

TOROT Guidelines for Annotation

Hanne Martine Eckhoff

20th April 2015

Contents

1 Foreword	5
2 Lemmatisation	5
3 Parts of speech	6
4 Morphological issues	7
5 Introducing the syntactic annotation scheme	7
6 The idea of dependency grammar	8
7 Sentence boundaries	8
8 Tokenisation	9
9 The verb	11
10 Exclamations	17
11 Non-verbal sentence-level grammatical relations	17
11.1 Subject	17
11.2 Object	20
11.3 Obliques	25
11.4 AG: Demoted subjects in passive constructions	27
11.5 Adverbials	33
11.6 Predicative complements	33
11.7 Adverbial appositions	37
11.8 Supertags	42
11.9 Auxiliary words	45
12 Noun phrases without nouns	45

13 Adnominal tags	48
13.1 General	48
13.2 Attributes	50
13.3 Appositions	50
13.4 Partitives	56
13.5 Arguments of nouns	60
14 Embedded predication	64
14.1 The basic types of subordinate clauses	64
14.1.1 Substantival clauses	64
14.1.2 Adjectival clauses	66
14.1.3 Adverbial clauses	68
14.2 Subordinate clauses with subjunctions	68
14.3 Finite subordinate clauses without subjunctions	68
14.4 Relative clauses	68
14.5 Indirect questions	79
14.5.1 Indirect questions vs. relative clauses vs. clauses with subjunctions	79
14.6 Non-finite predication: Accusative with infinitive and dative with infinitive	84
14.7 Infinite predication: absolute constructions	86
14.8 Open predications	86
14.8.1 XOBJ	88
14.8.2 XADV	91
14.9 Supertags	95
15 Coordination	97
15.1 Coordination of constituents/single-rooted elements	98
15.2 Sentence coordination	101
15.3 Coordination of non-constituents/multi-rooted elements	106
15.4 In which conjunct does an element belong?	106
15.5 Other issues	106
16 Gapping and ellipsis	106
16.1 Elided copulas	106
16.2 Gapping	111
17 The article	111
17.1 OCS mock article	111
18 Slash notation	115
18.1 Introduction	115
18.2 Open functions: XADV and XOBJ	116
18.3 Slashes in gapping and coordination	119
18.3.1 Motivation	119
18.3.2 Slashes in sentence coordination	120

18.3.3	When is an adverbial shared?	124
18.3.4	Slashes in gapping/multi-rooted conjuncts	124
18.3.5	In which conjunct does a shared element belong?	124
19	Other constructions	125
19.1	‘all’, ‘self’ and similar words	125
19.2	Headedness of compound verb forms	125
19.3	Comparison	125
19.4	Verbs of preference	125
19.5	‘Approximately’	125
19.6	Particle coordination	125
19.7	Scope	125
19.8	Impersonal verbs	126
19.9	esse + participle	126
19.10	Gerunds	126
19.11	Infinitives	126
19.12	Reflexives	126
19.13	Datives	127
19.14	Compound subjunctions	127
19.15	Appositive names	127
19.16	Copulae, copula-like verbs and location verbs: <i>byti</i> etc.	127
19.17	Vocatives	127
19.18	And it came to pass...	127
19.19	Copula, verb of position, verb of existence	127
19.20	Resumptive pronouns and topicalizations	127
19.21	AUX on pronouns – <i>iže ašte</i>	127
19.22	Adnominal infinitives	127
19.23	Impersonal temporal expressions	128
19.24	It happened to him...	128
19.25	‘if not’	128
19.26	COMP and OBJ in the same sentence	128
19.27	Questions with alternatives	128
20	General issues	128
20.1	ADV or OBL? adjunct or argument?	128
20.2	ATR or APOS? attribute or apposition?	129
20.3	ATR or XADV?	129
20.4	XADV or SUB?	129
20.5	More on the XOBJ relation	129
20.6	Empty nodes	129
21	Error messages during syntactic annotation	129
21.1	‘Must have or inherit one outgoing slash edge’	129
21.2	‘May not be a daughter of the root node’	130
21.3	‘Subgraphs overlap’	130
21.4	‘Slashes are not contained by subgraph’	130

21.5	‘The head of a PARPRED/VOC relation must be the root node or a valid coordination’	130
21.6	‘Slash must target the node’s head or a node dominated by the head’	131
21.7	‘The head of a PRED relation must be the root node, a subjunction or a valid coordination’	131
21.8	‘A subjunction may only be the dependent in a COMP, ADV or APOS relation’	131

1 Foreword

The TOROT corpus (Tromsø Old Russian and Old Church Slavonic Treebank) is a member of the larger PROIEL family of treebanks of ancient languages. The corpus continues the work done on Slavic in the PROIEL project itself. While that work was my responsibility in the PROIEL project period, all our Slavic annotation is wholly in line with the general PROIEL guidelines for annotation¹, which were developed by the PROIEL team and supervised and written up by Dag Haug. This text is greatly indebted to the PROIEL guidelines text, and sometimes quotes long passages verbatim. The purpose of a separate set of Slavic guidelines is primarily to document better the issues pertaining specifically to Slavic, both concerning lemmatisation, morphology and syntax, and to give Slavic examples throughout. For issues pertaining to the other PROIEL languages, we refer to the original PROIEL guidelines.

2 Lemmatisation

In Old Church Slavonic, lemmatisation is fairly straightforward. We follow the orthography of Cejtin et al.’s *Staroslavjanskij slovar’*, and consult the four-volume *Slovník jazyka staroslověnského* whenever necessary. When Cejtin et al. give several forms for an entry, we use the first one, unless the two forms belong to different categories (e.g. gender, conjugation class).

Old Church Slavonic

Old and Middle Russian provide much more of a challenge, since the dictionary resources are poorer and the material much bigger. We have therefore been required to do work of an essentially lexicographical nature. We have decided to lemmatise Old and Middle Russian together to ease diachronic studies. However, we have settled for a fairly conservative lemma orthography, with jers in situ, which clearly suits the Old Russian material better than the Middle Russian one, with some fairly artificial lemma forms as a result.

Old and Middle Russian

We lemmatise Old Russian after Sreznevskij’s *Materialy dlja slovarja drevnerusskogo jazyka*. This means that we adhere to Sreznevskij in spelling (with simplifications, see below), and also use the dictionary as a support in deciding whether something should be a separate lemma at all. However, Sreznevskij does omit large groups of words (such as proper names and names of ethnic groups), and in these cases we must attempt to establish lemmata that follow Sreznevskij’s spelling principles. *Slovar’ russkogo jazyka XI–XVII vv.* should be consulted as well, but note that the lemma spellings will often not comply to our principles.

Old Russian orthography

We have decided to simplify Sreznevskij’s orthography in the following ways:

- the ja ligature is always written я
- Ѧ, ѩ are also simplified to я
- all ou ligatures are simplified to y

¹http://folk.uio.no/daghaug/syntactic_guidelines.pdf

- ѣ is simplified to ю
- ы is simplified to ы
- ѐ is simplified to е

We have decided to try to reflect at least some of the South Slavic/East Slavic variability in the Old Russian (and hence also Middle Russian) lemma inventory. The rule of thumb is that we follow Sreznevskij if he has an entry for the word form in question (with the orthographical modifications listed here). There are some exceptions and modifications to this rule:

**East and
South Slavic
variation**

1. If Sreznevskij lists a South Slavic-style lemma and an East Slavic-style lemma, we should use the East Slavic-style lemma if there is clear evidence that the stem is East Slavic in form (e.g. has polnoglasie, East Slavic palatalisation reflexes). If there is no evidence of East Slavic features in the stem, use the South Slavic lemma.
2. If Sreznevskij lists a single lemma, we use that lemma unless there is evidence of South/East Slavic stem variation. If we find such variation, we posit a South Slavic-style lemma and an East Slavic-style lemma, and only use the latter if the stem has clear East Slavic features.
3. If an attested form has a stem with both South and East Slavic features, we prefer the South Slavic lemma.
4. The form of case/number/gender/person/tense inflectional endings should not affect our choice of lemma.

This rule will mean that we overgeneralise the use of South Slavic-style lemmas. One advantage of doing this is that the lemma inventory for Old Russian will be more readily comparable with the OCS inventory. We will clearly not capture all the variation in the material by this rule, but further refinements may be done by additional lemma- and token-level tagging

The Old Russian aspect system is in flux during the period under consideration, and there are a number of tense forms where aspectual pair verbs can be indistinguishable. In such cases we should always go with the most likely analysis. Prefixed perfective verbs will rarely occur in the imperfect or as present participles, conversely derived imperfective verbs will rarely occur in the aorist or as past participles, so these rare combinations should not be posited in ambiguous cases. So, for example, in *егда избавашеть ихъ. паки оукланяхуться. на бѣсслуженье* (PVL 97.4–5, 152806) the form *избавашеть* should be analysed as an imperfect form of *избавляти* and not of *избавити*.

**Aspectual
pairs**

3 Parts of speech

Each lemma has four features: form, part of speech, language and (when necessary) variant number. A difference in at least one of these features will result in

two separate lemmas. Thus, we can have two lemmas with the same form, but different part of speech, two lemmas with the same form and part of speech, but different language, two lemmas with the same form, part of speech and language, but different variant numbers. The part of speech is determined by syntactic function. For example, the OCS lemma *jako* occurs both as a subjunction (introducing complement and adverbial clauses), a relative adverb (“in the same way as”) and as an adverb, the latter with two different variant numbers (introductory “for” vs. the comparison word “as”).

4 Morphological issues

We disambiguate morphology by syntactic and semantic function. However, when we see unexpected forms, e.g. agreement mismatches, our policy is to give the form in question its face value.

5 Introducing the syntactic annotation scheme

The system of annotation presented here is based on dependency grammar enriched with secondary dependencies (slashes, see section 18) reminiscent of the structure sharing mechanism in Lexical-Functional Grammar. Much of the scheme is adapted from the *Guidelines for the Syntactic Annotation of Latin Treebanks (v. 1.3)*² and its ultimate source, the *Guidelines for Annotations at Analytical Level of the Prague Dependency Treebank*.³ However, there are several important differences.

Since we use a dependency grammar, word order is not modelled at all in our syntactic trees. The information about word order is stored in a separate layer instead, where each word is given a linearisation index corresponding to its position in the sentence. This means that there is no left-to-right ordering in our trees. It is decided by the dependency tree grapher and is purely conventional. If an annotation is changed in some minor way, this may result in the dependency grapher deciding to change the graphical representation of the analysis tree.

Word order

This also means that annotators can work in the sequence they prefer. Dependents of the same node can be added in any order: it is not necessary to add the subject, the object and other arguments in any particular order.

Note that in many cases it is not obvious what the correct syntactic analysis might be. Not infrequently, we lack the information to weigh one analysis properly against another. This means that we have settled for a number of solutions that are not necessarily ideal, possibly even incorrect. In these cases, our main concern is retrievability: that data on a particular construction are easy to retrieve, is more important than achieving the ideal analysis. Another concern is the need for consistent annotation: If a distinction, even though well-motivated, is difficult for the annotators to tackle in a consistent way, we may

²<http://nlp.perseus.tufts.edu/syntax/treebank/1.3/docs/guidelines.pdf>

³<http://ufal.mff.cuni.cz/pdt2.0/doc/manuals/en/a-layer/html/index.html>

choose to ignore that distinction.

6 The idea of dependency grammar

Dependency grammar sets out to model the syntactic structure of sentences as a set of *relations* between words. These relations are called *dependencies*. It is notoriously hard to define a syntactic dependency, but as a basis for an intuitive grasp of the concept, a word X is dependent on a word Y just in case, if you remove Y from the sentence, you also have to remove X.

These relations must be *antisymmetric*, which is why they are referred to as dependencies: if X depends Y, Y cannot depend on X. Sometimes it would be tempting to analyse words as mutually dependent. For example, it is the case in many languages that a preposition cannot occur without its complement; nor can the complement occur without the preposition. Most dependency grammarians do not tolerate such mutual dependencies.

Note that in many cases, a dependent does not simply modify its mother node, but has scope over the entire subtree headed by that node. This is typically the case for negations, and also for instance for modifiers of complex numeral constructions. In the text we often refer to the head as a shorthand for the entire subtree.

7 Sentence boundaries

The division of the text into sentences is determined by the annotators. The text is imported into the annotation interface with sentence units based on punctuation in the source text, but in most cases this punctuation corresponds to units considerably smaller than the sentence.

The interface therefore allows for changes in the “Change tokenization or sentence division” interface, where you can choose to merge the sentence with the following sentence, or split the sentence by starting a new sentence from any given token.

