home issue tracker

This page pertains to UD version 2.

Universal Dependencies

The following table lists the 37 universal syntactic relations used in UD v2. It is a revised version of the relations originally described in *Universal Stanford Dependencies: A cross-linguistic typology* (de Marneffe *et al.* 2014).

The upper part of the table follows the main organizing principles of the UD taxonomy:

- Rows correspond to functional categories in relation to the head:
  - Core arguments of clausal predicates
  - Non-core dependents of clausal predicates
  - Dependents of nominals
- Columns correspond to structural categories of the dependent:
  - Nominals
  - Clauses
  - Modifier words
  - Function words

The lower part of the table lists relations that are not dependency relations in the narrow sense:

- Relations used to analyze coordination
- Relations used to analyze multiword expressions (MWE)
- Loose joining relations
- Special relations for ellipsis, disfluencies, and orthographic errors
- Special relations for clausal heads, punctuation and other relations

	Nominals	Clauses	Modifier words	Function Words

Core arguments	`nsubj`
	`obj`
	`iobj`
		`csubj`
		`ccomp`
		`xcomp`		

Non-core dependents	`obl`
	`vocative`
	`expl`
	`dislocated`
		`advcl`	`advmod`*
			`discourse`
				`aux`
				`cop`
				`mark`

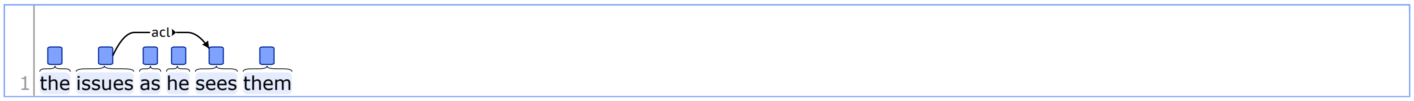
Nominal dependents	`nmod`
	`appos`
	`nummod`
		`acl`	`amod`	`det`
				`clf`
				`case`

Coordination	MWE	Loose	Special	Other

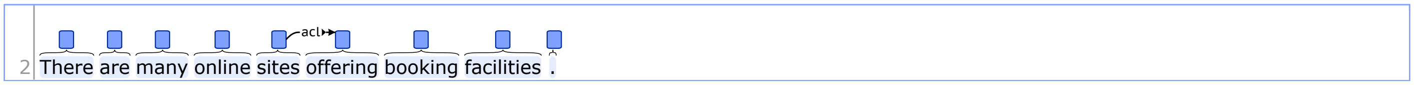
`conj`
`cc`
	`fixed`
	`flat`
	`compound`
		`list`
		`parataxis`
			`orphan`
			`goeswith`
			`reparandum`
				`punct`
				`root`
				`dep`

`acl`: clausal modifier of noun (adjectival clause)

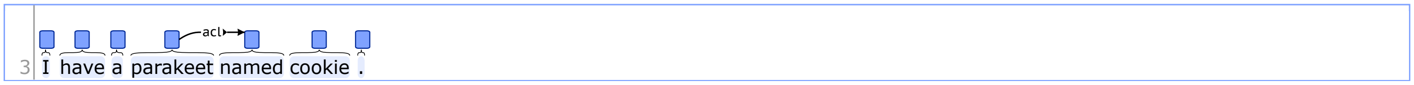
`acl` stands for finite and non-finite clauses that modify a nominal. The `acl` relation contrasts with the `advcl` relation, which is used for adverbial clauses that modify a predicate. The head of the `acl` relation is the noun that is modified, and the dependent is the head of the clause that modifies the noun.



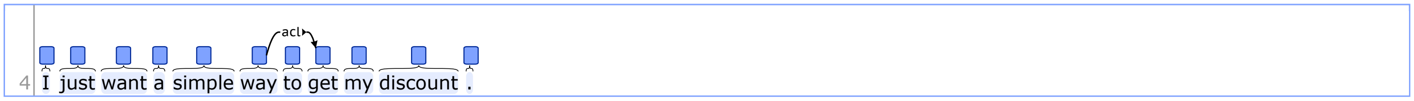
1 the issues as he sees them



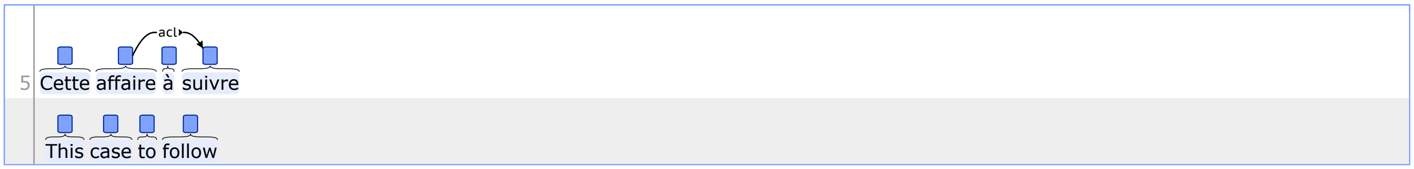
2 There are many online sites offering booking facilities .



3 I have a parakeet named cookie .



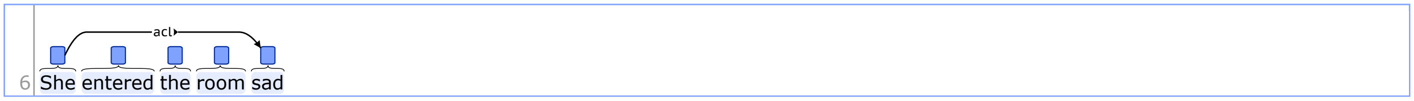
4 I just want a simple way to get my discount .



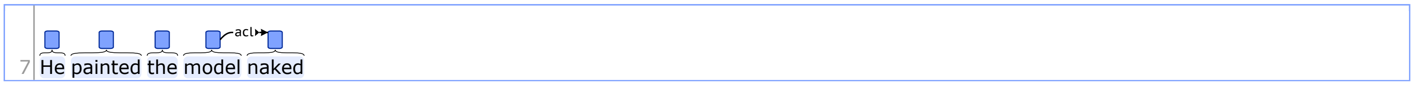
5 Cette affaire à suivre

This case to follow

This relation is also used for optional depictives. The adjective is taken to modify the nominal of which it provides a secondary predication. See `xcomp` for further discussion of resultatives and depictives.

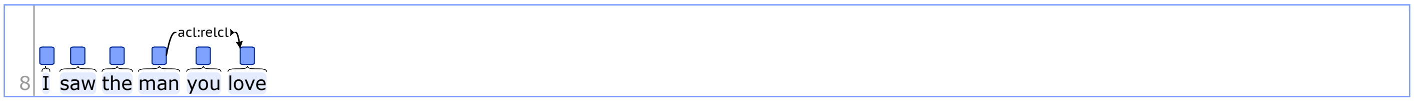


6 She entered the room sad



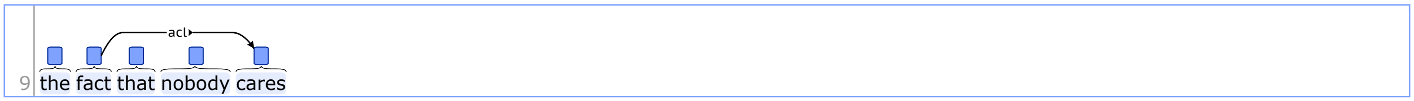
7 He painted the model naked

A relative clause is an instance of `acl`, characterized by finiteness and usually omission of the modified noun in the embedded clause. Some languages use a language-particular subtype `acl:relcl` for the traditional class of relative clauses.



8 I saw the man you love

Some languages allow finite clausal complements for nouns with a subset of nouns like *fact* or *report*. These look roughly like relative clauses, but do not have any omitted role in the dependent clause. This is the class of “content clauses” in Huddleston and Pullum 2002). These are also analyzed as `acl`.

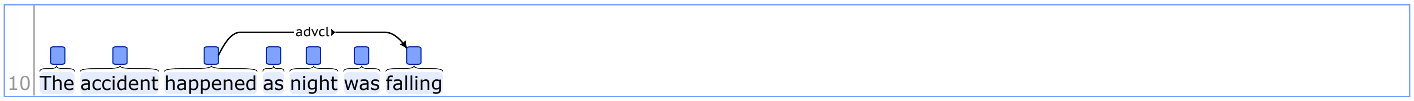


9 the fact that nobody cares

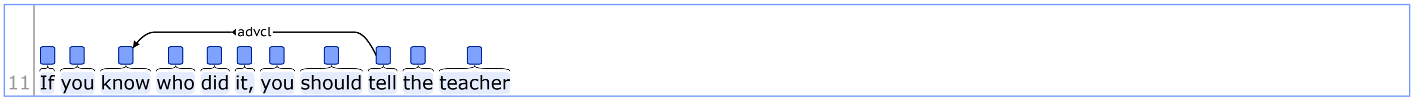
edit acl

`advcl`: adverbial clause modifier

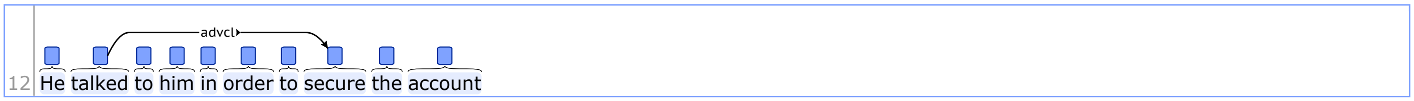
An adverbial clause modifier is a clause which modifies a verb or other predicate (adjective, etc.), as a modifier not as a core complement. This includes things such as a temporal clause, consequence, conditional clause, purpose clause, etc. The dependent must be clausal (or else it is an `advmod`) and the dependent is the main predicate of the clause.



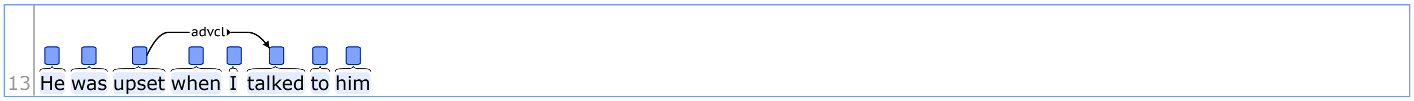
10 The accident happened as night was falling



11 If you know who did it, you should tell the teacher



12 He talked to him in order to secure the account



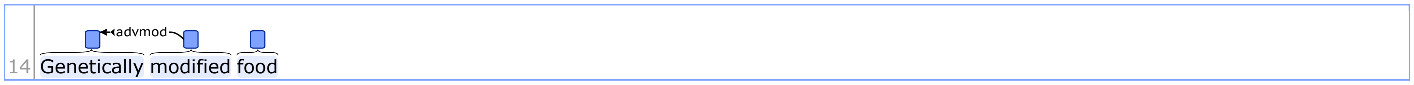
13 He was upset when I talked to him

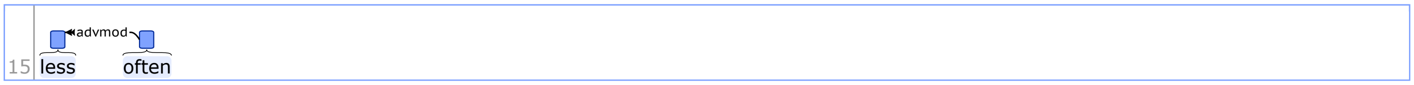
edit advcl

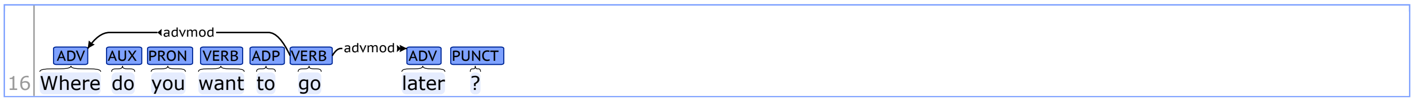
advmod: adverbial modifier

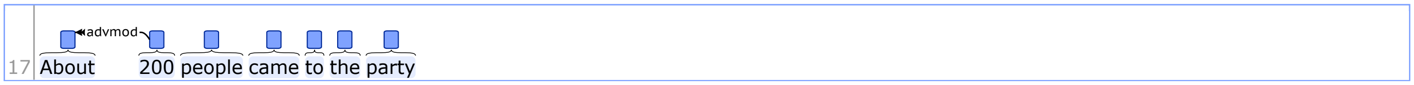
An adverbial modifier of a word is a (non-clausal) adverb or adverbial phrase that serves to modify a predicate or a modifier word.

Note that in some grammatical traditions, the term *adverbial modifier* covers constituents that function like adverbs regardless whether they are realized by adverbs, adpositional phrases, or nouns in particular morphological cases. We differentiate adverbials realized as adverbs *(advmod)* and adverbials realized by noun phrases or adpositional phrases (obl). However, we do not differentiate between modifiers of predicates (adverbials in a narrow sense) and modifiers of other modifier words like adjectives or adverbs (sometime called qualifiers). These functions are all subsumed under advmod.

14


15


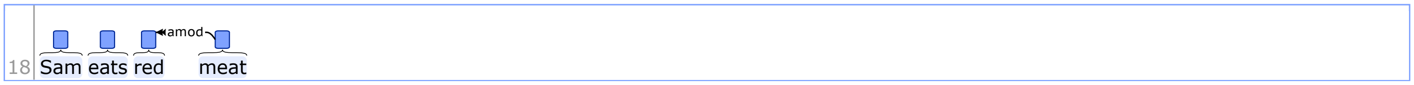
16

advmod
advmod
ADV AUX PRON VERB ADP VERB ADV PUNCT
Where do you want to go later ?

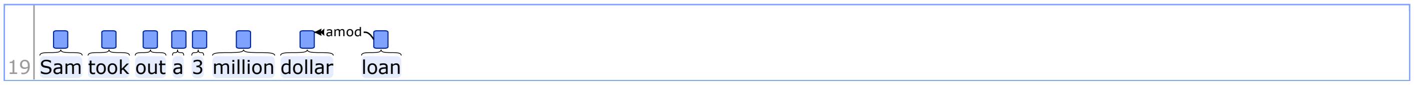
17


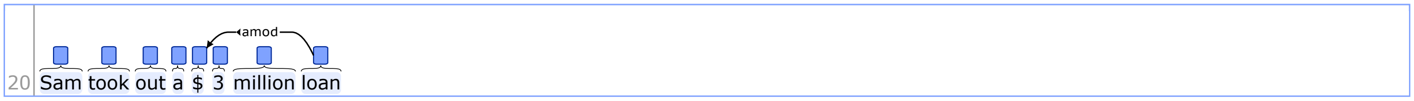
edit advmod

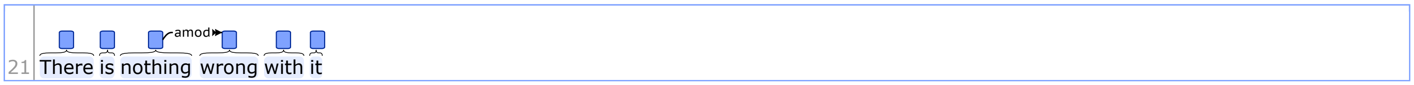
amod: adjectival modifier

An adjectival modifier of a noun (or pronoun) is any adjectival phrase that serves to modify the meaning of the noun (or pronoun).

18


19


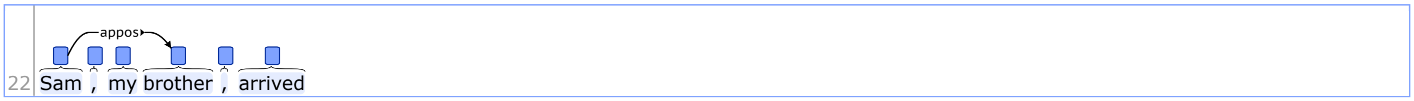
20


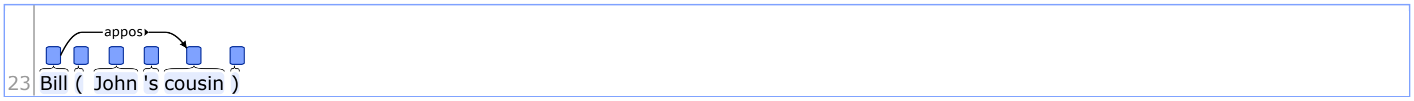
21


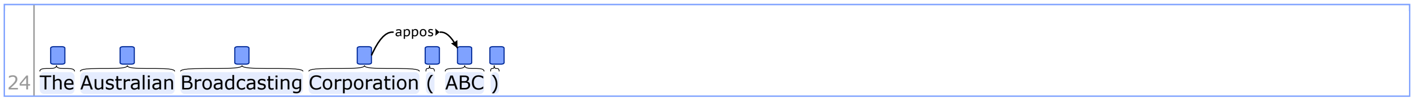
edit amod

appos: appositional modifier

An appositional modifier of a noun is a nominal immediately following the first noun that serves to define, modify, name, or describe that noun. It includes parenthesized examples, as well as defining abbreviations in one of these structures.

