _front_ of people.)
<b>near</b>	“near”	<i>vedle+2</i>	<i>Řadí tradiční <u>vedle banálního</u>, kříží vysoký i nízký styl.</i> (=He puts the traditional next to the banal, ...)
<b>opp</b>	“opposite”	<i>proti+3</i>	<i>Útok nebyl namířen <u>proti lidem</u>.</i> (=The attack was not aimed at the people.)
<b>target</b>	target	<i>po+6</i>	<i>Házel <u>po něm</u> kamením.</i> (=He was throwing stones at him.)
<b>to</b>	“to”	<i>k+3</i> <i>za+7</i>	<i>Dotyčný půjde jinam, ke konkurenci.</i> (=He will go somewhere else, to the competition.) <i>Zašli jsme do galerie <u>za Petrem</u>.</i> (=We went to the gallery to see Petr.)

Table 8.22. Subfunctors for the functor LOC

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	“where”	adverb adjective <i>na</i> +6 <i>v</i> +6	<i>Zde už můžeme váhat.</i> (=Here, we can be unsure.) <i>první pražské jednání</i> (=first negotiations in Praha) <i>ubytování v hotelu</i> (=hotel accommodation)
<b>above</b>	“above”	<i>nad</i> +7	<i>V neprofesionálním prostředí si lidé opálí jehlu <u>nad plaménkem</u> v karmě.</i> (=...above the flame..)
<b>abstr</b>	“in the area/field of”	<i>v oblasti</i> +2 <i>v oboru</i> +2	<i>Podle Redla je Hradišťan <u>v oblasti folkloru</u> kapelou ostře sledovanou.</i> (=...a band closely watched in the area of folklore.)
<b>along</b>	“along”	<i>podél</i> +2 <i>podle</i> +2	<i>Epicentrum otřesů bylo lokalizováno <u>podél západního pobřeží</u> ostrova.</i> (=...along the west coast.)
<b>around</b>	“around”	<i>kolem</i> +2 <i>okolo</i> +2	<i>Z pálení ukradených penězenek se <u>kolem jejich chalupy</u> linul čmoud.</i> (=...around their cottage...)
<b>behind</b>	“behind”	<i>za</i> +7	<i>Po kolizi jezdili závodníci pět okruhů <u>za vodicím vozem</u>.</i> (=...behind the car.)
<b>below</b>	“below”	<i>pod</i> +7	<i>Moc často jsem tu dírku - v umyvadle, <u>pod vodou</u> - hledal a pak lepil.</i> (=...under the water...)
<b>betw</b>	“between”	<i>mezi</i> +7	<i>Byl <u>mezi dvěma ohni</u>.</i> (=...between two fires)
<b>elsew</b>	“outside something”	<i>mimo</i> +4 <i>stranou</i> +2 <i>vně</i> +2	<i>Stranou strkanice nezůstává ani tisk.</i> (=...outside the conflict.) <i>vně složitých diskuzí</i> (=outside the complicated discussions)
<b>front</b>	“in front of”	<i>před</i> +7 <i>tváři v tvář</i> +3	<i>Tykal mu <u>před lidmi</u>.</i> (=...in front of people) <i>Stojí <u>tváři v tvář</u> problému.</i> (=...facing the problem)
<b>in</b>	“inside”	<i>uvnitř</i> +2	<i>jednání <u>uvnitř koncernu</u></i> (=negotiations within the concern)
<b>mid</b>	“in the middle”	<i>prostřed</i> +2 <i>uprostřed</i> +2 <i>vprostřed</i> +2	<i>Biskup sloužil mše v táborové kapli, kterou se stal indiánský stan teepee <u>uprostřed tábora</u>.</i> (=...in the middle of the camp)
<b>near</b>	“near”	<i>blízko</i> +2 <i>blízko</i> +3 <i>k</i> +3 <i>nedaleko</i> +2 <i>poblíž</i> +2 <i>po boku</i> +2 <i>při</i> +6 <i>u</i> +2 <i>v blízkosti</i> +2 <i>vedle</i> +2	<i>Hrát se má u Alžbětina mostu, <u>blízko zastávky autobusu č 78.</u></i> (=...near the bus stop..) <i>spor o oboru <u>poblíž Hrádečku</u></i> (=a dispute over a park near Hrádeček) <i>V poslední době vystupujete často <u>po boku ministrů</u>.</i> (=..next to the ministers...) <i>na jednom z ostrovů <u>při jižním pobřeží Aljašky</u></i> (=...by the south coast of Alaska) <i>Procházíme po odkrytém prostranství <u>nedaleko hotelu</u>.</i> (=...near the hotel)
<b>opp</b>	“opposite”	<i>naproti</i> +3 <i>proti</i> +3 <i>přes</i> +4	<i>Leží <u>proti oknu</u>.</i> (=It is situated opposite the window) <i>Bydlí <u>přes dvůr</u>.</i> (=...across the yard)

Table 8.23. Subfunctors for the functor **TWHEN**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	“when”	adverb genitive accusative temporal modifications without a case $k+3$ $když+vfin$ $na+4$ $na+6$ $o+6$ $při+6$ $při příležitosti+2$ $u příležitosti+2$ $v+4$ $v+6$ $v době+2$ $v období+2$ $za+2$	<i>Uvedená cena bude předána laureátům dne 19. září 1994 v Hotelu Hilton. (=...on September 19th 1994..)</i> <i>Autor (nar. 1943) je hispanista, překladatel, působí v Českém rozhlasu. (=...born 1943..)</i> <i>Přijdu k ránu. (=...at daybreak)</i> <i><u>Na jaře</u> skončí výroba. (=In spring...)</i> <i>Rota nováčků <u>při příležitosti sjezdu</u> vyčistila vojenský prostor. (=...at the occasion of the congress...)</i> <i>Jejich genocida <u>v době</u> druhé světové <u>války</u> uzavřela pokusy o společné soužití. (=...in the time of the WW II...)</i>
<b>after</b>	“after”	$po+6$ $poté, co+vfin$ $za+4$	<i>Stát ceny schválí <u>poté, co je prověří</u>. (=..after they are examined..)</i> <i><u>Za chvíli</u> přišel pan Mitrofanov. (=In a short while, Mr. M. came.)</i>
<b>approx</b>	“approximately”	$kolem+2$ $okolo+2$	<i><u>Okolo</u> přelomu roku je relativní klid. (=Around the end of the year it is calm here.)</i>
<b>before</b>	“before”	$než+vfin$ $před+vfin$	<i>Jakou povahu jsi měl, <u>než jsi přišel</u> o nohu? (=..before you lost your leg?)</i>
<b>begin</b>	“at the beginning”	$počátkem+2$ $začátkem+2$	<i>Olejomalbu jste mohli <u>začátkem února</u> koupit za 34 600 korun. (=...at the beginning of February...)</i>
<b>betw</b>	“between”	$mezi+7$	<i>hudba v přestávkách <u>mezi</u> <u>gamy</u> (=music between the games)</i>
<b>end</b>	“at the end”	$koncem+2$ $na závěr+2$ $v závěru+2$ $závěrem+2$	<i><u>Koncem</u> roku bylo bez práce 185 000 osob. (=At the end of the year..)</i> <i><u>V závěru</u> poločasu dala Sparta branku. (=Towards the end of the half..)</i>
<b>flow</b>	“in the course”	$postupem+2$ $v průběhu+2$	<i>Omezování centrální cenové regulace bude <u>v průběhu</u> roku zřejmě pokračovat. (=...in the course of the year..)</i>
<b>mid</b>	“in the middle”	$uprostřed+2$	<i>Ještě <u>uprostřed</u> minulého týdne jsem si nebyla jistá, zda budu vůbec startovat. (=...in the middle of the last week...)</i>

**Table 8.24. Subfunctors for the functor EXT**

Subfunctor	Meaning	Possible forms	Examples
<b>basic</b>	the basic notion of extent	adverb <i>do+2</i> <i>na+4</i> <i>po+6</i> <i>v+6</i> <i>za+4</i>	<i>Bez soukromého vlastnictví pozbyvá kapitálový trh částečně smyslu.</i> (=... loses its sense to a certain degree.) <i>I zanedbatelná záležitosť naroste v dětské mysli <u>do</u> tragickej <u>rozmery</u>.</i> (=...grows into tragical dimensions.) <i>Obyvatelstvo proudí <u>po</u> tisícach.</i> (=Thousands of people are streaming...)
<b>approx</b>	“approximately”	<i>kolem+2</i> <i>okolo+2</i>	<i>Letos by se měl obrat pohybovat <u>okolo</u> 1,2 miliardy korun.</i> (=...around 1.2 billion crowns.)
<b>less</b>	“less than”	<i>pod+4</i>	<i>Lidí s IQ <u>pod</u> 50 jsou již jen desetiny procenta.</i> (=..with the IQ less than 50...)
<b>more</b>	“more than”	<i>přes+4</i> <i>nad+4</i>	<i>Výnosy činily <u>přes</u> 16 miliard korun.</i> (=The sales were more than 16 milliard CZK.)

## 8.13.2. Attribute for the meaning of “state”

The verb *být*, but also full verbs and nouns are often modified by a number of prepositional phrases that have a very general meaning of state. Together with the governing verb, they express meanings like “being in some state” or “getting into some state”. For the time being, the solution is that the prepositional phrase is assigned a functor closest in the meaning (usually LOC, TWHEN, MANN) and a new attribute is used, called *is\_state*, with the values 0 and 1 (see Section A.2.15, “*is\_state*”). Those modifications that have the meaning of state are assigned the value 1. The *is\_state* attribute is specified for all nodes with adjunct functors.

Examples:

*Je v krizi.LOC [is\_state=1]* (=He is in crisis.)

*Byl po operaci.TWHEN [is\_state=1]* (=He was after operation.)

*člověk ve špatné náladě.RSTR [is\_state=1]* (=person in a bad mood.)

*Zůstal na živu.MANN [is\_state=1]* (=He stayed alive.)

*Upadl do nesnází.DIR3 [is\_state=1]* (=He fell into trouble.)

---

# Chapter 9. Coreference

The term **coreference** is used both for cases of endophoric and exophoric (deictic) reference.

When referring to a preceding element we speak of **anaphoric reference**. The expression to which it is referred is called **antecedent**. **Cataphoric reference** refers to certain following utterances or their parts; which are called **postcedent**. Apart from these already established terms, also the pair of terms **coreferring element - coreferred element** are used. It is more general and it ignores the relative location of the words in the text - as both the antecedent and postcedent can be coreferred elements.

There is:

grammatical coreference (see Section 9.2, “Grammatical coreference”),  
textual coreference (see Section 9.3, “Textual coreference”).

## 9.1. Representing coreference in the tectogrammatical trees

Three attributes have been introduced for coreference representation:

- a. attribute of the type reference `coref_gram.rf` (see Section A.2.3, “`coref_gram.rf`”) that serves for indicating grammatical coreference (see Section 9.2, “Grammatical coreference”).
- b. attribute of the type reference `coref_text.rf` (see Section A.2.5, “`coref_text.rf`”) that serves for indicating textual coreference in cases the coreferred element is overt (viz Section 9.3, “Textual coreference”).
- c. enumerative attribute `coref_special` (see Section A.2.4, “`coref_special`”) that serves for representing special types of textual coreference in which the coreferred node is not a particular node or subtree of a tree. These are cases of exophoric reference and reference to a segment (see Section 9.3, “Textual coreference”).

A node for a coreferred element is assigned a value in one of these attributes.

**Types of reference** There are several types of reference:

- a. **reference to a leaf (of a tree)**. The `coref_gram.rf` or `coref_text.rf` attribute contains the identifier of the target leaf. For example:

*Vlasta šla do divadla, kde na ni čekal Marek. (=Vlasta went to the theatre where Mirek was already waiting for her.)*

- b. **reference to a root of a subtree**. The `coref_gram.rf` or `coref_text.rf` attribute contains the identifier of the target root of a subtree. It is assumed that the coreferred element is not just this single node but the whole subtree. For example:

*Můj o dva roky mladší bratr, kterého ještě neznáš, přijde zítra. (=My two years younger brother, which you don't know yet, is coming tomorrow.)*

- c. **reference to more than one node (of a tree)**. The `coref_gram.rf` or `coref_text.rf` attribute contains several identifiers. For example:

*Marie vzala Vlastu do divadla, kde na ně čekal Marek. (=Marie took Vlasta to the theatre, where Marek was waiting for them.)*

- d. **reference to a segment**. The `coref_special` attribute is assigned the value `segm`. The coreferred node is a larger segment of the text, which is not further specified. For example:

*Rozprava o podobě reformy veřejných financí bude zahájena ve středu. Všechna jednání proběhnou za za-vřenými dveřmi. Lidovým novinám to sdělil včera ministr financí. (=The discussion.. will be started on Wednesday....The minister of Finance told this to LN.)*

- e. **extratextual reference**. The `coref_special` attribute is assigned the value `exoph`. The coreferred element is an unspecified phenomenon outside the text. For example:

*V období vrcholícího léta roku 1939 již málokdo v Evropě mohl uvěřit nadějeplným slovům britského ministerského předsedy Chamberlaina, proneseným z balkonu Buckinghamského paláce po návratu z Mnichova: Myslím, že je to mír na celou naši dobu. (=... I think this would mean peace...)*

## 9.2. Grammatical coreference

**Grammatical coreference** is such a kind of coreference in which it is possible to pinpoint the coreferred expression on the basis of grammatical rules.

**Types of grammatical coreference.** For a survey of types of grammatical coreference see Table 9.1, “**Types of grammatical coreference**”. (For more on relative clauses see Section 6.3.3, “Dependent verbal clauses (complex sentences)”. For more details on constructions with the relative pronoun *což* see Section 6.3.3.3, “False dependent clauses”. For more on the predicative complement see Section 6.1.1.1, “Predicative complement”. For more on control see Section 9.2.1, “Control”. For more on quasi-control see Section 7.1.1.4.1, “Quasi-control with complex predicates”. For more reciprocity see Section 6.6.1.2, “Ellipsis of a dependent meaning unit”).

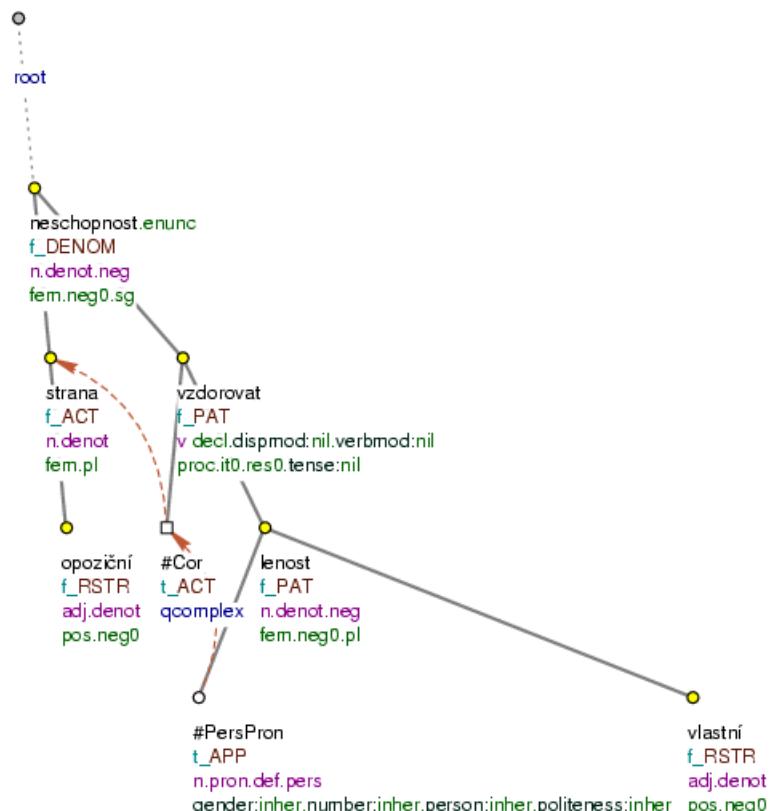
**Table 9.1. Types of grammatical coreference**

Type	Rule	T-lemma	Example
<b>Reflexive pronouns</b>	Reflexive pronouns corefer with the modification in the closest subject position.	#PersPron	<i>Neschopnost opozičních stran {#Cor.ACT} vzdorovat své vlastní lenosti. (=Inability of the opposition parties to resist their (lit. self's) own laziness.) Fig. 9.1</i> <i>Pavla se prohlíží v zrcadle. (=Pavla watches herself in the mirror.)</i>
<b>Relative elements</b>	Relative elements corefer with the noun modified by the relative clause.	který jaký kdo co kde kam kdy	<i>Ti, co kroutí hlavami, nerozumí a nechápou, zároveň instinktivně varují. (=Those who shake their heads...) Fig. 9.2</i>
<b>Relative element <i>což</i></b>	<i>Což</i> corefers with that part of the sentence which is paratactically connected to the clause introduced by <i>což</i> .	co	<i>Mužstvo získalo tři body, což je maximum. (=The team got three points, which is the maximum.) Fig. 9.3</i>
<b>Predicative complement expressed by a verb form</b>	There is a grammatical coreference relation between an argument (corresponding to the subject) of the predicative complement (expressed by a verb form) and the noun such that the predicative complement is in the second dependency relation with it.	#Cor	<i>{#PersPron.ACT} Odcházela poražena.COMPL {#Cor.PAT} (=She was leaving defeated.) Fig. 9.4</i>
<b>Argument expressed by a verb form in the position parallel to that of the predicative complement</b>	There is a grammatical coreference relation between an argument (corresponding to the subject) of the Patient or Effect (expressed by a verb form) and the noun such that the Patient/Effect agrees with it.	#Cor	<i>Mužstvo zůstává neporaženo.PAT {#Cor.PAT} i po tomto napínavém zápase. (=The team stays undefeated also after this match.)</i>
<b>Control</b>	Control is an obligatory or optional coreference relation between clearly defined modifications of certain verbs (control verbs).	#Cor	<i>Podnik plánoval zvýšit {#Cor.ACT} výrobu. (=The company planned to increase the production.)</i>

<b>Quasi-control</b>	Quasi-control is a coreference relation between a valency modification of the nominal part and a valency modification of the verbal part of a multi-word predicate.	#QCor	<i>Karel podal {#QCor.ACT} stížnost policii. (=K. submitted a complaint.)</i>
<b>Reciprocity</b>	A valency modification missing as a result of taking part in a reciprocal relation corefers with the valency position in which both the modifications are expressed simultaneously.	#Rcp	<i>Sultáni se vystřídali {#Rcp.PAT} na trůnu. (=lit. Sultans REFL changed on throne.) Fig. 9.5</i>