It is not always easy to decide whether two main sentences with и or a similar conjunction between them have been conjoined or not. Both in OCS, Old and Middle Russian, sentences are regularly introduced with such conjunctions, and since topical subjects (and other arguments) are also pervasively dropped, it is very difficult to decide whether or not we are dealing with a “real” coordination. We have decided not to try to keep a sharp distinction, but encourage annotators to keep to small units unless there are good arguments in favour of positing a true coordination. For instance, it is a good idea to posit a coordination when the two verbs in question share dependents, since this can then be indicated by slash notation (see section 18), information that we otherwise lose. Adversative and disjunctive coordinations should also preferably be preserved.

Indirect speech should depend on, and be grouped with, the sentence containing the verb of speech. Direct speech should be separated off as main clauses.

**Grouping of
main clauses**

**Direct and in-
direct speech**

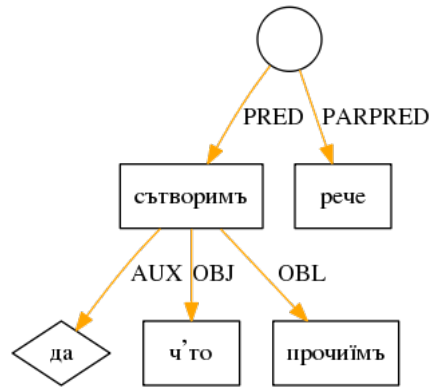


Figure 1: ‘and what, he said, should we do to the others?’ (Supr. 4, 83897)

Sentences introduced by subjunctions should be considered indirect speech if this is possible, ie. if there is no evidence from e.g. pronouns (first or second person pronouns if the subject of the speech verb is in the third person) or other deictic elements to indicate that the quote is direct. If there is such evidence, however, the subjunction should be dependent on the verb of speech (normally via the relation COMP), but not itself have any dependent, and the direct speech should be split off as a separate sentence.

**‘Mixed’
speech**

There are cases where one main clause is entirely included in another main clause as a parenthetical sentence. There is a separate relation tag for parenthetical predications, PARPRED, which should always be placed directly under the sentence root. Since we consider direct speech as main clauses, this tag must be used whenever a speech verb is inserted parenthetically into a sequence of direct speech, as in 1.

**Parenthetical
sentences**

Conversely, when direct speech is wedged into another main clause, as *съде хъ* is in 2, the direct speech is analysed as a PARPRED:

But there are also longer parenthetical sentences. For the purpose of the sentence boundaries, the whole should be treated as one sentence. In analysing the syntactic dependencies, the verb of the parenthetical sentence should be made dependent directly on the sentence root (and not on the main clause) via the relation PARPRED (parenthetical predication).

Notice that only the PRED, PARPRED and VOC (vocative) relations are allowed directly under the root.

8 Tokenisation

The words to be analyzed and represented in the dependency tree do not always

Clitics

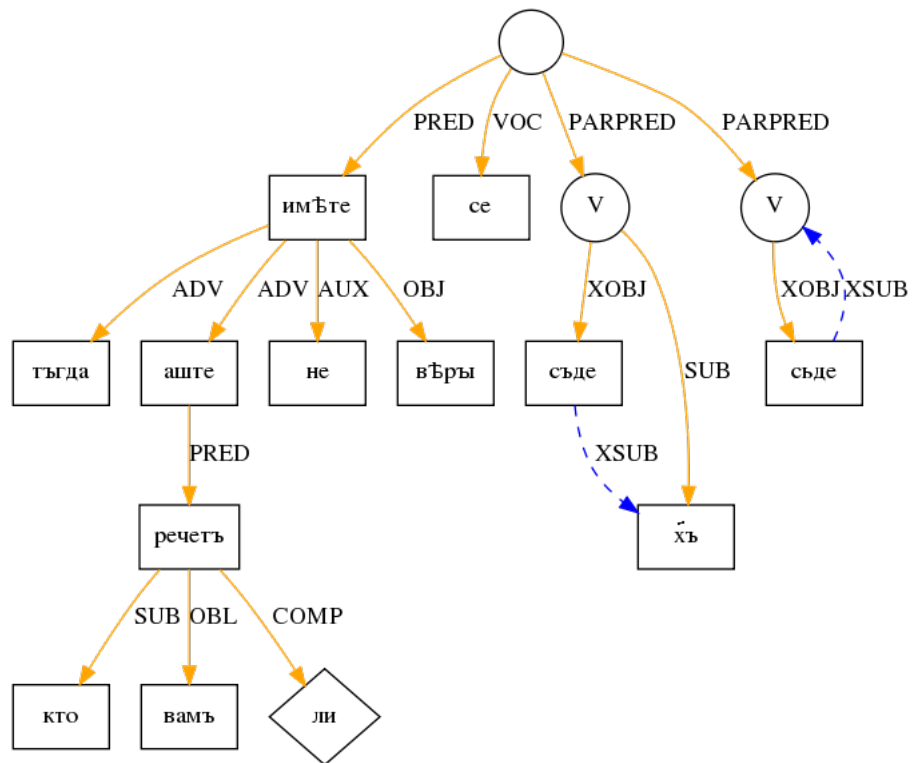


Figure 2: ‘If someone tells you ‘Christ is here’ or ‘Here’, do not believe him.’
(Mar. Matt. 24.23, 39478)

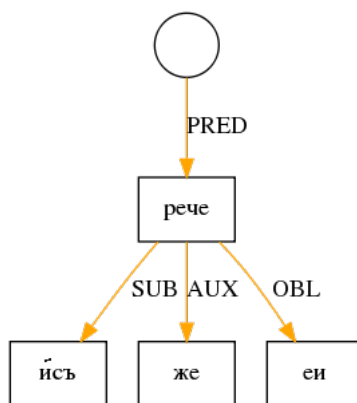


Figure 3: ‘But Jesus said to her’ (Mar. Mark 5.34, 50315)

match those in the text. Sometimes elements which need to be treated as separate tokens in the syntactic model are run together in the text. The main Slavic example of this is the reflexive marker *сѧ*, *сѧ*, which we always split from the verb (if the edition merges them). We sometimes also split *не* off from verbs, but never in the case of *не быти*, which instead has a separate lemma.

9 The verb

The verb is the central element of every predication. We model this by taking the verb to be the head of all other elements in its sentence (except subjunctions, see section 14.2). We also let the verb “stand in” for the sentence as a whole, which means that it is annotated with the function of the whole sentence (unless, again, a subjunction is present). This structure is apparent in example 2: in the main clause, the main verb *имѣте* dominates all the other elements of its sentence.

Main clauses do not have a function within some larger sentence, therefore their verbs are attached to the root via the relation PREDication 3.

Main clauses

In subordinate clauses introduced by a subjunction, the subjunction is assigned the relation which corresponds to the function of the clause, and the verb is attached to the subjunction via the relation PRED 4.

Subordinate clauses

For more on subordinate clauses with subjunctions, see section 14.2. Sometimes the subjunction is missing even in sentences where we would expect one, see section 14.3.

In other contexts (relative clauses, accusative with infinitives, absolute constructions, dominant participles, conjunct participles and governed infinitives),

Other embedded predications

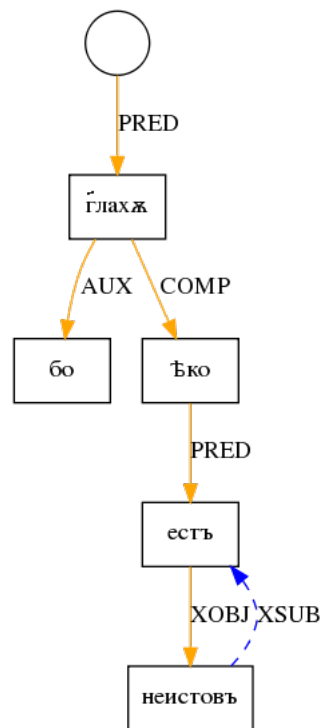


Figure 4: ‘They said that he was out of his mind’ (Mar. Mark 3.21, 36489)

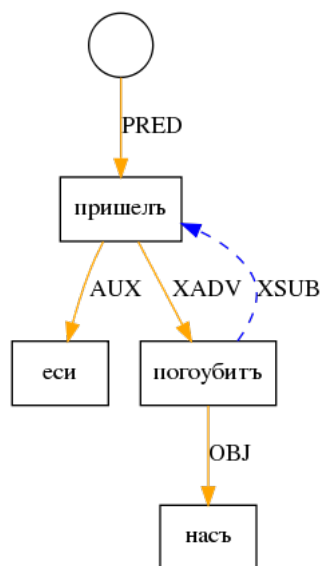


Figure 5: ‘You have come to destroy us’ (Mar. Luke 4.34, 58224)

verbs will be annotated with the function of the whole construction, see section 14. This means that a verb should always be present: if it is not in the text, it will have to be inserted, see section 16.

Auxiliary verbs are attached to the main verb via the relation AUX. Note that we only recognise one type of auxiliary in OCS, *byti* with l-participles (5).

**Auxiliary
verbs**

In Old and Middle Russian we also recognise *byti* as an auxiliary if it is only there to serve as a tense/mood marker for an infinitive (6). Thus, while we only use AUX for auxiliaries of l-participles in OCS, we also use it for auxiliaries of infinitives in the East Slavic part of the corpus.

Since we only allow PRED, PARPRED and VOC under the root, it is necessary to supply an empty verb whenever none is present. For example, when Jesus asks the disciples how many baskets of bread they distributed, the disciples answer ‘seven’ (7):

**Non-verbal
predication**

On the other hand, there are cases where a participle is coordinated with a finite verb. Since we annotate syntactic function, and not form, these are treated as PREDs, corresponding to their function (8).

There are also cases where a dative absolute is seemingly coordinated with a finite verb (9). In these cases we always insert a null verb over the dative absolute, mostly for the sake of retrievability.

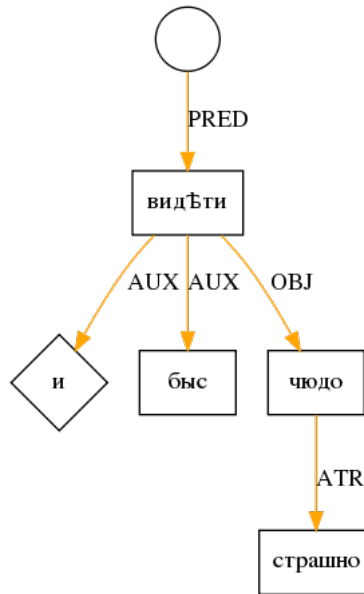


Figure 6: ‘And a terrible miracle could be seen.’ (PVL 44.29, 124910)

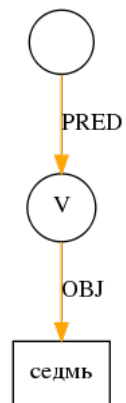


Figure 7: ‘Seven.’ (Mar. Mark 8.5, 50383)

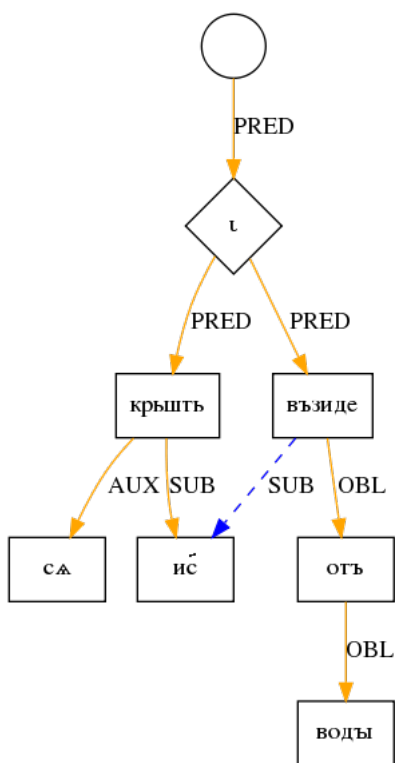


Figure 8: ‘Jesus was baptised and came up from the water’ (Zogr. Matt. 3.16, 75522)

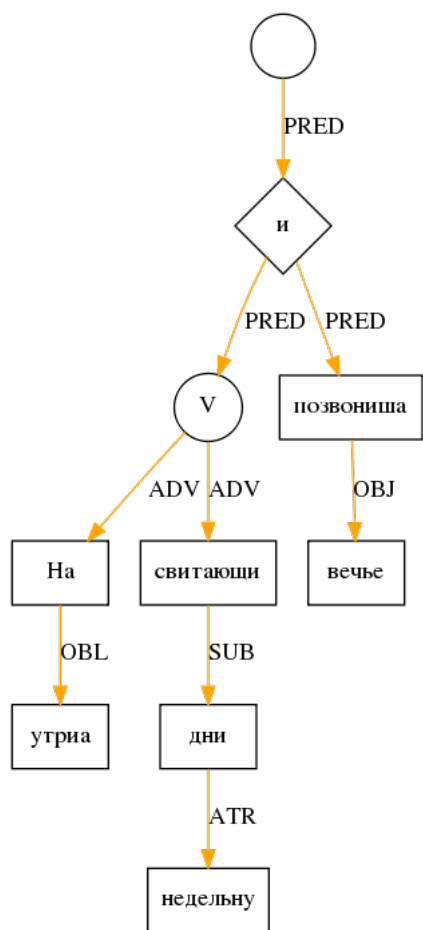


Figure 9: ‘When Sunday was dawning in the morning, they rang for the assembly’ (The Taking of Pskov 140415)

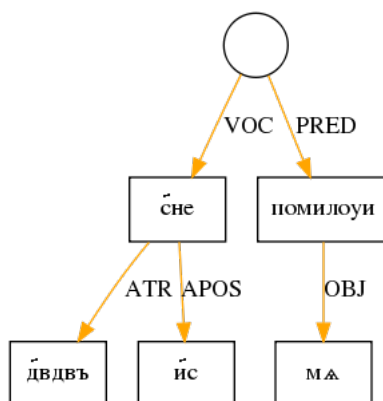


Figure 10: ‘Son of David, Jesus, have mercy on me’ (Mar. Mk. 10.47, 36955)

10 Exclamations

We use the relation VOC for all kinds of exclamations. These are placed directly under the root, as they are external to the sentence. In the syntactic annotation we annotate function and not form, so this relation is not only used for vocative nouns (10), but also for nominative and accusative exclamations as well as for various interjections (11) etc.