22


23


24


appos is intended to be used between two nominals. In general, modulo punctuation, the two halves of an apposition can be switched. For example, you could also say *My brother, Sam, arrived.* There are somewhat similar constructions with titles where the title is less than a full nominal, such as *state senator Paul Mnuchin*, where reversal is impossible or would require insertion of a determiner to make a full nominal.

Some grammatical traditions, descending from Latin, call *state senator* in such cases a “fixed (or close) apposition” and take the name as the head. However, we seem to have only one nominal not two here. For example:

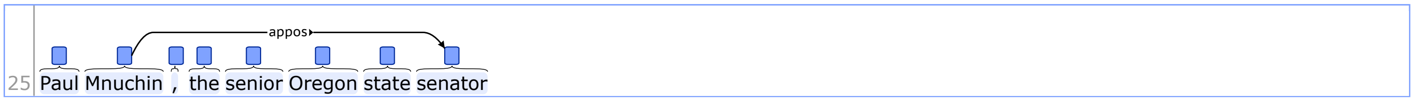
President Obama

*Obama President

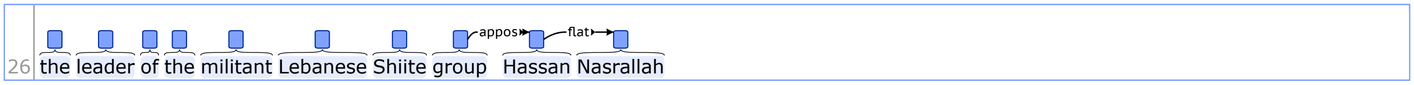
state senator Paul Mnuchin

*Paul Mnuchin state senator

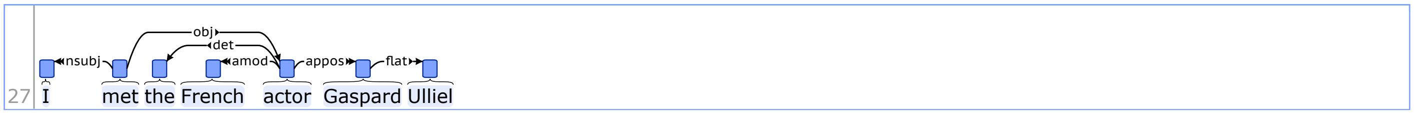
`appos` should not be used in such cases. However, the examples can usually be rendered in a fuller form, corresponding to “loose (or wide) apposition” in the Latin tradition, where there are two full phrases. Then the relation `appos` is appropriate, for example:

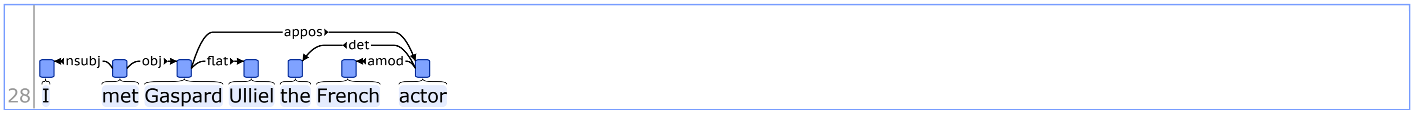


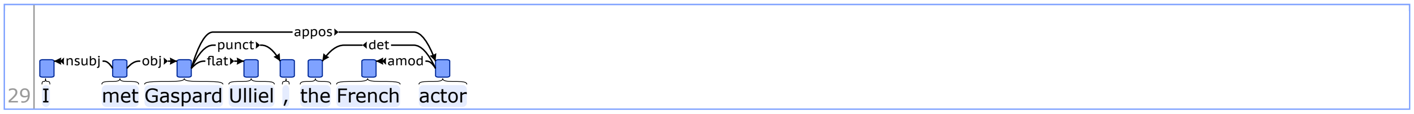
As is often the case, there are borderline cases. In formal writing, punctuation is usually a good signal of apposition, but there are certainly cases of apposition where no punctuation is used:

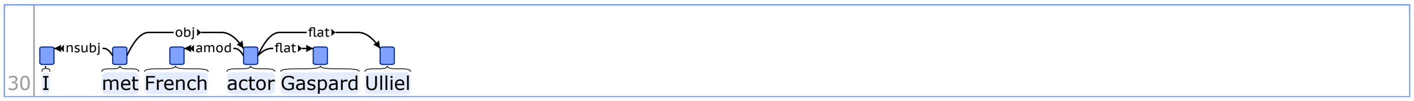

26
the leader of the militant Lebanese Shiite group Hassan Nasrallah
appos
flat

Good tests include to ask whether the two halves are full nominals, whether the two halves can be swapped or not, and whether there is case or agreement concord (in a language with rich morphology). So we have:

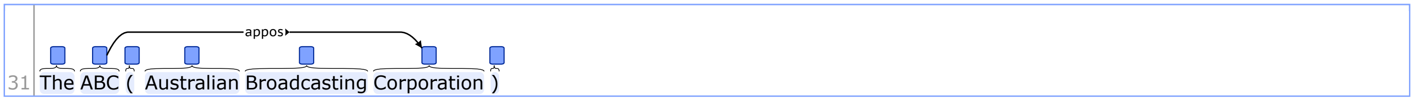




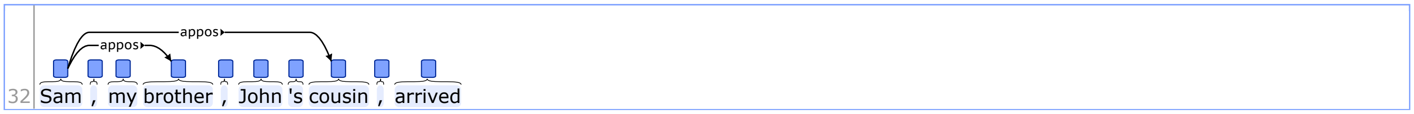




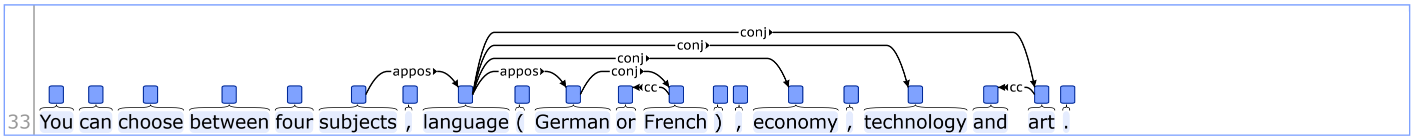
While items like abbreviations are generally reversable, the determiner test suggested above doesn’t quite work there, since the determiner seems to belong with the main item:



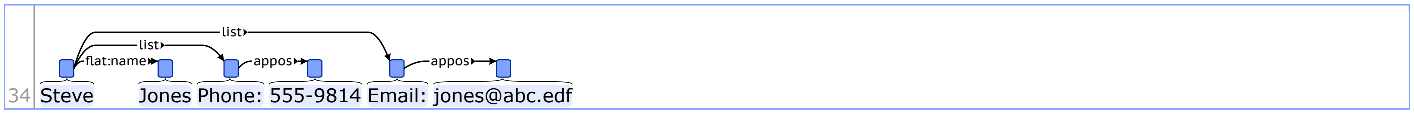
In the rare cases of more than one appositive nominal, all nouns should be marked as modifying the first noun, rather than being chained:



Note however that nested apposition cannot be completely excluded. It may occur in combination with coordination:


33
You can choose between four subjects , language (German or French) , economy , technology and art .
appos
appos
conj
conj
conj
conj
cc
cc

`appos` is also used to link key-value pairs in addresses, signature blocs, etc. (see also the `list` label):

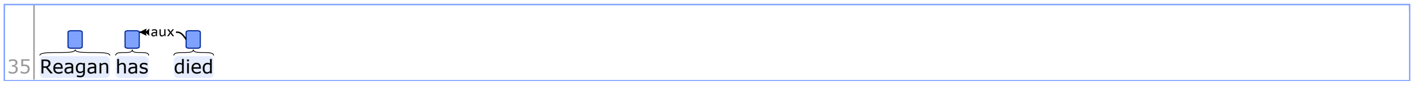


edit appos

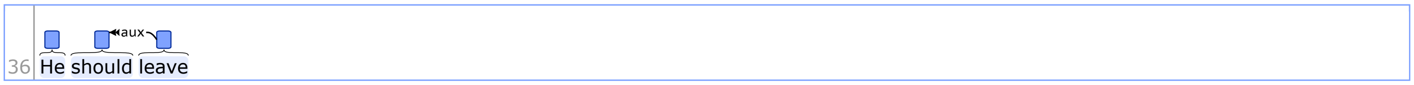
`aux`: auxiliary

An `aux` (auxiliary) of a clause is a function word associated with a verbal predicate that expresses categories such as tense, mood, aspect, voice or evidentiality. It is often a verb (which may have non-auxiliary uses as well) but many languages have nonverbal TAME markers and these are also treated as instances of `aux`.

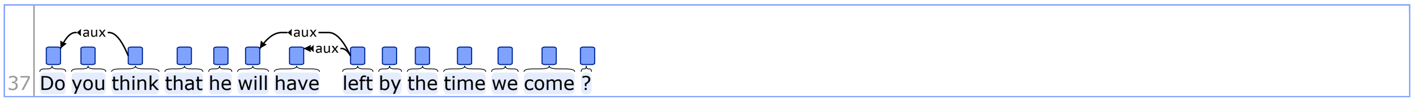
New from v2: Auxiliares used to construct the passive voice are now also labeled `aux`, although we strongly encourage the use of the subtype `aux:pass` in language that have a grammaticalized (periphrastic) passive.



35 Reagan has died



36 He should leave

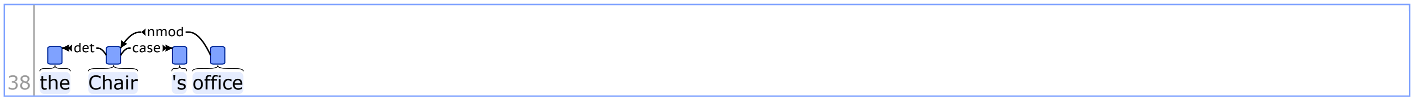


37 Do you think that he will have left by the time we come ?

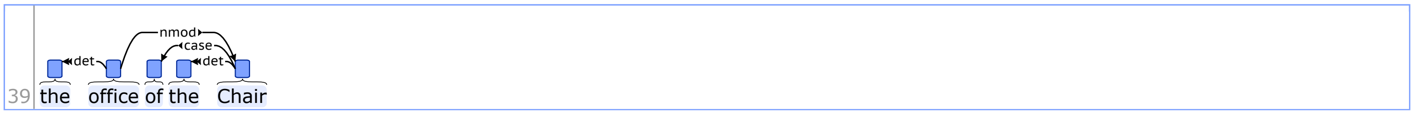
edit aux

`case`: case marking

The `case` relation is used for any case-marking element which is treated as a separate syntactic word (including prepositions, postpositions, and clitic case markers). Case-marking elements are treated as dependents of the noun they attach to or introduce. (Thus, contrary to SD, UD abandons treating a preposition as a mediator between a modified word and its object.) The `case` relation aims at providing a more uniform analysis of nominal elements, prepositions and case in morphologically rich languages: a nominal in an oblique case will receive the same dependency structure as a nominal introduced by an adposition.

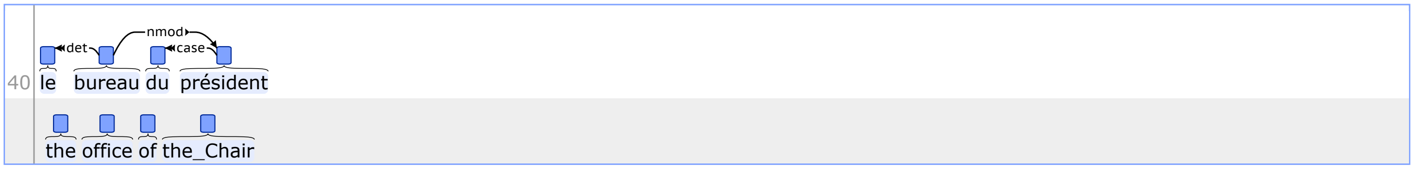


38 the Chair 's office



39 the office of the Chair

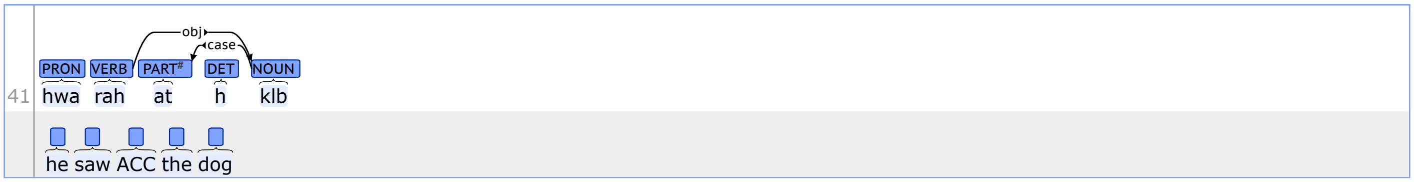
French:



40 le bureau du président

the office of the_Chair

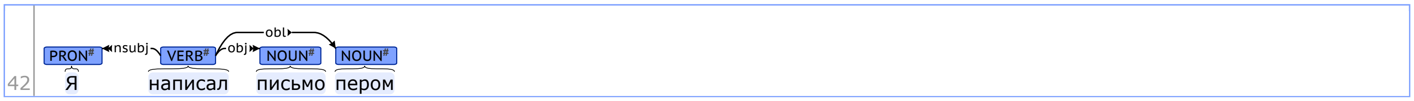
Hebrew:



41 hwa rah at h klb

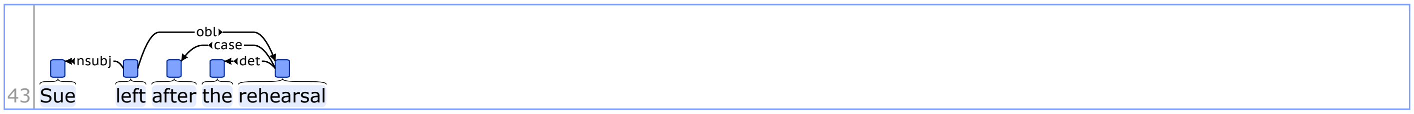
he saw ACC the dog

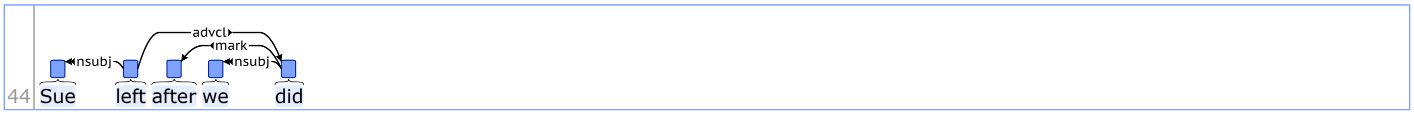
When case markers are morphemes, they are not divided off the noun as a separate case dependent, but the noun as a whole is analyzed as `obl` (if dependent on a predicate) or `nmod` (if dependent on noun). To overtly mark case, POS tags and features are included in the representation as shown below on a Russian example (put your mouse pointer over the words to see additional morphosyntactic features).