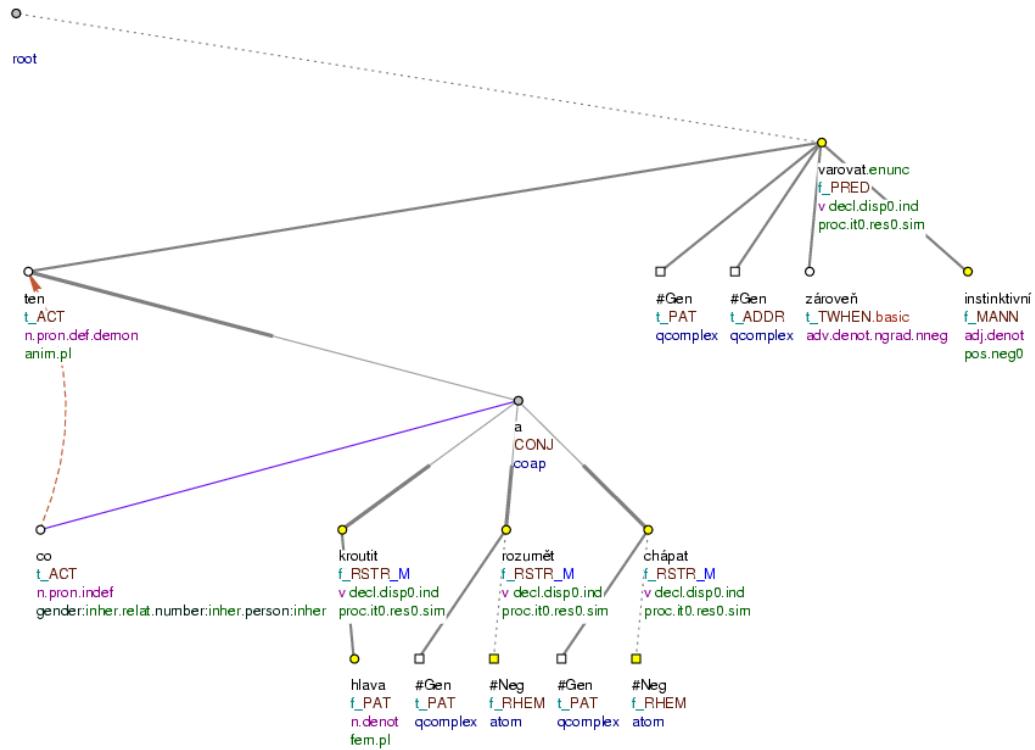
**Figure 9.1. Coreference with reflexive pronouns**

*Neschopnost opozičních stran vzdorovat své vlastní lenosti. (=lit. Inability (of) oppositional parties to\_resist self's own laziness.)*

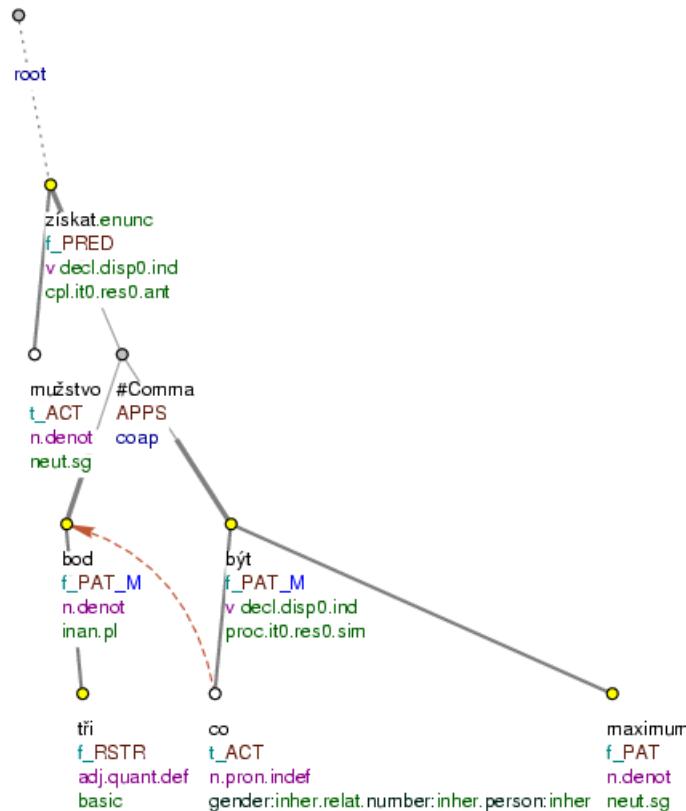


**Figure 9.2. Coreference in relative clauses**

*Ti, co kroutí hlavami, nerozumí a nechápou, zároveň instinktivně varují. (=lit. Those, who shake heads, not\_understand and not\_comprehend, simultaneously instinctively warn.)*

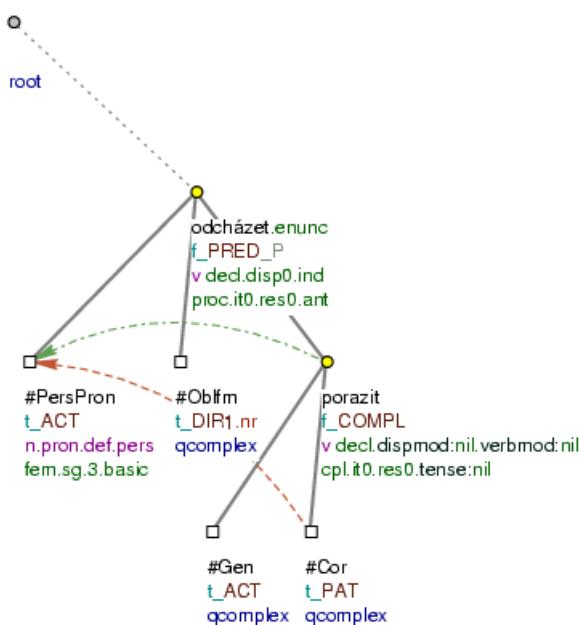
**Figure 9.3. Coreference of the relative element což**

*Mužstvo získalo tři body, což je maximum. (=lit. Team got three points, which is maximum.)*



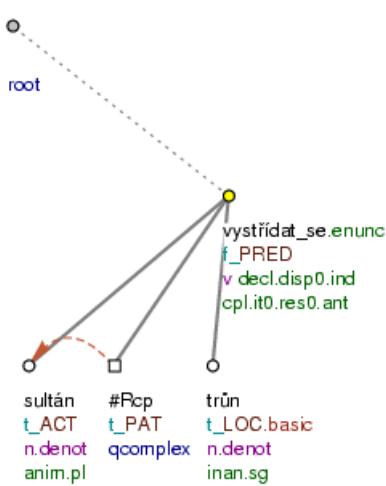
**Figure 9.4. Coreference with the predicative complement expressed by a verb form**

*Odcházela poražena. (=lit. (She) was\_leaving defeated.)*



**Figure 9.5. Coreference in reciprocal constructions**

*Sultáni se vystřídali na trůnu. (=lit. Sultans REFL changed on throne.)*



## 9.2.1. Control

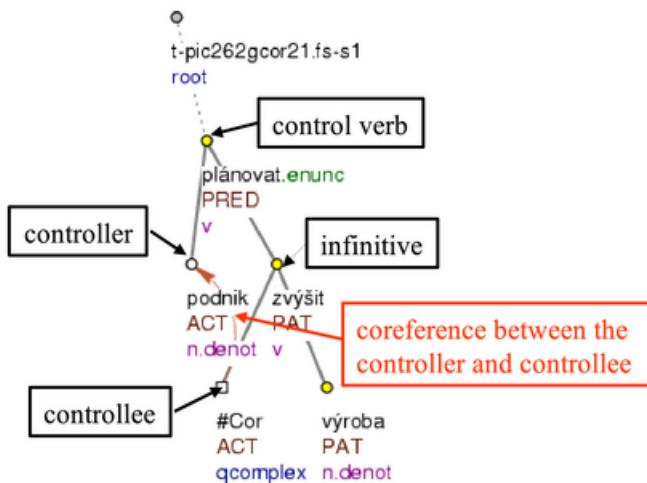
**Control** is a type of grammatical coreference that arises with certain verbs, called **control verbs**. It is an either obligatory or optional coreference relation between the controller and controllee and it holds that:

- **controlling element (controller)** is one of the members of the valency frame of the governing control verb: ACT, PAT, ADDR, ORIG, or an obligatory adjunct with the LOC functor; in certain constructions in which the infinitive depends on the verbal part of a verbonominal (control) predicate as its Actor, the controller is an adjunct with the BEN functor.
- An adjunct as a controller is a specific type of modification. The node for such a modification is always inserted in the tree if not present at the surface level. It is the only case when a non-obligatory adjunct is added to a tectogrammatical tree. Since it is always a Beneficiary, a special t-lemma substitute was introduced for such a controller: #Benef.
- **controlled element (controllee)** is a member of the valency frame of the infinitive (or deverbal noun) dependent on the control verb. It is usually the non-expressed subject of the infinitive (i.e. the Actor with active infinitives and Patient or Addressee with passive infinitives). The controllee's reference is obligatorily identical to that of the controller and it cannot be expressed at the surface level.
  - **infinitive (or a deverbal noun), a valency modification (subject) of which is controlled** is a valency (or typical) modification of the control verb. Usually, it has an argument functor (most often PAT), however, adjunct functors are sometimes also possible (most often INTT).

It is possible to find control also in constructions in which the control verb or the infinitive is nominalized. Nominalized control verbs are **control nouns** and **control adjectives**.

**Representing control in the tectogrammatical trees.** The controllee is assigned the t-lemma substitute #Cor. The grammatical coreference between the controller and controllee is marked in the following way: the coref\_gram\_rf attribute of the controllee contains the identifier of the controller.

For an illustration of the way control is represented in tectogrammatical trees, see Fig. 9.6.

**Figure 9.6. Structure of control constructions**

*Podnik plánoval zvýšit výrobu.  
(=lit. Company planned to \_increase production.)*

Examples:

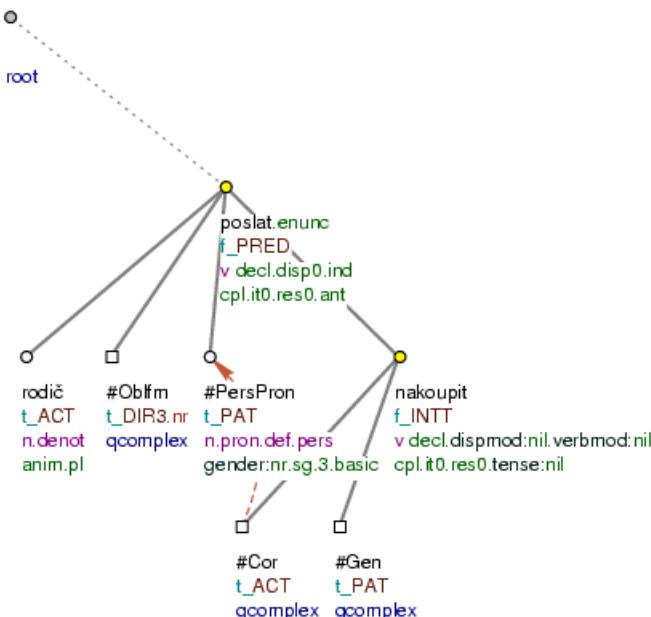
*Rodiče ho.PAT poslali {#Cor.ACT} nakoupit. (=His parents told him to go shopping.) Fig. 9.7*

*Je škoda {#Benef.BEN} {#Cor.ACT} ochudit se o tolik vzácnych látek. (=It is a pity to be losing so many valuable substances.) Fig. 9.8*

*{#PersPron.ACT} Potřebujete poradit {#Cor.ADDR} ? (=Do you need a piece of advice?) Fig. 9.9*

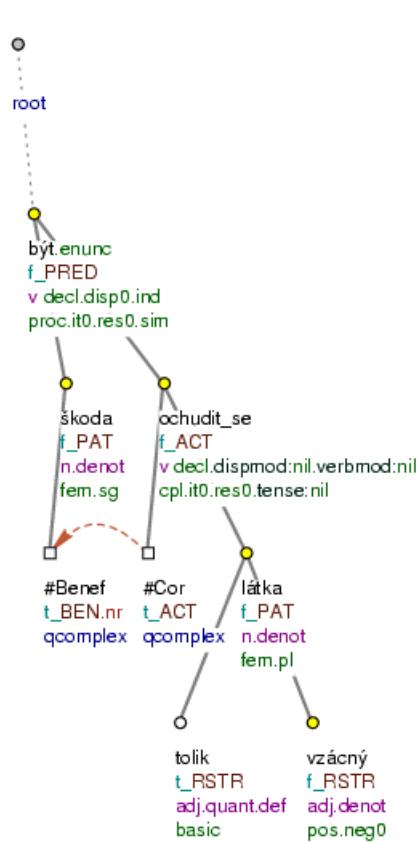
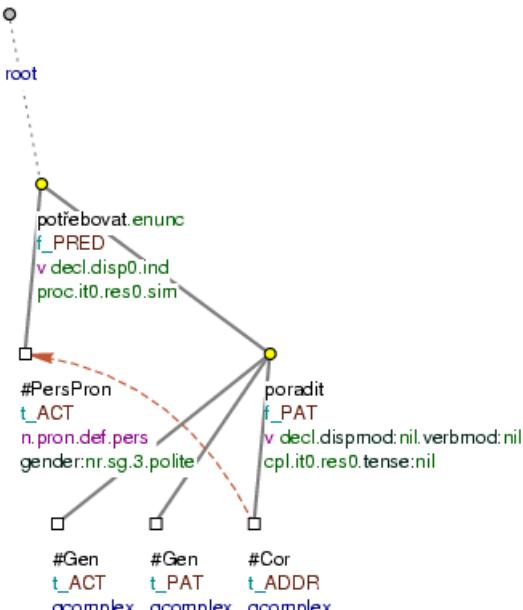
**Figure 9.7. Control**

*Rodiče ho poslali nakoupit. (=lit. Parents him sent to \_shop.)*



**Figure 9.8. Control**

*Je škoda ochudit se o tolik vzácných látek. (=lit. Is pity Potřebujete poradit? (=lit. (Do you) need to\_advise?) to\_lose REFL - so\_many valuable substances.)*

**Figure 9.9. Control**

**Control verbs.** Both one-word and multi-word predicates can be control predicates (see Section 7.1.1, “Multi-word predicates”). As for multi-word predicates, there are some complex predicates that belong to control predicates (see Section 7.1.1.4, “Complex predicates”), as well as some verbonominal predicates (see Section 7.1.1.5, “Verbonominal predicates (the copula *být*)” and verbal idioms (phrasemes) (see Section 7.1.2, “Idioms”). Control is given by the lexical properties (meaning) of the main verb. It is necessary to distinguish between the individual meanings of the verb in question. Control is always connected to certain meanings of the verb (represented by different valency frames).

Cf.:

*Dala dítě spát.* (=She put the child to sleep.)

The valency frame for this meaning of *dát*: ACT(.1) PAT(.4) EFF(.f) In this meaning, *dát* is a control verb. The Actor of the infinitive *spát* is controlled by the Patient of the verb *dát*.

*Dala dítě léčit.* (=She had the child treated/sent the child to the doctor.)

The valency frame for this meaning of *dát*: ACT(.1) PAT(.f) In this meaning, the verb *dát* is not a control verb (the noun in the accusative belongs to the valency frame of the dependent infinitive).

**Subject of the infinitive.** The controllee is usually the non-expressed subject of the infinitive. The controllee has obligatorily the same reference as the controller and it is also obligatorily non-expressed. Therefore, not every subject of an infinitive is a controllee. Non-expressed subjects of infinitives can also have the t-lemma substitute #Gen or #PersPron. Moreover, subjects of infinitives can sometimes be expressed at the surface level. For a survey of possible t-lemmas of expressed as well as non-expressed subjects of infinitives, see Table 9.2, “**Subjects of infinitives: possible t-lemmas**”.

**Table 9.2. Subjects of infinitives: possible t-lemmas**

T-lemma	Surface form	Definition	Example
#Cor	cannot be expressed	The subject of the infinitive is in a control relation with a modification of the governing verb	Rodiče <i>hoposlali</i> {#Cor.ACT} <i>nakoupit.</i> (=His parents told him to go shopping.)
#Gen	not expressed	The subject of the infinitive is a general argument (see Section 6.6.1.2, “Ellipsis of a dependent meaning unit”).	Je co číst {#Gen.ACT} . (=There are things to read.)
#PersPron	not expressed	It is possible to find the coreferred element (antecedent) of the subject but it is not grammatical but rather textual coreference (see Section 9.3, “Textual coreference”).	{#PersPron.ACT} Čte si. (=He is reading.)
noun #PersPron	expressed	Specific cases of infinitives expressing a condition.	Já tam být, nestalo se to. (=If I had been there this wouldn't have happened.)

**NB!** The t-lemma substitute #Cor was introduced primarily for the controllee in control constructions. So far, it has been used also for non-expressed subjects of non-finite verb forms in constructions with dual dependency. Such constructions involve grammatical coreference but not control since the dependent non-finite verb form is neither a valency nor typical modification of the governing verb.

### 9.2.1.1. Types of control constructions

In most cases of control constructions, it is possible to nominalize the governing verb, as well as the infinitive. The following types of control constructions can be found:

1. an infinitive depends on a verbal control predicate.  
Subtypes:  
an infinitive depends on the nominal part of a complex control predicate,  
an infinitive depends on a verbal idiom (which is a control predicate),  
an infinitive depends on the non-verbal part of a verbonominal control predicate,  
an infinitive depends on the verbal part of a verbonominal control predicate,  
an infinitive depends on the predicate *lze*; control in the constructions of the type *Je vidět Sněžku*.
2. an infinitive depends on the nominalized control predicate, i.e. on a deverbal noun or adjective.
3. a deverbal noun (a modification of which is controlled) depends on a verbal control predicate.
4. a deverbal noun (a modification of which is controlled) depends on a nominalized control predicate.

Most control verbs can be found in all the four types of construction. Cf.:

*Slíbil napsat dopis.* (=He promised to write a letter.)

The infinitive depends on a verbal control predicate (type 1).

*slib napsat dopis* (=the promise to write a letter)

The infinitive depends on a nominalized control predicate, i.e. on a deverbal noun (type 2).

*Slíbil napsání dopisu.* (=He promised to write a letter; lit. writing.)

The deverbal noun, i.e. a nominalized infinitive depends on a verbal control predicate (type 3).