**Vocatives,
interjections**

Some sentences consist of only an exclamation. In such cases there is no need to introduce an empty PRED-node (12).

Notice in particular that *ce* belongs here. In many cases, this interjection is used in presentation constructions, which should be analysed as involving an empty PRED (13).

11 Non-verbal sentence-level grammatical relations

In this section we describe sentence-level grammatical relations, ie. relations that have a verbal node as a head. These are SUBject, OBJect, OBLique, AGent, XOBJ (argument with external subject, for instance predicative complements), ADVerbial and XADV (adverbial with external subject, such as conjunct participles), as well as some supertags described in section 11.8.

11.1 Subject

In the typical case, SUB relates a nominative noun to its verb (14).

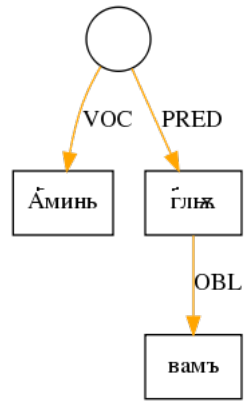


Figure 11: ‘Verily I say onto you’ (Mar. Mark 6.11, 50333)

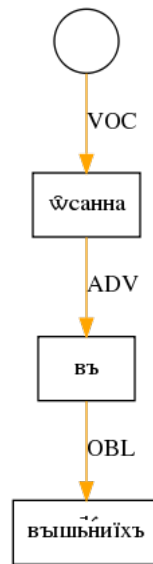


Figure 12: ‘Hosanna in the highest’ (Supr. 28, 155009)

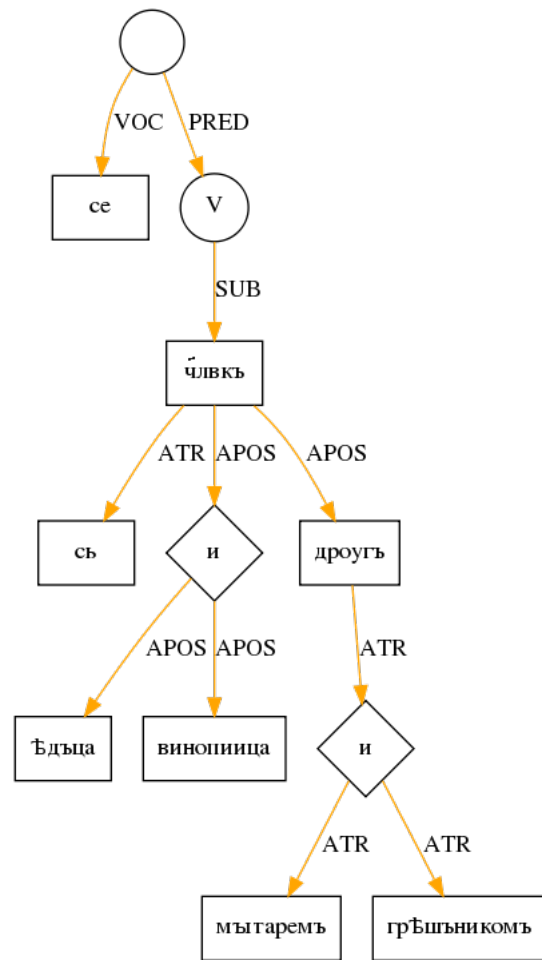


Figure 13: ‘Behold, a gluttonous man and a drunkard, a friend of tax collectors and sinners!’ (Mar. Matt. 11.19, 60029)

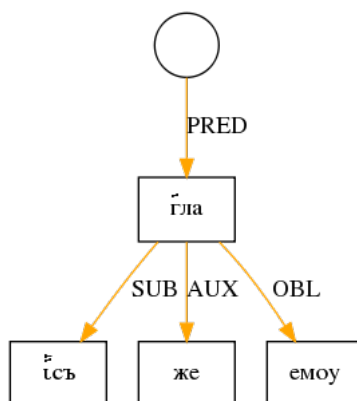


Figure 14: ‘but Jesus said to him’ (Mar. Matt. 8.22, 50730)

But on some occasions, we also find partitive expressions (genitive nouns and prepositional phrases) which are subjects and must be given the relation SUB (15). **Partitive subjects**

Genitive partitive subjects are most notably (but not only) found in negative existential constructions (“there wasn’t”, 16).

We also analyse datives in modal infinitive constructions as subjects (17). Note that we take the infinitives in these constructions as PREDS, both in main and subordinate clauses. **Dative subjects**

Note that infinitives are never subjects, but rather COMPs, XOBJs, XADVs or NARGs, see section 19.11. In OCS infinitives can marginally be nominalised using the mock article construction, see section (17.1). Such infinitives may be analysed as SUBs.

Subjects can also be nouns in oblique cases in an absolute construction, see section 14.7. **Subjects in absolute constructions**

11.2 Object

In the typical case, OBJ relates an accusative noun to the verb (18).

Note especially the policy for genitive-accusatives. We solve this problem using the interplay between the morphological and syntactic annotation layers: Morphologically, genitive-accusatives (indeed any form that looks like a genitive) are annotated as genitives, including pronominal forms such as *mene*, *tebe*, *sebe*. In the syntax, however, a genitive-accusative direct object should be analysed as an OBJ, not an OBL (19). **Genitive-accusative**

In some cases there may be doubt whether the verb requires a genitive or an accusative, or may occur with either. In such cases the supertag ARG should

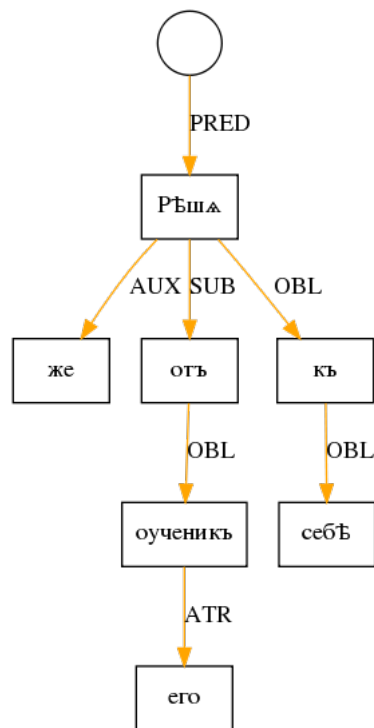


Figure 15: ‘Some of his disciples said to themselves’ (Mar. John 16.17, 42726)

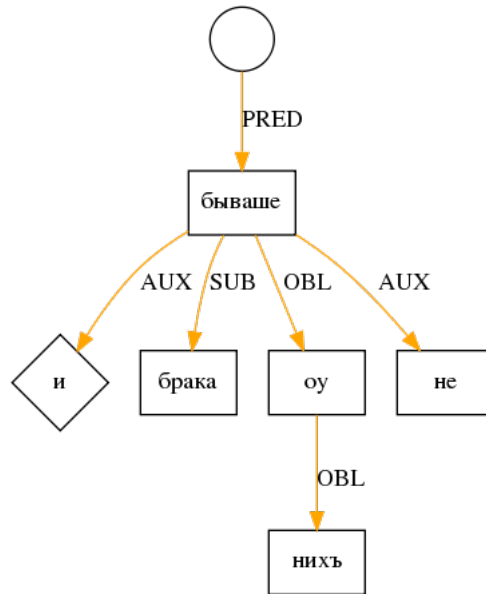


Figure 16: ‘And marriage didn’t exist among them’ (PVL 13.18, 123595)

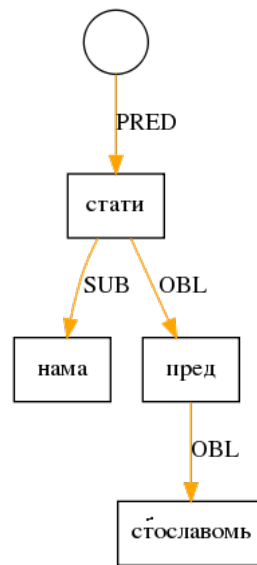


Figure 17: ‘We should stand before Svyatoslav.’ (PVL 177.22–23, 130143)

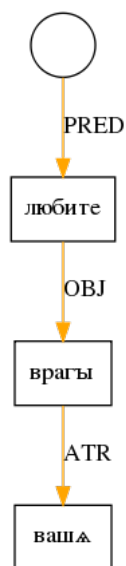


Figure 18: ‘Love your enemies’ (Mar. Matt. 5.44, 50678)

be used, see section 11.8. Note especially that this is the policy used with perception verbs such as *viděti* ‘see’ (20).

Negated objects regularly occur in the genitive. They should be analysed as OBJs (21). Again, if there is doubt whether the verb requires the accusative or the genitive, use the supertag ARG. **Negated objects**

As with subjects, we sometimes find partitive expressions which are objects. They are given the relation OBJ, even if they are PPs (22). Partitive objects raise a problem which does not occur with partitive subjects, since they must be distinguished from normal governed genitives and PPs. The relation OBJ should only be assigned if an accusative could be substituted. If in doubt, consult section 11.8. In OCS, it can be hard to distinguish partitive genitive objects (OBJs) from genitive objects required by the verb (OBLs). If in doubt, use the supertag ARG. **Partitive objects**

Notice also that the supine requires genitive objects. If the verb requires an accusative object in other forms, the genitive object of the supine should still be an OBJ. If there is doubt whether the verb requires the accusative or the genitive, the supertag ARG should again be used. **Supines**

We consider verbs with two direct objects to be marginally possible in OCS and Old Russian. In such cases, annotators should first check if both accusatives (or genitives, see above) qualify as arguments (see section 20.1). If they do, it is possible to take both as OBJs, but this should only be done if both **Double objects**

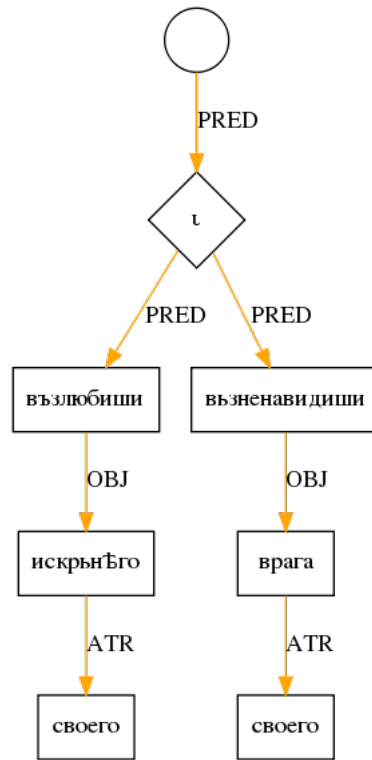


Figure 19: ‘You shall love your neighbour and hate your enemy’ (Mar. Matt. 5.43, 50676)

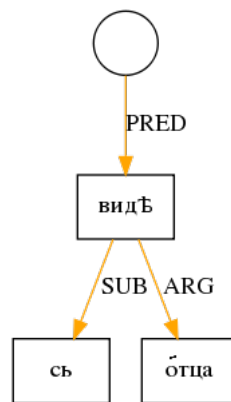


Figure 20: ‘He has seen the Father’ (Mar. John 6.46, 63042)

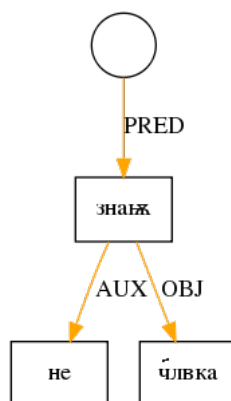


Figure 21: ‘I don’t know the man’ (Mar. Matt 26.72, 62963)

could reasonably become subjects in a passive construction. If only one of the accusatives (or genitives) can be the subject in a passive construction, the other must become an OBL (or an ADV, if it does not qualify for argumenthood). Very often, it is not possible to determine whether a given argument can become the subject in the passive or not; in such cases, annotators should assume that they can be, and annotate them as OBJ. Thus the ‘default case’ is that two accusative arguments of a verb should both be treated as OBJects. We stress again that this is extremely rare: the vast majority of verbs take only a single OBJ.