42 Я написал письмо пером

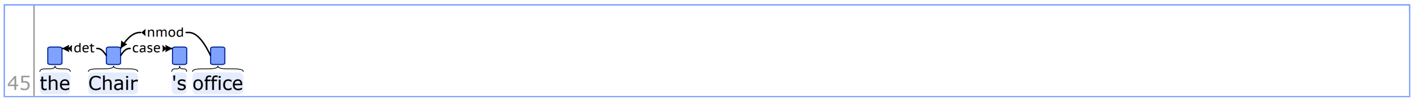
This treatment provides parallelism between different constructions across and within languages. A good result is that we now have greater parallelism between prepositional phrases and subordinate clauses, which are often introduced by a preposition in some languages (but note that the relation should be `mark` in those cases):

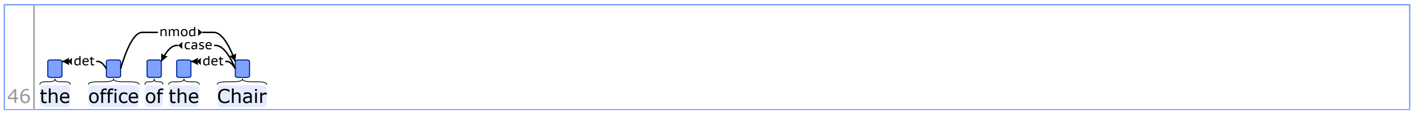




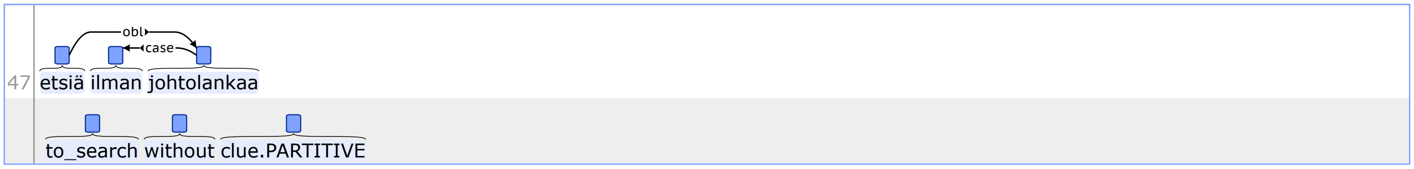
We also obtain parallel constructions for

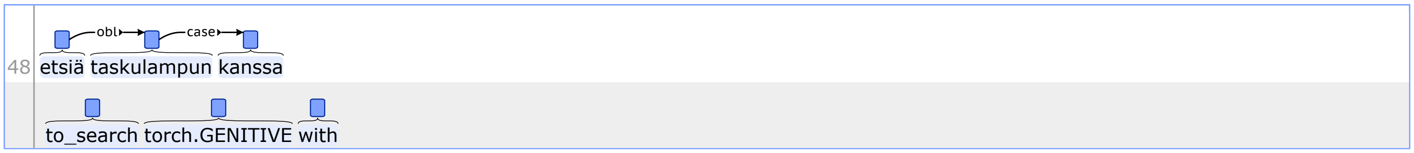
- the possessive alternation

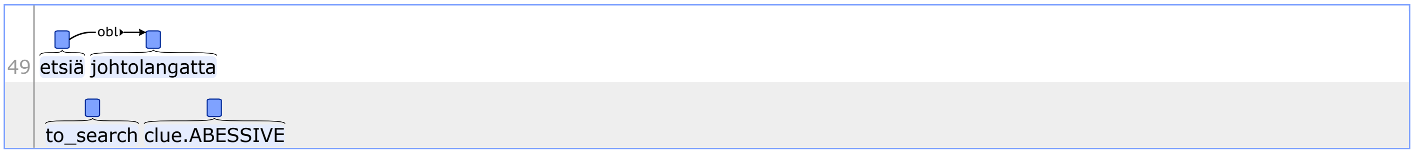




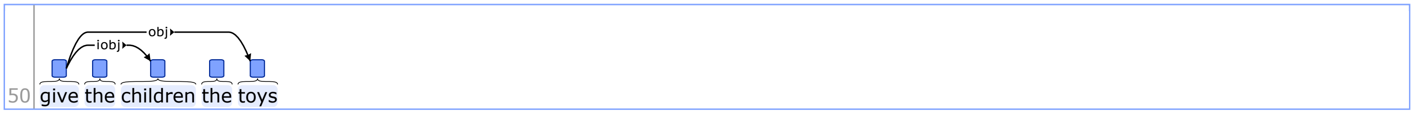
- variant forms with case, a preposition or a postposition in Finnish

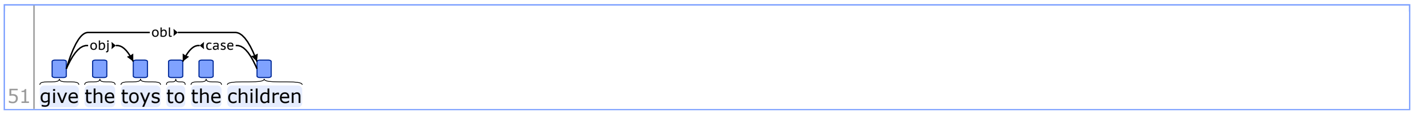


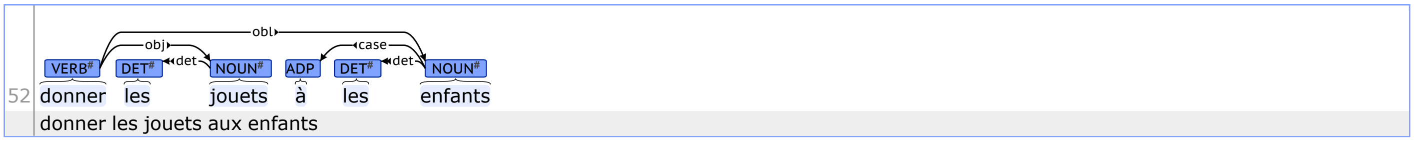




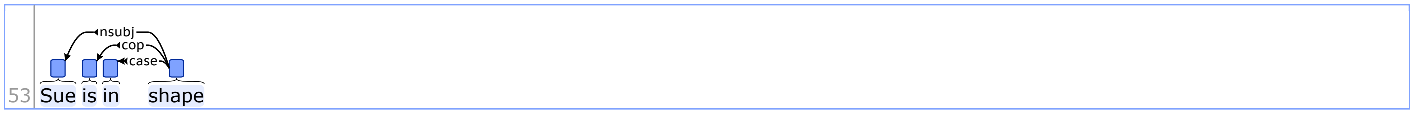
- the dative alternation where the prepositional construction gets a similar analysis to the double object construction



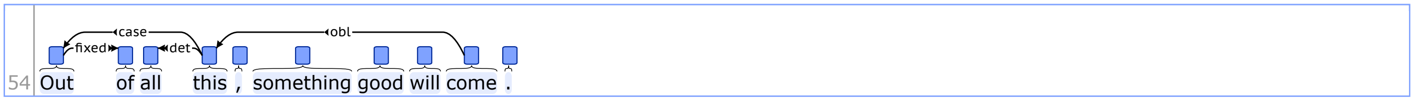



52
obl
obj
case
det
det
VERB#
DET#
NOUN#
ADP
DET#
NOUN#
donner les jouets à les enfants
donner les jouets aux enfants

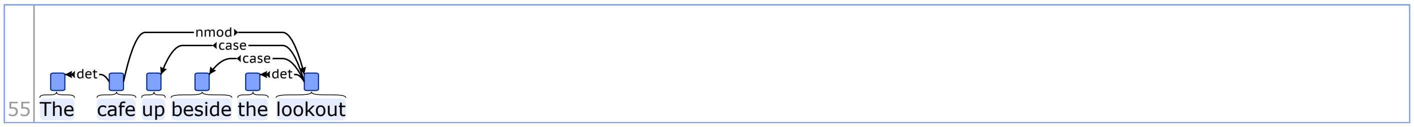
Another advantage of this new analysis is that it provides a treatment of prepositional phrases that are predicative complements of “be” that is consistent with the treatment of nominal predicative complements:



When prepositions are stacked (that is, there is a sequence of prepositions), there are two possible analyses. If the sequence is a frozen combination with a specific meaning, then the best analysis is as `fixed`. An English example of this is *out of*.



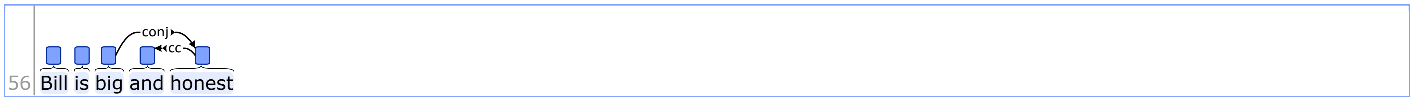
However, if various combinations of prepositions can be used to express different meaning combinations or nuances, then each preposition is independently analyzed as a case dependent. Examples of this in English include *up beside* (which can alternate with *down beside* or *up* *near*) or *except during* which can alternate with *as during* or *except after*.



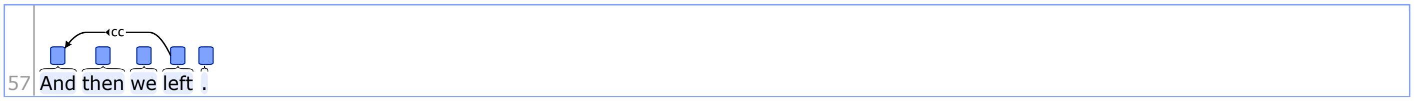
edit case

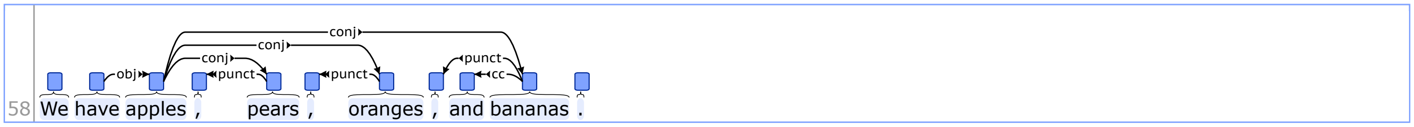
cc: coordinating conjunction

A cc is the relation between a conjunct and a preceding coordinating conjunction.



A coordinating conjunction may also appear at the beginning of a sentence. This is also called a cc, even though there is no preceding conjunct (except implicitly or in a preceding sentence).

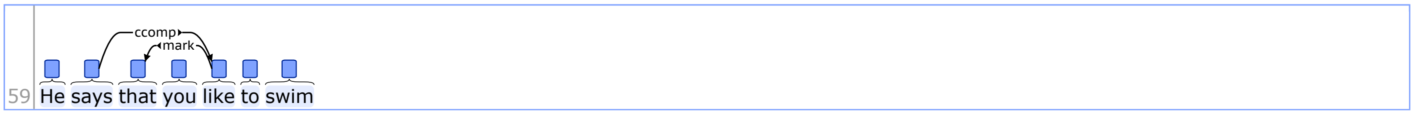


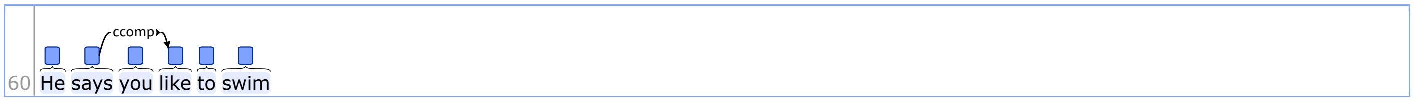


edit cc

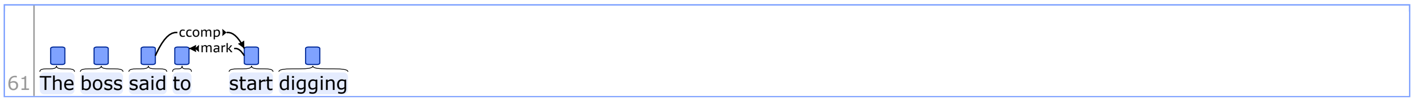
ccomp: clausal complement

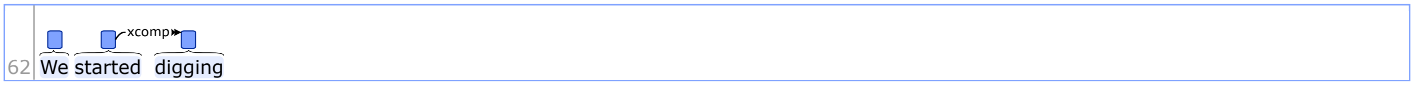
A clausal complement of a verb or adjective is a dependent clause which is a core argument. That is, it functions like an object of the verb, or adjective.





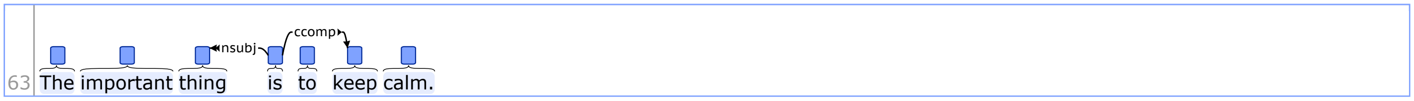
Such clausal complements may be finite or nonfinite. However, if the subject of the clausal complement is controlled (that is, must be the same as the higher subject or object, with no other possible interpretation) the appropriate relation is xcomp.

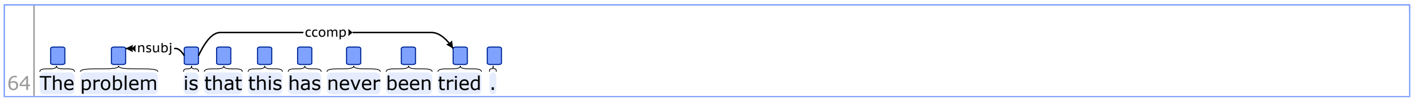




The key difference here is that, while it is possible to interpret the first sentence to mean that the boss will not be doing any digging, in the second sentence it is clear that the subject of **digging** can only be **we**. This is what distinguishes ccomp and xcomp.

Additionally, ccomp is used with copulas in equational constructions involving full clauses.





(In these cases, the copula is treated as a head to preserve the integrity of clause boundaries and prevent one predicate to be assigned two subjects. This is not an optimal solution given the analysis of equational constructions involving nominals, where one of the nominals is treated as the head, but it is the preferred solution for now.)

Note: In earlier versions of SD/USD, complement clauses with nouns like *fact* or *report* were also analyzed as `ccomp`. However, we now analyze them as `acl`. Hence, `ccomp` does not appear in nominals. This makes sense, since nominals normally do not take core arguments.

edit ccomp

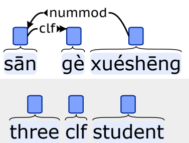
`clf`: classifier

A `clf` (classifier) is a word which accompanies a noun in certain grammatical contexts. The most canonical use is numeral classifiers, where the word is used with a number for counting objects. A classifier generally reflects some kind of conceptual classification of nouns, based principally on features of their referents. Etymologically, classifiers are normally historically nouns, and the words may still also be used as independent nouns, but in their classifier use they have scant semantics left. In most cases, the most appropriate UPOS to give classifiers will still be NOUN, though you may wish to give the words a feature indicating their special status as a classifier. (There is at present no Universal feature for classifiers, but `NounType=Clf` might be apt.) The `clf` function is intended for languages which have highly grammaticalized systems of classifiers. The greatest density of such languages is in Asia. As well as core classifiers, there are often also other words, sometimes called “massifiers” that are used in counting with similar behavior to classifiers. These typically include words for containers (“cup”, “box”) and units (“month”, “inch”), such as Chinese 袋 ‘bag’ in 一袋米 [one bag rice] ‘a bag of rice’. In a classifier language, it is usually most appropriate to also analyze these words as classifiers. Most other languages also count things with units, however, for these languages, such as English, `clf` is not used and rather standard noun phrase relations are still used (despite there also being incipient grammaticalization in many cases, including English). See the examples for English at the end.

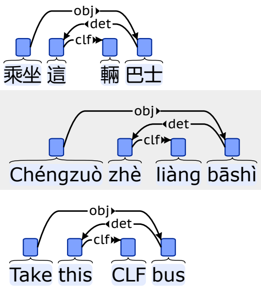
Here are some examples from Mandarin/Putonghua Chinese:

- 三个学生 (三個學生) sān gè xuéshēng = “three students”, literally “three [human-classifier] student”
- 三棵树 (三棵樹) sān kē shù = “three trees”, literally “three [tree-classifier] tree”
- 三只鸟 (三隻鳥) sān zhī niǎo = “three birds”, literally “three [bird-classifier] bird”
- 三条河 (三條河) sān tiáo hé = “three rivers”, literally “three [long-wavy-classifier] river”

Syntactically, the classifier groups with the numeral rather than the noun and we therefore treat classifiers as functional dependents of numerals (or possessives) using the new `clf` relation. (This is one of Greenberg’s universals and is true in almost all cases. A couple of exceptions are noted in Aikhenvald (2000: 105) *Classifiers*, OUP, but it is noticeable that in those languages the putative head noun is in the genitive case.)