*slib napsání dopisu* (=the promise to write a letter; lit. of writing of letter.)

The deverbal noun, i.e. a nominalized infinitive, depends on a nominalized control predicate (type 4).

Some control verbs (e.g.: *stíhat* (=prosecute), *podezírat* (=suspect), *obvinit* (=accuse), *omluvit se* (=apologise), *přisoudit* (=asccribe/attribute), *osocít* (=smear/malign)) cannot be modified by an infinitive at all. However, they are modified by a prepositional phrase with a deverbal noun one valency modification of which is controlled by a valency modification of the governing verb. They can only be found in constructions of type 3 and 4; e.g.:

*Stíhají ho.PAT pro falšování.CAUS {#Cor.ACT} dokladů.* (=They are prosecuting him for falsifying the documents.)

The Actor of the noun *falšování* (which is in the position of CAUS) is controlled by the Patient of the verb *stíhat*.

In exceptional cases, no nominalization of the infinitive (a modification of which is controlled) is possible and no nominalization of the governing verb either. Such control verbs can, then, only occur as type 1, e.g.:

*Viktor se zdá být chytrý.* (=Viktor seems to be clever.)

Examples:

(My.ACT) máme záměr.CPHR {#Cor.ACT} vyklidit prostory. (=We have the intention to clear out the premises.)

Type 1: The Actor of the infinitive *vyklidit* (dependent on the nominal part of the predicate *mít záměr*) is controlled by the Actor of the verbal part of the complex control predicate (*my*), which is identical in reference with the non-expressed Actor of the noun *záměr*. Fig. 9.10

*Petr má v plánu.DPHR {#Cor.ACT} vystudovat fakultu.* (=Petr plans (lit. has in plan) to finish his studies at the faculty.)

Type 1: The Actor of the infinitive *vystudovat* (dependent on the verbal part of the idiom, i.e. on *mít*) is controlled by the Actor of the verbal part of the idiom, the noun *Petr*.

*Pavel je ochoten {#Cor.ACT} přijít.* (=Pavel is willing to come.)

Type 1: The Actor of the infinitive *přijít* (dependent on *ochoten*) is controlled by the Actor of the verbal part of the predicate, namely by the noun *Pavel*. Fig. 9.11

*Je povinností koalice.ACT {#Cor.ACT} nalézt.ACT cestu.* (=It is the coalition's duty to find a way.)

Type 1: The Actor of the infinitive *nalézt* (dependent on the verbal part of the verbonominal predicate) is represented as controlled by the Actor of the noun *povinnost* (by *koalice*).

*Je {#Benef.BEN} nutné {#Cor.ACT} přejít.ACT* (=It is necessary to cross the street.)

Type 1: The Actor of the infinitive *přejít* (dependent on the verbal part of the verbonominal predicate) is controlled by the non-expressed Beneficiary of the adjective *nutný*. Fig. 9.12

{#Benef.BEN} Je {#Cor.ACT} vidět.ACT Sněžku. (=It is possible to see Sněžka.)

Type 1: The Actor of the infinitive *vidět* (dependent on *být*) is controlled by the non-expressed Beneficiary of *být*.

*jeho.ACT odhodlání {#Cor.ACT} přijít.PAT včas* (=his determination to come in time)

Type 2: The Actor of the infinitive *přijít* (dependent on *odhodlání*) is controlled by the Actor of the noun *odhodlání*, by the pronoun *jeho*. Fig. 9.13

*student usilující {#Cor.ACT} dokončit fakultu* (=a student trying to finish the study)

Type 2: The Actor of the infinitive *dokončit* (which is the Patient of the adjective *usilující*) is controlled by the noun governing the adjective *usilující*, by the noun *student*. Fig. 9.14

*Rodiče zakázali synovi {#Cor.ACT} práci ve výškách.* (=The parents have forbidden their son to work high above the ground.)

Type 3: The Actor of the noun *práce* is controlled by the Addressee of *zakázat*, the noun *synovi*. Fig. 9.15

{#PersPron.ACT} Měl zájem o {#Cor.ACT} studium na vysoké škole. (=He was interested in studying at a university.)

Type 3: The Actor of the noun *studium* (depending on the nominal part of the complex predicate *mít zájem*) is controlled by the Actor of the verbal part of the complex predicate (identical in reference with the non-expressed Actor of the noun *zájem*).

*Petrovo úsilí o {#Cor.ACT} zajištění Pavlovy přítomnosti na semináři.* (=Petr's effort to secure (lit. (of) securing) Pavel's presence in the class)

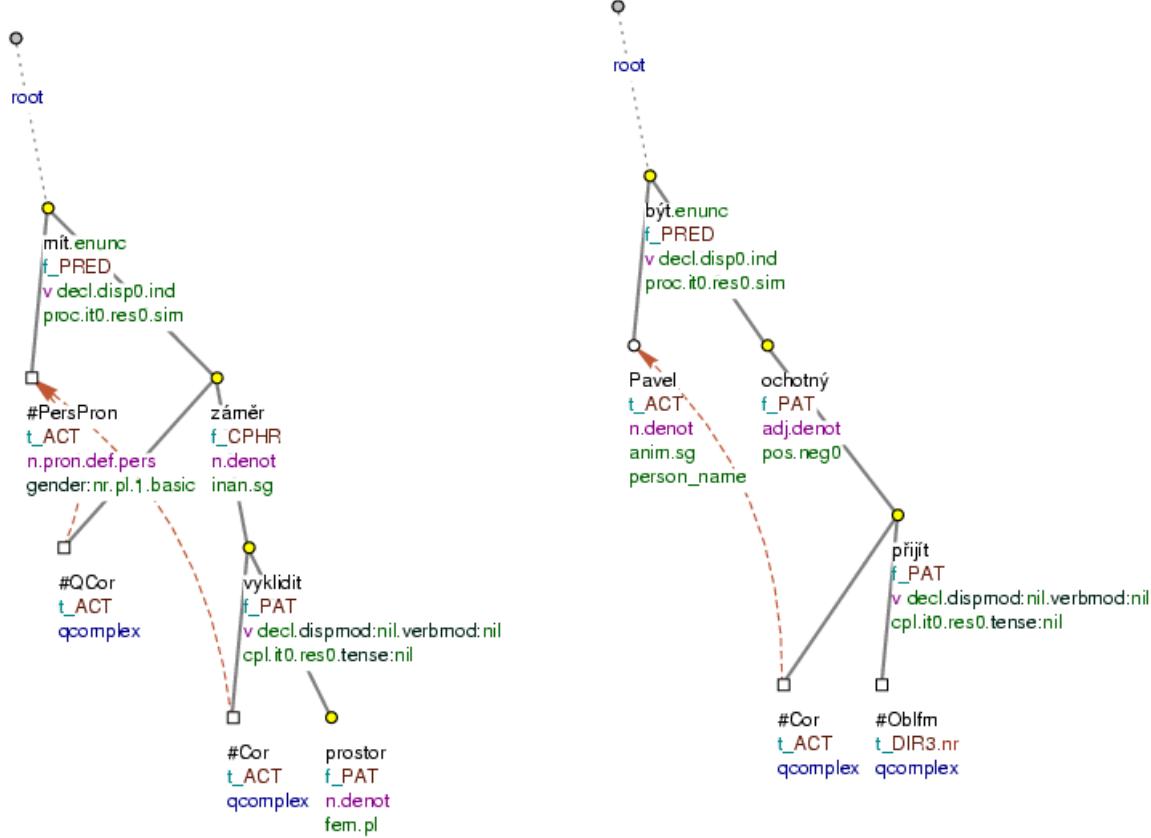
Type 4: The Actor of the noun *zajištění* is controlled by the Actor of the noun *úsilí*, by the possessive adjective *Petrovo*. Fig. 9.15

*student usilující o {#Cor.ACT} dokončení fakulty* (=a student trying to finish (lit. at finishing) the faculty)

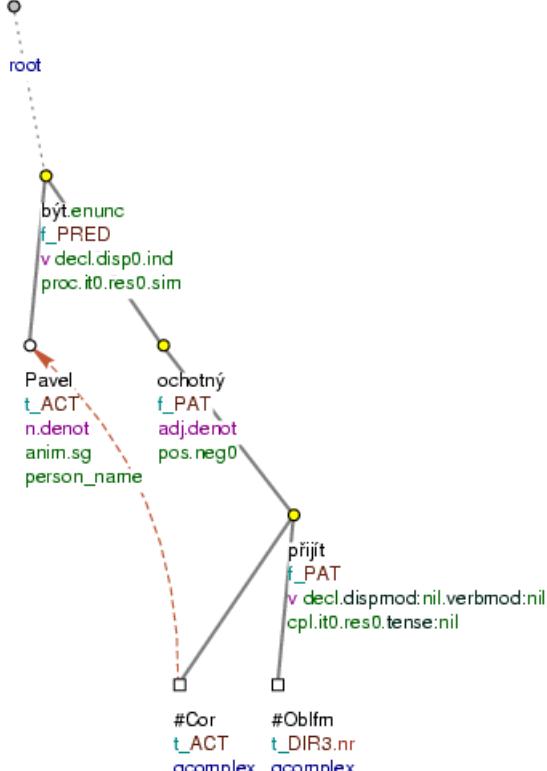
Type 4: The Actor of the noun *dokončení* is controlled by the noun modified by the adjective *usilující*, i.e. by the noun *student*.

**Figure 9.10. Type 1 control (complex control predicates)**

Máme záměr vyklidit prostory. (=lit. (We) have intention Pavel je ochoten přijít. (=lit. Pavel is willing to\_come.) to\_clear\_out premises.)

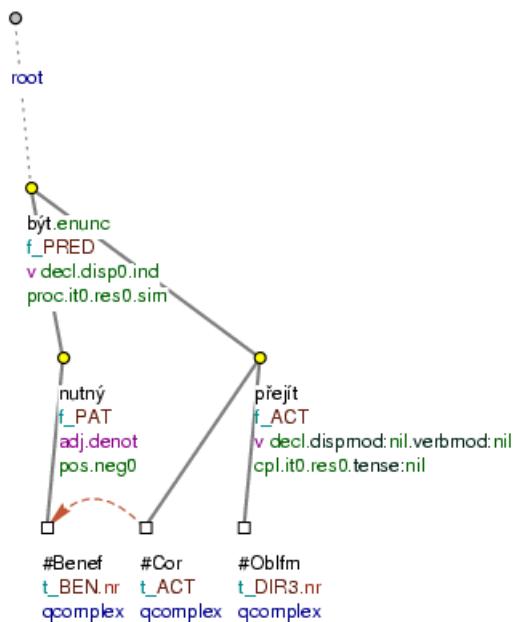


**Figure 9.11. Type 1 control (verbonominal control predicates)**



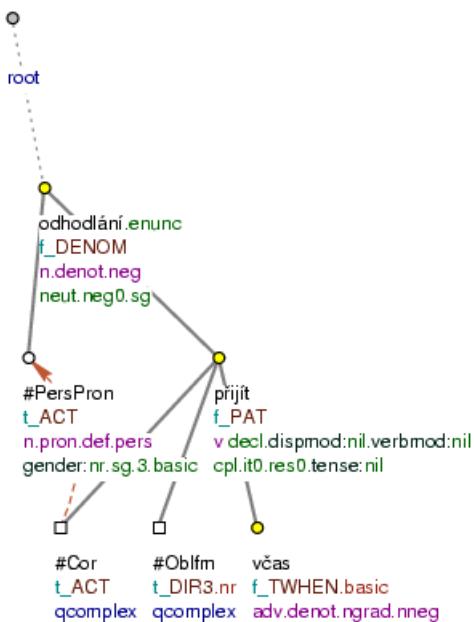
**Figure 9.12. Type 1 control (verbonominal control predicates)**

Je nutné přejít. (=lit. Is necessary to\_cross\_over.)



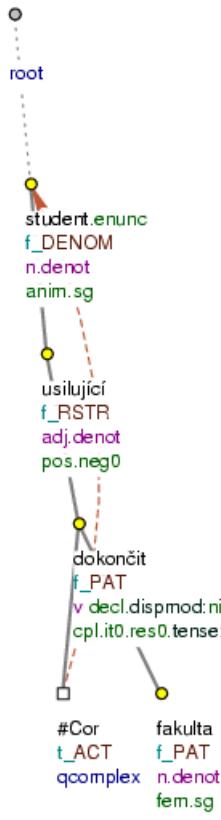
**Figure 9.13. Type 2 control (control nouns)**

jeho odhodlání přijít včas (=lit. his determination to\_come in\_time)

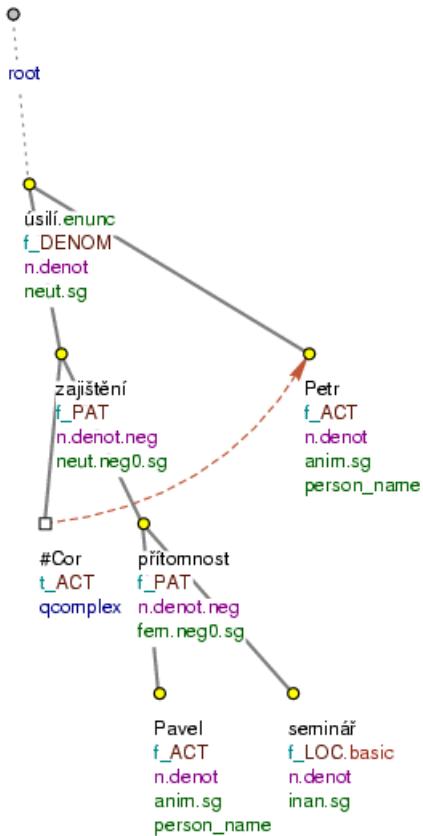


**Figure 9.14. Type 2 control (control adjectives)**

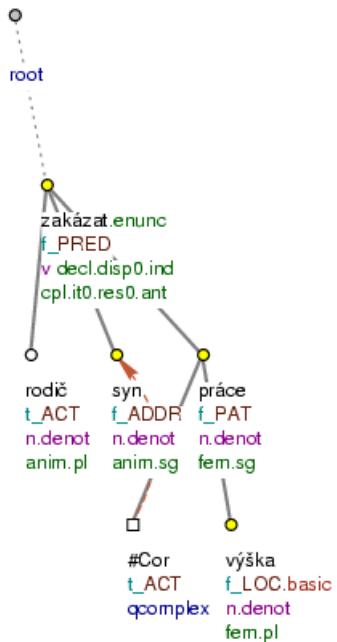
*student usilující dokončit fakultu* (=lit. student trying to \_finish faculty)

**Figure 9.15. Type 4 control (control nouns)**

*Petrovo úsilí o zajištění Pavlovy přítomnosti na semináři.*  
 (=lit. Petr's effort in securing Pavel's presence in class.)

**Figure 9.16. Type 3 control (one-word control predicates)**

*Rodiče zakázali synovi práci ve výškách.* (=lit. Parents have \_forbidden son work in heights.)



## 9.3. Textual coreference

**Textual coreference** is generally taken to mean the use of various linguistic means (pronouns, synonyms, generalising nouns etc.) which function as anaphoric (occasionally cataphoric) reference devices. This type of reference is not realised by grammatical means alone, but also on the basis of the context.

Textual coreference devices are vague by nature and the identification of a coreferred constituent based purely on context is problematical, and therefore our approach is to concentrate for the time being only on the most frequent textual coreference devices, i.e pronouns. Textual coreference is marked with:

- a. personal and possessive third person pronouns; first and second person pronouns are set aside. (Personal and possessive pronouns are assigned a unified t-lemma, namely #PersPron, in the tectogrammatical trees.)
- b. the demonstrative pronouns *ten*, *ta*, *to*.
- c. contextual ellipsis, where a new node with the t-lemma substitute #PersPron is added to the tectogrammatical tree (textual co-reference is not marked here when the added node represents a pronoun in the first or second person).

Coreference is for the time being unrepresented with the pronominal adverbs (*tam* (=there/thither), *sem* (=here/hither), *tady* (=here), *tak* (=thus) etc.) and with other pronominal expressions.

**Types of textual coreference.** For a survey of types of textual coreference see Table 9.3, “**Types of textual coreference**”. The t-lemma of the coreferring node in the case of textual coreference is always #PersPron or *ten*.

**Table 9.3. Types of textual coreference**

Type	Definition of the coreferred element	Example
<b>Explicit coreferred element</b>	particular subtree or leaf in the tree for the given sentence or some of the preceding or following ones	<i>Myslíte, že rozhodnutí NATO, zda se {#PersPron.ACT} rozšíří, či nikoli, bude záviset na postoji Ruska?</i> (=Do you think that NATO's decision whether it will expand or not will depend on Russia's attitude?) Fig. 9.17
<b>Segment</b>	a larger segment of text	<i>Rozprava o podobě reformy veřejných financí bude zahájena ve středu. Všechna jednání proběhnou za zavřenými dveřmi.</i> Lidovým novinám <i>to</i> sdělil včera ministr financí. (=The discussion... will be started on Wednesday....The minister of Finance told this to LN.)
<b>Exophor</b>	extra-textual situation	<i>V období vrcholícího léta roku 1939 již málokdo v Evropě mohl uvěřit nadějným slovům britského ministerského předsedy Chamberlaina, proneseným z balkonu Buckinghamského paláce po návratu z Mnichova: Myslím, že je <i>to</i> mír na celou naši dobu.</i> (=... I think this would mean peace...)

**Textual coreference is not represented.** Sometimes the pronouns (*ten*, *on*, *jeho*) do not corefer. In some cases, there is no coreferred element and coreference is therefore not represented with them. This holds especially for the following cases:

- a. idioms, set expressions; e.g.:

*Tak je tomu i v těch případech, kdy dosavadní domovníci užívali byty na základě dohod s bytovými podniky nebo domovními správami, podle kterých jim byl přidělen byt po dobu výkonu domovnických prací.* (=This is the case (lit. it) also when...)

- b. intenzifiers, pronouns assigned the ATT functor and pronouns used for emphasis; e.g.:

*To ale prší!* (=Boy, it's raining so much!)

- c. semantically empty uses of the pronouns. These are those cases in which the pronoun is used for emphasis or just padding; therefore, no coreferred element can be found:

*To máte těžké.* (=You know, it's hard.)

- d. the pronoun *ten* in the attributive position; e.g.:

*Tento velký problém není možné vyřešit za hodinu.* (=This problem cannot be solved in an hour.)