Note that infinitives are never OBJs, but rather COMPs, XOBJs, XADV or NARGs, see section 19.11. In OCS infinitives can marginally be nominalised using the mock article construction, see section (17.1). Such infinitives may be analysed as OBJs.

11.3 Obliques

We use the relation OBLique to attach those arguments of the verb which are not subjects or objects to the clausal node. By *argument* we mean any syntactic element seen as required by a verb. This could be a genitive, as with *bojati se* ‘fear’; a dative, as with *pomošti* ‘help’; an instrumental, as with *vlasti* ‘rule’; a locative, as with *kosnoti se* ‘touch’; a prepositional phrase, as with *gněvati se* ‘be angry with’ and in general with motion verbs; and even an adverb, frequently found with motion and position verbs. It is not always clear whether a noun phrase is an oblique argument or not; or whether a genitive is an object (animate? negated? partitive?) or an oblique argument. If in doubt consult section 11.8. Oblique arguments include non-accusative ‘objects’ as well as prepositional

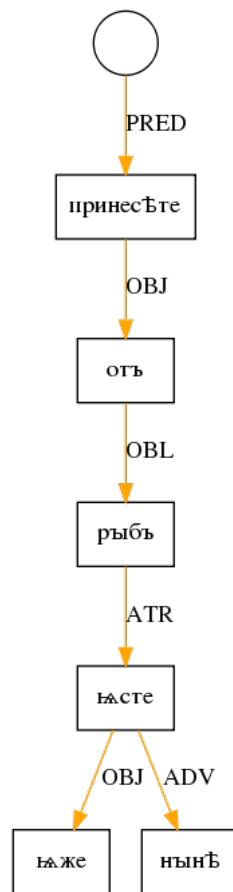


Figure 22: ‘Bring some of the fish you have just caught.’ (Mar. John 21.10, 43008)

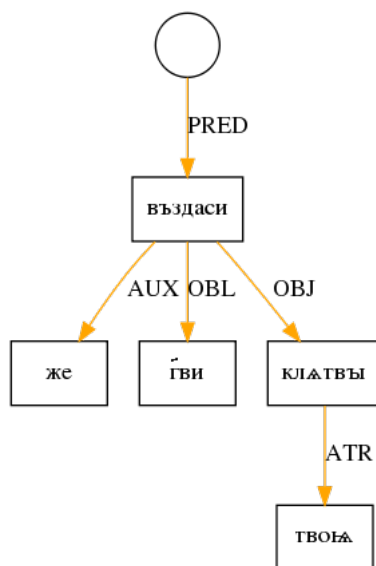


Figure 23: ‘Give your oaths to the Lord.’ (Mar. Matt. 5.33, 50672)

arguments. Note that accusative OBLs are also possible, see section (11.2).

OBL is the relation used for indirect objects (23).

With motion verbs, we consider all directional expressions (goal and source) to be OBLs. This goes both for prepositions and adverbs, so e.g. *туда* should very often be an OBL. Other prepositions that are necessary to the meaning of the verb, such as in e.g. *избавити отъ* ‘deliver from’, also belong here, see (24), which also exemplifies a goal argument with a motion verb.

Path expressions, on the other hand, are normally not oblique arguments, but rather adjuncts (ADV). Exceptions occur, however, for example in cases where the path argument is ‘required’ by a prefix.

A restricted group of adjectives such as *достойнъ* also take complement nouns. We relate these nouns to their adjectives via the relation OBL (25).

As example (25) illustrates, the complement of the preposition is also considered an oblique argument (of the preposition), no matter the function of the phrase as a whole (26).

**Indirect ob-
ject
Prepositional
arguments**

**Arguments
of adjectives
Objects of
prepositions**

11.4 AG: Demoted subjects in passive constructions

In passive constructions, and some rare active ones, we use the relation AGent to relate a non-nominative demoted subject to the verb. The AG is typically expressed by a prepositional construction (27) or a bare case form (usually

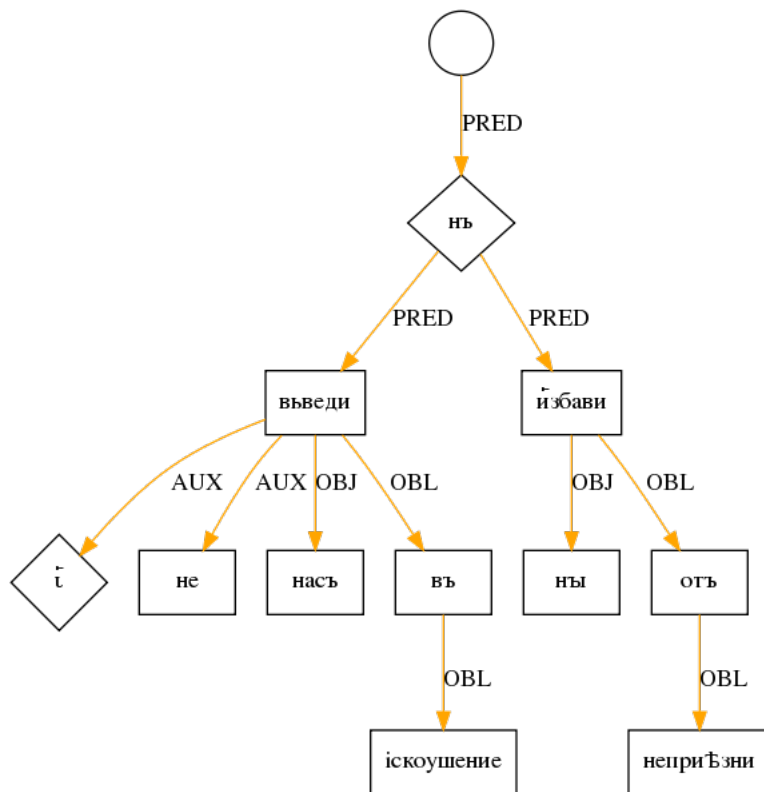


Figure 24: ‘And lead us not into temptation, but deliver us from evil.’ (Mar. Luke. 11.4, 40646)

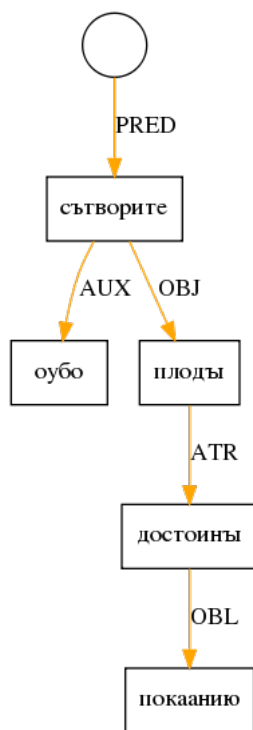


Figure 25: ‘Bring forth fruit worthy of repentance.’ (Mar. Luke. 3.8, 40068)

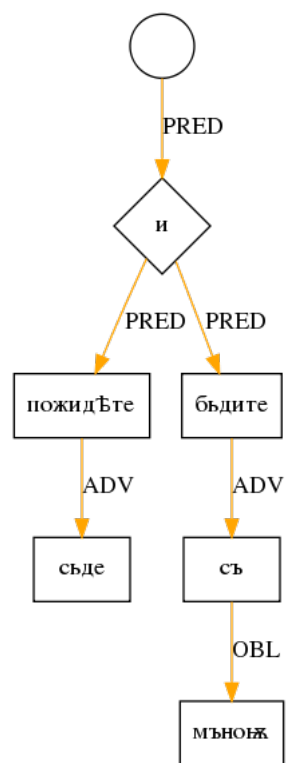


Figure 26: ‘Stay here and keep watch with me.’ (Mar. Matt. 26.38, 39639)

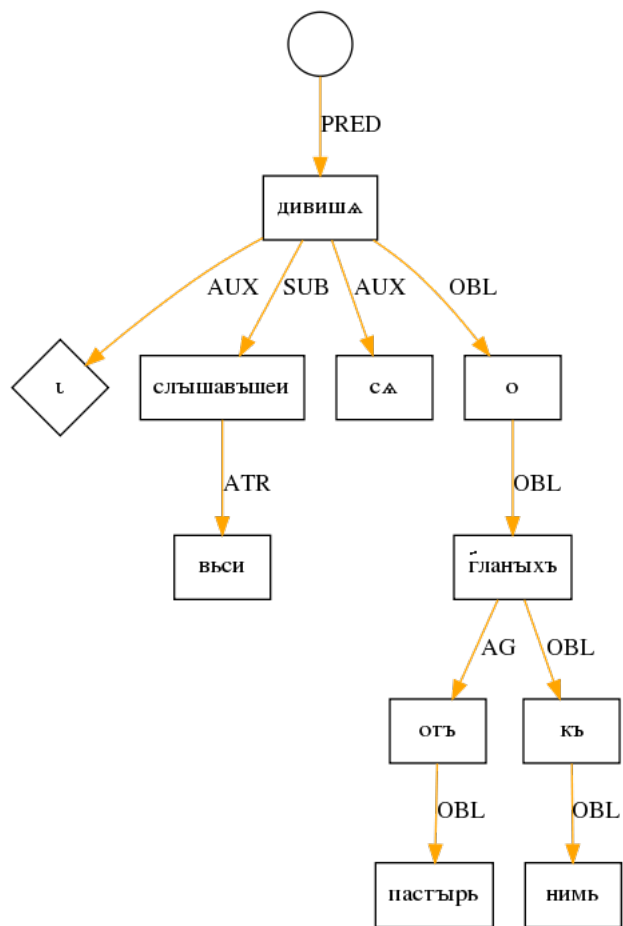


Figure 27: ‘And all who heard it wondered at the things which were told them by the shepherds.’ (Mar. Luke. 2.18, 40023)

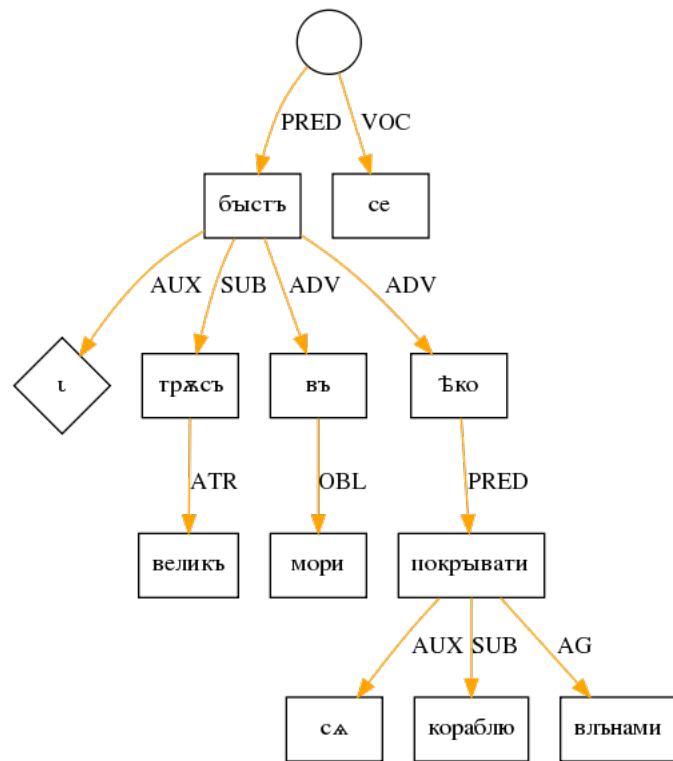


Figure 28: ‘And behold, there arose a great storm on the sea, so that the boat was being swamped by the waves.’ (Mar. Matt. 8.24, 38525)

instrumental) (28). Note that the AG does not have to be an agent semantically (cf. example 28), as long as it could have been the subject in the corresponding active construction.

Sometimes, deverbal nouns take arguments with the typical AG morphological marking (*otb* + genitive, bare instrumental. We use the relation AG for such dependents as well, see the analysis of the *otb* dependent of *goneniju* ‘persecution’ in (29).

**Adnominal
AG**

The AG tag is ordinarily used to express the demoted agent with a *passive* verb. In Greek, however, some intransitives are regularly used instead of the expected passive form of a verb. Verbs occurring with this construction include *paskhō*, *piptō*, *pheugō*, *apothnēiskō*. Since these verbs are functionally equivalent to passives and have supplanted the passive verbs forms, the PROIEL annotation allows for prepositional phrases with *hupo* to be marked as AG with these verbs. Note, however, that this is only allowed with a restricted set of verbs and is most often found in Greek. However, the same type of construction occurs marginally in OCS, e.g. with *iscělēti* ‘be healed (by)’ in Mar. Luke 8.43 (40449). We allow the AG analysis in these rare cases as well.