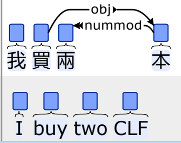
65


Sometimes a classifier is inserted between a demonstrative and a noun (instead of numeral and noun) [zh]:

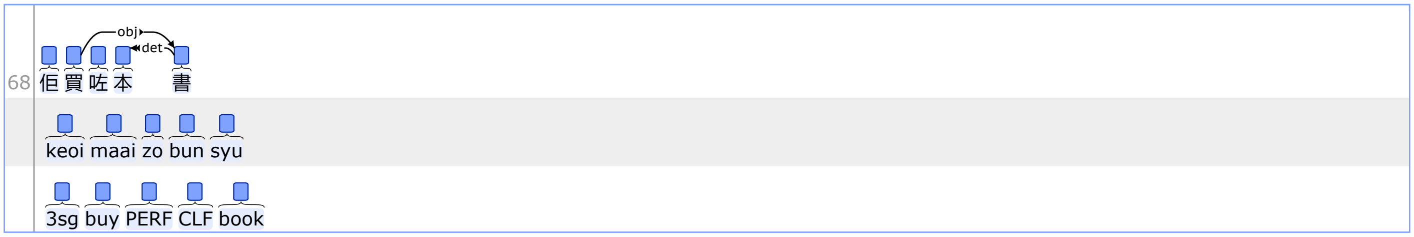
66


Classifier words also occur in various other constructions, and so it is important to distinguish the word in a particular language from the universal classifier function proposed in UD. We go through here some further examples with Chinese classifiers.

No noun may appear with the number and classifier. In this case, the classifier takes the role of the missing noun, and we promote the classifier to be the head. So 我 買 兩 本 “I am buying two” is regarded as “I am buying two [books-CLF]”.

67


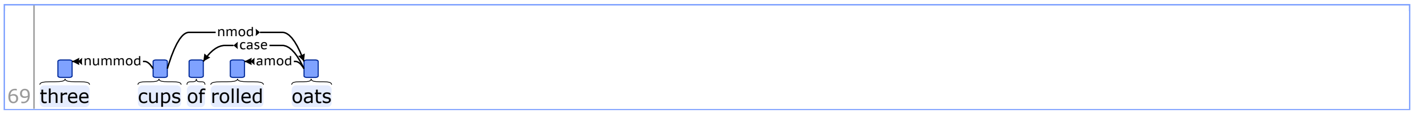
In some languages, including Chinese, a classifier can also appear without a number, and frequently then has some sort of determinative function. We use the relation `det` for such uses of a classifier. For instance, in Cantonese ‘She bought a/the book’:

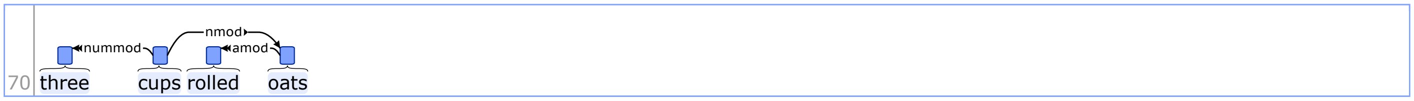

68 佢 買 咗 本 書

keoi maai zo bun syu

3sg buy PERF CLF book

For languages without highly grammaticalized classifier systems, standard nominal modification relationships are used even when things are being counted in groups (with “massifiers”). For example, in English:


69 three cups of rolled oats

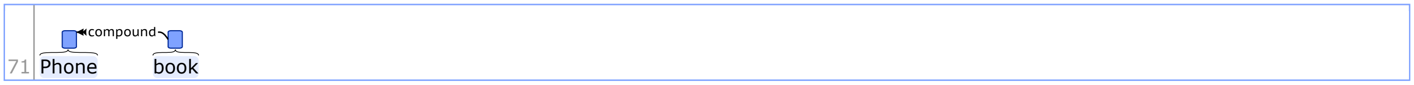

70 three cups rolled oats

edit clf

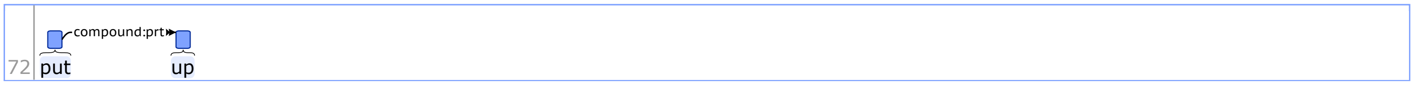
`compound`: compound

The `compound` relation is one of three relations for multiword expressions (MWEs) (the other two being `fixed` and `flat`). It is used for

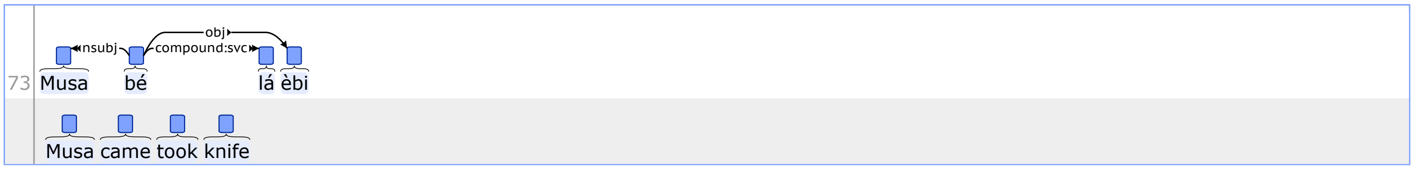
- any kind of X^0 compounding: noun compounds (e.g., *phone book*), but also verb and adjective compounds that are more common in other languages (such as Persian or Japanese light verb constructions).


71 Phone book

- for particle verbs (with the subtype `compound:prt`):


72 put up

- for serial verbs (with the subtype `compound:svc`):


73 Musa bé lá èbi

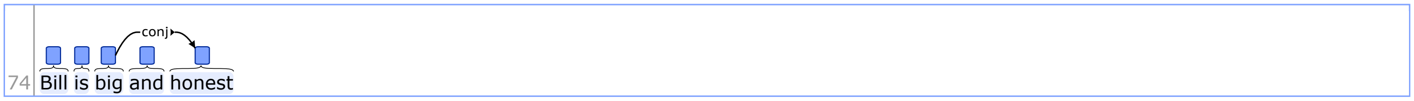
Musa came took knife

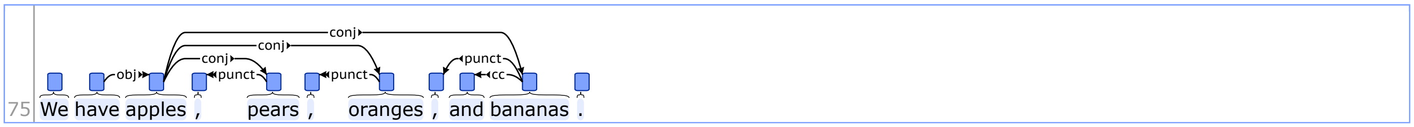
The `compound` relation (nor any subtype thereof) is not used to link an inherently reflexive verb with the reflexive morpheme, despite the similarity of this construction to particle verbs. The current UD guideline is to use an appropriate subtype of the `expl` relation.

edit compound

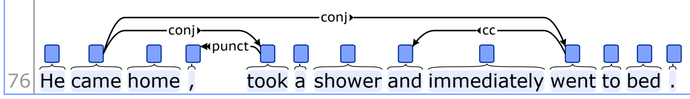
`conj`: conjunct

A conjunct is the relation between two elements connected by a coordinating conjunction, such as *and, or,* etc. We treat conjunctions asymmetrically: The head of the relation is the first conjunct and all the other conjuncts depend on it via the `conj` relation.

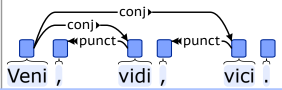

74 Bill is big and honest


75 We have apples , pears , oranges , and bananas .

Coordinated clauses are treated the same way as coordination of other constituent types:

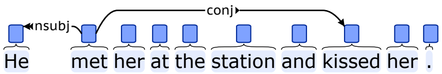
76


Coordination may be *asyndetic,* which means that the coordinating conjunction is omitted. Commas or other punctuation symbols will delimit the conjuncts in the typical case. Asyndetic coordination may be more frequent in some languages, while in others, conjunction will appear between every two conjuncts *(John and Mary and Bill).*

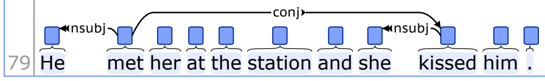
77


Shared Dependents and Effective Parents in Coordination

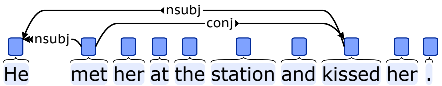
Note that the current basic annotation scheme cannot distinguish between a dependent of the first conjunct and a shared dependent of the whole coordination:

78


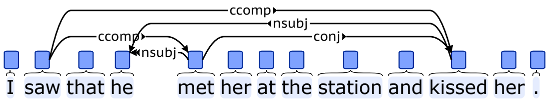
vs.

79


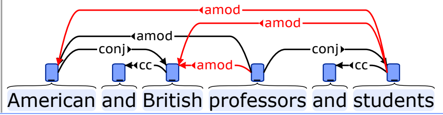
In contrast, the additional dependencies in the enhanced representation can be used to encode the fact that in the first case, *he* is also subject of *kissed:*

80


Furthermore, the enhanced representation can also capture the relation of each conjunct to the parent of the coordination. Nevertheless, the effective parents can be found algorithmically and showing them explicitly is for convenience only, while the information about shared dependents is otherwise not available.

81


If a dependent is shared among conjuncts, the basic representation always links it to the first conjunct (coordination head), while the enhanced representation shows all dependencies. In the following example, relations that are only part of the enhanced representation are shown in red.

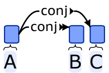
82


Nested Coordination

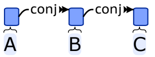
Note further that the basic annotation scheme has only a limited capability to capture nested coordination such as *apples and pears or oranges and lemons.* Consider coordinations

- A, B, C
- (A, B), C
- A, (B, C)

The first two cases, i.e. (A, B, C) and ((A, B), C), lead to the same tree:

83


Only the right-nesting case (A, (B, C)) can be distinguished because its tree is different:

84 

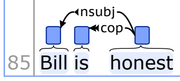
edit conj

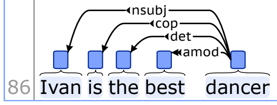
cop: copula

A cop (copula) is the relation of a function word used to link a subject to a nonverbal predicate. It is often a verb but nonverbal (pronominal) copulas are also frequent in the world's languages. Verbal copulas are tagged AUX, not VERB. Pronominal copulas are tagged PRON or DET.

The cop relation should only be used for pure copulas that add at most TAME categories to the meaning of the predicate, which means that most languages have at most one copula, and only when the nonverbal predicate is treated as the head of the clause.

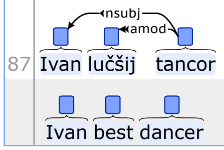
As a concrete example, in many European languages the equivalent of the English verb *to be* is the only word that can appear with the cop relation. In Spanish and related languages, both *ser* and *estar* can be copulas. In Czech and related languages, both *být* and *bývat* are copulas (because they are morphological variants of the same lexeme, and the reason they have two lemmas is that aspect-related morphology is treated as derivational in these languages). In contrast, the equivalents of *to become* are not copulas despite the fact that traditional grammar may label them as such. Existential *to be* can be copula only if it is the same verb as in equivalence clauses *(John is a teacher)*. If a language uses two different verbs, then the existential one is not a copula. Some more discussion of the topic is archived here.

85 

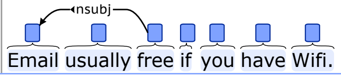
86 

The copula *be* is not treated as the head of a clause, but rather the nonverbal predicate, as exemplified above.

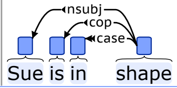
Such an analysis is motivated by the fact that many languages often or always lack an overt copula in such constructions, as in the the following Russian example:

87 

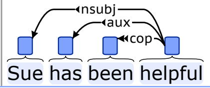
In informal English, this may also arise.

88 

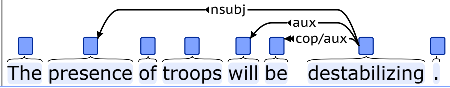
This analysis is adopted also when the predicate is a prepositional phrase, provided that the same copula (or absence thereof) is used here, in which case the nominal part of the prepositional phrase is the head of the clause.

89 

If the copula is accompanied by other verbal auxiliaries for tense, aspect, etc., then they are also given a flat structure, and taken as dependents of the lexical predicate:

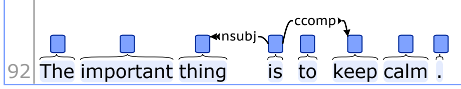
90 

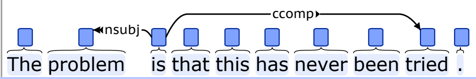
The motivation for this choice is that this structure is parallel to the flat structure which we give to auxiliary verbs accompanying verbs. In particular, in languages such as English, it is often very difficult to decide whether to regard a participle as a verb or an adjective. Perhaps the following sentence is such a case:

91 

While a part of speech (and associated deprel: `cop` vs. `aux`) has to be decided in such cases, it would be unfortunate if the choice of part of speech also changed the dependency structure. Note, however, that the exact distribution of the copula construction is subject to language-specific variation.

Finally, the `cop` relation is not used when the nonverbal predicate has the form of a clause, which typically occur in equational constructions like the following:

92 

93 

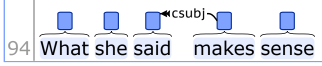
If we took the predicate of the clause as the head, instead of the copula verb, it would have two subjects, which would be unworkable. Examples like the above *could* be analyzed reversed with the initial noun phrase as the predicate, but in addition to this seeming undesirable, it would fail to be a solution if there were a clause on both sides of *be*, such as in: *(For us) to not attempt to solve the problem is (for us) to acknowledge defeat*. (Note: This solution is not perfect and refining it is a possible direction for the future.)

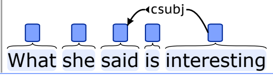
edit cop

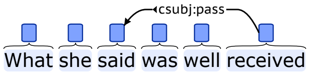
`csubj`: clausal subject

A clausal subject is a clausal syntactic subject of a clause, i.e., the subject is itself a clause. The governor of this relation might not always be a verb: when the verb is a copular verb, the root of the clause is the complement of the copular verb. The dependent is the main lexical verb or other predicate of the subject clause. In the following examples, *what she said* (that is, *said*) is the clausal subject of *makes* and *interesting*, respectively.

New from v2: The `csubj` relation is also used for the clausal subject of a passive verb or verb group. For languages that have a grammaticalized passive transformation, it is strongly recommended to use the subtype `csubj:pass` in such cases.

94 

95 

96 

edit csubj

`dep`: unspecified dependency

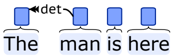
A dependency can be labeled as `dep` when it is impossible to determine a more precise relation. This may be because of a weird grammatical construction, or a limitation in conversion or parsing software. The use of `dep` should be avoided as much as possible.

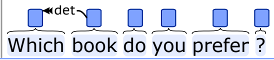
edit dep

`det`: determiner

The relation determiner (`det`) holds between a nominal head and its determiner. Most commonly, a word of POS `DET` will have the relation `det` and vice versa. The known exceptions at present are:

- In some of the datasets, a possessive determiner like [en] *my* is currently given the POS tag `DET` but the relation `nmod`, so that it is parallel with other possessive constructions. This is not yet completely parallel across languages; in some languages, it is much more clear than in English how possessive determiners relate to adjectives, and the `nmod` relation is out of question.

97 

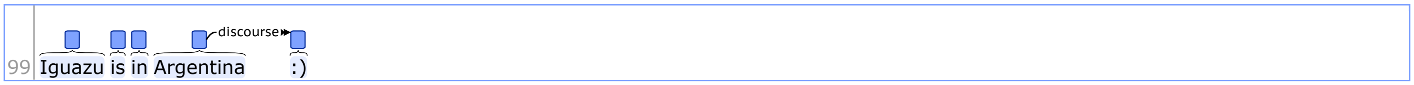
98 

edit det

discourse: discourse element

This is used for interjections and other discourse particles and elements (which are not clearly linked to the structure of the sentence, except in an expressive way). We generally follow the guidelines of what the Penn Treebanks count as an INTJ. They define this to include: interjections (*oh*, *uh-huh*, *Welcome*), fillers (*um*, *ah*), and non-adverbial discourse markers (*well*, *like*, but not *you know* or *actually*).

These discourse elements are attached to the head of the most relevant nearby clause, which is why they are grouped with non-core clausal dependents even though they are normally not dependents of the predicates as such.