- e. the pronoun *ten* modified by a relative clause, e.g.:

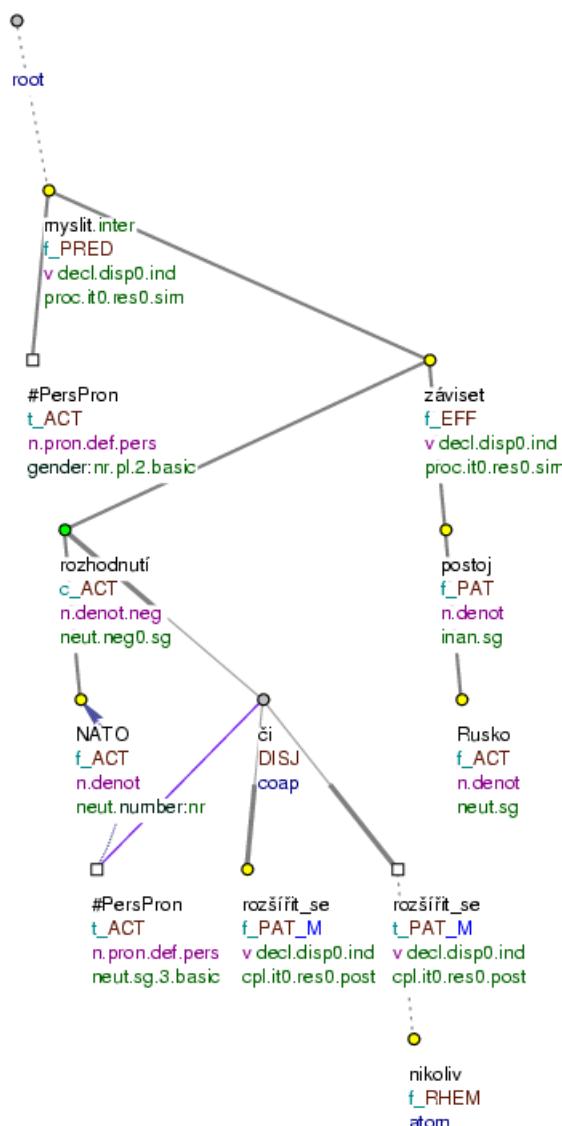
*Srovnáme-li současný plán rozvoje dálniční sítě s tím, jaký byl přijat v roce 1991, výrazně se změnil.* (=If we compare the current plan with that (=lit. that what was) accepted in 1991, the difference is big.)

- f. the pronoun *ten* modified by an adjective, e.g.:

*Je to možná tragedia dell' arte o nepřátelích, kteří jsou posedlí touhou být spolu, a dokonce být tím druhým.* (=...enemies that are obsessed by the desire to be together, even to be the other one (lit. that other).)

**Figure 9.17. Explicit coreferred element**

*Myslíte, že rozhodnutí NATO, zda se rozšíří, či nikoli, bude záviset na postoji Ruska?* (=lit. Do you think that decision (of) NATO whether (it) REFL will\_expand or not will depend on attitude (of) Russia?)



---

# Chapter 10. Topic-focus articulation

**Topic-focus articulation** (TFA) is represented by the annotation of two phenomena:

contextual boundness (see Section 10.1, “Contextual boundness”),  
communicative dynamism (see Section 10.2, “Communicative dynamism”).

TFA of the Czech sentence is signalled mainly by the surface word order and intonation.

**Surface word order.** The boundary between contextually bound and contextually non-bound expressions (see Section 10.1, “Contextual boundness”) is signalled by the position of the governing verb in the Czech surface word order. In unmarked cases (in the cases of the objective ordering), direct modifications of the governing verb appearing before it in the surface word order are contextually bound.

Particularly two types of constructions constitute exceptions to these general tendencies: constructions with the verb on “the second position” in the sentence and constructions with so-called subjective ordering. The governing verb sometimes does not appear on the boundary between contextually bound and contextually non-bound expressions (depending directly on the verb), but it appears immediately after the first sentence part, on “the second position” in the sentence (whether the boundary is or is not situated in that position). In this case contextually bound expressions can appear to the right from the verb. With the subjective ordering, modifications that are contextually non-bound can appear before the governing verb. For pragmatic reasons, the more dynamic part of the sentence is placed at the beginning.

**Intonation.** Information concerning TFA cannot always be obtained from the written context only. The annotation of TFA has to take into account also the spoken form of the sentence. In a particular context, every sentence has a natural pronunciation, and we suppose that as people can spontaneously produce sentences with appropriate word order and intonation, they are able to comparatively well assign the correct intonation to a written sentence. It is important to identify especially:

- a. **the intonation centre**, i.e. the word (prosodic unit) that carries the “sentence” stress . It is the most important prosodic unit of the sentence, usually placed at the end. It is characterized by a falling pitch contour and increased volume, but it is constituted also by other factors (rate of speech, voice timbre and others). Every complete sentence contains an intonation centre.

We assume that the intonation centre in Czech sentences signals its focus proper (see Section 10.2.1, “Basic rules for the nodes ordering in tectogrammatical trees”). If the sentence ends with a noun phrase the intonation centre can be placed on its last member instead of the focus proper.

- b. **contrastive stress**, i.e. a specific stress characterized by a rising pitch contour. In the sentence, it signals contrastive contextually bound expressions. Contrastive contextually bound expressions do not have to be signalled by a contrastive stress, contrastive stress is optional (the presence of contrastive stress is governed by other factors, primarily by the rate of speech and the carefulness in pronunciation).

We assume that an expression on which a contrastive stress can be placed in the spoken form of a sentence is contrastive contextually bound.

## 10.1. Contextual boundness

**Contextual boundness** is a property of an expression (be it expressed or absent in the surface structure of the sentence) which determines whether the speaker (author) uses the expression as given for the recipient, i.e. uniquely determined by the context.

**Context.** Context is in the annotation of contextual boundness understood very widely. Context comprises not only immediate textual context (“cotext”), but also wider contextual layers, including all shared or commonly known information, whose sharing may be conditioned by the situation, perception, culture, other texts, or other factors.

**Representing contextual boundness.** The information on the contextual boundness of individual expressions is contained in the attribute `tfa` (see Section A.2.21, “`tfa`”). Every node (relevant for the topic-focus articulation of the sentence) is assigned one of three possible values of the attribute `tfa` (`t`, `f`, or `c`).

The attribute `tfa` is not relevant for:

the technical root node of the tectogrammatical tree (`nodetype = root`),  
paratactic structure root nodes (`nodetype=coap`),  
nodes with the CM functor,  
nodes with the FPHR functor (`nodetype = fphr`).

**Types of expressions on the basis of contextual boundness.** According to their contextual boundness or non-boundness, we distinguish three types of expressions:

- a. **non-contrastive contextually bound expressions** ( $tfa = t$ ), i.e. expression (both expressed and absent in the surface structure of the sentence) that represent some “known information”. Such expressions are repeated from the preceding text (not necessarily verbatim), they are deducible from it (e.g. using coreferential or inferential relationships), or somehow related to a wider context.
- b. **contrastive contextually bound expressions** ( $tfa = c$ ), i.e. expressions that can be identified on the basis of the following properties:
  - such an expression is usually a choice from a set of alternatives. This set need not be explicitly specified in the text. A contrastive contextually bound expression can refer to a larger text segment and does not have to be deducible from the immediately preceding textual context.
  - the occurrence of a contrastive contextually bound expression is primarily determined by the thematic progression of the text. Contrastive contextually bound expressions usually occur in enumerative passages, at the beginning of paragraphs etc.
  - in the spoken form of the utterance the contrastive contextually bound expression carries an optional contrastive stress.
- c. **contextually non-bound expressions** ( $tfa = f$ ), i.e. expressions (both expressed and absent in the surface structure of the sentence) that represent in the text some unknown, new facts, or introduce known facts in new relations, i.e. they express information not deducible from the context. Contextually non-bound expressions can carry the intonation centre of the sentence.

Examples:

(Dnes nesu jen dopis.) *Tu knihu* [ $tfa = c$ ] *ti přinesu zítra* [ $tfa = f$ ] (=Today I'm only bringing you a letter)  
As for the book, I'll bring it tomorrow.)

*Janu* [ $tfa = c$ ] *Marie* [ $tfa = t$ ] *neviděla* [ $tfa = f$ ] (=As for Jana, Marie didn't see her.)

**Contextual boundness of expressions absent in the surface structure of the sentence.** Certain lexical units are absent in the surface structure of the sentence precisely because they are considered to be deducible from the context. Newly established nodes (representing elements not expressed in the surface structure of the sentence) are therefore in most cases assigned the  $tfa$  value  $t$ .

**Exceptions:** Newly established nodes can be assigned the  $tfa$  value  $f$  or  $c$  in the following cases:

- a. contextual ellipsis of the governing noun in sentence-part paratactic constructions and in the case of contextual ellipsis of the governing noun in binary relations of the type “from-to”. The first occurrence of a noun absent in the surface structure of the sentence can be contextually non-bound or contrastive contextually bound (depending on the context), while its second occurrence expressed in the surface structure is contextually bound. For example:

*Pil červené* { *víno* [ $tfa = f$ ] } *a bílé víno* [ $tfa = t$ ] (=He was drinking red and white wine) Fig. 10.1

*Přemaloval to z černé* { *barvy* [ $tfa = f$ ] } *na červenou barvu* [ $tfa = t$ ] (=He repainted it from black to red color.)

- b. ellipsis of a noun following an expressed preposition. For example:

*Neexistuje argument <pro>*. {#PersPron.PAT [ $tfa = f$ ] } (=There is no pro argument.)

- c. contextual ellipsis of a governing node when the copied node is different from the original node in a value of some grammateeme. For example:

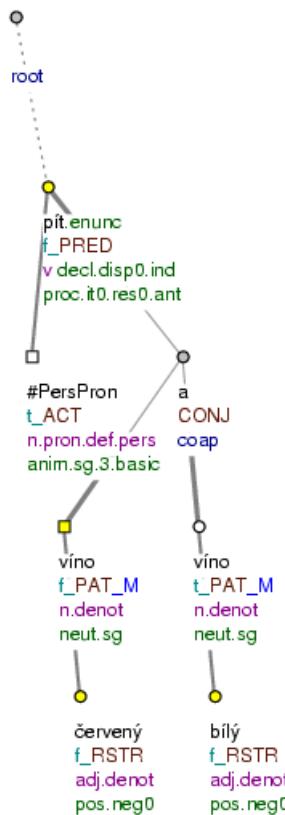
*Nechtěli nebo nemohli odklad platby povolit.* (=They didn't want to and couldn't allow the payment postponement.)

= *Odklad platby* { *nechtěli povolit* [deontmod = vol;  $tfa = f$ ] } *nebo nemohli povolit* [deontmod = poss;  $tfa = f$ ]

- d. newly established node for syntactic negation, represented as a rhematizer ( $t\_lemma = \#Neg$  and  $functor = RHEM$ ).
- e. newly established node for the root of a list structure for foreign-language expressions ( $t\_lemma = \#Forn$ ).

**Figure 10.1.** Contextually non-bound expression absent in the surface structure of the sentence

*Pil červené a bílé víno. (=lit. (He) was\_drinking red and white wine.)*



## 10.2. Communicative dynamism

**Communicative dynamism** is a property of an expression that reflects its relative degree of importance in comparison to other expressions in the sentence attributed to it by the speaker; we consider contextually non-bound expressions to be more dynamic than expressions contextually bound (be they non-contrastive or contrastive).

In tectogrammatical trees, communicative dynamism is represented by the so called underlying word order. Information about the underlying order of the nodes is stored in the attribute `deepord` (see Section A.2.6, “`deepord`”).

The degree of communicative dynamism is always determined with respect to the governing node and the sister nodes, i.e. for each level of the tectogrammatical tree. Nodes on the individual levels of the tree are ordered according to the increasing communicative dynamism (see Section 10.2.1, “Basic rules for the nodes ordering in tectogrammatical trees”). Also the relative order between individual (sister) contextually bound and contextually non-bound nodes is set. The order is naturally different in verb phrases (see Section 10.2.2, “Ordering of nodes in verb phrases”) and noun phrases (see Section 10.2.3, “Ordering of nodes in noun phrases”). The same rules as for nodes in verb phrases hold also for nodes in adjective phrases.

### 10.2.1. Basic rules for the nodes ordering in tectogrammatical trees

There are three basic rules for the ordering of nodes in a tectogrammatical tree:

1. nodes representing contextually bound expressions (nodes with the values `t` or `c` in the attribute `tfa`) are placed to the left from their governing node and nodes representing contextually non-bound expressions (nodes with the value `f` in the attribute `tfa`) are situated to the right from their governing node.

**Exceptions.** There are only a few exceptions to this rule:

- a. a node representing the quasi-focus (for the definition see below) depends to the right from its governing node.

- b. a node representing a rhematizer whose scope contains its governing verb which is contextually non-bound has the  $tfa$  value  $f$ , but is placed to the left from the node representing the verb (see also Section 10.4, “Rhematizers”).
  - c. the effective root of a syntactically non-incorporated parenthesis ( $functor = PAR$ ) has usually the  $tfa$  value  $f$ , but it stays at the same place as in the surface word order, i.e. even if it is to the left from its governing node.
  - d. nodes representing predicates of certain types of dependent clauses (mostly causal) can have the  $tfa$  value  $f$  and at the same time be to the left from the governing predicate node (see also Section 10.2.1.1, “Paratactic structures and dependent clauses”).
  - e. in the case of ellipsis of the governing noun in sentence-part paratactic constructions (of the type “červené a bílé víno (=red and white wine)”), the second (expressed) noun has the  $tfa$  value  $t$ , but it still depends to the right from the paratactic structure root even if the  $tfa$  value of the first member of the paratactic structure is  $f$  (see Section 10.1, “Contextual boundness”).
2. in the underlying word order, the focus proper is placed on the rightmost path from the effective root of the tectogrammatical tree, even if it is at a different position in the surface structure. If the focus proper is constituted by an expression represented as the effective root of the tectogrammatical tree (i.e. the governing predicate is the focus proper), there is no right path from the effective root.
  3. the tectogrammatical tree is projective (see Section 10.2.4, “Projectivity of tectogrammatical trees”).

Examples:

*Černý [tfa = f] kocour [tfa = t] se napil [tfa = f] ze své [tfa = t] misky [tfa = f]* (=The black cat drank from his bowl.) Fig. 10.2

*Taky [tfa = f] KAREL [tfa = f] se doma [tfa = t] ukázal [tfa = t]* (=Also Karel has shown up at home.) Fig. 10.3

**Focus proper** is the most dynamic, communicatively significant contextually non-bound part of the sentence. In the spoken form of a sentence, the focus proper carries the intonation centre. For example:

*(Mám rád červené tulipány, ale) v Holandsku jsem viděl i tulipány [tfa = t] modré [tfa = f]* (=I like red tulips) In Holland I have also seen blue ones.) Fig. 10.4

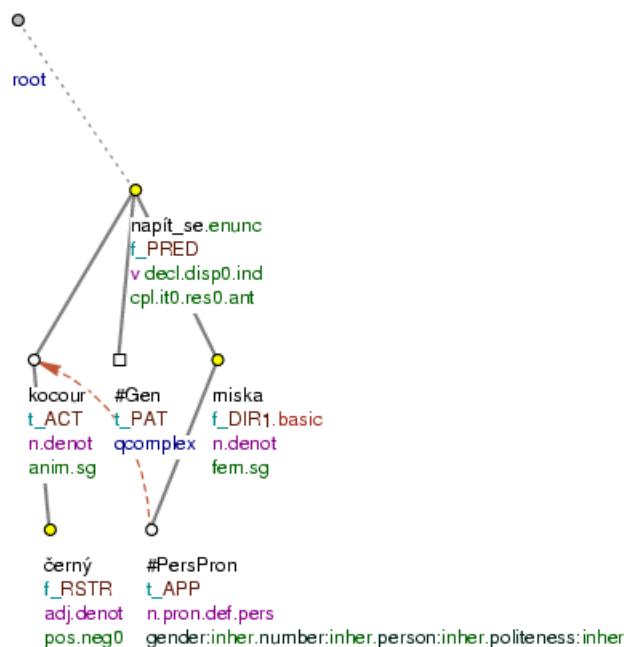
If a noun phrase is the focus proper, the intonation centre is usually placed at the last word of the noun phrase, even if it is not its most dynamic member. This is caused by the highly grammaticalized word order in noun phrases (see also Section 10.2.3, “Ordering of nodes in noun phrases”).

**Quasi-focus** is constituted by (both contrastive and non-contrastive) contextually bound expressions, to which the focus proper is subordinated. The focus proper can immediately depend on the quasi-focus, or it can be a more deeply embedded expression. For example:

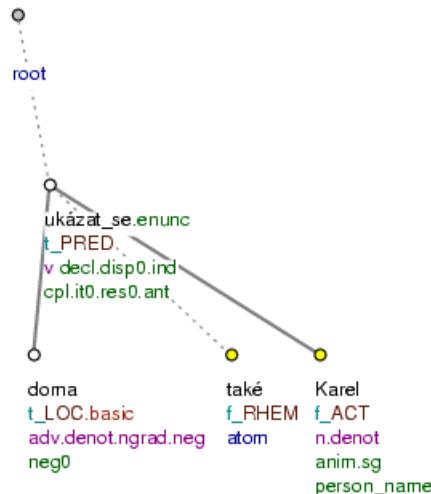
*(Kterého učitele jsi potkal?) Potkal jsem učitele [tfa = t] chemie [tfa = f]* (=Which teacher did you meet?) I met the chemistry teacher.) Fig. 10.5

**Figure 10.2. Nodes ordering in tectogrammatical trees (objective ordering)**

*Černý kocour se napil ze své misky.* (=lit. *(The) black tomcat REFL drank from its bowl.*)

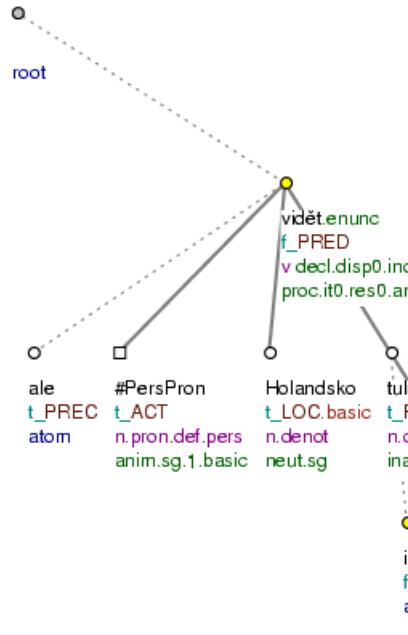
**Figure 10.3. Nodes ordering in tectogrammatical trees (subjective ordering)**

*Taky Karel se doma ukázal.* (=lit. *Also Karel REFL home showed\_up.*)

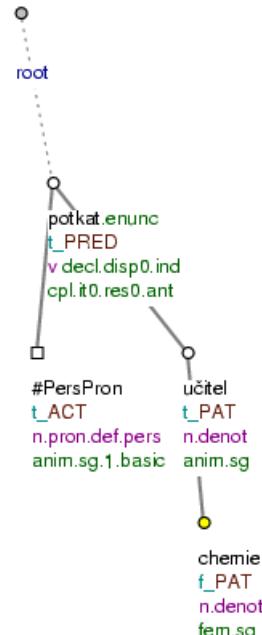


**Figure 10.4. Focus proper**

(Mám rád červené tulipány.) Ale v Holandsku jsem viděl i tulipány modré. (=lit. (I) like - red tulips) But in Holland AUX chemie. (=lit. (Which teacher AUX (you) met?) (I) met AUX teacher (of) chemistry.)