**AG with act-
ive verbs**

11.5 Adverbials

We use the relation ADverbial to attach adverbial expressions to the sentence. Such expressions can take various forms: adverbs (30), prepositional phrases (31), nouns (in oblique cases) (32) and others. In some cases, it is not clear whether they are adverbials or oblique arguments, and in that case, sections 11.8 and 20.1 should be consulted.

We also consider predicative/conjunct participles and adjectives (see sections 20.3 and 14.8.2), as well as gerunds, to be adverbial, but they are special because they are cases of embedded predications with their own argument structure. Moreover, they are special in that they cannot take a subject dependent because they have an external subject. They are further described in section 14.8. Participles in absolute constructions (see section 14.7) take subject dependents and therefore get the ADV relation.

Adverbial accusatives are also ADVs. Note that this also holds for the use of *čto* when it is used to mean ‘why’ (33). Note that we generally assume pronouns, adjectives etc. that function as ADVs to occur in the accusative, even though the form may as well indicate the nominative.

**Adverbial
accusative**

The relation ADV is also used for sub-sentence-level modifiers of adjectives (34), (subtrees headed by) prepositions (35), (subtrees headed by) numerals (36) and other adverbs (37).

**ADV
modifying
non-verbal
elements**

11.6 Predicative complements

The relation XOBJ is used for subject and object complements (as well as other functions as explained in section 20.5) which are introduced by verbs like *byti*, *naricati se*, *mьnēti se* (subject complements) and *tvoriti* (object complements), as well as in verbless absolute constructions (see section 14.7). The relation

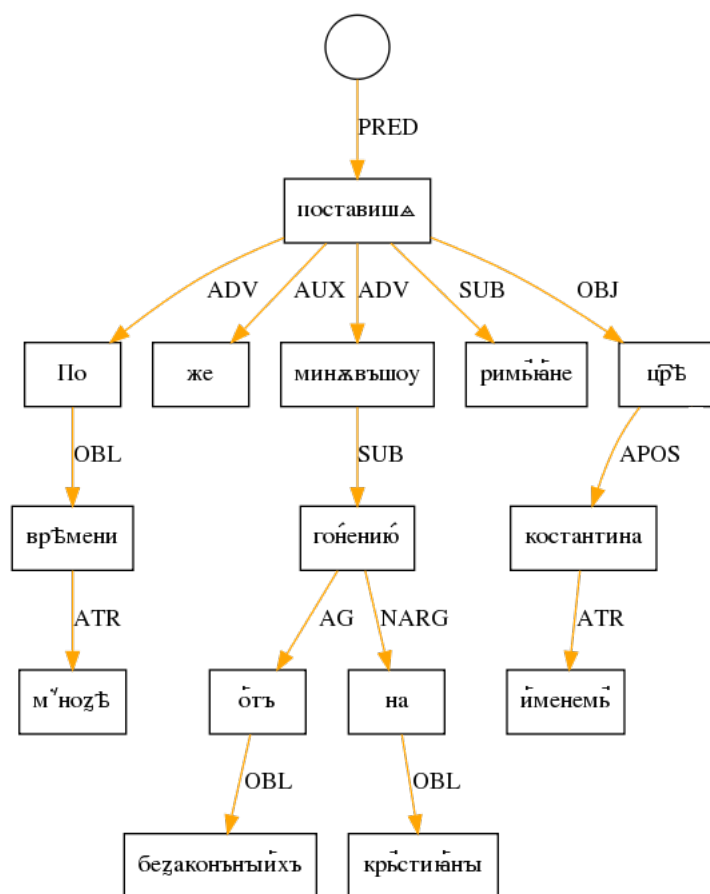


Figure 29: ‘After a long time, when the persecution by the lawless ones against the Christians was over, the Romans appointed an emperor by the name of Constantine [...]’ (Supr. 47, 170624)

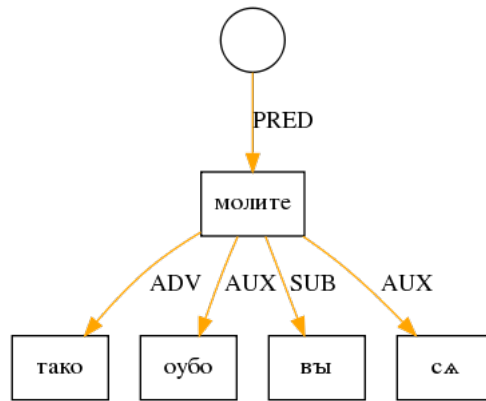


Figure 30: ‘Pray then like this.’ (Mar. Matt. 6.9, 38420)

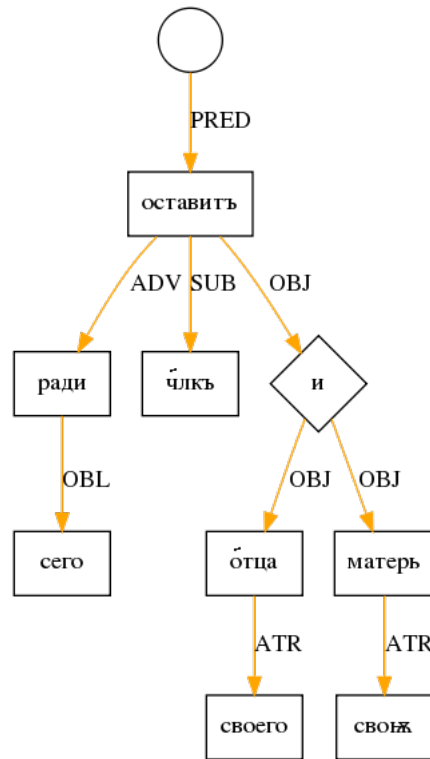


Figure 31: ‘For this reason a man will leave his father and mother.’ (Mar. Matt. 19.5, 50923)

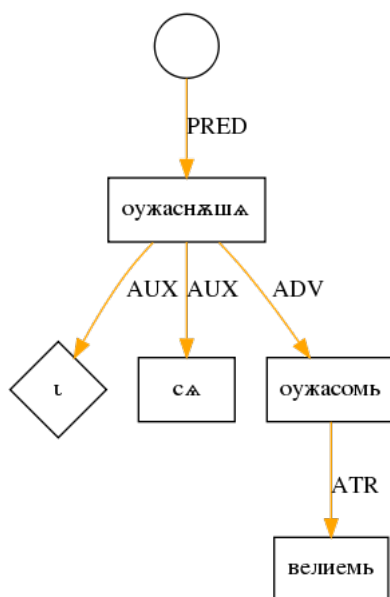


Figure 32: ‘And they were astonished with a great astonishment.’ (Mar. Mark 5.42, 36617)

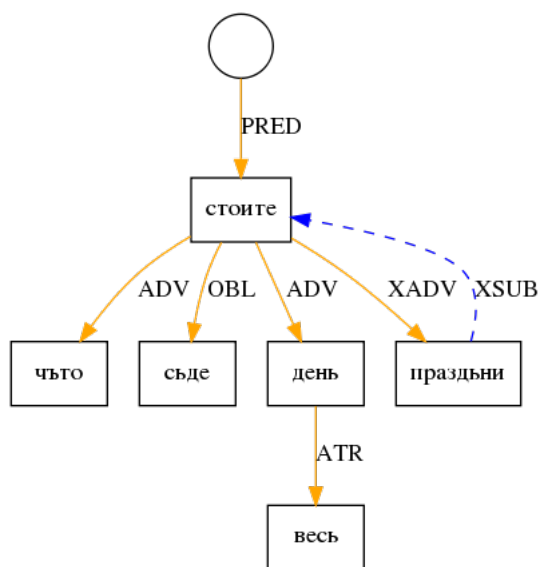


Figure 33: ‘Why are you standing here idle all day?’ (Mar. Matt. 20.6, 50937)

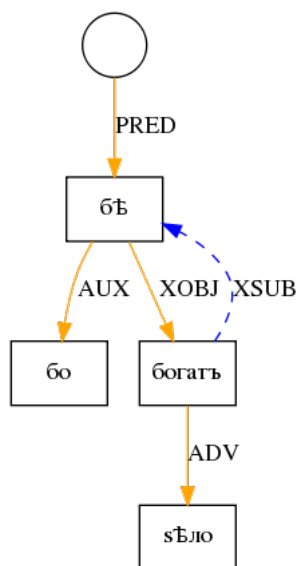


Figure 34: ‘For he was very rich’ (Mar. Luke 18.23, 51612)

itself does not make clear whether we are dealing with a subject predicative or an object predicative. Instead, we use the slash notation to mark this, see section 18. Predicative complements can be instrumental-marked (rare in OCS, but increasingly common in Old and Middle Russian). They should still be analysed as XOBJs, not as OBLs. The same goes for complements introduced by a preposition, such as in *byti vъ*.

Note that predicatives can be of many different syntactic categories: adjectives (38), nouns (39) and preposition phrases (40) are typical examples.

11.7 Adverbial appositions

Although appositions (see section 13.3) are normally dependents of sentence constituents, there is at least one case where we use sentence level appositions, namely in the case of ‘distributive elaborations’, i.e. cases where a predication structure (normally without a finite verb) contains a distributive element or a list. A classical example is (41).

Since the appositive element is not a constituent, we need to embed the separate elements under an empty verbal node, which is then given the relation APOS to the finite verb.

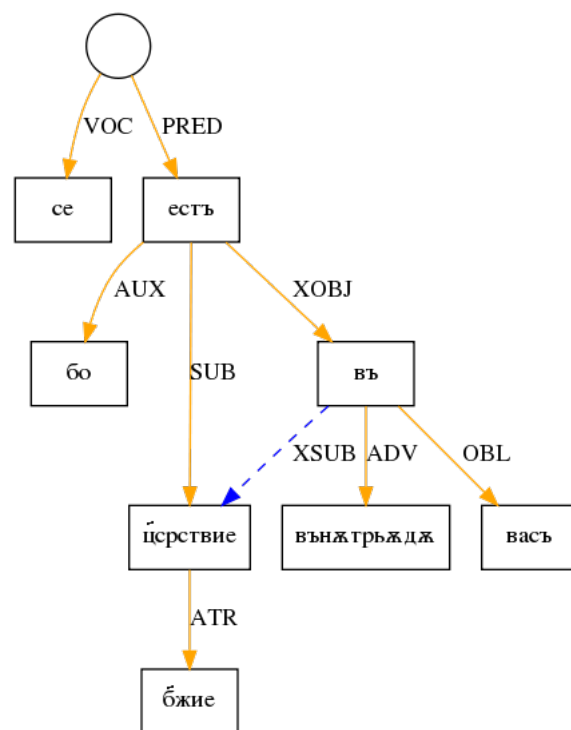


Figure 35: ‘For, behold, the kingdom of God is within you.’ (Mar. Luke 17.21, 41054)

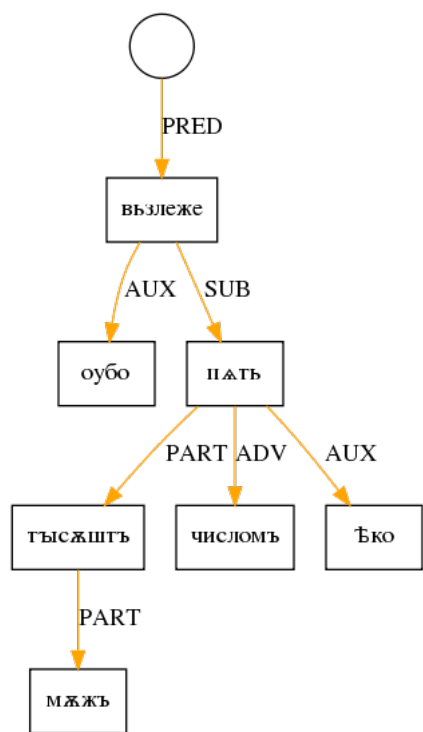


Figure 36: ‘About five thousand men in number sat down’ (Mar. John 6.10, 41928)