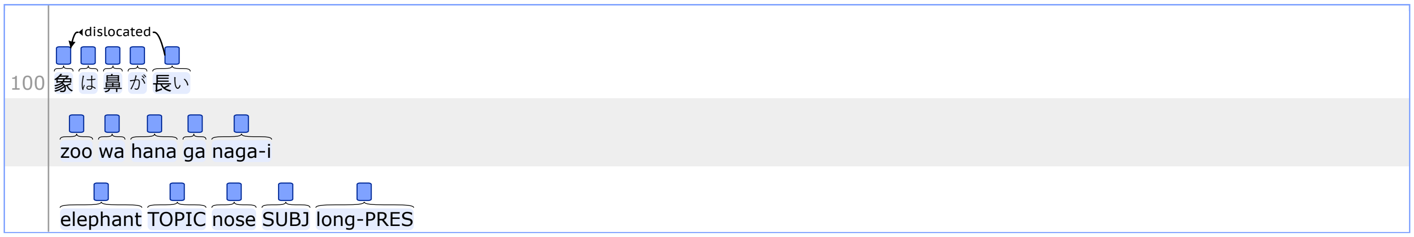
99 Iguazu is in Argentina :)

edit discourse

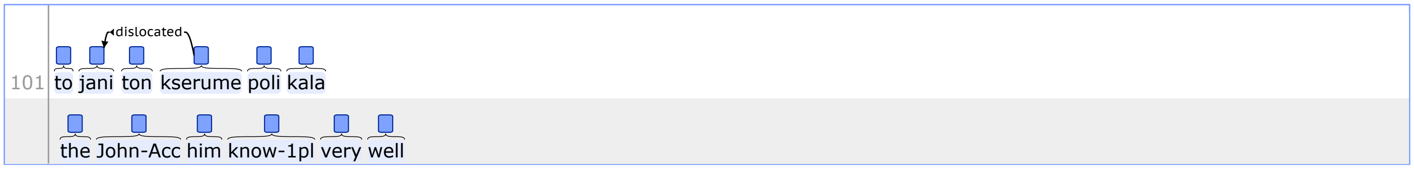
dislocated: dislocated elements

The `dislocated` relation is used for fronted or postposed elements that do not fulfill the usual core grammatical relations of a sentence. These elements often appear to be in the periphery of the sentence, and may be separated off with a comma intonation.

It is used for fronted elements that introduce the topic of a sentence, as in the following Japanese and Greek examples. The dislocated element attaches to the head of the clause to which it belongs:



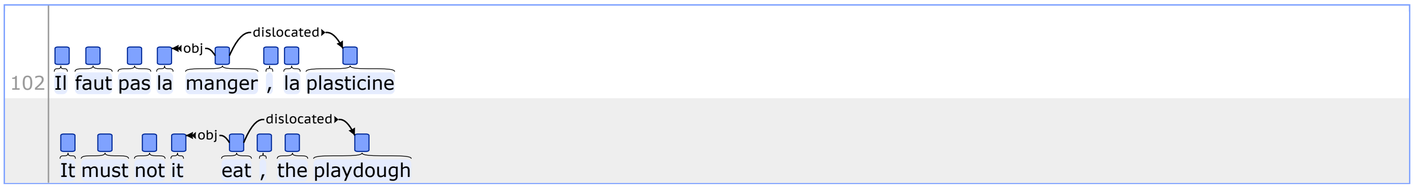
100 象 は 鼻 が 長い
zoo wa hana ga naga-i
elephant TOPIC nose SUBJ long-PRES



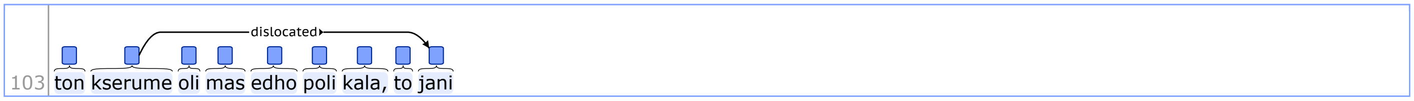
101 to jani ton kserume poli kala
the John-Acc him know-1pl very well

However, it would not be used for a topic-marked noun that is also the subject of the sentence; this would be an `nsubj`.

It is also used for postposed elements. The dislocated elements attach to the same governor as the dependent that they double for. Right dislocated elements are frequent in spoken languages. French and Greek examples follow.



102 Il faut pas la manger , la plasticine
It must not it eat , the playdough

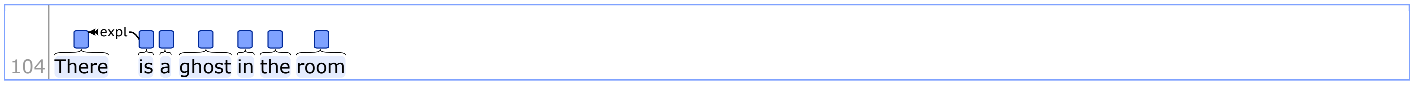


103 ton kserume oli mas edho poli kala, to jani

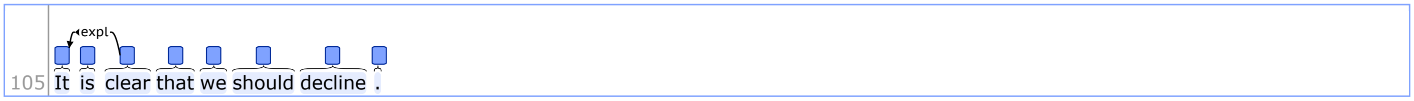
edit dislocated

expl: expletive

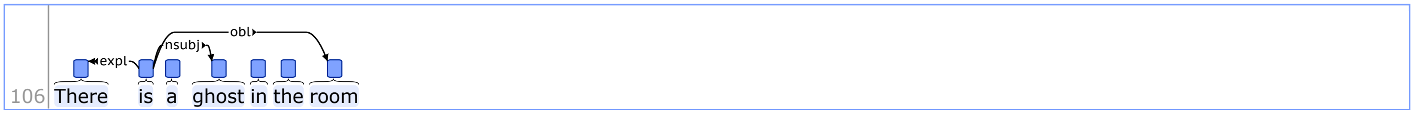
This relation captures expletive or pleonastic nominals. These are nominals that appear in an argument position of a predicate but which do not themselves satisfy any of the semantic roles of the predicate. The main predicate of the clause (the verb or predicate adjective or noun) is the governor. In English, this is the case for some uses of *it* and *there*: the existential *there*, and *it* when used in extraposition constructions. (Note that both *it* and *there* also have non-expletive uses.)

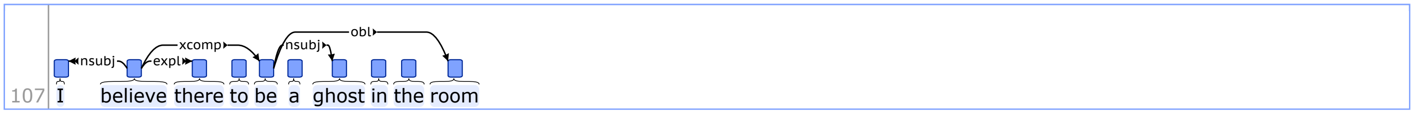


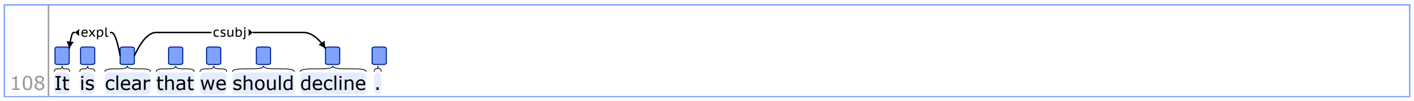
104 There is a ghost in the room

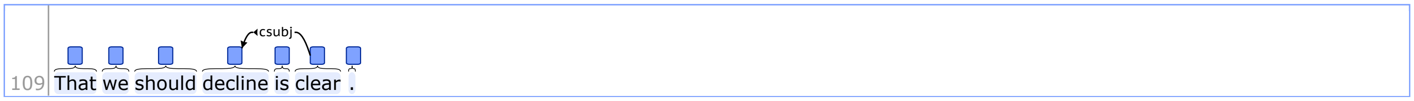


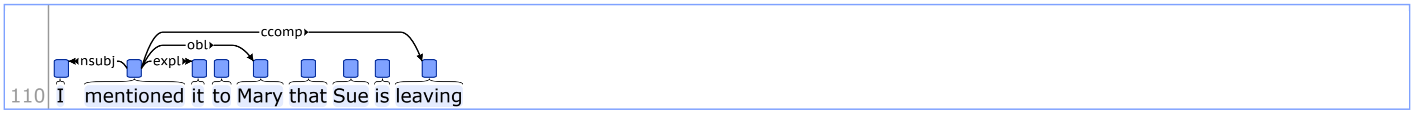
Some languages do not have expletives of the English sort, including most languages with free pro-drop (the ability to use zero anaphora rather than overt pronouns). In languages with expletives of this sort, they can be positioned where normally a core argument appears: the subject and direct object (and even indirect object) slots, as in the examples below. Note that in the analysis of these examples, we treat the postposed subject or clausal argument as a regular core argument, and mark the expletive with `expl`.



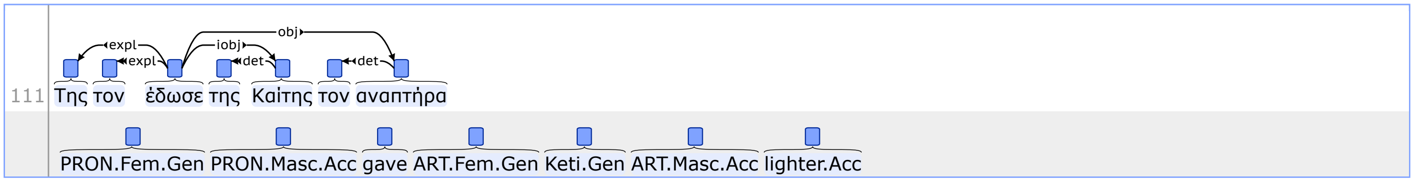


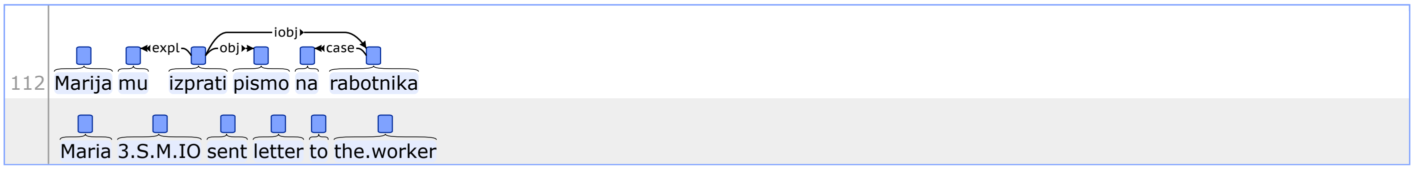






A second, related, use of the `expl` relation is for cases of true clitic doubling. For languages in which clitics and lexical nominals are ususally in complementary distribution – languages, such as French, which obey “Kayne’s generalization” – then whichever of a clitic or a lexical nominal occurs will get the appropriate role, such as `obj` or `iobj`. In such languages, when doubling does occur, such as in spoken French, the right analysis is to regard the lexical nominal as `dislocated` (see the examples there). As such, the analysis will be the same as when a noun phrase doubles another noun phrase or a regular pronoun that fills a nominal argument position. However, other languages, such as Greek and Bulgarian, standardly allow doubling of a lexical nominal and a pronominal clitic, with the former still appearing in its regular role as an argument of the predicate. In these cases, if only one of the lexical nominal and the clitic appear in a clause, then whichever appears will be given the grammatical role of `obj`, `iobj`, etc. – parallel to the treatment of lexical nominals and pronouns in other languages, modulo the clitic pronoun having a different position in the sentence. However, if both occur, the lexical nominal will be given the grammatical role of `obj`, `iobj`, etc., and the clitic will be treated as a pronominal copy, which does not receive its own semantic role, and hence will get the role `expl`. Modulo the different word order, this is fairly parallel to the treatment of *it* and *there* in English mentioned above, where another phrase satisfies the semantic role of the predicate. Examples from Greek and Bulgarian follow:


111
expl
expl
iobj
obj
det
det
Της τον έδωσε της Καίτης τον αναπτήρα
PRON.Fem.Gen PRON.Masc.Acc gave ART.Fem.Gen Keti.Gen ART.Masc.Acc lighter.Acc


112
expl
iobj
obj
case
Marija mu izprati pismo na rabotnika
Maria 3.S.M.IO sent letter to the.worker

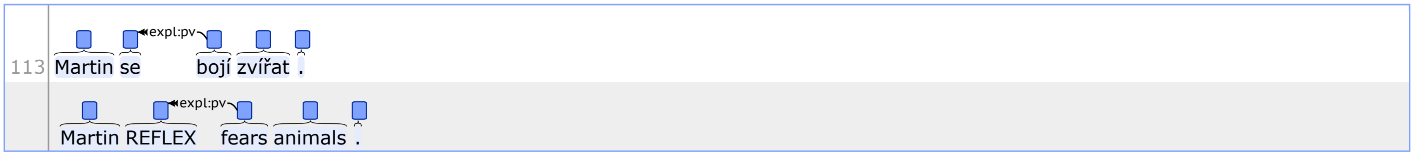
Reflexives

The expletive relation is also used for reflexive pronouns (see the feature `Reflex`) attached to inherently reflexive verbs, i.e. verbs that cannot occur without the reflexive pronoun and thus the pronoun does not play the role of a normal object (otherwise it would be possible to substitute it with an irreflexive pronoun or other nominal).

UD recognizes several functions of reflexive pronouns (clitics) that are usually distinguished with the help of subtypes of the `expl` relation (see also the report from the 2015 Uppsala discussion of clitics where this approach was approved):

- `expl:pv` for reflexive clitics attached to inherently reflexive verbs (also called pronominal verbs in some grammars)
- `expl:pass` for reflexive clitics attached to transitive verbs and acting as a voice marker (passive or mediopassive)
- `expl:impers` for impersonal usage (works also with intransitive verbs)

A Czech example:


113 Martin se boží zvířat .

Martin REFLEX fears animals .

Further general discussion of expletives can be found in Postal, P. M., and G. K. Pullum (1988) “Expletive Noun Phrases in Subcategorized Positions,” *Linguistic Inquiry* 19(4): 635–670. The status of clitic doubling, and arguments for the lexical nominal being an argument with the clitic a kind of pronominal copy, appear inter alia in Boris Harizanov (2014) Clitic doubling at the syntax-morphology interface: A-movement and morphological merger in Bulgarian. *Natural Language and Linguistic Theory*.

edit expl

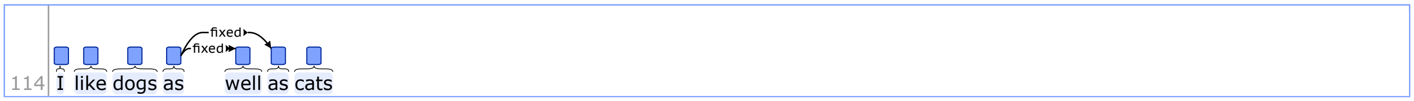
`fixed`: fixed multiword expression

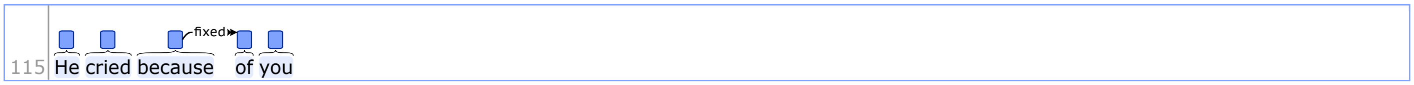
The `fixed` relation is one of the three relations for multiword expressions (MWEs) (the other two being `flat` and `compound`). It is used for certain fixed grammaticized expressions that behave like function words or short adverbials.

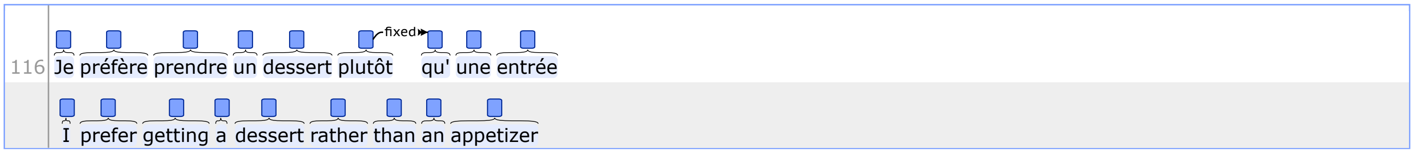
New from v2: The `fixed` relation replaces the old `mwe` relation to prevent misunderstanding regarding its scope.

The scope of `fixed` MWEs corresponds roughly to the *fixed expressions* category of Sag et al. and excludes any semi-fixed or flexible MWEs.

