Annotation of English on the tectogrammatical level

Reference book

Silvie Cinková
Jan Hajič
Marie Mikulová
Lucie Mladová
Anja Nedolužko
Petr Pajas
Jarmila Panevová
Jiří Semecký
Jana Šindlerová
Josef Toman
Zdeňka Urešová
Zdeněk Žabokrtský

On the basis of Mikulová et al.: Annotation on the tectogrammatical level in the Prague Dependency Treebank. The abbreviated version of the Annotation manual. TR-2005-28, ÚFAL MFF UK, Prague, 2005, which was translated into English by Kateřina Součková.

Document version: 1.0.1

Annotation of English on the tectogrammatical level: Reference book

by Silvie Cinková, Jan Hajič, Marie Mikulová, Lucie Mladová, Anja Nedolužko, Petr Pajas, Jarmila Panevová, Jiří Semecký, Jana Šindlerová, Josef Toman, Zdeňka Urešová, and Zdeněk Žabokrtský

Table of Contents

Acknowledgements

The document was written with the financial support of the grants GA UK 489/2004, GA CR 405/06/0589, LC 536, MSM 0021620838 and 1ET201120505.

Introduction

This manual presents a short description of sentence representation on the tectogrammatical level in the Prague English Dependency Treebank (PEDT), which is the English counterpart of the Prague Czech-English Dependency Treebank. It is based on the abbreviated version of the annotation manual for the Czech corpus Prague Dependency Treebank (hence "PDT manual") and on a preliminary manual for English tectogrammatical annotation (TR-2004-21), which was only available in Czech and has been significantly revised and extended since 2004. This manual seeks to keep the structure of the original PDT manual to the highest possible degree and we are going to complete the missing topics as soon as their annotation has been launched. So far, the annotation of PEDT has been concerned with the tree structure and with functor assignment as well as with creating the valency lexicon of verbs.

As the tectogrammatical annotation of PEDT is still at its initial stage, not all issues described by the PDT annotation manual could be adopted (e.g. Topic-Focus Articulation, Coreference, Grammatemes etc.). The current manual mentions them but does actually neither define nor explain them. Readers willing to get consistent information on these issues are kindly requested to consult the annotation manual for Czech. The next version of this manual is supposed to cover the complete range of FGD-related topics.On the other hand, a chapter dealing with issues specific to the annotation of English data has been added.

The chapters are organized in a way parallel to the way of sentence representation on the tectogrammatical level. In Chapter 1, ***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 2, ***Node types***, divides the tectogrammatical nodes into different types. The next chapter, Chapter 3, ***Tectogrammatical lemma (t-lemma)***, is devoted to the attributes specifying the lexical unit (represented by a node). A description of the sentence structure follows (Chapter 4, ***Sentence representation structure***). Chapter 5, *Specific phenomena* contains a description of certain specific phenomena. A separate chapter is devoted to the functors and sub-functors (Chapter 6, ***Functors and subfunctors***).

1. Typographical conventions[1]

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. They are usually presented without context. The given annotation corresponds to the most common context in which the sentence can be used.

Example sentences have mostly been taken from PEDT or BNC, from Grammar textbooks or at least from Google (English-speaking domains). In few emergency cases they were made-up. Despite all efforts, odd or even grammatically inacceptable example sentences may occur. Native English speakers among the readers are encouraged to report weird or incorrect sentences to `cinkova@ufal.mff.cuni.cz`. Alternative suggestions are welcome!

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:

He ran into trouble. `CPHR`

Top price.`DENOM`[`is_member=1`] *and*.`CONJ` *top performance*.`DENOM`[`is_member=1`]

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:

He came and {`#PersPron.ACT`} *said(=He came and he said)*

{`#PersPron.ACT`[`tfa=t`]} *comes. (=He comes)*

[1] Some attributes mentioned here are specific to PDT but they are intended for the next version of this manual.

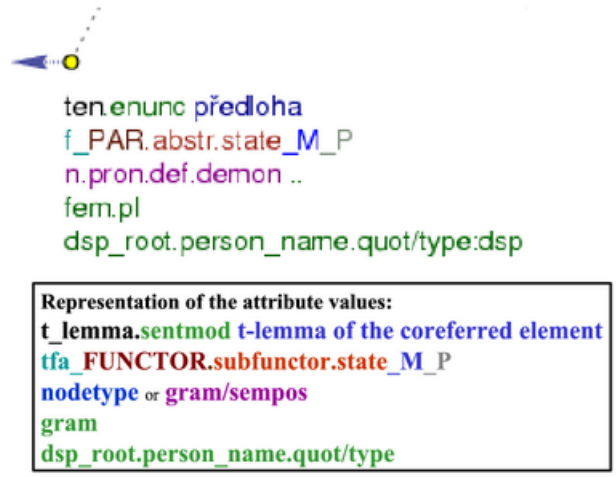
If it is necessary to stress that certain words are not represented by a separate node in the tectogrammatical tree, they are given in angle brackets < >. For example:

He <has been>sleeping.PRED

Example trees. For a number of example sentences, the corresponding example trees are included as well. Every example tree is equal to the complete analysis of the given sentence at the current annotation stage. This means that there are no grammatemes, TFA and very little coreference. In the figures it is only the tree structure, the shape of edges and nodes and functors that is the relevant information.

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; instead, the value of the `gram/sempos` attribute is given directly.

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

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 4.1.2, “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 4.4, “Parataxis”.)

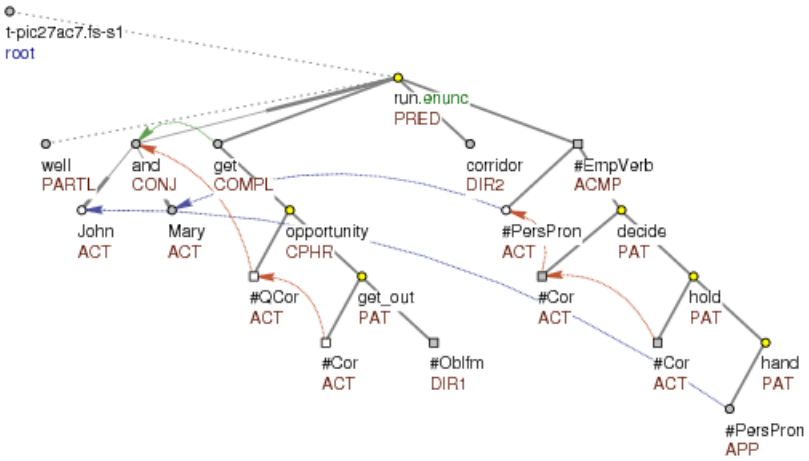
References. Attributes of the type reference, marking co-referential relations, are represented as arrows going from one node to another. They are actually ignored in the current annotation, though sometimes coreference arrows are inserted into the figures to give a better illustration of the semantic and pragmatic relations in the given sentence. Grammatical coreference (the `coref_gram.rf` attribute) is represented by an orange 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 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 4.1.1.1, “Predicative complement”) is represented by a green 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 is to be found in Fig. 2. The illustrating sentence is artificial and designed in a way to show as many types of arrows and edges as possible:

Well, John and Mary, having got the opportunity to get out, were running through the corridors, with her decided to hold his hand.

Figure 2. Edges and references



NB! In the example tectogrammatical trees (just like in the PDT trees; see Section 4.2.5, “Representing valency in the tectogrammatical trees”) the valency of nouns is not represented in full!

2. Building the Prague English Dependency Treebank and the Prague Czech-English Dependency Treebank

This manual describes the tectogrammatical annotation of English data. The data used is original PennTreebank data, which comprises about 50,000 sentences. About one half (approx. 22,000 sentences) were translated into Czech. In 2004, the English data as well as the Czech translated texts were converted into the PDT 1.0-shape and released by the LDC as Prague Czech-English Dependency Treebank. The tectogrammatical annotation level was generated upon the both data, and a tiny fraction of the English tectogrammatical data was manually annotated. In 2006 the data was converted into the PDT 2.0 shape. The second half of the PTB texts was translated into Czech. The original English manual tectogrammatical annotation got lost in the conversion. The current annotation of the English data was launched in the fall 2006. First we had converted the English valency lexicon PropBank into EngValLex to obtain a reference and supporting tool for the annotation of verbal frames. For more information about EngValLex see Section 4.2.4.3, “Valency lexicon”.

Chapter 1. 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 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 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 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, than 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 and the node for do as pro-form (dummy-do) in positive sentences (see Section 4.6, “Ellipsis” and Section 8.17, “Dummy-do (Verbal Pro-form)”). 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.

Chapter 2. 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. 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 2.1, "**Node types**".

Table 2.1. Node types

Node type	`nodetype`	Definition of the node type	Properties of the immediate daughter nodes
Technical root node of a tectogrammatical tree	`root`	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.

Atomic node	`atom`	Node with the functor:
		`ATT`
		`CM`
		`INTF`
		`MOD`
		`PARTL`
		`PREC`
		`RHEM`
			As a rule, it has no immediate daughters.

Paratactic structure root node	`coap`	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 (`functor` = `RHEM`)
			- conjunction modifier (`functor` = `CM`)
			- root node of a(n embedded) paratactic structure (`nodetype` = `coap`)

List structure root node	`list`	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 `#Idph` t-lemma form a tree structure. The items of a list under the node with the `#Forn` t-lemma are sisters with respect to each other and their functor is `FPHR`.

Node representing a foreign-language expression	`fphr`	Node with the functor: `FPHR`	It has no immediate daughters.
Node representing the dependent part of an idiom	`dphr`	Node with the functor: `DPHR`	As a rule, it has no immediate daughters.
Quasi-complex node	`qcomplex`	Node with the t-lemma: `#Amp`	Any node except for the types `root` and `fphr`.

		`#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`
			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.

Complex node	`complex`	All other nodes	Any node except for the types `root` and `fphr`. (Only) complex nodes have grammatemes.

Chapter 3. 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 3.1, "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),

b. the basic form of the word from which the word represented by the node was derived (see Section 3.2, "T-lemma of derived expressions (representing derivation in tectogrammatical trees)"),

c. a so called complex t-lemma, consisting of more words (see Section 3.3, "Complex t-lemma"),

d. an "artificial" value, a so called t-lemma substitute (see Section 3.4, "T-lemma substitute").

3.1. Basic form of the word

With nodes representing words present at the surface level and with so called copied nodes (see Section 4.6, "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 1, ***Relation between the tectogrammatical level and the lower levels***).

The basic word form is its representative form; i.e. the articleless non-genitive singular form of nouns, simple present active infinitive of verbs without *to* (incl. particle in phrasal verbs, see Section 3.3, "Complex t-lemma"), the positive degree of gradable adjectives and adverbs and positive forms of words that can be negated). Morphological meanings are captured by grammatemes, which have not yet been systematically studied for English.

Exceptions. Exceptions to this rule:

a. **Personal pronouns and their possessive counterparts** are represented by a node with the t-lemma substitute `#PersPron` (see also Section 3.4, "T-lemma substitute"). For example:

   My brother said: I am not going there. [`t_lemma` = `#PersPron`]

   For her I will do anything. [`t_lemma` = `#PersPron`]

   I don't share your opinion. [`t_lemma` = `#PersPron`]

   She is a good friend of mine. [`t_lemma` = `#PersPron`]

   This pen is hers. [`t_lemma` = `#PersPron`]

b. **Frozen finite verb forms, frozen participles, gerunds and infinitives** (with adverbial functions and in parenthesis; see Section 4.3.3.1, "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:

   To think we might have been in Rome now! [`t_lemma` = *think*]

   It was annoying, to say the least. [`t_lemma` = *say*]

   He is a good player, granted, but no better than William was. [`t_lemma` = *granted*]

   Judging with the Minister, it is going to be a problem. [`t_lemma` = *judging*]

c. **Foreign-language words** (with the `FPHR` functor; see Section 5.4, "Foreign-language expressions") are also represented by a node with the t-lemma corresponding to the actual word form.

3.2. T-lemma of derived expressions (representing derivation in tectogrammatical trees)

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 1, ***Relation between the tectogrammatical level and the lower levels***).

Word-formation (derivational) relations are represented in the original Czech corpus 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.

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

   The basic form of the derivative follows from the combination of the t-lemma and the assigned grammateme(s).

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).

This issue has not yet been systematically treated in the tectogrammatical representation of English. For more detail on derivation representation in tectogrammatical trees see the manual of the tectogrammatical annotation of Czech. The category of possessive adjectives is certainly not relevant for English. Semantically it is equivalent to the Saxon genitive.

3.3. Complex t-lemma

Complex t-lemma is assigned to nodes representing some multi-word lexical units (see Section 5.1, "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 complex t-lemmas:

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

   This concerns:

   a. **verbs with particles (phrasal verbs)**. The t-lemma of a node representing a verb with a particle), consists of the infinitive of the verb and the *particle*, connected by the underscore mark. For example:

      The plane took off. [`t_lemma` = *take_off*]

   b. **verbs with inherent reflexive pronouns**. The t-lemma of the node representing a verb with the appropriate form of the reflexive pronoun -*self*), consists of the infinitive of the verb and the reflexive *oneself*, connected by the underscore mark. For example:

      She always prides herself on her academic background. [`t_lemma` = *pride_oneself*]

   c. **complex conjunctions and conjunction pairs**. The t-lemma of a node representing a complex 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. For example:

      Either you will come, or I will. [`t_lemma` = *either_or*]

      Both kids and adults were there. [`t_lemma` = b*oth_and*]

d. **complex operators**. The t-lemma of a node representing a complex 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. For example:

It takes a long time, from spring to winter. [`t_lemma` = *from_to*]

e. **numbers with the function of a "label"**. The t-lemma of a node representing a (compound) number with the function of a "label" consists of all parts of the given number, linked by the underscore marks, in the surface word order. For example:

Call number 737 677 228. [`t_lemma` = *737_677_228*]

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

41-year old woman [`t_lemma` = *41_year*]

5-hour waiting time [`t_lemma` = *5_hour*]

g. **surnames containing *van*, *von*, *de*, *O'*, *Mc'*, *Mac*, 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. The prefixes O', Mc', Mac etc, that are attached to the surname without space inbetween are not separated by the underscore. Apostrophes are preserved. For example:

the composer Ludwig van Beethoven [`t_lemma` = *van_Beethoven*]

von Ryan's father [`t_lemma` = *von_Ryan*]

O'Reilly [`t_lemma` = *O'Reilly*]

B. the complex 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, `nodetype` = `dphr`; see Section 5.1.3, "Idioms") consists of the actual surface forms of the dependent parts (including prepositions) as present at the surface. For example:

We can talk now, I'm all ears [`t_lemma` = *all_ears*]

They appeared out of the blue. [`t_lemma` = *of_the_blue*]

3.4. T-lemma substitute

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

a. **newly established nodes that are not copies of other nodes** (see Section 4.6, "Ellipsis").

T-lemma substitutes: `#AsMuch`, `#Benef`, `#Cor`, `#EmpNoun`, `#EmpVerb`, `#Equal`, `#Forn`, `#Gen`, `#Idph`, `#Neg`, `#Oblfm`, `#PersPron`, `#QCor`, `#Rcp`, `#Separ`, `#Some`, `#Unsp`, `#Total` (see Table 4.8, "**Survey of newly established nodes**" in Section 4.6, "Ellipsis").

b. nodes representing **personal pronouns and their possessive counterparts, reflexive pronouns and the verbal pro-form (dummy-*do*, only in non-negated utterances)**.

T-lemma substitutes: `#PersPron`, `#VerbPron`, `#Rcp` (when substituting *each other/one another*)

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

For their t-lemma substitutes see Table 3.1, "**T-lemma substitutes of non-alphabetical/numerical symbols**".

Table 3.1. 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 3.2, "**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 3.2. 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 1	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, or *each other/one another*
#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
#VerbPron	complex	1	node representing the verbal pro-form (dummy-*do*) in positive (non-negated) utterances. It has the original *do*-node attached as auxrf and it inherits its grammatemes.

Chapter 4. Sentence representation structure

4.1. 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 ot 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 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 4.1.1, “Dual dependency”);
- existence of non-dependency edges (see Section 4.1.2, “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 4.1.3, “Ambiguous dependency”).

4.1.1. Dual dependency

The term ***dual dependency*** is used for 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 4.1.1.1, “Predicative complement”.

4.1.1.1. Predicative complement

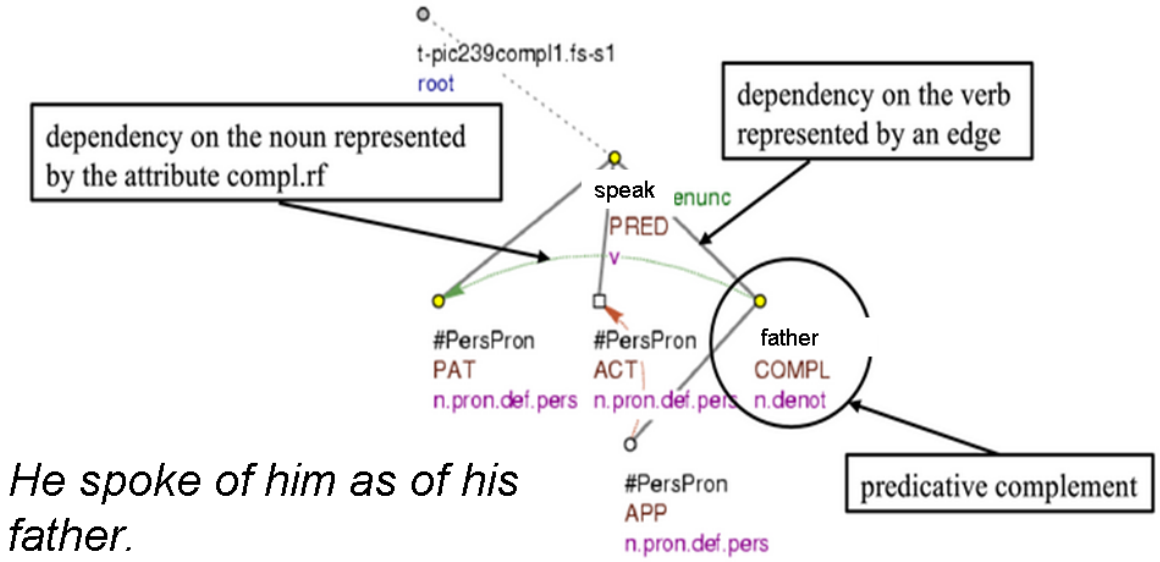
Predicative complements are (optional) adjuncts that have two semantic dependency relations, i.e. they simultaneously modify a noun and a verb (which can be nominalized).

Representing predicative complements in tectogrammatical trees. Nodes representing predicative complements always have the `COMPL` functor (see Section 6.11, “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`, the value of which is the identifier of the modified noun.

Representing predicative complements in tectogrammatical trees is illustrated in Fig. 4.1.

Figure 4.1. Structure containing a predicative complement



A predicative complement can be a noun, adjective or a non-finite verb form (participle, gerund). A predicative complement expressed by a noun (or an adjective) can be introduced by the conjunction as or it can be non-prepositional. Adverbial modifications and prepositional phrases are never considered predicative complements.

Examples:

They found their friend ill.COMPLFig. 4.2

I say it only as a layman.COMPL.

She was leaving defeated.COMPL (Fig. 4.3)

The manager approached us full.COMPL *of apologies.*

Full.COMPL *of apologies, the manager approached us.*

He left whistling.COMPL

She lay quiet.COMPL

He found her lying.COMPL*on the floor.*

She married young.COMPL

I like cofee black.COMPL

He was born blind.COMPL

He lives alone.COMPL

Beer is best cool.COMPL

He came running.COMPL

He went off staggering.COMPL

He died a beggar.COMPL

The door banged shut.COMPL

Julia, being.COMPL*a nun, spent much of her life in prayer and meditation.* (Fig. 4.5)

John, knowing.COMPL *that his wife was expecting a baby, started to take a course on baby care.*

He stood there, {#EmpVerb.COMPL} *his hands in his pockets.*

Construction of the type "He stood there, his hands in his pockets" The group of gerundial predicative complements includes also cases like "He stood there, his hands in his pockets". These are constructions in which the verb is followed by a gerundial construction with the governing gerund of the verb *have*, which is not present on the surface and we regard it as elided. The ellipsis is represented by a newly established node for an empty verb (t-lemma #EmpVerb, functor = COMPL).

He stood there, {#EmpVerb.COMPL} *his hands in his pockets.* (Fig. 4.4)

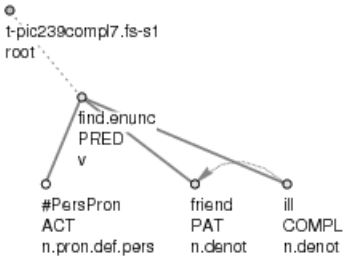
NB! Such a newly established node is used also in some cases of direct speech (t_lemma = #EmpVerb, functor = COMPL, here, this concerns the gerund form of the verb *say*. For more details see Section 5.5, "Direct speech".

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

He won the last election with a program designed as the negation.COMPL of the former *program*. (Fig. 4.6)

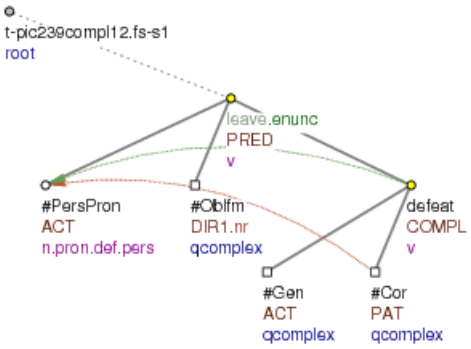
AirTran mails $100 vouchers as apology.COMPL *for delays.*

Figure 4.2. Predicative complement



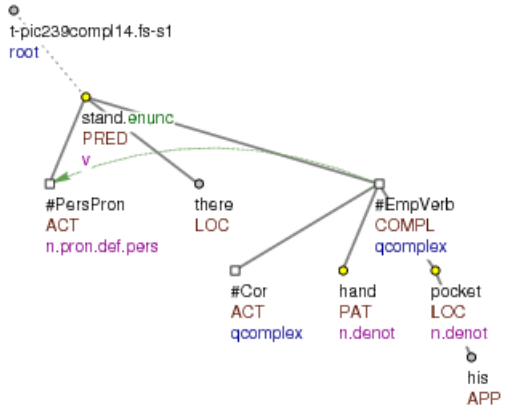
They found their friend ill.

Figure 4.3. Predicative complement



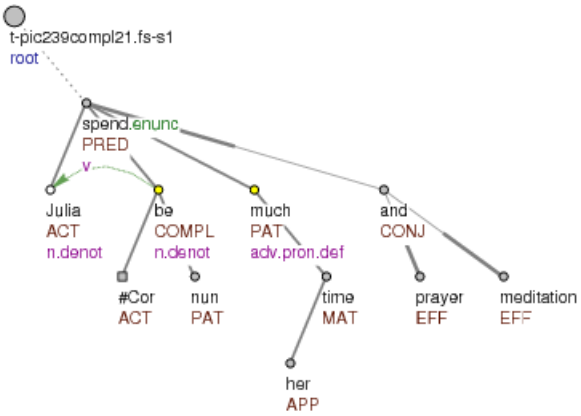
She was leaving defeated.

Figure 4.4. Predicative complement



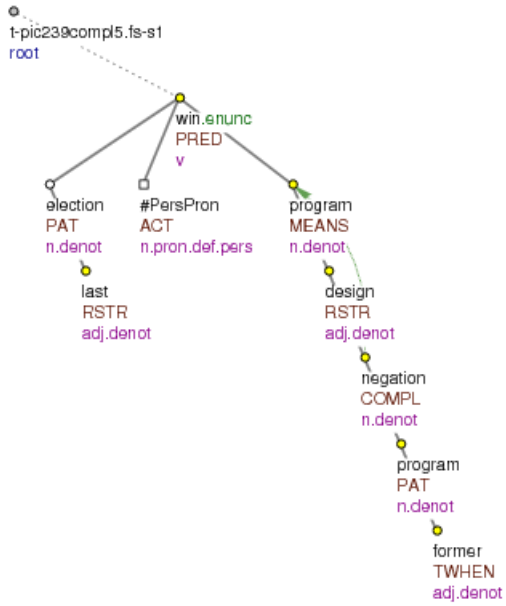
He stood there, his hands in his pockets.

Figure 4.5. Predicative complement



Julia, being a nun, spent much of her time in prayer and meditation.

Figure 4.6. Predicative complement



He won the last election with a program designed as the negation of the former program.

4.1.1.2. Constructions not regarded as predicative complements

As the English conception of predicative complement sometimes even includes nominal components of verbo-nominal predicates with copula verbs and copula-like verbs such as *seem, appear* etc. and postponed non-restrictive attributes, we find it appropriate to list types of constructions which this annotation does not regard as double-dependency predicative complements.

- **Obligatory subject/object complementations (members of verbal valency frames)**:

  They regard the case as closed.EFF.

  This made him happy.EFF

  He seems fit.PAT

  They elected him president.EFF

  Some obligatory subject/object complementations are easy to confuse with resultative constructions. Cf.:

  The door banged shut.COMPL (=*The door banged - when it became shut*)

  The screw worked loose.PAT (≠*The screw worked - when it become loose. The adjective is obligatory: *The screw worked.*)

- **Gerund/infinitive following a perception verb:**

  John heard Mary crying/cry.PAT (extra case marking, *Mary* is the Actor of *cry*) (Fig. 4.7)

- **Resultative constructions:**

  He painted the wall green.RESL

  She could laugh herself ill.RESL

  He walked his soils flat.RESL

- **Dependent parts of phrasemes:**

  This drives me nuts.DPHR.

- **Postponed non-restrictive attributes:**

  John, tired.DESCR, *went home.*

  Pierre Vinken, 50 years old.DESCR, *was elected chairman.*

  Julia, who was.DESCR *a nun, spent much of her life in prayer and meditation.*

  Only gerundial forms are regarded as predicative complements:

  Julia, being.COMPL *a nun, spent much of her life in prayer and meditation.*

- **Manner adverbials (even apparent adjectives):**

  The door fits tight.MANN (How does the door fit? *What is the door like when it fits? Besides, *tight* can be adjective as well as adverb, synonymous with *tightly*)

  She looks good.PAT (The valency frame for *look* in the sense 'to appear' requires a subject complementation)

  They sat quiet.COMPL (=They sat and were quiet).

  They sat quietly.MANN

  She married young.COMPL

  She dresses young.MANN (≠She is young when she dresses)

  When the complementation is optional, it clearly has adjectival form and the sentence can be paraphrased as "the subject was x when doing something", it is regarded as predicative complement (COMPL). When the complementation is obligatory in a valency frame it has an inner-participant functor. When the complementation is an adjunct and it has a form that is common for adjectives as well as adverbs it is regarded as adverb (MANN). When the complementation is an adjunct and it is an adjective (not having any adverbial counterpart) and the

sentence cannot be paraphrased with “the subject was x when doing something”, it is regarded as manner adverbial (MANN).

- **Nouns postponed to the end of the sentence, separated by a comma :**

  The subject or the object of a sentence can be further modified by a nominal phrase attached at the end of the sentence, separated by a comma. A sentence like

  He went out, an ugly old man.

  or

  They pushed him out, an ugly old man.

  is to be annotated differently from

  He went out, desperate.COMPL

  or

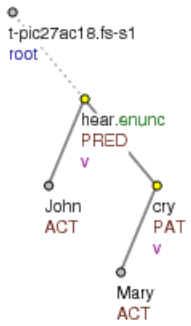
  He went out, whistling.COMPL

  Due to the large span between the two nominal phrases (the subject/object and this attached modifier) resolving this construction as an apposition appears somewhat awkward. Therefore such constructions are to be interpreted as coordinations:

  He.ACT *went*.PRED[is_member=1] *out* {#Comma.CONJ} {#EmpVerb}.PRED[is_member=1] *an ugly old man*.PAT (Fig. 4.8)

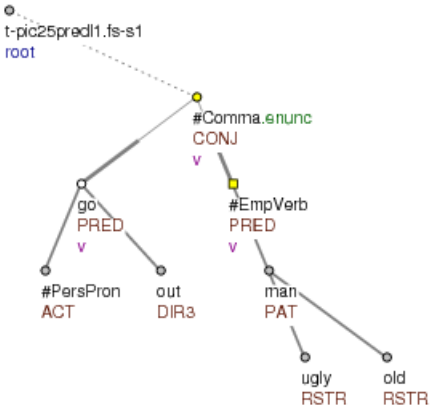
  This construction is thus neither to be interpreted as “..., being an ugly old man.COMPL”, nor “he, an ugly old man, went out” (*he* and *an ugly old man* in apposition). Instead, the interpretation is “He went out, he was an ugly old man”. #EmpVerb stands for a copula verb. The noun phrase is clearly the nominal part of a verbonominal predicate and therefore it gets the functor PAT. The Actor is not to be inserted since generally no valency frame is to be assigned to #EmpVerb.

  Figure 4.7. Gerund following a perception verb

  John heard Mary crying/John heard Mary cry.

Figure 4.8. Noun postponed to the end of the sentence, separated by a comma



He went out, an ugly old man.

4.1.2. 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 6.1, "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 4.4, "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 a 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 5.4, “Foreign-language expressions”. For more on list structures for identifying expressions see Section 5.3.1, “Identification structure”.

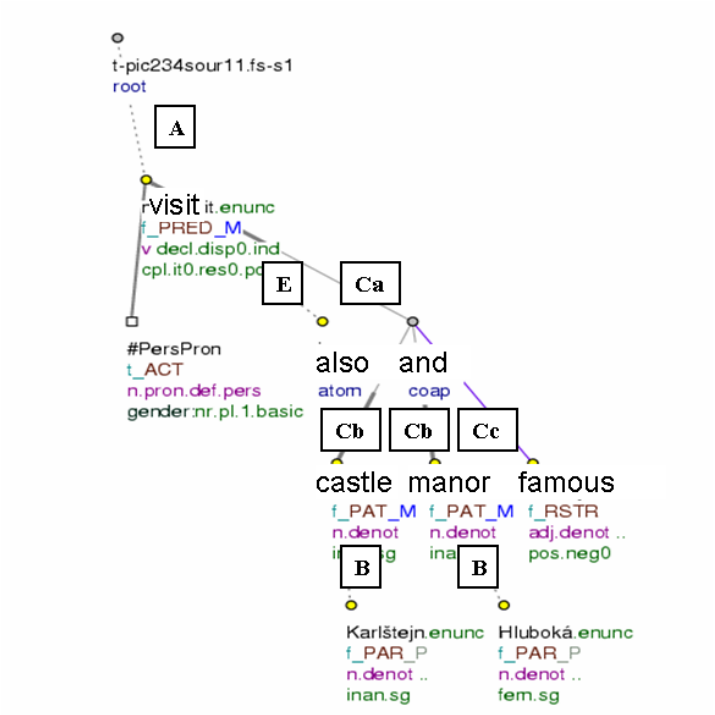
E. **the edge between an atomic node and its mother node**.

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 6.7, “Functors for rhematizers, sentence, linking and modal adverbials”.

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

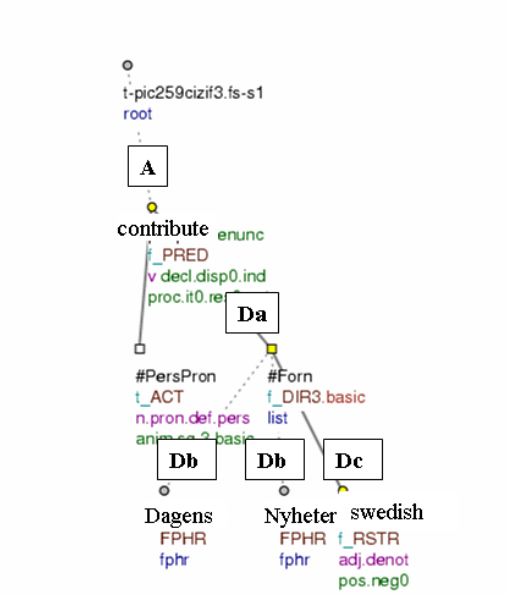
A node the functor of which is `DPHR` or `CPHR` expresses the fact that it constitutes a single lexical item together with its mother node (such a lexical item is usually represented by a single node). Therefore, the edge expresses the fact that the expressions belong together rather than dependency. For more details see Section 5.1, “Multi-word lexical units”.

Figure 4.9. Examples of non-dependency edges I



We are also going to visit the famous castle Karlstejn and manor Hluboka.

Figure 4.10. Examples of non-dependency edges II



He was contributing to the Swedish Dagens Nyheter.

4.1.3. 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 folllows:

- 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:

There was an old man coming in a shabby coat. (Fig. 4.11)

He was restlessly pacing around the room.

and even:

She was quick to shut the door.

Dual function of a single modification - except for verbal complement

In this respect, the current annotation of the English data differs from the original annotation convention for the Czech data. For 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), the Czech annotation makes the modification depend on the node representing the **lowest** modification it can modify and it is assigned the functor adequate to its real position. For instance (Czech data):

I bought a belt for seventy CZK.

= *For seventy CZK*.REG *I bought a belt for seventy CZK* . RSTR. (Fig. 4.13)

He repaid his debts to the insurance company.

= *To the insurance company*.ADDR *he repaid his debts to the insurance company*.ADDR

and also

= *To the insurance company*.ADDR *he repaid his debts to the insurance company*.ADDR (*debt* is regarded as nominalization - see Section 8.10, “Nominalizations” - and that's why its modification has a functor for inner participants)

The automatic pre-annotation of English data follows the original Penn Treebank annotation, whose guidelines say that the modification is to be put to the position most sensible in relation to the context, or, when all solutions would be semantically equally good, the modification is to be placed as **high** as possible. Our manual annotation respects the pre-annotation as far as the position of the modification is semantically acceptable.

There is no explicit indication that the modification has a dual function (unless it is a verbal complement, see Section 4.1.1.1, “Predicative complement”) in either the Czech or the English annotation. The original Penn Treebank annotation guidelines had introduced a special annotation resolution for marking multiple dependencies (“Permanent Predictable Ambiguity”). However, it has only been used 40 times in the entire corpus as it was very time-consuming and the annotators had clearly preferred to make one decision about the position of the modification instead. Those 40 occurrences of PPA are not reflected by our automatic pre-annotation at all.

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.:

   He stayed at home, in Prague.

b. **one temporal or locative/directional modification dependent on another**. The following cases are considered cases of one modification dependent on another:
   - **one modification is introduced by the preposition *of***. For example:

     He arrived on Thursday 5th (of) January.APP*1997.* (Fig. 4.15)

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

     It is located two kilometers.LOC *from the river*.DIR1. (Fig. 4.16)

     She had got dressed half.TWHENan hour.MAT *before the performance started*.TWHEN.

   - 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:

     He arrived soon.TWHEN *after Christmas*.TWHEN

     It happened far.LOC *from Moscow*.DIR1

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:

   We will meet at the Central station.LOC *in the hall*.LOC (Fig. 4.17)

   They arrived in February.TWHEN *1999*.TWHEN.

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.:

Give it to our director.`RSTR`*Novak*. (Fig. 4.18)

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.:

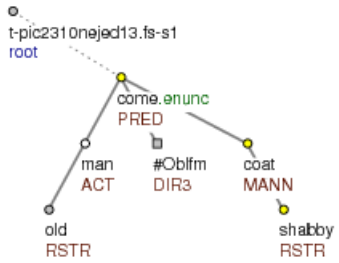
the most desired speaker **.** `DENOM` *Colonel*.`RSTR` *Jan*.`RSTR` *Landgerman*.`RSTR`

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 commnon noun and has the `ID` functor; e.g.:

There are steamboats on the Vltava.`ID` river. (Fig. 4.19)

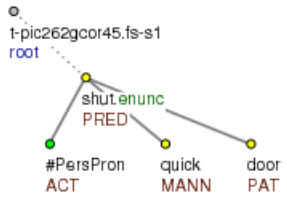
For more information on proper nouns with descriptors (common nouns) which are not integral parts of the name see Section 5.3, “Identifying expressions”).

Figure 4.11. Ambiguous dependency



The old man came in a shabby coat.

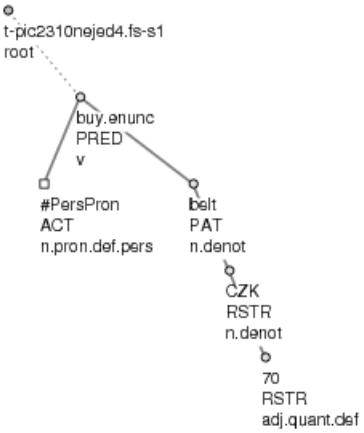
Figure 4.12. Dual function of a single modification



She was quick to shut the door.

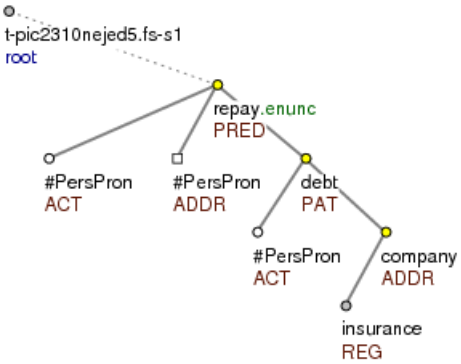
.

Figure 4.13. Dual function of a single modification



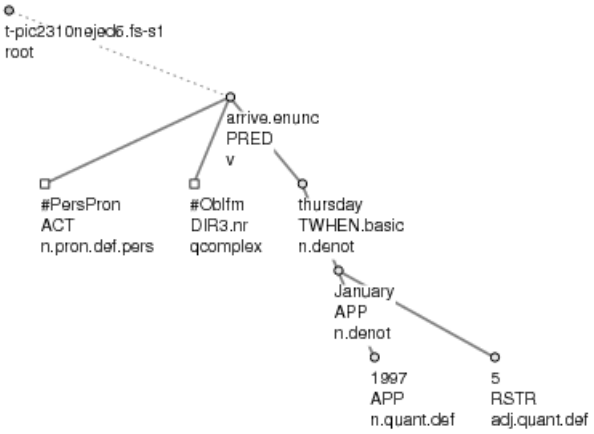
I bought a belt for CZK 70.

Figure 4.14. Dual function of a single modification



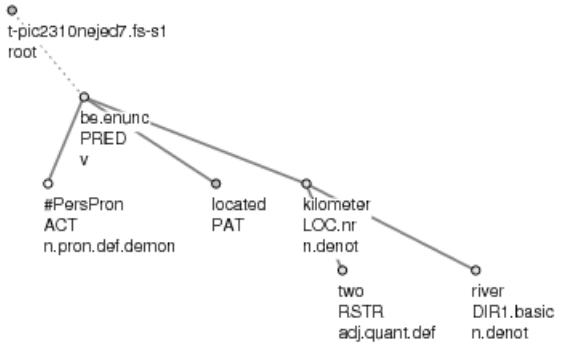
He repaid his debts to the insurance company.

Figure 4.15. Mutual relation of two temporal modifications



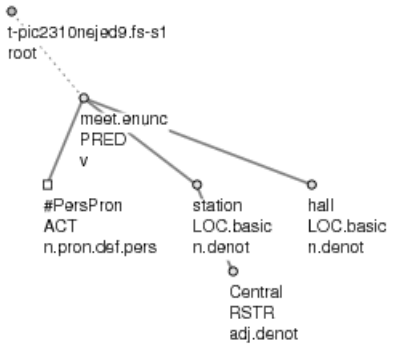
He arrived on Thursday 5th (of) January 1997.

Figure 4.16. Mutual relation of two locative/directional modifications



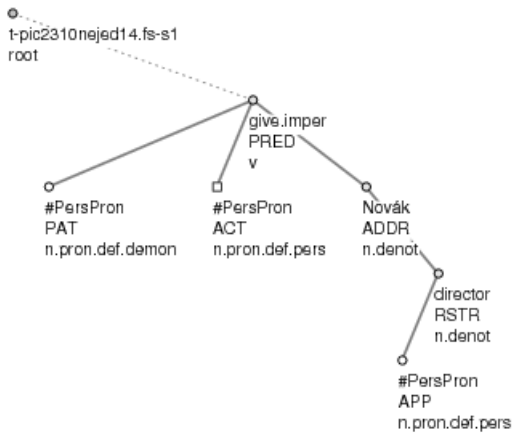
It is located two kilometers from the river.

Figure 4.17. Mutual relation of two locative/directional modifications



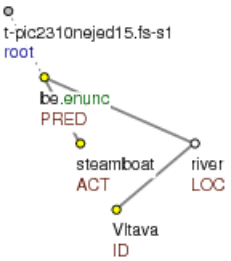
We will meet at the Central station in the hall.

Figure 4.18. Multi-word names



Give it to our director Novak.

Figure 4.19. Multi-word names



There are steamboats on the Vltava river.

4.2. 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 4.2.1, “Criteria for distinguishing different kinds of modifications”). These modifications are always specified in the valency frame of the lexical unit (see Section 4.2.4, “Valency frames and the way they are recorded in the valency lexicon”).

4.2.1. 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 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 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 6, ***Functors and subfunctors***).

Obligatory and optional modifications. A certain type of modification is either an argument, or an adjunct in all its occurences. 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 determing 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 necessarilly 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: *When he saw it he bought it.*

B: *Who?*
A: **I don't know.*

b. A: *When he saw it he bought it.*
B: *For whom?*
A: *I don't know.*

c. A: *When he saw it he bought it.*
B: *From whom?*
A: *I don't know.*

The verb *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 *Who?* by saying *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 *For/to whom?* and *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: *My friends have come.*
B: *Where?*
A: **I don't know.*

b. A: *My friends have come.*
B: *Where from? Why?*
A: *I don't know.*

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

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 4.1, "**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 4.1. Structure of a valency frame

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

4.2.2. 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 6.2, "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 (in English indicated only in personal pronouns). Only if the second argument is introduced by the preposition *to*, the semantics of the arguments comes into account. If the argument introduced by the preposition *to* refers to the Experiencer (or Agent), we consider this argument the Actor and the first - prepositionless- argument the Patient; cf.:

I.`ACT` *liked the book*.`PAT`

It seemed to John.ACT *that Mary was*.PAT asleep.

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. 4.20):

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.

Water.ACT*was boiling.*

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*has dug a hole*.PAT

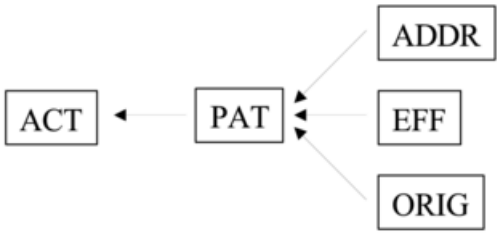
     Why did the arctic **.** ACT *turn from a greenhouse*.ORIG *into an icehouse*.PAT?

   - 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:

     The teacher.ACT*called the pupil*.PAT

     A beautiful oak.ACT*has grown from the acorn*.PAT

Figure 4.20. 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 4.2.1, "Criteria for distinguishing different kinds of modifications") is always assigned an adjunct-like functor; e.g.:

Petr.ACT c*ame to Prague*.DIR3.

The argument shifting applies to valency frames of all verbs, with the exception of complex predicates (for the discussion see Section 4.2.4.2, "Valency frames of idiomatic expressions and complex predicates"). Neither does the argument shifting apply to valency frames of nouns and adjectives not referring to events and to valency frames of adverbs.

4.2.3. 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 an adjunct or among different types of adjuncts. There are two ways to deal with these cases:

A. 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 *behave*:

   ACT(.1) MANN(*)|CRIT(*)|CPR(*)

   he behaves in a kind way.MANN; *he behaves in accordance with the rules*.CRIT; he behaves as if he hadn't seen.CPR *us*

B. in other cases of competition:

   competition of the Addressee and directional modification (*send a letter to grandma*.ADDR - *send a letter to the office*.DIR3,
   competition between locative and directional modifications (?p*lace a picture onto the notice board*.DIR3 - *?place a picture on the notice board*.LOC)
   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.

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

A valency frame record in the valency lexicon (EngValLex) is a sequence of records of the individual valency modifications (types of dependents), separated by spaces. Valency modification alternatives (see Section 4.2.3, “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.

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 the surface slot filler of the given modification (see Section 4.2.4.1, “Surface form of a valency modification (Surface frame slot filler)”). The question mark preceding a functor for inner participant indicates optionality; if the question mark is not present, the inner participant is obligatory. Free modifications that are not required (semantically obligatory) but which are typical for the given verb (noun, adjective) in the given meaning are also recorded (mainly originating from Argument roles in the PropBank Lexicon, from which EngValLex was converted). They are also marked with question mark.

Cf. the valency frame for one of the meanings of the verb *acknowledge*:

John.ACT *acknowledged to Mary.*?ADDR that he was.PAT *an idiot.*

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.