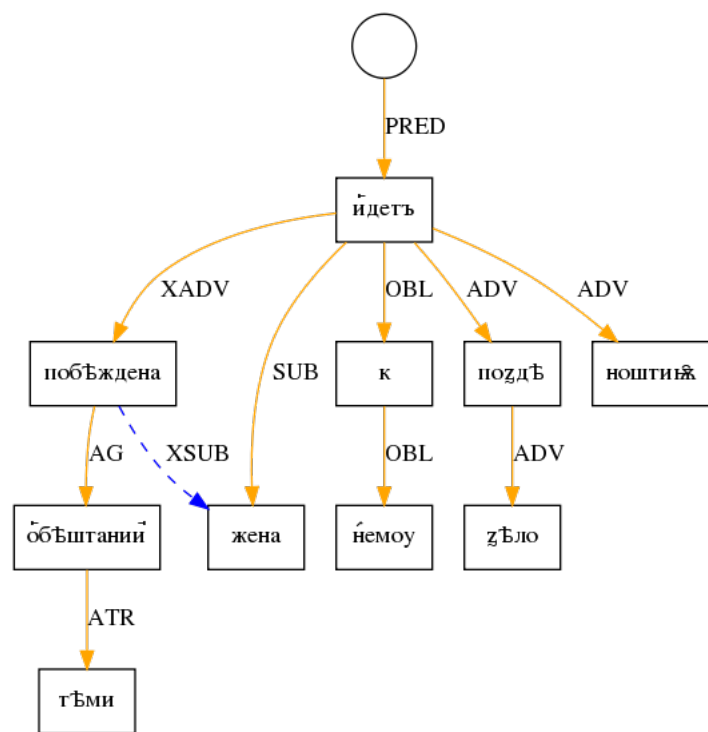


Figure 37: ‘Won over by those promises, the woman comes to him very late at night’ (Supr. 46, 164856)

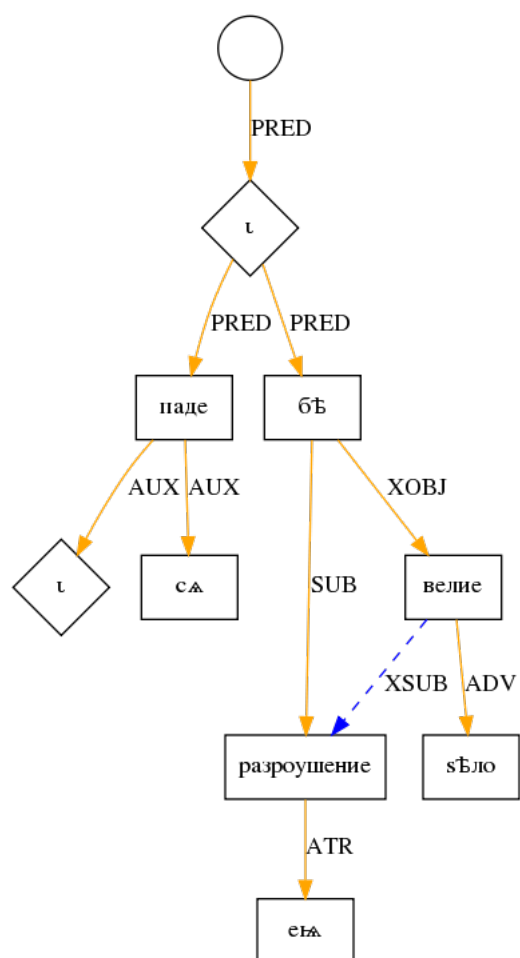


Figure 38: ‘And it fell, and its destruction was very great.’ (Mar. Matt. 7.27, 50698)

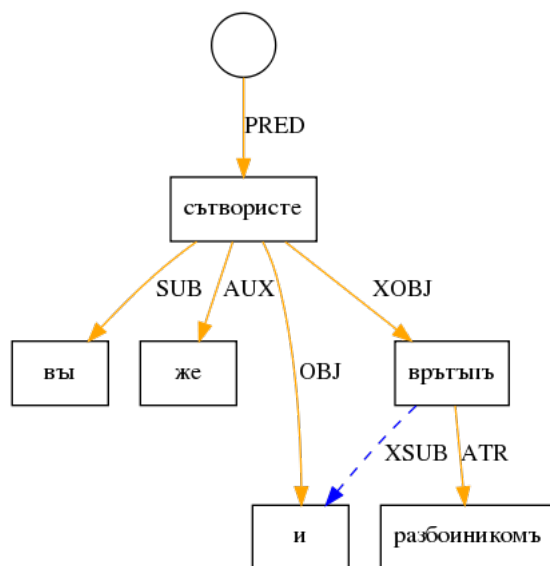


Figure 39: ‘But you made it a den of robbers’ (Mar. Matt. 21.13, 39262)

11.8 Supertags

The distinctions that we attempt to make in our syntactic model are not always crystal clear. Section 20 offers some supplementary advice on the use of some relations. In order to preserve the quality of the data, we also provide certain supertags which the annotators should use when they are in doubt, instead of simply choosing one of the alternatives. These are:

- PER for peripheral (not subject or object) elements, ie. OBL or ADV. This should be used whenever it is not possible to decide whether an element is an argument or an adjunct.
- ARG for arguments, to be used whenever it is not possible to decide whenever an element is an OBJECT or an OBLIQUE. (Note that this is the correct choice when an argument is ambiguous between the genitive-accusative and regular genitive.)
- NONSUB for non-subjects, ie. elements that are either OBJECTs, OBLIQUEs or ADVERBIALs⁴

⁴This case is extremely rare, of course, but the tag can also be used to flag anomalies which the reviewers should look at.

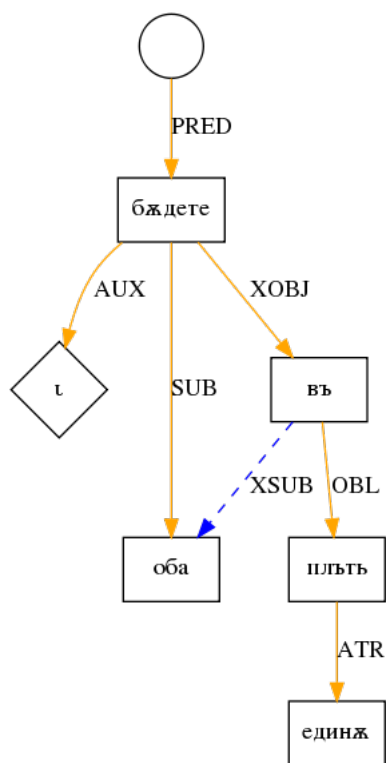


Figure 40: ‘And the two will become one flesh’ (Zogr. Mark 10.8, 176656)

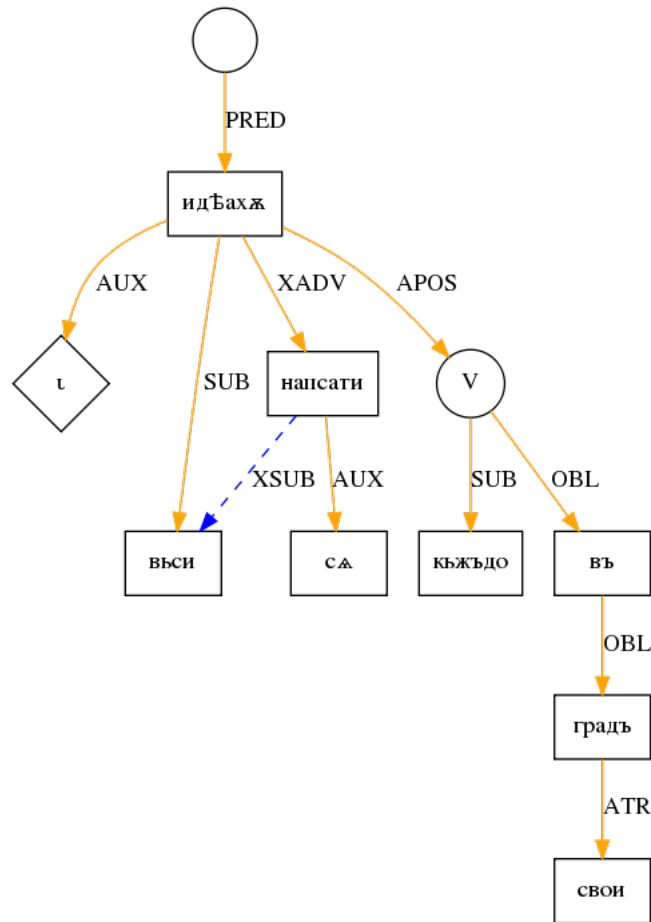


Figure 41: ‘And all went to be registered, each to his own town.’ (Mar. Luke 2.3, 40007)

In addition there are supertags for verbal functions, see section 14.9 and a supertag for adnominal functions, see section 13.

11.9 Auxiliary words

Items that are not covered by these tags are simply given the tag AUX, which serves to mark auxiliary verbs, modal particles, focus particles, negation etc. Information about these items is always recoverable from the categorial information in their morphology. The intuition behind the relation AUX is that it serves to mark off ‘grammatical words’ as opposed to ‘lexical words’. There are certain adverbs whose meaning can sometimes be so weakened that they appear as grammatical words, so the distinction between AUX and ADV is not always clear.

It can be a bit confusing that the AUX relation is used for such diverse phenomena. However, they have something in common: Auxiliary words form a unit with their head, i.e. they provide functional information about the head or the entire sentence, but do not have a proper syntactic function. Consequently, an AUX can not be a head for anything else.⁵ An article specifies the definiteness of its head noun; a negation negates its head word, a focus particle gives information about the information status of its head, auxiliary verbs give tense and aspect information relating to the verb, reflexive markers give voice information about the verb etc.

It is important to attend to scope issues: auxiliary verbs should be attached via AUX to their verbs; focus particles and constituent negation should be attached to the items they take scope over. In general, a scoping item is considered to scope over its mother node and all nodes dominated by its mother, and it should be placed accordingly. For more on this, see section 19.7.

12 Noun phrases without nouns

Sometimes a sentence-level function is not filled by a noun, but by an adjective (42), participle (43), numeral or adjectival pronoun. Such elements should be given the appropriate function in the sentence, and should *not* be related to an empty node representing the ‘elliptical’ subject, object or otherwise.

Before opting for such an analysis, the annotators should make sure that the adjective/participle is not predicative. Example (43) means ‘Those tending the pigs ran off’. If the sentence had meant ‘Tending the pigs, they ran off’, *pasqšti* would have been an XADV, see section 14.8.2. (XADVs will almost exclusively occur in short form.)

As we have seen in the sections on SUBjects (11.1) and OBJects (11.2) even prepositional phrases can be subjects and objects, usually when they express

⁵Note that the current policy for words meaning ‘only’ is at odds with this principle. Such words can clearly be negated (‘not only Peter, but also John’), and since we have an AUX policy both for them and negations, this is the one case where we’ve allowed stacked AUXes. However, this is probably rather an indication that the ‘only’ words are better analysed as ADV.

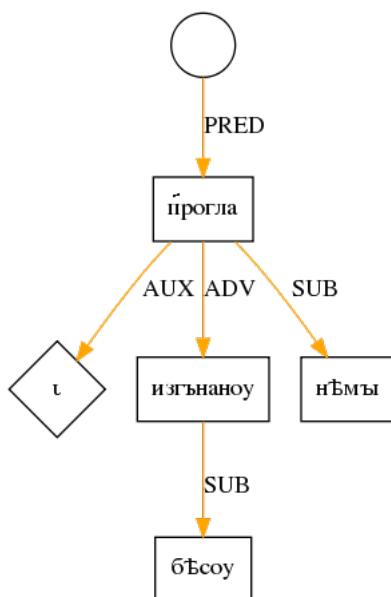


Figure 42: ‘And when the demon had been cast out, the mute man spoke.’
(Mar. Matt 9.33, 38591)

partitivity. Note that in Greek, nominalisation of prepositions, adverbs and infinitives is quite common, by way of adding the definite article in constructions like *hoi par’ autou* ‘the ones around him’ or *hoi nun* ‘people nowadays’. In such cases, the preposition should be the head and the article an AUX. Note that such nominalisations are also (marginally) found in OCS, see section 17.1 for details.

Sometimes there are several adjectives, participles and/or numerals. In such cases annotators will have to make a choice as to what is the head of the construction and what is the modifier/attribute. In general, the element which is most central to establishing the referent should be the head.