Fixed MWEs are annotated in a flat structure, where all subsequent words in the expression are attached to the first one using the `fixed` label. The assumption is that these expressions do not have any internal syntactic structure (except from a historical perspective) and that the structural annotation is in principle arbitrary. In practice, however, it is highly desirable to use a consistent annotation of all fixed MWEs in all languages.


114 I like dogs as well as cats


115 He cried because of you


116 Je préfère prendre un dessert plutôt qu' une entrée

I prefer getting a dessert rather than an appetizer

edit fixed

`flat`: flat multiword expression

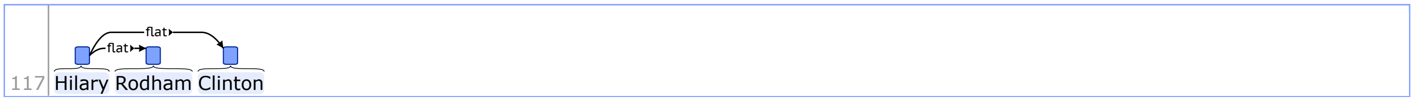
The `flat` relation is one of three relations for multiword expressions multiword expressions (MWEs) in UD (the other two being `fixed` and `compound`). It is used for exocentric (headless) semi-fixed MWEs like names (*Hillary Rodham Clinton*) and dates (*24 December*). It contrasts with `fixed`, which applies to completely fixed grammaticized (function word-like) MWEs (like *in spite of*), and with `compound`, which applies to endocentric (headed) MWEs (like *apple pie*).

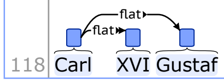
Flat MWEs are annotated with a flat structure, where all subsequent words in the expression are attached to the first one using the `flat` label. The assumption is that these expressions do not have any internal syntactic structure and that the structural annotation is in principle arbitrary. In practice, however, it is highly desirable to use a consistent annotation of all flat MWEs in all languages.

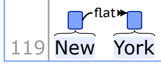
Below we describe some of the most common uses of `flat` across languages. Note that semantically equivalent expressions in different languages (or even in the same language) may require a different analysis if sometimes there is and sometimes there is not a regular compositional syntactic structure.

Names

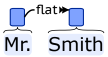
In many languages, there are multiword proper names with no clear internal syntactic structure and no clear evidence that one of the words is the syntactic head. Such names are annotated using the `flat` relation, with the optional subtype `flat:name`.

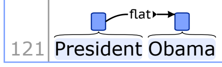

117 Hilary Rodham Clinton

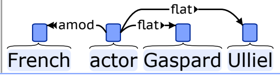
118


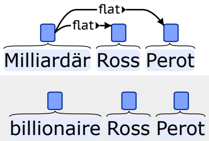
119


Titles/honorifics are also analyzed using the `flat` relation. Note that some titles are complex and have their own internal syntactic structure. Such structure is shown with regular relations embedded under `flat`:

120


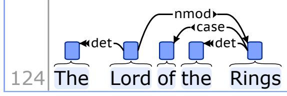
121


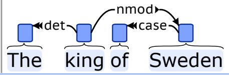
122


123


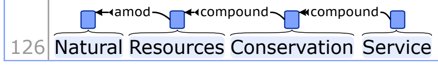
However if the two halves of a descriptive title and a name appear to be two separate nominals, then analysis with `flat` is not appropriate, and appos is appropriate. These cases are often set off by punctuation, such as a comma, but no punctuation may appear in more informal text. You can generally test for such examples by asking if the two halves can be reversed; if they can, it is probably an `appos`; see the examples there.

In contrast to the above, names that have a regular syntactic structure, like *The Lord of the Rings* and *Captured By Aliens*, should be annotated with regular syntactic relations.

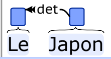
124


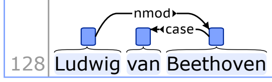
125


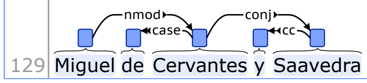
For organization names with clear syntactic modification structure, the dependencies should also reflect the syntactic modification structure using regular syntactic relations, as in:

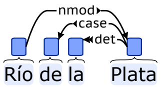
126


In addition, regular syntactic relations are used: (i) for a modifying determiner or similar function word and (ii) to connect together the words of a description or name which involve embedded prepositional phrases, sentences, etc., when these relations are (i) recognized in the language being annotated (i.e., the analyses below are for French, German, and Spanish, not English) and (ii) deemed not to be grammaticalized to the extent that the original role of the function words has been lost.

127


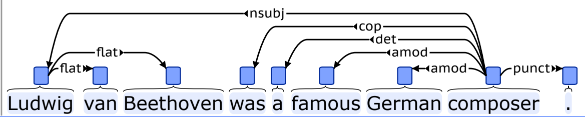
128


129


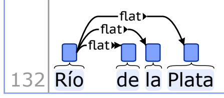
130


The above analyses of *Ludwig van Beethoven* and *Miguel de Cervantes y Saavedra* assume that *van* resp. *de* are prepositions. This is true in the languages of the names' origin, but it can be expected to change when the name is used in foreign text or when sufficient grammaticalization has taken place. For example, when names like this are annotated in English, the appropriate analysis is as a `flat` name:

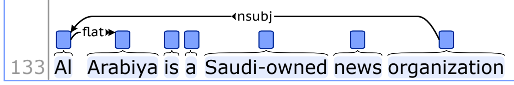
131



132

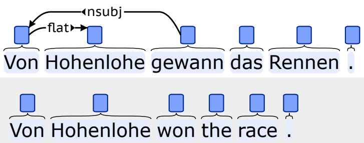


133



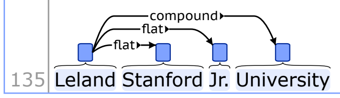
And in Modern German or French, these prepositions have generally just become a fossilized part of a family name and regularly appear without the given name. Again, here, analysis as `flat` seems correct:

134


nsubj
flat
Von Hohenlohe gewann das Rennen .
Von Hohenlohe won the race .

In the case of proper entities named after people, e.g. *Leland Stanford Jr. University*, the `flat` relation should only be used inside the person name, with the rest of the construction analyzed compositionally using normal syntactic relations:

135



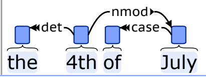
Some further notes on relations for names

This paragraph briefly records some of the arguments that have been made in the past on relations for name structure. It is an issue over which there has historically been variation and about which there is some continuing debate. Examples like *French actor Gaspard Ulliel:* Some treebanks have used `nmod` for titles and honorifics like *Mr.* or *French actor*. Most people think this is inappropriate, since an `nmod` dependent should be a full phrase, which will typically take its own case as a modifier in a cased language. In contrast, these titles seem to be part of the same phrase as the name that follows them; they show case agreement concord in a cased language. Some grammatical traditions, descending from Latin, call *French actor* in such cases a "fixed (or close) apposition" and take the name as the head. UD has restricted the `appos` relation to following appositives (corresponding to "loose (or wide) apposition" in the Latin tradition). The relation `appos` is only used when you have two full nominals, typically joined loosely, and often separated by a punctuation mark like a comma. So `appos` is not correct for these cases. Sometimes the relation `compound` has been used, but this does not seem right. It implies headedness, and titles do not usually behave like compounds: in German, they are not joined to the following words, as compounds are normally joined in German, and they appear at the beginning of names in both German and Hebrew, even though German compounds are head last and Hebrew compounds are head first. So `compound` does not seem appropriate either. Some UDv1 treebanks used `name` for honorifics like *Mr.*, although some felt that was wrong and `name` should be restricted to joining the proper nouns of multi-word names. In UDv2, `name` was removed and replaced by `flat`, which allowed a broader notion of a chunk of unheaded material. In the UDv2 guidelines, cases of both titles and honorifics are joined to names with `flat`.

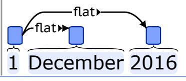
Dates and Complex Numerals

Date expressions come in many shapes and forms across languages. In some cases, they have a very clear syntactic structure, as in *the 4th of July*, and should be annotated with regular dependency relations. In other cases, they have a flat structure with no clearly discernible head, as in *1 December 2016*, in which case the `flat` relation should be used.

136

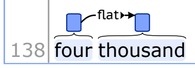


137



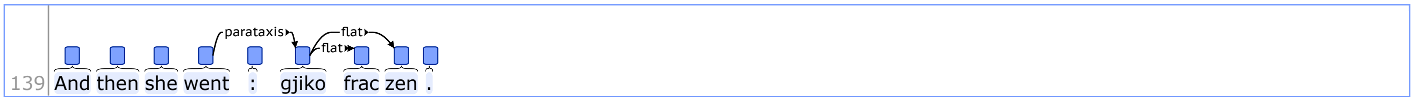
The `flat` relation can also be used for other numerals and other numerical expressions that lack phrasal structure.

138



Foreign Phrases

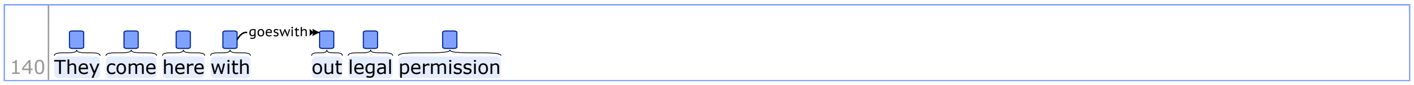
The `flat` relation, with the optional subtype `flat:foreign` should also be used when a foreign phrase cannot be given a compositional analysis. In this case, it replaces the `foreign` relation, which was used in v1 but is no longer part of the relation taxonomy.

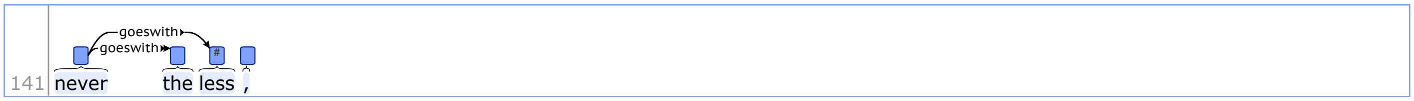


edit flat

`goeswith`: goes with

This relation links two or more parts of a word that are separated in text that is not well edited. These parts should be written together as one word according to the orthographic rules of a given language. The head is always the first part, the other parts are attached to it with the `goeswith` relation (for consistency, similarly as in `flat`, `fixed` and `conj`). Note that only the last part may be annotated with `SpaceAfter=No`.

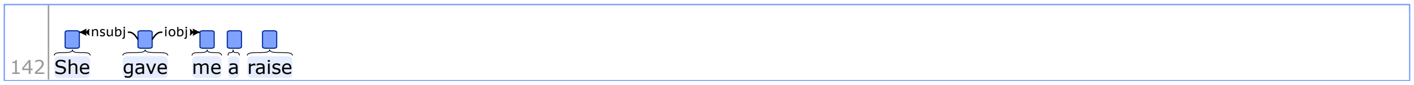




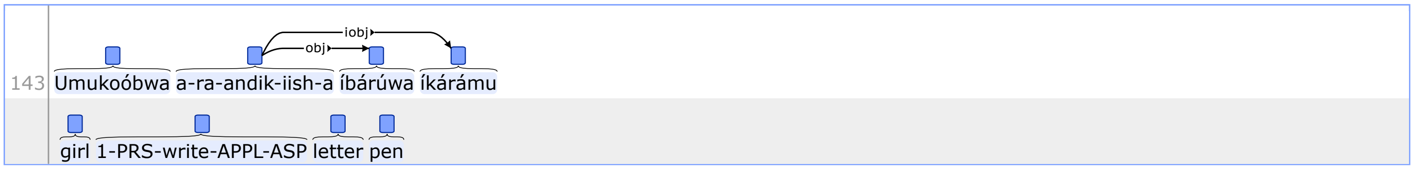
edit goeswith

`iobj`: indirect object

The indirect object of a verb is any nominal phrase that is a core argument of the verb but is not its subject or (direct) object. The prototypical example is the recipient of ditransitive verbs of exchange:

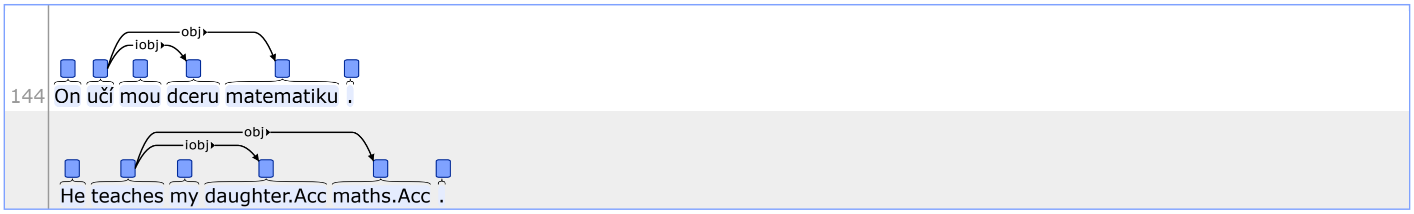


However, many languages allow other semantic roles as additional objects. The most common case is allowing benefactives, but some languages allow other roles. Examples include instruments, such as in the Kinyarwanda example below, or comitatives. At the other extreme, some languages lack all indirect objects.

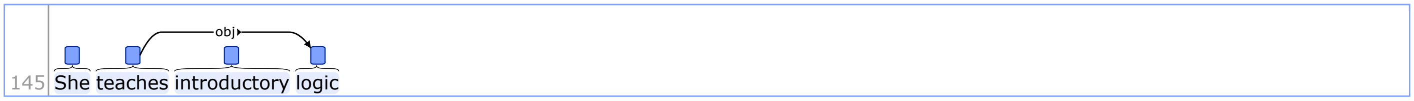

143
Umukoóbwa a-ra-andik-iish-a íbárúwa íkárámu
iobj
obj
girl 1-PRS-write-APPL-ASP letter pen

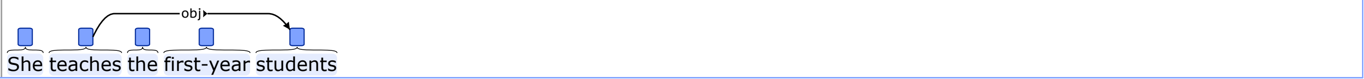
In languages distinguishing morphological cases, the indirect object will often be marked by the dative case. However, verb valency may occasionally dictate that the direct object is in dative, or that the indirect objects shall take various other forms.

In the following Czech example, the verb takes two arguments, both are nouns in the accusative case. One of them is direct object (patient), the other is indirect (addressee). It is parallel to how the English translation would be annotated (where there is no morphological case marking) and also to verbs of giving (consider a similar sentence, *he gave my daughter a class of maths*).


144
On učí mou dceru matematiku .
obj
iobj
He teaches my daughter.Acc maths.Acc .
obj
iobj

In general, if there is just one object, it should be labeled `obj`, regardless of the morphological case or semantic role. For example, in English, *teach* can take either the subject matter or the recipient as the only object, and in both cases it would be analyzed as the `obj`:



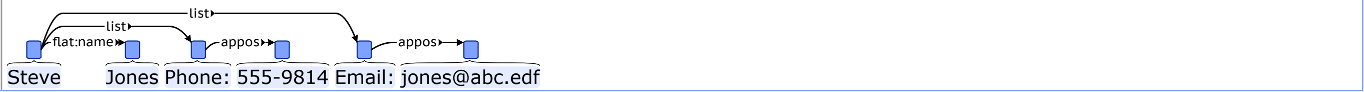
146


This is consistent with the analysis of Huddleston and Pullum (2002) “The Cambridge Grammar of the English Language”, chapter 4 section 4 (p. 251). As they note, it is no different to the same semantic role being sometimes the subject and sometimes the object in intransitive/transitive alternations.

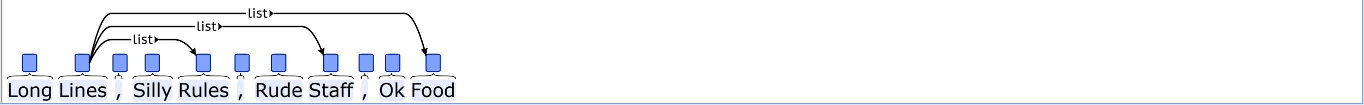
edit iobj

`list`: list

The `list` relation is used for chains of comparable items. In lists with more than two items, all items of the list should modify the first one. Informal and web text often contains passages which are meant to be interpreted as lists but are parsed as single sentences. Email signatures often contain these structures, in the form of contact information: the different contact information items are labeled as `list`; the key-value pair relations are labeled as `appos`.