**Figure 10.5. Quasi-focus**

(Kterého učitele jsi potkal?) Potkal jsem učitele chemie. (=lit. (Which teacher AUX (you) met?) (I) met AUX teacher (of) chemistry.)



### 10.2.1.1. Paratactic structures and dependent clauses

**Placement of the shared modifier.** The effective root of a shared modifier (i.e. the whole subtree) of a paratactic structure is placed as the leftmost immediate daughter of the root of the paratactic structure if the effective root of the shared modifier is contextually bound; if the effective root of the shared modifier is contextually non-bound, it is placed as the rightmost immediate daughter of the root of the paratactic structure. For example:

Ten nůž má nerezovou [tfa = f] čepel [tfa = f] i rukojet' [tfa = f] (=The knife has both a stainless blade and handle.) Fig. 10.6

If there are more than one shared modifiers, the rules for the ordering of nodes in verb phrases apply (if verbs or adjectives are concerned; see Section 10.2.2, “Ordering of nodes in verb phrases”), or the rules for the ordering of nodes in noun phrases apply (if nouns are concerned; see Section 10.2.3, “Ordering of nodes in noun phrases”).

**Topic-focus articulation of paratactically connected independent clauses.** Paratactically connected independent clauses are considered to have separate topic-focus articulation. Expressions in the second clause may be contextually bound with respect to the preceding clause (if they are repeated or linked to expressions in the first clause). The effective roots of paratactically connected independent clauses can have differing tfa values. For example:

Tom [tfa = c] přinesl [tfa = f] knihy [tfa = f] a pak [tfa = t] Jirka [tfa = c] odnesl [tfa = f] noviny [tfa = f] (=Tom brought some books and then Jirka took away the newspaper.) Fig. 10.7

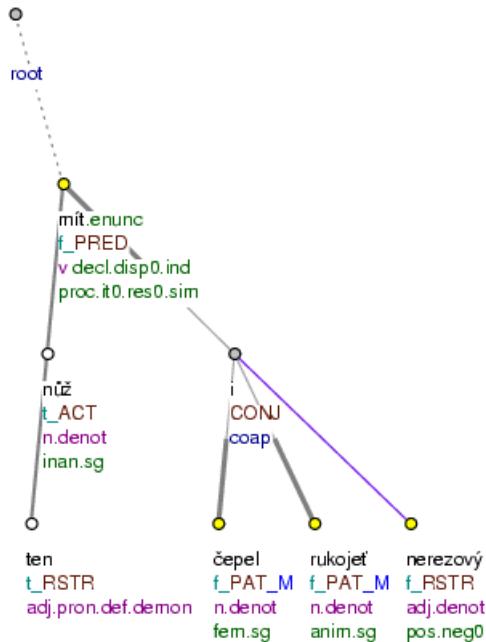
**Topic-focus articulation of dependent verbal clauses.** The placement of a dependent clause reflects in most cases the contextual boundness of the clause: if the governing predicate of the dependent clause is contextually bound, the clause is to the left from the governing verb; if the predicate is contextually non-bound, it is to the right from the governing verb. Some complex sentences with dependent adverbial clauses (esp. causal and temporal) are more like compound sentences (paratactic connection); then it holds that:

- if the dependent adverbial clause introduces a relatively independent new event and if it precedes the governing clause (for semantic reasons, like temporal or causal relations between the events), the subtree for the dependent clause stays to the left from the governing node and its effective root has the tfa value f. For example:

*Jestliže se nám podaří[t fə = f] zasadit strom, můžeme se těšit [t fə = f] na jablka. (=If we manage to plant a tree, then we can look forward to the apples.) Fig. 10.8*

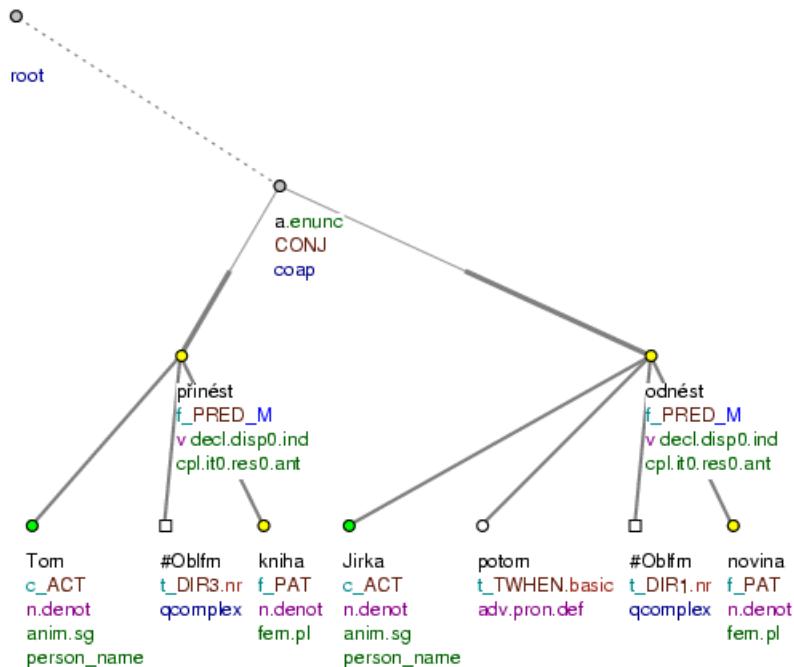
**Figure 10.6. Placement of the shared modifier**

*Ten nůž má nerezovou čepel i rukojet'. (=lit. The knife has stainless blade as\_well\_as handle.)*



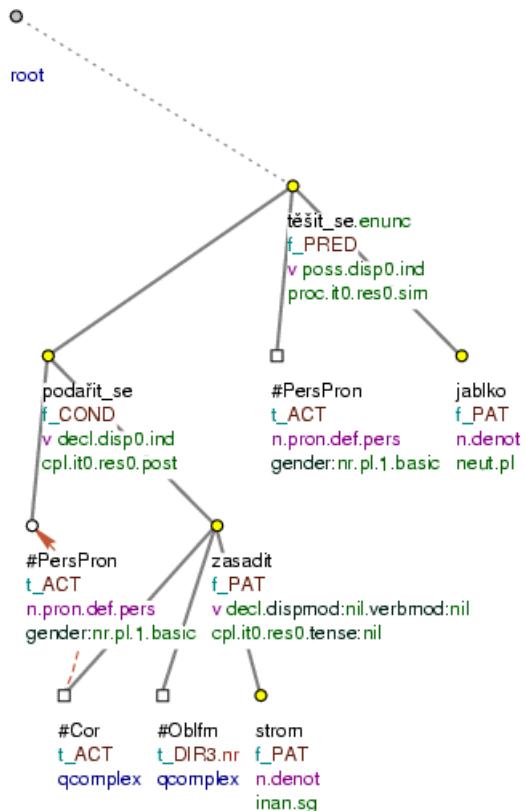
**Figure 10.7. Topic-focus articulation of paratactically connected independent clauses**

*Tom přinesl knihy a pak Jirka odnesl noviny. (=lit. Tom brought books and then Jirka carried\_away newspapers.)*



**Figure 10.8. Topic-focus articulation of a dependent adverbial clause**

*Jestliže se nám podaří zasadit strom, můžeme se těšit na jablka.* (=lit. If REFL we manage to \_plant (a) tree, we can look\_forward to apples.)



## 10.2.2. Ordering of nodes in verb phrases

**Ordering of contextually non-bound nodes.** Nodes representing contextually non-bound expressions (nodes with the value  $f$  in the attribute  $tfa$ ) are ordered according to the surface word order, so that potential deviations from systemic ordering can be easily detected and their causes studied.

**Exception.** The only exception is the node for the focus proper. If the focus proper is immediately dependent on the verb, its node is always placed rightmost in the underlying order, even though it is in a different position in the surface structure (see also Section 10.2.1, “Basic rules for the nodes ordering in tectogrammatical trees”).

**Ordering of contextually bound nodes.** The communicative dynamism of (both contrastive and non-contrastive) contextually bound nodes ( $tfa = t$  or  $tfa = c$ ) is given by their function in the topic-focus articulation. The individual types of contextually bound nodes are ordered (according to their increasing communicative dynamism) in the following way:

<b>1st node</b>	is_generated = 0	$tfa = t$	functor = VOCAT	
<b>2nd node</b>	is_generated = 0	$tfa = t$	functor = PREC	
<b>3rd node</b>	is_generated = 0	$tfa = t$	functor = ATT	
<b>4th node</b>	is_generated = 0	$tfa = t$	functor = RHEM	if a node of type 5 follows
<b>5th node</b>		$tfa = c$		with the exception of types 8 and 9
<b>6th node</b>	is_generated = 0	$tfa = t$		
<b>7th node</b>	is_generated = 1			
<b>8th node</b>	is_generated = 0	$tfa = t$	$t\_lemma = \#PersPron$	
<b>9th node</b>	is_generated = 0		locative or temporal functor	

If there are more nodes of a single type present, the nodes are ordered as in the surface structure.

Example:

*Jirka ho totiž bohužel včera v Praze prodal.* (#Gen.ADDR) *(=Jirka unfortunately sold it in Praha yesterday.)*

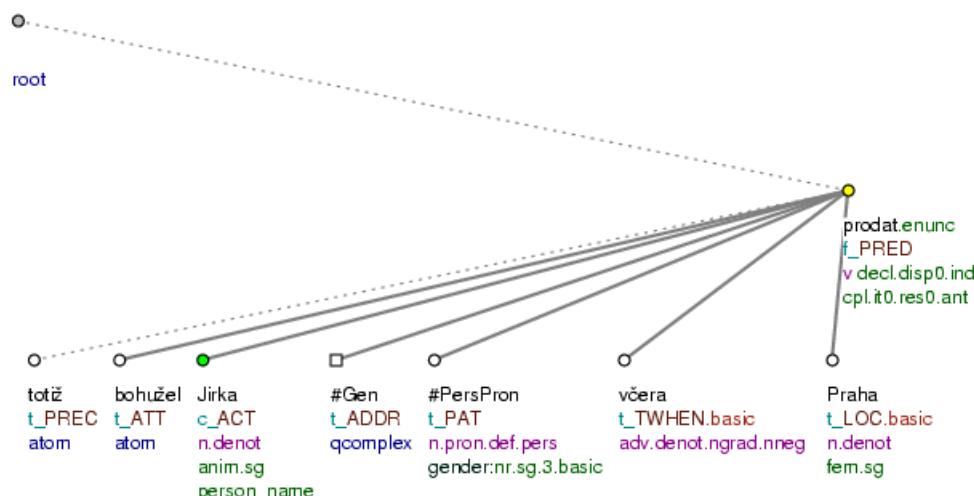
The contextually bound expressions will be ordered in the following way:

1. the node with the functor PREC representing *totiž*,
2. the node with the functor ATT representing *bohužel*,
3. the contrastive contextually bound node representing the expression *Jirka*,
4. the newly established node representing the non-expressed Addressee,
5. the expressed node with the t-lemma substitute #PersPron representing the pronoun *ho*,
6. the node representing the temporal modification *včera*,
7. the node representing the locative modification *v Praze*.

Fig. 10.9.

**Figure 10.9. Ordering of contextually bound nodes in verb phrases**

*Jirka ho totiž bohužel včera v Praze prodal.* (=lit. *Jirka it in\_fact unfortunately yesterday in Praha sold.*)



### 10.2.3. Ordering of nodes in noun phrases

The surface word order in noun phrases is governed by word-order rules of Czech (e.g. agreeing attributes stand before the noun, non-agreeing attributes after it), which are not related to communicative dynamism – the word order in noun phrases is to a large extent grammaticalized. In the underlying word order, however, we reorder the modifications of a noun according to their increasing communicative dynamism.

**Ordering of contextually non-bound nodes.** Nodes representing contextually non-bound modifications of a noun ( $tfa = f$ ) are to the right from the governing node, ordered from the nodes representing tightly connected expressions to nodes representing loosely connected expressions. The types and their ordering:

<b>1st node</b>	functor = DPHR nebo functor = CPHR
<b>2nd node</b>	functor = ID
<b>3rd node</b>	argument or functor = APP
<b>4th node</b>	agreeing adjectival modifications: functor = RSTR
<b>5th node</b>	effective roots of restrictive relative clauses: functor = RSTR
<b>6th node</b>	non-valency non-agreeing modifications
<b>7th node</b>	effective roots of non-restrictive relative clauses: functor = RSTR

**Ordering of contextually bound nodes.** When ordering nodes for contextually bound modifications of a noun ( $tfa = t$  or  $tfa = c$ ) we proceed in the same way as with contextually non-bound nodes but the order is from 7 to 1.

If there are several nodes of the same type, we order them according to the surface word order, which is free within the individual types of modifications, and thus semantically relevant. If the position of nodes in the underlying word order is unchanged with respect to the surface word order, we leave unchanged also the relative order of nodes within the individual types of nodes. If in the underlying word order, we reorder the nodes from the position before the noun to a position after it or vice versa, the underlying word order is the mirror image of the surface word order.

Example:

*Vzala jsem i ty dvoje krátké zelené šaty po sestře, které mi jsou malé.* (=I also took the two short green dresses after my sister; which are too small for me.)

The only contextually bound modification of the noun *šaty* is the demonstrative pronoun *ten*. The node representing this pronoun is placed to the left from its governing node. The remaining modifications of the noun *šaty* are contextually non-bound. The nodes representing these modifications are placed (as sister nodes) to the right from their governing node in the following order from left to right:

1. the node representing the agreeing adjectival modification *zelené*,
2. the node representing the agreeing adjectival modification *krátké*,
3. the node representing the agreeing adjectival modification *dvoje*,
4. the node representing the effective root of the restrictive relative clause,
5. the node representing the non-valency modification *po sestře*.

## 10.2.4. Projectivity of tectogrammatical trees

**Projectivity** is defined as follows: if two nodes M and N are connected by an edge and M is to the left from N, then all nodes to the right from M and to the left from N are connected with the root via a path that passes through at least one of the nodes (M or N). More briefly: between a parent and its immediate daughter there can only be daughters of the parent.

We suppose that non-projectivity in the surface structure of sentences is caused by word-order movements and that such non-projective surface realizations correspond to projective tectogrammatical structures. In the tectogrammatical annotation, we therefore projectivize such constructions. By **projectivization** we mean modifying the underlying order of a node causing a non-projectivity so that no node of the resulting tectogrammatical tree violates the definition of projectivity.

When carrying out the projectivization we take into account the motivation for the particular word-order movement. Presently we distinguish three types of motivations:

- a. **non-projectivity motivated by word-order rules.** Non-projectivity can be motivated by the fixed position of an expression in the surface word order. This applies mainly to noun phrases, which can be non-projective due to the position of a more deeply embedded modification, to dependent clauses introduced by adjectival relative elements that are moved away from their governing nouns, and to other set expressions that have frozen non-projective word-order. For example:

*Ptal jsem se ho, JAKÉHO [tfa = f] si koupil psa.* (=I asked him what kind of dog he bought.)

*Měl plné [tfa = t] kapsy peněz [tfa = f]* (=His pockets were full of money.)

*(Měl plné ruce peněz.) Ne, měl plné [tfa = t] KAPSY peněz [tfa = t]* (=His hands were full of money)  
No, his pockets were full of money.)

- b. **non-projectivity motivated by prosodic causes.** In Czech, there is a set of unstressed expressions (so-called clitics) that always take “the second position” in the sentence (i.e. the position after the first prosodic unit). In case these expressions are not immediately dependent on the governing verb of the clause, they can cause non-projectivity. Clitics are usually contextually bound, they are therefore assigned the *tfa* value *t* and placed at projective positions according to the rules for the ordering of contextually bound nodes. For example:

*Konečně [tfa = f] se to [tfa = t] podařilo [tfa = f] uskutečnit [tfa = f]* (=Finally, we managed to put it into effect.)

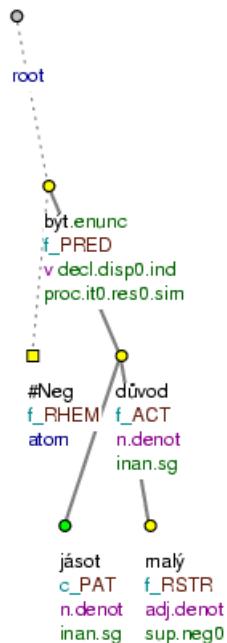
- c. **non-projectivity motivated by topic-focus articulation.** In Czech, contrastive contextually bound expressions have a strong tendency towards standing in the initial position in the sentence, that means that they move to the left although they are more deeply embedded, and so can cause non-projectivity. A node representing an

expression that is placed non-projectively to the left in the surface word-order due to its contrastive use ( $\text{tfa} = \text{c}$ ) is placed according to the rules in Section 10.2.3, “Ordering of nodes in noun phrases” and Section 10.2.2, “Ordering of nodes in verb phrases” projectively leftmost possible. For example:

*K jásotu [tfa = c] není [tfa = f] nejmenší [tfa = f] důvod [tfa = f]* (=There is not a slightest reason to be so cheerful) Fig. 10.10

**Figure 10.10. Projectivization**

*K jásotu není nejmenší důvod.* (=lit. For jubilation is \_not slightest reason.)



**Non-projectivity with unclear motivations (constructions with multi-word predicates).** In addition to cases of non-projectivity with clear motivations, there are also cases where the motivation of the word-order movement is not obvious, and furthermore in some of these cases the non-projective surface word-order is the unmarked option. This applies mainly to dependent modifications of multi-word predicates: a contextually bound expression standing to the left from a multi-word expression and dependent on the dependent part of this expression causes non-projectivity (even though it is not contrastive).