4.2.4.1. Surface form of a valency modification (Surface frame slot filler)

This section shows how surface syntax representations of the respective valency frame members are described. When referring to the frame members and their surface syntax representations, we will henceforth adopt the terminology coined by Zdeněk Žabokrtský (his PhD thesis):

1. (deep or surface) valency frame is a sequence of frame (deep or surface) slots
2. (deep or surface) frame slot contains a set of constraints on what can (or must) be filled into this slot,
3. (deep or surface) frame evoker is a part of the (deep or surface) sentence representation, which represents the frame-evoking lexical unit,
4. (deep or surface) frame slot filler is a part of the (deep or surface) sentence representation, which "saturates" one of the frame slots of the frame evoked by the frame evoker,
5. (deep or surface) frame instance is a part of the (deep or surface) sentence representation, in which the frame usage is manifested; frame instance consists of a frame evoker and frame slot fillers.

Deep frame slot fillers are the functors. We are now presenting the rules for recording their corresponding surface frame slot fillers.

In EngVALLEX, each deep frame slot filler (functor) corresponds to a set of typical/usual/possible surface frame slot fillers (surface representations), which are listed in round brackets following the functor name. Surface frame slot fillers are basically defined by PTB-tags and lemmas in the order 'lemma.tag', with lemma and tag separated by a dot without spaces inbetween. Often only lemma or only tag is enough to define the given surface frame slot filler. A surface frame slot filler defined only by lemma does not include the dot while the dot is obligatory when defining a surface frame slot filler by a tag or by an abbreviation. In some rare cases (idioms) a surface frame slot filler can be defined by its form instead of the lemma. The form is distinguished from the lemma by being enclosed by apostrophes ('), e.g. *'em*" (the colloquial form of *them* referring to the lemma *they*).

To save typing, we have introduced several abbreviations that substitute groups of regularly co-occurring PTB-tags. Some abbreviations are only defined by tags (first-degree abbreviations) but others are defined by tags as well as by tag-defined abbreviations (second-degree-abbreviations), and even others may be defined by tags and first- as well as second-degree abbreviations (third-degree abbreviations). We have not imposed any limitation on the 'degree' of an abbreviation but the highest degree used here is the third degree.

Apart from containing only the daughter node of the given verb listed as an EngVALLEX entry, which is the commonest case, the surface frame slot filler can be defined as a more complex – analytical - subtree, whose top-most node is the daughter of the given verb node listed as an EngVALLEX entry.

The surface frame slot filler is the form as found at the analytical level. The specification involves:

A. **indication of the syntactic dependency**. Square brackets ([ ]) are used to indicate descendant nodes. Square brackets allow nesting to indicate the dependency relations among the nodes of a given subtree. For example, the frame describing the idiom *come out of one's shell* will consist of the deep frame slot fillers with the functors `ACT` and `DPHR` respectively. `ACT` can be rendered as any noun or a personal pronoun in subject case or a wh-word/wh-clause, while the `DPHR` functor will at the analytical layer correspond to a prepositional group with certain lemma restrictions. When we regard *of one's shell* as analytical subtree of *come out*, the mother node defined by the preposition *of* would have the daughter *shell* (noun in basic case, singular only), which in turn will be the mother of the possessive pronoun without any lemma restriction. This relation will be indicated by the square brackets around *shell* containing another pair of square brackets with the possessive pronoun: `DPHR(of[shell.NNP[PP$]]).` NB that the possessive pronoun has no lemma restriction as it varies in morphological agreement with the subject of the sentence.

   Sister nodes are separated by a comma and listed from left to right; e.g.: `(House[the,White&]).` Daughter nodes that occur on the left side of the parent are separated from those that occur on the right side of the parent by ampersand (&). When all members of such list occur on the left side of the parent the list also ends with ampersand, e.g. `(House[the,White&]).`

B. **indication of the part-of-speech and morphemic properties, lemmas and forms**. A surface frame slot filler is either defined by its lemma or its tag. These two positions are separated by a dot. The dot is not obligatory with lemma-based definitions but it obligatorily precedes the tag when the surface slot filler is defined by a tag only. The notation is thus `(lemma.tag)` or `(lemma)` or `(.tag)` If a surface-level category is not specified, it means that the given valency modification may get any value of the category.

   The description allows indication of alternatives. Three types of alternatives are distinguished:

   I. Alternatives of the entire surface slot filler, e.g. n (nouns) or 'p' (personal pronouns). Such alternatives are separated by semicolon (;).

   II. Alternatives of lemmas or tags within a given surface representation. Such alternatives are separated by colon (:). E.g. `.VB:.VBG` means 'verb basic form OR gerund'.

   III. Open list of lemmas: lemmas are enclosed in curly brackets ({}), separated by commas. Three dots without space follow the last comma to indicate that the list is not complete. This is typically used with predicate nouns of support verbs in support verb constructions.

   A surface form of an obligatory adjunct is usually not specified, which expresses the assumption that the deep frame slot filler follows its usual morphosyntactic constraints. This is indicated by an asterisk (*) used instead of the explicit specification of the surface frame slot filler. With arguments, the surface frame slot fillers are always specified.

   The PTB tags are either used as they are, or, when appropriate, abbreviations grouping the tags are employed. The abbreviations have been introduced in order to

   - save the annotator's typing time
   - minimize the risk of typing errors
   - add additional information not included in the tags, e.g. the subject/object case in personal pronouns.

   The abbreviations have no semantics of their own. The abbrevation must be placed into its context in exactly the same way as the unabridged expression would. The abbreviations are defined in the file *engvallex.form-abbreviations*.

NB! The annotation of surface slot fillers in EngVALLEX has not been performed yet.

Table 4.2. Types of surface slot filler notation

Type	Example	Abbreviation	Notes
Alternation of various tags	`.JJ:.JJR:.JJS:.VBN:.VBG`	`a`	*lexemes that can act as adjectives: adjectives in positive, comparative and superlative, past and present participles*
Specification of *in* as a preposition (i.e. not used as particle in a phasal verb)	`in.IN`		
Specification of a tag and the requirement of a daughter with a given lemma and tag	`.VBN[with.IN]`		past participle governing its daughter by the preposition *with*
Lemma alternation	`{on,at,in}`		
Use of an abbreviation	`n[the.DET]`		`n` stands for noun - includes all tags for nouns. It would be the same to write `.NN:.NNS:.NP:.NPS:.CD:.DT:.WP:.WRB:.WDT`. It is a requirement for a noun with the definite article. The abbreviation for nouns has in fact a slightly wider definition than appropriate for this particular case; e.g. the tags beginning with W (interrogative pronouns) cannot be preceded by an article. To be absolutely precise, the annotator should list the N-starting tags as alternatives instead.)

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:

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

The company.ACT *built a house*.PAT.

Passive: *A house*.PAT *was built by the company*.ACT.

The valency frame of the verb *build*: ACTPAT

Sandra.ACT gave a book.PAT to John.ADDR.

Passive: John.ADDR was given a book.PAT by Sandra.ACT.

The valency frame of the verb *give*: ACT PAT ADDR

They.ACT *spoke* {#Rcp.ADDR} *about Martin*.PAT.

Prepositional passive: Martin.PAT was spoken about by them.ACT {#Rcp.ADDR}.

The valency frame of the verb *speak*: ACT PAT ADDR

People.ACT *believed that the earth was flat*.PAT./*People*.ACT *believed* the earth to be.PAT flat.

Passive: *It was believed by the people*.ACT *that the earth was flat*.PAT.

Passive with extra-case marking (ECM): *The earth was believed to be flat*.PAT *by the people*. ACT.

The valency frame of the verb *believe*: ACT PAT

George Washington.ACT *slept in this bed*.LOC

Prepositional passive: *This bed*.LOC *was slept in by George Washington*.ACT.

The valency frame of the verb *sleep*: ACT

The examples above suggest that it is necessary to specify which complementation has been passivized and in which way.

The following attribute values have to be filled in:

i. pass_functor (which complementation of the verb has been passivized)

ii. pass_phrase_type:

   a. NP_pass (a prepositionless noun has been passivized - the type *A house has been built* as well as *Sandra was given a book.* These two are distinguished by means of the pass-functor attribute value.)

   b. VP_pass (a verb phrase without extra case marking has been passivized - the type *It was believed that the earth was flat*)

   c. ECM_pass (a verb phrase with extra case marking has been passivized - the type *The earth was believed to be flat*)

   d. PP_pass (a prepositional phrase has been passivized - the type *Martin was spoken about* and *This bed was slept in*).

NB! Many verbs in active render the "dative" object by either a direct object or by a prepositional object introduced by the preposition to. It depends then on the given verb from which of the two possible forms the passive form has been derived. E.g. a passive sentence with *give* having the attribute values [pass_functor=ADDR] and [pass_phrase_type=NP_pass] would return the active constructions *X gives Y.ADDR Z* as well as *X gives Z to Y*, while verbs allowing only one of these forms would return only the appropriate one. This information will be later added into the lexicon when filling in possible surface frame slot fillers.

b. **gerundial clause** (see Section 8.2, “Gerundial clause”). Unlike gerundial clauses, action nominals (i.e. gerunds using the preposition *of* to express the direct object of a transitive verb) and nominalizations will have their own frames in the valency lexicon. Gerunds in existential constructions are regarded as nominalizations.

c. **causative constructions with *to have* and a past participle.** The surface form variants that are the result of a verb occuring in a resultative construction are not specified in the verb's valency frame. For example:

Someone.ACT*cleaned the windows*.PAT*for Sasha*.BEN

Causative: *Sasha*.BEN*had the windows*.PAT *cleaned* {#Gen.ACT}.

The valency frame of the verb *clean*: ACTPAT

d. **object-subject transpositions**. The surface form variants that are the result of a verb occuring in object-subject transpositions are not specified in the verb's valency frame. For example:

People.ACT*can break crystal*.PATeasily.MANN

Object-subject transposition (Middle Alternation): {#Cor.ACT}*Crystal*.PAT breaks *easily*.MANN {#Benef.BEN}

The valency frame of the verb *break*: ACTPAT

There are five types of object-subject transposition resolved in a very similar way (middle alternation - manner/condition, tough movement and object-subject transposition with a lexical verb). See Section 8.8, “Object-Subject Transposition *(Crystal breaks easily., John is difficult to please., This flat must have cost a lot to furnish.)*” for figures and more detailed analysis.

e. **exceptional case marking** (“object raising”, “nominative-accusative constructions”, see Section 8.1.2, “Infinitive-accusative constructions/exceptional case marking *(John expects Mary to leave)*” and Section 8.1.3, “Infinitive-accusative constructions/exceptional case marking with verbs of perception *(John hears Mary cry)*”). Object raising (*John*.ACT *expects Mary to leave*.PAT, *John*.ACT *hears Mary leave*.PAT) is not considered in the valency frame, i.e. the surface frame slot filler corresponding to the deep frame slot filler with the functor PAT. In the future the lexicon will contain the information which verbs enter this alternation.

f. **reciprocity**. The fact that there is a reciprocal meaning in the sentence is sometimes signalled by the presence of *each other* or *one another*. Another typical form used for expressing reciprocity is the form *between* and *among*. 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:

Anna met Jane./Ann met with Jane. → *Anna and Jane*.ACT *met*{#Rcp.PAT}.

Jill bantered with Sarah. → Jill and Sarah.ACT bantered {#Rcp.PAT}.

Anna kissed John. → John and Anna kissed <each other> {#Rcp.PAT[auxrf:each, other]}

a negotiation of the Prime Minister with the President) → a *negotiation between the Prime Minister and the President*

The valency frame of the noun n*egotiation* (for better exemplification, the surface slot fillers are explicitly named without using the PTB tags or abbrevaitions): ACT(genitive, possessive) PAT(of) ADDR(with)

his.ACT negotiation *of a new employment package*.PATwith another firm.ADDR

g. **understood reflexive object alternation**. Some intransitive verbs or verbs with an intransitive frame can employ the reflexive pronoun oneself. The action described by the intransitive variant is understood to be directed to the subject of the verb. The reflexive variant is associated to the intransitive frame, though the reflexive pronoun acquires the functor PAT in the tectogrammatical representation.

Jill dressed hurriedly. → Jill dressed herself hurriedly.

The (intransitive) valency frame of *dress*: ACT

h. **PRO-arb object alternation - characteristic property of the agent**. When a normally transitive verb is used without the object and it is supposed to describe a property characteristic of the agent, such as in:

The dog bites.

it is associated with the transitive frame. In the data the missing Patient acquires the t-lemma #Unsp.(e.g. the dog bites not generally everything but only strangers entering his master's lot etc.)

i. **PRO-arb object alternation - characteristic property of the instrument**. When a normally transitive verb is used without the object and it is supposed to describe a property characteristic of the inanimate agent, which in fact must be an instrument, such as in:

This knife doesn't cut.

it is associated with the transitive frame. In the data the missing Patient acquires the t-lemma #Gen.(e.g. generally everything that can be cut with a knife.)

j. ***way* object alternation**. A large number of intransitive motion verbs with a directional adverbial can add the object *one's way*:

They pushed past. → They pushed their way through the crowd.

These constructions refer to the frame: ACT DIR1|DIR2|DIR3 and *one's way* will have the functor PAT in the data. They are definitely to be held apart from any possible transitive frame as e.g.

They pushed the table to the window. (ACT PAT DIR1|DIR2|DIR3) or *They pushed the button.* (ACT PAT)

k. **conative alternation (partitive)**. The conative alternation makes an intransitive verb with a prepositional object (with the prepositions *at/on*) out of a transitive verb. The intransitive variant has the meaning of an "attempted" action without specifying whether the action was actually carried out with any considerable effect. The intransitive variant refers to the transitive frame and the prepositional object acquires the functor PAT in the data. The verb acquires the attribute value [partitive=1].

The mouse nibbled the cheese. → The mouse nibbled on/at the cheese.

l. ***with*-PAT preposition drop in ACT-PAT-ADDR verbs** ("fulfilling" alternation).

The judge.ACT *presented a prize*.PAT *to the winner*.ADDR. → The judge.ACT *presented the winner*.ADDR *with a prize*.PAT.

The first variant has been arbitrarily chosen as the basic one, i.e. its surface slot fillers will be listed in the lexicon. The other variant will refer to the same frame in the data but its surface form will be derived by a rule.

m. *together* **reciprocal alternation**. This alternation refers to the basic frame ACT PAT DIR3:

I.ACT *creamed the sugar*.PAT *into the butter*.DIR3

and in the data it can also be represented as

I.ACT creamed the sugar and the butter.PAT together.MANN.

The same goes for the intransitive variant with the basic frame ACT PAT:

The eggs.ACT *mixed with cream*.PAT

The eggs and the cream.ACT *mixed together*.MANN

n. ***apart* reciprocal alternation**. This alternation refers to the basic frame ACT PAT DIR1:

I.ACT broke the twig.PAT off/of the branch.DIR1

and in the data it can also occur as

I.ACT broke the twig and the branch.PAT broke apart.MANN

The same goes for the intransitive variant with the basic frame ACT PAT:

The twig.ACT broke *off/of the branch*.PAT

The twig and the branch.ACT *broke apart*.MANN

o. **substitution** of "what was said or thought" **with the pro-forms *so* and *not*** in verbs of saying and thinking (see Section 8.17.2, “The pro-forms *so* and *not* in verbs of saying and in cognitive verbs”).

p. **location subject alternation in transitive verbs**. See i

Regular changes in the deep form (not having their own frames in the lexicon and referring to a "basic" frame):

a. **induced-action alternation**. Some basically intransitive verbs can become transitive when entering the induced-action alternation, e.g.

The horse jumped across the fence. → Sylvia jumped the horse across the fence.

They marched out of the boss' office. → The boss marched them out of his office.

The rats ran through the maze. → The scientist ran the rats through the maze.

The baby burped. → The father burped the baby.

The valency frame of *jump*: ACT

The valency frame of *run*: ACT

The valency frame of *march*: ACT

The valency frame of *burp*: ACT

The PropBank-Lexicon, on which EngVALLEX is based, too, indicates cases of verb transitivizing by means of the induced-action alternation by assigning the object the argument label Arg0, which is normally reserved for agentive subjects, while assigning the formal subject of the clause an argument label reserved especially for induced-action agents (ArgA). However, the boundary between the induced-action-alternation pairs and common transitive/intransitive alternation pairs (e.g. *open*) is not sharp. E.g. verb uses marked in the PropBank-Lexicon annotation as induced (or “impelled”) action are sometimes described in the large learner's dictionary Macmillan as transitive without any further remarks (*march, diversify*), while the most typical induced-action example, the one of *jump* named above, has no transitive record in Macmillan.

The most practical solution would apparently be to ignore the induced-action alternation in general and to add all transitive uses of intransitive verbs as separate transitive frames into the lexicon (which has been the universally applied resolution of transitive/intransitive frames in EngVALLEX). This has basically how the induced-action alternation frames from the PropBank-Lexicon were treated during the manual correction of EngVALLEX. We assume, though, that the induced-action alternation is a productive mechanism which can potentially be applied in many more verbs than listed as such in lexicons, mainly in occasional uses. Therefore we have preserved the induced-action alternation as one of the possible syntactic transformations verbs can undergo.

b. **location subject alternation and "induced" location subject alternation in intransitive verbs.** Levin lists so-called fit verbs as those entering the location subject alternation. The typical example of this alternation is, according to Levin:

We sleep five people in each room.

Each room sleeps five people.

The *fit* verbs include transitive verbs (*carry, contain, fit, feed, hold, house, seat, serve, store, take, use*) as well as intransitive verbs (sleep). We consider the transitive verbs irrelevant for this case (they will be mostly split into two frames) and will only concentrate on apparently intransitive verbs like *sleep, sit* etc. As a rule, these verbs' frame comprises only the functor ACT:

*Five people.*ACT *sleep in each room.*

*Five people.*ACT *sit in the cabine.*

When they happen to occur in transitive uses in the data, they will refer to this intransitive frame. Two alternations are thus possible for them:

i. the location subject alternation (narrower defined by location in the position of the syntactic subject):

*Each room.*LOC *sleeps five people.*ACT

ii. the "induced" location subject alternation (narrower defined by an agentive subject marked with the functor ACT, the location preserved as LOC and the original Actor renamed to PAT):

*We.*ACT *sleep five people.*PAT *in each room.*LOC

Alternations whose surface slot fillers are listed in the lexicon (within one common deep frame slot filler):

a. ***With* preposition drop in ACT-PAT verbs**. A small set of verbs that involve potentially reciprocal actions take either the direct object or the prepositional object introduced by the preposition *with*. As the name of the alternation suggests, it is the variant with the prepositional object that is considered to be the basic one. The transitive variant refers to the intransitive frame. Both surface slot fillers match the same functor (usually PAT):

Jill **.** ACT *met with Sarah* **.** PAT. → *Jill* **.** ACT *met Sarah* **.** PAT.

b. **dative alternation**. This applies to transitive verbs whose direct object gets the functor PAT and the alternating object has the semantic function of ADDR (*send a letter to Tom* but not **send a letter to London*). As for the dative alternation, there are three groups of verbs:

- verbs that only can have double object (*ask, bear, deny, envy...*)
- verbs that only can have the prepositional object with to as the second object (*address, administrate, surrender something to an authority, deliver*)
- verbs that allow the alternation (*advance, carry, give, send...*):

Bill sold a car to Tom **.** ADDR. → *Bill sold Tom* **.** ADDR *a car.*

The possible surface frame slot fillers are listed in the lexicon within the ADDR slot.

c. **benefactive alternation**. Benefactive as understood by Levin is resolved as optional ADDR or typical BEN in EngVALLEX. As for the benefactive alternation, there are two groups of verbs:

- verbs that allow the alternation of the direct object and the prepositional object introduced with the preposition *for* (*bake, build, design...*)
- verbs that require only *for* (*compose, invent, manufacture, confiscate...*)

The possible surface frame slot fillers are listed in the lexicon within the ADDR/BEN slot (when marked in the lexicon).

d. ***search* alternation - *for* drop**. The *search* alternation contrasts the frame

*Ida.*ACT *hunted the woods.*PAT *for deer.*EFF

with another - separately listed - frame

*Ida.*ACT *hunted deer.*PAT *in the woods.* or *Ida.*ACT *hunted for deer.*PAT *in the woods.*

It is only the second frame that is of interest here. Only some verbs accept the direct object as well as the prepositional object introduced by *for*. Some verbs accept only *for* and some accept only the direct object. These variants of surface slot fillers are to be explicitly stated.

4.2.4.2. Valency frames of idiomatic expressions and complex predicates

Idiomatic expressions (see Section 5.1.3, “Idioms”) and complex predicates (see Section 5.1.2.5, “Support verb constructions”) 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 *to be all ears*:

`DPHR([ears.NNS[all]])`

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:

John.`ACT`*gave Mary*.`ADDR` her *salary*.`PAT`

The valency frame for one of the meanings of the predicate *to give*: `ACT`(sub) `PAT(obj1)ADDR(obj2;[to[obj2;wh+v]])`

John.`ACT`*gave Mary*.`ADDR`*the order*.`CPHR`*to come.*

The valency frame for the complex predicate *to give an order*: `ACT(sub)CPHR`({order;chance;right;...}) `ADDR(obj2;[to[obj2;wh+v]])`

He has made this doll.`PAT`

The valency frame of the predicate in the unmarked use *make*: `ACT`(sub) `PAT`(obj1).

He made a favorable impression.`CPHR` *on me*.`ADDR`

The valency frame of the complex predicate *to make an impression*: `ACT`(sub) `CPHR`({impression,...}.4) `ADDR`(on[obj1]).

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.

4.2.4.3. Valency lexicon

The valency lexicon (EngVALLEX) contains valency frames of verbs. In the future, valency of nouns, adjectives and adverbs will be included. Individual valency frames are clustered on the basis of what t-lemma they are related to (for the discussion of t-lemmas see Chapter 3, ***Tectogrammatical lemma (t-lemma)***).

The valency lexicon of verbs is based on the PropBank-Lexicon, which was automatically converted into an FGD-compliant shape and manually corrected. Further corrections are still necessary and are being performed during the annotation. In the future, NomBank will be converted in a similar way to obtain valency frames for nouns.

As the PropBank-Lexicon was developed for the annotation of the Penn Treebank - Wall Street Journal section, generally only those verbs- i.e. those of their meanings - are included, which occured in the analyzed data (though sometimes made-up sentences had been added by the creators of the PropBank-Lexicon). 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 basic information on transforming the PropBank-Lexicon into EngVALLEX as well as on its main features in comparison to the PDT-VALLEX is given below.

The PropBank - Lexicon

The verb entries of the PropBank-Lexicon are divided into rolesets. Rolesets roughly correspond to senses. Each roleset includes a set of labeled arguments ("roles") and one or more example sentences, in which combinations of the roles are rendered by the surface syntax. Rolesets are numbered within the respective entries ("roleset-ID's"). In addition, each roleset has a definition-like description attached ("roleset names"). E.g. the verb *to yell* has only one roleset, `yell.01`, which is labeled as "to cry out loudly". The verb *to abandon*, on the other hand, has three rolesets (`abandon.01.-03`). They are labeled as "leave behind", "exchange", "surrender, give_over", respectively. The PropBank-Lexicon comprises about 2000 roleset names, in which about 4600 rolesets are grouped. (There are 3323 verb entries in the PropBank-Lexicon. Some of them also include phrasal verbs. Phrasal verbs do not have entries of their own but are displayed as rolesets).

PropBank's conception of valency derives from Levin's assumption that the syntactic alternations verbs participate in are not arbitrary but reflect underlying semantic components of the events denoted by each given verb. Semantically related verbs can be grouped into classes according to which alternations they enter. The roleset names group semantically and syntactically related verb senses into classes like the Levin classes. The PropBank classes go somewhat across the Levin verb classes in accordance with the valency behavior of the living data in PTB-WSJ.

Each roleset introduces an enumeration of arguments (roles). The arguments are divided into "numbered arguments" and "adjuncts". The numbered arguments are arguments that take part in the syntactic alternations analyzed by Levin and can become syntactic subjects. The adjuncts are optional, often rendered by prepositional groups and adverbs. Each argument has two parts: the argument number and a semantic descriptor specific to the given roleset. E.g. *to yell* would acquire the following arguments:

`Arg0:`Yeller

`Arg1:` Utterance

`Arg2:` Hearer

The first roleset of *to abandon* (`abandon.01` "leave behind") will have the following arguments:

`Arg0:` abandoner

`Arg1:`thing abandoned, left behind

`Arg2:`attribute of arg1

The arguments do not have to be all present on the surface of a sentence at the same time. Thus the first example sentence

And they believe the Big Board`.Arg0`, *under Mr. Phelan, has abandoned their interest*`.Arg1`

contains only `Arg0` and `Arg1`, while the second example sentence contains all three:

John`.Arg0` *abandoned his pursuit*`.Arg1` *of an Olympic gold medal as a waste*`.Arg2` *of time.*

Considering the syntactic alternations as pairs of alternation realizations, one can often identify alternation realizations in the rolesets. Each example sentence is provided with a supplementary comment, which even sometimes suggests which alternation realization the given sentence represents. Yet these comments are not formalized, nor is it explicitly stated by which alternations the respective rolesets are defined. About one half of PropBank entries are mapped onto VerbNet, in which relevant alternations are listed for each verb entry. However, the linking between the PropBank-Lexicon and VerbNet does not reach as deep as to the respective example sentences. Besides that, the PropBank verb classes (i.e. the roleset names) do not correspond to the VerbNet verb classes (i.e. the

original Levinian verb classes), and thus the example sentences in the PropBank-Lexicon do not necessarily show the same alternation patterns as the corresponding entry in VerbNet.

The structure of (PDT-)VALLEX

The Czech valency lexicon describes the valency behavior of a given lexeme (verb, noun, adjective or adverb) in form of valency frames, which roughly correspond to senses. Like rolesets in the PropBank-Lexicon, the valency frames primarily rely on syntactic criteria though the syntactic criteria are "softened" with regard to the semantics of the given verb. A valency frame in the strict sense consists of inner participants and obligatory free modifications. Free modifications are prototypically optional and do not belong to the valency frame in the strict sense though some frames require a free modification (e.g. direction in verbs of movement). Both the obligatory and the optional inner participants belong to the valency frame in the strict sense. Like the free modifications, the inner participants have semantic labels according to the cognitive roles they typically enter: `ACT` (Actor), `PAT` (Patient), `ADDR` (Addressee), `ORIG` (Origin) and `EFF` (Effect). However, if a verb only has one inner participant, it is automatically labeled with `ACT`. A two-participant verb always has an `ACT` and a `PAT` (see Fig. 4.20).

The valency lexicon of Czech verbs has two branches – the PDT-VALLEX and VALLEX. PDT-VALLEX is represented by the lists of valency frames being created and used by annotators during their work. It contains valency frames of words (verbs and nouns) in their particular meanings (as they appear in PDT) and serves for consistency of annotation. VALLEX is represented by the valency lexicon, in which the words are analyzed in all their meanings. Rich syntactic annotation is assigned to particular valency frames, including e.g. control and reciprocity.

PDT-VALLEX and VALLEX are very similar in structure: each lexeme corresponds to one entry. The entry is divided into valency frames. A valency frame is modeled as a sequence of frame slots. Each frame slot corresponds to one complementation of the verb in question. Each slot is assigned a functor according to its semantic relation towards the governing verb. Each slot includes an enumeration of its surface forms (surface frame slot fillers). Each frame includes at least one example sentence. PDT-VALLEX notes only the valency frames in the strict sense, i.e. obligatory or optional inner participants and obligatory free modifications), while VALLEX also lists optional free modifications typical of the given frame.

When delimiting the respective valency frames, syntactic as well as semantic criteria are adopted. Therefore a verb can have two valency frames with identical distribution of functors. Compared to the PropBank-Lexicon, the distinction of the respective valency frames appears to be somewhat more fine-grained in the annotations of VALLEX and PDT-VALLEX, which, in any case, are well aware of the absence of reliable semantic criteria, and also prefer syntactic distinction criteria to the semantic. The enormous word formation potential of Czech makes it difficult to build a list of surface syntax alternations like Levin's as their realizations are rendered by different verb derivates.

Similarities and differences

Both the PropBank and the VALLEX-style approaches assume that sense distinctions are reflected in varying valency frames and both look into their underlying syntax, employing semantic judgments. Both provide the opened valency slots with labels. The philosophy of labeling is yet different. While FGD employs rather general semantic labels, the PropBank-Lexicon combines the sheer numbering of arguments with verb-specific semantic descriptions. These verb-specific descriptions in their turn follow some regularities within the respective verb classes (as they are suggested by the names of rolesets). This implies that there is no straightforward matching between a given Arg and a given functor but relations within an entire group of verbs have to be taken into account to assign correct functors to the given Args when building EngVALLEX.

The most substantial difference between the PropBank-Lexicon and VALLEX lies in the very conception of the frame. The PropBank-Lexicon annotation observes alternation patterns in verbs to merge them into rolesets while VALLEX does not display any relations between frames within one lemma. The PropBank-Lexicon example sentences within a given roleset often represent realizations of a given alternation pair, e.g. "Intransitive - Inchoative" and "Transitive - Causative". Both sentences would refer to a roleset with the listed arguments `Arg0` and `Arg1`, although `Arg0` never emerges in the inchoative sentence. Vallex would separate such instances into two frames, the inchoative having only the "Actor" label, the causative having an Actor and a Patient. In FGD, agentivity is basically not an issue. All first inner participants (typically represented by syntactic subjects) are Actor-labeled.

Conversion of the PropBank-Lexicon into EngVALLEX

Pre-EngVALLEX

Pre-EngVALLEX is the product of the automatic conversion of the PropBank-Lexicon into an FGD-compliant form. It has the following features:

1. The PropBank-rolesets were automatically split into their respective example sentences. Each example sentence got a header with functors, which it inherited from the list of roles located at the beginning of the particular PropBank-roleset it used to belong to. The example sentences are now considered as preliminary FGD-valency frames.

2. The argument labels (Args) were turned into functors by means of a handful of simple rules.
3. Each preliminary valency frame has its own ID saying which roleset the original sentence used to belong to.
4. Unlike in the PropBank-Lexicon, the respective Pre-EngVALLEX files are not grouped as lemma files but according to the respective roleset names.

Assigning Functors to Args

Due to the difference in theoretical approaches no straightforward mapping could be performed. We had to make use of all hints the xml-data was offering, mainly the non-formalized attribute “role descr”. Even the mapping of the PropBank-Lexicon to other lexical sources was exploited. About 50% of the PropBank-Lexicon is mapped onto VerbNet, which uses its own semantic labeling (the attribute "vntheta" in the xml data). The semantic labels from VerbNet mainly helped to classify adjuncts (ArgM's) and Args with higher numbers.

Rules for ArgM's and higher Args typically looked like this: If <role descr="low point" n="3", functor: `DIR1` If <role descr="instrument" n="5", functor `MEANS` If <role descr="medium" n="5", functor `DIFF`. `CAU` → `CAUS` `PRP` → `AIM` `MNR` → `MANN` (but whenever the frame contains an `ArgA` it should be `EFF`). All slots corresponding to arguments with the original "agentive-subject" (`Arg0`) got the Actor functor, unless the `ArgA` was present. The `ArgA` always got the Actor-functor. The `Arg1` always got the Patient functor, unless the `ArgA` was present etc. The annotation work has proved that `Arg-M`s, which roughly correspond to free modifications, were assigned quite correctly. Yet the re-labeling rules originally assumed that the frames in Pre-EngVALLEX would be defined by the Args used in the example sentences. The more recent technically-motivated decision to use the entire list of Args from the roleset beginning and to ignore the example-sentence annotation made the set of rules less powerful. E.g. the manual correction revealed that Actor had been systematically interchanged with Patient in all intransitive sentences that had a transitive counterpart within the same roleset.

However, it was clear already at the beginning that the functors would have to be manually corrected anyway. The rules were only meant to save the annotators' typing time, and they proved powerful enough to serve this particular purpose. Therefore no evaluation was performed.

Links between the PropBank-Lexicon and Pre-EngVALLEX

Pre-EngVALLEX seeks to preserve as much original information of the PropBank-Lexicon as possible. Therefore each preliminary valency frame has its own ID saying which roleset it used to belong to as its example sentence. This ID remains even if the given preliminary valency frame is later merged with another preliminary valency frame during the manual adjustment. This ID ensures that each preliminary valency frame bears the same links as the original PropBank-roleset, including the links to other lexical sources as well as the links to the original corpus annotation.

Each functor within the given frame is linked to the original Arg or "role" of its original PropBank roleset. NB that the functors are linked to the list of roles at the beginning of each roleset, not to the corresponding Args in the annotation of the corresponding example sentences, which is not preserved in Pre-EngVALLEX. This linking is of special importance for transferring the alternation information from the PropBank-Lexicon into EngVALLEX. To illustrate the linking policy, lets take up the Causative-Inchoative alternation case again: the (made-up) pair of sentences

John opened the door.

and

The door opened.

give two preliminary valency frames of the verb *to open*; the former transitive and the latter intransitive. The original roleset had the following Args: `Arg0` (the agent opening the patient) and `Arg1` (the thing opened). The transitive original example sentence had both the `Arg0` (John) and `Arg1` (the door), while the intransitive example sentence only had `Arg1` (the door). According to our linguistic conventions set before, the frames will be kept separate in EngVALLEX, as they are now in Pre-EngVALLEX. As noted above, the valency theory of FGD requires the first argument of a verb to be `ACT`. Therefore, in the Causative-Inchoative Alternation pair, both the functor `PAT` in the transitive frame and the functor `ACT` in the intransitive frame will be linked to the original `Arg1`. As the annotation seeks to treat all major alternations as consequently as possible, this shifting of `PAT` to `ACT` will be characteristic of this type of alternation, as shifts of other functors will be characteristic of other alternation types.

Rearrangement of the XML-Files for manual editing

The editing tool FrameEditor opens a file with a roleset name, in which all rolesets (verb senses) called one particular name are gathered, no matter which lemma they actually belong to. This rearrangement was made in order to ensure that the annotators keep the consistency of the original PropBank-verb-class annotation.

Transforming rolesets into valency frames

Pre-EngVALLEX split the PropBank-rolesets into preliminary FGD-valency frames consisting of one example sentence each. The first task of the annotators is to merge certain types of frames. The frames are merged when:

1. the sentences in question have the same surface syntax structure.
2. a pair of sentences in question belong to an alternation whose realizations are to be merged in compliance with our linguistic conventions. E.g. a sentence with Unspecified Object is to be merged with a Transitive sentence, and so is a sentence with a Reciprocal Object, etc.

In some regular cases, complementations regarded as Args by the PropBank-Lexicon would neither have been classified as inner participants nor as obligatory free modifications, i.e. would not have been regarded as members of the frame in the strict sense, and therefore should not be listed in the valency lexicon. In order not to lose this part of the ready-made PropBank annotation, we included them as "typical free modifications" (indicated by a question mark in front of the name of the given functor), which has been an approved practice of VALLEX.

Preserving the alternation information

Linguistic rules were set for the commonest pairs of alternation realizations as to whether EngVALLEX would list each realization as a separate frame, or whether both should be merged into one frame. When annotators recognize a sentence pair as an alternation pair or a single sentence as an unpaired realization of a particular alternation, they are supposed to follow the appropriate rule.

The following situations can occur in EngVALLEX during the processing of the example sentences (i.e. the frames automatically generated from a PropBank roleset):

1. The frames are merged as the sentences only give variants of surface representations, irrelevant to the verb frame. E.g.: Unspecified Object, Instructional Imperative and Reciprocity.
2. The frames are merged as either alternation realization makes an Arg to an optional free modification (FGD). This makes it fit into the frame of the other alternation realization. E.g. Instrumental Subject.
3. One of the alternation realizations is regarded as a derivation of the other. Its tectogrammatical tree structure looks different from the lexicon frame but it refers back to it. Rules are stored to generate such trees from the lexicon frames. E.g. Induced Action, Middle Alternation, Location Subject Alternation.
4. The frames remain split and each frame acquires functors of its own. E.g.: Causative-Inchoative alternation, Substance/Source emission alternation.

4.2.5. 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 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 4.6.1.2, “Ellipsis of a dependent meaning unit”.

The present state of valency representation in PEDT. Not every case of valency requirements is represented properly in the tree structures. Valency is only represented properly with nodes with verb t-lemmas.

At the end of the annotation and after all necessary checks for all checked notes will 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).

4.3. Clauses (governing, dependent, verbal, verbless)

Sentences (represented by tectogrammatical trees) are formed by one or more clauses. Annotation of clauses differs depending on whether the clause is verbal or verbless (see Section 4.3.1, “Verbal and verbless clauses”), and on the kind of dependency (see Section 4.3.2, “Dependent and independent clauses (clause connecting)”).

4.3.1. Verbal and verbless clauses

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

verbal clauses (see Section 4.3.1.1, “Verbal clauses”),
verbless clauses (see Section 4.3.1.2, “Verbless clauses”).

4.3.1.1. Verbal clauses

A ***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 4.5, “Parenthesis”), its effective root has the `PAR` functor (see Section 6.1, “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**.

  Father is sleeping.`PRED`

- **non-finite verb form**.

  Manchester United defeated.`PRED`*!*

  We don't know where to go.`PAT`

- **contextual ellipsis of a predicate** (see Section 4.6.1.1, “Ellipsis of a governing meaning unit”).

  Father is sleeping, Mother {*is sleeping*.`PRED`} too.

- **grammatical ellipsis of a predicate**. Also those clauses are considered verbal in which there is no verb but which suggest a verb missing (and not being known from the context as in cases of textual ellipsis). 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 4.6.1.1, “Ellipsis of a governing meaning unit”). It is especially the following construction types:

  1. Lexicalized verbless glosses to preceding utterances, typically adverbs:

     {`#EmpVerb.PRED`} *Of course.*`ATT`

     {`#EmpVerb.PRED`}*Undoubtedly*.`ATT`

     {`#EmpVerb.PRED`} Very well.`ATT` !

  2. Noun phrases in the sense of imperative (followed by *please* or a manner or a time adverbial or any other evidence that it is the case of imperative). The nouns are supposed to be Patients of the hypothesized missing verb:

     {`#EmpVerb.PRED`} *Attention*.`PAT`*, please! (= pay attention)*

     {`#EmpVerb.PRED`} *Water, quickly! (= bring water)*

     {`#EmpVerb.PRED`} *Surgeon, immediately! (= bring/call a surgeon)*

     {`#EmpVerb.PRED`} *Another coffee, if you don't mind. (= make/bring another coffee)*

     {`#EmpVerb.PRED`} *Scissors, somebody! (= bring scissors)*

  3. Noun phrases acting as questions in the sense of offers and invitations. The nouns are supposed to be Patients of the hypothesized missing verb:

     {`#EmpVerb.PRED`} *Cigarette*.`PAT`?

     {`#EmpVerb.PRED`} *Another coffeee*.`PAT`?

     {`#EmpVerb.PRED`} *Tea*.`PAT` *or*.`DISJ` *coffee*.`PAT`?

4. Noun phrases with the force of wh-questions:

   {#EmpVerb.PRED} *Your name*.ACT?

5. Exclamatory prepositional phrases beginning with *of all* and expressing strong disapproval:

   {#EmpVerb.PRED} *Of all*.EXT *the impudence*.PAT !

   {#EmpVerb.PRED} *Of all*.EXT *the stupid*.RSTR *things*.PAT *to say*.RSTR!

6. Exclamatory constructions with *what a*:

   What.RSTR *a beautiful*.RSTR *day*.ACT {#EmpVerb.PRED} !

7. Exclamatory adjective phrases expressing approval or disapproval:

   {#EmpVerb.PRED} Very interesting.PAT !

8. Adjective phrases with the force of yes-no question:

   {#EmpVerb.PRED} Boring.PAT?

9. Adverbials:

   {#EmpVerb.PRED} *In Praha, at five o'clock.* (Fig. 4.22)

10. Some formulaic expressions:

    {#EmpVerb.PRED} *Your money*.PAT *or*.DISJ *your life*.PAT! (= Give me your money or I'll take your life!)

    {#EmpVerb.PRED} *Hands*.PAT *up*.DIR3! (= raise your hands up!)

    {#EmpVerb.PRED} Good evening/night/afternoon/morning.PAT ! (= have a good evening, I wish you a good evening) NB: *Goodbye, bye* are interjections.

    {#EmpVerb.PRED} Merry Christmas.PAT

    {#EmpVerb.PRED} All the best {#EmpNoun}.PAT

    Not at all: {#EmpVerb.PRED} #Neg.RHEM at all.EXT

    {#EmpVerb.PRED} No matter.PAT (= This is no matter)

    Certainly not: {#EmpVerb.PRED} *Certainly*.ATT #Neg.RHEM

    {#EmpVerb.PRED} *Glad*.PAT *to meet*.CAUS *you.*

Some types of noun phrases that are **not** regarded as grammatical ellipses are listed in Section 4.3.1.2, “Verbless clauses”.

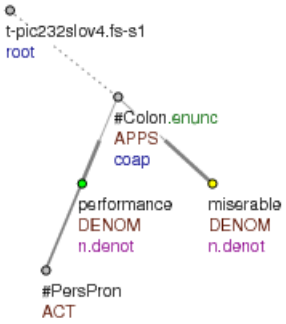
- **punctuation mark**. A punctuation mark is the governing node of a verbal clause in those cases in which it occupies the place of the missing verb and has its function. It is always possible to insert a simple verb (like be) into the clause and it is also possible to determine the function of the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

  Your performance: miserable. [#Colon.PRED] (Fig. 4.21)

- **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.

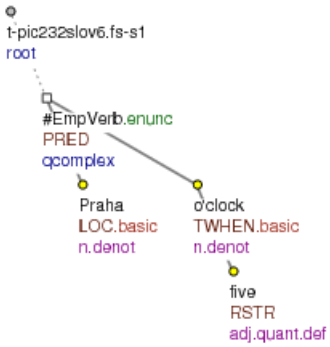
  But... {#Period3.PRED}

Figure 4.21. Verbal clauses with a punctuation mark in the role of the predicate



Your performance: miserable.

Figure 4.22. Verbal clauses with ellipsis of the predicate



In Praha, at five o'clock.

4.3.1.2. Verbless clauses

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

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

   If a subject-case clause is independent, its effective root has the `DENOM` functor. If an independent subject-case clause is a parenthesis (see Section 4.5, “Parenthesis”), its effective root has the `PAR` functor (see Section 6.1, “Functors for effective roots of independent clauses”). In those specific cases in which the subject-case clause is dependent (direct speech, *names introduced or postmodified by a descriptor that is not integral part of the name*: (see b)), its effective root has the functor according to its relation to the governing node.

   Several types of verbless clauses are listed or listed and narrower described below:

   i. Miscellaneous:

      An important event.`DENOM` (Fig. 4.23)

      1989.`DENOM`

      10 years.`DENOM`

      reuters.`DENOM`

   ii. Exclamatory clauses modified by a restrictive relative clause:

      The clothes.`DENOM` *she wears*.`RSTR`! (= Look at the terrible/beautiful clothes she wears!)

   iii. Scornful exclamatory clauses consisting of a noun phrase, generally a personal pronoun, followed by *and* and another noun phrase with a matching possessive pronoun:

You.ACT *and*.DENOM *your statistics*.ACT! (= It's funny how much you rely on and over-use statistics!)

NB! Object case is also used in these constructions with personal pronouns; it is even preferred: *Him and his malicious gossip!*

iv. Exclamatory noun and adjective phrases expressing approval or disapproval:

Charming couple.DENOM!

*Excellent meal.*DENOM !

v. Noun phrases functioning as assertions or conveying of information, which in verbal phrases would be inner participants :

False alarm.DENOM

No luck.DENOM

No news.DENOM

but not e.g. That way.DIR3, which in fact is an adverbial (resolved above as an empty-verb clause, in 9).

vi. Noun phrases acting as *yes-no* questions (except offers,invitations - see above, 3).

Any luck.DENOM, *Ron* **.** VOCAT ?

New hat.DENOM ?

Good flight.DENOM?

vii. Exclamatory noun phrases conveying a warning or e.g. alarm or frustration after a period of forgetfulness:

Fire.DENOM!

The police.DENOM!

The cake.DENOM!

My husband's birthday.DENOM!

NB! When there is any evidence that the noun phrase could mean an imperative, e.g. *Police!* not to be interpreted as *The police is coming!* but *Call the police!*, the construction is supposed to be an empty-verb clause (see 2).

b. ***vocative clauses***. The governing node of a vocative clause is a noun that clearly is meant to address someone (e.g. marked with an exclamation mark).

The effective root of a vocative clause has always the VOCAT functor (see Section 6.1, “Functors for effective roots of independent clauses”); the only exception is cases in which the vocative clause stands in the position of a name introduced or postmodified by a descriptor that is not integral part of the name*:* (see b).

Examples:

George.VOCAT ! Fig. 4.24

You.RSTR *idiot*.VOCAT !

They asked: Citizen.VOCAT*, do you want to breathe fresh air..?*

But: *the inscription Brothers and Sisters*.ID

c. ***interjectional clauses***. The governing node of an interjectional clause is an interjection, yes-no particle and *well* introducing a clause, separated by a comma .