This means that numerals should not be the head if there are other available heads, so *drugaě dvъa* ‘two others’ should be headed by *drugaě* and not the other way around (44). It has proven difficult to annotate constructions with several potential heads in a consistent way, since intuitions tend to be unstable, annotators are therefore asked to use the following hierarchy:

- (1) adjectives, participles >> nominalised prepositional phrases >>
demonstratives >> indefinite pronouns >> numerals >> relative clauses

Notice that since relative clauses are lowest on the hierarchy, sentences like

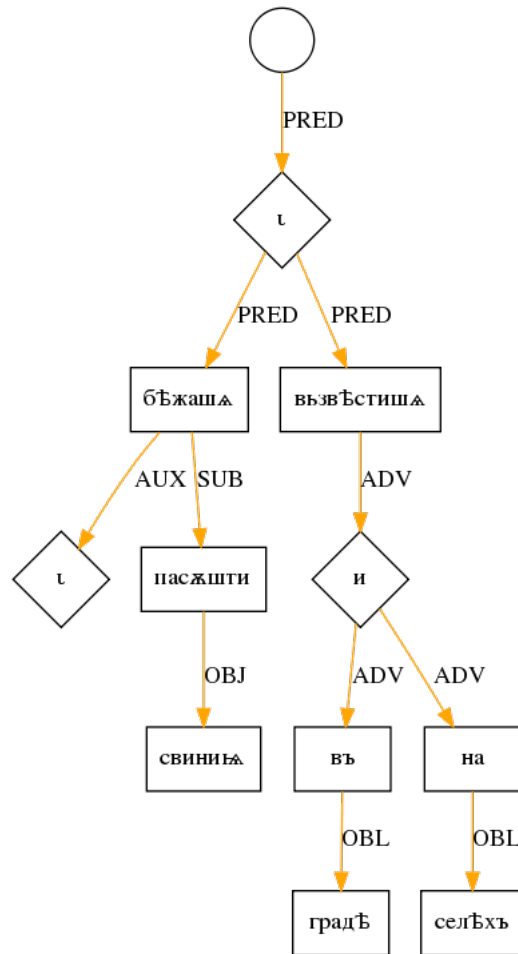


Figure 43: ‘Those tending the pigs ran off and reported this in the town and countryside.’ (Mar. Mark 5.14, 36580)

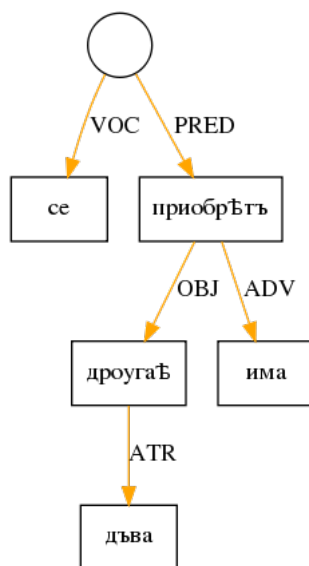


Figure 44: ‘I have gained two more using them’ (Mar. Matt. 25.22, 39542)

(45) will be analysed as having the relative clause modifying the quantifier .

13 Adnominal tags

13.1 General

Dependents of nouns can be of various types. There are negations, emphatic particles etc. which are related via AUX, but with due consideration of scope issues (see section 19.7). In general, ‘grammatical words’ will bear the relation AUX, whereas ‘lexical words’ which are dependents of nouns will have various relations depending on their function. We recognise 4 main types:

ATR - attributes are elements which serve to restrict the reference of a noun.

For example, in *odežda běla* the adjective *běla* serves to restrict the possible reference of the noun *odežda*. Attributes can be adjectives, prepositional phrases, relative clauses, participles, genitives, adnominal datives, numerals - but these categories can also have other adnominal functions, so it is important to pay attention here.

APOS - appositions are elements which serve to further elaborate on a nominal referent, without restricting the reference. Examples are *česarъ* in *Isusъ*

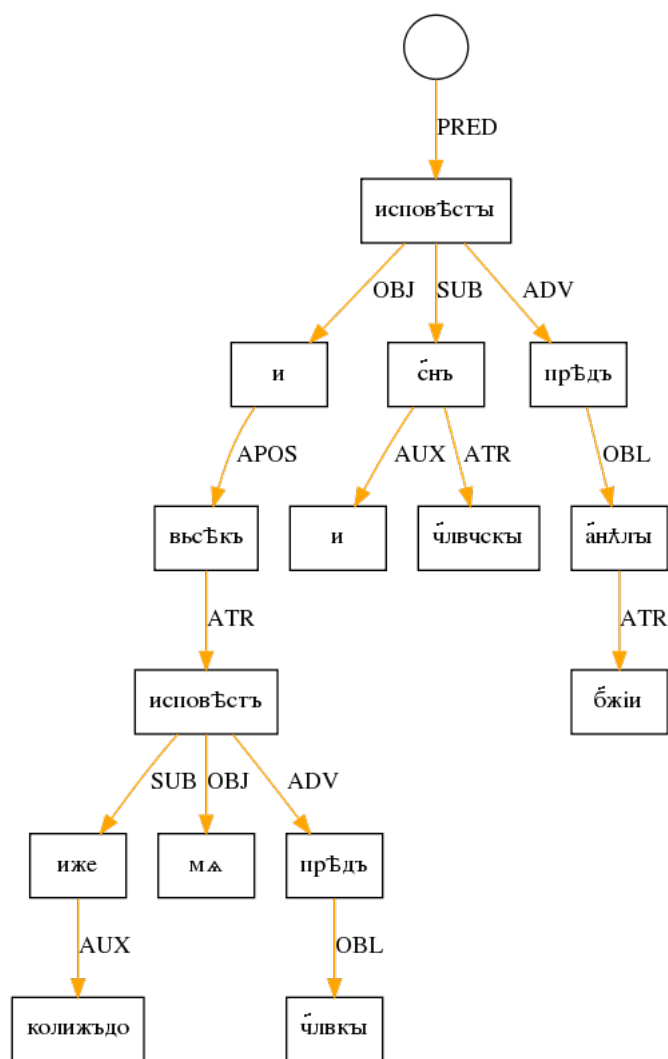


Figure 45: ‘Everyone who acknowledges me publicly here on earth, the Son of Man will also acknowledge in the presence of God’s angels.’ (Mar. Luke 12.8, 51468)

cěsarъ ijudeiskъ, synъ in Xristosъ synъ božii. Appositions are mostly nouns in the same case as their head, and non-restrictive relative clauses.

PART - partitives are elements which tell us to which group or whole the noun belongs. They are typically realised by genitives or by prepositions, predominantly *otъ*.

NARG - some elements can be said to be *arguments* of nouns. The most clear example is the objective genitive or dative, as in *lovitva rybъ* ‘the catching of fish’ or *lovъсь člověkomъ* ‘a catcher of men’. Arguments can also be realised as prepositional phrases, as in e.g. *sъrasenie otъ vragъ našixъ* ‘salvation from our enemies’, and sometimes even as denominal adjectives, as in *straхъ ijudeiskъ* ‘fear of the Jews’. Note that subjective genitives or adjectives are counted among the normal possessive genitives, as attributes (ATR). Arguments of adjectives are *not* NARGs but OBLs.

These categories are further explained and exemplified in the following sections. There is also supplementary information in section 20. If in doubt, annotators should use the supertag ADNOM. Note that several of the adnominal relations are essentially ‘the same’ as certain adverbial relations. For instance, ATR is very much the adnominal version of ADV. This means that non-nominals should never have ATR dependents (unless they’re clearly nominalised), and, conversely, nominals should never have ADV dependents. Likewise, NARG is very much the adnominal counterpart of OBL, this means that nominals should not have OBL dependents, and conversely, that non-nominals should never have NARG dependents. Finally, note that PART is a strictly adnominal relation. Partitive subjects and objects should be analysed as SUB and OBJ.

13.2 Attributes

Attributes are given the tag ATR. Note that ATRs are not stacked. Example (46) has an adjective and a possessive pronoun.

Descriptive genitives are also taken to be ATRs. Example (47) has a descriptive genitive noun ATR which in its turn has a numeral ATR, ‘a daughter of about twelve years’.

We use the tag ATR also for some constructions which are not traditionally called attributes, e.g. because we do not allow nouns to take adverbial dependents. Therefore it is useful to think of the ATR as not only comprising traditional attributes but all kinds of modifiers of nouns.

**Descriptive
genitives**

**Non-
traditional
attributes**

13.3 Appositions

This section deals with adnominal appositions. We also use the APOS tag in a variety of sentential contexts, see section 11.7.

Appositive nouns are attached to their head noun via the relation APOS (48). Such nouns are never restrictive (if they are, they are attached via the relation

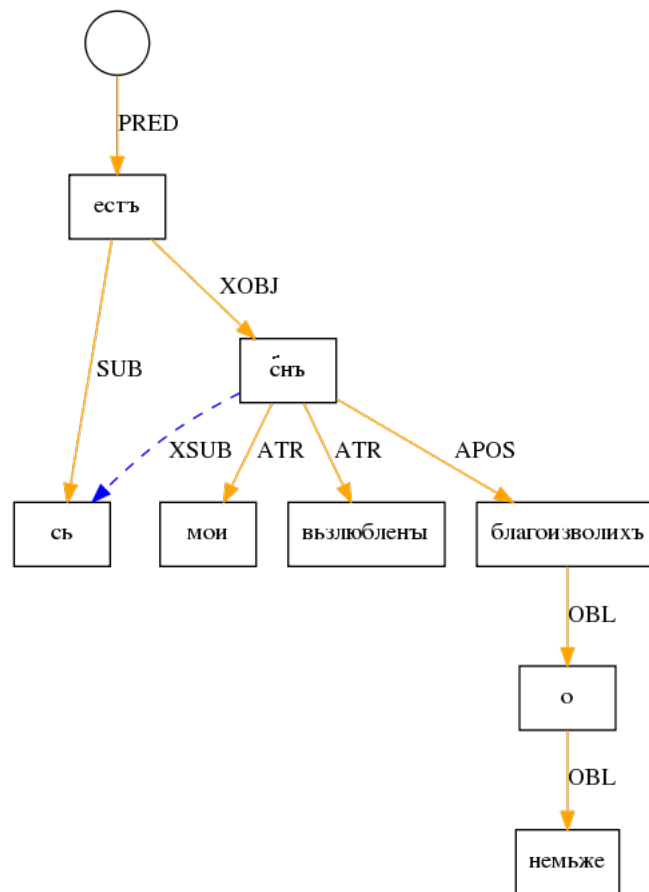


Figure 46: ‘This is my beloved son, with whom I am well pleased.’ (Mar. Matt. 17.5, 39039)

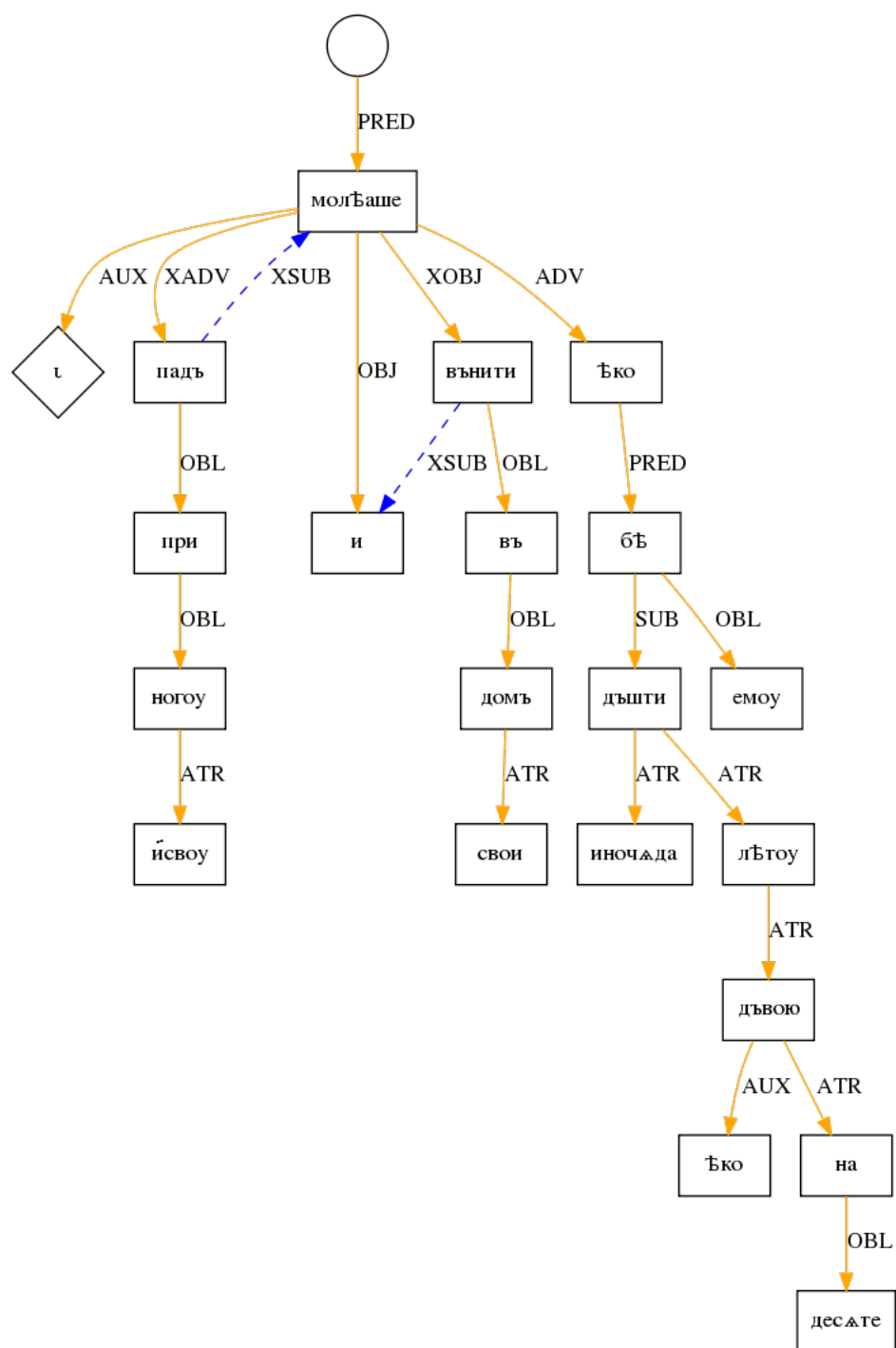


Figure 47: ‘Falling at Jesus’ feet, he implored him to come to his house, because his only daughter, a girl of about twelve, was dying.’ (Mar. Luke 8.41–42, 40446)