Presently we know about the following types of non-projective non-contrastive contextually bound nodes:

- the node causing non-projectivity depends on the node representing the nominal part of a complex predicate (on the node with the functor CPHR; for complex predicates see Section 7.1.1.4, “Complex predicates”),
- the node causing non-projectivity depends on the node representing the infinitive of the autosemantic verb in a modal or phase predicate represented by two nodes (see Section 7.1.1.1, “Modal predicates” and Section 7.1.1.2, “Phase predicates”),
- the node causing non-projectivity depends on the node representing the non-verbal part of a quasi-modal or quasi-phase predicate (see Section 7.1.1.3, “Quasi-modal and quasi-phase predicates”),
- the node causing non-projectivity depends on the node representing the non-verbal part of a verbonominal predicate (see Section 7.1.1.5, “Verbonominal predicates (the copula *být*)”),
- the node causing non-projectivity depends on the node representing the dependent part of an idiom (see Section 7.1.2, “Idioms”).

In order to preserve the information about the original position of the expression, the node representing this non-projective non-contrastive contextually bound expression will stay in the position corresponding to its surface word-order position, it is not moved to the right and the resulting tectogrammatical tree therefore contains a non-projective edge. For example:

(*V galerii V. Špály ode dneška vystavuje A. Born.*) *Výstavu* [ $\text{tfa} = \text{t}$ ] *je možné navštívit do dvacátého srpna.* (=In the VŠ gallery, there is an exhibition of works by AB) The exhibition is possible to see until 20th August.)

## 10.3. Topic and focus

We assume that each complete sentence divides from the point of view into two complex categories: topic and focus. The complex categories of topic and focus are not annotated explicitly. Below we describe a hypothesized procedure for determining topic and focus based on the annotated tectogrammatical trees.

**Topic** consists of that part of the sentence that connects it to the preceding context – its meaning provides linking to the expressions that have already appeared in the text, to the facts deducible from them or facts assumed to be known to the recipient. We assume that in a tectogrammatical tree, the topic consists of the following nodes:

- a. the effective root of the sentence (the node representing the governing verb) if its `tfa` value is `t`;
- b. nodes with the value `t` in the attribute `tfa` dependent on the effective root of the sentence which are not ancestors of the focus proper (see Section 10.2.1, “Basic rules for the nodes ordering in tectogrammatical trees”), and all their descendant nodes.
- c. nodes with the value `c` in the attribute `tfa` dependent on the effective root of the sentence and all their descendant nodes.

**Contrastive topic.** Connecting to the context using contrast is a specific type of contextual boundness. A certain part of the topic of a new sentence is put into contrast with a fact known from the preceding context. We call the contrastive contextually bound part of the topic contrastive topic in short. We assume that in a tectogrammatical tree, a contrastive topic consists of the following nodes:

- a. nodes with the value `c` in the attribute `tfa` and all their descendant nodes.

**Focus** consists of that part of the sentence that introduces new information not deducible from context, which is more communicatively important than the topic and cannot be omitted from the sentence. We assume that the focus is necessarily present in each sentence. We assume that in a tectogrammatical tree, the focus consists of the following nodes:

- a. the effective root of the sentence (the node representing the governing verb) if its `tfa` value is `f`;
- b. nodes with the value `f` in the attribute `tfa` dependent on the effective root and all their descendant nodes;
- c. those more deeply embedded nodes with the value `f` in the attribute `tfa` that depend on a node with the value `t`, provided one of them is the focus proper.

## 10.4. Rhematizers

**Rhematizers** are expressions the function of which is to signal topic-focus articulation categories, namely the communicatively most important categories – the focus and contrastive topic (see Section 10.3, “Topic and focus”). Nodes representing rhematizers are assigned the functor `RHEM`.

**Expressions with the function of a rhematizer.** Rhematizers are mostly particles and adverbs. Most expressions that fulfill the function of a rhematizer are functionally homonymous (they fulfill also other functions). The role of a rhematizer is often played by particles and adverbs that primarily express adverbial modifications. Rhematizers are also functionally homonymous with some conjunction modifiers. Prototypical rhematizers include: *pouze* (=only), *jen* (=only), *jenom* (=only), *zejména* (=in\_particular), *zvláště* (=especially), *především* (=primarily), *obzvlášt'* (=especially), *hlavně* (=mainly), *jedině* (=only), *například* (=for example), *tolik* (=just), *výhradně* (=exclusively), *výlučně* (=exclusively) etc. Also *ne* (=no) and *ano* (=yes)) are considered rhematizers.

**Scope of a rhematizer.** The placement of rhematizers in the surface word order is quite free, however they almost always stand right before the expressions they rhematize, i.e. the expressions whose being part of the focus or contrastive topic they signal. We say that the part of the sentence that is rhematized is in the scope of the rhematizer. A rhematizer signaling the focus has in principle in its scope all contextually non-bound modifications (including their modifications) standing to the right from it in the surface structure of the sentence. A rhematizer signaling the contrastive topic has in principle in its scope the first contrastive contextually bound modification (including its modifications) standing to the right from it. There are, however, complex constructions, mainly within noun phrases, where the scope of a rhematizer over modifications of the governing expressions (contextually non-bound or contrastive contextually bound) is questionable. It seems that the “strength” of the scope of a rhematizer diminishes downwards in the structure of the sentence.

**Representing rhematizers in tectogrammatical trees.** The rules for positioning of rhematizers in tectogrammatical trees are simple:

- a. the node representing a rhematizer is placed as the closest left sister node of the first node of the expression that is in its scope.
- b. if a rhematizer has scope over the governing predicate, the rhematizer is placed as the closest left daughter of the node representing the predicate.
- c. if a rhematizer constitutes the focus proper, it is placed according to the rules for placement of the focus proper (see Section 10.2.1, “Basic rules for the nodes ordering in tectogrammatical trees”) - on the rightmost path from the effective root of the tectogrammatical tree.

**Contextual boundness of rhematizers.** For rhematizers, the value of the attribute `tfa` is determined by the function of the rhematizer, i.e. by which category of the topic-focus articulation it signals. The value of the attribute `tfa` is therefore closely related to the placement of the rhematizer in the tectogrammatical tree. A node representing a rhematizer can be assigned either the `tfa` value `t` or `f`. A rhematizer cannot be assigned the value `c`.

For a survey of possible situations, the corresponding `tfa` values and placement of the rhematizers, see Table 10.1, “Placement of rhematizers and their contextual boundness”.

**Table 10.1. Placement of rhematizers and their contextual boundness**

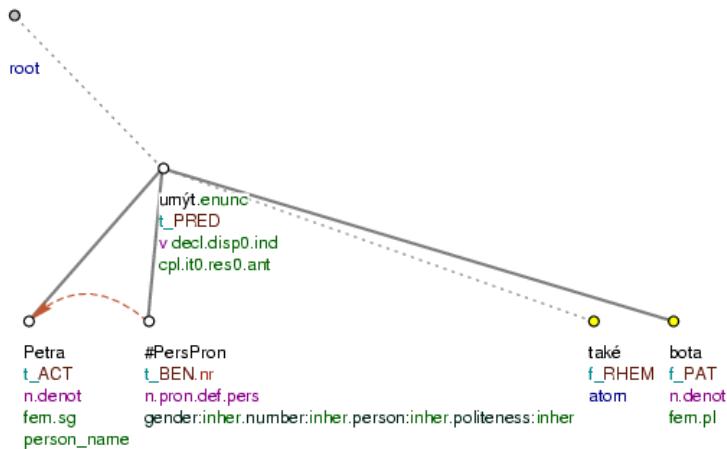
Situation	<code>tfa</code> of the nodes in the scope	<code>tfa</code> of RHEM	Position of RHEM	Example
<b>Signaling the focus</b> The predicate is not in the scope.	<code>f</code>	<code>f</code>	the closest left sister of the first node in the scope	<i>Petra si umyla také.RHEM [tfa = f] boty [tfa = f]</i> (=Petra also washed her shoes.) Fig. 10.11
<b>Signaling the focus</b> The predicate is in the scope.	<code>f</code>	<code>f</code>	the rightmost left daughter of the predicate	<i>Petra si také.RHEM [tfa = f] umyla [tfa = f] boty [tfa = f]</i> (=Petra also washed her shoes.) Fig. 10.12
<b>Signaling the contrastive topic</b>	<code>c</code>	<code>t</code>	the closest left sister of the first node in the scope	<i>Také.RHEM [tfa = t] Petra [tfa = c] si umyla boty.</i> (=Also Petra washed her shoes.) Fig. 10.13
<b>Scope over the non-contrastive contextually bound expressions</b> The predicate is not in the scope.	<code>t</code>	<code>t</code>	the closest left sister of the first node in the scope	<i>(Karel si umyl také boty. Spletl ses.) PETRA si umyla také.RHEM [tfa = t] boty [tfa = t]</i> (=Karel also washed his shoes. No, you're wrong) PETRA also washed her shoes.)
<b>Scope over the non-contrastive contextually bound negated predicate</b> The rhematizer is the predicate's negation morpheme ( <code>t_lemma = #Neg</code> ).	<code>t</code>	<code>t</code>	the rightmost left daughter of the predicate	<i>(Petr neodešel domů.) Hanka neodešla do školy.</i> (=P. didn't go home) H. didn't go to school.)
<b>Rhematizer in the focus proper</b>	<code>Ø</code>	<code>f</code>	on the rightmost path from the effective root of the tectogrammatical tree	<i>Petra si umyla boty také.RHEM [tfa = f]</i> (=Petra washed her shoes as well.) Fig. 10.14

**Rhematizer in a paratactic structure.** The scope of a rhematizer can consist of a whole paratactic structure (connecting either clause or sentence parts). If the rhematizer has scope over the whole paratactic structure, the node representing the rhematizer is placed as the closest left sibling of the root of the paratactic structure. For example:

*Byli tam jenom.*RHEM Petr a Pavel. (=Only Petr and Pavel were there.) Fig. 10.15

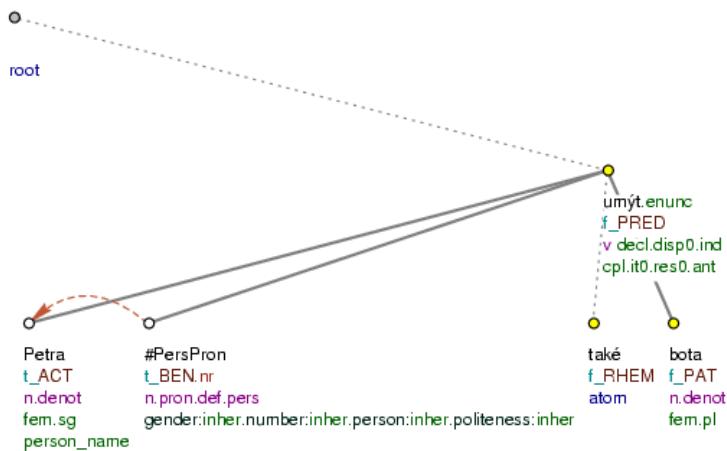
**Figure 10.11. Rheumatizers**

*Petra si umyla také boty. (=lit. Petra REFL washed also shoes.)*



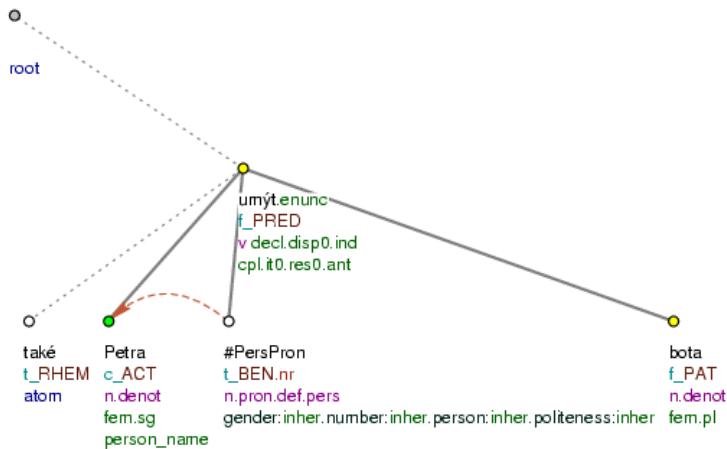
### Figure 10.12. Rheumatizers

*Petra si také umyla boty. (=lit. Petra REFL also washed shoes.)*



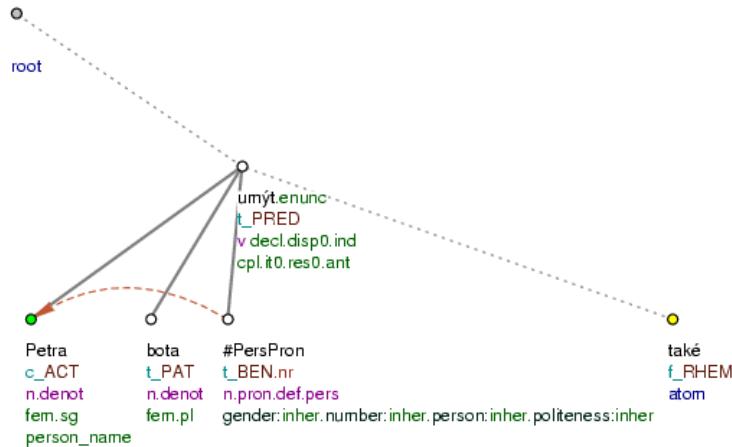
### **Figure 10.13. Rheumatizers**

Také Petra si umyla boty. (=lit. Also Petra REFL washed shoes.)

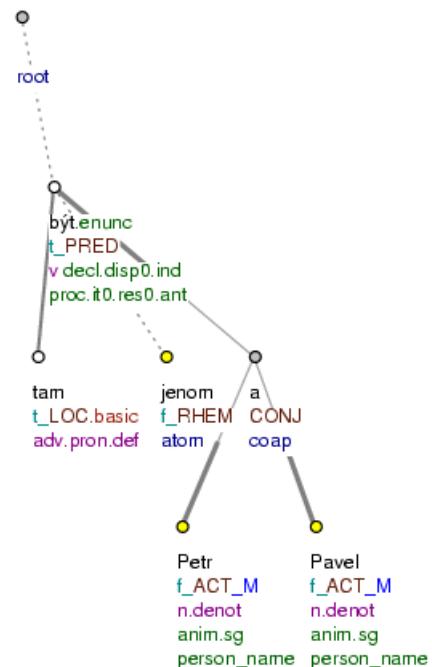


**Figure 10.14. Rhematizers**

*Petra si umyla boty také.* (=lit. *Petra REFL washed shoes as\_well.*)

**Figure 10.15. Rhematizers**

*Byli tam jenom Petr a Pavel.* (=lit. *Were there only Petr and Pavel.*)



---

# Appendix A. Attributes of nodes in tectogrammatical trees

## A.1. Attributes of the technical root

The technical root of a tectogrammatical tree is assigned four attributes with special values - see Table A.1, “Attributes of the technical root”.

Table A.1. Attributes of the technical root

Attribute	Obligatory	Value (type)	Note
<b>atree.rf</b>	NO	PML reference	This attribute links the tectogrammatical level to the analytical level by means of a reference to the corresponding analytical tree.
<b>deepord</b>	NO	0	This attribute determines the position of the node in the linear ordering in the tree. Unlike the other nodes, it carries no linguistic information.
<b>id</b>	YES	identifier	The unique identifier of the tree in PDT 2.0.
<b>nodetype</b>	NO	root	This attribute is present at the root just for the sake of the user. The technical root node is (as a type) different from all other nodes.

## A.2. Attributes of the other nodes

### A.2.1. a

Obligatory: NO.

This attribute binds the tectogrammatical nodes with units of lower levels. Its value is a structure with the attributes `lex.rf` and `aux.rf`. For more details see Chapter 2, *Relation between the tectogrammatical level and the lower levels*.

#### A.2.1.1. a/`lex.rf`

Obligatory: NO.

Table A.2. Values of the attribute a/`lex.rf`

PML reference	identifier of the node at the analytical level from which the tectogrammatical node got its lexical meaning (or its biggest part)
---------------	---

#### A.2.1.2. a/`aux.rf`

Obligatory: NO.

Table A.3. Values of the attribute a/`aux.rf`

a list every element of which is a PML reference	identifiers of the analytical nodes that influence in some way or other the value of the functor, subfunctor or grammatemes of the tectogrammatical node. These nodes frequently carry grammatical words (i.e. prepositions, subordinating conjunctions, auxiliary verbs, supporting words etc.) and together with the node referred to in the attribute a/ <code>lex.rf</code> they constitute one autosemantic expression.
--	--

## A.2.2. `compl.rf`

Obligatory: NO.

This attribute serves for representing the second dependency with predicative complements. It is assigned to nodes with the `COMPL` functor. For more details see Section 6.1.1.1, “Predicative complement”.

**Table A.4. Values of the attribute `compl.rf`**

PML reference	identifier of the node (usually in the same tree) with which the predicative complement is in the second dependency relation
---------------	--

## A.2.3. `coref_gram.rf`

Obligatory: NO.

This attribute serves for representing grammatical coreference. For more details see Chapter 9, *Coreference*.

**Table A.5. Values of the attribute `coref_gram.rf`**

PML reference	a list every element of which is a identifier of the nodes (usually in the same tree) with which the given node is in a grammatical coreference relation
---------------	--

## A.2.4. `coref_special`

Obligatory: NO.

This attribute serves for representing some special types of textual coreference when the coreferred node is not a particular node or subtree. For more details see Chapter 9, *Coreference*.

**Table A.6. Values of the attribute `coref_special`**

<code>exoph</code>	the coreferred element is an extra-textual situation
<code>segm</code>	the coreferred element is a larger text segment

## A.2.5. `coref_text.rf`

Obligatory: NO.

This attribute serves for representing textual coreference; the coreferred node is explicit (it is a particular node or subtree). For more details see Chapter 9, *Coreference*.

**Table A.7. Values of the attribute `coref_text.rf`**

PML reference	a list every element of which is a identifier of the nodes with which the given node is in a textual coreference relation
---------------	---

## A.2.6. `deepord`

Obligatory: YES.

This attribute assigns numbers to the nodes of a tree to indicate the deep word order. The order determined by the `deepord` attribute is the same order as the one displayed in the graphical programmes (from left to right). For more details see Section 10.2, “Communicative dynamism”.

**Table A.8. Values of the attribute `deepord`**

non-negative integer	order of the nodes in a tectogrammatical tree
----------------------	---

## A.2.7. functor

Obligatory: YES.