The effective root of an interjectional clause has always the PARTL functor (see Section 6.1, “Functors for effective roots of independent clauses”); the only exception are cases in which the vocative clause stands in the position of a name introduced or postmodified by a descriptor that is not integral part of the name: (see b).

Examples:

Sorry.PARTLFig. 4.25

Pardon **.** PARTL

Goodbye/bye/cheers/cheerio.PARTL

Hi/Hello.PARTL

Yes.PARTL

Well.PARTL, ...

Ooh.PARTL

Pooh.PARTL

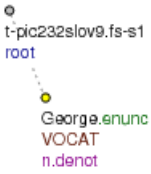
Oops.PARTL

But: *the inscription* *Yes*.ID

Figure 4.23. Nominative clauses

An important event.

Figure 4.24. Vocative clauses



George!

Figure 4.25. Interjectional clauses

Sorry.

4.3.2. Dependent and independent clauses (clause connecting)

Verbal and verbless 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 verbless.
- ***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 verbless.

Non-dependency. Several types of non-dependency connections of verbal and verbless 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 4.4, "Parataxis"):

verbal clause + verbal clause	*The cat is*.PRED [is_member = 1] *a mammal, but even the whale is*.PRED [is_member = 1] *a mammal too.*
nominative clause + nominative clause	*John Smith*.DENOM [is_member = 1] , *London*.DENOM [is_member = 1] Fig. 4.26
vocative clause + vocative clause	*Dear George*.VOCAT [is_member = 1] , *dear Peter*.VOCAT [is_member = 1]*!*
interjectional clause + interjectional clause	*Uh*.PARTL [is_member = 1] , *uh*.PARTL [is_member = 1]
verbal clause + nominative clause	*Book review*.DENOM [is_member = 1] *: The Hours is*.PRED [is_member = 1] a new book. (Fig. 4.27)
vocative clause + interjectional clause	*Oh*.PARTL [is_member = 1] , *George*.VOCAT [is_member = 1]*!*

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	*Oops*.PARTL, *I didn't mean*.PRED *that.* Fig. 4.28
verbal clause + vocative clause	*Sir*.VOCAT, *I'm not staying*.PRED *any longer.* Fig. 4.29
nominative clause + interjectional clause	*Oh*.PARTL, *that impudent youth*.DENOM
nominative clause + vocative clause	*He called: George*.VOCAT, *water*.PAT*!*

c. **parenthesis**. Another case of non-dependency relations is the case of syntactically non-incorporated parenthesis. For the annotation rules see Section 4.5, "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:

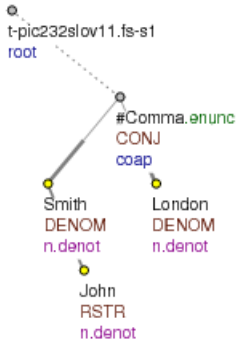
governing verbal clause + dependent verbal clause	*I don't know*.PRED, *why he left*.PAT

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 4.3.3, "Dependent verbal clauses (complex sentences)".

Specific cases of dependency relations:

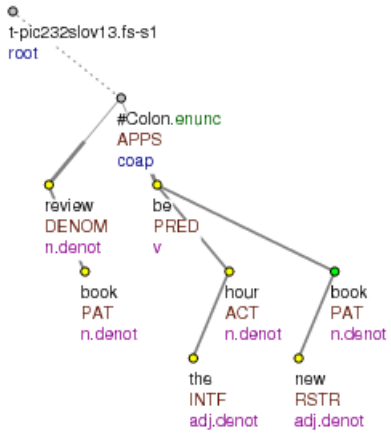
- **dependent direct speech**. All types of clauses can be used in the position of direct speech. For the annotation rules see Section 5.5, "Direct speech".
- **names introduced or postmodified by a descriptor that is not integral part of the name**: All types of clauses can occur in the position of a name introduced or postmodified by a descriptor that is not integral part of the name. The effective root of any clause in the position of a name introduced or postmodified by a descriptor that is not integral part of the name has the ID functor. For detailed rules see b.

Figure 4.26. Paratactic connection of two nominative clauses



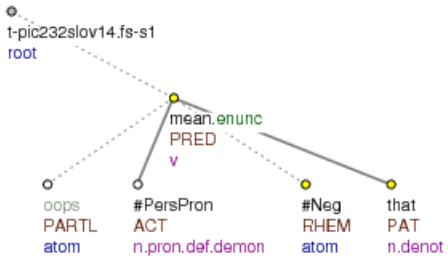
John Smith, London.

Figure 4.27. Paratactic connection of a nominative clause and a verbal clause



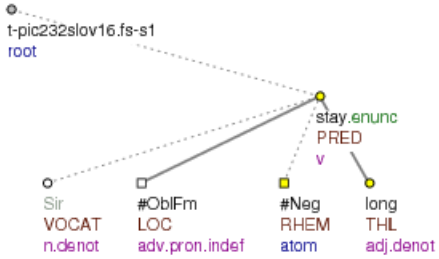
A book review: The Hours is a new book.

Figure 4.28. Connection of an interjectional clause and a verbal clause



Ooops, I didn't mean that.

Figure 4.29. Connection of a vocative clause and a verbal clause



Sir, I'm not staying any longer.

4.3.3. 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 4.3, "**Types of dependent verbal clauses**").

Table 4.3. Types of dependent verbal clauses

Dependent clause	Definition	Connective	Examples
Content clause	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.	*He said, <that> he would come*.EFF *He asked, who was coming*.PAT
Relative clause	It further specifies, modifes a noun phrase in the governing clause The effective root of the dependent clause has the RSTR functor.	relative expression connective *co* Relative elements introducing relative clauses corefer with the modified noun.	*A question that was not answered*.RSTR *needs to be answered.* *The district, which*.LOC *John lives in*.RSTR*, is a quiet and clean one.*
Adverbial clause	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	*<If> the weather is nice*.COND *we'll go out.* *He went where he was told to*.DIR3

4.3.3.1. Dependent verbal clauses without a finite verb form

Dependent verbal clauses with no finite verb form include:

dependent infinitival constructions,
dependent past participial constructions,
gerundial constructions.

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

a. **an argument**. It is the predicative-complement-like position (i.e. there is dual dependency involved, see Section 4.1.1, "Dual dependency").
b. **a predicative complement** (for the annotation rules see Section 4.1.1.1, "Predicative complement").
c. **an 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.

For more detail and more examples on annotation resolutions of non-finite verb forms see Section 8.1, "Infinitive clause" and Section 8.2, "Gerundial clause"

Examples:

The professor, inspired.COMPL *by the article, gave a lecture on the new issues.*Fig. 4.31

He remains inspired.PAT *by the article.*

The house, altough indebted.CNCS*, was sold very quickly.*

You have two possibilities, how to get.PAT *money* (Fig. 4.30)

He told her to leave.PAT

He left, being.COMPL*sorry.*

Incongruent infinitival constructions.

Incongruent gerundial/participial constructions. Incongruent gerundial/participial constructions are analyzed as conditional clauses and regarded as syntactically incorporated parentheses (see Section 4.5, “Parenthesis” and Section 8.2.8, “Gerund as free modification - COND”), which means that all members of the parenthesis have the attribute value [is_parenthesis = 1]. For example:

Judging.COND *by his face*.CRIT*, he was angry* [is_parenthesis = 1]Fig. 4.32

Judged.COND *by modern standards*.CRIT*, this was a cruel thing to do.* [is_parenthesis = 1]

He does very well, considering.COND how old he is. [is_parenthesis = 1]

Frankly speaking.COND*, she has chosen a wrong partner.* [is_parenthesis = 1] Fig. 4.33

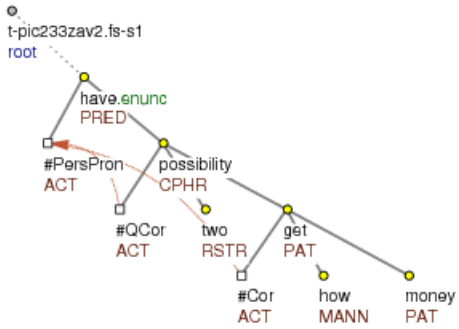
NB! Certain gerundial/participial constructions have frozen to such an extent that they are considered to be subordinators. For their complete list (taken from Quirk et al.) see Section 7.1.1, “One-word subordinators and the functors assigned to them”.

They invited all relatives <excepting> his brother.RESTR

They invited all relatives <including> his brother.ACMP

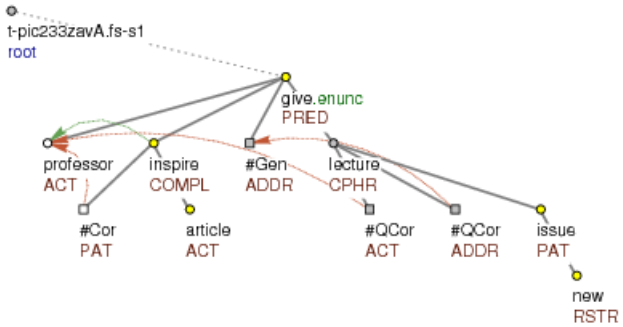
<Assuming/Given> your calculations are.COND correct, we should travel northeast.

Figure 4.30. Dependent infinitival constructions



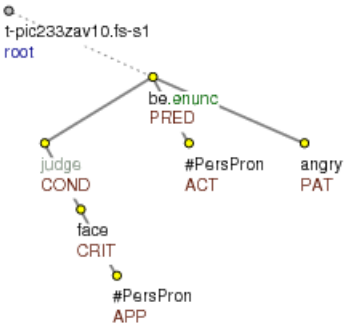
You have two possibilities how to get money.

Figure 4.31. Dependent participial constructions



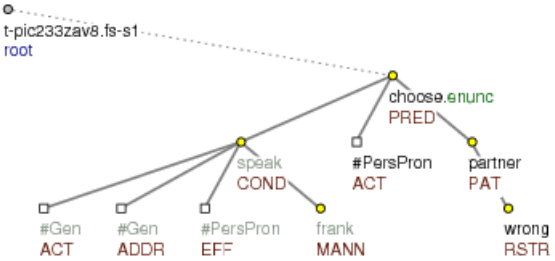
The professor, inspired by the new article, gave a lecture on the new issues.

Figure 4.32. Incongruent participial constructions



Judging by his face, he was angry.

Figure 4.33. Incongruent participial constructions

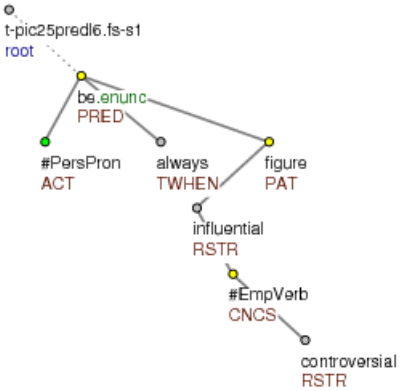


Frankly speaking, she has chosen a wrong partner.

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:

He has always been an influential, <though> {`#EmpVerb.CNCS`} *controversial*`.RSTR` *figure*`.PAT`Fig. 4.34

Figure 4.34. Adjective introduced by a subordinating conjunction



He has always been an influential, though controversial figure.

NB: the constructions with adjectives introduced by subordinating conjunctions are resolved in a very different way compared to constructions with gerunds/participles introduced by subordinating conjunctions, which take no empty-verb node and the appropriate functor is assigned directly to the gerund/participle (see above, this section).

4.3.3.2. Extraposition with *it*

See Section 8.3, “The personal pronoun *it* ”.

4.3.3.3. 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, infinitive or conjunctional 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, it is the form rather than the content that 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:

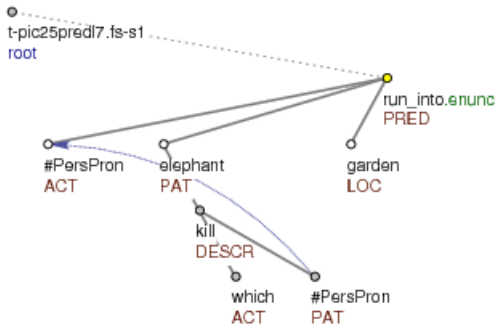
He run into an elephant in the garden, which killed.DESCR him.Fig. 4.35

NB: the decision whether a wh-clause introduced with which refers to a modification or to the entire predicate is in ambiguous cases up to the annotator. Here, the interpretation "an elefant that killed him" is preferred to "he run into an elephant, and this event killed him".

He left, never to come.AIM back again.Fig. 4.36

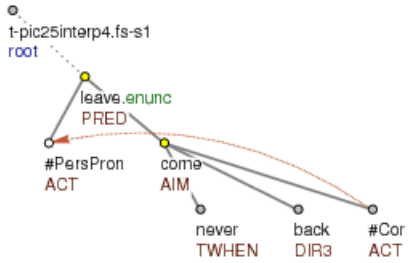
<If> I seem.COND angry sometimes, it's . PRED usually <because> I'm.CAUS very tired. (see Fig. 4.37 obrazek)

Figure 4.35. False dependent clause: which as DESCR



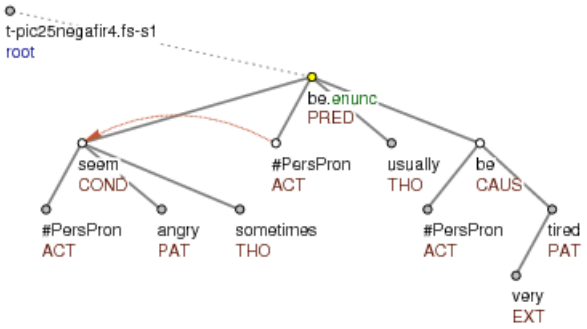
He run into an elephant in the garden, which killed him.

Figure 4.36. False dependent clause: a false purpose clause



He left, never to come back again.

Figure 4.37. False dependent clause



If I seem angry sometimes, it's usually because I'm very tired.

NB! If the connective could be considered a coordinating conjunction, the construction is analyzed as paratactic. The following constructions are discussed whether to be considered paratactic (coordination, apposition) or hypotactic:

a. **constructions with *which* coreferring with the entire clause.** When a clause 2 is attached to a clause 1 by means of the anaphorical construction with the connective *which* (neither introduced by a preposition nor combined with *be* + *wh-word*) co-referring with the predicate of the clause 1, the two clauses are regarded as paratactically connected. The connective *which* (even when combined with prepositions) 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:

   She had not arrived.PRED[is_member=1] #Comma.CONJ which.ACT made.PRED [is_member=1] him sad.Fig. 4.38

   She had missed.PRED [is_member=1] the plane #Comma.CONJ which.PAT he didn't know.PRED [is_member=1] .

   We have not answered.PRED[is_member=1] *your question completely, <for> which*.CAUS *we apologize.* PRED[is_member=1] (Fig. 4.39)

   They had not answered.PRED [is_member=1] #Comma.CONJ <for> which.CAUS they were punished.PRED [is_member=1]

b. **Constructions with *which +be* only coreferring with a valency modification**. Constructions with *which* + *copula be* are resolved as apposition (see Section 4.4.1, “Coordination, apposition, mathematical operations and intervals”):

   He got five points.PAT[is_member=1] #Comma.APPS *which*.ACT *was*.PAT[is_member=1] *the maximum*.PATFig. 4.40

c. **constructions with relative clauses comprising (preposition +) *which* + noun**. These constructions introduce clauses which get the functors RSTR/DESCR as common subordinate clauses.

   Eleanor's pen, which.ACT *had been lying*.DESCR *on the table, suddenly moved.*Fig. 4.41

   Guidance.DENOM on the Circumstances.PAT <in> which.COND Parents may Choose.RSTR to Educate their Children at Home. (Fig. 4.42)

d. **constructions with *which* + *be* + *why* correlating with the predicate**. The construction *which is/was why* correlating with another clause's predicate is resolved as coordination. The root of the paratactic structure is the node for the present punctuation mark and has the functor CSQ. The construction *which is/was why* is represented as the node for *why* with the auxrfs *which* and *is/was*. The node has the functor CM (conjunction modifier).

   She.ACT 's.PRED[is_member=1]poor.PAT #Comma.CSQ <which is> why.CM she.ACT has.PRED[is_member=1] to come to Arthur. (Fig. 4.43)

This construction has the same tectogrammatical representation as that's why (see Section 8.7.3, “Anaphorical *that* in sentences of the type *that's +wh-word*”).

e. **constructions with non-restrictive postmodifying clauses *which + be + wh-words* except *why* correlating with the predicate.** The construction *which + be + wh-words*. These constructions are resolved as paratactic conjunctions like w*hich +be + why*.

In the process, the uranium (loses, or) is depleted(,) of almost half its radioactivity, which is how depleted uranium gets its name.Fig. 4.44

NB: Especially in expert texts the non-restrictive postmodifying clause need not modify the effective root of the sentence but can "dive" much deeper - and it can even modify a gerundial predication:

*Another thing the microscope revealed was the difference between forming a laminate using vacuum pressure, which is how 3DL is made, and forming a laminate using pressure from rollers, which is how (a few of) the typical (‘look-alike’) sails are made.*Fig. 4.45

f. **constructions with non-restrictive postmodifying clauses *which + be + wh-words* except *why* not correlating with the entire predicate but only with a valency modification.** Such constructions are - according to the way they corefer - either resolved as non-restrictive attribute:

This is where you will find Ogden Utah, <which is> where.`DESCR` *the Central Pacific and Union Pacific railroads met in 1869.* (Fig. 4.46)

or as apposition (when the non-restrictive postmodifying clause actually paraphrases the valency modification in the preceding clause).

NB: Two types of apposition can occur:

i. the valency modification is modified by two restrictive relative clauses (`RSTR`). The **second** restrictive relative clause **paraphrases** the **first** restrictive relative clause:

We have reached the point where we want.`RSTR[is_member=1]` *to sum up.*`PAT` a*n infinite number of differential amounts* `#Comma.APPS` *<which is> when.*`TWHEN` we *integrate.*`RSTR[is_member=1]` Fig. 4.47

Only the wh-word keeps its tectogrammatical representation in this case, *which* and *be* are attached as its auxrfs.

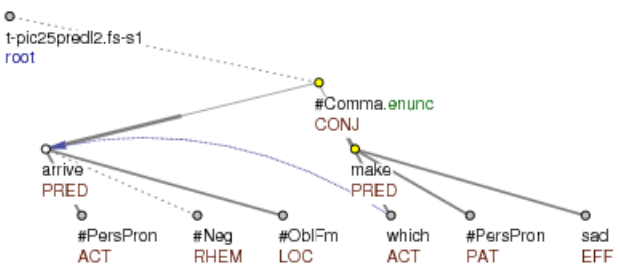
ii. only a **noun phrase is paraphrased** with the non-restrictive postmodifying clause:

To be successful, the formula is this reach combined.`RSTR` *with frequency.*`EFF[is_member=1]` `#Comma.APPS` *which.*`ACT` *is.*`EFF[is_member=1]` *how often viewers will receive.*`PAT` (Fig. 4.48)

N is for Never.`SUBS[is_member=1]` `#Comma.APPS`, *which.*`ACT` *is.*`SUBS[is_member=1]` *how often he's.*`PAT` *sincere when he says he cares.*Fig. 4.49

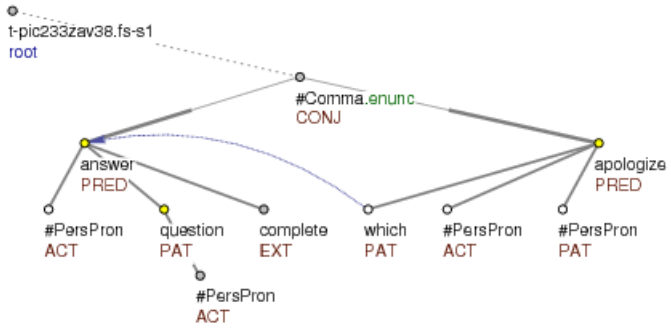
NB: This case is different in that *which* and *how* (or any other wh-word if ever found with this function) **do have** their **own tectogrammatical nodes**!

Figure 4.38. Constructions with which referring to the predicate



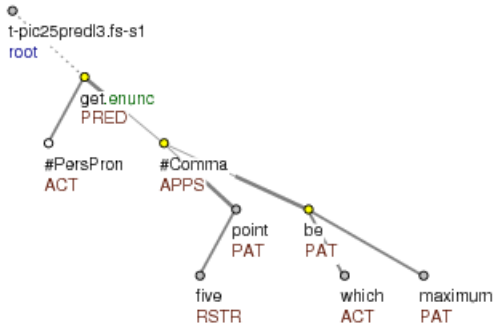
She had not arrived, which made him sad.

Figure 4.39. Constructions with prep+which



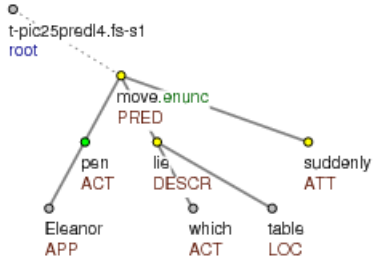
We have not answered your question completely, for which we apologize.

Figure 4.40. Which as apposition



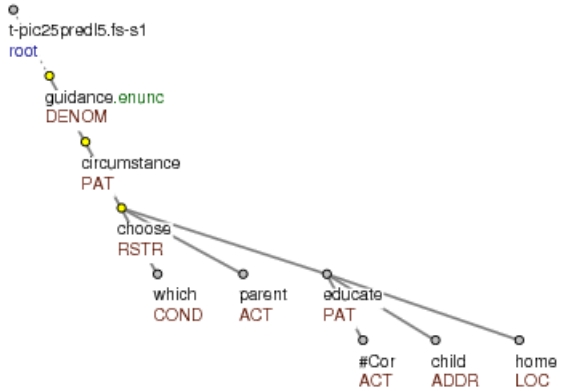
He's got five points, which was the maximum.

Figure 4.41. Which as non-restrictive attribute



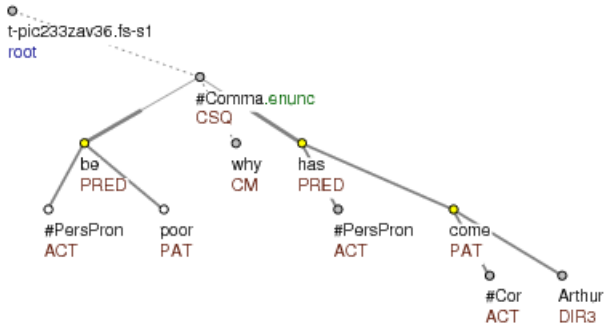
Eleanor's pen, which had been lying on the table, suddenly moved.

Figure 4.42. Which as restrictive attribute



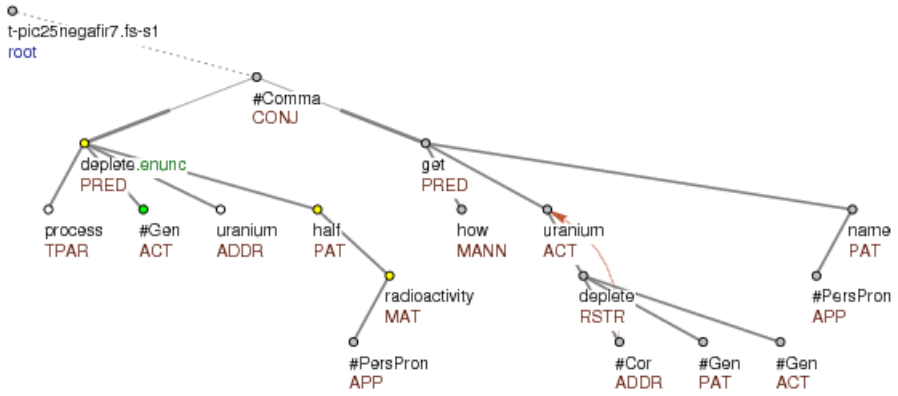
Guidance on the Circumstances, in which the Parents may Choose to Educate their Children at Home.

Figure 4.43. Constructions with which + be+ why



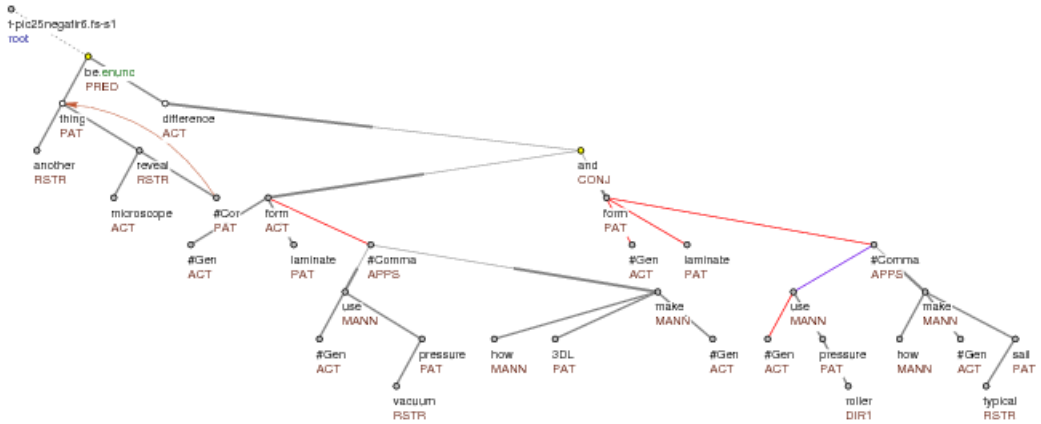
She's poor, that's why she had to come to Arthur.

Figure 4.44. Constructions with which + be + wh word except why



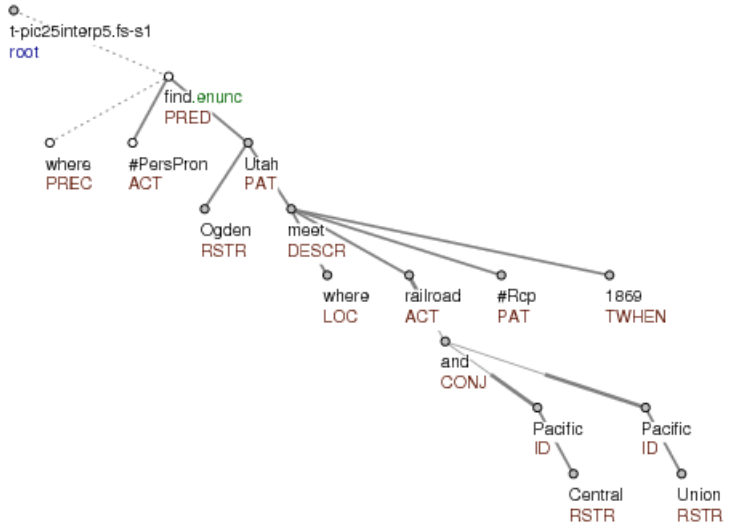
In the process, the uranium (loses, or) is depleted(,) of almost half its radioactivity, which is how depleted uranium gets its name.

Figure 4.45. Constructions with which + be + wh word except why



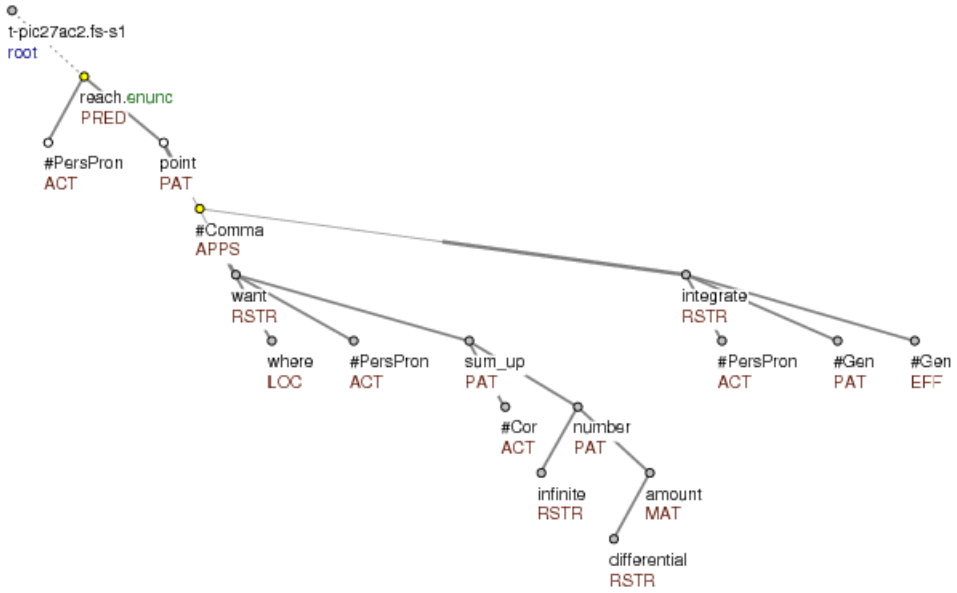
Another thing the microscope revealed was the difference between forming a laminate using vacuum pressure, which is how 3DL is made, and forming a laminate using pressure from rollers, which is how (a few of) the typical ('look-alike') sails are made.

Figure 4.46. Constructions with which + be + wh word except why



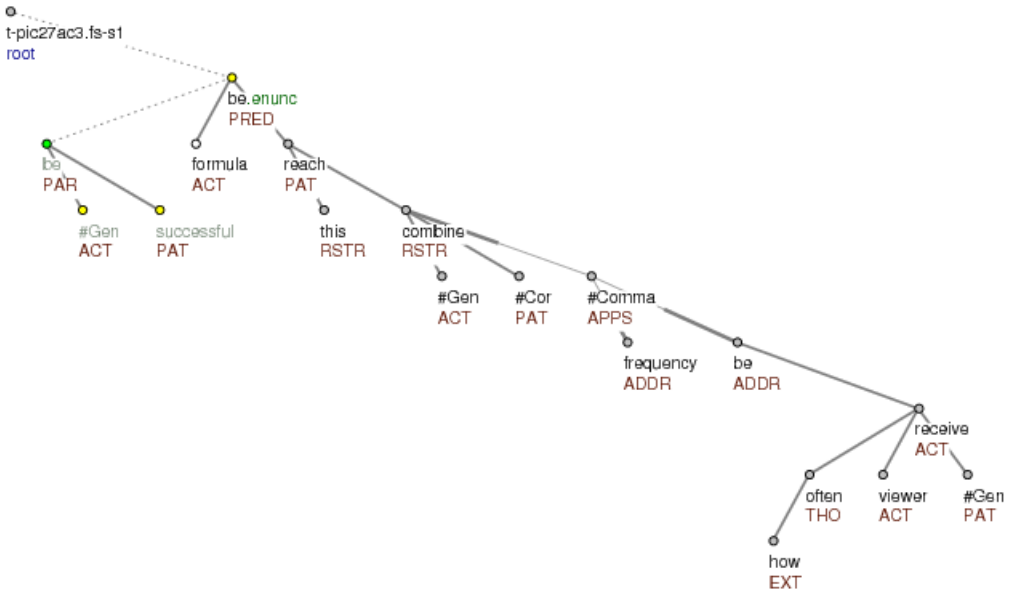
This is where you find Ogden Utah, which is where the Central Pacific and Union Pacific railroads met in 1869.

Figure 4.47. Constructions with which + be + wh word except why



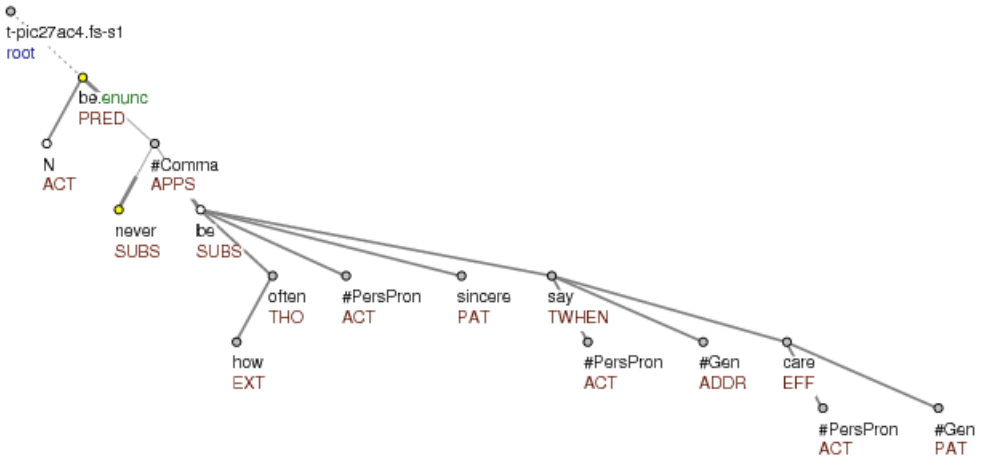
We have reached the point where we want to sum up an infinite number of differential amounts, which is when we integrate.

Figure 4.48. Constructions with which + be + wh word except why



To be successful, the formula is this reach combined with frequency, which is how often viewers will receive.

Figure 4.49. Constructions with which + be + wh word except why



N is for never, which is how often he's sincere when he says he cares.

4.4. Parataxis

Parataxis is a non-dependency connection of two or more elements (modifications or clauses) that are on the same level and 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 4.1.2, “Non-dependency edges”).

Representing parataxis in the tectogrammatical trees. Paratactic connections are represented by a ***paratactic structure***(see Fig. 4.50 and Fig. 4.51). *A* ***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`.

He was troubled by insects.`ACT[is_member=1]` `#Separ.CONJ` etc.`ACT[is_member=1]` Fig. 4.52

Paratactic structure root nodes are assigned the value `coap` (see Chapter 2, ***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 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 4.50. Paratactic structure I

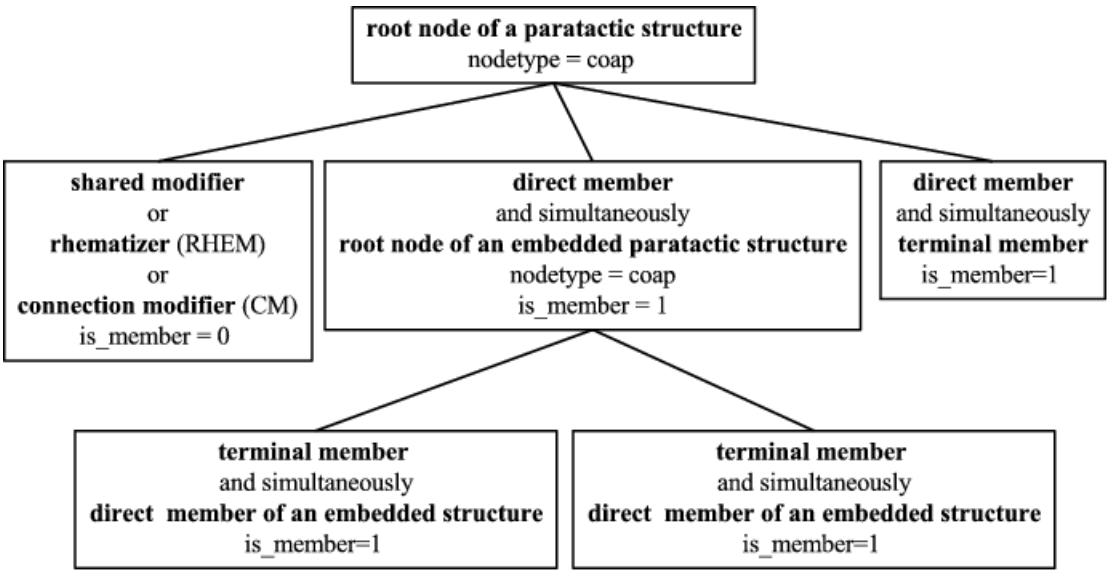
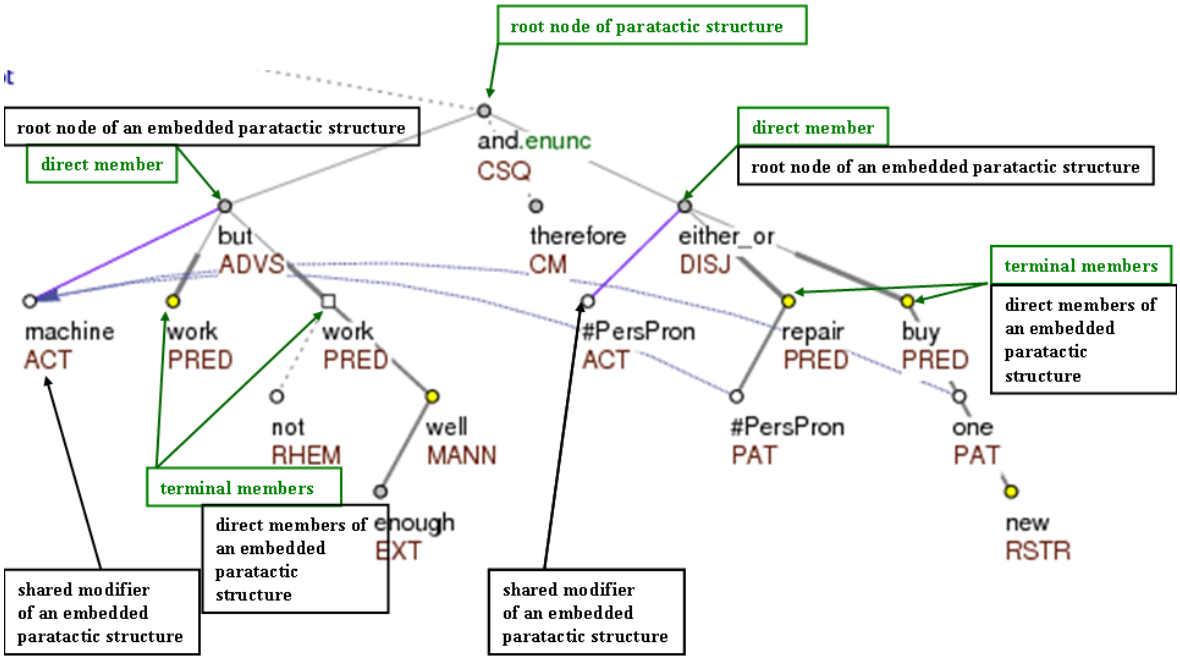


Figure 4.51. Paratactic structure II



The machine does work,but not well enough, and therefore we either must repair it or buy a new one.

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:

*I saw and heard Mary sing.*Fig. 4.53

Peter had been working on his dissertation and preparing for an English exam the whole day but in the evening he was doing nothing more. (Fig. 4.54)

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:

*John was stressed and difficult to get along with.*Fig. 4.55

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 to clausal parataxis.

Nevertheless, it is not always possible to represent the construction as parataxis of sentence parts. All cases which do not fullfil 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 4.6.1.1, “Ellipsis of a governing meaning unit”). For example:

Peter arrived and Paul probably as well.

= *Peter came and Paul probably came as well.* The Actors *Peter* and *Paul* cannot be captured as being in constituent coordination; the expression *probably* modifies the absent predicate. The construction is therefore represented as clausal coordination. Fig. 4.56.

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 coordination or apposition of non-obligatory adjuncts. For example:

   eight-hour.RSTR [is_member = 1] *working*.RSTR[is_member=0] *time a*nd.CONJ *without break*.ACMP [is_member = 1]

   They did it with pleasure.ACMP [is_member = 1] , *i.e. well*.MANN [is_member = 1] [#Comma.APPS]

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.:

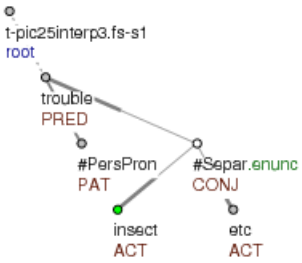
   Little is known about other interesting places.PAT [is_member = 1], *such as.* APPS *#EmpVerb*.PAT [is_member = 1] *Hogwarts.*

   If the paratactic connection is a connection of a verbal clause and a verbless clause, the functor assignment follows the rules in Section 4.3.1, “Verbal and verbless clauses”. For example:

   The topic.DENOM [is_member = 1][#Colon.APPS] *What I am doing*.PRED [is_member = 1] at the moment

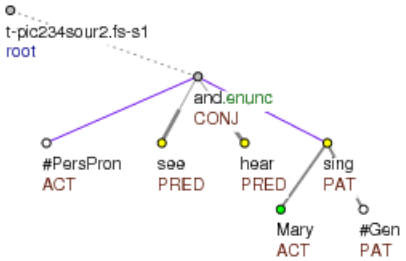
Semantic types of paratactic connections. The functor assigned to the root of a paratactic structure expresses the semantic relation between the connected elements. All functors (and their definitions) for paratactic structure root nodes are in Section 6.12, “Functors expressing the relations between members of paratactic structures”.

Figure 4.52. Paratactic structure without punctuation



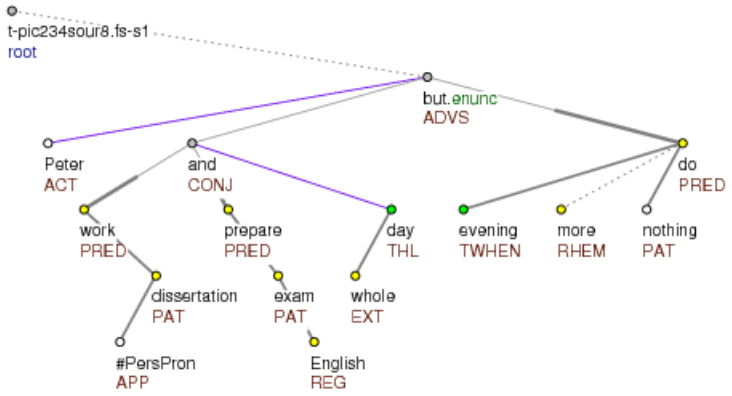
He was troubled by insects etc.

Figure 4.53. Paratactic structure



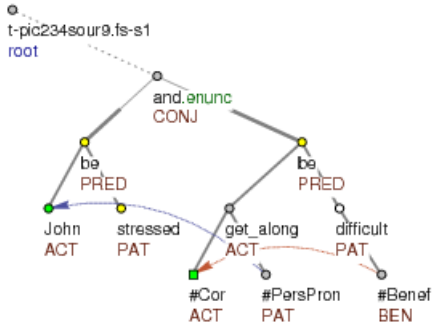
I saw and heard Mary sing.

Figure 4.54. Paratactic structure



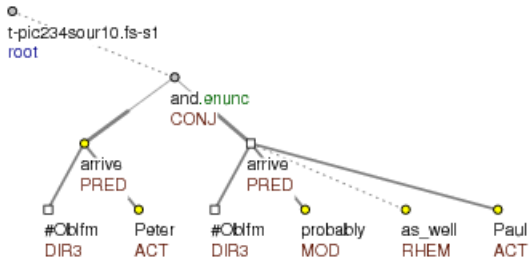
Peter had been working on his dissertation and preparing for an English exam the whole day but in the evening he was doing nothing more.

Figure 4.55. Paratactic structure



John was stressed and difficult to get along with.

Figure 4.56. Paratactic structure



Peter arrived and Paul probably as well.

4.4.1. 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.: Karel, the Czech Emperor; Prague, the capital). If there is no comma (e.g.: *the* Czech Emperor Karel; the capital city *Praha*; it happened in Prague on the Old Town Square; *in February 1999; 10:50 a.m. EST*) the connection is not analyzed as apposition (see especially Section 4.1.3, “Ambiguous dependency”).

Also some specific constructions are represented as coordination or apposition: constructions with the abbreviations *etc.*, *e.g., i.e.*; constructions in which a modification follows expressions like *such as.* Also constructions *which* (see also Section 4.3.3.3, "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:

the match ended in a 5:0 draw. (Fig. 4.57)

10 minus 2 is 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:

It kept snowing from Christmas.`TSIN` *to Easter*.`TTILL`

It happened between Monday.`TWHEN` [`subfunctor=betw`] *and Wednesday*.`TWHEN` [`subfunctor = betw`]

I know it here from Aš.`DIR1` *over Prague*.`DIR2` *to Brno*.`DIR3`

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:

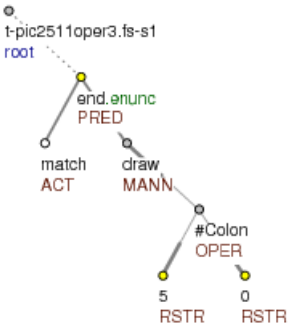
In the years 1995 - 1999 I was attending the high school.

Everybody was watching, from kids to adults. Fig. 4.59

a three to ten-year follow-up study [1]

An accident occurred on the route London-Sydney.

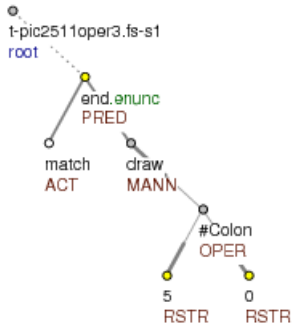
Figure 4.57. Mathematical operations



The match ended in a 5:0 draw.