147


Another place where `list` has been used is for a sequence of attributes or descriptive terms used as the title line of a review (such as product or restaurant reviews, etc.:

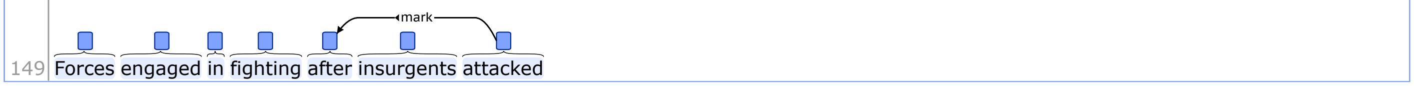
148


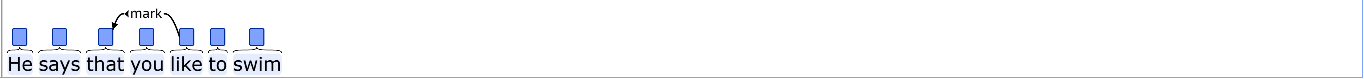
However, `list` should not be over-used. If a construction can be easily analyzed using the grammatical relations of standard sentences, such as when there is overt coordination, then it should be analyzed with these more standard relations, even if it is laid out as a list typographically.

edit list

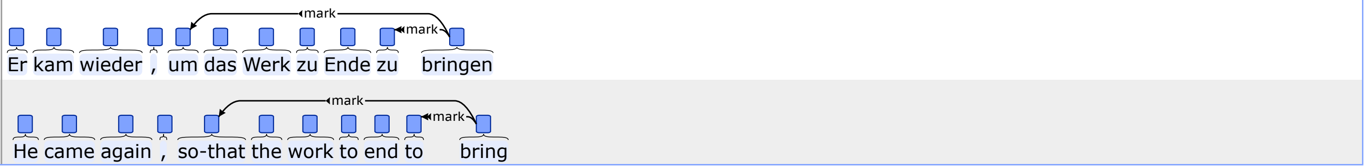
`mark`: marker

A marker is the word marking a clause as subordinate to another clause. For a complement clause, this is words like [en] *that* or *whether.* For an adverbial clause, the marker is typically a subordinating conjunction like [en] *while* or *although.* The marker is a dependent of the subordinate clause head. In a relative clause, it is a normally uninflected word, which simply introduces a relative clause, such as [he] *še.* (In this last use, one needs to distinguish between relative clause markers, which are `mark`, from relative pronouns such as [en] *who* or *that*, which fill a regular verbal argument or modifier grammatical relation.)

149


150


Infinitive markers (e.g. English *to,* German *zu*) in infinitival clauses are also attached as `mark`:

151

mark
mark
Er kam wieder , um das Werk zu Ende zu bringen
mark
mark
He came again , so-that the work to end to bring

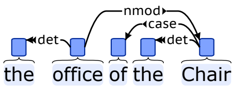
edit mark

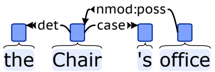
`nmod`: nominal modifier

The `nmod` relation is used for nominal dependents of another noun or noun phrase and functionally corresponds to an attribute, or genitive complement.

New from v2: The `nmod` relation was previously used also for nominal dependents of verbs, adjectives, and adverbs. The latter are now covered by the new `obl` relation.

In conjunction with the `case` relation, `nmod` provides a uniform analysis for the possessive alternation (with the option of a subtype like `nmod:poss` to distinguish non-adpositional case):

152


153


edit nmod

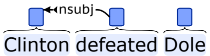
`nsubj`: nominal subject

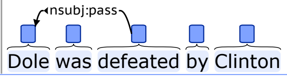
A nominal subject (`nsubj`) is a nominal which is the syntactic subject and the proto-agent of a clause. That is, it is in the position that passes typical grammatical test for subjecthood, and this argument is the more agentive, the do-er, or the proto-agent of the clause. This nominal may be headed by a noun, or it may be a pronoun or relative pronoun or, in ellipsis contexts, other things such as an adjective.

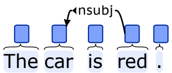
New from v2: The `nsubj` relation is also used for the nominal subject of a passive verb or verb group, even though the subject is then not typically the proto-agent argument due to valency changing operations. For languages that have a grammaticalized passive transformation, it is strongly recommended to use the subtype `nsubj:pass` in such cases.

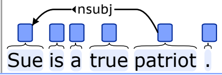
The governor of the `nsubj` relation might not always be a verb: when the verb is a copular verb, the root of the clause is the complement of the copular verb, which can be an adjective or noun, including a noun marked by a preposition, as in the examples below.

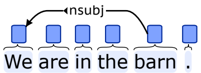
The `nsubj` role is only applied to semantic arguments of a predicate. When there is an empty argument in a grammatical subject position (sometimes called a pleonastic or expletive), it is labeled as `expl`. If there is then a displaced subject in the clause, as in the English existential *there* construction, it will be labeled as `nsubj`.)

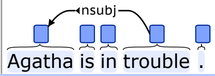
154


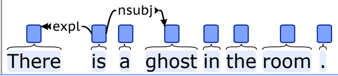
155


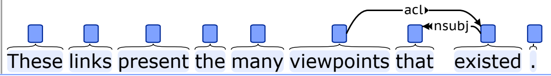
156


157


158


159


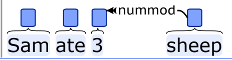
160


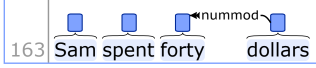
161


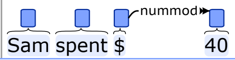
edit nsubj

`nummod`: numeric modifier

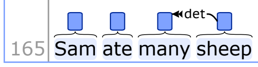
A numeric modifier of a noun is any number phrase that serves to modify the meaning of the noun with a quantity.

162 

163 

164 

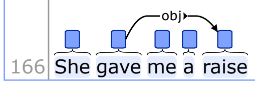
Note that indefinite quantifiers such as *few, many* are tagged `DET` rather than `NUM`. Therefore their relation to the quantified noun is not `nummod` but `det`:

165 

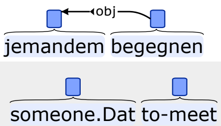
edit nummod

`obj`: object

The object of a verb is the second most core argument of a verb after the subject. Typically, it is the noun phrase that denotes the entity acted upon or which undergoes a change of state or motion (the proto-patient).

166 

In languages distinguishing morphological cases, the object will often be marked by the accusative case. However, verb valency may occasionally dictate a different form, such as the dative case in the following German example:

167 

In general, if there is just one object, it should be labeled `obj`, regardless of the morphological case or semantic role that it bears. If there are two or more objects, one of them should be `obj` and the others should be `iobj`. In such cases it is necessary to decide what is the most directly affected object *(patient).*

There is more discussion of constructions with multiple objects on the page for `iobj`. If possible, language-specific documentation should be available to help identify the primary (or direct) object.

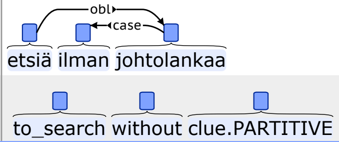
edit obj

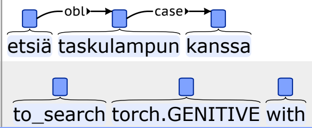
`obl`: oblique nominal

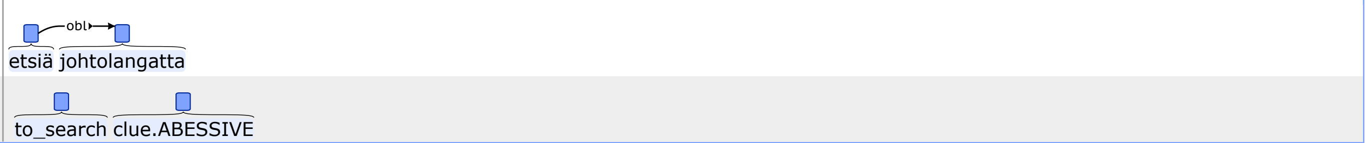
The `obl` relation is used for a nominal (noun, pronoun, noun phrase) functioning as a non-core (oblique) argument or adjunct. This means that it functionally corresponds to an adverbial attaching to a verb, adjective or other adverb.

The `obl` relation can be further specified by the case. In conjunction with the `case` relation, it provides a uniform analysis for:

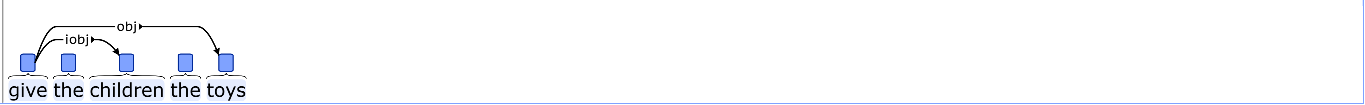
- variant forms with case, a preposition or a postposition, as in Finnish for example:

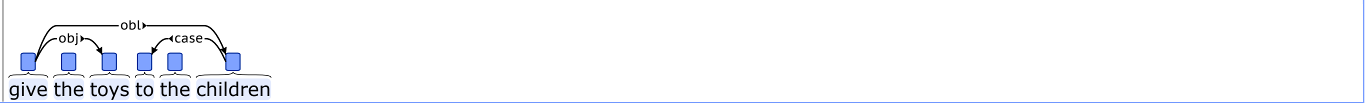
168 

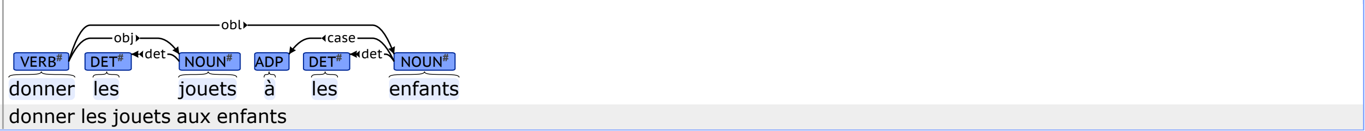
169 

170


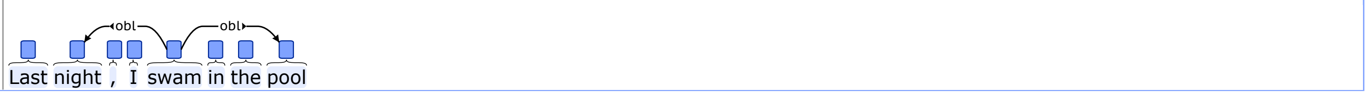
- the dative alternation where the prepositional construction gets a similar analysis to the double object construction:

171


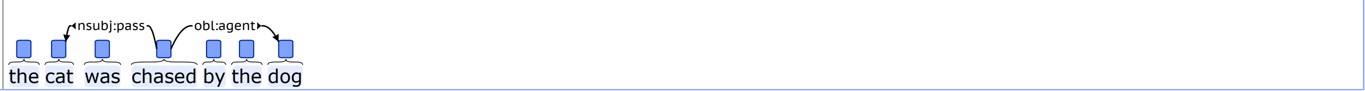
172


173

obl
obj
case
det
det
VERB
DET
NOUN
ADP
DET
NOUN
donner les jouets à les enfants
donner les jouets aux enfants

`obl` is also used for temporal and locational nominal modifiers:

174


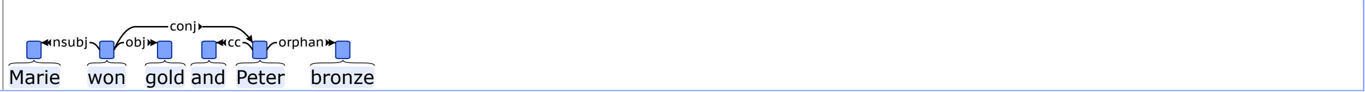
and for the agent of a passive verb (with the optional subtype obl:agent):

175


edit obl

`orphan`: orphan

The ‘orphan’ relation is used in cases of head ellipsis where simple promotion would result in an unnatural and misleading dependency relation. The typical case is predicate ellipsis where one of the core arguments has to be promoted to clausal head.

176


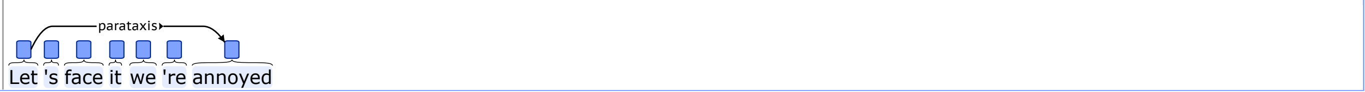
In this example, the subject *Peter* is promoted to the head position in the second conjunct. Attaching the object *bronze* to the subject is necessary to preserve the integrity of the clause, but using the standard relation `obj` would be misleading because *bronze* is not the object of *Peter*. Therefore, the `orphan` relation is used to indicate that this is a non-standard attachment. By contrast, the coordinating conjunction *and* performs essentially the same function as in the non-elliptical case and therefore retains its normal relation `cc`.

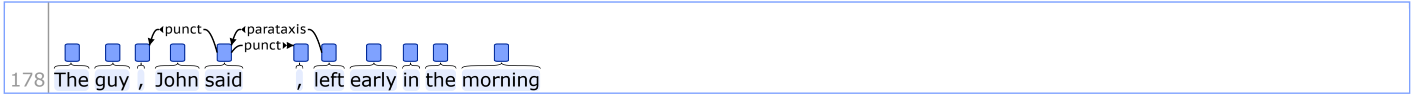
See further discussion of ellipsis.

edit orphan

`parataxis`: parataxis

The parataxis relation (from Greek for “place side by side”) is a relation between a word (often the main predicate of a sentence) and other elements, such as a sentential parenthetical or a clause after a “:” or a “;”, placed side by side without any explicit coordination, subordination, or argument relation with the head word. Parataxis is a discourse-like equivalent of coordination, and so usually obeys an iconic ordering. Hence it is normal for the first part of a sentence to be the head and the second part to be the parataxis dependent, regardless of the headedness properties of the language. But things do get more complicated, such as cases of parentheticals, which appear medially.

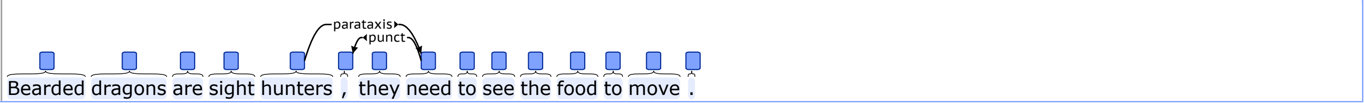
177


178


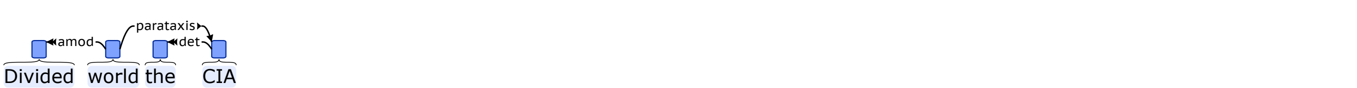
An inventory of constructions to which parataxis has been applied

Side-by-side sentences ("run-on sentences")

The relation parataxis is used for a pair of what could have been standalone sentences, but which are being treated together as a single sentence. This may happen because sentence segmentation of the sentence was done primarily following the presence of sentence-final punctuation, and these clauses are joined by punctuation such as a colon or comma, or not delimited by punctuation at all. In a spoken corpus, it may happen because what is labeled as a sentence is more commonly an utterance turn. Even if the treebanker is doing the sentence division, it may happen because there seems to be a clear discourse relation linking two clauses. Sometimes there are more than two sentences joined in this way. In this case we make all the later sentences dependents of the first one, to maximize similarity to the analysis used for conjunction.

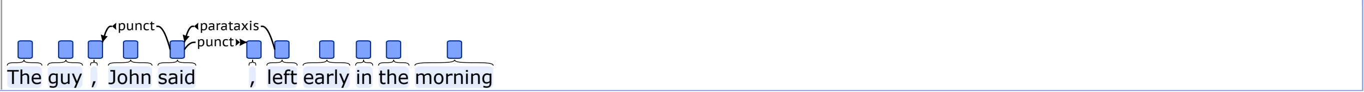
179

parataxis
punct
Bearded dragons are sight hunters , they need to see the food to move .