The value is the functor of the node. For more details see Chapter 8, *Functors and subfunctors*.

**Table A.9. Values of the attribute `functor`**

<b>ACMP</b>	adjunct expressing accompaniment (in the broad sense of the word)	<i>tatínek s maminkou.ACMP (=Father with Mother)</i>
<b>ACT</b>	argument - Actor	<i>Otec.ACT pracuje. (=Father is working)</i>
<b>ADDR</b>	argument - Addressee	<i>Poslal dárek příteli.ADDR (=He sent a present to a friend.)</i>
<b>ADVS</b>	paratactic structure root node - adversative relation	<i>Viděl, ale.ADVС neslyšel. (=He saw me but didn't hear me.)</i>
<b>AIM</b>	adjunct expressing purpose	<i>Cvičí, aby zhubla.AIM (=She is doing exercises in order to lose weight.)</i>
<b>APP</b>	adnominal adjunct expressing appurtenance	<i>můj.APP hrad (=my castle)</i>
<b>APPS</b>	appositional structure root node	<i>substantivum, neboli.APPS podstatné jméno (=substantive, or noun)</i>
<b>ATT</b>	atomic expression expressing the speaker's attitude	<i>Je to samozřejmě.ATT pravda. (=Of course, this is true.)</i>
<b>AUTH</b>	adnominal adjunct referring to the author (of sth)	<i>Nezvalovy.AUTH verše (=Nezval's poems)</i>
<b>BEN</b>	adjunct expressing that sth is happening for the benefit (or disadvantage) of sb/sth	<i>Pracuje pro firmu.BEN (=He is working for the company.)</i>
<b>CAUS</b>	adjunct expressing the cause (of sth)	<i>Z důvodu nemoci.CAUS zavřeno. (=Closed for illness.)</i>
<b>CNCS</b>	adjunct expressing concession	<i>Navzdory studijním úspěchům.CNCS se v praxi neuplatnil. (=Although he was a good student he wasn't very successful in practice.)</i>
<b>CM</b>	conjunction modifier	<i>otec a také.CM syn (=Father as well as his son)</i>
<b>COMPL</b>	adjunct - predicative complement	<i>Vrátila se unavená.COMPL (=She returned tired.)</i>
<b>COND</b>	adjunct expressing a condition (for sth else to happen)	<i>Když spí.COND, nezlobí. (=When he sleeps, he is good.)</i>
<b>CONFR</b>	paratactic structure root node - confrontation	<i>Pavel se zlepšuje, kdežto.CONFR Jan dostává čtyřky. (=P. is getting better; while J. is getting Ds.)</i>
<b>CONJ</b>	paratactic structure root node - simple coordination/conjunction	<i>Pavel a.CONJ Jan (=Pavel and Jan)</i>
<b>CONTRA</b>	paratactic structure root node - two entities are in conflict (fight etc.)	<i>otec versus.CONTRA syn (=Father vs. son)</i>
<b>CONTRD</b>	adjunct expressing confrontation	<i>Zatímco mzdy klesají.CONTRD, ceny se zvyšují. (=While the wages are going down, the prices are increasing.)</i>
<b>CPHR</b>	the nominal part of a complex predicate	<i>mit plán.CPHR (=have a plan)</i>
<b>CPR</b>	adjunct expressing comparison	<i>víc než tisíc.CPR korun (=more than thousand crowns)</i>
<b>CRIT</b>	adjunct expressing a criterion/measure/standard	<i>Serád' slova podle abecedy.CRIT (=Put the words in the alphabetical order.)</i>
<b>CSQ</b>	paratactic structure root node - consequential relation	<i>Pracoval nezodpovědně, a.CSQ proto dostal výpověď. (=He was irresponsible, therefore he was fired.)</i>
<b>DENOM</b>	effective root node of an independent nominative clause (which is not parenthetical)	<i>Základní škola.DENOM (=Primary school)</i>

<b>DIFF</b>	adjunct expressing a difference (between two entities, states etc.)	<i>Je vyšší <u>o</u> dva centimetry.</i> DIFF (=He is two centimeters taller.)
<b>DIR1</b>	directional adjunct - answering the question “where from”	<i>Přijel <u>z</u> Prahy.</i> DIR1 (=He came from Praha.)
<b>DIR2</b>	directional adjunct - answering the question “which way”	<i>Jdou <u>lesem</u>.</i> DIR2 (=They are walking through the forest.)
<b>DIR3</b>	directional adjunct - answering the question “where to”	<i>Přišel <u>domů</u>.</i> DIR3 (=He came home)
<b>DISJ</b>	paratactic structure root node - disjunctive relation	<i>Pojedu já, <u>nebo</u>.</i> DISJ ty. (=Either I will go, or you will.)
<b>DPHR</b>	the dependent part of an idiomatic expression	<i>křížem <u>krážem</u>.</i> DPHR (=criss cross)
<b>EFF</b>	argument - Effect	<i>Jmenovali ho <u>předsedou</u>.</i> EFF (=They appointed him as a chairman.)
<b>EXT</b>	adjunct expressing extent	<i>V nádobě je <u>přesně</u>.</i> EXT litr vody. (=There is exactly one litre of water in the pot.)
<b>FPHR</b>	part of a foreign-language expression	<i><u>cash</u>.</i> FPHR <i><u>flow</u>.</i> FPHR
<b>GRAD</b>	paratactic structure root node - gradation	<i>Běžel, <u>ba</u>.</i> GRAD utikal. (=He not only ran, he ran helter-skelter.)
<b>HER</b>	adjunct expressing inheritance	<i>štátek <u>po matce</u>.</i> HER (=scarf after my mother)
<b>ID</b>	nominative of identity and explicative genitive	<i>hrad <u>Karlštejn</u>.</i> ID (=the Karlstejn castle); <i>trest <u>smrti</u>.</i> ID (=death penalty)
<b>INTF</b>	atomic expression referring to the “false (expletive) subject”	<i><u>Ono</u>.</i> INTF prší. (=It is raining.)
<b>INTT</b>	adjunct expressing intention	<i>Šel <u>nakoupit</u>.</i> INTT (=He went shopping.)
<b>LOC</b>	locative adjunct - answering the question “where”	<i>Pracuje <u>v Praze</u>.</i> LOC (=He works in Praha.)
<b>MANN</b>	adjunct expressing manner (of doing sth)	<i>Mluví <u>hlásitě</u>.</i> MANN (=He talks loudly.)
<b>MAT</b>	adnominal argument referring to the content of a container	<i>sklenice <u>vody</u>.</i> MAT (=a glass of water)
<b>MEANS</b>	adjunct expressing a means (of doing sth)	<i>Píše <u>perem</u>.</i> MEANS (=He is writing with a pen.)
<b>MOD</b>	atomic expression with a modal meaning	<i>Pracuje <u>asi</u>.</i> MOD na půl úvazku. (=He is probably working part-time.)
<b>OPER</b>	paratactic structure root node referring to a mathematical operation or interval	<i>pět <u>až</u>.</i> OPER deset hodin (=five to ten hours)
<b>ORIG</b>	argument - Origo	<i>Vyrábí nábytek <u>ze dřeva</u>.</i> ORIG (=He makes furniture out of wood.)
<b>PAR</b>	effective root node of a parenthetic (verbal or nominative) clause	<i>Přijedu 13. prosince (<u>pátek</u>).</i> PAR. (=I'm coming on 13th of December (Friday).)
<b>PARTL</b>	effective root node of an independent interjectional clause	<i>Hurá.</i> PARTL, vyhráli jsme! (=Hurray, we have won!)
<b>PAT</b>	argument - Patient	<i>Varí <u>oběd</u>.</i> PAT (=He's cooking dinner.)
<b>PREC</b>	atomic expression referring to the preceding context	<i>A.</i> PREC pak odešel. (=And then he left.)
<b>PRED</b>	effective root node of an independent verbal clause (which is not parenthetical)	<i>Pavel <u>dal</u>.</i> PRED kytku Martině. (=Pavel gave a flower to Martina)
<b>REAS</b>	paratactic structure root node - causal relation	<i>Dostal výpověď, <u>nebot'</u>.</i> REAS pracoval nezodpovědně. (=He was fired since he was an irresponsible worker.)
<b>REG</b>	adjunct expressing with regard to what sth is asserted	<i>Vzhledem k počasí.</i> REG nelze nic plánovat. (=It is impossible to plan anything due to the weather.)
<b>RESL</b>	adjunct expressing the result/effect of something	<i>Mluví tak potichu, <u>že</u> mu nerozumíme.</i> RESL (=He speaks so softly that we cannot understand him.)
<b>RESTR</b>	adjunct expressing an exception (to sth)	<i>Kromě tebe.</i> RESTR tam byli všichni. (=Except for you everybody was there.)

<b>RHEM</b>	atomic expression - rhematizer	<i>Jen.RHEM Karel odešel. (=Only Karel left)</i>
<b>RSTR</b>	adnominal adjunct modifying its governing noun	<i>velký.RSTR dům (=a big house)</i>
<b>SUBS</b>	adjunct expressing that sb/sth substitutes for sb/sth else	<i>Za otce.SUBS jednal strýc. (=My uncle took action in place of my father.)</i>
<b>TFHL</b>	temporal adjunct - answering the question “for how long”	<i>Přijel na měsíc.TFHL (=He came for a month.)</i>
<b>TFRWH</b>	temporal adjunct - answering the question “from when”	<i>Přeložil jednání ze soboty.TFRWH na dnešek. (=He shifted the negotiations from Saturday to today.)</i>
<b>THL</b>	temporal adjunct - answering the questions “how long” and “after how long”	<i>Stihnu to za týden.THL (=He managed to do it in a week.)</i>
<b>THO</b>	temporal adjunct - answering the questions “how often” and “how many times”	<i>Pracuju na tom každý den.THO (=I'm working on it every day.)</i>
<b>TOWH</b>	temporal adjunct - answering the question “to when”	<i>Přeložil jednání ze soboty na dnešek.TOWH (=He shifted the negotiations from Saturday to today.)</i>
<b>TPAR</b>	temporal adjunct - answering the questions “in parallel/simultaneously with what” and “during what time”	<i>Během naší dovolené.TPAR ani jednou nepršelo. (=During our holiday it didn't rain once.)</i>
<b>TSIN</b>	temporal adjunct - answering the question “since when”	<i>Budu pracovat od zítra.TSIN (=I'll be working from tomorrow.)</i>
<b>TTILL</b>	temporal adjunct - answering the question “until when”	<i>Udělám to do pátku.TTILL (=I'll do it by Friday.)</i>
<b>TWHEN</b>	temporal adjunct - answering the question “when”	<i>Přijdu zítra.TWHEN (=I'll come tomorrow.)</i>
<b>VOCAT</b>	effective root node of an independent vocative clause	<i>Hanko.VOCAT, podej mi to. (=Hanka, hand it to me.)</i>

## A.2.8. gram

Obligatory: NO.

The value is a structure with the attributes-grammatemes: `sepos`, `gender`, `number`, `degcmp`, `verbmod`, `deontmod`, `tense`, `aspect`, `resultative`, `dispmod`, `iterativeness`, `indeftype`, `person`, `numbertype`, `politeness`, `negation`. This structure is only assigned to complex nodes, i.e. nodes whose value of the `nodetype` attribute is `complex`. For more details see Chapter 5, *Complex nodes and grammatemes*.

### A.2.8.1. gram/aspect

Obligatory: NO.

For more on the aspect grammeme, see Section 5.2, “Grammatemes”.

Table A.10. Values of the grammeme `aspect`

<b>proc</b>	process, imperfective aspect
<b>cpl</b>	complex, perfective aspect
<b>nr</b>	all basic values are possible in the given case; none of them can be excluded

### A.2.8.2. gram/degcmp

Obligatory: NO.

For more on the degree grammeme, see Section 5.2, “Grammatemes”.

**Table A.11. Values of the attribute `degcmp`**

<code>pos</code>	positive
<code>comp</code>	comparative
<code>sup</code>	superlative
<code>acomp</code>	elative
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded.

### A.2.8.3. `gram/deontmod`

Obligatory: NO.

For more on the deontic modality grammategeme, see Section 5.2, “Grammatemes”.

**Table A.12. Values of the grammategeme `deontmod`**

<code>deb</code>	the event is understood as “necessary”
<code>hrt</code>	the event is understood as “obligatory (an obligation)”
<code>vol</code>	the event is understood as “wanted/intended”
<code>poss</code>	the event is understood as “possible”
<code>perm</code>	the event is understood as “permitted”
<code>fac</code>	the event is understood as “an ability (to do sth)”
<code>decl</code>	the basic (unmarked) modality
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded

### A.2.8.4. `gram/dispmad`

Obligatory: NO.

For more on the dispositional modality grammategeme, see Section 5.2, “Grammatemes”.

**Table A.13. Values of the grammategeme `dispmad`**

<code>disp0</code>	no dispositional modality
<code>disp1</code>	the predicate expresses dispositional modality
<code>nil</code>	none of the basic values of the grammategeme (i.e. the grammategeme as such) is relevant.
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded

### A.2.8.5. `gram/gender`

Obligatory: NO.

For more on the gender grammategeme, see Section 5.2, “Grammatemes”.

**Table A.14. Values of the grammategeme `gender`**

<code>anim</code>	masculine animate
<code>inan</code>	masculine inanimate
<code>fem</code>	feminine
<code>neut</code>	neuter
<code>inher</code>	the value of the grammategeme follows from the value of the corresponding grammategeme of the coreferred node
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded.

## A.2.8.6. gram/`indeftype`

Obligatory: NO.

For more on the indefiniteness grammatememe, see Section 4.2, “T-lemma of derived expressions”.

**Table A.15. Values of the grammatememe `indeftype`**

<code>relat</code>	relative pronoun / adverb / numeral ( <i>kdo, jaký; kdy, jak; kolik</i> )
<code>indef1</code>	indefinite pronoun / adverb / numeral of the type <i>někdo, nějaký / někde, nějak / několik</i>
<code>indef2</code>	indefinite pronoun / pronominal adverb of the type <i>kdosi, jakýsi / kdesi, jaksi</i>
<code>indef3</code>	indefinite pronoun / pronominal adverb of the type <i>kdokoli, jakýkoli / kdekoli, jakkoli</i>
<code>indef4</code>	indefinite pronoun / pronominal adverb of the type <i>ledakdo, ledajaký / ledakde, ledajak</i>
<code>indef5</code>	indefinite pronoun / pronominal adverb of the type <i>málokdo / málokde, kdovíkdo / kdovíkde</i>
<code>indef6</code>	indefinite pronoun of the type <i>kdekdo, kdejaký;</i>
<code>inter</code>	interrogative pronoun / pronominal adverb/ numeral
<code>negat</code>	negative pronoun / pronominal adverb ( <i>nikdo, nijaky; nikde, nijak</i> )
<code>total1</code>	totalizing pronoun / pronominal adverb (referring to a whole) ( <i>všichni; všude</i> )
<code>total2</code>	totalizing pronoun (referring to individuals) ( <i>každý</i> )
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded.

## A.2.8.7. gram/`iterativeness`

Obligatory: NO.

For more on the iterativity grammatememe, see Section 5.2, “Grammatemes”.

**Table A.16. Values of the grammatememe `iterativeness`**

<code>it0</code>	no iterative meaning present
<code>it1</code>	iterated, multiple event
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded

## A.2.8.8. gram/`negation`

Obligatory: NO.

For more on the negation grammatememe, see Section 5.2, “Grammatemes”.

**Table A.17. Values of the grammatememe `negation`**

<code>neg0</code>	affirmative
<code>neg1</code>	negative
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded.

## A.2.8.9. gram/`number`

Obligatory: NO.

For more on the number grammatememe, see Section 5.2, “Grammatemes”.

**Table A.18. Values of the grammateme number**

<b>sg</b>	singular
<b>pl</b>	plural
<b>inher</b>	the value of the grammateme follows from the value of the corresponding grammateme of the coreferred node
<b>nr</b>	all basic values are possible in the given case; none of them can be excluded.

## A.2.8.10. **gram/numertype**

Obligatory: NO.

For more on the numeral type grammateme, see Section 4.2, “T-lemma of derived expressions”.

**Table A.19. Values of the grammateme numertype**

<b>basic</b>	cardinal numeral ( <i>tři, šest, kolik</i> )
<b>frac</b>	fraction numeral ( <i>třetina, šestina</i> )
<b>kind</b>	sort numeral ( <i>trojí, šestery, kolikery</i> )
<b>ord</b>	ordinal numeral ( <i>třetí, šestý, kolikátý</i> )
<b>set</b>	set numeral ( <i>troje, šestery, kolikery</i> )
<b>nr</b>	all basic values are possible in the given case; none of them can be excluded.

## A.2.8.11. **gram/person**

Obligatory: NO.

For more on the person grammateme, see Section 5.2, “Grammatemes”.

**Table A.20. Values of the grammateme person**

<b>1</b>	first person (speaker)
<b>2</b>	second person (hearer)
<b>3</b>	third person (what is talked about)
<b>inher</b>	the value of the grammateme follows from the value of the corresponding grammateme of the coreferred node
<b>nr</b>	all basic values are possible in the given case; none of them can be excluded.

## A.2.8.12. **gram/politeness**

Obligatory: NO.

For more on the politeness grammateme, see Section 5.2, “Grammatemes”.

**Table A.21. Values of the grammateme politeness**

<b>basic</b>	the common use
<b>polite</b>	the polite form
<b>inher</b>	the value of the grammateme follows from the value of the corresponding grammateme of the coreferred node
<b>nr</b>	all basic values are possible in the given case; none of them can be excluded.

## A.2.8.13. **gram/sempos**

Obligatory: YES.

The *sempos* attribute contains the information regarding the membership of a complex node in a subgroup of a semantic part of speech. For more details see Section 5.1, “Semantic parts of speech”.

**Table A.22. Values of the grammategeme `sempos`**

<code>n.denot</code>	denominating semantic noun
<code>n.denot.neg</code>	denominating semantic noun with separately represented negation
<code>n.pron.def.demon</code>	definite pronominal semantic noun: demonstrative
<code>n.pron.def.pers</code>	definite pronominal semantic noun: personal
<code>n.pron.indef</code>	indefinite pronominal semantic noun
<code>n.quant.def</code>	definite quantificational semantic noun
<code>adj.denot</code>	denominating semantic adjective
<code>adj.pron.def.demon</code>	definite pronominal semantic adjective: demonstrative
<code>adj.pron.indef</code>	indefinite pronominal semantic adjective
<code>adj.quant.def</code>	definite quantificational semantic adjective
<code>adj.quant.indef</code>	indefinite quantificational semantic adjective
<code>adj.quant.grad</code>	gradable quantificational semantic adjective
<code>adv.denot.ngrad.nneg</code>	nongradable denominating semantic adverb, impossible to negate
<code>adv.denot.ngrad.neg</code>	nongradable denominating semantic adverb, possible to negate
<code>adv.denot.grad.nneg</code>	gradable denominating semantic adverb, impossible to negate
<code>adv.denot.grad.neg</code>	gradable denominating semantic adverb, possible to negate
<code>adv.pron.def</code>	definite pronominal semantic adverb
<code>adv.pron.indef</code>	indefinite pronominal semantic adverb
<code>v</code>	semantic verb

### A.2.8.14. `gram/verbmod`

Obligatory: NO.