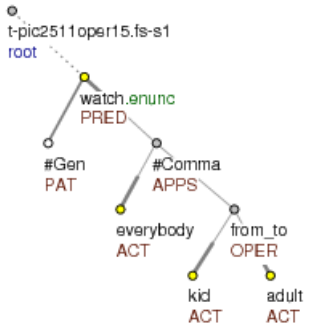
[1] NB: this would currently pose problems as nodes with hyphens are not yet analyzed in more detail.

Figure 4.58. Interval as a paratactic structure



In the ears 1995-1999 I was attending the high school.

Figure 4.59. Interval as a paratactic structure



Everybody was watching, from kids to adults.

4.5. 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`. 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.:

*Peter (and Paul) came there.*Fig. 4.60

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:

   I can see our house (and our garden.PAT*)*

   You can (if you dare.COND*) follow John.* (Fig. 4.61)

b. **syntactically non-incorporated parenthesis**. Syntactically non-incorporated parenthesis corresponds - both semantically and in its form - to some kind of independent clause (see Section 4.3, "Clauses (governing, dependent, verbal, verbless)"). The effective root of a parenthesis is assigned a functor on the basis of what kind of independent clause it is, see Table 4.4, "**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:

The team ended up third (last year they lied.PAR first). (Fig. 4.62)

John Smith (London.PAR).

He has again failed to get a job (damn.PARTL).

Table 4.4. 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 hat it has become a particle. The lexicalized parenthesis is formed by a frozen verb form, which can keep part of the original valency, or by a one-word noun phrase that does not enter any dependency relations in the sentence (*the hell/the fuck* etc. emphasizing a wh-word in a wh-clause).

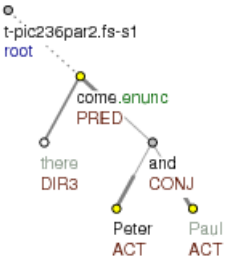
The effective root of a lexicalized parenthesis is assigned the ATT functor (nodetype = atom). A multi-word lexicalized parenthesis is usually considered a verbless idiom (see Section 5.1.3, “Idioms”). The effective root of a lexicalized parenthesis is an immediate daughter of the effective root of the clause in which the parenthesis is inserted. For example:

To think.ATT[is_parenthesis=1] *we might have been in Rome now!* (Fig. 4.63)

*This Conference should, by God's.*DPHR[is_parenthesis=1]*help*.ATT[is_parenthesis=1]*, be of fortunate omen for the opening century.*

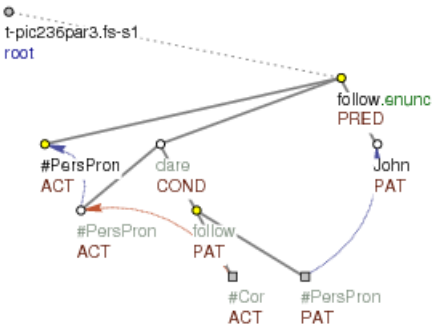
*What the hell.*ATT[is_parenthesis=1] *is XML?*

Figure 4.60. Discontinuous parenthesis



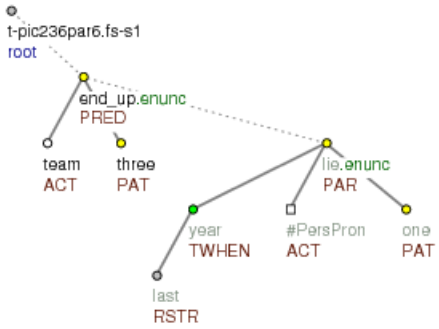
Peter (and Paul) came there.

Figure 4.61. Syntactically incorporated parenthesis proper



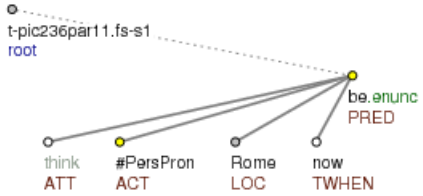
You can (if you dare) follow John.

Figure 4.62. Syntactically non-incorporated parenthesis proper



The team ended up third (last year they lied first).

Figure 4.63. Lexicalized parenthesis



To think we might have been in Rome now!

Speaker's comments such as "just to make things clear" Clauses introduced by the conjunction *aby* in constructions such as *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:

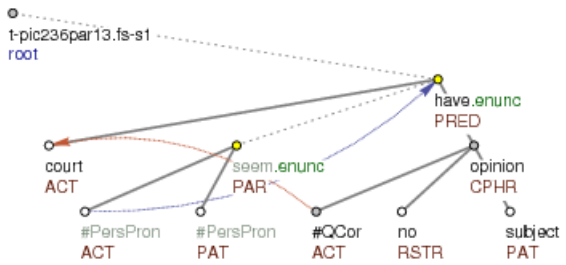
Just to make.PAR*things clear: I'm the boss here.*

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 (*The court, it seems, has no opinion on the subject*), or a clause introduced by the connective *as* (*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 *as*, if present, is assigned no node in the tree. For example:

The court, <as> it seems.PAR to me*, has no opinion on the subject. / The court, it seems*.PAR*to me, has no opinion on the subject.*Fig. 4.64

Figure 4.64. Clauses with a reversed syntactic relation



The court, as it seems to me, has no opinion on the subject.

4.6. 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`

Types of newly established nodes:

a. nodes for omitted meaning units (see Section 4.6.1, “Ellipsis”),
b. auxiliary nodes for representing more complex syntactic structures (see Section 4.6.2, “Newly established nodes in more complex syntactic structures”),
c. nodes representing negation with verbs (see Section 4.6.3, “Nodes representing negation with verbs”).

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

A. ***copied nodes***, i.e. newly established nodes that have the values of certain atributes 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 4.6.1.1, “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:

`#EmpNoun` `#EmpVerb`	grammatical ellipsis of governing meaning units (see Section 4.6.1.1, “Ellipsis of a governing meaning unit”)

`#AsMuch`
`#Equal`
`#Total`
	ellipsis of governing meaning units in specific constructions
	(see Section 4.6.1.3, “Specific elliptical constructions”)

`#Benef`
`#Cor`
`#Gen`
`#Oblfm`
`#PersPron`
`#QCor`
`#Rcp`
`#Some`
`#Unsp`
	ellipsis of dependent meaning units
	(see Section 4.6.1.2, “Ellipsis of a dependent meaning unit”)

`#Forn`
`#Idph`
`#Separ`
	auxiliary nodes for representing more complex syntactic structures
	(see Section 4.6.2, “Newly established nodes in more complex syntactic structures”)

`#Neg`	verbal negation by means of the morpheme *ne-* (see Section 4.6.3, “Nodes representing negation with verbs”)

4.6.1. Ellipsis

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

ellipsis of a governing meaning unit (see Section 4.6.1.1, “Ellipsis of a governing meaning unit”),

ellipsis of a dependent meaning unit (see Section 4.6.1.2, “Ellipsis of a dependent meaning unit”).

4.6.1.1. 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:

a. **ellipsis of the governing predicate**, i.e. omission of the governing predicate in verbal clauses.
b. **ellipsis of the governing noun**, i.e. omission of the governing noun with adjectival modifications.
c. **ellipsis of the governing clause**, i.e. omission of the governing clause in cases where there is a dependent clause.
d. **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:

a. ***contextual ellipsis***, i.e. 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 4.5, “**Representing ellipsis of a governing meaning unit**”.

Table 4.5. Representing ellipsis of a governing meaning unit

Types of ellipsis		**Contextual ellipsis**	**Grammatical ellipsis**
Ellipsis of the predicate		copied node	`#EmpVerb`
Ellipsis of the governing noun		copied node	`#EmpNoun`
Ellipsis of the governing clause	for a dependent content or adverbial clause	copied node	`#EmpVerb`
	for a dependent relative clause	copied node	`#EmpNoun`
Specific types of ellipsis	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:

Mary prepared the lunch and John {prepared} the dinner. (Fig. 4.65)

(You will go.) Why {`#EmpVerb.PRED`} *me?*Fig. 4.66

{`#EmpVerb.PRED`} *Of course.*

Though {`#EmpVerb.PRED`} *tired, he kept on working.*

Examples of ellipsis of the governing noun (see also Section 8.6, “*One* as pro-form ”):

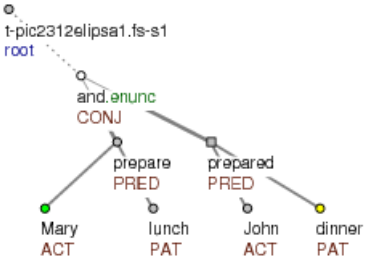
The blue slippers are torn, the green {slippers} {are} dirty. Fig. 4.67

Only the younger {`#EmpNoun.ACT`} *came.* (Fig. 4.68)

Examples of ellipsis of the governing clause:

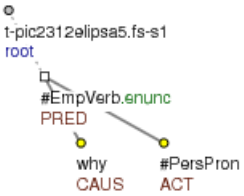
(She didn't leave.) {*leave*.PRED} *Because she wouldn't be there in time.*Fig. 4.69

Figure 4.65. Contextual ellipsis of the governing predicate



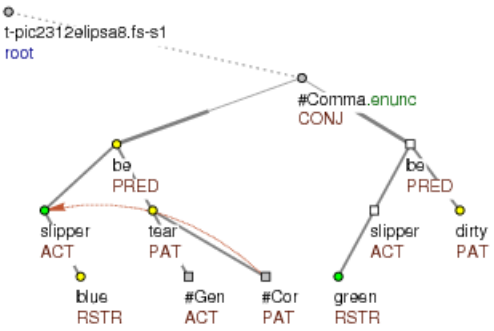
Mary prepared the lunch and John the dinner.

Figure 4.66. Grammatical ellipsis of the governing predicate



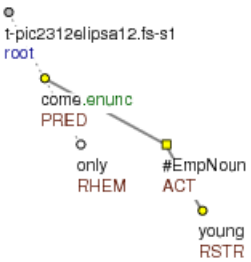
Why me?

Figure 4.67. Contextual ellipsis of the governing noun



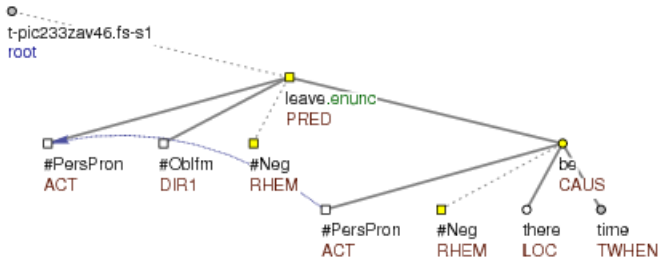
The blue slippers are torn, the green dirty.

Figure 4.68. Grammatical ellipsis of the governing noun



Only the younger came.

Figure 4.69. Ellipsis of the clause governing an adverbial clause



(She did not leave.) Because she wouldn't be there in time.

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

(George is all ears.) John too. = (George is all ears.) John is all ears, too.

(George has no chance to pass the exam.) John neither. = (George has no chance to pass the exam.) John has neither any chance to pass the exam. The negation *no* is not to be copied as the negation is already contained in *neither*.

NB: *Too* epressed as *So does X* is resolved in a different way. See below and Section 8.17.1, “The pro-form *to do so*”.

Contextual ellipsis of a nominal part of a verbonominal predicate. The omitted component is copied to its place:

I am happy if you are. = I am happy if you are happy.

You must be a member of the party since he is. = You must be a member of the party since he is a member of the party.

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 grammatemes. The inner participants of the omitted verb are to be filled in as well. Those realized on the surface get the t-lemma `#PersPron`. Free modifications are not to be copied. For example:

Will you be learning? I will. = (Will you be learning?) I will be learning.

Do you know him? "I <don't> {know} {`#PersPron`}.

NB: This does not apply to positive forms of do as pro-form (dummy-do) and for the pro-form *to do so* (different from the use of *to do so in the sense "too, as well")*. For more detail see Section 8.17, “Dummy-do (Verbal Pro-form)” and Section 8.17.1, “The pro-form *to do so*”.

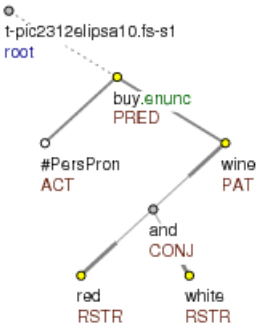
(Do you have to go?) I do. = (Do you have to go?) I `#VerbPron.PRED`

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. Nevertheless, the parataxis is anyway placed to the lowest possible place. For example:

*He bought red and white wine. = He bought red wine and white wine.*Fig. 4.70

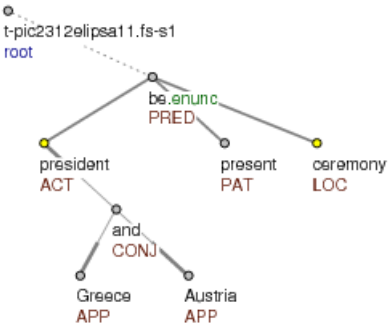
The presidents of Greece and of Austria were present at the ceremony. (Fig. 4.71)

Figure 4.70. Contextual ellipsis of the governing noun in constituent coordination



He bought red and white wine.

Figure 4.71. Contextual ellipsis of the governing noun in constituent coordination



The presidents of Greece and of Austria were present at the ceremony.

4.6.1.2. 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 4.6, "**Representing ellipsis of a dependent meaning unit**".

Table 4.6. Representing ellipsis of a dependent meaning unit

Types of ellipsis		**T-lemma of the newly established node**
Ellipsis of an obligatory modification	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	`#Oblfm`
Ellipsis of a non-obligatory modification	Non-expressed Beneficiary in control constructions	`#Benef`
Specific type	nominal 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 quasi-control see Section 5.1.2.5.1, “Quasi-control with support verb constructions” For more on comparative constructions see Section 4.6.1.3.1, “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. Examples:

(Has the shop assistant wrapped the book?) He <has> {wrap} {#PersPron.PAT}. Fig. 4.72

(The company was supposed to deliver the goods to the customer.) Yet the delivery {#PersPron.ACT} {#PersPron.PAT} {#PersPron.ADDR} *did not take place.* (Fig. 4.73)

{#PersPron.ACT} *Go! (when the speaker addresses someone the addressed person is regarded as known)*

Please, {#PersPron.ACT} *let us* .PAT *go to the movies without you.*

but the imperative regarding the first person plural (*let us go*) is treated in a different way: *let* is attached to the lexical verb as auxrf and *us* becomes the #PersPron with the appropriate functor (ACT in active, PAT or possibly ADDR in passive infinitives (see Section 4.2.4.1, “Surface form of a valency modification (Surface frame slot filler)”, especially a)). The semantic interpretation of the imperative with *let us* is up to the annotator.

Ellipsis of an obligatory adjunct: Example:

The boss has left {#Oblfm.DIR1} Fig. 4.74

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:

Houses are built from bricks. {#Gen.ACT} Fig. 4.75

This oven is good for baking. {#Gen.ACT} {#Gen.PAT}

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 explicitely 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:

The dog bites. {#Unsp.PAT} *(=People who penetrate his territory/who tease him)*

It.PAT *was* {#Unsp.ACT} *announced on the radio.*Fig. 4.76

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:

John and Mary met. = John met Mary and (simultaneously) Mary met John.)

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.

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. The reciprocal pronouns *each other* and *one another*, possibly present at the surface level, are not represented by separate nodes 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:

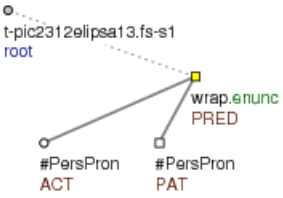
John.ACT*and Mary*.ACT*kissed.* {#Rcp.PAT} Fig. 4.77

John.ACT*and Mary*.ACT*kissed* {#Rcp.PAT} *<each other>.*

The trading <among> the EU states.ACT {Gen.PAT} {#Rcp.ADDR} *(= The EU states trade with each other.)*

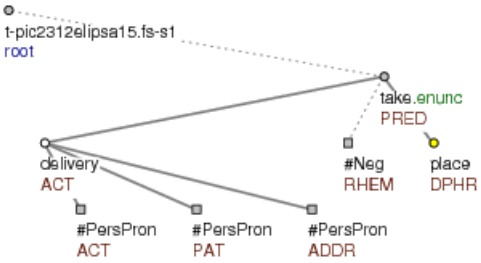
They compared Germany.PAT *and Korea*.PAT {#Rcp.EFF}

Figure 4.72. Contextual ellipsis of an obligatory argument



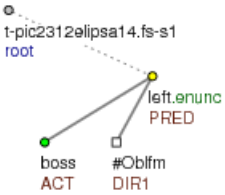
(Has the shop assistant wrapped the book?) He has.

Figure 4.73. Contextual ellipsis of an obligatory argument



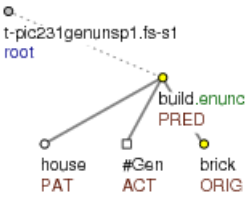
(The company was supposed to deliver the goods to the customer.) The delivery did not take place.

Figure 4.74. Ellipsis of an obligatory adjunct



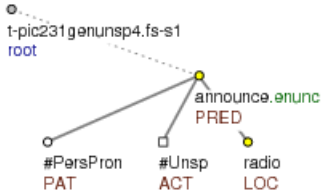
The boss has left.

Figure 4.75. General argument



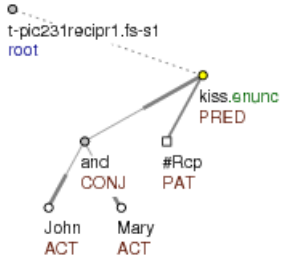
Houses are built from bricks.

Figure 4.76. Unspecified Actor



It was announced on the radio.

Figure 4.77. Reciprocity



John and Mary kissed (each other).

4.6.1.3. Specific elliptical constructions

Special cases of ellipsis of a governing meaning unit are:

comparative constructions (comparison of two events, see Section 4.6.1.3.1, “Comparative constructions: comparison of two events”),
constructions with the meaning of a restriction (see Section 4.6.1.3.2, “Constructions with the meaning of a restriction”),
constructions with consecutive clauses (see Section 4.6.1.3.3, “Constructions with consecutive clauses”).

4.6.1.3.1. Comparative constructions: comparison of two events

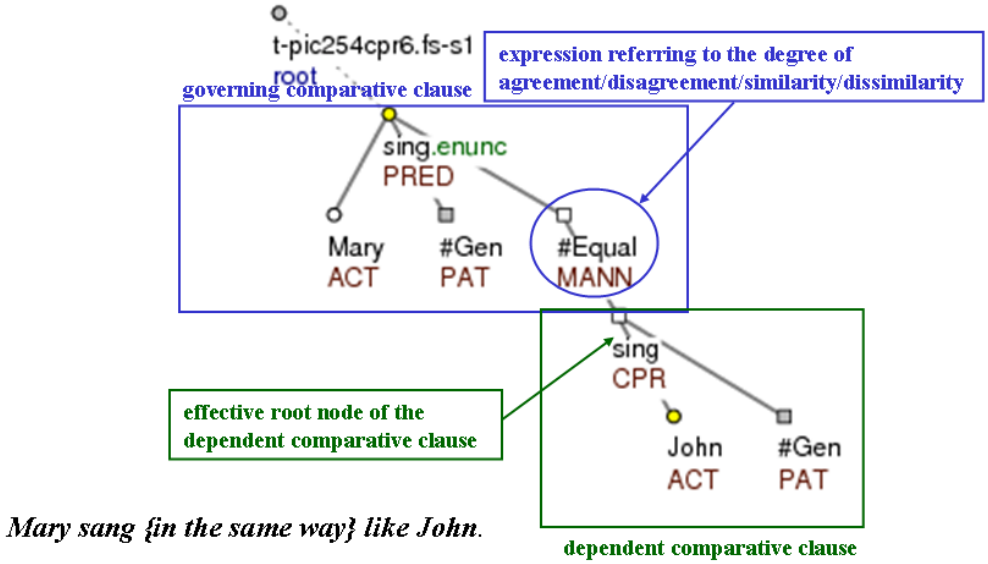
Constructions with the meaning of comparison of two events are constructions in which two events or states are compared. There is always a property or degree of agreement/disagreement/similarity with respect to which the events are compared.

Two basic types are distinguished, depending on the form of comparison:

a. comparison based on similarity or identity, by means of the conjunctions *as-as*, *as* and *like* (*John runs as fast as Mary does.*),
b. comparison on the basis of dissimilarity, by means of the conjunction *than (He arrived earlier than Jirka* did.).

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. 4.78.

Figure 4.78. 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 *like*), a new node is added into the structure, with the t-lemma substitute `#Equal`. For example:

Mary sang {`#Equal.MANN`} *like John. = Mary sang in the same way as John sang.*

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 4.6.1.1, "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 4.6.1.1, "Ellipsis of a governing meaning unit"). For example:

Mary sang like John. = Mary sang in the same way as John sang `.CPR`Fig. 4.79

When the dependent clause contains dummy *do* (see Section 8.17, "Dummy-do (Verbal Pro-form)"), *a new node with* the t-lemma substitute `#VerbPron` is generated onto the place of *do*. *Do* is attached to it as its auxrf. The node is regarded as generated (`[is_generated=1]`):

Mary sang like John <*did*>{`#VerbPron}.CPR`. (Fig. 4.80)

Mary arrived earlier than John. = Mary arrived earlier than John {arrived}`.CPR`

Mary had arrived earlier than John did. = Mary had arrived earlier than John <did>{`#VerbPron}.CPR`

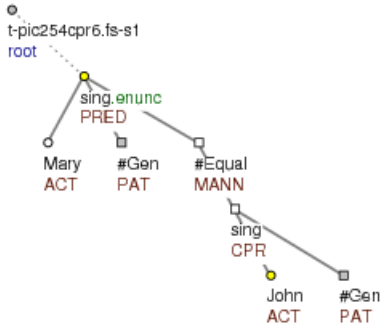
He is fit as a fiddle = He is as fit as a fiddle is.`CPR` fit.Fig. 4.81

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 nominal 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. *the same, similar*, *different*, *more*, *less*). For example:

Mary is the same as John. = Mary is the same as John is `#Some.PAT` (Fig. 4.82)

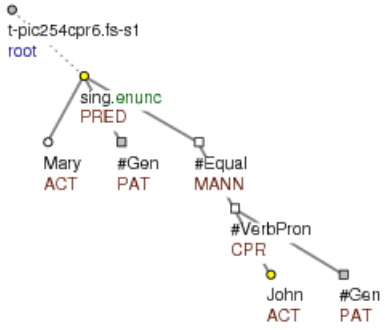
The situation in the army is different than the one at the Ministry. = The situation in the army is different than the situation at the Ministry {is} `{#Some.PAT}`

Figure 4.79. Comparative constructions: comparison of two events



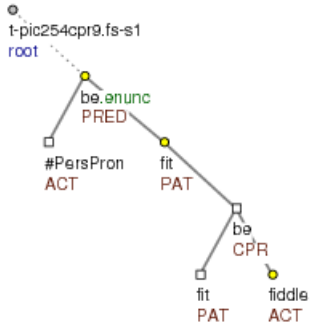
Mary sang like John.

Figure 4.80. Comparative constructions: comparison of two events



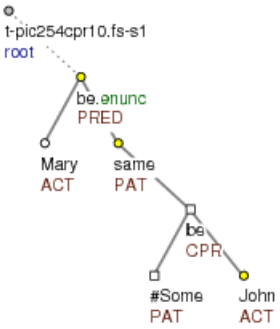
Mary sang like John did.

Figure 4.81. Comparative constructions: comparison of two events



He is fit as a fiddle.

Figure 4.82. Comparative constructions: comparison of two events



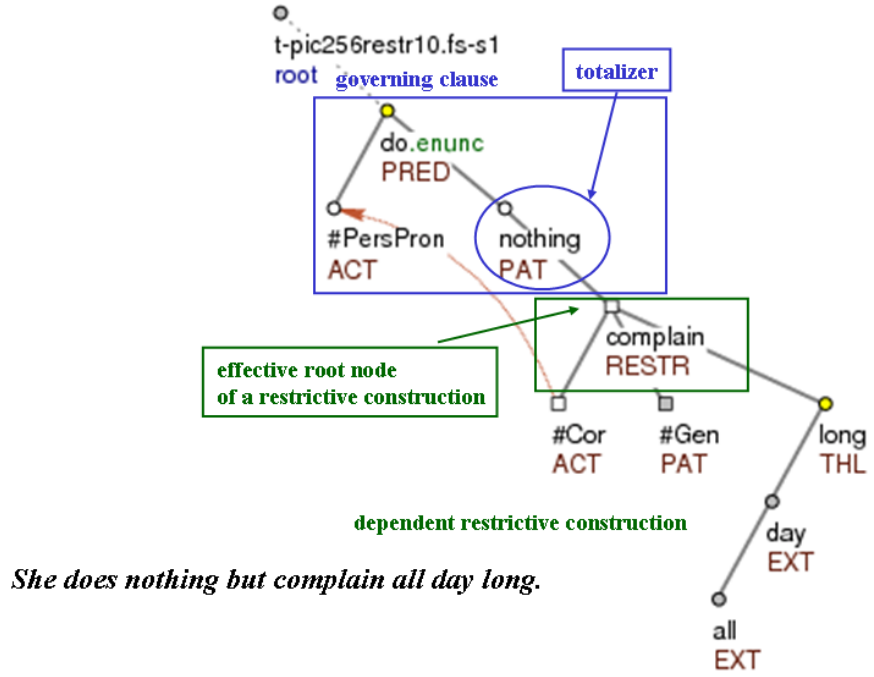
Mary is the same as John.

4.6.1.3.2. 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 *every, each*, *whole*, entire, *all*, *nothing*, *nobody, never, nowhere, none* 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. 4.83.

Figure 4.83. Constructions with the meaning of a restriction



Examples of restrictive constructions with totalizers

<With the exception of> James.`RESTR`, *none of us had the money.*

<Beyond the> press release.`RESTR`, there were no further comments.

All <but/except(ing)> the captain.`RESTR` *were rescued.*

The worst period of my life, apart from the war.RESTR*, was when I was out of work.*

She does nothing <but> complain.RESTR all day long.Fig. 4.84

You can leave any time <but> now.RESTR = *You can leave any time except now.* (*any* is regarded as part of the adverbial *any time* - therefore it is *time* as the governing node that is regarded as totalizer)

This car is anything.PAT *<but> slow*.RESTR

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 (*every, each*, *whole*, entire, *all*, adjectives/adverbs in superlatives, *any* etc.) or a negative totalizer (*nothing*, *nobody, never, in no way, nowhere, none, no* +noun, *not a* +noun etc) or *hardly* + indefinite pronoun/adverb: *hardly anyone, hardly ever. Not* governed by the verb (represented as #Neg.RHEM) is not a totalizer in this sense. Also *beyond* is sometimes used as except in non-assertive contexts and should be treated the same way. The added node for the totaliser has a functor corresponding to the position in which it was added. For example:

<Except for> dates.RESTR*,* {#Total.RSTR} *ordinals should be written in words.*

= *Except for dates, all ordinals should be written in words.* Fig. 4.85

We can <but> guess.RESTR *at the extent of the problem.*

Several Japanese companies already operate in the U.S.: Nissan, Honda, Fujitsu, to name.PAR {#Total.PAT}*<but> a few* #PersPron.RESTR

He didn't help {#Total.MANN}*, beyond showing*.RESTR *a mild interest.*

Effective root of a restrictive construction. The effective root of a restrictive clause and its position in the tectogrammatical tree are determined depending on what means are used for expressing the restriction and what discourse function the restriction has. All restrictive constructions in the first group get the functor RESTR. The second group includes some constructions that are likely to originate in restrictive constructions but their meaning and function have further developed so that the original restrictive meaning has bleached to such an extent that they are assigned a different functor.

A. **Ordinary restrictive construction introduced by a preposition**. As for constructions introduced by the prepositions e*xcept (for)*, *excepting*,bar, barring, *apart from, save*, *unless with the exception of* and the functional word *but* (to be held apart from the coordinating conjunction) there are several types:

   a. **direct restriction**: The effective root node of the restrictive construction is the node representing the governing (syntactic) noun in the prepositional phrase, an adverb or *do nothing but* +infinitive without *to, or a gerund*. No totalizer is needed since the governing clause contains an element expressing "normality", "regular state", etc. The restrictive construction is then governed directly by this expression. For example:

      <Except> this week.RESTR*I'll be teaching regularly.*Fig. 4.86

      We had a pleasant time, <except for the> weather.RESTR

      Even the construction of the type

      Who should turn up but our old friend Tom. (= It was no-one else but our old friend Tom who turned up)

      which serves moving the noun phrase into the focus, is resolved as an ordinary restrictive construction:

      Who should turn up <but> our old friend Tom.RESTR

   b. **restriction with ellipsis of the verb**. The following constructions are regarded as ellipsis of a governing verb. The main verb node is copied on the position of the effective root of the dependent clause. When the governing clause is positive and the meaning of the governed clause is interpreted as restrictive (opposite to exceptional conjoining, see Section 4.6.1.3.2, “Constructions with the meaning of a restriction” [81]), the restrictive construction is governed by a totalizer (an originally present or an inserted one) and the empty-verb node gets a negation (#Neg.RHEM). When the governing clause is negative and the meaning of the governed clause is interpreted as restrictive the restrictive construction is governed by a totalizer (an originally present or an inserted one) and the empty-verb node gets no negation.

i. The preposition with restrictive meaning is followed by another prepositional phrase

Lane closures will effect Castle Boulevard and Castle bridge Road the 5th February up to the 25th March, access is maintained apart from to the Castle

= *Access is maintained everywhere apart from that access is not maintained to the Castle.*Fig. 4.87

= Access is maintained to all lanes except the lane that leads to the Castle.

Reformists have no access to free media except to the Internet.

= *Reformists have no access to free media except that they have access to the Internet*.Fig. 4.88

I will go nowhere but to Prague.

= I will go nowhere (but/)except that I will go to Prague.

= I will only go to Prague.

ii. The preposition with restrictive meaning is followed by *when, where, what*, provided the wh-clause itself is not to be governed by an empty noun:

I hardly ever get a chance to study except when the children have gone to bed.

= I hardly ever get a chance to study except that I do get the chance to study when the children have gone to bed.

but:

The result of the scan looks very good apart from where the error occured!

= The result of of the scan looks very good apart from the places where the error occured!

= The result of of the scan looks very good <apart from>`#EmpNoun.RESTR` *where the error occured*.`RSTR`!

The restrictive constructions described in this group are often ambiguous and their interpretation is up to the annotator. When one of the clauses is negated the construction means restriction. When both the main clause and the dependent clause are positive or both are negative the restrictive construction most likely has the meaning of exceptional cojoining and has to be annotated in a different way (see below).

c. **restrictive construction introduced by *unless*** and *apart from the fact that*. The effective root of the restrictive constructions is the effective root of the dependent clause, which is governed by the predicate of the governing clause. For example:

I won't go there <unless> they ask.`RESTR` *me to.*

We do not share email addresses with third parties <unless> required.`RESTR` *to do so by law.*Fig. 4.89

<Apart from the fact that> she sings.`RESTR` *well she can do nothing.*Fig. 4.90

This resolution is also applied to constructions with negative statements with *but* + clause, used for saying that something does not happen without something else happening or being true (which are to be held apart from adversative coordinations!)

I never take a bath <but> the phone rings.`RESTR[is_member=1]` *or*.`DISJ` *someone knocks*.`RESTR[is_member=1]` *on the door.*Fig. 4.91

B. **Formally similar constructions that have to be resolved in a different way**:

I. **negative irreal condition** with *but for* + noun phrase (used for saying that something would have happened if something else or someone had not prevented it).

*But for John, we would have lost this match.(= Had it not been for John, we would have lost this mach.)*Fig. 4.92

An `#EmpVerb.COND` (see the rules in Section 4.6.1.1, “Ellipsis of a governing meaning unit”) is to be inserted into the clause with *but for* and a `#Neg.RHEM` is to be inserted as its daughter with John as `ACT`. The parts of the expression *but for* are attached to `#EmpVerb` as its auxrfs.

II. **negative real condition** with *barring* + noun phrase (used for saying that something will happen if something or someone does not prevent it, perhaps even introduced with other subordinators with the meaning of restriction).

<*Bar/barring*> #EmpVerb.RESTR #EmpVerb.COND *accidents*.ACT, *we'll be there on time.*Fig. 4.93

The preposition with restrictive meaning is followed by a noun typically denoting an event or a state. The noun phrase *X* can be paraphrased with "except when X happens", except when *X* occurs". The effective root of the restrictive clause is a newly established node for a verb (see the rules in Section 4.6.1.1, “Ellipsis of a governing meaning unit”). It governs another empty-verb node with the functor COND, which stands for the hypothesized conditional clause. The governing noun of the noun phrase *X* is the ACT of the #EmpVerb.COND. No negation is added, unlike the negative irreal condition in which the original restriction clause is not captured any more.

III. **intensifying** *all but* is resolved as idiom:

He did all but strangled me. (= He almost strangled me. = He did all except that he strangled me)

i.e. *do* is kept as the root of the verbal phraseme *do all* and it is at the same time the effective root of the sentence. *Do all* is treated as a verbal predicate, which is why *all* gets the functor DPHR. This DPHR-frame of *do* requires the free modification RESTR in the valency lexicon. Attaching the function word *but* to the predicate of the dependent clause as auxrf at least slightly recalls the relation of this construction to restrictive constructions.

He did.PRED all.DPHR <but> {#Cor.ACT} strangled.RESTR me.PAT (Fig. 4.94)

IV. **Relativizing restrictive construction introduced by clause coordination.** This construction is mainly used in spoken language to make the previous statement seem less true or less possible. This type of construction is resolved as a coordination of the main clause with a hypothesized clause represented by #EmpVerb. The effective root of the restrictive construction is its predicate verb.

I would have told.PRED[is_member=1] *them the truth* #Comma.ADVDS *except/except_that*.CM *they wouldn't have believed*.PRED[is_member=1] *me*.Fig. 4.95

I'd be glad to help, except that I'm going to be away this weekend.

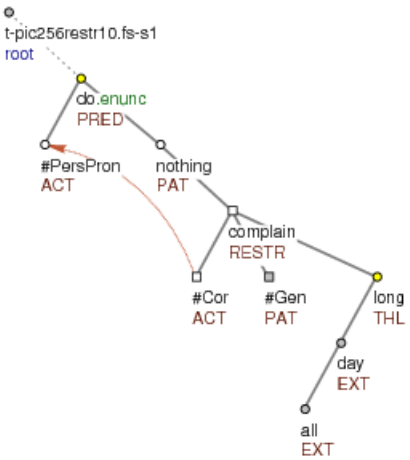
The beast is starving -- except that it isn't a beast.

We were handling it (the gun) safely, except that it was loaded.

I have become rather like King Midas, except that everything turns not into gold but into a circus.

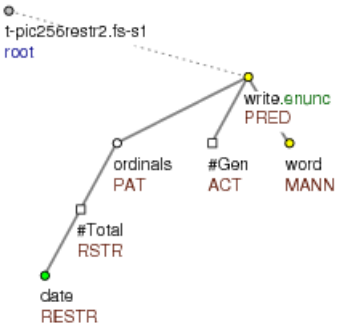
It looked like a state funeral, except that many of the state's top figures weren't there.

Figure 4.84. Constructions with the meaning of a restriction



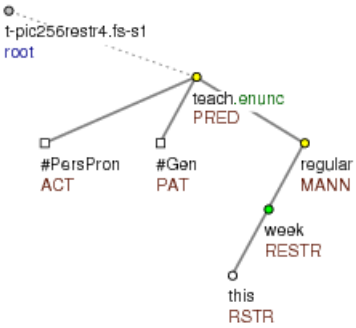
She does nothing but complain all day long.

Figure 4.85. Constructions with the meaning of a restriction



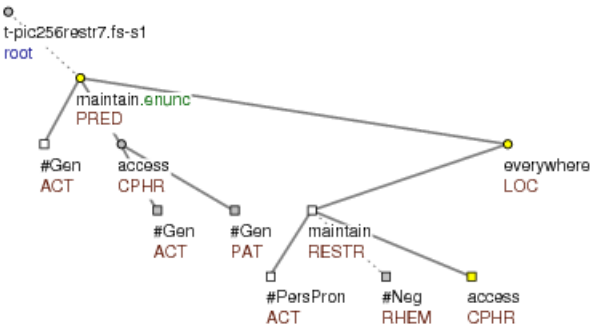
Except for dates, all ordinals should be written in words.

Figure 4.86. Constructions with the meaning of a restriction



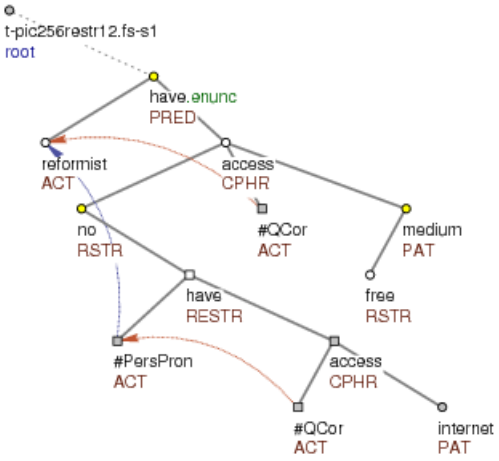
Except this week I'll be teaching regularly.

Figure 4.87. Constructions with the meaning of a restriction



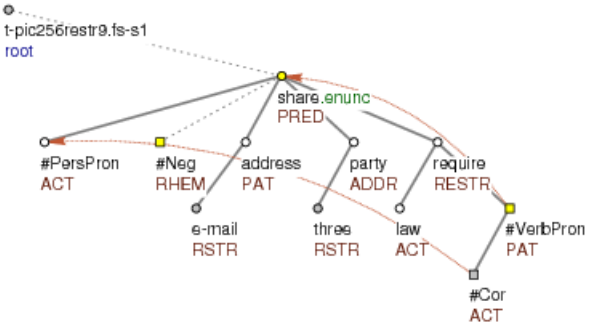
(Lane closures will effect Castle Boulevard and Castle bridge Road the 5th February up to the 25th March), access is maintained everywhere apart from to the Castle.

Figure 4.88. Constructions with the meaning of a restriction



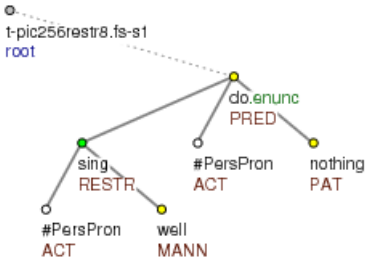
Reformists have no access to free media except to the Internet.

Figure 4.89. Constructions with the meaning of a restriction



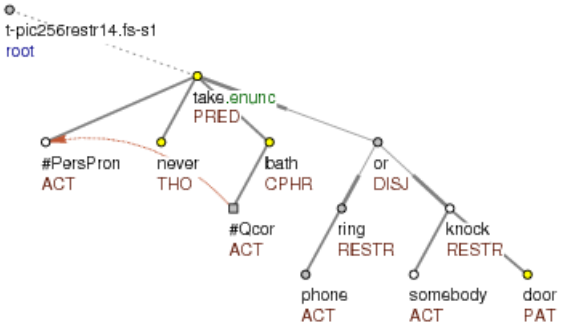
We do not share email addresses with third parties unless required to do so by law.

Figure 4.90. Constructions with the meaning of a restriction



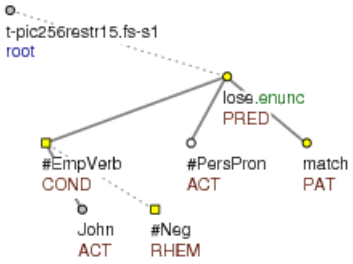
Apart from the fact that she sings well she can do nothing.

Figure 4.91. Constructions saying that something hardly ever happens without something else to happen at the same time



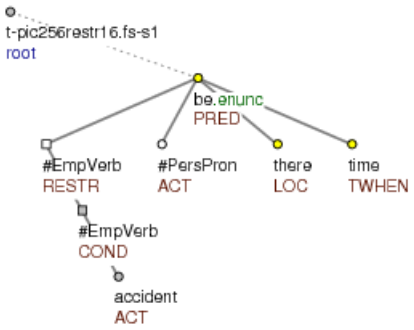
I never take a bath but the phone rings or someone knocks the door.

Figure 4.92. Negative irreal condition



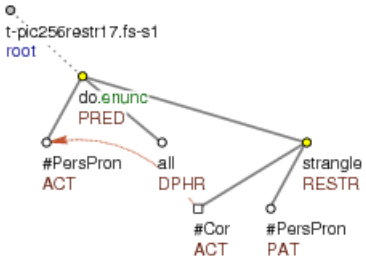
But for John, we would have lost this match.

Figure 4.93. Negative real condition



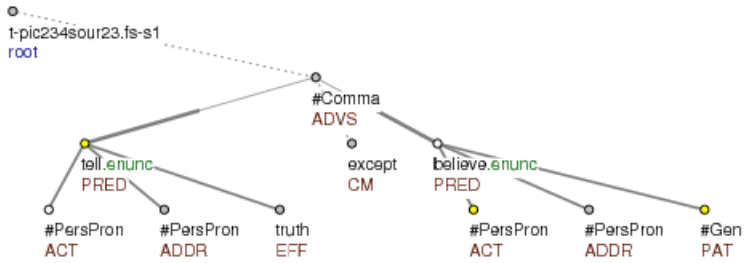
Barring accidents, we'll be there on time.

Figure 4.94. Intensifying all but resolved as idiom



He did all but strangled me.

Figure 4.95. Relativizing restrictive construction introduced by clause coordination



I would have told them the truth, except/except that they wouldn't have believed me.

Constructions with the meaning of exceptional conjoining. Constructions with the meaning of exceptional conjoining are different from standard restrictive constructions (with the prepositions *apart from* and *except*); they do not express a simple restriction nor simple conjoining. They are typically introduced with *besides*. They seem to help the speaker select the actual object of interest out of two (or more):

Besides to Rome, they also wanted to go to Venice.

= *Besides wanting to go to Rome they also wanted to go to Venice.* Fig. 4.96

(= "They wanted to go to Rome but that's not what I have in mind. I mainly want to say that they wanted to go to Venice.")

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 (not on a totalizer).

NB: Even constructions introduced with *except* and *apart from* can acquire the meaning of exceptional conjoining:

You cannot enter this trial if you have breast cancer that has spread to another part of the body, apart from to the lymph nodes.

= You cannot enter this trial if you have breast cancer that has spread to another part of the body beyond the lymph nodes.

(= You can enter this trial when you have breast cancer than has not spread yet. You can still enter the trial even when the breast cancer has spread to your lymph nodes. However, if it has spread elsewhere, you cannot enter the trial.)

And have you been in the hospital as a day patient apart from when you were having a baby?

= *And have you been in the hospital as a day patient excluding times that you have been in the hospital as a day patient to have a baby?*

= *("We are only interested in your hospital stays when you were treated for some illness. We are not interested in deliveries in the hospital you might have been through.")*

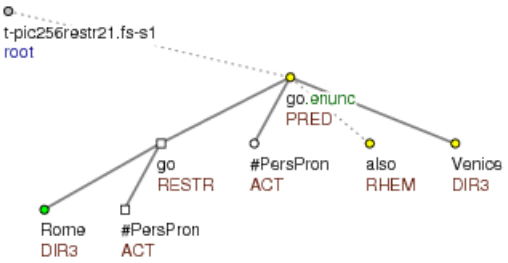
The lieutenant-governor is therefore the province's vice-regal representative, but the occupant rarely exercises the executive role apart from with the advice of the relevant provincial Premier.

= The lieutenant-governor is therefore the province's vice-regal representative, but the occupant rarely exercises the executive role apart from that he exercises the executive role with the advice of the relevant provincial Premier.

(="The lieutenant-governor is therefore the province's vice-regal representative, but the occupant rarely exercises the executive role alone. Mostly he is just acting on the advice of the relevant provincial Premier.")

The constructions with the meaning of exceptional conjoining are governed by the main predicate of the sentence (i.e. not by a totalizer). Their effective root gets the functor `RESTR`.

Figure 4.96. Constructions with the meaning of exceptional conjoining



Besides to Rome, they also wanted to go to Venice.

4.6.1.3.3. 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 *so much*, *enough*, *too*, *so*, *such*, *sufficient(ly)*, *to such a degree/to such an extent*, which introduces the consecutive clause.