This relation may happen with units that are smaller than sentences:

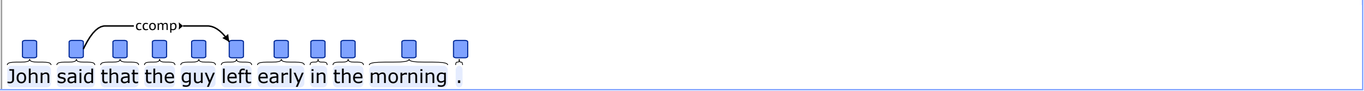
180


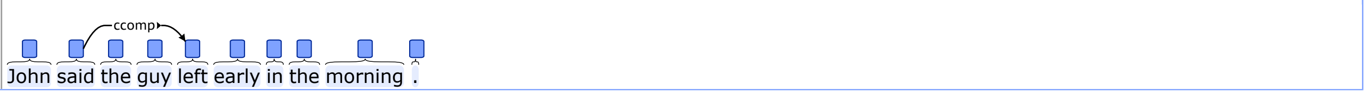
Treatment of reported speech

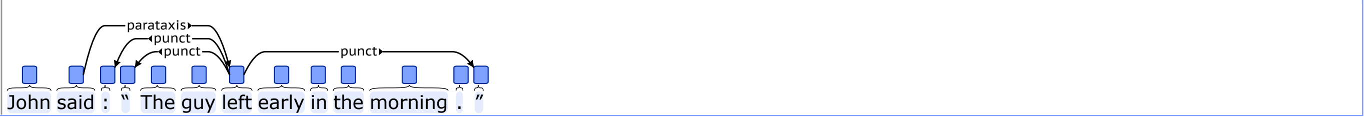
For this reported speech example:

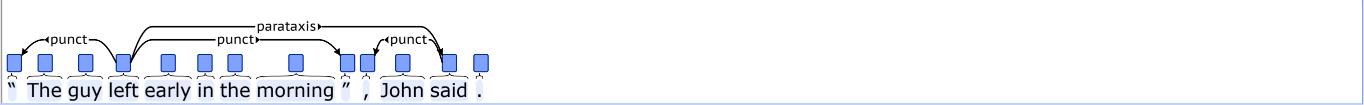
181


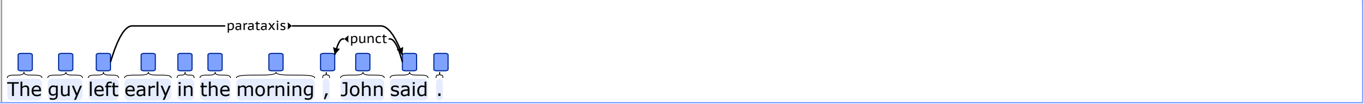
there are paraphrases that convey essentially the same meaning but with a different syntactic structure. When the reported speech is embedded in a subordinate clause (with or without an overt complementizer *that*), the subordinate clause is a ccomp of the speech verb. When the reported speech follows the speech verb and is separated by a colon, the reported speech forms a main clause that attaches to the preceding main clause with a parataxis relation, hence with the speech verb as its head. However, when the speech verb occurs as a medial or final parenthetical, the relation is reversed and the speech verb is treated as a parataxis of the reported speech. This analysis is not uncontroversial but follows many authorities, such as Huddleston and Pullum (2002), *The Cambridge Grammar of the English Language* (see chapter 11, section 9).

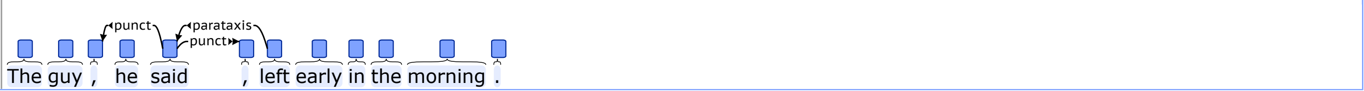
182


183


184


185


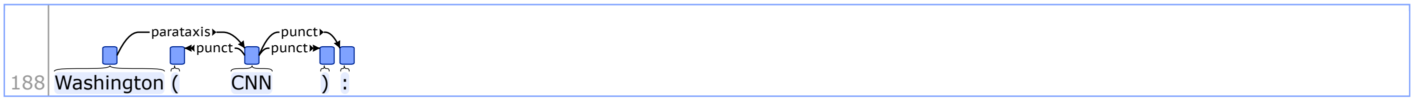
186


187


An argument for this analysis is that in the cases analyzed as embedding, the entire clause can be further embedded (*I was taken aback when John said the guy left early in the morning.*), while this is not possible with medial or final placement of the speech verb (**I was taken aback when the guy left early this morning, John said.*).

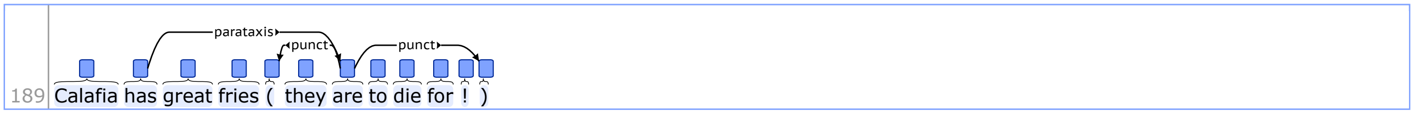
News article bylines

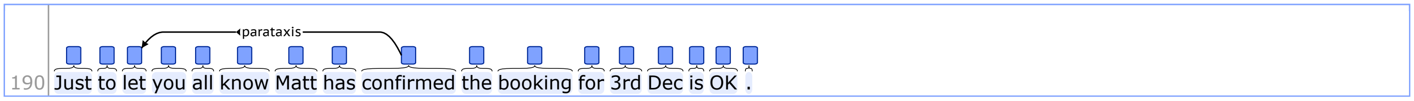
We have used the parataxis relation to connect the parts of a news article byline. There does not seem to be a better relation to use.



Interjected clauses

Single word or phrase interjections are analyzed as discourse, but when a whole clause is interjected, we use the relation parataxis.

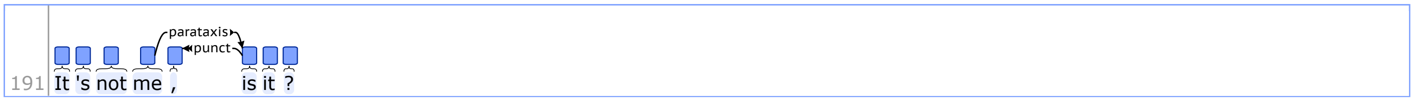



190
Just to let you all know Matt has confirmed the booking for 3rd Dec is OK .
parataxis

In the second example, we treat the second half as the head of the dependency because the first half feels like a whole clause interjection, not like the main clause of the utterance.

Tag questions

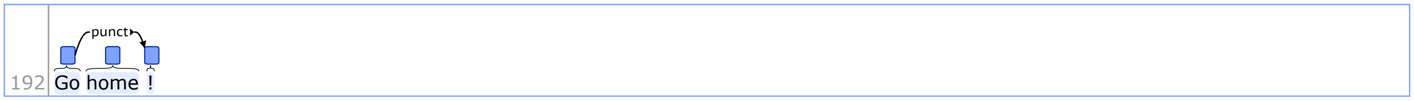
We also use the parataxis relation for tag questions such as *isn't it?* or *haven't you?*.



edit parataxis

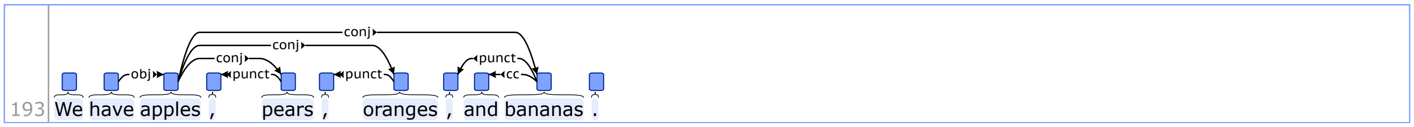
punct: punctuation

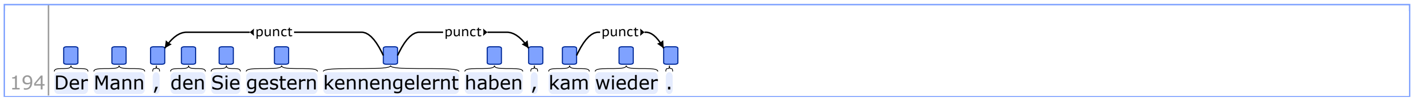
This is used for any piece of punctuation in a clause, if punctuation is being retained in the typed dependencies. Note that symbols tagged SYM are not punctuation and cannot be attached via the punct relation.

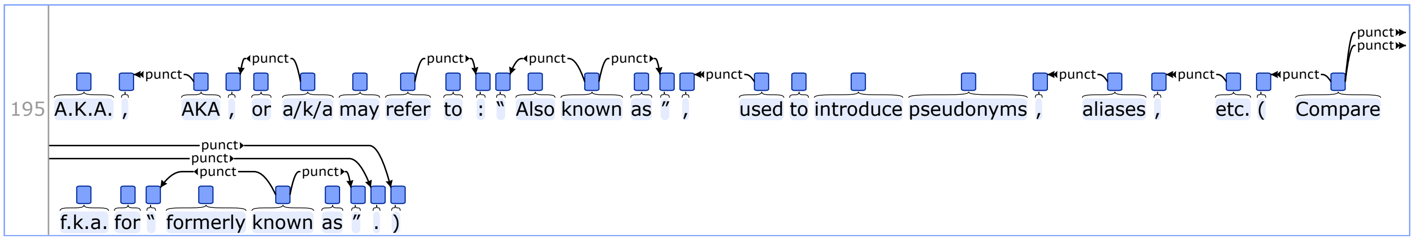


Tokens with the relation punct always attach to content words (except in cases of ellipsis) and can never have dependents. Since punct is not a normal dependency relation, the usual criteria for determining the head word do not apply. Instead, we use the following principles:

1. A punctuation mark separating coordinated units is attached to the following conjunct.
2. A punctuation mark preceding or following a dependent unit is attached to that unit.
3. Within the relevant unit, a punctuation mark is attached at the highest possible node that preserves projectivity.
4. Paired punctuation marks (e.g. quotes and brackets, sometimes also dashes, commas and other) should be attached to the same word unless that would create non-projectivity. This word is usually the head of the phrase enclosed in the paired punctuation.




194
Der Mann , den Sie gestern kennengelernt haben , kam wieder .
punct
punct
punct

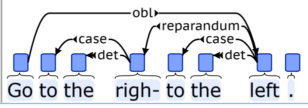

195
A.K.A. , AKA , or a/k/a may refer to : " Also known as " , used to introduce pseudonyms , aliases , etc. (Compare
f.k.a. for " formerly known as " .)
punct
punct
punct
punct
punct
punct
punct
punct
punct
punct
punct
punct
punct
punct
punct
punct

See also examples at parataxis.

edit punct

`reparandum`: overridden disfluency

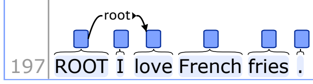
We use `reparandum` to indicate disfluencies overridden in a speech repair. The disfluency is the dependent of the repair.

196


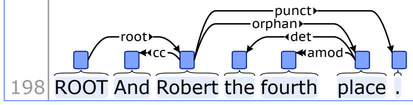
edit reparandum

`root`: root

The `root` grammatical relation points to the root of the sentence. A fake node `ROOT` is used as the governor. The `ROOT` node is indexed with 0, since the indexing of real words in the sentence starts at 1. (The `ROOT` node is not represented explicitly in CoNLL-U.)

197


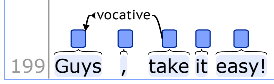
New from v2: There should be just one node with the `root` dependency relation in every tree. If the main predicate is not present (due to ellipsis) and there are multiple orphaned dependents, one of these is promoted to the head (root) position and the other orphans are attached to it. (This rule has in practice been followed since release v1.2 but was not explicitly stated in the original v1 guidelines.)

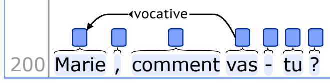
198


edit root

`vocative`: vocative

The *vocative* relation is used to mark a dialogue participant addressed in a text (common in conversations, dialogue, emails, newsgroup postings, etc.). The relation links the addressee's name to its host sentence. A vocative commonly co-occurs with a null subject, as in the first example below. If the nominal is clearly vocative in intent, the preference is to use the vocative relation.

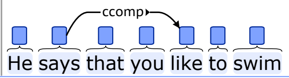
199


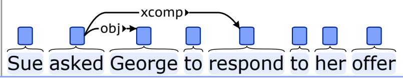
200


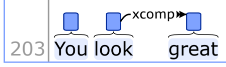
edit vocative

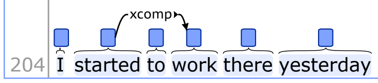
`xcomp`: open clausal complement

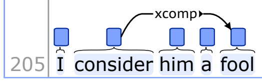
An open clausal complement (`xcomp`) of a verb or an adjective is a predicative or clausal complement without its own subject. The reference of the subject is necessarily determined by an argument external to the xcomp (normally by the object of the next higher clause, if there is one, or else by the subject of the next higher clause). This is often referred to as *obligatory control*. These clauses tend to be non-finite in many languages, but they can be finite as well. The name `xcomp` is borrowed from Lexical-Functional Grammar.

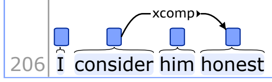
201


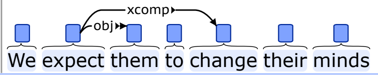
202


203


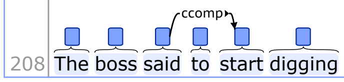
204


205


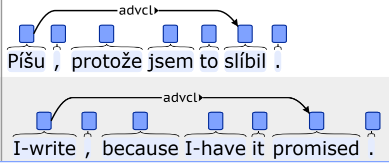
206


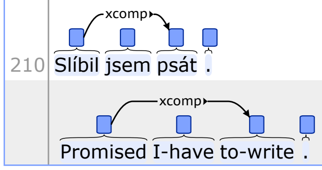
207


Note that the above condition “without its own subject” does not mean that a clause is an `xcomp` just because its subject is not *overt.* The subject must be necessarily inherited from a fixed position in the higher clause. That is, there should be no available interpretation where the subject of the lower clause may be distinct from the specified role of the upper clause. In cases where the missing subject may or must be distinct from a fixed role in the higher clause, `ccomp` should be used instead, as below. This includes cases of arbitrary subjects and anaphoric control.

208


Pro-drop languages have clauses where the subject is not present as a separate word, yet it is inherently present (and often deducible from the form of the verb) and it does not depend on arguments from a higher clause. Thus in neither of the following two Czech examples is there any overt subject, yet only the second example contains an `xcomp`.

209

advcl
Píšu , protože jsem to slíbil .
advcl
I-write , because I-have it promised .

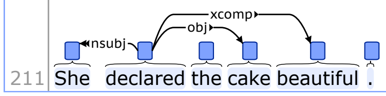
210


Secondary Predicates

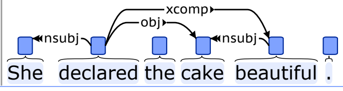
The `xcomp` relation is also used in constructions that are known as *secondary predicates* or *predicatives.* Examples:

- *She declared the cake beautiful.*
- *She declared the cake a success.*

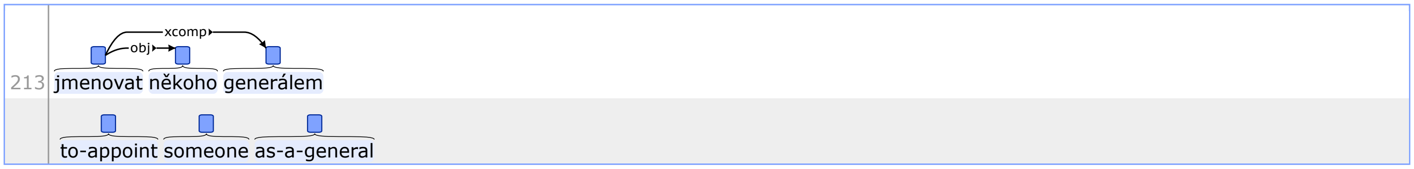
We could paraphrase the sentence using a subordinate clause: *She declared that the cake was beautiful.* There are two predicates mixed in one clause: 1. she declared something, and 2. the cake was beautiful (according to her opinion). The secondary predicate will be attached to the main predicate as an `xcomp`:

211


In the enhanced representation, there is an additional subject link showing the secondary predication:

212


A Czech example:



Remember that `xcomp` is used for core arguments of clausal predicates so it will not be used for other instances of secondary predication. For instance, in *She entered the room sad* we also have a double predication (she entered the room; she was sad). But *sad* is not a core argument of *enter:* leaving it out will neither affect grammaticality nor significantly alter the meaning of the verb. On the other hand, leaving out *beautiful* in *she declared the cake beautiful* will either render the sentence ungrammatical or lead to a different interpretation of *declared.*

The result is that in *She entered the room sad, sad* will depend on *She* and the relation will be `acl` instead of `xcomp`.

edit xcomp