For more on the verbal modality grammategeme, see Section 5.2, “Grammatemes”.

**Table A.23. Values of the grammategeme `verbmod`**

<code>ind</code>	indicative
<code>imp</code>	imperative
<code>cdn</code>	conditional
<code>nil</code>	none of the basic values of the grammategeme (i.e. the grammategeme as such) is relevant.
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded

### A.2.8.15. `gram/resultative`

Obligatory: NO.

For more on the resultativity grammategeme, see Section 5.2, “Grammatemes”.

**Table A.24. Values of the grammategeme `resultative`**

<code>res0</code>	no resultative meaning
<code>res1</code>	resultative meaning (aspect)
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded

### A.2.8.16. `gram/tense`

Obligatory: NO.

For more on the tense grammategeme, see Section 5.2, “Grammatemes”.

**Table A.25. Values of the grammateme `tense`**

<code>sim</code>	simultaneous
<code>ant</code>	preceding
<code>post</code>	subsequent
<code>nil</code>	none of the basic values of the grammateme (i.e. the grammateme as such) is relevant.
<code>nr</code>	all basic values are possible in the given case; none of them can be excluded

## A.2.9. `id`

Obligatory: YES.

The value is the unique identifier of the tree in PDT 2.0.

## A.2.10. `is_dsp_root`

Obligatory: NO.

This attribute marks the root of direct speech subtrees. If no value is assigned, the assumed value is 0. For more details see Section 7.5, “Direct speech”.

**Table A.26. Values of the attribute `is_dsp_root`**

<code>0</code>	the node is not the root of a subtree representing direct speech
<code>1</code>	the node is a the root of a subtree representing direct speech

## A.2.11. `is_generated`

Obligatory: NO.

This attribute marks generated (newly established) nodes. If no value is assigned, the assumed value is 0. For more details see Section 6.6, “Ellipsis”.

**Table A.27. Values of the attribute `is_generated`**

<code>0</code>	the node represents an expression present in the surface structure
<code>1</code>	the node represents an expression absent in the surface structure

## A.2.12. `is_member`

Obligatory: NO.

This attribute is only assigned to immediate daughters of paratactic structure root nodes (`nodetype = coap`) and marks the roots of the paratactically connected elements. For other nodes, the attribute is not relevant. For more details see Section 6.4, “Parataxis”.

**Table A.28. Values of the attribute `is_member`**

<code>0</code>	the node is not a root of a paratactically connected element
<code>1</code>	the node is the root of a paratactically connected element

## A.2.13. `is_name_of_person`

Obligatory: NO.

This attribute serves for marking proper names of people. If no value is assigned, the assumed value is 0. For more details see Section 7.3, “Identifying expressions”.

**Table A.29. Values of the attribute `is_name_of_person`**

<b>0</b>	the node represents an expression that is not part of a proper name of a person
<b>1</b>	the node represents an expression that is part of a proper name of a person

## A.2.14. `is_parenthesis`

Obligatory: NO.

This attribute marks nodes representing expressions that are part of a parenthesis. If no value is assigned, the assumed value is 0. For more details see Section 6.5, “Parenthesis”.

**Table A.30. Values of the attribute `is_parenthesis`**

<b>0</b>	the node represents an expression that is not part of a parenthesis
<b>1</b>	the node represents an expression that is part of a parenthesis

## A.2.15. `is_state`

Obligatory: NO.

This attribute marks nodes representing modifications with the meaning of state. If no value is assigned, the assumed value is 0. For more details see Section 8.13.2, “Attribute for the meaning of “state””.

**Table A.31. Values of the attribute `is_state`**

<b>0</b>	the node represents a modification that has no state meaning
<b>1</b>	the node represents a modification with a state meaning that cannot be captured just by the functor and subfunctors

## A.2.16. `nodetype`

Obligatory: YES.

This attribute defines a nodetype. For more details see Chapter 3, *Node types*.

**Table A.32. Values of the attribute `nodetype`**

<b>atom</b>	atomic node
<b>coap</b>	paratactic structure root node
<b>list</b>	list structure root node
<b>fphr</b>	node representing a foreign-language expression
<b>dphr</b>	node representing the dependent part of an idiomatic expression
<b>complex</b>	complex node
<b>qcomplex</b>	quasi-complex node

## A.2.17. `quot`

Obligatory: NO.

The attribute marks nodes representing a text segment “in quotation marks”. The value of the attribute is a list every item of which is a structure with the attributes `type` and `set_id`. For more details see Section 7.6, “Text segments marked by graphic symbol”.

### A.2.17.1. `quot/set_id`

Obligatory: YES.

The attribute serves as the identifier unambiguously marking a set of nodes representing a text segment in quotation marks

**Table A.33. Values of the attribute `quot/set_id`**

arbitrary sequence	identifier unambiguously marking a set of nodes representing a text segment in quotation marks
--------------------	--

## A.2.17.2. `quot/type`

Obligatory: YES.

The attribute indicates the type of quotation mark use.

**Table A.34. Values of the attribute `quot/type`**

<b>citation</b>	the node represents an expression that is part of quotation (marked by quotation marks)
<b>dsp</b>	the node represents an expression that is part of direct speech (marked by quotation marks)
<b>meta</b>	the node represents an expression that is part of a metalinguistic expression (marked by quotation marks)
<b>title</b>	the node represents an expression that is part of a proper name or title (marked by quotation marks)
<b>other</b>	the node represents an expression that is part of a text segment in quotation marks and it has none of the above mentioned functions

## A.2.18. `sentmod`

Obligatory: NO.

The attribute indicates the modality of the sentence. For more details see Section 5.3, “The `sentmod` attribute”.

**Table A.35. Values of the attribute `sentmod`**

<b>enunc</b>	the indicative mood
<b>excl</b>	exclamation
<b>desid</b>	the optative (desiderative) mood
<b>imper</b>	the imperative mood
<b>inter</b>	the interrogative mood

## A.2.19. `subfunctor`

Obligatory: NO.

The value of the attribute is a subfunctor further specifying the meaning of the assigned functor. For more details see Section 8.13.1, “Subfunctors”.

**Table A.36. Values of the attribute `subfunctor`**

<b>Value</b>	<b>Definition</b>	<b>Functors to which it is assigned</b>
<b>above</b>	the meaning “above”	DIR3, LOC
<b>abstr</b>	the abstract meaning “in the area”	LOC
<b>across</b>	the meaning “across”	DIR2
<b>after</b>	the meaning “after”	TWHEN
<b>agst</b>	the meaning of disadvantage	BEN
<b>along</b>	the meaning “along”	DIR2, LOC
<b>approx</b>	the meaning “approximately”	EXT, TWHEN
<b>around</b>	the meaning “around”	DIR2, LOC
<b>basic</b>	the basic meaning of the functor	all functors
<b>before</b>	the meaning “before”	TWHEN
<b>begin</b>	the meaning “at the beginning”	TWHEN
<b>behind</b>	the meaning “behind”	DIR3, LOC
<b>below</b>	the meaning “below”	DIR3, LOC
<b>betw</b>	the meaning “between”	DIR2, DIR3, LOC, TWHEN
<b>circ</b>	the meaning of circumstance	ACMP
<b>elsew</b>	the meaning “outside”	DIR3, LOC
<b>end</b>	the meaning “at the end”	TWHEN
<b>ext</b>	the meaning “to what extent”	DIR3
<b>flow</b>	the meaning “in the course”	TWHEN
<b>front</b>	the meaning “in front of”	DIR3, LOC
<b>incl</b>	inclusion	ACMP
<b>in</b>	the meaning “inside”	LOC
<b>less</b>	the meaning “less than”	EXT
<b>mid</b>	the meaning “in the middle”	LOC, TWHEN
<b>more</b>	the meaning “more than”	EXT
<b>near</b>	the meaning “near”	DIR2, DIR3, LOC
<b>nr</b>	there is no subfunctor for the given surface form	all functors
<b>opp</b>	the meaning “opposite”	DIR3, LOC
<b>target</b>	the meaning of target	DIR3
<b>than</b>	comparison on the basis of difference	CPR
<b>to</b>	the meaning “to”	DIR3
<b>wout</b>	the meaning of negative accompaniment	ACMP
<b>wrt</b>	the meaning “entity something is compared to”	CPR

## A.2.20. `t_lemma`

Obligatory: YES.

The value of the attribute is an arbitrary sequence. For more details see Chapter 4, *Tectogrammatical lemma (`t_lemma`)*.

## A.2.21. **tfa**

Obligatory: NO.

The attribute represents contextual boundness. For more details see Section 10.1, “Contextual boundness”.

**Table A.37. Values of the attribute **tfa****

<b>c</b>	the node represents a contrastive contextually bound expression
<b>f</b>	the node represents a contextually non-bound expression
<b>t</b>	the node represents a non-contrastive contextually bound expression

## A.2.22. **val\_frame.rf**

Obligatory: NO.

The attribute serves as a reference into the valency lexicon. For more details see Section 6.2, “Valency”.

**Table A.38. Values of the attribute **val\_frame.rf****

PML reference	identifier of the valency frame describing the meaning and valency of the given node
---------------	--

---

# Subject index

## A

Actor, 45, 121  
  unspecified, 81  
Addressee, 45, 121  
adjective  
  control, 143  
  introduced by a subordinating conjunction, 60  
  semantic, 24  
adverb  
  predicative, 100  
  semantic, 25  
antecedent, 139  
apposition, 70  
argument, 45, 121  
  contextually elided, 81  
  controlled, 81  
  general, 81  
  quasi-controlled, 81  
  unspecified, 81  
argument shifting, 45  
aspect, 31  
attribute, 1, 129  
  of the type reference, 2

## B

brackets, 120

## C

clause  
  adverbial, 57  
  content, 57  
  dependent, 54, 57  
  dependent consecutive, 90  
  false dependent, 63  
  governing, 57  
  independent, 54  
  interjectional, 54  
  nominative, 53  
  non-verbal, 53  
  relative, 57  
  verbal, 51  
  vocative, 53  
clauses with a reversed syntactic relation, 74  
communicative dynamism, 154  
competing valency modifications, 46  
construction  
  control, 146  
  with a dependent consecutive clause, 90  
  with the meaning of a restriction, 86  
  with the meaning of comparison of two events, 83  
  with the meaning of exceptional conjoining, 87  
context, 152  
contextual boundness, 152  
control, 140, 143  
controllee, 143  
controller, 143  
coordinating connective  
  complex, 94  
coordination, 70

coreference, 139  
  grammatical, 140  
  textual, 150

## D

dashes, 120  
degree, 28  
dependency, 34, 55  
  ambiguous, 40  
  dual, 34  
derivation, 7  
dialogue test, 44  
direct speech, 55, 115  
dual function of a single modification, 40

## E

edge, 2, 34  
  non-dependency, 37  
Effect, 45, 121  
effective root, 3  
element  
  controlled, 143  
  controlling, 143  
  coreferred, 139  
  coreferring, 139  
  dependent, 34  
  governing, 34  
ellipsis, 76  
  contextual, 76  
  grammatical, 77  
  of a dependent meaning unit, 80  
  of a governing meaning unit, 76  
  of a governing meaning unit, in specific constructions, 76  
  of the governing clause, 76  
  of the governing noun, 76  
  of the governing predicate, 76  
empty verb, 52  
exophor, 150  
explicative genitive, 110  
explicit coreferred element, 150  
expression, 3  
  contextually non-bound, 153  
  contrastive contextually bound, 153  
foreign-language, 113  
identifying, 109, 120  
in the function of a rhematizer, 163  
modifying a coordinating connective, 94, 130  
non-contrastive contextually bound, 153  
numeral, 106  
used metalinguistically, 120

## F

focus, 163  
  proper, 155  
focus proper, 155  
foreign-language expression, 113  
functors, 121  
  adnominal, 129  
  for arguments, 121  
  for causal relations, 125  
  for effective roots of independent clauses, 121

for expressing manner and its specific variants, 126  
 for foreign-language expressions, 128  
 for multi-word lexical units, 128  
 for relations between members of paratactic structures, 130  
 for rhematisers, sentence, linking and modal adverbials, 127  
 for the predicative complement, 130  
 locative and directional, 124  
 temporal, 122

**G**

gender, 28  
 grammatem, 27

**I**

identification structure, 111  
 items of an identifying expression, 112  
 modifier, 112  
 root, 111  
 idiom, 104  
 idiomatic expression, 104  
 non-verbal, 104  
 verbal, 104  
 immediate daughter, 2  
 left, 2  
 right, 2  
 infinitival construction  
 dependent, 57  
 infinitival constructions  
 frozen, 58  
 infinitive  
 a valency modification of which is controlled, 143  
 inner participant, 44  
 interval, 70  
 intonation, 152  
 intonation centre, 152, 155  
 iterativity, 31

**L**

lemma  
 tectogrammatical, 7  
 lexical unit  
 multi-word, 94  
 linear order, 2  
 list structure, 5  
 item of a list, 113  
 modifier of a list, 113  
 root, 5, 113

**M**

mathematical operation, 70  
 member  
 direct, 66  
 terminal, 66  
 modality  
 deontic, 31  
 dispositional, 31, 48  
 sentential, 32  
 verbal, 32  
 modification  
 adnominal, 129

free, 44  
 locative and directional, 124  
 obligatory, 44  
 optional, 44  
 temporal, 122  
 valency, 44, 121  
 modification with the meaning of  
 accompaniment, 126  
 appurtenance, 129  
 attitude, 127  
 author, 129  
 beneficiary, 128  
 cause, 125  
 comparison, 126  
 concession, 125  
 condition, 125  
 confrontation, 128  
 container, 129  
 criterion, 126  
 difference, 126  
 exception, 126  
 extent, 126  
 identification, 129  
 inheritance, 128  
 instrument/means, 126  
 manner, 126  
 modality, 127  
 purpose, 125  
 regard, 126  
 result/effect, 126  
 substitution, 128  
 modifier  
 of a list, 113  
 of a paratactic structure, 67  
 of an identification structure, 112  
 mother, 2

**N**

negation, 29, 92  
 node, 1  
 added (newly established), 1, 75  
 atomic, 5, 127  
 complex, 6, 20  
 contextually non-bound, 153  
 contrastive contextually bound, 153  
 copied, 75  
 dependent, 2, 34  
 governing, 2, 34  
 newly established, with a t-lemma substitute, 76  
 node type, 5  
 non-contrastive contextually bound, 153  
 quasi-complex, 6  
 node attributes, 167  
 nominative  
 of identity, 55  
 nominative of identity, 110  
 non-dependency, 54, 121  
 noun  
 control, 143  
 semantic, 21  
 noun phrase, 41  
 number, 29, 105  
 numeral, 105

in the role of an attribute, 105  
used as labels, 105  
with an adverbial meaning, 106  
with the function of a container, 105

**O**

ordering  
objective, 152  
subjective, 152  
Origo, 45, 121

**P**

paratactic structure, 66  
modifier, 67  
root, 5, 66  
parataxis, 66  
parenthesis, 72  
lexicalized, 73  
proper, 72  
proper, syntactically incorporated, 72  
proper, syntactically non-incorporated, 73  
part of speech  
semantic, 20  
participial construction  
dependent, 57  
non-agreeing, 58  
passive, 48  
path  
left, 2  
leftmost, 2  
right, 2, 155  
rightmost, 2, 155  
Patient, 45, 121  
person, 30  
politeness, 30  
postcedent, 139  
predicate  
complex, 100  
modal, 95  
multi-word, 95  
phase, 99  
quasi-modal, 99  
quasi-phase, 99  
verbonominal, 103  
predicative complement, 34, 130  
preposition  
frozen, derived from a transgressive, 58  
projectivity, 161  
projectivization, 161  
proper name, 109  
punctuation, 52, 66

**Q**

quasi-control, 141  
with complex predicates, 101  
quasi-focus, 155  
quotation, 120  
quotation marks, 119

**R**

reciprocity, 81, 141  
reference

anaphoric, 139  
cataphoric, 139  
exophoric, 150  
to a segment, 150  
relation  
adversative, 131  
causal, 131  
confrontational, 131  
conjoining, 131  
consecutive, 131  
disjunctive, 131  
gradation, 131  
resultativity, 32  
rezultativ, 48  
rhematizer, 163  
root  
effective, 3  
expression, 3  
list structure, 5  
of a list structure for foreign-language expressions, 113  
of a paratactic structure, 66  
of an identification structure, 111  
paratactic structure, 5  
sentence, 3  
subtree, 3  
technical root node of a tectogrammatical tree, 2, 5, 167

**S**

scope of a rhematizer, 163  
segment, 150  
sentence  
complex, 55  
compound, 66  
shared modifier, 67  
sister, 2  
sister node  
left, 2  
right, 2  
state, 138  
stress  
contrastive, 152  
sentence, 152  
structure  
identification, 111  
list structure for foreign-language expressions, 113  
paratactic, 66  
paratactic, embedded, 66  
subfunctor, 132  
subtree, 3  
supporting element, 60

**T**

t-lemma, 7  
multi-word, 16  
representative, 17  
substitute, 18  
technical root node of a tectogrammatical tree, 2, 5, 167  
tectogrammatical trees, 1  
tense, 32  
the value of the attribute, 2

title, 109  
topic, 163  
    contrastive, 163  
topic-focus articulation, 152  
transgressival construction, 57  
transgressival constructions  
    frozen, 58

## V

valency, 44  
valency frame, 45  
    empty, 47  
    of a complex predicate, 49  
    of an idiom, 49  
valency lexicon, 50  
verb  
    být, 103  
    control, 143, 145  
    empty, 52  
    inherently reciprocal, 16  
    inherently reflexive, 16  
    modal, 95  
    phase, 99  
    quasi-modal, 99  
    quasi-phase, 99  
    reflexive, 16  
    semantic, 27

## W

word form, 7  
word order  
    deep, 2, 154  
    surface, 152