Herdwicks (start lambing a year after other breeds because they) need to grow enough.EXT *to stand*.RESL *the harsh climate.*Fig. 4.98

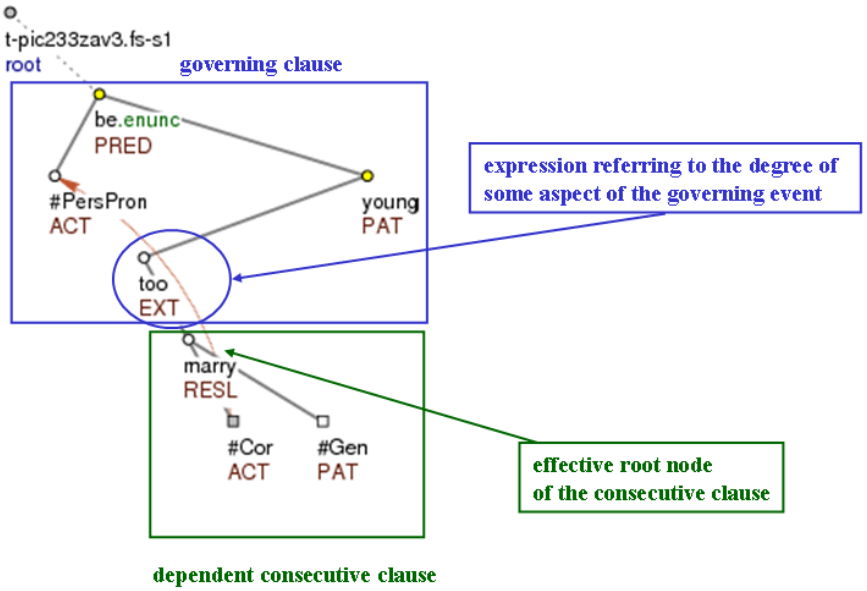
Representing constructions with consecutive clauses. The node for the expression referring to the degree of some aspect of the governing event is assigned a functor according to its position; if it is a non-valency modification, it is usually assigned the EXT functor. The effective root of the consecutive 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. 4.97. Depending on the context, even other expressions than too, so, such, enough etc. can act as quantifiers, e.g. *necessary*:

They have the necessary **.** EXT *votes to block*.RESL *the law.* = They have the necessary amount of votes = They have enough votes.

However, this does not apply to postmodifying adjectives:

T*hey have the votes necessary*.RSTR *to block*.AIM *the law.*

Figure 4.97. Representing constructions with consecutive clauses



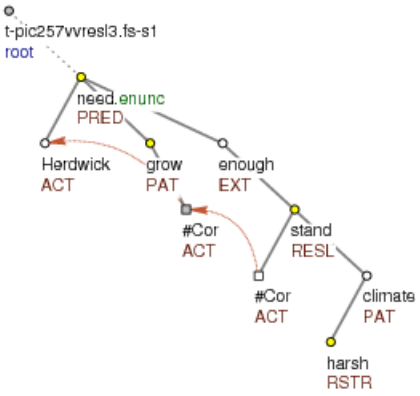
I'm too young to marry.

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 the 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 (*too, enough*). For example:

They have `#AsMuch.EXT` *the votes to block the law.*

= T*hey have enough votes to block the law.*

Figure 4.98. Dependent consecutive clause



Herdwicks (start lambing a year after other breeds because they) need to grow enough`.EXT` *to stand*`.RESL` *the harsh climate.*

***Enough* as part of a copula-predicate.** When depending on the copula verb to be, *enough* has the functor `EXT` as usually.

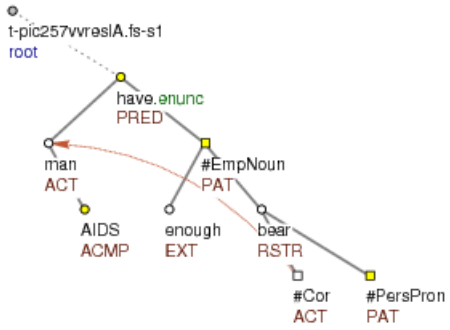
Fumes are often enough`.EXT` *to activate*`.RESL` *the alarm.*

Enough **acting as adjective**. When *enough* depends on an adverb or an adjective it acts as adverb. When depending on a noun, it acts as adjective. The only case it is governed by a verb is when acting as part of a copula-predicate (see above). Grammatical as well as contextual ellipses of the governing noun in the adjectival uses of *enough* are possible:

People with AIDS have enough.EXT `#EmpNoun.PAT` *to bear*.RSTRFig. 4.99

(... daisy flowers.) Given the variety available there are enough.EXT `#PersPron.ACT` to fill.RESL (most of) our summer with colour.

Figure 4.99. Infinitive after a quantifier like enough



People with AIDS have enough to bear.

4.6.2. 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 4.7, "**Newly established nodes in more complex syntactic structures**").

Table 4.7. 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 4.4, "Parataxis". For more on list structures for foreign-language expressions see Section 5.4, "Foreign-language expressions". For more on identification structures, see Section 5.3.1, "Identification structure".

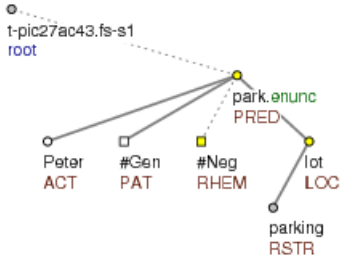
4.6.3. Nodes representing negation with verbs

A specific type of newly established node is a node for negation expressed by *not* depending on a verb. *Not* with negated verbs is represented by a newly established node with the t-lemma substitute `#Neg` (the node is either a rhematizer (`functor` = `RHEM`), or conjunction modifier (`functor` = `CM`; see Section 5.1, "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:

Peter didn't park.`RHEM` *in the parking lot.* {`#Neg`} Fig. 4.100

Figure 4.100. Nodes representing negation with verbs



Peter didn't park in the parking lot.

The current annotation has no special resolution for *no*.

I have no.RSTR *idea*.

Neither negated adverbs and pronouns (never, nobody) indicate negation yet. This issue is going to be subject to significant changes in the next future.

4.6.4. Survey of newly established nodes (ordered according to their t-lemmas)

The following table (Table 4.8, "**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 4.8. 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	*They have* #AsMuch *the votes to block the law.*
#Benef	ellipsis of the Beneficiary in control constructions	*It is difficult* {#Benef.BEN} *to leave.*
#Cor	controlled element	*He decided* {#Cor.ACT}*to leave.*
#EmpNoun	grammatical ellipsis of the governing noun	*Only the poor* {#EmpNoun.ACT} *came.*
#EmpVerb	grammatical ellipsis of the predicate	{#EmpVerb.PRED} *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	*He did it* {#Equal.MANN} like Tony did.CPR *it.*
#Forn	root node of a list structure for foreign-language expressions	*This was a* {#Forn.PAT} *faux pas.*
#Gen	general argument	*Mike is reading* {#Gen.PAT}.
#Idph	root node of an identification structure	*I'm reading* {#Idph.PAT} *The Hours.*
#Neg	verbal negation (*not*)	{#Neg.RHEM} *I'm not leaving.*
#PersPron	contextual ellipsis of an obligatory argument	{#PersPron.ACT} *Go!*
#QCor	quasi-controlled element	*I have made* {#QCor.ACT} *the decision to study.*
#Rcp	obligatory modification missing due to its participation in a reciprocal relation	*The members of the Parliament discussed the new law* {#Rcp.ADDR}.
#Separ	paratactic structure root node	{#Separ.CONJ} *Purchase Sales Midpoint*
#Some	verbless part of an elided verbonominal predicate, which cannot be represented by a copied node due to semantic reasons	*He is just* like me {*be*} {#Some.PAT}.
#Total	ellipsis of a totalizer governing a dependent restrictive clause	*Except for dates* {#Total.RSTR} *ordinals*.RESTR *are written in words.*
#Unsp	unspecified Actor	{#Unsp.ACT} *It was announced on the radio.*
copied node (verb)	grammatical ellipsis of the predicate	*(George is all ears.) John* {*be*.PRED} *{all ears* **.** DPHR*} too.*
copied node (noun)	contextual ellipsis of the governing noun	*He likes red* {*wine*.PAT} *as well as white wine. (=He likes both red and white wine.)*

Chapter 5. Specific phenomena

5.1. Multi-word lexical units

5.1.1. Types and representation

Multi-word lexical units are 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 3.3, “Complex t-lemma”.

B. **single node + grammateme**: 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 grammatemes. This holds for:

   modal predicates (the `deontmod` grammateme; see Section 5.1.2.2, “Modal predicates”)

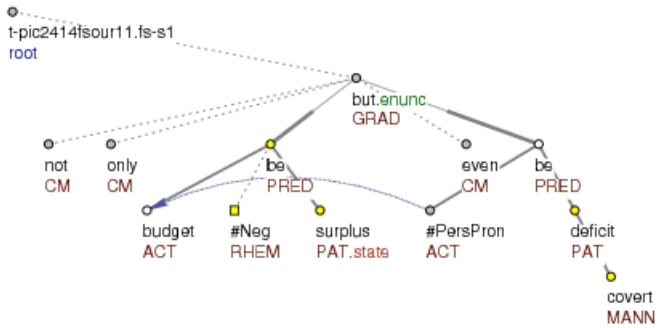
C. **more than one node+ special functor**: the multi-word lexical unit is represented by several nodes and the fact that these form a single unit is captured by using a special functor. This holds for:

   support verb constructions - (also called **light verb constructions**, the `CPHR` functor; see Section 5.1.2.5, “Support verb constructions”)
   idioms (the `DPHR` functor; see Section 5.1.3, “Idioms”)
   multi-word (coordinating) connectives (the `CM` functor)
   sequences 'preposition-noun-prepostition' acting as prepositions (here preliminarily called subjunction modifiers, the `SM` functor)

   Multi-word coordinating connectives. When analyzing coordinating connectives, the following two are distinguished: **coordinating conjunction** and **conjunction modifier**. A multi-word co-ordinating connective is represented by at least two nodes: a node for the conjunction (`nodetype = coap`; see also Section 4.4, “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:

   Not.`CM` *only*.`CM`*is the budget not in surplus, but it is* even.`CM`*covertly deficit.*Fig. 5.1

   Figure 5.1. Multi-word coordinating connectives



   Not only is the budget not in surplus, but it is even covertly deficit.

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:

   titles with no governing noun (see Section 5.3.1, “Identification structure”)
   sequences of foreign text (see Section 5.3.1, “Identification structure”)

5.1.2. 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 alone but where the predicate additionally incorporates the meaning of other words - verbs, nouns, adverbs. Multi-word predicates include:

auxiliary predicates (see Section 5.1.2.1, “Auxiliary predicates”),
modal predicates (see Section 5.1.2.2, “Modal predicates”),
phase predicates (see Section 5.1.2.3, “Phase predicates”),
quasi-modal and quasi-phase predicates (see Section 5.1.2.4, “Quasi-modal and quasi-phase predicates”),
support verb constructions (see Section 5.1.2.5, “Support verb constructions”),
copula predicates (see Section 5.1.2.6, “Copula predicates (the copula “be”)”),
verbal idioms (see Section 5.1.3, “Idioms”).

5.1.2.1. Auxiliary predicates

The following verbs and verb forms are regarded as auxiliary verbs and their nodes do not have their own nodes in the tectogramatical representation:

be going to + infinitive (future)
be + *-ing* (progressive tense, progressive infinitive)
be + past participle (passive)
will + infinitive (future)
would + infinitive (future in the past, conditional, politeness, iterativeness)
shall + infinitive (future)
have + past participle (perfect tense, past infinitive)
have been + *-ing* (past progressive infinitive)

(negated) auxiliary verb + non-negated present infinitive of a full verb.	*Peter will come to the concert.* *Peter <will> come to the concert.*Fig. 5.3
(negated) auxiliary verb + paratactic connection of non-negated present infinitives of full verbs	*Peter won't come and stay.* *Peter <will not> come and stay* .Fig. 5.4
paratactic connection of (negated) auxiliary verbs + non-negated present infinitive of a full verb.	*Peter should not, and will not, come to the concert.* *Peter <should not> {come} and <will not> come to the concert.* Fig. 5.5
paratactic connection of (negated) auxiliary verbs + paratactic connection of non-negated present infinitives of full verbs	*So they should not and will not confirm or deny it.* *So they <should not> {potvrdit} and <will not> confirm or {deny}* it.Fig. 5.6

5.1.2.2. 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. Only the "core" modal auxiliaries and *ought to* are treated as modals in the English annotation. All other predicates with modal meaning (e.g. *to have to, to be able to, be bound to* etc.) are resolved as quasi-modal predicates.

The following are distinguished:

- ***modal verb***: *can*; *could; may*; *might*; *shall*; *should*; *must; ought to; will* (even in the rare cases when used to express volitionality).
- **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 grammateme (`deontmod`). Prospectively, evident cases of epistemic modality will be marked separately. The following combinations are represented in the basic way (by copying the nodes) :

(negated) modal verb + non-negated present infinitive of a full verb.	*Peter can come to the concert.*

	Peter <can> come to the concert. [deontmod = poss] Fig. 5.3
(negated) modal verb + paratactic connection of non-negated present infinitives of full verbs	*Peter cannot come and stay.* *Peter <cannot> come* [deontmod = vol] *and stay* [deontmod = poss]. (Fig. 5.4)
paratactic connection of (negated) modal verbs + non-negated present infinitive of a full verb.	*Peter could not, and cannot, come to the concert.* *Petr <could not> {come}* [deontmod = poss] *and <cannot> come* [deontmod = poss] *to the concert.* Fig. 5.5
paratactic connection of (negated) modal verbs + paratactic connection of non-negated present infinitives of full verbs	*So they could not and cannot confirm or deny it.* *So they <could not> {potvrdit}* [deontmod = poss] *and <cannot> confirm* [deontmod = poss] *or {deny}*[deontmod = poss] it [deontmod = poss].Fig. 5.6

The basic annotation cannot be used for modal and auxiliary predicates with a negated infinitive of the full verb (predication negation) and cases of more modal meanings within one modal predicate (layering of modal meanings). This concerns the following combinations:

(negated) modal verb + negated present infinitive of a full verb	*I can't not obey her. (=It is impossible for me not to obey her.)*
(negated) auxiliary verb + negated infinite form of a full verb	*He hasn't ever not understood a lecture. (= He hasn't ever failed to understand a lecture..)*

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 grammateme deontmod 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.

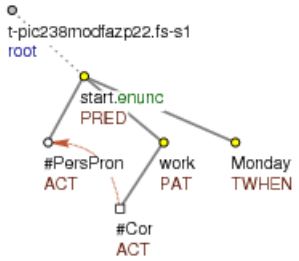
5.1.2.3. Phase predicates

Phase predicate is a multi-word predicate consisting of a phase verb like *begin*, *start*, *stop, become, turn, grow*, get, *cease, keep, come to* 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:

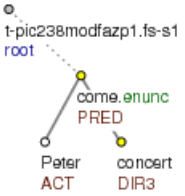
I <will> start.PRED *to work*.PAT *on Monday.* (Fig. 5.2)

Figure 5.2. Auxiliary predicate



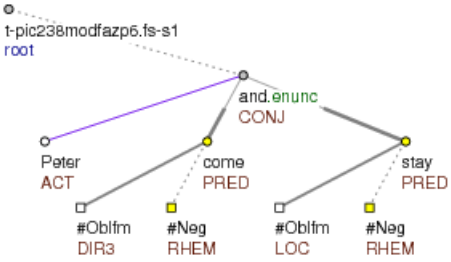
I will start to work on Monday.

Figure 5.3. Modal or auxiliary predicate



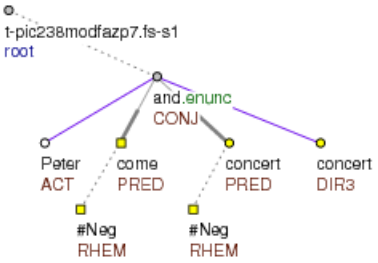
Peter can come to the concert./Peter will come to the concert.

Figure 5.4. Modal or auxiliary predicate



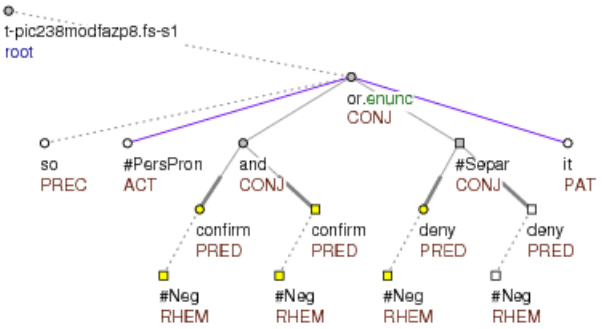
Peter cannot come and stay./Peter will not come and stay.

Figure 5.5. Modal or auxiliary predicate



Peter could not, and cannot, come to the concert./Peter should not, and will not, come to the concert.

Figure 5.6. Modal or auxiliary predicate



So they could not and cannot confirm or deny it./So they should not and will not confirm or deny it.

5.1.2.4. 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. **a quasi-modal or quasi-phase verb**, i.e. by a multi-word synonym of a modal verb (semantically empty verb + noun or adjective 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, two 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 5.1, "**Representing quasi-modal and quasi-phase verbs**".

b. **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:

   have the right to be treated properly → have the right to proper treatment

   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 grammateme.

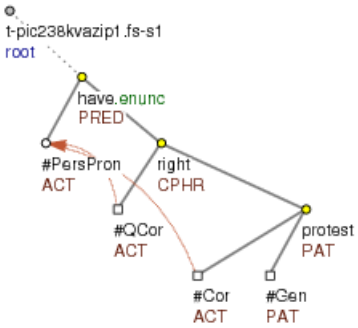
Table 5.1. Representing quasi-modal and quasi-phase verbs

Definition of the group	Representation	Examples
verb (not *be*) + noun	following the same rules as in the case of support verb constructions (see Section 5.1.2.5, "Support verb constructions")	*have the right*
být + adjective or noun	following the same rules as in the case of copula predicates (see Section 5.1.2.6, "Copula predicates (the copula "be")")	*be able* *be necessary* *be obliged*

Example:

He had the right.CPHR *<to> protest*. (Fig. 5.7)

Figure 5.7. Quasi-modal predicate



He had the right to protest.

5.1.2.5. Support verb constructions

Support verb constructions or ***light verb constructions*** 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 support verb construction 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.:

talk → have a talk
to claim → to make a claim
to limit → to impose a limitation
to be interested → to show interest

The existence of an adequate synonymous one-word expression is not however necessary for regarding a certain collocation of a semantically empty verb and a meaning-bearing noun as a support verb construction.

The following parts of complex predicates are distinguished:

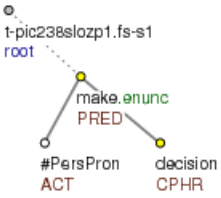
- **the verbal part of a support verb construction** (support verb).

- **the nominal part of a support verb construction** (predicate noun), i.e. the dependent noun, which carries the main lexical meaning of the entire predicate.

Basic annotation rules for support verb constructions. A support verb construction is represented in the tectogrammatical tree by two nodes: by a node representing the verbal component of the support verb construction and by a node representing the nominal component of the support verb construction. The node representing the verbal part is assigned a functor according to the function of the entire support verb construction 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:

She made a decision.`CPHR`Fig. 5.8

Figure 5.8. Support verb construction



She made a decision.

5.1.2.5.1. Quasi-control with support verb constructions

The fact that a support verb construction 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 support verb constructions share certain valency modifications. This is called **quasi-control**, a specific type of grammatical coreference

Representing quasi-control with support verb constructions. 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 support verb construction. If the valency modification occurs in the valency frames of both parts of the support verb construction 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 support verb construction. A simple convention has been accepted for these cases: the shared modification is represented by a node dependent on the verbal part.

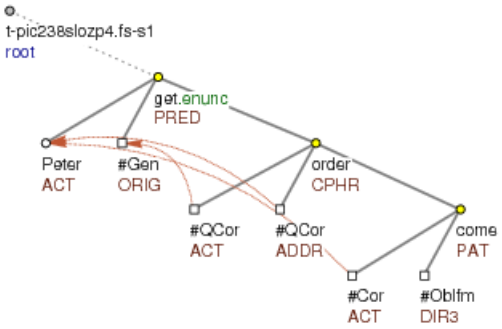
In place of the other (shared) valency modification that is omitted at the surface level (as a rule it is a valency modification of the nominal part of the support verb construction), a new node is added to the tectogrammatical tree, with the t-lemma substitute `#QCor`. In addition to the special t-lemma, the 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 support verb construction 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 4.6.1.2, “Ellipsis of a dependent meaning unit”); thus: `#Gen`, `#PersPron`, possibly `#Unsp`.

Examples:

Peter got {`#Gen.ORIG`} {`#QCor.ACT`} *the order to come.* (Fig. 5.9)

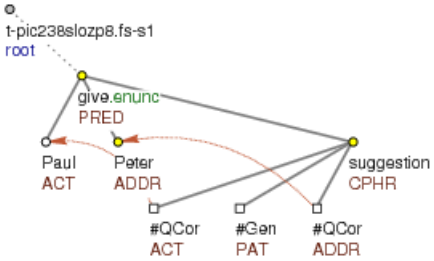
Paul.`ACT` *gave Peter*.`ADDR` {`#QCor.ACT`} {`#QCor.ADDR`} *a suggestion*.`CPHR`. (Fig. 5.10)

Figure 5.9. Quasi-control with support verb constructions



Petr got the order to come.

Figure 5.10. Quasi-control with support verb constructions



Pavel made Petr a suggestion.

5.1.2.6. Copula predicates (the copula “be”)

Copula predicate is a collocation of the verb *be* or *become* (whose function is primarily to carry the grammatical meanings) and another word, carrying the main lexical meaning of the predicate. Also verbs like *seem, appear*, *turn, grow, get* etc. can act as copula verbs. Copula predicates can express a wide range of meanings: identity of the Actor and Patient, description, classification or quantification. With copula predicates, we distinguish:

- **the verbal part**, i.e. the copula *be*.
- **the non-verbal part**, i.e. the semantic adjective or noun in the nominative (or possessive *hers, yours, ours* and *theirs*). The non-verbal part can also be expressed by a noun in the genitive, by an infinitive, a dependent clause, an adverb or interjection.

Representing copula predicates in the tectogrammatical trees. Copula 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 *be*). For example:

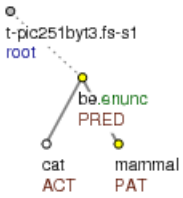
The cat is a mammal.`PAT`Fig. 5.11

Eating is a ritual.`PAT`

Jirka is kind.`PAT` .

This is mine.`PAT`

Figure 5.11. Copula predicate



The cat is a mammal.

5.1.3. 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.: *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` (`nodetype` = `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 3.3, “Complex t-lemma”).

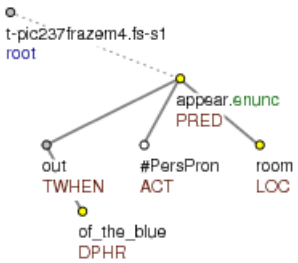
Example of a non-verbal idiom:

Out.`TWHEN`*of the blue*.`DPHR`*they appeared in our room.* (Fig. 5.12)

Example of a verbal idiom:

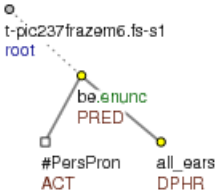
We can talk now, I'm.`PRED`*all ears*.`DPHR`. Fig. 5.13

Figure 5.12. Non-verbal idiom



Out of the blue they appeared in our room.

Figure 5.13. Verbal idiom



I'm all ears.

5.2. Numbers and numerals

5.2.1. Main rules

There are several groups of **numerals** and they are distinguished in the annotation:

a. **numerals in the role of an attribute (RSTR)**: numerals in the combination with countable objects (except for the numerals with the function of a container). The governing node is the node of the countable noun, on which the numeral depends. The node of the numeral is assigned the RSTR functor. For example:

I've got five.RSTR *houses and three*.RSTR *cars.* (Fig. 5.14)

our third.RSTR *presentation*

b. **cardinal numerals without the countable object**. The numeral is considered a syntactic noun. The node for the numeral can get different functors. For example:

They elected three.PAT*out of five vice-chairs.*Fig. 5.15

How much.PAT*will you give me?*

c. **numerals with the function of a "container"**: numeral expressions in plural: *milions, thousands, hundreds* but mainly words like *(one) third*, *quarter*, *dozen*, *heap*, *row*, *a lot*, *pile*, *crowd*, *couple*, *trio*, *a bit* with a countable object introduced by the preposition *of*. The governing node is the node of the numeral-container, the node of the countable object depends on it. The node of the countable object is assigned the functor MAT. For example:

Billions of people.MAT*live here*.

Also *half* belongs to this group when **followed** by a determiner (i.e. acting as a predeterminer):

We live halv a mile.MAT *up the road.*

The rebel army now controls over half the country.MAT.

I have to spend half my time.MAT *taking care of the children*

or when followed by the preposition *of* (i.e. when acting as a governing noun or pronoun):

the second half of life.MAT

You can keep half of the money.MAT.

Half of us.MAT *are still unemployed.*

Half is **never** a container numeral when acting as number, adverb or adjective:

Emma was talking well by the age of two.RSTR *and*.CONJ *a half*.RSTR. (number)

She spoke half.EXT *in Italian and*.CONJ *half*.EXT *in English*. (adverb)

I only half.EXT *understood what she was saying*. (adverb)

We ordered a half.EXT [1] *bottle of red wine*. (adjective)

[1] see Section 8.9.1, "Adjectives as noun modifiers" why half as adjective is not RSTR.

They sold 250 tickets in the first half.EXT *hour.* (adjective)

Unlike in the predeterminer position, in the adjectival position *half* is **preceded** by a determiner.

d. **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 ID functor. For example:

The new Golf 500.ID *is already on the market.*Fig. 5.17

e. **numerals with adverbial meanings**: numerals: *five times*,*several times*, *once*, *for the second time* etc. The node for the numeral is assigned an adjunct functor corresponding to the position in the structure. For example:

We only won twice.THOFig. 5.18

I won't do it for the second.RSTR*time*.TWHEN.

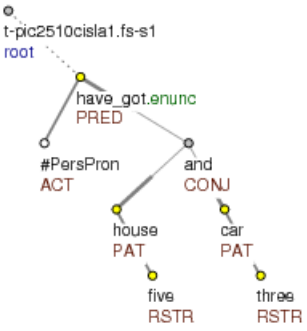
NB! Adjectival numerals and numeral expressions *more*, *less*, *much/many*, *little/few*, *enough* used separately, without a countable object, are considered syntactic adjectives (attribute function). A new node for the governing noun is added to the structure, following the rules in Section 4.6.1.1, "Ellipsis of a governing meaning unit". For example:

We have missed the third.RSTR {#EmpNoun.PAT} Fig. 5.19

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.:

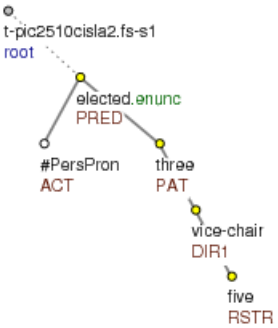
I've got 38 234.RSTR *crowns*.PAT

Figure 5.14. Numerals as attributes



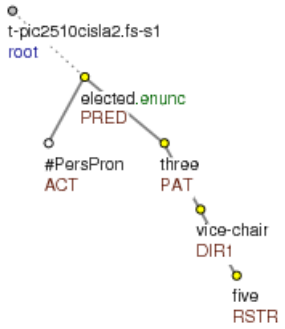
I've got five houses and three cars.

Figure 5.15. Cardinal numerals without countable object



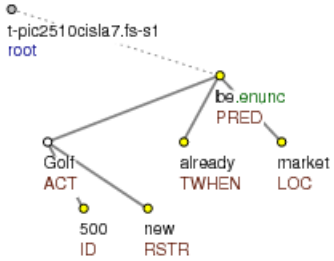
They elected three out of five vice-chairs.

Figure 5.16. Numerals with the function of a "container"



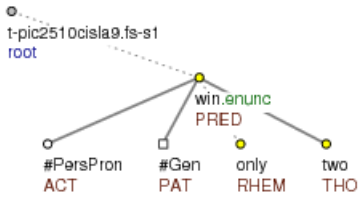
Billions of people live here.

Figure 5.17. Numerals used as "labels"



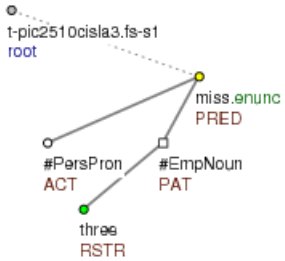
The new Golf 500 is already on the market.

Figure 5.18. Numerals with adverbial meanings



We only won twice.

Figure 5.19. Numerals - syntactic adjectives



We have missed the third.

5.2.2. Specific constructions with numerals *(next best, second-best, a three-month stay, $50 a share, five times, for the first time, 5 years old, 5 m long)*

The combination *next best* is resolved as two sister nodes with the functor `RSTR`, governed by the noun. Its synonym *second-best*, however, written with dash, is one node with the t-lemma *second-best*. Generally, all multi-word

lexical units connected with dash are rendered as one node with the t-lemma corresponding to the entire expression in the basic form (singular, infinitive). Therefore, a three-month.THL stay.

Constructions of the type $50 a share have $ as the governing node on which the number depends as RSTR and *a share* as REG.

Constructions of the type *five times, for the first time* have *time* as governing node and the given number as its RSTR daughter. The type *how many times* gets the functor THO, the type *for which time* gets the functor TWHEN.

In the constructions of the type *5 years old, 5 m long*, it is the adjective (*old, long*, etc.) that is the governing node in the subtree. The unit (*years, meters*, etc.) gets the functor EXT and the number gets the functor RSTR.

5.3. Identifying expressions

Identifying expressions are expressions used for identification (proper names, and titles, meta-language use).

Identifying expressions are divided into five basic groups according to the following criteria:

1. do they only consist of a proper name? *(Prague, the Kremlin)*
2. do they consist of a proper name and a common noun (a descriptor)? *(Lake Michigan, the novel The Hours)*
3. when comprising a descriptor, is the descriptor integral part of the name *(Lake Michigan)* or not *(Kent cigarettes)*?
4. being a company's name, does the identifying expression seem to be governed by an abbreviation?
5. when without descriptor, does the name include a noun not governed by any preposition that can be the governing node of the entire name *(Amadeus, Angela's Ashes, Two Towers, Hair)* or does the name lack such an apparent governing element *(On the Road, On Mice and Men, Breaking the Waves, Guess Who's Coming to Dinner)*? Coordinated structures are governed by the coordination node, and therefore they are regarded as those lacking the governing noun (e.g. *Harry Potter and the Goblet of Fire*)

The following types of identifying expressions are thus distinguished in the annotation:

a. **Names comprising proper nouns without descriptors or names comprising proper nouns with descriptors as integral components**. The effective root of an identifying expression has a functor according its position in the structure. Nodes dependent on the effective root are analyzed according to the standard annotation rules. Articles that are obviously part of the identifying expression are governed by the nouns and get their own node with the functor INTF. The annotation considers the article to be part of the identifying expression when:

   - it immediately precedes a proper noun: *the Kremlin, the Koran, The Hague.*
   - when it precedes a name in general: *The Times, The Paris Peace Talks*

   On the other hand, it is treated in the regular way when occurring inside of an identifyig expression: *Kerouac's*.AUTH {#Idph.DENOM} *<On> <the> Road*.ID

   The article is always the left sister of adjectival and nominal modifiers and the right sister of possible predeterminers.

   The effective root of the identifying expression of this type gets the appropriate functor according to itsfunction in the sentence. For example:

   I'm reading The.INTFHours.PAT

   The.INTF *United*.RSTR*Nations*.APP*Organization*.DENOM*for Education*.BEN, *Science*.BEN*and*.CONJ*Culture*.BENFig. 5.20

   His.APP *Hamlet*.ACT *is all torn.* (Fig. 5.21)

b. **Names introduced or postmodified by a descriptor that is not integral part of the name**: these are represented as identification structures (see Section 5.3.1, “Identification structure”). For example:

   the bookThe.INTF*Hours*.ID

   the Kent.ID*cigarettes* (*the* is attached as auxrf to *cigarettes*)

c. **Names of companies including an abbreviation like *Ltd.*, *GmbH.*, etc.:** names of companies are not analyzed at this annotation stage. All members of the subtree get the attribute value [is_name=1], and they will be annotated together later.

d. **Names lacking a governing noun in a prepositionless case** *(On the Road, On Mice and Men, Breaking the Waves, Guess Who's Coming to Dinner)*: Like names introduced or postmodified by a descriptor that is not integral part of the name, this type of identifying expression is represented as identification structure (see Section 5.3.1, "Identification structure"). An artificial governing node with the t-lemma substitute #Idph is inserted as the effective root of the subtree.

Descriptor, which is not integral part of the name: The descriptor as non-integral part of the name is a (common) noun written in small letters, introducing (or following) a proper noun, title, an expression used metalinguistically or an expression quoted word for word. It can also occur with the preposition *of*:

the city <of> Prague.`ID`

the notice "Danger!".`ID`

Explicative *of*-attribute: an expression consisting of an *of*-phrase modifying a common or a proper noun while the following transformation is applicable: *the concept of time → time is a (kind of) concept. the person of Christ → Christ is a (kind of) person*. This specific type of identifying expressions is analyzed with the help of an identification structure (see Section 5.3.1, "Identification structure"). For example:

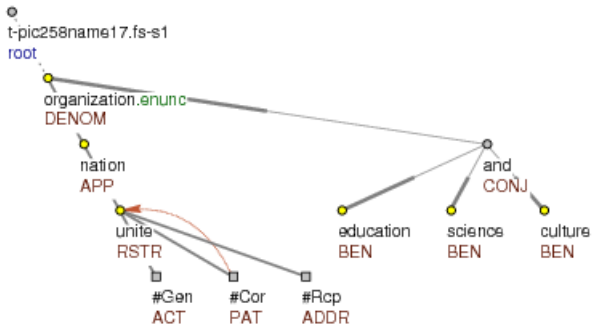
the concept <of> time.`ID`

the issue <of> impeachment.`ID`

the person <of> Christ.`ID`

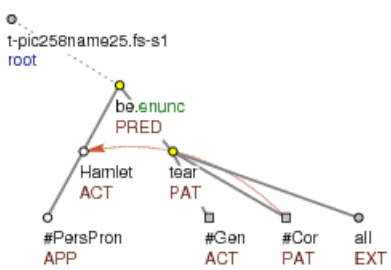
Proper names of people (the `is_name_of_person` attribute). At all nodes representing expressions which are constituents of proper names of people (nodes representing first name or surname) the value `1` is entered in the attribute `is_name_of_person`.

Figure 5.20. Identifying expression



The United Nations Organization for Education, Science and Culture.

Figure 5.21. Identifying expression



His Hamlet is all torn.

5.3.1. Identification structure

Identification structure is illustrated in Fig. 5.22 and Fig. 5.23.

Figure 5.22. Identification structure I

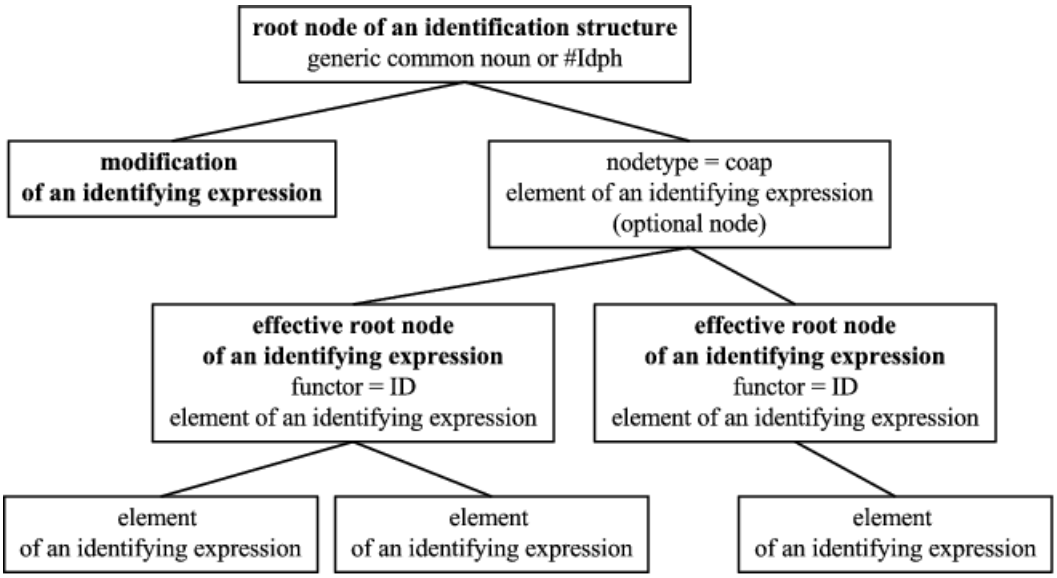
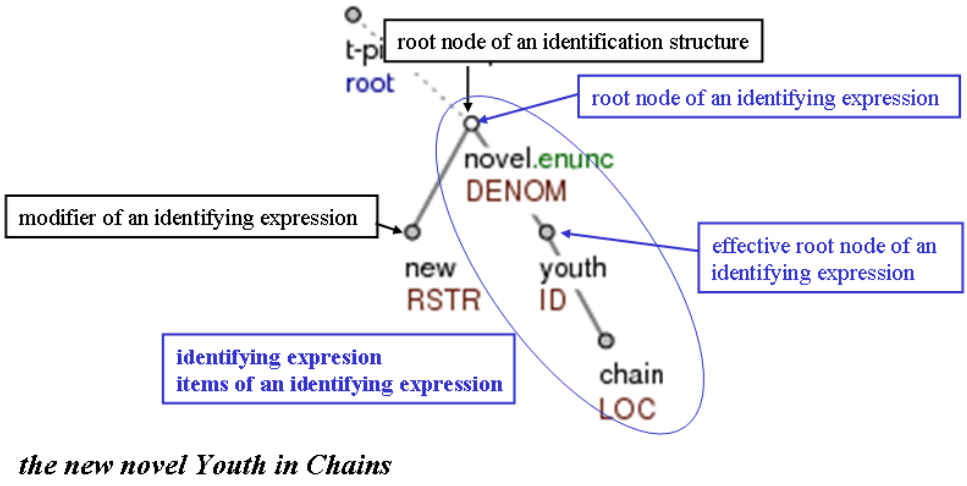


Figure 5.23. Identification structure II



Root node of an identification structure. The root of an identification structure is the node representing the descriptor. If the identifying expression does not contain a noun in prepositionless case and it occurs at the surface level without any descriptor, 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 the 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 with **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: when 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.

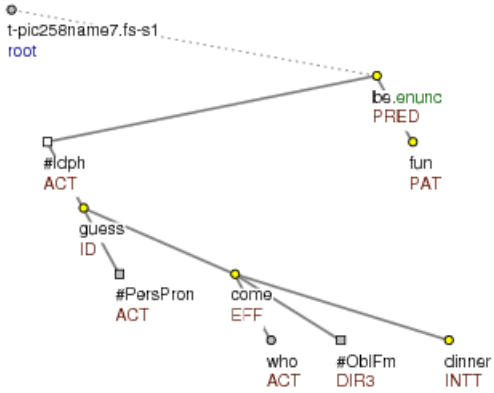
Modifiers of identification structures. Identifying expressions can be further modified (as a whole). The root of the modifier is always an immediate daughter node of the root of the identification structure.

Examples:

{#Idph.ACT} Guess.ID {#PersPron.ACT} *Who's Coming*.PAT *to Dinner is fun*.PAT (Fig. 5.24)

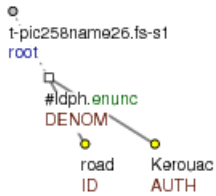
Kerouac's.AUTH {#Idph.DENOM} <*On the*> *Road*.ID (Fig. 5.25)

Figure 5.24. Identifying expression



Guess Who's Coming to Dinner is fun.

Figure 5.25. Identifying expression



Kerouac's On the Road

5.4. Foreign-language expressions

Foreign-language expressions those text segments that are written in a language other than English.

List structure for foreign-language expressions. Foreign-language expressions are represented as a list structure (see Fig. 5.26 and Fig. 5.27). **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 an English expression. The root of the modifier is always an immediate daughter node of the root of the list structure.

Figure 5.26. List structure for foreign-language expressions I

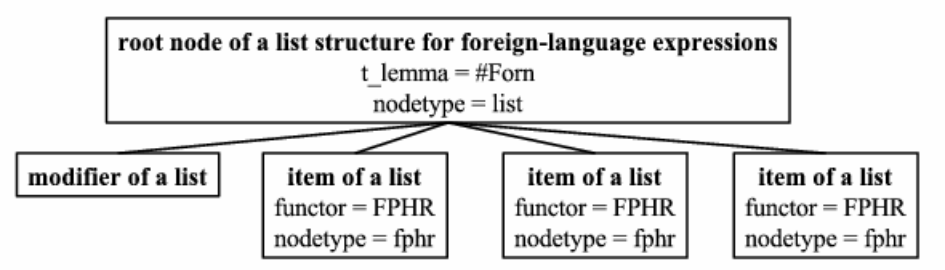
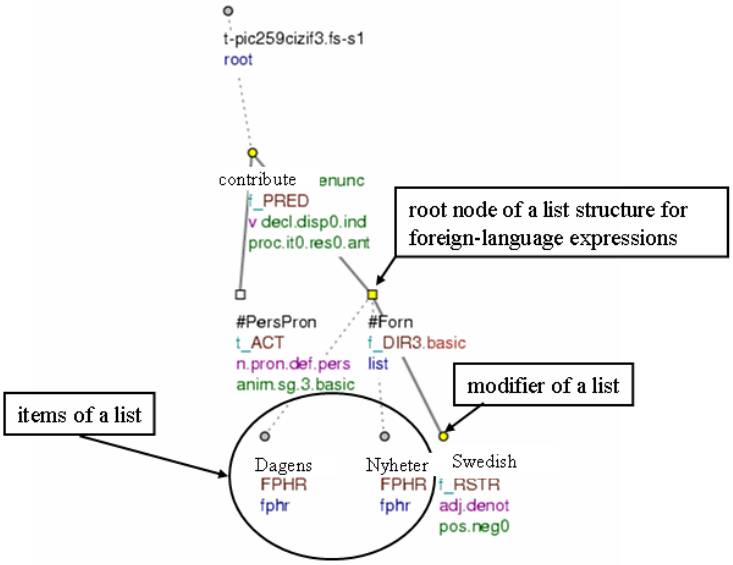


Figure 5.27. List structure for foreign-language expressions II



He was contributing to the Swedish Dagens Nyheter.

Examples:

the company {`#Forn.ID`} *Sady*`.FPHR` [`#Comma.FPHR`] *Lesy*`.FPHR` [`#Comma.FPHR`] Zahradnictví`.FPHR`Fig. 5.28

It was a tremendous {`#Forn.PAT`} *faux*`.FPHR`*pas*`.FPHR`

He was contributing to the Swedish {`#Forn.ID`} *Svenska*`.FPHR`*Dagbladet*`.FPHR`Fig. 5.29

NB! Common loan words, made-up words and foreign-language words when consisting of one token declined/conjugated/used with articles or particles are not represented by means of a list structure:

So many taboos`.PAT` *we keep!*

They were tolshocking`.PRED` *a sexy devotschka*`.PAT` *while I was enjoying my moloko*`.PAT`.

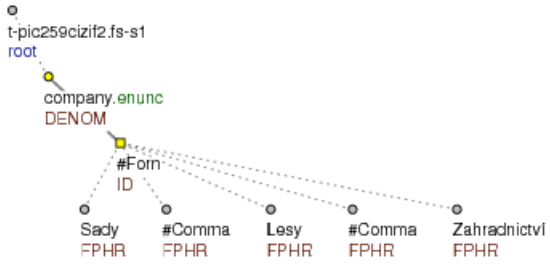
This is a typical chucpe`.PAT`

But: There is certainly a competition between identifying expressions and foreign-language expressions. Whenever possible, foreign-language expression is preferred. Meta-uses and names without descriptor are therefore rather annotated as foreign-language expressions:

the French "oui" - *French*.RSTR#Forn.DENOM*oui*.FPHR.

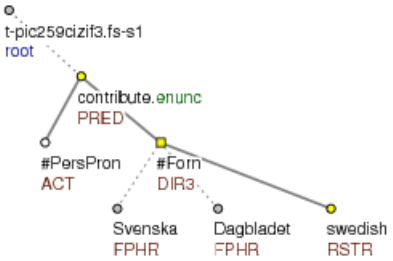
the joy in An die Freude - *the joy in*#Forn.LOC*An*.FPHR*die*.FPHR*Freude*.FPHR

Figure 5.28. Foreign phrase



the company Sady, Lesy, Zahradnictví

Figure 5.29. Foreign phrase



He was contributing to the Swedish Svenska Dagbladet.

5.5. 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`, 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 5.6, “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 4.3, “Clauses (governing, dependent, verbal, verbless)”; i.e. it is represented as a verbal or non-verbal clause. For example:

   „We have.PRED *several hundred members.“* [is_dsp_root=1]

   „Wow.PARTL *!“* [is_dsp_root=1]

   „A new car.DENOM *!“* [is_dsp_root=1]

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:

His answer was: „I will come.PAT*“.*

He said he wouldn't come.EFF [is_member = 1]*: „I will certainly not arrive*.EFF [is_member = 1]*.“*Fig. 5.30

The coach: „It was not.PAT *too bad.“* [#Colon.PRED] Fig. 5.31

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:

I don't share his opinion: „They are going to win.PAT*.“*

His words „I envy.RSTR *him.“ were surprising.*Fig. 5.32

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. *to say*), **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:

But he started {#EmpVerb.PAT} *: „You have to explain*.EFF *a few details.“*

d. if it is possible to attach the direct speech to the reporting clause with the help of a 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 gerund is added to the tree**, with the t-lemma substitute #EmpVerb and the functor COMPL (the node stands for expressions like *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:

He pointed at me {#EmpVerb.COMPL} *: „You*.EFF".Fig. 5.33

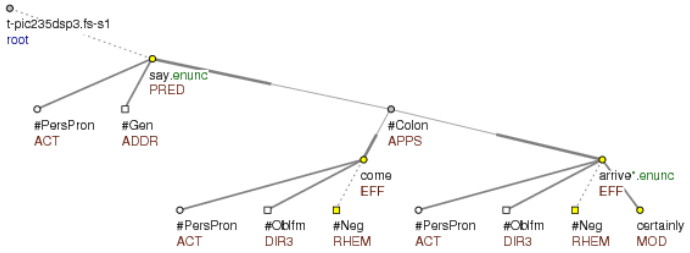
e. if the direct speech is not a modification of the reporting clause, we represent it as **paratactically connected to the reporting clause**. The root node of the paratactic structure is usually a node representing a colon. For example:

The mood is getting.PRED [is_member = 1] *better: „We will beat*.PRED [is_member = 1] *them.“*Fig. 5.34

NB! In case the direct speech is a verbal or a 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:

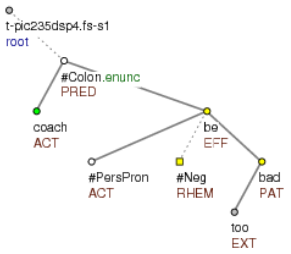
„ {#EmpVerb.EFF [is_dsp_root=1]} *Wow*.PARTL *!“, he exclaimed.*

Figure 5.30. Direct speech as an argument of the reporting verb



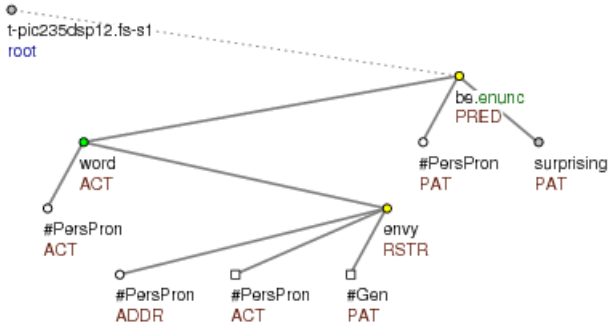
He said he wouldn't come: "I will certainly not arrive".

Figure 5.31. Direct speech as an argument of the reporting verb



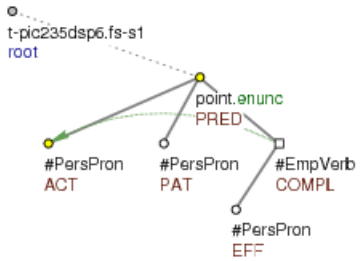
The coach: "It was not too bad."

Figure 5.32. Direct speech as a noun modifier



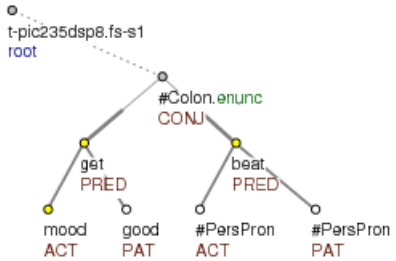
His words "I envy him" were surprising.

Figure 5.33. Direct speech paraphrasable with a verb of saying



He pointed at me: "You".

Figure 5.34. Direct speech is not a modification of the reporting clause



The mood is getting better: "We will beat them".

5.6. 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, 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 5.2, "**Types of quotation mark use**".

Table 5.2. Types of quotation mark use

Type	`quot/type`	Definition	Examples
quotation	`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).	*He added that the FRG did not want to „pressurise"but to "be supportive".* *He said that „it is not necessary."*
direct speech	`dsp`	direct speech (see Section 5.5, "Direct speech"); the quoted speech is a coherent (complete) unit	*„We have several hundred members."* *His answer was: „I will come".*
m e t a - l a n - guage	`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	*the notice "Danger!"* *The verb "to be" is a copula verb.* *Skin full of tatoos means: "I am the boss here".*
title	`title`	proper name, title (identifying expression)	The vessel was called "Admiral Nelson".
other cases	`other`	other types of use (irony, metaphor etc.)	*The company respects the influence of "destiny".* *One district "weighs" 100 kilograms.*

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 4.5, "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 4.5, "Parenthesis"); in other cases the dashes are treated as coordinating connectives.

Chapter 6. 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`.

Subfunctors are to be described in Section 6.13, “Further specification of the meaning of a functor”. This manual version does not include subfunctors yet.

6.1. 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 6.1, “**Functors for effective roots of independent clauses**”.

Table 6.1. Functors for effective roots of independent clauses

Functor	Definition	Example
DENOM	effective root node of an independent nominative clause (which is not parenthetical)	*Our readers' opinions.*
PAR	effective root node of an independent verbal or nominative clause (parenthesis)	*I'll arrive on December 13th (Friday).*
PARTL	effective root node of an independent interjectional clause	*Watch out!*
PRED	effective root node of an independent verbal clause (which is not parenthetical)	*Paul gave a flower to Mary.*
VOCAT	effective root node of an independent vocative clause	*Sir, I intend to introduce some new elements.*

6.2. Argument functors

The basic definitions of the arguments are in Section 4.2, “Valency”. For a brief survey of argument functors and their definitions see Table 6.2, “**Argument functors**”.

NB! The modification with the `MAT` functor is also an argument; it is described in Section 6.10, “Specific adnominal functors”.

Table 6.2. Argument functors

Functor	Definition	Examples
ACT	functor for the first argument In those cases when there is no argument shifting, the modification with the `ACT` functor refers to the human or non-human originator of the event, the bearer of the event or a quality/property or the experiencer.	*The novel appealed to me.* *It appears to me that....* *mafia's crimes on Sicily* *his performance*
ADDR	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.	*He gave the child a toy.* *He took the toy from the child.* *He addressed the court with a problem.*
EFF	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.	*He considered Paul a professional.* *They made each other's life bearable.* *She changed her hairstyle from curly hair to straight hair.*

		He said he wouldn't come.

ORIG	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.	*They produced furniture made out of wood.* *He heard about the accident from the teacher.* *a puddle from* melting *snow*

PAT	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).	*They pitched the tents.* *I was astonished at the quality of the cheeses.* *We forgot to breathe.* *They recommend to organize* a competition. *a book on dinosaurs* *the people responsible for the deportations* of Jews.

6.3. 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 6.3, "**Temporal functors**".

Table 6.3. Temporal functors

Functor	Definition	Examples
TFHL	for how long?	*He came for* two *days.* *I'll be forever grateful.* *containers used for long-time storage of burn-out fuel*

TFRWH	from when?	*sweets from Christmas* *I don't remember anything from my childhood.*

THL	how long? in what time?	*He was doing it for three years.* *He wrote it in two hours.* *He was writing the homework for two hours.* *They haven't seen each other for* a long *time.* *He was working until he felt tired.* a *forty-five minute performance*

THO	how often? how many times?	*Drivers in Austria have to stop every six hours.* *They met daily.* *annual accounts*

TOWH	to when?	*He called the meeting for six o'clock.* *We'll put it off until next time.*

TPAR	during what time? in parallel with what?	During our *holiday it didn't rain once* *Simultaneously with the lecture, there were some seminars.* *On our way to Nuernberg, we stopped three times.)* *While she was sleeping, I was thinking.*

TSIN	since when?	*It didn't rain since Saturday.* *I never liked this town.*
TTILL	till when?	*I'll have arrived in Praha by the evening.* *By this date, 2173 applications were submitted.* *Do it before I come back.*
TWHEN	when?	*I'll come in the evening.* *We'll see each other at the end of the year.* *In the rainy period, the elephants stay in the mountains.* *A week ago, she stopped commuting to work.)* *We'll meet on February 2nd.* *There will soon be a faster means of transport.* *As soon as he learned one language, he started another one.*

6.4. 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 6.4, "**Locative and directional functors**".

Table 6.4. Locative and directional functors

Functor	Definition	Examples
DIR1	where from?	*He made a step from the wall.* *In the direction from Praha, there are more cars.* *one of the boys* *There was some noise coming from inside.*
DIR2	which way?	*They passed near the house.* *He passes hishand across his forehead.* *through the valley* *We can't go this way.*
DIR3	where to?	*She made him come back over the sea.* *We are not going anywhere.*
LOC	where?	*Danger can be everywhere around us.* *Stay at home.* *a match in a foreign country*

6.5. 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 6.5, "**Functors for causal relations**".

Table 6.5. Functors for causal relations

Functor	Definition	Examples
AIM	purpose, aim	*I am here to help you.* *For the sake of improvement of the process...* *He worked in order to make some money.*
CAUS	cause proper	*I am not going to tell you because you wouldn't believe me anyway.* *Because of you I'm in a bad mood.* *Closed for illness.* *He died of thirst.*
CNCS	concession	*Although his manners are normally good...* *Fouad Siniora says he won't step down, despite three days of protests in Beirut.*
COND	condition	*If he fights, we'll use the force.* *If he doesn't arrive there will be no meeting.* *If it were not for you I wouldn't be here.*
INTT	intention	*The people are coming here for skiing.* *He rushed to make the picture.* *He stayed for lunch.* *The Izraeli Prime Minister is in Prague on a visit.*

6.6. 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 6.6, "**Functors for expressing manner and its specific variants**".

Table 6.6. Functors for expressing manner and its specific variants

Functor	Definition	Examples
ACMP	accompaniment	*He works without his glasses.* *Father with Mother* *In connection with the growth in wages* *The parents bought two sets with the idea that they'll give the sons each one.*
CPR	comparison	*smaller and easier than the last one.* *just like it was in America.* *Compared to you I'll always be better.*
CRIT	criterion	*He is judging people on the basis of their clothes.* *In the light of these facts...* *This was not registered which is not in accordance with the rules.*

DIFF	difference	*The goods were delivered four days later.* *The older the wine the better it is.* *it eased a fraction of a percentage point*

EXT	extent	*When people by thousands were coming west during the Pike's Peak excitement, he decided to join the tide of emigration that moved westward.* *The costs were almost one billion crowns.* *The weather is quite nice today.* *The shares fell as much as 50 %.*

MANN	manner proper	*He works slowly.* *They enjoyed the victory in silence.* *in the form of an illuminated manuscript.* *He acted as the law required.* *electric control*

MEANS	means	*Using the warning shots, the military scattered the crowd.* *The thief left on the bike.* *The shares can be bought through a broking house.*

REG	regard	*It is necessary to prepare the spot with regard to the number of participants.* *As for this law, nobody proposed voting...* *mentally handicapped* *life insurance* *diabetes in children*[a] *decrease in rates* *problems with verbs* *yields on money-market mutual funds*

[a]This is a temporary solution. In the future annotation, the functor RSTR is being considered for the cases X in Y that can be paraphrased as *Y does X* or *Y has X*.

RESL	result	*They pushed the car so.EXT that it ended up in a ditch.* *The author is trying to make it such.RSTR that anybody can read it.*

RESTR	exception/restriction	*Apart from its current job, the office is going to issue...* *The society is not evil, except for certain individuals.* *He wants nothing but a good job.* *He won't do it unless he gets well paid.*

6.7. 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 6.7, "**Functors for rhematizers, sentence, linking and modal adverbials**".

Table 6.7. Functors for rhematizers, sentence, linking and modal adverbials

Functor	Definition	Examples
ATT	evaluating or emotional attitude of the speaker to the content of the utterance	*Fortunately it's over now.* *This is without exaggeration the new Honda.* *Technically the war was over but there were still fights.* *so-called business* *Frankly, she shouldn't have said it.*
INTF	expression intensifying the subject (so called "false subject")	It is raining.
MOD	modal characteristics of the content of the utterance	*We'll probably come.* *This would apparently lead to...* *Such workers hardly exist.* *possible failure*
PREC	expression linking the clause to the preceding text	*Hence, I'm happy.* *But this is not our case.* *However, isolated research cannot have good results.*
RHEM	rhematizer	*Only he didn't know anything.* *As late as a week ago she started commuting to Brno.*

6.8. Functors for complex lexical units and foreign-language expressions

Functors for complex lexical units and foreign-language expressions are functors assigned to nodes representing a foreign-language expression (see Section 5.4, "Foreign-language expressions"), or the dependent part of a multi-word predicate (see Section 5.1.2, "Multi-word predicates"). For a survey of the functors for complex lexical units and foreign-language expressions and their definitions see Table 6.8, "**Functors for complex lexical units and foreign-language expressions**".

Table 6.8. Functors for complex lexical units and foreign-language expressions

Functor	Definition	Examples
CPHR	nominal part of complex predicates or a predicative adverb with quasi-modal verbs	*They got the order not to leave the tents.* *It is necessary to leave.*
DPHR	the dependent part of an idiomatic expression	*He's getting on my nerves.* *criss cross*
FPHR	foreign-language expression	*The foreigner called: "To je pravda!"*

6.9. Functors for some specific 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 6.9, "**Functors for some specific (new) modifications**".

Table 6.9. Functors for some specific (new) modifications

Functor	Definition	Examples
BEN	adjunct expressing to whose advantage or disadvantage something happens	*He did it for their sake.* *a chance for rich tenants* *This is another argument against the existence...*
CONTRD	adjunct expressing a contrasting background (confrontation) for the governing event	*While in the past...*, *in the future ..*
HER	adjunct referring to a person (group of people, institution, time) which is the source of an inherited object	*She got a pension from JF.* *The prize is named for the late Mexican writer Juan Rulfo.*
SUBS	adjunct expressing that something was substituted for something else	*...in the name of the company...* *...fathers instead of their sons.* *In place of my father...*

6.10. 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 adverbal modifications. For a survey of the specific adnominal functors and their definitions see Table 6.10, "**Specific adnominal functors**".

Table 6.10. Specific adnominal functors

Functor	Definition	Examples
APP	adjunct referring to the person or thing something or someone belongs to	*organization of the hearing-impaired* *my hat*
AUTH	author, creator, originator of artefacts	*young poets' poetry books* *Schindler's List by Spielberg*
ID	effective root of an identifying expression represented by an identification structure	*in the Lewinsky case* *notion of time* *Kent cigarettes* *the Tate gallery*
MAT	adnominal argument referring to the content (material etc.) of something	*a box of paper* *one half of the cake* *billions* of people
RSTR	adnominal adjunct more closely specifying its governing noun	*harsh weather* *five children* *referee Severýn* *a vessel with a displacement of 9700 tons* *He even hit a man who didn't deserve it.*
DESCR	nonrestrictive attribute in postposition	*Mr. Campbell, who is a lawyer, was here last night.* *Pierre Vinken , 61 years old , will join the board as a nonexecutive director Nov. 29 .*

6.11. 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 6.11, "**Functor for the predicative complement**").

Table 6.11. Functor for the predicative complement

Functor	Definition	Examples
COMPL	predicative complement	*They found their friend ill.* *The players were leaving the field undefeated.* *He watched him how he behaved to his younger classmates.*

6.12. 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 4.4, "Parataxis"). For a survey of the functor for relations between members of paratactic structures see Table 6.12, "**Functors for coordination**", Table 6.13, "**Functor for apposition**" and Table 6.14, "**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 6.15, "**Functor for conjunction modifiers**".

Table 6.12. Functors for coordination

Functor	Relation between the coordinated elements	Examples
ADVS	adversative	*The weather seemed nice but it started raining.* *There will be no reference to a law, and still it will be highly relevant.* *He came to Praha, not to Brno.*
CONFR	confrontational	*B. is in England, while G. is in Scotland.* *Bachelors often have a mess all around them but married men, on the other hand, have a mess in their souls.*
CONJ	simple conjoining	*Eyesight and hearing and touch belong to the senses.* *She stands out due to her thoroughness, as well as her stringiness.*
CONTRA	coordination of two fighting, competing subjects	*Academy of Science contra universities* *Bukač versus Hlinka in JH's eyes* *the match Russia - Sweden*
CSQ	consecutive	*He was irresponsible, therefore he was fired.* *He was sick so that's why he didn't come.* *You only need 1/4 of the water so a shower costs 0.46 Kc.*
DISJ	disjunctive	*They will be either for or against the proposal.)* *He had two possibilities - to be outrun, or to cause a crash.*

		Are they right, or not? *At least one parent has to come, or at least a member of the family.)*

GRAD	gradation	*The state will not offer a discount, not even any other financial help.* *The factory fulfilled the plan; it even overfulfilled it.*

REAS	causal	*This step is dangerous since it lets the jungle into our economy.* *We'll fulfil the task, for it is not difficult.*

Table 6.13. Functor for apposition

Functor	Relation between the appositioned elements	Examples
APPS	apposition	*Božena Němcová, the author of Babička* *one year, or two semesters* *soil environment, i.e. soil humidity*

Table 6.14. Functor for mathematical operations and intervals

Functor	Definition	Examples
OPER	paratactic connection of operands of mathematical operations or intervals	*10:1* *at the age of 34-44* from *the main* points *to the tiniest details*

Table 6.15. Functor for conjunction modifiers

Functor	Definition	Examples
CM	expressions modifying coordinating connectives	*Not only is the budget not in surplus, but it is even covertly deficit.*

Table 6.16. Functor for subjunction modifiers

Functor	Definition	Examples
SM	preposition-noun-preposition sequences acting as complex prepositions	*in the aftermath of yesterday's decrease*

For more detail on SM, see Section 7.2, “Multi-word expressions acting as prepositions or subordinators” in Chapter 7, *Subordinators*.

6.13. 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 [1]
the `is_state` attribute (see Section 6.13.1, “Attribute for the meaning of “state””).

6.13.1. 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

[1] not yet defined for English tectogrammatical representation

values 0 and 1. 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:

He is in crisis.LOC [is_state=1]

He was under treatment.TPAR [is_state=1]

person in a bad *mood*.RSTR [is_state=1]

He fell into trouble.DIR3 [is_state=1].

Chapter 7. Subordinators

7.1. Subordinators

Subordinators (subordinating conjunctions) connect finite clauses, *to*-infinitive clauses, bare infinitive clauses, *-ing* clauses, *-ed* clauses and verbless clauses. Many of them can also act as prepositions (not marked here). The tectogrammatical representation regards subordinators as function words, i.e. they usually do not have their own nodes. Though, as a preliminary solution, we keep a tectogrammatical node for the noun in preposition-noun phrase-preposition sequences that act as complex prepositions (e.g. *in the aftermath of*). For more detail and a list see Section 7.2, “Multi-word expressions acting as prepositions or subordinators”.

According to their form, subordinators can be divided into several groups:

- "core" subordinators consisting of one word
- complex subordinators consisting of several words but acting like single-word subordinators
- correlative subordinators

7.1.1. One-word subordinators and the functors assigned to them

Table 7.1. One-word subordinators - gerunds, past participles

notwithstanding	CNCS
concerning	REG
regarding	REG
touching	REG
excepting	RESTR
excluding	RESTR
including	RESTR
barring	RESTR
assuming	COND
given	COND
provided	COND
providing	COND
supposing/suppose	COND
granting/granted	COND
seeing	CAUS

Table 7.2. One-word subordinators

except	RESTR
but (in postmodification)	RESTR
bar	RESTR
Re	REG
beyond (initial position, negative main clause)	RESTR
besides	RESTR
after	TWHEN
as	TPAR, CAUS
before	TWHEN
once	TWHEN
when	TWHEN, COND
whenever	THO, COND
where(ver)	LOC, DIR3, COND
while	TPAR, COND, CNCS, CONFR
whilst	TPAR, COND, CNCS
since	TSIN, CAUS
till	TTILL, COND
until	TTILL, COND
when	TWHEN, COND, CNCS
immediately	TWHEN
directly	TWHEN
if	COND, CNCS
unless	COND
although	CNCS
though[a]	CNCS
whereas	CONFR
adj/noun, 'as', subject, 'be', main clause (Naked as I was, I braved the storm)	CNCS
because	CAUS
for	CAUS
so	AIM, RESL
like	CPR
with	ACMP
without	ACMP

[a]When separated by a comma (*She died, though.*) it has its own node and gets the functor `PREC`.

7.1.2. Fixed complex subordinators acting like one-word subordinators

Table 7.3. Complex subordinators

in addition to	RESTR
as soon as*	TWHEN
as long as*	TPAR, REAS
so long as	TPAR, COND
in the event that	COND
on condition that	COND
in order to/in order that	AIM
in spite of	CNCS
in the event (that)	COND
but that	RESTR
in that	REG
such that	RESL
save that	RESTR
now that	TPAR, REAS
in case	COND
just so (that)	COND
on condition (that)	COND
even though	CNCS
even if	CNCS
whether or not	DISJ
so that	AIM, RESL
such that	RESL, MANN
in that	CAUS
with/for all	CNCS, CAUS
but for	COND
inasmuch as, insomuch as	COND
as if	CPR
as though	CPR
whether or not	DISJ ("Subordinate clause coordinator")

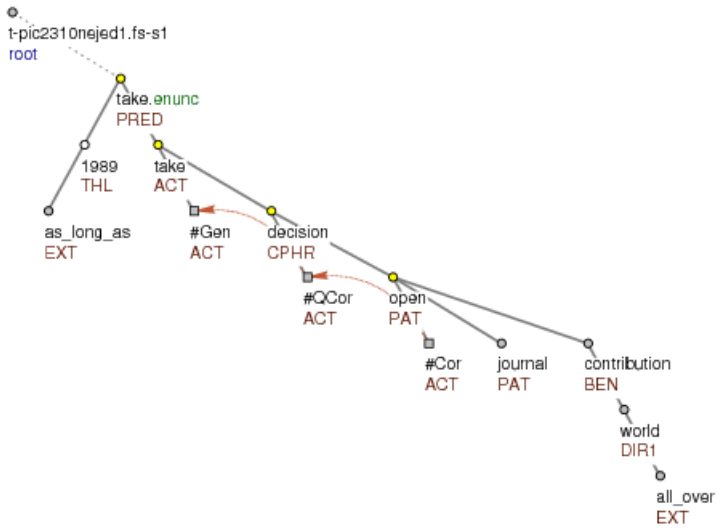
The expressions *as soon as, as long as, as far as* etc. can also act as adverbials:

It took as long as to 1989 until the decision was taken to open the Journal to contributions from all over the world.Fig. 7.1

They get the functor `EXT`. When combined with a numeral expression, they are governed by the numeral with the functor `RSTR`, e.g.:

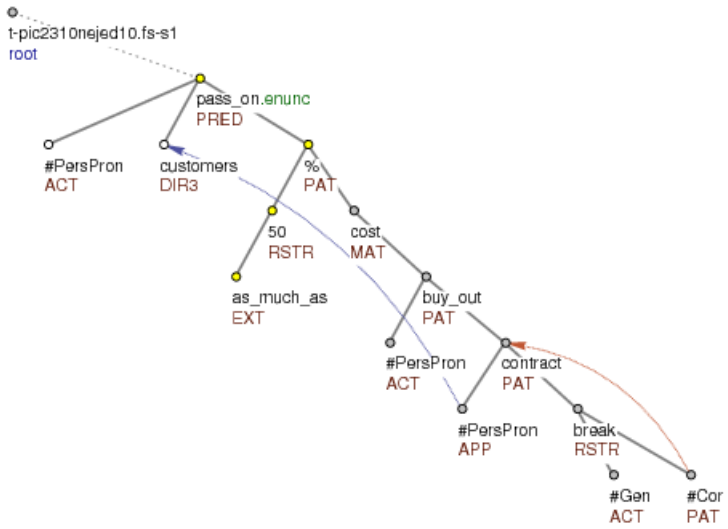
They passed on to customers as much as.`EXT` *50*.`RSTR` *% of the costs of buying out their broken contracts.* (Fig. 7.2)

Figure 7.1. as long as as an adverbial



It took as long as to 1989 until the decision was taken to open the Journal to contributions from all over the world.

Figure 7.2. as much as as an adverbial



They passed on to customers as much as 50% of the costs of buying out their broken contracts.

7.1.3. Correlative subordinators

Correlative subordinators consist of two parts, distributed one in the subordinate clause and the other one in the governing clause (usually the main clause). The main predicate of the subordinate clause gets one component of the correlative subordinator (the actual subordinating conjunction) associated as auxrf (it is hidden). Constructions with correlative subordinators (e.g. comparative and consecutive clauses) often require specific tree structures. The following table lists correlative subordinators. The components in angle brackets are attached to predicates as auxrf's.

Table 7.4. Correlative subordinators

so.EXT.... <as>.RESL		She was so kind as to drive us home.
<as>.... <as>	CPR	He is as old as I am.
such.EXT... <as>.RESL	RESL	
so.EXT... <that>.RESL	RESL	
such... <that>.RESL	RESL	
less.EXT... <than>	CPR	
more.EXT... <than>	CPR	
<as>... <so>	DIFF	As the strenght of the defenders failed.DIFF, so the courage of the attackers grew.PRED[auxrf: as, so].
<the>.DIFF...<the>	DIFF	The older the wine is, the better it tastes.
no sooner.TWHEN <than>	CPR	

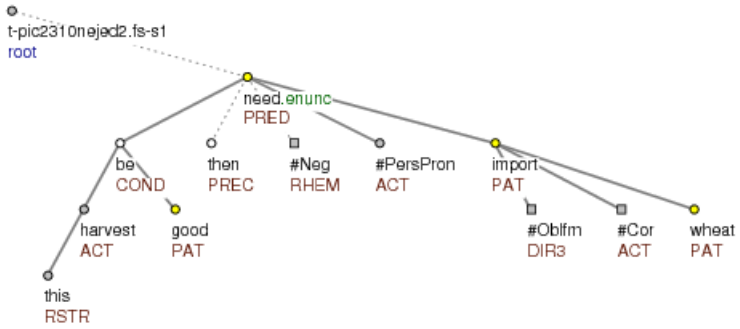
In the correlative subordinators listed above both components must be present on the surface. Yet there are also a few correlative subordinators whose second components are optional. Their second component, present in the governing clause, is not to be hidden but gets the functor PREC:

Though the workers were.CNCS *unhappy with some aspects of the proposed new contract, nevertheless*.PREC *they overwhelmingly voted in favour of it.*

If this harvest is.COND *good, then*.PREC *they will not need to import wheat.*Fig. 7.3

Because you have not replied.CAUS *to my formal letter of May 1, I am therefore*.PREC *withdrawing my offer.*

Figure 7.3. Correlative subordinators



If this harvest is good, then they will not need to import wheat.

7.2. Multi-word expressions acting as prepositions or subordinators

This section concentrates on multi-word expressions that have the ability to act as prepositions, coordinators and subordinators. We discriminate between two-word expressions consisting of a prepositon preceded by an adverbial (*outside of*), a subordinator (*because of*) or a preposition/subordinator (*except for*) and three-word expressions comprising a sequence of preposition-noun-preposition (on the verge of, in face of, in spite of). The two-word as well as the three-word unit group have in common that they comprise a wide range of expressions varying from established prepositions/subordinators/coordinators to collocations that in many contexts act as prepositions.

It was - to some extent - possible to make a list of two-word expressions which are to be regarded as prepositions. They are only expressed by their own nodes in the a-layer. In the tectogrammatical representation they are associated to their governing node as auxrf's. The list is given below.

Table 7.5. List of two-word sequences regarded as prepositions and the functors typically assigned to their parent nodes

as per	CRIT
except for	RESTR
but for	RESTR
apart from	RESTR
away from	RESTR
aside from	RESTR
as from	TSIN
ahead of	TWHEN
back of	LOC, DIR3
exclusive of*	RESTR
instead of	SUBS
outside of	LOC, DIR3
off of	DIR1
upwards of	LOC, DIR3
as of	TSIN
because of	CAUS
inside of	LOC, DIR3
irrespective of	REG
out of	LOC, DIR1
regardless of	REG
according to	CRIT
due to	CAUS
next to	LOC, RESTR
owing to*	CAUS
preparatory to*	TWHEN
prior to*	TWHEN
subsequent to*	TWHEN
as to/for	REG
contrary to*	CPR
close to*	LOC, EXT (except the case named in the next table)
near to	LOC, DIR3
nearer to	LOC, DIR3
preliminary to*	TWHEN
previous to*	TWHEN
pursuant to*	CRIT
thanks to	CAUS
along with	ACMP
together with	ACMP
devoid of*	ACMP
void of*	ACMP

NB that for a two-word sequence to be recognized as preposition none of its components may be a verb particle! This is to be checked in the phrase-structure PTB-tree.

Janice has been together with John four years now. (Fig. 7.4)

The sequences marked with an asterisk comprise adjectives. They are **not** to be regarded as prepositions when they can be regarded as attributes, e.g.:

Sources close to the CEO say he is ready to make a deal. (Fig. 7.5)

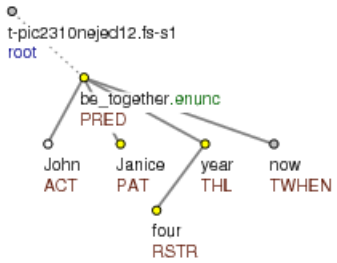
As the line between collocations with prepositional function and complex prepositions is especially fine in the three-word combinations, we treat all these expressions in the same way, unless they already had been tagged as prepositions in the Penn Treebank and thus automatically made to auxrf's during the conversion. Instead of explicitly listing expressions recognized as complex prepositions, which then would not be present as nodes in the tectogrammatical representation, we decided to give the nominal part of the three-word expression its own tectogrammatical node. The surrounding prepositions are attached to it as its auxrf's. The node will get the functor `SM` ("subjunction modifier"). The node governed by `SM` gets the appropriate functor according to the meaning of `SM`. Only the obvious prepositions *in spite of* and *in order to* have been treated as regular prepositions/subordinators.

For instance, in the sentence

But maintaining U.S. influence will be difficult in the face of Japanese dominance in the region.

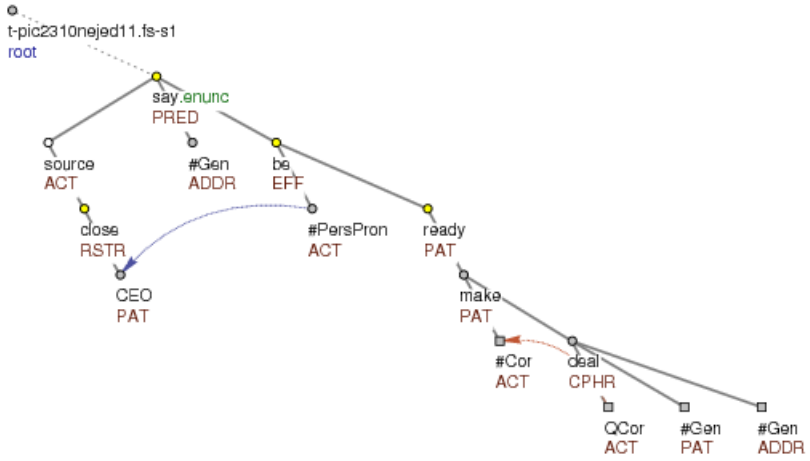
the expression *in the face of* will be governed by the effective root (*be*) and will have the functor `SM`. The word *dominance*, governed by this `SM`, will have the functor `REG`, according to `SM`'s meaning. This is a preliminary annotation whose purpose is to gather all expressions with the structure preposition-noun-preposition in the data that appear as ad hoc prepositions in order to be able to find a systematic resolution for this particular phenomenon.

Figure 7.4. Contexts in which two-word expressions are not to be regarded as prepositions



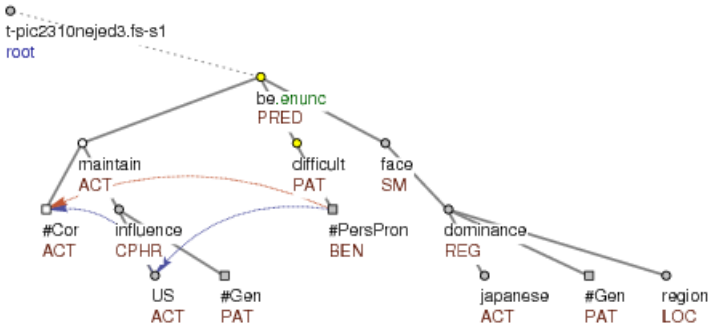
Janice has been together with John four years now.

Figure 7.5. Contexts in which two-word expressions are not to be regarded as prepositions



Sources close to the CEO say he is ready to make a deal.

Figure 7.6. Subjunction modifier (SM)



But maintaining U.S. influence will be difficult in the face of Japanese dominance in the region.

A few criteria have been set to identify `SM`'s:

1. Does the expression act as a preposition? (yes = `SM`)
2. Can the expression be paraphrased by a preposition? (yes = `SM`)
3. Can it be regarded as the non-verbal part of a copula-predicate? (no = `SM`)
4. Can it be, together with the governing verb, paraphrased by a verbal expression?(*to be on guard against something* = *guard against*)(no = `SM`)
5. Is it coordinated with a preposition? (yes = `SM`)

Analogically to two-word sequences, it is mainly to be checked whether the given expression really acts as preposition. It is not the case in multi-word predicates, e.g.:

X stayed in contact.`PAT.[is_state=1]` *with Y.*

Could you put.`PRED` *me in contact*.`CPHR` *with a good piano teacher?*

Our work brings.`PRED` *us into contact*.`CPHR` *every day.*

The verbs *stay, put* and *bring* are used as support verbs and the word contact denotes an event of X contacting Y. Therefore it is expected to keep its valency frame, which results in and therefore it cannot be regarded as preposition. The phrase in contact, into contact can yet be used as preposition when it does not denote an event, typically with inanimate complementations and verbs that usually do not apear as support verbs:

It was placed in contact with the other one (=it was placed how?/where?: so that it touched the other one, it was placed *(on)to* the other one).

In concrete cases the decision whether to regard the given sequence as a preposition or as part of a multi-word predicate is up to the annotator. Ambiguity arises especially when the preposition-like expression seems to fit into the verbal frame. Sequences that fit into the governing verb's frame are expected to be borderline cases. The typical SM's are not eligible for being inner participants of a verb:

Her son was happily playing.`PRED` *at the family home in contradiction of*.`SM` *a prognosis*.`CRIT` *12 years ago that he was not 'worth treating'.* (Fig. 7.7)

In some cases they adopted patterns of behavior in direct.`RSTR` contradiction to.`SM` *the expectations*.CRIT *and desires of the dominant group.*

Table 7.6. List of three-word sequences that act as prepositions (SM's)and the functors typically assigned to their daughter nodes (selected examples)

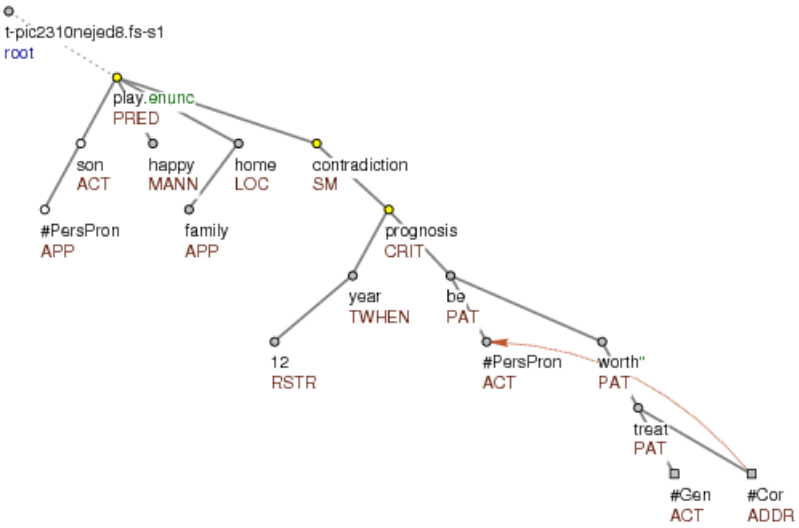
in aid of	BEN
in behalf of	SUBS, BEN
in (the) face of	REG
in front of	LOC, DIR3
in lieu of	SUBS
in place of	SUBS
in quest of	AIM, INTT
in search of	AIM
in view of	REG
in back of	LOC, DIR3
in case of	COND
in consequence of	CAUS
in favour of	BEN
in (the) light of	REG
in need of	CAUS
in (the) process of	TPAR
in respect of	REG
with respect to	REG
in accordance with	CRIT
in comparison with	CPR
in conformity with	CRIT
in line with	CRIT
in common with	ACMP
in compliance with	CRIT
in contact with	LOC, DIR3
by dint of	CAUS, MEANS
by virtue of	CAUS
by means of	MEANS
by way of	MEANS, MANN, REG
on account of	CAUS
on behalf of	SUBS, BEN
on (the) ground of	CAUS
on (the) grounds of	CAUS
on pain of	COND
on the strenght of	CAUS
on the matter of	REG
on the part of	ACT
on someone's part	ACT
on top of	RESTR
at the expense of	BEN
at a cost of	BEN
for (the) sake of	BEN, AIM
in exchange for	SUBS
in addition to	RESTR

with/in regard to	REG
at variance with	CRIT
at the hands of	ACT
for want of	CAUS
from want of	CAUS
in return for	SUBS
in relation to	ACMP, REG
with/in reference to	ACMP, REG
with the exception of	RESTR
in the aftermath of	CAUS

This list has been primarily compiled on the basis of A Comprehensive Grammar of the English Language (Quirk, Greenbaum, Leech and Svartvik, 2004). Besides, we use the lexicon of complex prepositions by Klégr, 2002: English Complex Prepositions of the Type *in spite of*and Analogous Sequences - a Study & Dictionary. Yet not all preposition-noun-preposition sequences listed by Klégr are automatically assigned the functor SM in all contexts (as already mentioned with multi-word predicates). The following example shows a sentence where the preposition-noun-preposition sequence is semantically related to the event presented by the verb as well as to the subject of the sentence. For more details on ambiguous dependencies see Section 4.1.3, “Ambiguous dependency”.

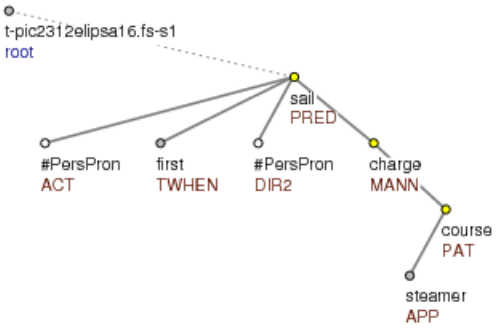
... he last sailed upon it in charge of .MANN *a steamer's course*.PAT.Fig. 7.8

Figure 7.7. Subjunction modifier



Her son was happily playing at the family home in contradiction of a prognosis 12 years ago that he was not 'worth treating'.

Figure 7.8. Subjunction modifier



He last sailed upon it in charge of a steamer's course.

7.2.1. Multi-word expressions acting as adverbs of extent

Some multi-word expressions similar to those marked with the SM functor (see above Section 7.2, “Multi-word expressions acting as prepositions or subordinators”) express the extent of the event/state/feature denoted by their governing node. It is typically structures listed below. They differ from the preposition-like expressions in that they do not determine the functor of the node they govern but they only modify the extent of the event/state/feature. These expressions are not to be assigned the functor `SM` but the `LOC` and `DIR3`, possibly `MANN`). Their daughter nodes get the functor `APP`.

Examples of multi-word expressions with the meaning of extent:

- *beyond the shadow of*
- *to the extent of*
- *to the brink of*
- *to the point of*
- *to/on the verge of*
- *within an ace/a whisker/a hair's breadth/inch of*

The university is staggering on the verge **.** `LOC` *of collapse*.`APP`

She was on the verge **.** `LOC[is_state=1]` *of crying*.`APP`

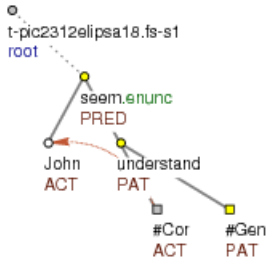
Chapter 8. English-annotation specific phenomena

8.1. Infinitive clause

8.1.1. Infinitive-nominative constructions/Raised subject *(John seems to understand)*

Constructions with copula-like verbs like *to seem, to appear, to prove, to turn out* etc. have different valency frames for the prop-*it* construction (seeSection 8.3, “The personal pronoun *it* ”) and the raised-subject constrution respectively. The reason is that although *John seems to be tired* can be paraphrased with *It seems that John is tired*, it is also possible to say *John seems tired*. The adjective as well as the infinitive clause are regarded as surface variants of the subject complement (the nominal part of a complex predicate). See more in section Section 8.1.8, “Verb-controlled infinitive *(His remarks seem to irritate her, He refused to cooperate)*”. Fig. Fig. 8.1 shows a sentence with raised subject in a copula-like verb.

Figure 8.1. Raised subject



John seems to understand.

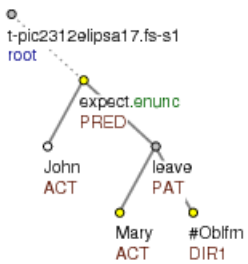
8.1.2. Infinitive-accusative constructions/exceptional case marking *(John expects Mary to leave)*

Infinitive-accusative constructions do not affect the valency frame of the governing verb, either. In a sentence like

John.ACT expects Mary to leave.PAT (Fig. 8.2).

the valency frame of *to expect* will not be affected, i.e. *Mary* will not be part of its frame. Instead, *the "accusative" noun Mary* is regarded as the Actor of the governed infinitive *to leave* while the governed infinitive *to leave* is an inner participant (PAT) of the governing verb *to expect*.

Figure 8.2. Exceptional case marking



John expects Mary to leave.

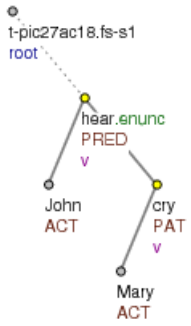
8.1.3. Infinitive-accusative constructions/exceptional case marking with verbs of perception *(John hears Mary cry)*

Infinitives governed by a verb of perception (*to see, to hear*) are also regarded as infinitive-accusative constructions. In the sentence

John hears Mary cry.(Fig. 8.3)

it is *John* and *cry* that are governed by the predicate *to hear*.

Figure 8.3. Exceptional case marking with verbs of perception



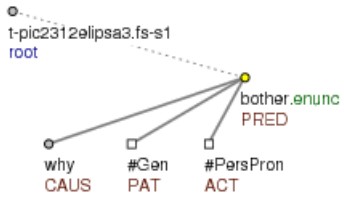
John hears Mary cry.

8.1.4. Infinitive as governing predicate *(How to raise money)*

When the infinitive is the governing verb in the sentence it gets the functor `PRED`.

Why bother.`PRED`? (Fig. 8.4)

Figure 8.4. Infinitive as governing predicate



Why bother?

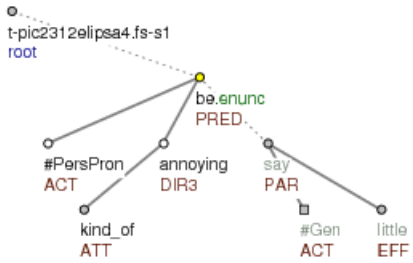
8.1.5. Absolute purpose clause *(To get back to what I was saying, bla bla...)*

Absolute purpose clause is an infinitive clause that clearly expresses purpose but is not governed by any superordinate clause. The governing verb of the absolute purpose clause will get the functor PAR. The verb with the functor `PAR` as well as all its dependent nodes get the `1` value of the `is_parenthesis` attribute attached to their functors. When the absolute purpose clause is attached to another clause by a comma or a colon, its governing predicate will have the functor `PAR` (and the `1` value of the `is_parenthesis` attribute and it will depend on the predicate (`PRED`) of the other clause.

But to get`.PAR [is_parenthesis=1]` *back to what I was saying: we must do*`.PRED` *it right away.*

It's been kind of annoying, to say`.PAR[is_parenthesis=1]` *the least.*(Fig. 8.5)

Figure 8.5. Absolute purpose clause



It's been kind of annoying, to say the least.

8.1.6. Infinitive in parenthesis

Infinitives can occur as cases of lexicalized parenthesis:

To think`.ATT[is_parenthesis=1]` *we might have been in Rome now!* (Fig. 8.6)

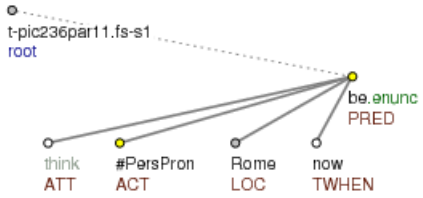
Having lost their valency, these parenthetical infinitives act as modal particles rather than clauses. They will be assigned the functor `ATT` and their `is_parenthesis` attribute will have the value `1`.

This type of parenthesis is expected to be relatively rare. An infinitive clause introducing a main clause is most likely to be a purpose clause with the underlying subject coreferring with the agent of the governing clause, although it will often rather help the main clause refer back to the preceding text:

To conclude`.AIM[is_parenthesis=0]`, *let us enumerate the most important points.*

To illustrate`.AIM[is_parenthesis=0]` *the latter approach we can imagine a black chamber...*

Figure 8.6. Infinitive in parenthesis



To think we might have been in Rome now!

8.1.7. Infinitive as inner participant *(To leave early would be impolite)*

Infinitives can also occur as inner participants in a verbal frame, typically as `ACT` or `PAT`:

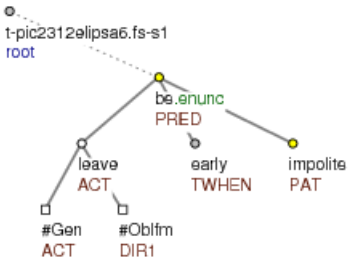
To leave`.ACT` *early would be impolite*.(Fig. 8.7)

The order was for everybody to stay indoors`.ACT`

He refused to cooperate`.PAT` (Fig. 8.8)

As follows from the last example sentence, an infinitive clause typically gets the functor `PAT` in cases of verb control.

Figure 8.7. Infinitive as Actor



To leave early would be impolite.

8.1.8. Verb-controlled infinitive *(His remarks seem to irritate her, He refused to cooperate)*

Some verbs have the ability to control other verbs. The constructions with control are represented in the frames of the valency lexicon. This section lists a few types of constructions with infinitives which are to be represented as governed by a control verb. The governed infinitive is usuallly assigned the functor `PAT`. (In transitive verbs it is naturally `EFF`)

1. a lexical verb listed as control verb + infinitive

   He refused.`PRED` *to cooperate*.`PAT`.(Fig. 8.8)

   I was invited.`PRED` *to contribute*.`PAT` *to the volume.* (Fig. 8.9)

2. infinitive-nominative construction (actually a subset of 1)

   John seems.`PRED` *to understand*.`PAT` (Fig. 8.1)

3. *to be* + (active or passive) infinitive with *to* meaning obligation

   I am.`PRED` *to blame*.`PAT`. (Fig. 8.11)

   <It> is.`PRED` *to be stressed*.`PAT` *that Peter was living in London in 1976.* (Fig. 8.12)

4. infinitive in causative constructions

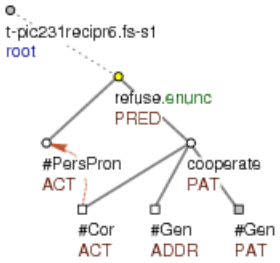
   Also verbs *to have, to make, to get* govern infinitives (with or without *to*) in causative constructions like

   They made.`PRED` *him*.`PAT` *stay*.`EFF`. (Fig. 8.13)

   They got.`PRED` *him*.`PAT` *to change*.`EFF` *his mind*.

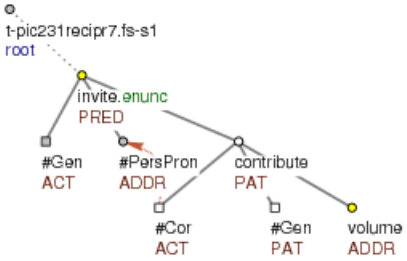
   We will have.`PRED` *someone*.`PAT` *clean*.`EFF` *the windows.*

Figure 8.8. Verb-controlled infinitive



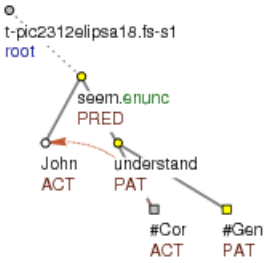
He refused to cooperate.

Figure 8.9. Verb-controlled infinitive



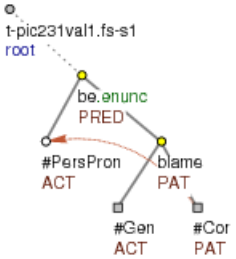
I was invited to contribute to the volume.

Figure 8.10. Verb-controlled infinitive - raised subject



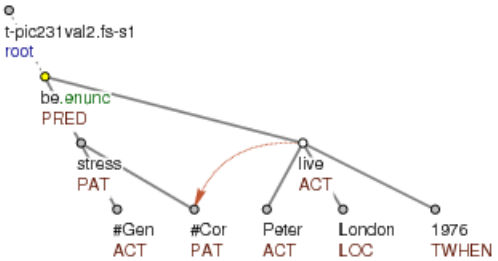
John seems to understand.

Figure 8.11. Verb-controlled infinitive - be+obligation



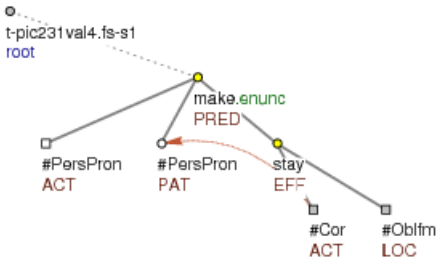
I am to blame.

Figure 8.12. Verb-controlled infinitive - be+obligation



It is to be stressed that Peter was living in London in 1976.

Figure 8.13. Verb-controlled infinitive - causative constructions with make, have, get



They made him stay.

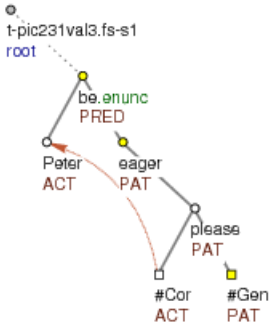
8.1.9. Infinitive controlled by a predicate adjective *(John is eager to please)*

An infinitive can be governed by the adjectival part of the complex predicate. In a sentence like

John is eager to please.(Fig. 8.14)

the infinitive *to please* is governed by the adjective *eager* as `PAT`. The infinitive *to please* will be assigned the proper valency frame from the valency lexicon. The `ACT` of *to please* will have the t-lemma substitute `#Cor` and will co-refer with *John.* This case is to be held apart from cases of *Object-Subject Transformation*, which will be described in the following section. Also, multi-word predicates with adjectives denoting a property that can both be associated with the subject of the governed infinitive clause and with the entire event denoted by the infinitive, are to be treated in a different way (see Section 8.1.14, “Infinitive governed by the copula verb with certain adjectives *(She was lucky to get that job, She was kind to let us come)*”).

Figure 8.14. Infinitive controlled by a predicate adjective



John is eager to please.

8.1.10. Infinitive in *(John is difficult to please)*

Similar in form but different in meaning are sentences with objects transformed into subjects in sentences with complex predicates formed by a copula verb and an evaluative adjective, such as in the following sentence:

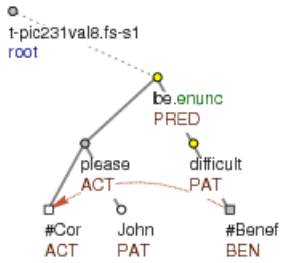
John is difficult to please.(Fig. 8.15)

These structures can be paraphrased as

It is difficult to please John.

The object-subject-transformed and the non-transformed sentences will have identical tectogrammatical representations:

Figure 8.15. Tough-movement



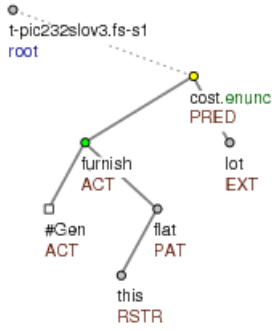
John is difficult to please/It is difficult to please John.

In the structures with objects transformed into subjects the predicate will have the value `osub3` of the attribute `osub` (*object-subject transposition*, see Section 8.8, “Object-Subject Transposition *(Crystal breaks easily., John is difficult to please., This flat must have cost a lot to furnish.)*”). The adjective governs a generated Benefactor (`BEN`) node in both cases and the `ACT` of *to please* will have the t-lemma substitute `#Cor;`

8.1.11. Infinitive of a lexical verb in object-subject transposition *(This flat must have cost a lot to furnish)*

Object-subject transformations occur also with lexical verbs. The tectogrammatical representation of such structures is not identical with Section 8.1.10, “Infinitive in *(John is difficult to please)*” in that there is no artificial Benefactor added. (Fig. 8.16) The governing predicate will obtain the value `osub4` in the attribute `objectsubject`.

Figure 8.16. Object-subject transposition with lexical verbs



This flat must have cost a lot to furnish.

8.1.12. Infinitive governed by a "manner-adverbial" adjective *(She was quick to shut the door, Bob is slow to react)*

In sentences like

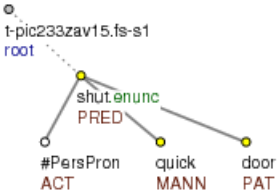
She was quick to shut the door.(Fig. 8.17)

the adjectival part of the complex predicate expresses manner in which the event denoted by the governed infinitive takes place. Such a sentence can be paraphrased as

She quickly shut the door.

The tectogrammatical representation of both variants is identical. In the structure comprising the infinitive governed by the adjective the predicate *to be* will be deleted and its `ACT` becomes the `ACT` of the infinitive. The adjective will also be governed by the infinitive and will be assigned the functor `MANN`.

Figure 8.17. Infinitive governed by a "manner-adverbial" adjective

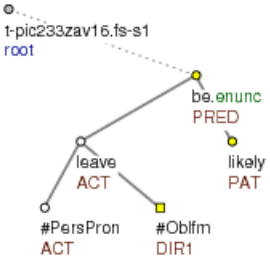


She was quick to shut the door.

8.1.13. Infinitive governed by *likely (She was likely to leave)*

The structure *X is likely to y* is to be treated as *That X will do y is likely.* The adjective *likely* is governed by the predicate *to be* and is assigned the functor `PAT`. The infinitive clause will be governed by the predicate *to be* as `ACT`. (Fig. 8.18)

Figure 8.18. Infinitive governed by likely



She was likely to leave.

8.1.14. Infinitive governed by the copula verb with certain adjectives *(She was lucky to get that job, She was kind to let us come)*

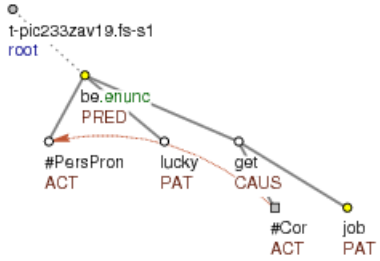
This type of structure is very close to the one decribed in Section 8.1.9, “Infinitive controlled by a predicate adjective *(John is eager to please)*”. In the "eager-structure" the infinitive clause is considered to be governed by the adjective. Assumingly, there is a list of adjectives that have infinitives in their valency frames and the combination of the predicate adjective and the infinitive cannot be paraphrased with a prepositional phrase (**John is eager for pleasing somebody*, but *John is sorry because he heard that*). Such infinitives get the functor `PAT`. On the other hand, we believe that there are adjectives whose semantic nature lies somewhere between "eager-adjectives", which in complex predicates govern the infinitives as their Patients, and "quick-adjectives" (see Section 8.1.12, “Infinitive governed by a "manner-adverbial" adjective *(She was quick to shut the door, Bob is slow to react)* ”) in complex predicates, which we represent in the same way as manner adverbials. The "lucky-adjectives" differ in that a complex predicate with a "lucky-adjective" governing an infinitive clause can be paraphrased as "It was x that y happened to Z" or "It was x of Z that Z did y", or even "Z was x because Z did y".

She was lucky to get that job. (= It was a luck for her that she got that job. She was lucky because she got that job.)

She was kind to let us come. (=It was kind of her that she let us come. She was kind because she let us come.)

"lucky"-adjectives do not govern the infinitive clause. Instead, the infinitive clause is governed by the main predicate *to be* as `CAUS`.

Figure 8.19. Infinitive governed by



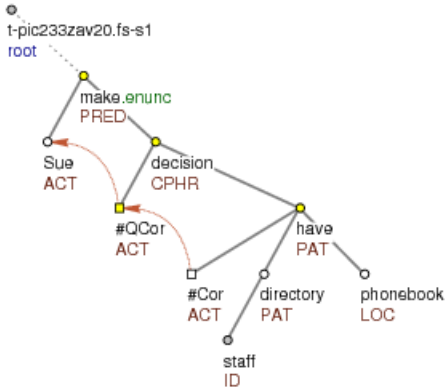
She was lucky to get that job.

See Section 8.2.4, “Gerund (with a preposition) governed by the copula verb with certain adjectives *(You were fortunate in winning his support)*” with the same resolution for gerunds.

8.1.15. Infinitive in support verb constructions/ (CPHR) and in idioms (DPHR) *(Peter made the decision to leave. Peter did not have the nerve to leave)*

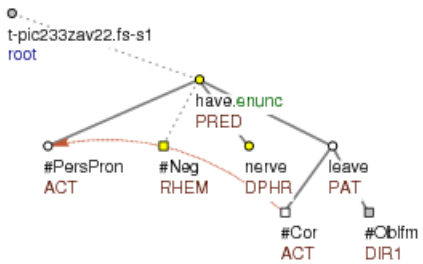
Even support verb constructions can govern an infinitive. Then the infinitive is governed by the predicate noun (the nominal component of the predicate, assigned the functor CPHR, see Fig. 8.20). Dependent parts of verbal idioms marked with the functor DPHR, on the other hand, cannot govern any node themselves. Therefore, they get the infinitive as sister, assigned the "lowest" possible inner-participant functor (see Fig. 8.21).

Figure 8.20. Infinitive governed by CPHR



Sue made the decision to have a staff-directory in the phonebook.

Figure 8.21. Infinitive governed by DPHR

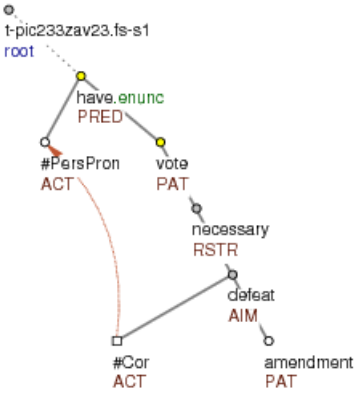


She did not have the nerve to leave.

8.1.16. Infinitive governed by a postponed attribute *(They had the votes necessary to defeat the amendment)*

An infinitive governed by a postponed adjectival attribute will preliminarily get the functor `PAT`. The `ACT` of the infinitive (except the passive infinitive) will be controlled by the governing predicate (Fig. 8.22).

Figure 8.22. Infinitive governed by a postponed attribute



They had the votes necessary to defeat the amendment.

8.1.17. Purpose clause (`AIM`) *(He stopped to have a drink)*

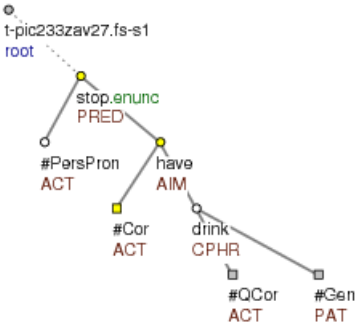
The infinitive clause is often used in purpose clauses:

He stopped to have.`AIM` *a drink* (Fig. 8.23).

Purpose clauses can be paraphrased by the subordinating conjunction *in order to*:

He stopped in order to have a drink.

Figure 8.23. Purpose clause



He stopped to have a drink.

8.1.18. "False" purpose clause (`REG`) *(The firms move faster to create new adds than in 1987)*

Sometimes the sentence is unlikely to express purpose although its surface syntax suggests it. In the following case the infinitive clause will get the functor `REG`:

The brokerage firms are moving faster to create.`REG` *new adds than they did in the fall of 1987*.(Fig. 8.24)

The reason is that the sentence can be interpreted as

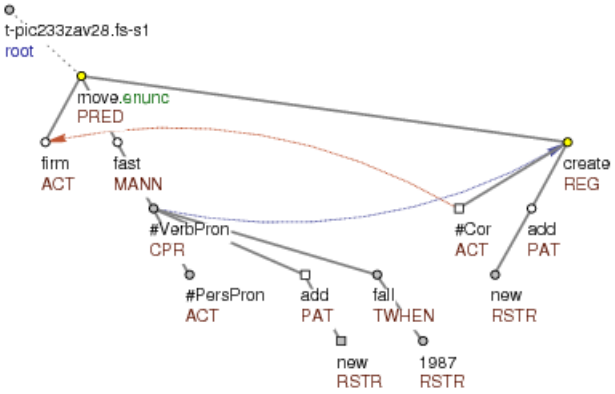
The brokerage firms are moving faster as in creating new adds than they did in the fall of 1987.

rather than

*?*The brokerage firms are moving faster in order to create new adds than they did in the fall of 1987.*

In fact, it is the creation of new adds itself, not any physical moving to anywhere, that is the event described here, and is proceeding faster compared to fall 1987.

Figure 8.24. False purpose clause (REG)



The firms are moving faster to create new adds than they did in the fall of 1987.

8.1.19. Intention clause (`INTT`) *(She has gone to see her sister)*

Like the `REG` infinitive, intention can be told from purpose by the "in-order-to"-paraphrase test. It occurs mainly with verbs of motion when the spatial adverbial (functors `DIR1, DIR2, DIR3` and `LOC`), the obligatory free modification for their valency frame, is not present on the surface.

She has gone.`PRED` *to see*.`INTT` *her sister*. (Fig. 8.24)

She has sent.`PRED` *me to collect*.`INTT` *the luggage*.

I've just been.`PRED` *to look*.`INTT` *at the puppies.*

He left.`PRED` *me to contemplate*.`INTT` *my predicament*.

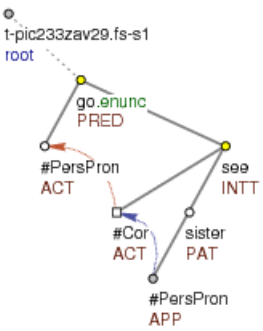
When the relevant spatial adverbial is present, the infinitive clause will be assigned the functor `AIM`.

Sometimes verbs of directional motion (e.g. *to go*) occur with an infinitive clause and an adverbial, which is expressed as location:

We go to hear.`INTT` *Archibald Menyies preach at Dull*`.LOC`.

Then the infinitive clause is also marked as `INTT` and not as `AIM`.

Figure 8.25. Intention clause (INTT)



She has gone to see her sister.

8.1.20. Consecutive clause (RESL) *(I'm too young to marry, Will you be so kind as to call me a taxi?)*

An infinitive clause typically gets the functor RESL when an inner participant of the main predicate is modified with the quantifiers *too, so (...that)* or *enough* etc. and it describes the extent to which the event denoted by the main predicate is real, related to the circumstance rendered by the infinitive clause. (See also Section 4.6.1.3.3, “Constructions with consecutive clauses”) The infinitive clause expresses the expected extent/intensity of the event or its generally accepted character. E.g. in the sentence

My hands are so numb with cold that I can't stretch them.

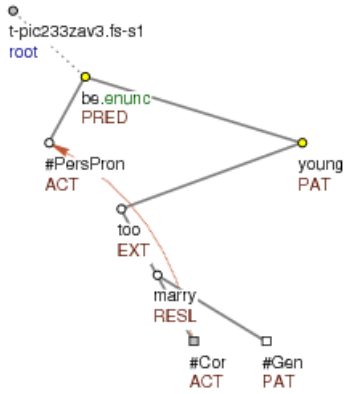
the actual extent/intensity of the event of having hands numb with cold is in balance with the expectation of unability to stretch the hands. In a sentence like

I'm too young to marry.

the actual event of being young has higher intensity than the intuitively felt norm of the right age for getting married. Consecutive clauses of this type are to be held apart from purpose clauses (AIM). Infinitive clauses with the functor RESL are governed by the given quantifier while infinitive clauses with the functor AIM are governed by the main predicate.

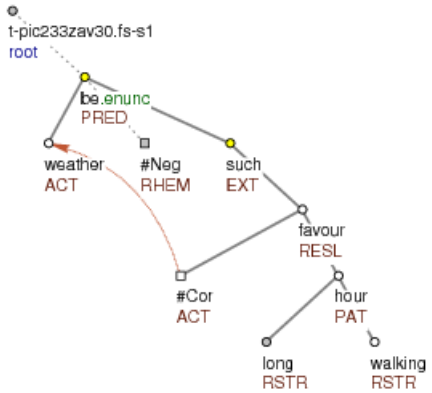
When the word expressing extent is such in the position of a predicate adjective, it does not get the functor EXT but PAT.

Figure 8.26. Consecutive clause (RESL).



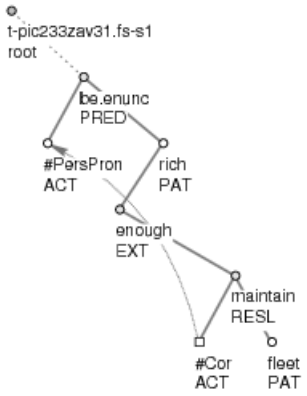
I'm too young to marry.

Figure 8.27. Consecutive clause (RESL) - with such as predicate adjective



The weather was not such as to favour long walking hours.

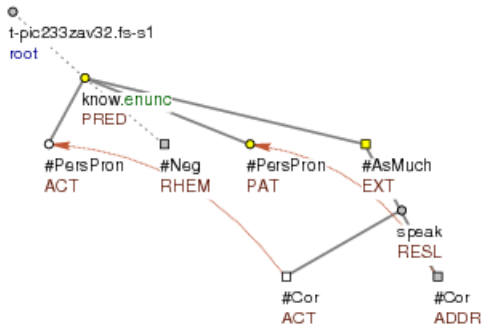
Figure 8.28. Consecutive clause (RESL)



He was rich enough to maintain a fleet.

Sometimes the quantifier is not present on the surface but the sentence clearly has the consecutive meaning. Then the infinitive clause is governed by an artificial node with the t-lemma substitute `#AsMuch` (see Section 4.6.1.3.3, “Constructions with consecutive clauses”):

Figure 8.29. I don't know her to speak to.



I don't know her to speak to.

8.1.21. Other free modifications rendered by infinitive

The English infinitive clause can have many different readings depending on the context as well as on the common knowledge. Using infinitive clauses allows for significant underspecification in an utterance, which might not be transferrable to other languages. E.g. a Czech translation would typically use subordinate clauses with desambiguating subordinators in place of infinitive clauses. Therefore we seek to specify the meaning of the infinitive clause as much as we can, with the annotators individually exploiting the context and the common knowledge.

Examples:

I shudder to think.`THO` *about it.* (=whenever)

I couldn't do it to save.`CNCS` *my life*. (I couldn't have done it even if I had been able to save my life by that)

He raised his hand as to beckon.`CPR` *me*. (= as though he beckoned me)

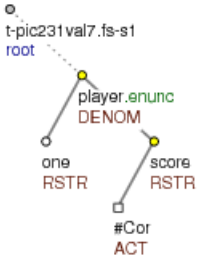
8.1.22. Attributive infinitive (`RSTR`)

Attributive infinitives usually follow the noun or pronoun it modifies. It can often be paraphrased by a relative clause:

There was no one to consult (= There was no one whom one might consult)

Examples:

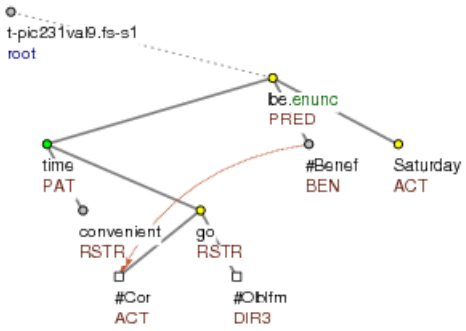
Figure 8.30. Attributive infinitive (RSTR)



The first player to score

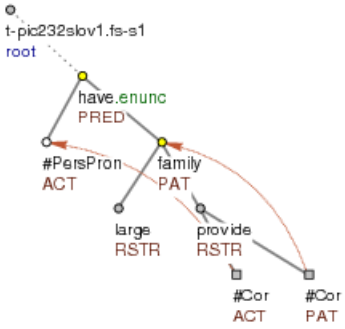
Sometimes it is very close to a purpose clause, depending on which paraphrase is selected. The interpretation is up to the annotator:

Figure 8.31. Attributive infinitive (RSTR)



The most convenient time to go is Saturday.

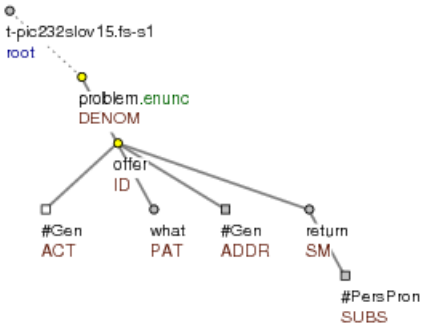
Figure 8.32. Attributive infinitive (RSTR)



He has a large family to provide for.

When the attributive infinitive is preceded by the preposition *of* it gets the functor `ID` since it is then classified as a combination of an identifying expression with a descriptor (see b):

Figure 8.33. The border between attributive infinitive and an identifying expression



the problem of what to offer in return

NB: the attributive infinitive diffuses into the governed predicate of control. In this case it would have the functor `PAT`. This happens with nominalizations:

his ability to solve.`PAT` *the task*

Some constructions are ambiguous and the decision is up to the annotator:

She wants someone to look.`RSTR` *after her.* (= She wants a person that would look after her. The infinitive is governed by the pronoun *someone*.)

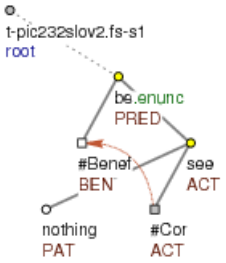
or

She wants someone.`ACT` *to look*.`PAT` *after her* (= She wants that a person would look after her. The infinitive is governed by the predicate *to want* and the pronoun *someone* has become the Actor of *to look*.)

8.1.23. Infinitive of perception verbs in existential constructions

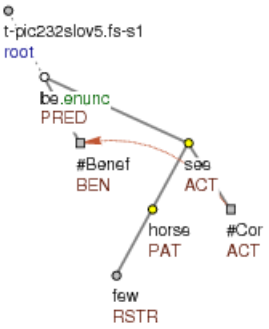
This type of structure has en resolved analogically to the equivalent construction in Czech (the variant *Je vidět Sněžku*). It can be subject to changes if met repeatedly in the data. These constructions occur with active as well as with passive infinitives of perception verbs. Both are resolved in the same way.

Figure 8.34. Infinitive of perception verbs in existential constructions



There is nothing to be seen.

Figure 8.35. Infinitive of perception verbs in existential constructions



There are so few horses to see.

8.2. Gerundial clause

This section deals with gerunds as *-ing* forms that have partially retained verbal features (unlike action nominals)). The Penn Treebank tags them as `VBG` and discriminates them from nouns (`NN`). In the tectogrammatical representation they have verbal t-lemmas. Besides following the original PTB tagging, the annotators can identify a gerund with the following criteria:

1. With transitive verbs it does not employ the preposition *of* to express object. In verbs that require prepositional objects (e.g. *to dispose of, to look forward to* etc.) the decision gerundial clause vs. action nominal is up to the annotator. Preferably they are captured as gerundial clauses.
2. It is not modified by adjectives.
3. It combines with adverbials.
4. It accepts auxiliaries.
5. It is (on the tectogrammatical layer) always preceded by subject in a possessive or common-case form. Gerunds with missing subject get the relevant inner participant (mostly ACT) completed with t-lemma substitutes `#Cor` or `#Gen`.

Typical examples:

His leaving no address was most inconvenient.

I count on his parents forbidding it.

You may depend on my not mentioning it.

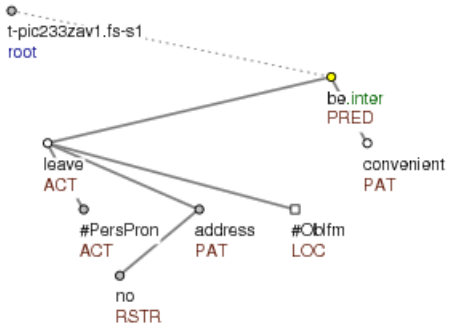
Mother always insisted on us having a party at Christmas.

The similar *-ing* form that is not to be mixed up with gerund is the action nominals. They are usually lemmatized as nouns, and they act as such. At the current annotation stage we do not even complete their valency frames with generated nodes. This section is exclusively dedicated to gerunds.

8.2.1. Gerund as inner participant

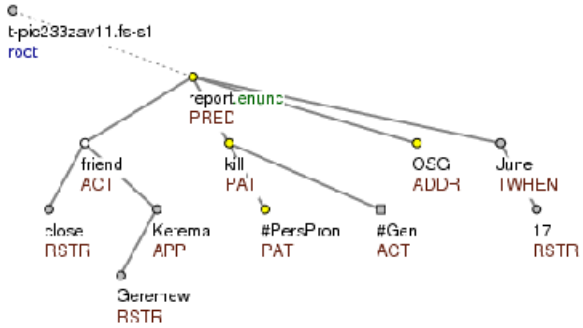
When a sentence contains an *-ing* form identified as gerund that is modified by a personal pronoun in the object case or a noun in its basic form (the so-called fused participle), it is always the gerund and not the noun/pronoun that is governed by the predicate. The noun/pronoun itself is governed by the gerund as its ACT (with active infinitives) or PAT (with passive infinitives). This applies both to gerunds acting as sentential subjects and sentential objects alike.

Figure 8.36. Gerundial clause as inner participant



His leaving no address was most inconvenient.

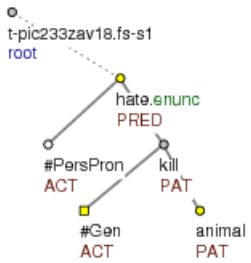
Figure 8.37. Gerundial clause as inner participant



A close friend of Geremew Ketema reported his being killed to OSG on 17 June.

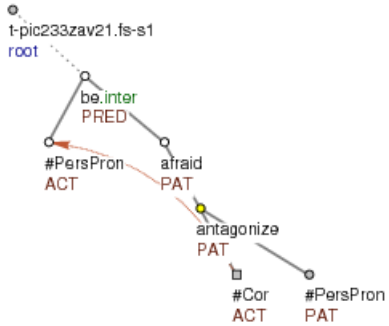
Even the verbal frame of a gerund, however, can have a generalized subject:

Figure 8.38. Gerundial clause as inner participant



I hate killing animals.

Figure 8.39. Gerundial clause as inner participant



I'm afraid of antagonizing him.

This missing participant is under no circumstances to be added into the frame of the finite verb but always into the frame of the gerund!

8.2.2. Gerund-accusative constructions/exceptional case marking with verbs of perception *(John hears Mary crying)*

The structure "Noun/Personal pronoun in subject case - Verb of perception - Noun/Personal pronoun in object case - gerund" is to be regarded as Exceptional Case Marking, analogically to the structure "Noun/Personal pronoun in subject case - Verb of perception - Noun/Personal pronoun in object case - Infinitive" (see Section 8.1.3, “Infinitive-accusative constructions/exceptional case marking with verbs of perception *(John hears Mary cry)*”), i.e. the Patient of the perception verb is the gerund and the apparent object of the sentence is its `ACT`.

8.2.3. Gerund with a preposition as nominal part of a complex predicate *(The pain was past bearing)*

Gerund is sometimes used with a preposition, governed by *to be* and several other verbs acting as copula. Together with the verb it forms a complex predicate expressing a state or a change of state. Then it is assigned the functor according to the meaning of the given preposition (typically a locative or a directional functor) and it acquires the attribute value `[is_state=1]`.

Examples:

The pain was almost past bearing.`LOC[is_state=1]`.

He is not above lending a hand.`LOC[is_state=1]`.

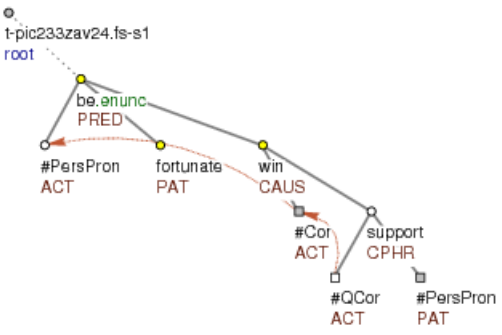
8.2.4. Gerund (with a preposition) governed by the copula verb with certain adjectives *(You were fortunate in winning his support)*

Similarly to how infinitive clauses are treated in Section 8.1.14, “Infinitive governed by the copula verb with certain adjectives *(She was lucky to get that job, She was kind to let us come)*”, a gerund with a preposition that occurs in a sentence with predicate consisting of a copula verb an an adjective is governed by the copula verb as `CAUS` when:

1. the adjective expresses the subject's characteristics in connection with a given event *(lucky, fortunate, kind, etc.)*
2. the sentence can be paraphrased as "It is x of Y that Y did z" or "It is x (for Y) that z happened to Y".

You were fortunate in winning.`CAUS` *his support.* (= You were fortunate to win his support. It was fortunate (for you) that you had won his support.)

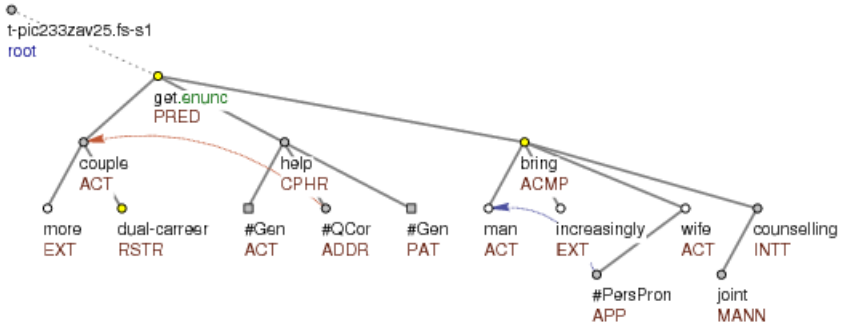
Figure 8.40. Gerund (with a preposition) governed by the copula verb



You were fortunate in winning his support.

8.2.5. Gerund as free modification - ACMP

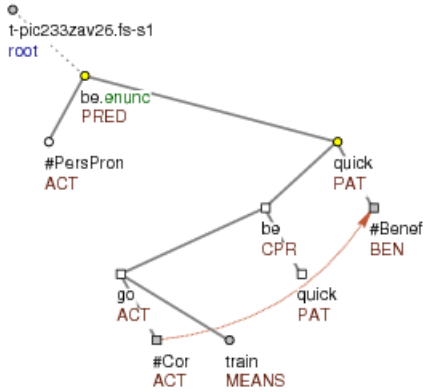
Figure 8.41. Gerund as free modification - ACMP



More dual-carreer couples are also getting help, with men increasingly bringing their wives for joint counselling.

8.2.6. Gerund as free modification - CPR

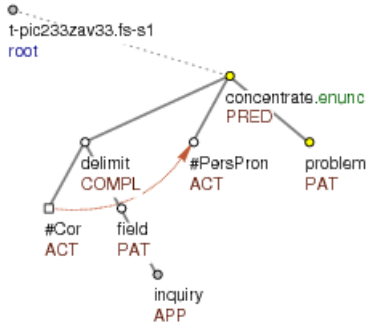
Figure 8.42. Gerund as free modification - CPR



It is quicker than going by train.

8.2.7. Gerund as free modification - COMPL

Figure 8.43. Gerund as free modification - COMPL



Having delimited the field of inquiry, we shall concentrate on the problem...

After perception verbs the gerund behaves like infinitive clauses with perception verbs. (see Section 8.1.3, “Infinitive-accusative constructions/exceptional case marking with verbs of perception *(John hears Mary cry)*”):

8.2.8. Gerund as free modification - COND

A gerund in a dependent clause, whose subject is not represented in the surface syntax and whose underlying subject does not agree with the subject of the governing clause is usually marked with the functor COND and it is regarded as syntactically incorporated parenthesis (see Section 4.5, “Parenthesis”), which implies that it has the attribute value [is_parenthesis = 1] :

Frankly speaking.COND*, she has chosen a wrong partner.* [is_parenthesis = 1] (Fig. 4.33)

8.2.9. Gerund as free modification - INTT

Like with infinitive intention clauses (see Section 8.1.19, “Intention clause (INTT) *(She has gone to see her sister)*”), gerund expressing intention occurs mainly with verbs of motion when the spatial adverbial (functors DIR1, DIR2, DIR3 and LOC) is not present on the surface. Unlike infinitive intention clauses, it does not matter for gerundial intention clauses whether the obligatory free modifications are present on the surface or not.

8.2.10. Gerund as preposition

Some gerunds have frozen to such an extent that they are regarded as prepositions and they have no tectogrammatical node, e.g. *excluding, including, excepting* (see Chapter 7, *Subordinators*).

8.2.11. Gerund in existential constructions *(There'll be singing...)*

Gerund in existential clauses denotes the actual event that is taking place. In fact, the predicate *to be* merely adds the tense and mode information and potentially negation. Negated sentences usually acquire the modal meaning of *impossibility* in addition (*There was no stopping him. = It was impossible to stop him.*).

Using existential clauses with gerunds allows for underspecification of inner participants, and it possibly affects topic-focus articulation. As the gerundial form in this case can have objects introduced by the preposition *of* in verbs that normally take direct objects and manner adverbials are expressed by adjectives (and not with adverbs) we regard it as a deverbal noun, which implies that it is lemmatized as noun and missing inner participants of the corresponding verb are not inserted. It gets, though, the whole range of "verbal" functors according to the meaning of the given modifications.

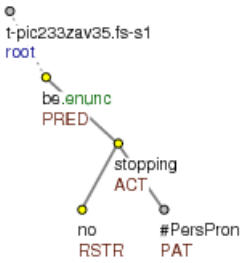
Seemingly, it could be easily paraphrased with an ordinary passive sentence. However, we have decided to keep the existential clause in the tectogrammatical representation anyway:

Examples:

Figure 8.44. Gerund in existential constructions

There will be singing.

Figure 8.45. Gerund in existential constructions



There was no stopping him.

The negative modal meaning is to be indicated in the grammateme `deontmod` (which acquires the value `[deontmod=poss]` and `[negation=neg1]`) in the main predicate. The gerund will get its relevant valency frame.

The reason for preserving the existential structure in the tectogrammatical representation is that these gerunds also occur in coordination with concrete nouns:

There will be dancing and refreshments.

Because of the concrete noun *refreshments* the existential structure would have to be artificially inserted anyway.

When the gerund governs a prepositional `ACT` (introduced by the preposition of) or a `REG` (introduced by the preposition *in*), it is also represented as the usual existential sentence.

There will be significant ageing of the population.`ACT`.

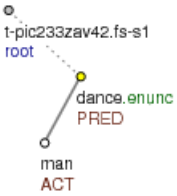
?There will be significant ageing in the population.`REG`.

Yet sentences where the `ACT` of the gerund is expressed by direct case, e.g.

There will be people dancing.

have the gerund as main predicate and *There* and *be* are associated to the gerund's node as auxrf's. The verb *to be* is regarded as the progressive tense auxiliary.

Figure 8.46. Gerund in existential constructions



There will be people dancing.

8.3. The personal pronoun *it*

The personal pronoun *it* has a variety of functions:

- anaphorical
- anticipatory
- deictic
- exclamative
- prop

The **anaphorical** *it* refers to a preceding noun denoting an inanimate entity or a not personalized animal, or it refers to a preceding clause or its part. The anaphorical *it* is assigned a functor according to its syntactic and semantic function in the clause.

I bought a new hat but my husband did not like it.`PAT`.

The **anticipatory** *it* can occur in subject as well as in object position:

<It> is no good bothering.ACT *about it. (=Bothering about it is no good.)* (Fig. 8.47)

<It> is feared that the ship was wrecked.PAT. *(=People feared that the ship was wrecked.*) (Fig. 8.48)

I.ACT *owe <it> to you that I have been able*.PAT *to get back to my studies. (=I owe getting back to my studies to you)*

I find <it> impossible to suggest.PAT *solutions in the abstract. (I find suggesting solutions in the abstract impossible.)*

He made <it> his business.EFF *to secure*.PAT *financial aid*. (Fig. 8.49)

We would greatly appreciate <it> if you would lend.PAT *us your support.*

The **deictic** *it* belongs to deictic personal pronouns in general. It is used refor deixis out of the language:

Is it.ACT *your suitcase*.PAT *(over there)?*

The deictic pronoun as well as the copula verb must be in morphological agreement with the entity it refers to:

Are they/these/those.ACT *your parents*.PAT?(Fig. 8.50)

The need of number agreement is typical of the deictic *it*. The **exclamative** *it* is also used in deictic contexts but it refers to a situation implicitly known in the discourse rather than immediately to the given entity:

(Knock knock knock...) "It' s me, open the door"! The exclamative *it* is to be annotated in the same way.

The **prop** *it* has little or no semantic content. It occurs in clauses which do not require any subject. It is typically clauses signifying time, atmospheric conditions and distance where the copula verb *to be* is regarded as (see Section 8.4.1, "Existential *to be (There is no smoke without fire.)* "):

<It> is not far.EXT *to New York*. (Fig. 8.51)

<It> is 5 o'clock.EXT

<It> is our wedding anniversary.ACT *next month.*(Fig. 8.52)

<It> is Sunday.ACT.

<It> is just one more stop.ACT *to Toronto.*

When the complementation is expressed by a non-prepositional case of a noun it gets the functor ACT.

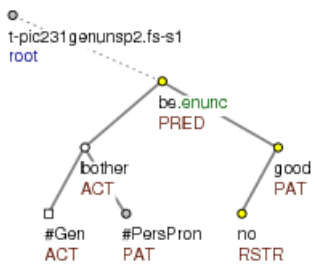
NB that in gerundial clauses *to be* is not regarded as existential but merely as auxiliary to the lexical verb in gerund. The verbs themselves have then empty frames:

<It is> raining. (empty frame)

<It is> freezing outside.LOC.(Fig. 8.53)

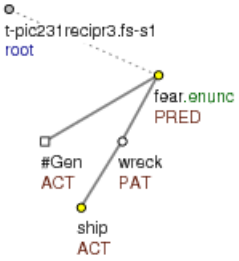
Both *it* and *to b*e are associated with the lexical verb as auxrfs.

Figure 8.47. Anticipatory it



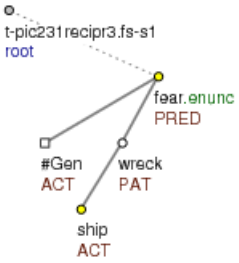
It's no good bothering about it.

Figure 8.48. Anticipatory it



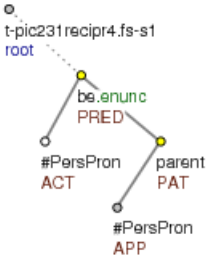
It is feared that the ship was wrecked.

Figure 8.49. Anticipatory it



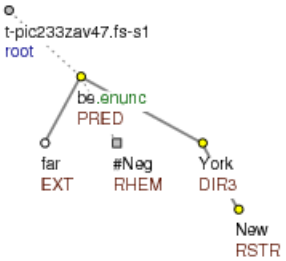
He made it his business to secure financial aid.

Figure 8.50. Deictic it



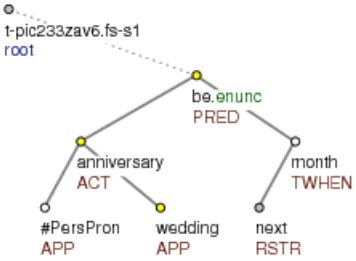
Are they your parents?

Figure 8.51. Prop it



It is not far to New York.

Figure 8.52. Prop it



It is our wedding anniversary next month.

Figure 8.53. Prop it

It's freezing outside.

8.4. Constructions with *to be*

8.4.1. Existential *to be (There is no smoke without fire.)*

Existential constructions comprise the existential *there*, the copula verb *to be* and an `ACT`.

There is.`PRED` *no smoke*.`ACT` *without fire*.`ACMP`.(Fig. 8.54)

The existential *there* is to associated to the verb *to be* as auxrf. Existential-locative constructions like

There is a girl in the waiting-room.
There was a farewell party on Friday.

are treated in the same way.

Figure 8.54. Existential to be

There's no smoke without fire.

8.4.2. Existential clauses with gerund or infinitive *(There will be singing., There's nowhere to go.)*

For resolution of the type

There will be singing.

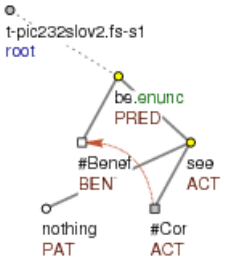
see Section 8.2.11, “Gerund in existential constructions *(There'll be singing...)*”.

Constructions with an infinitive, like

There is nowhere to go.

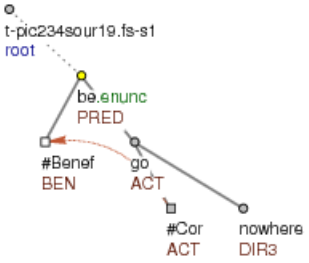
are resolved in the same way as existential constructions with perception verbs infinitives (see Section 8.1.23, “Infinitive of perception verbs in existential constructions”), i.e. there is a generated node with the t-lemma substitute `#Benef` with the functor `BEN` governed by the verb *to be*, which controls the inserted `ACT` of the verb in infinitive:

Figure 8.55. Infinitive of perception verbs in existential constructions



There is nothing to be seen.

Figure 8.56. Infinitive in existential constructions



There is nowhere to go.

The infinitive is the `ACT` of the verb *to be*.

8.4.3. The verb *to be* as copula verb (in multi-word predicates)

The nominal (or adjectival) component of the multi-word predicate usually gets the functor `PAT`.

The whole thing.`ACT` *is*.`PRED` *a fraud*.`PAT`.(Fig. 8.57)

The difference.`ACT` *is*.`PRED` *negligible*.`PAT`

The garden.`ACT` *is*.`PRED` *theirs*.`PAT`

On the other hand, components rendered by a prepositional group get the functor closest to their most literal meaning and they have a positive value in the attribute [is_state] (see also Section 6.13.1, “Attribute for the meaning of “state”” and Section 8.2.3, “Gerund with a preposition as nominal part of a complex predicate *(The pain was past bearing)*”):

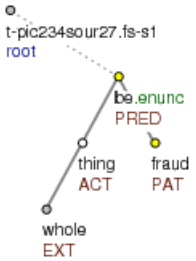
We.`ACT` are.`PRED` in trouble.`LOC[is_state=1]` now.`TWHEN`.

Several specific constructions are resolved as follows:

- *He is 6*.`RSTR` *years*.`EXT` *old*.`PAT` (*Old* is the adjectival component of the complex predicate and governs *years* as `EXT`, which in turn is modified by the number having the functor `RSTR`.)(Fig. 8.58)

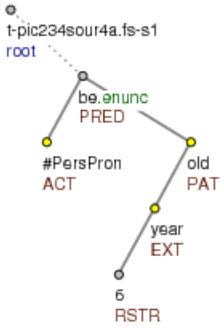
- *GM sank to $ 600 a share*. The governing node is the one denoting currency. The number gets the functor RSTR and *a share* (etc.) is its sister, who gets the functor REG.

Figure 8.57. To be as copula verb



The whole thing is a fraud.

Figure 8.58. To be as copula verb



He is 6 years old.

8.4.4. Special cases *(Who is to blame? It is to be stressed that...)*

To be is used in several other constructions which we rather regard as its valency frames to be captured by the valency lexicon (see e.g. Section 8.1.7, "Infinitive as inner participant *(To leave early would be impolite)*".)

8.4.5. Expressing personal experience with *to be* and an adverb/adjective with adverbial function *(He is fine.)*

Sentences consisting of a mostly human, or at least animate Actor, the copula verb to be and an adjectival/adverbial component of the complex predicate denote the way the Actor is feeling at the given moment. Therefore the adjectival as well as adverbial components will be marked as MANN, regardless the part of speech:

He.ACT *is*.PRED *cold*.MANN.

She.ACT *is*.PRED *well*.MANN.

8.4.6. Nominalizations ending with *-er* with adjectival modifications as nominal components of complex predicates *(She is a clever liar.)*

The *-er* nominalizations as nominal components of complex predicates in sentences with human subjects sometimes denote a feature of the sentence subject:

She is a liar. = She often lies.

An adjectival modification of the nominalization then expresses the manner or the extent to which the subject has the given feature:

She is a clever liar. = She lies in a clever way.

This type of adjectival modification gets the functor MANN.

For more details on nominalizations and on functors given to their modifications see Section 8.10, “Nominalizations”.

8.5. The existential *there*

The existential *there* (tagged as EX in the original Penn Treebank annotation) is associated as auxrf with the copula verb *to be*).

8.6. *One* as pro-form

Apart from being used as cardinal number, *one* can act as noun pro-form in noun phrases as well as in the nominal components of complex predicates. Besides, it has a generic meaning *(One has to care for his kids)*.

8.6.1. *One* as pro-form in the noun phrase *(The blue one suits her better.)*

One (*ones* in plural) can take adjectival modifications. This pro-form type must always have a determiner (an article or a demonstrative pronoun):

She wore a red dress but a blue one suits her better.

I like that one.

When coreference is being annotated, a coreferential arrow is supposed to point from *one/ones* towards its antecedent. When *one* has been omitted, the relevant antecedent node is to be copied in its place, as usual in ellipsis resolution:

She wore a red dress but a blue {dress}.ACT *suits her better*.(Fig. 8.59)

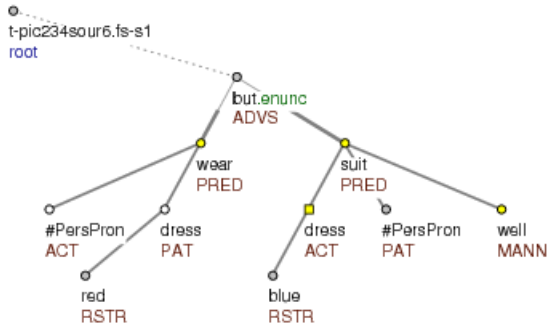
In uncount nouns *one* cannot be inserted (it is never regarded as 'missing') but they follow the general ellipsis rule anyway, i.e. the appropriate antecedent node is to be copied. In general, every adjective has to have a governing node except very few denominal adjectives which are regarded as nouns. It is typically labels for humans of certain origin *(Chinese)* or convinction *(Presbytarian)*. E.g. we do not regard e.g. *the hungry and poor* as nouns. When they do not have any antecedent in the surrounding text they have to be governed by a newly inserted node with the t-lemma #EmpNoun:

the {#EmpNoun}.DENOM *hungry*.RSTR[is_member=1] *and*.CONJ *poor*.RSTR[is_member=1]

the {#EmpNoun}.DENOM[is_member=1] *hungry*.RSTR *and*.CONJ *the* {#EmpNoun}.DENOM[is_member=1] *poor*.RSTR (Fig. 8.60)

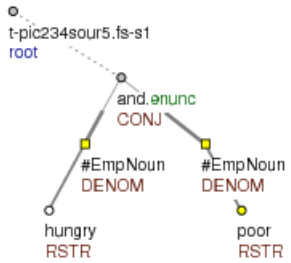
a move from the relatively simple{#EmpNoun}.DIR1 *to the relatively complex*{#EmpNoun}.DIR3

Figure 8.59. Contextual ellipsis of the governing noun



She wore a red dress but a blue suits her better.

Figure 8.60. Grammatical ellipsis of the governing noun



The hungry and the poor

8.6.2. *One* as pro-form in the nominal component of the multi-word predicate *(The new financing structure is a very highly leveraged one.)*

In multi-word predicates *one* gets the appropriate functor according to its function in the sentence.

Pemex managers.ACT *are the ones*.PAT *most thrilled*.RSTR *by the contract.*

The new financing structure.ACT *is still a very highly leveraged*.RSTR *one*.PAT

The problem is one.PAT *of great difficulty*.APP

8.6.3. Generic *One (One likes having his work appreciated.)*

The generic *one* as noun gets the appropriate functor according to its function in the sentence:

One.ACT *likes having*.PAT *his work appreciated.*

8.7. The pronouns *this/that*

The pronouns *this* and *that* have many different functions:

- deictic
- anaphorical
- adverbial (extent)

8.7.1. Deictic *this/that*

When *this/that* is used as deictic pronoun alone and it can be replaced with a personal pronoun it acts as a noun. In the coreference annotation a coreference arrow points from *this/that* towards its antecedent in the surrounding text. In connection with *one*, *this/that* is the daughter and gets the functor RSTR. In coreference annotation, the coreference arrow points from one towards its antecedent in the surrounding text.

That.ACT *was nice of you.*

This.ACT *is Mr. Jones.*

Come and have a look at this.PAT.

Those.ACT *who try hard*.RSTR *will succeed.*

When there is a nominal antecedent specified in the text, this/that acts as adjective and the antecedent nominal node must be copied as mother node of the demonstrative pronoun.

This chair is more comfortable than that.RSTR *{chair}.*

Take a pair of socks. These/those.RSTR *ones*.ACT *are nice.*

Take a pair of socks. These/those.RSTR *{socks}*.ACT *are nice.*

8.7.2. Anaphorical *this/that (This chair is more comfortable than that., That's where we live.)*

In the following case the demonstrative pronouns act as nouns:

The appearance.ACT *of brick and stone is better*.PAT *than {is}*.CPR *{good} that*.ACT *of other building materials.*

8.7.3. Anaphorical *that* in sentences of the type *that's +wh-word*

The construction *that is* in combination with a wh-word refers either to a noun phrase:

Hot soup! That's what I would like just now! (I would like a hot soup just now.)(Fig. 8.61)

or to a prepositional phrase:

There is more freedom in Canada and that is where I will go. (I will go to Canada)

or to an entire sentence:

She is poor. That's why she has to come to Arthur. (She has to come to Arthur because she is poor.)

in any case, the nodes *that* and *to be* of the construction *that is + wh-word* are to be attached to the wh-word as its auxrfs.

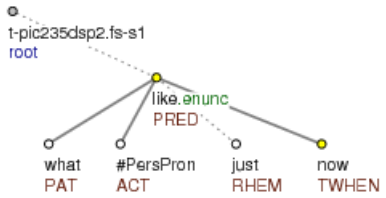
When the two clauses are joined in one sentence, they stand in coordination. The coordinating node is usually a comma or *and*. All wh-word combinations except *that's why* get the functor appropriate to their meaning in the sentence and the coordinating node gets the functor CONJ:

- that's where: LOC, DIR3
- that's when: TWHEN
- that's how: MANN
- that's what... like: MANN
- that's how long it takes: THL

That's why has the same meaning as the conjunctions therefore, thus and hence. In a complex sentence with *that's why* the second clause denotes a consequence of the event denoted by the first sentence. The node *why* gets the functor CM and the coordinating node gets the functor CSQ.

She.ACT *'s*.PRED[is_member=1] *poor*.PAT #Comma.CSQ *<that's> why*.CM *she*.ACT *has to come*.PRED[is_member=1] *to Arthur.*

Figure 8.61. that's + wh word



(Hot soup!) That's what I would like just now!

8.7.4. *This/that* as adverbials of extent *(Is it that far?)*

Especially spoken English employs *this* and *that* as adverbials denoting extent:

I didn't think it was this.EXT *late*.TWHEN

It can't be that.EXT *far*.LOC.

8.8. Object-Subject Transposition *(Crystal breaks easily., John is difficult to please., This flat must have cost a lot to furnish.)*

The annotation discriminates between three types of object-subject transposition:

1. middle alternation - manner:

   Crystal breaks easily. (Fig. 8.62)

2. middle alternation - condition

   Crystal breaks at the slightest touch. (Fig. 8.63)

3. object-subject transposition- "*tough*"-movement

   John is difficult to please.(Fig. 8.64)

4. object-subject transposition with a lexical verb - no Benefactor added!

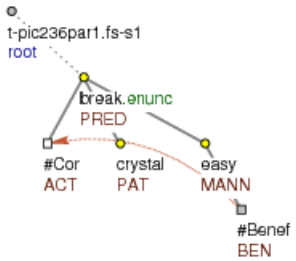
   This flat must have cost a lot to furnish. (Fig. 8.65. See also above, Section 8.1.11, “Infinitive of a lexical verb in object-subject transposition *(This flat must have cost a lot to furnish)*”)

Both middle alternations are associated to the transitive verb frame in the valency lexicon (Someone.ACT breaks crystal.PAT). The verbal attribute `objectsubject` will aquire different values according to the construction type as follows:

middle alternation - manner	osub1
middle alternation - condition	osub2
object-subject transposition - tough movement	osub3
object-subject transposition - with a lexical verb	osub4

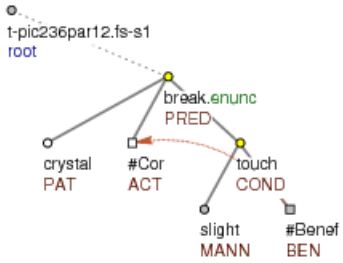
For the annotation resolution of the respective object-subject-transposition types see figures below.

Figure 8.62. osub1: Middle alternation - manner



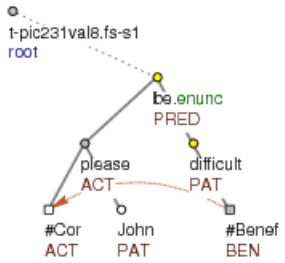
Crystal breaks easily.

Figure 8.63. osub2: Middle alternation - condition: Crystal breaks at the slightest touch.



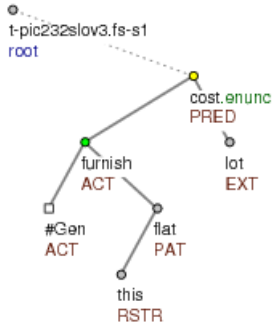
Crystal breaks at the slightest touch.

Figure 8.64. osub3: Object-subject transposition - tough movement



John is difficult to please.

Figure 8.65. osub4: Object-subject transposition with lexical verbs



This flat must have cost a lot to furnish.

8.9. Attributes (Modifiers)

This section is dedicated to adjectives and nouns as noun modifiers (attributes). They are each treated in a different way.

8.9.1. Adjectives as noun modifiers

Adjectives as noun modifiers get the functor RSTR (restrictive attribute) by default. Governing nouns recognized as nominalizations (see Section 8.10, “Nominalizations”) have all their adjective modifiers interpreted in detail whenever possible. E.g. the adjective *quick* in *a quick answer* would get the functor MANN exactly as *to answer quickly*.MANN. Some adjectives are interpreted in detail when occurring with any noun since their meaning inherently evokes assigning a certain functor. It is mostly adjectives denoting temporal relations, such as *frequent, rare, yearly* etc. A table of adjectives that always have the same functor is given below, including some warning examples of overinterpretation risk and brief explanation, exemplification when necessary/convenient for the Czech annotators. Sometimes, though, an adjective can have several readings. Typical collocates activating the given reading are listed in the second column.

Table 8.1. Adjectives as noun modifiers and their functors

next, last		RSTR	which in a row (ambiguous - space, time...)
old	bread	RSTR	
old	friend, scout	THL	
latest, recent, past		TWHEN	
former, current		TWHEN	
acting	chairman	TWHEN	currently occupying the post
constant		THL	
late		TWHEN	
long-term, temporary		TFHL	
instant, immediate		TWHEN	
frequent, occasional, rare		THO	
daily, monthly, yearly, weekly, quarterly		THO	
simultaneous		TPAR	
pharisaical, etc.		CPR	like Pharisees (hypocritical), like... etc.
complete, partial, entire, whole, half		EXT	
great, slight, significant		EXT	
perfect	stranger, nonsense	EXT	
utter	fool	EXT	
plain	nonsense, stupidity	EXT	absolute
plain	girl	RSTR	ugly
plain	design	RSTR	without decoration
plain old		RSTR	nothing special
plain speaking, thinking, face, truth		RSTR	honest, using simple, direct language, true
extreme	poverty, cold, right, extreme left of the picture	EXT	also: extreme right wing, extreme liberalism... whatever that can vary in intensity or whose intensity can be estimated (even in terms 'right or wrong', as with liberalism and right/left politics)
extreme cases	case, example, opinion	RSTR	no inherent intensity, it means rather 'unaceptable', 'unusual', 'unreasonable'.
virtual	a virtual monopoly, impossibility, they live like virtual prisoners	ATT	the speaker's estimation rather than objective extent
total, absolute		EXT	
firm	friend, commitment, believer	RSTR	
possible, impossible, sure, (un)certain, (un)likely, (im)probable, apparent, potential		MOD	
certain	person	RSTR	
so-called		ATT	
very	the very person I spoke to last night	RHEM	NB that rhematizers are left sisters of the highest node they have in their scope. Thus with the very person, very would be left sister of person
only	the only person I spoke to	RHEM	NB - rhematizer = left sister, see above

NB: Some nouns themselves bear temporal information of the type TWHEN, e.g.: *year, month, week, day*. Temporal adjectives like *past, future, recent* etc. do not answer the question "The days when?" "The past days" but rather "Which days?", the answer to which is best represented by the functor RSTR. (also: *a cold January*.RSTR *day*, though it can be paraphrased as *a day in January*. PDT has it as RSTR.) This applies only to TWHEN. E.g. the attribute *2-year period* would be asked about like: "how long period", not "which period", and therefore it will get the functor THL (unlike the Roman.RSTR period, of course). Long and short in periods and spans will get RSTR like last and next with year, week etc.

8.9.2. Participles as noun modifiers

The adjectival valency is not being focused at the current annotation stage. On the other hand, having a valency lexicon of verbs entitles us to associate valency frames to present as well as past participles in attributive positions, provided they have a verbal t-lemma. Valency frames are assigned to participle attributes in anteposition as well as in postposition. All their inner participants are inserted, **except the one that would corefer with the governing noun node**. E.g. in the word combination *a horrifying experience, horrifying* would acquire a newly generated PAT but not ACT as it is *experience* itself that horrifies someone.

When a participle obviously does not correspond to any verbal frame, it is to be noted. Assumingly, there is going to be a group of such participles, though many had already been filtered away by the original PTB annotation as nouns or adjectives.

The current PEDT annotation distinguishes restrictive and non-restrictive attributes to a limited extent. Non-restrictive attributes (i.e. also attributive sentences) are only considered in attributes in postposition, e.g.:

Mr. Vinken, 50 years old.DESCR*, was elected chairman.*

Mr. Vinken, who is.DESCR *50 years old, was elected chairman.*

Mr. Vinken, who is.DESCR *50, was elected chairman.*

8.9.3. Nouns as noun modifiers

Nouns as modifiers are interpreted according to their semantic relation towards the governing noun. The rule of thumb when assigning a noun modifier a functor is paraphrasing the noun phrase with a prepositional phrase. When any paraphrase is impossible and no other functor seems plausible, the default functor for a noun as modifier is APP (appurtenance) whenever one can paraphrase the semantic relation of X (the governing noun) to Y (the modifier) as: "X has Y". Discussions over concrete data after interannotator-agreement tests have resulted in a table of prototypical cases, which the annotators agreed to interpret with certain functors. We have attempted to group them. The groups have been build purely on the basis of analogy seen by intuition, and rather on the basis of family resemblances than one shared feature. The list is naturally incomplete and subject to constant changes. New data is to be annotated by analogy with this list, which in turn results in changes. The modifiers marked with bold font are the ones that determine the functor themselves, similarly to the adjective modifiers. In all other word combinations it is the governing noun that plays the more important role in assigning the functor of its noun modifier.

Table 8.2. Nouns as noun modifiers and their functors

yesterday, yesterday's	bread	TFRWH	
yesterday, yesterday's	meeting, earthquake	TWHEN	
yesterday, yesterday's	dancing partner	TWHEN	
every-year		THO	
one-centimeter	improvement	DIFF	
trade, consumption, computer, intellectual property, unfair-trade	law, tax, business, rights, investigations	REG	
shareholder	rights	ACT	
mortality, interest	rate, fraction, percentage	MAT	
Vermont, Boston, Boston's	University, Institute	APP	
board	meeting	ACT	
lobster, chocolate	consomme, mousse	RSTR	raw material + product
future		TWHEN	
red-carpet	treatment	RSTR	like CPR, but too implicit
trade	deficit	REG	
government	figures	APP	it is not sure that the government has made them but it has and uses them
trade, bankruptcy, commerce, licensing	Ministry, Institute, school, association, court, commission, group, company	REG	specializing in X, concerned with X, in charge of X's cases - but: deaf.APP organization because it is the deaf who organize themselves
"topic"	dispute, discussion	PAT	deverbal nouns
circulation, reorganization	plan	PAT	both X plans to do y and X plans the details/rates of y.
bankruptcy	proceedings	REG	but bankruptcy.PAT report - deverbal noun
electricity	demand	PAT	deverbal noun (homograph)
operating, trade	efficiency	REG	
trade	diplomacy	REG	
reorganization	process	RSTR	
outside	bidder	RSTR	not really a local meaning
US Treasury, XY-Company's	secretary, vicepresident	APP	
oil, power, chemicals	plant, factory, concern (company)	REG	producer of X - not deverbal noun
average	refund	RSTR	not EXT
construction	expense, cost (cost is not deverbal noun)	AIM	
year	start, end, begin, duration, course	APP	probably not to be regarded as deverbal nouns
year-to-year	increase	THO	
passenger	car	RSTR	product designed for x, to do x
control	device	RSTR	product designed for x, to do x
motor	vehicle	RSTR	product with its main feature
household	appliance	RSTR	product designed for x, to do x
desktop	computers	RSTR	product with its main feature
housekeeper	system for PCs	RSTR	
disk	drives	RSTR	

holding	company	RSTR	
watch, wish	list, survey, review, summary	RSTR	
stiff-trade	sanctions	CAUS	
fair-trade	award	CAUS	
video-viewing	parlors	AIM	
intellectual-property	case, instance, example, norm, pattern	RSTR	
videocassette	piracy	RSTR	
voting	stock	RSTR	
stock	holding	RSTR	a term
joint	venture	RSTR	a term
debt	load	RSTR	
operating	profit	RSTR	
consulting	carreer	REG	
composite	trading	RSTR	
high-yield, bank	debt	RSTR	all debts get RSTR
management	group	REG	like institute when management is focused. When group focused, then APP, when the number of people is focused, then MAT
faculty	member	APP	

8.10. Nominalizations

We define nominalization according to Quirk-Greenbaum-Leech-Svartvik: "A comprehensive Grammar of the English Language", 1985. Longman. 17.51, p. 1288: "A noun phrase [...] which has a systematic correspondence with a clause structure will be termed a NOMINALIZATION. The noun head of such a phrase is normally related morphologically to a verb or to an adjective".

Examples

his refusal to come ~ he refuses to come

the truth of her statement ~ her statement is true

her friendship for Chopin ~ she was a friend of Chopin

*the critics' hostile reception of the play ~ the critics received the play in a hostile manne*r

but also:

She is a good writer ~ She writes well.

He is a clever liar ~ He lies cleverly.

As suggested above, Quirk et al. do not confine nominalizations to deverbal event nouns. However, the deverbal nouns, as characterized below, make a separate group for which certain annotation rules apply. A survey of suffixes typical of (but not specific to) deverbal nouns together with their semantic description and a few examples is given below (based on Quirk et al., slightly modified).

Deverbal Event Nouns

- **-age:** action of, instance of

  breakage, leverage, drainage, wastage
- **-al:**the action or result of

  refusal, survival, dismissal, arrival
- **-ation:** the process or state of

  exploration, victimization, starvation, ratification
- **–ison, -ion**, when corresponding verbs are known

comparison, depletion, conclusion

- **-ing** not artifacts, just 'action nominals' (Prochazkova)
- **-ment**: the result of (also process – *treatment*)

  treatment, amazement, puzzlement (but also: *embodiment* when not meant as artifact)
- also event nouns that are verb homographs:

  care, discharge, lunch, breakfast (they have verbal matches. *To lunch, to breakfast*, marked as formal) but not e.g. *label* since *label* is a concrete noun.

Agentive Deverbal Nouns

-ant: agential inhabitant, informant, participant, lubricant

- **-ee:** one who is the object of the verb (or non-agentive Actor: *absentee*)

  appointee, employee, payee, refugee, absentee
- **-er, -or**

  writer, driver, employer, window-cleaner, high-flier, silencer accelerator, incubator, survivor but not *doctor, author*.... they have no corresponding verbal base

Annotation remarks: Verb nominalizations with the above-listed suffixes (nominalizations) will inherit their valency frames in the future (when possible). As NomBank has not yet been converted into EngValLex, missing participants are **not** manually completed at the current annotation stage, nor are the verbal valency frames assigned to nouns, **except** when a nominalization ending with *-ing* is a **gerundial clause,** e.g.

His.ACT *coming late* #Oblfm.DIR3 (reference to EngValLex: ev-w590f16).

We expect a number of nominalizations to require noun-specific frames, which will be different from the frames of the corresponding verbs.

Adjectival complementations will get the whole range of functors for free modifications. Participant functors are preferred to free modification functors with noun complementations, e.g. *Peter's*.ACT *hasty*.MANN *lab*.PAT *inspection*. *Peter's*.ACT *math*.PAT *studies*. Agentive verb nominalizations as listed above will mark their possible noun complementations with the same participant functor as the corresponding verb would do (it will be mainly PAT: *a novel*.PAT *writer*, but even CD.PAT *writer – to burn/write a CD = to record information onto a CD*).

8.11. t-lemmas of phrasal verbs

Phrasal verbs get the particle into their t-lemmas, e.g. the t-lemma of *take up* would be take_up. The particle is assigned to the verb as its auxrf. In numerous cases the particle has been tagged as RP in the original PTB annotation. In these cases the t-lemma as well as the auxrf-assignment should have proceeded automatically. However, the annotators are also allowed to change t-lemmas manually. Their decisions are mainly based on the Macmillan English Dictionary for Advanced Learners (Rundell et al. 2003).

8.12. Gradation of adjectives and adverbs

When adjectives or adverbs create comparative and superlative forms by means of *more* and *most*, *more* and *most* are associated to the given adjective or adverb as auxrf.

8.13. Noun determiners on t-layer

Only the articles *a(n)* and *the* are associated with their governing nouns as auxrfs (except when parts of identifying expressions). Others have their own nodes on the t-layer. Their t-lemmas will not be changed manually.

8.14. Assignment of verbal frames to past participles in attributive position *(an industrialized country)*

As past (as well as present) participles in attributive position are recognized as verbs whenever possible, and therefore they are assigned the corresponding verbal valency frame and their missing inner participants are completed, except the one that is rendered by the governing noun (see also Section 8.9.2, “Participles as noun modifiers”). This one will be inserted later automatically (it is merely a technical time-saving decision - the annotation ends up with complete frames). Past participles are yet ambiguous when the verb they are derived from has an intransitive as well as a transitive frame, e.g.:

in industrialized countries.

This construction can be assigned to either of the following frames:

1. *X industrializes a country*
2. *a country industrializes*

Whenever possible, the transitive frame assignment is favored.

8.15. Cleft sentences

There are two types of cleft sentences:

1. the true cleft sentences in

   It was John (that) broke the window.

   It was with a hammer (that) John broke the window.

   It was a window (that) John broke.

   It was today (that) John broke the window.

   Was it a window (that) John broke?

   etc.
2. pseudo-cleft sentences.

   What John did next was break the window.

   What John needs is a long rest.

   The person John needs is a doctor.

   The thing John needs is a long rest.

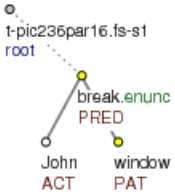
   etc.

The former type gets a special treatment in the annotation while the latter not (similarly to the annotation conventions followed in the Penn Treebank). The true cleft sentence is regarded as a topic-focus-articulation means, affecting only the TFA, which is to be reflected by the TFA annotation. Therefore the tree structure of the tectogrammatical representation of a cleft sentence is very identical with the tectogrammatical representation of an ordinary sentence with the same meaning. The tree diagrams of cleft sentences may differ from the ordinary sentence diagram in the tectogrammatical word order, according to which sentence element has been clefted. The anticipatory *it, be* and the (optionally present) connective *that* are attached to the clefted element as its auxrfs ; e.g.:

John.ACT *broke the window*.PAT/*<It was> John that*.ACT *broke the window*.PAT.

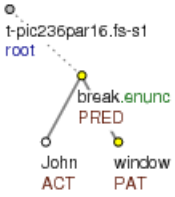
<It was a> window <that>.PAT *John*.ACT *broke*.

Figure 8.66. True cleft sentence/ordinary sentence



John broke a window. / It was John that broke the window.

Figure 8.67. True cleft sentence



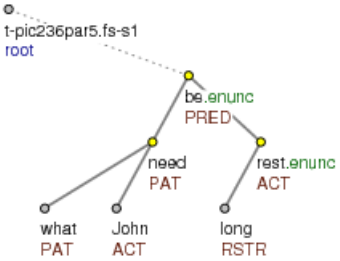
It was a window John broke.

The pseudo-cleft sentences are annotated according to their form, i.e. the wh word (or the word thing/person)introducing the relative clause is regarded as the non-verbal part of a copula-predicate, which acquires the functor `PAT`. The governing node of the subtree governed by the main predicate represented by *be* gets the functor `ACT`.

The thing John needs is a long rest.

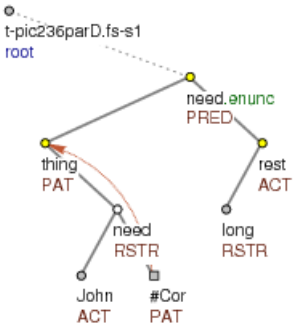
What John needs is a long rest.

Figure 8.68. Pseudo-cleft sentence



What John needs is a long rest.

Figure 8.69. Pseudo-cleft sentence



The thing John needs is a long rest.

8.16. Reflexivity

Reflexivity is regarded as a special case of transitivity, except a few cases of truly reflexive verbs (reflexivum tantum). These verbs, e.g. *pride oneself, absent oneself* etc., have their own valency frames and the reflexive pronoun is part of their lemma. In the data, the reflexive pronoun does not have its own tectogrammatical node but it is attached as auxrf to the node for verb. This resolution is similar to how particles of phrasal verbs are treated in the lexicon.

In all other cases, the reflexive pronoun gets the t-lemma substitute `#PersPron`. A coreference arrow points to its antecedent. Besides, reflexive pronouns will be further distinguished by a grammateme.

Reflexive pronouns typically occur in the following contexts:

1. as regular valency modifications: *She doesn't think of them.* → *She doesn't tink of herself.*

   The reflexive pronoun will get the t-lemma substitute `#PersPron` and the appropriate functor.
2. as optional verb objects in verbs allowing the understood reflexive object alternation: *She dressed* → *She dressed herself.*

   The reflexive pronoun will get the t-lemma substitute `#PersPron` and the functor `PAT` but the verb occurrence will refer to the intransitive frame.
3. as "ethical datives": *She bought herself some trifles*.

   The reflexive pronoun will get the t-lemma substitute `#PersPron` and the functor `BEN.`
4. unintentional actions: *to prick oneself, to cut oneself.*

   The reflexive pronoun will get the t-lemma substitute `#PersPron` and the functor `PAT`. If ever an intransitive verb occurs in this use the occurrence will refer to the intransitive frame unless is meaning is too different.
5. intentional actions: *to dry oneself, to recover oneself. She recovered from her illness* → *She sat and tried to recover herself.*

   Unlike with unintentional actions, optional reflexive usages of intransitive verbs/verbal frames deserve their own valency frames in the lexicon since their meaning undergoes a substantial change, as the different (Czech) translation equivalents suggest.
6. as predicative complement: *Do it yourself.*

   The reflexive pronoun will get the t-lemma substitute `#PersPron` and the functor `COMPL.`
7. as the dependent part of a phraseme: She was beside herself with joy.

   She was`.PRED` *beside*`_#PersPron.DPHR` *with joy*`.CAUS.`
8. as intensifier of a noun (always immediately folllowing the noun). Cf.: *I approached the chairman myself* → *I approached the chairman himself.*

   I approached the chairman `<myself> #PersPron.COMPL`

   I approached the chairman `<himself> #PersPron.RSTR`

   The Chairman `<himself> #PersPron.RSTR` *shook hands with him.*

   The reflexive pronoun when acting as a noun intensifier will get the t-lemma substitute `#PersPron` and the functor `RSTR.`

8.17. Dummy-do (Verbal Pro-form)

The **negative** form of do when used as present- or simple-past-tense pro-form is treated as textual ellips; i.e. the lexical verb is copied onto the place of the negative *do*-node. The inserted node inherits all relevant attribute values from the original verb. The negation is rendered as a `#Neg.RHEM` -node governed by the copied node. The node for *do* is attached to the newly copied node as its auxrf. The copied verbal node is assigned the same valency frame as the original occurrence of the given verb. Obligatory modifications are completed into the frame, having the appropriate t-lemma substitutes.

Do you know him? "I <don't> {know`[negation=neg1] #PersPron`*}.*

`#PersPron` corefers to "him". The copied node has the grammateme `[negation=neg1]`. The same applies to *will*, *would, should, be* as passive auxiliary, *have* as perfect-tense auxiliary etc. (even when negated).

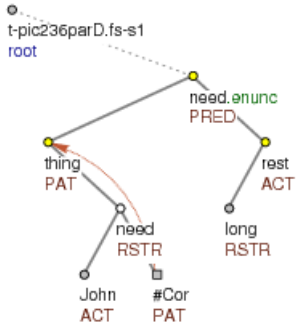
"Will you go there?" "I <will> {go `#OblFm.DIR3`}.

The **positive** form of *do*, used as pro-form, however, is a different case. E.g. in

Did Peter take a plane to New York today? His wife did.

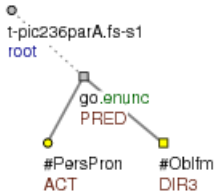
the lexical verb (*take*) is not to be copied into the second sentence as elided since mechanical insertion of the lexical verb on the surface behind the positive *do* might modify the meaning of the sentence (e.g. by moving the focus). Instead, *do* is attached as auxrf to a newly inserted node with the -lemma substitute `#VerbPron`. This new node corefers to the given verbal antecedent. **The valency frame of the antecedent is assigned to the** `#VerbPron` **node and its obligatory modifications are inserted**. All modifications governed by the original *do* naturally stay in their places, and they are assigned functors according to the valency frame inherited from the antecedent verb. Coreference arrows point from the present inner participants to their antecedents.

Figure 8.70. Dummy-do



(Do you know him?) I don't.

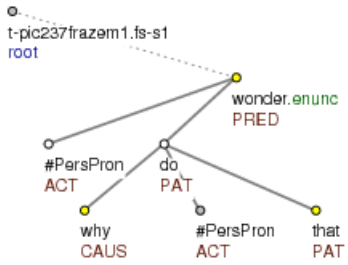
Figure 8.71. Ellipsis of the lexical verb



Will you go there? I will.

Whenever *do* has object it is treated as lexical verb:

Figure 8.72. Not-dummy-do



(That Ford slowed down as we passed.) I wonder why it did that/it.

8.17.1. The pro-form *to do so*

Sometimes a whole predication is substituted by *to do so*:

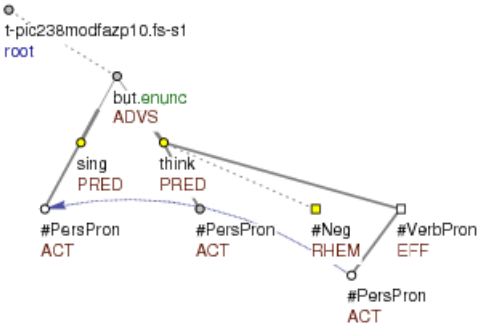
She might sing but I don't think she will do so. (= I don't think she will sing.)

This construction is also resolved by means of the t-lemma substitute `#VerbPron`. So is attached to `#VerbPron` as its auxrf.

Analogically:

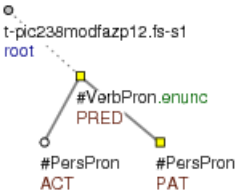
John buys his drinks at the local supermarket. So he does. (=He really buys his drinks at the local supermarket.)

Figure 8.73. to do so



She might sing but I don't think she will do so.

Figure 8.74. So X does.



(John buys his drinks at the local supermarket.) So he does.

NB: this construction is not identical with the ones meaning "as well" and "neither":

Peter likes tea. So.`RHEM` *<does>*{`#VerbPron`} *Paul*.

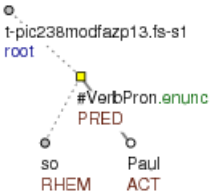
They have a cold. So.`RHEM` *have {a cold} I.*

She can't sleep. Neither.`RHEM` *<can> I {sleep}.*

She can't sleep. Me {sleep} neither.

So means *too*. It is assigned the functor `RHEM` and is governed by the verb. The verbs are resolved depending on whether they are lexical verbs, modals/auxiliaries or dummy-do's:

Figure 8.75. So does X in the meaning "X too"



(Peter likes tea.) So does Paul.

8.17.2. The pro-forms *so* and *not* in verbs of saying and in cognitive verbs

The pro-form *so* gets the functor appropriate to the valency modification it substitutes and a coreference arrow will point to its antecedent predication.

(John hasn't found a job yet). He told me so.EFF *yesterday.*

The pro-form *not* becomes the negative rhematizer #Neg.RHEM and is governed by the copied predication. Obligatory valency modifications are to be inserted and provided with coreference arrows pointing to their respective antecedents.

*(John says he.*ACT *will get a job .*PAT *tomorrow.) I think not.* = *I think* #PersPron.ACT *will* #Neg.RHEM *get* #PersPron.PAT.

Subject index

Symbols

A

C

D

E

F

G

H

I

K

L

M

N

O

P

List of example sentences

They found their friend ill.

She was leaving defeated.

He stood there, his hands in his pockets.

Julia, being a nun, spent much of her time in prayer and meditation.

He won the last election with a program designed as the negation of the former program.

John heard Mary crying/John heard Mary cry.

He went out, an ugly old man.

The old man came in a shabby coat.

She was quick to shut the door.

I bought a belt for CZK 70.

He repaid his debts to the insurance company.

He arrived on Thursday 5th (of) January 1997.

It is located two kilometers from the river.

We will meet at the Central station in the hall.

Give it to our director Novak.

There are steamboats on the Vltava river.

Your performance: miserable.

In Praha, at five o'clock.

An important event.

George!

Sorry.

John Smith, London.

A book review: The Hours is a new book.

Ooops, I didn't mean that.

Sir, I'm not staying any longer.

You have two possibilities how to get money.

The professor, inspired by the new article, gave a lecture on the new issues.

Judging by his face, he was angry.

Frankly speaking, she has chosen a wrong partner.

He has always been an influential, though controversial figure.

He run into an elephant in the garden, which killed him.

He left, never to come back again.

If I seem angry sometimes, it's usually because I'm very tired.

She had not arrived, which made him sad.

We have not answered your question completely, for which we apologize.

He's got five points, which was the maximum.

Eleanor's pen, which had been lying on the table, suddenly moved.

Guidance on the Circumstances, in which the Parents may Choose to Educate their Children at Home.

She's poor, that's why she had to come to Arthur.

In the process, the uranium (loses, or) is depleted(,) of almost half its radioactivity, which is how depleted uranium gets its name.

Another thing the microscope revealed was the difference between forming a laminate using vacuum pressure, which is how 3DL is made, and forming a laminate using pressure from rollers, which is how (a few of) the typical ('look-alike') sails are made.

This is where you find Ogden Utah, which is where the Central Pacific and Union Pacific railroads met in 1869.

We have reached the point where we want to sum up an infinite number of differential amounts, which is when we integrate.

To be successful, the formula is this reach combined with frequency, which is how often viewers will receive.

N is for never, which is how often he's sincere when he says he cares.

He was troubled by insects etc.

I saw and heard Mary sing.

Peter had been working on his dissertation and preparing for an English exam the whole day but in the evening he was doing nothing more.

John was stressed and difficult to get along with.

Peter arrived and Paul probably as well.

The match ended in a 5:0 draw.

In the ears 1995-1999 I was attending the high school.

Everybody was watching, from kids to adults.

Peter (and Paul) came there.

You can (if you dare) follow John.

The team ended up third (last year they lied first).

To think we might have been in Rome now!

The court, as it seems to me, has no opinion on the subject.

Mary prepared the lunch and John the dinner.

Why me?

The blue slippers are torn, the green dirty.

Only the younger came.

(She did not leave.) Because she wouldn't be there in time.

He bought red and white wine.

The presidents of Greece and of Austria were present at the ceremony.

(Has the shop assistant wrapped the book?) He has.

(The company was supposed to deliver the goods to the customer.) The delivery did not take place.

The boss has left.

Houses are built from bricks.

It was announced on the radio.

John and Mary kissed (each other).

Mary sang like John.

Mary sang like John did.

He is fit as a fiddle.

Mary is the same as John.

She does nothing but complain all day long.

Except for dates, all ordinals should be written in words.

Except this week I'll be teaching regularly.

(Lane closures will effect Castle Boulevard and Castle bridge Road the 5th February up to the 25th March), access is maintained everywhere apart from to the Castle.

Reformists have no access to free media except to the Internet.

We do not share email addresses with third parties unless required to do so by law.

Apart from the fact that she sings well she can do nothing.

I never take a bath but the phone rings or someone knocks the door.

But for John, we would have lost this match.

Barring accidents, we'll be there on time.

He did all but strangled me.

I would have told them the truth, except/except that they wouldn't have believed me.

Besides to Rome, they also wanted to go to Venice.

Herdwicks (start lambing a year after other breeds because they) need to grow enough.EXT to stand.RESL the harsh climate.

People with AIDS have enough to bear.

Peter didn't park in the parking lot.

Not only is the budget not in surplus, but it is even covertly deficit.

I will start to work on Monday.

Peter can come to the concert./Peter will come to the concert.

Peter cannot come and stay./Peter will not come and stay.

Peter could not, and cannot, come to the concert./Peter should not, and will not, come to the concert.

So they could not and cannot confirm or deny it./So they should not and will not confirm or deny it.

He had the right to protest.

She made a decision.

Petr got the order to come.

Pavel made Petr a suggestion.

The cat is a mammal.

Out of the blue they appeared in our room.

I'm all ears.

I've got five houses and three cars.

They elected three out of five vice-chairs.

Billions of people live here.

The new Golf 500 is already on the market.

We only won twice.

We have missed the third.

The United Nations Organization for Education, Science and Culture.

His Hamlet is all torn.

Guess Who's Coming to Dinner is fun.

Kerouac's On the Road

the company Sady, Lesy, Zahradnictví

He was contributing to the Swedish Svenska Dagbladet.

He said he wouldn't come: "I will certainly not arrive".

The coach: "It was not too bad."

His words "I envy him" were surprising.

He pointed at me: "You".

The mood is getting better: "We will beat them".

It took as long as to 1989 until the decision was taken to open the Journal to contributions from all over the world.

They passed on to customers as much as 50% of the costs of buying out their broken contracts.

If this harvest is good, then they will not need to import wheat.

Janice has been together with John four years now.

Sources close to the CEO say he is ready to make a deal.

But maintaining U.S. influence will be difficult in the face of Japanese dominance in the region.

Her son was happily playing at the family home in contradiction of a prognosis 12 years ago that he was not 'worth treating'.

He last sailed upon it in charge of a steamer's course.

John seems to understand.

John expects Mary to leave.

John hears Mary cry.

Why bother?

It's been kind of annoying, to say the least.

To think we might have been in Rome now!

To leave early would be impolite.

He refused to cooperate.

I was invited to contribute to the volume.

John seems to understand.

I am to blame.

It is to be stressed that Peter was living in London in 1976.

They made him stay.

John is eager to please.

John is difficult to please/It is difficult to please John.

This flat must have cost a lot to furnish.

She was quick to shut the door.

She was likely to leave.

She was lucky to get that job.

Sue made the decision to have a staff-directory in the phonebook.

She did not have the nerve to leave.

They had the votes necessary to defeat the amendment.

He stopped to have a drink.

The firms are moving faster to create new adds than they did in the fall of 1987.

She has gone to see her sister.

I'm too young to marry.

The weather was not such as to favour long walking hours.

He was rich enough to maintain a fleet.

I don't know her to speak to.

The first player to score

The most convenient time to go is Saturday.

He has a large family to provide for.

the problem of what to offer in return

There is nothing to be seen.

There are so few horses to see.

His leaving no address was most inconvenient.

A close friend of Geremew Ketema reported his being killed to OSG on 17 June.

I hate killing animals.

I'm afraid of antagonizing him.

You were fortunate in winning his support.

More dual-carreer couples are also getting help, with men increasingly bringing their wives for joint counselling.

It is quicker than going by train.

Having delimited the field of inquiry, we shall concentrate on the problem...

There will be singing.

There was no stopping him.

There will be people dancing.

It's no good bothering about it.

It is feared that the ship was wrecked.

He made it his business to secure financial aid.

Are they your parents?

It is not far to New York.

It is our wedding anniversary next month.

It's freezing outside.

There's no smoke without fire.

There is nothing to be seen.

There is nowhere to go.

The whole thing is a fraud.

He is 6 years old.

She wore a red dress but a blue suits her better.

The hungry and the poor

(Hot soup!) That's what I would like just now!

Crystal breaks easily.

Crystal breaks at the slightest touch.

John is difficult to please.

This flat must have cost a lot to furnish.

John broke a window. / It was John that broke the window.

It was a window John broke.

What John needs is a long rest.

The thing John needs is a long rest.

(Do you know him?) I don't.

Will you go there? I will.

(That Ford slowed down as we passed.) I wonder why it did that/it.

She might sing but I don't think she will do so.

(John buys his drinks at the local supermarket.) So he does.

(Peter likes tea.) So does Paul.

List of Figures

List of Tables