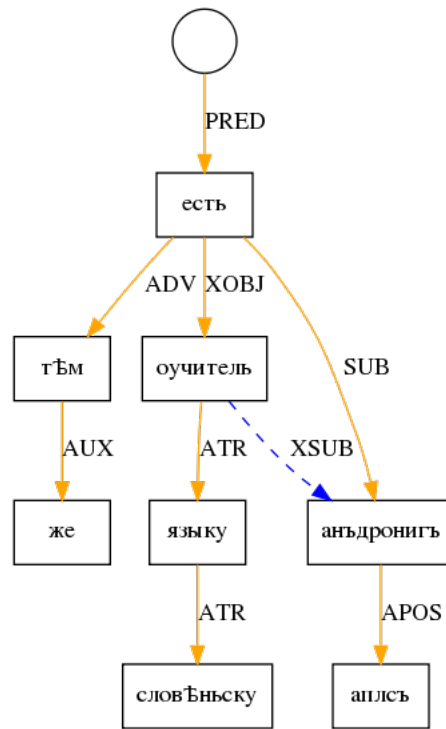


Figure 48: ‘Thus the teacher of the Slavic people is Andronicus the apostle’ (PVL 28.12–13, 124027)

ATR instead), and we do not attempt to keep predicative appositions apart from other appositionive nouns (doubtful cases are rare in the Slavic material).

Non-restrictive relative clauses are also considered appositions (as seen in example 46), see section 14.4 for more details.

Notice in particular that ‘second names’ and modifiers of names, as in *Mary Magdalene*, *Judas Iscariot* or *Jesus Christ*, are always considered appositions. In OCS, such ‘second names’ are often denominal adjectives.

There can sometimes be doubt whether a participle is an apposition or a predicative adjunct (XADV). In general, APOS should be chosen whenever a timeless property of the head is expressed, or at least a property which is not tied to the time of the governing verb, as in example (49). Being called “the Twin” is a property of Thomas that is not temporally connected with the event of him uttering something to his fellow disciples, so we use APOS instead of XADV.

**Names and
modifiers
of names
APOS
or XADV?**

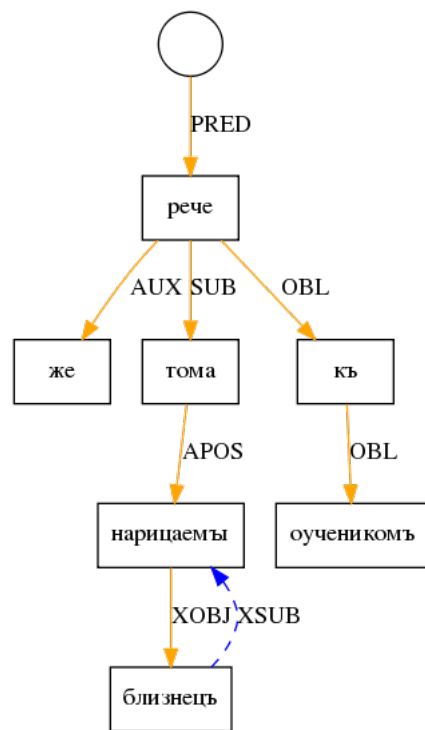


Figure 49: ‘So Thomas, called the Twin, said to his fellow disciples’ (Mar. John 11.16, 42407)

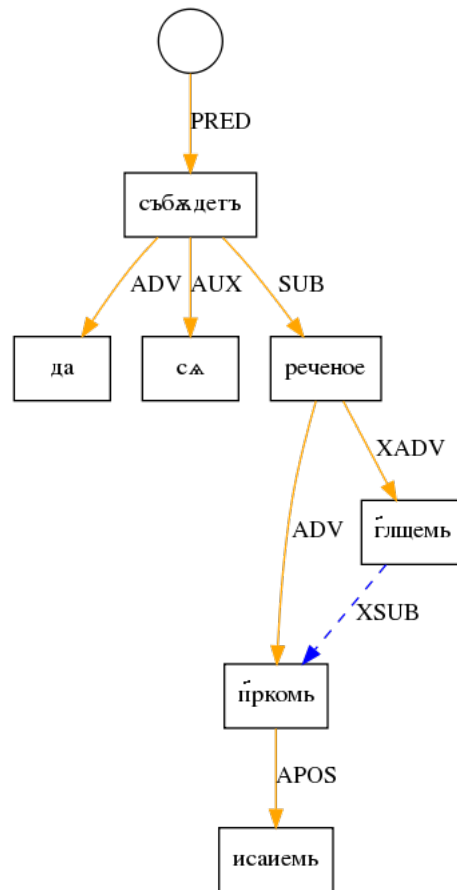


Figure 50: ‘This was to fulfill what was spoken through the prophet Isaiah’
(Mar. Matt. 12.17, 38724)

When multiple expressions with the same function occur, it is often difficult to decide which one to take as the head and which one as the apposition. In such cases we follow the surface word order: the expression which comes first in the sentence is taken as the head, while the one that follows is taken as an apposition. This policy quite frequently leads to proper noun appositions on common noun heads (50), which may seem a bit non-intuitive. However, an intuition-based approach appears to give quite inconsistent results.

**Appositions
and
word order**

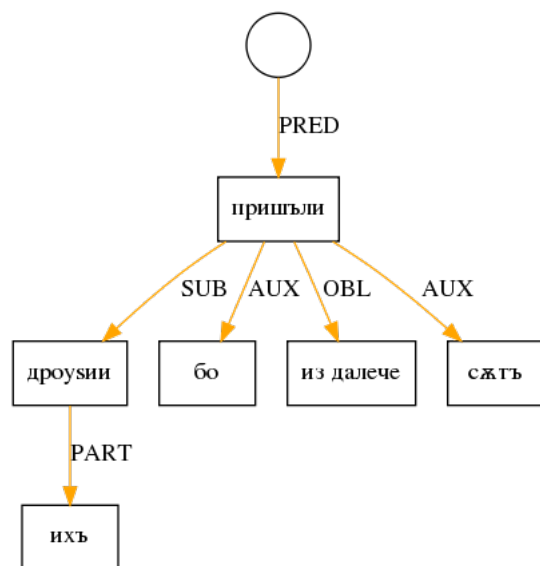


Figure 51: ‘For some of them have come a long distance.’ (Mar. Mark 8.3, 36748)

13.4 Partitives

Adnominal partitive expressions are typically realised as genitives (51) or headed by prepositions (overwhelmingly *отъ*) (52).

Notice that PART is reserved strictly for adnominal partitives. It should *not* be used for partitive objects. These are OBJs, and only the morphology signals that they are partitives, see section 11.2 for further details.

The PART relation is not limited to the meaning ‘one of/some of’. It is also used for other part-whole relations, as in example (53).

Genitives required by numerals should be analysed as PART. Note that numerals will be heads when they require PART genitives, but ATRs on the counted item when adjectival, as demonstrated in (54).

In general, PART should not be used for ‘components’ such as body part nouns, or ‘the root of the tree’ and similar: there should be a *contingent* part-whole relationship, not a permanent one. The head will generally be a measure noun of some sort.

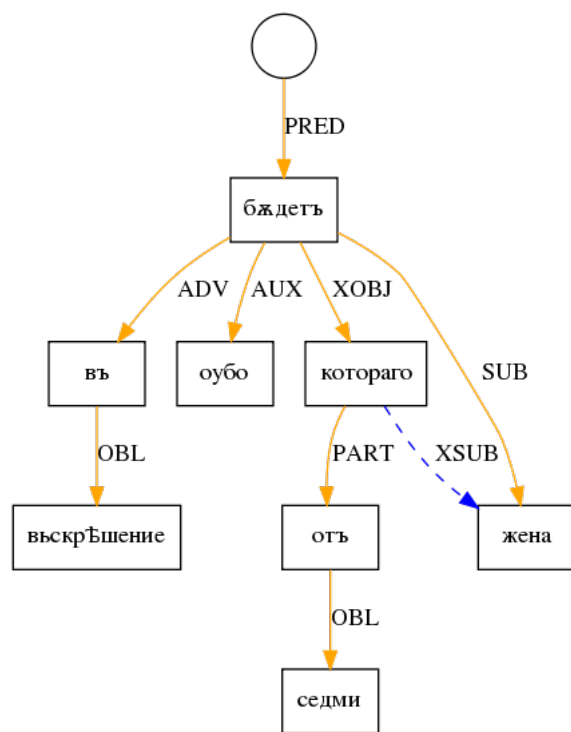


Figure 52: ‘Now then, at the resurrection, whose wife will she be of the seven?’
(Mar. Matt. 22.28, 39378)

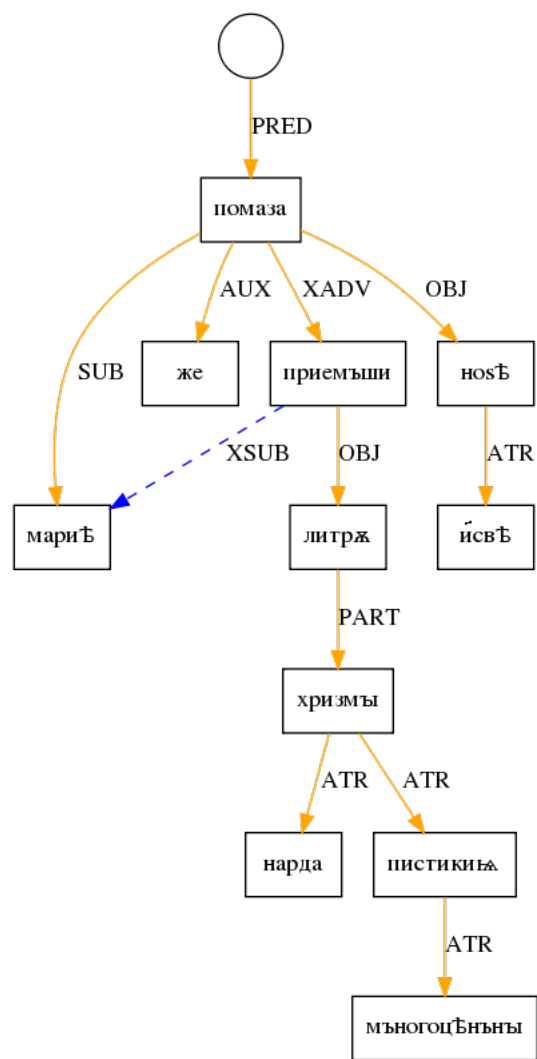


Figure 53: ‘Mary therefore took a pound of expensive ointment made from pure nard, and anointed the feet of Jesus.’ (Mar. John 12.3, 42476)

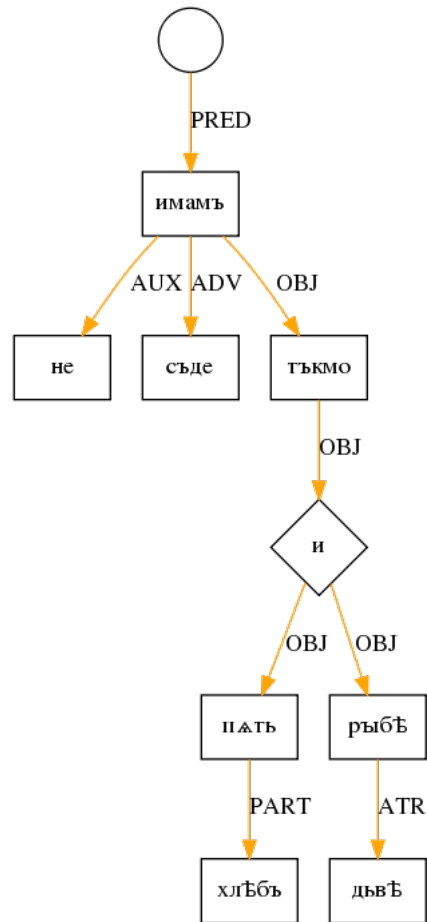


Figure 54: ‘We have only five loaves here and two fish.’ (Mar. Matt. 14.17, 50853)

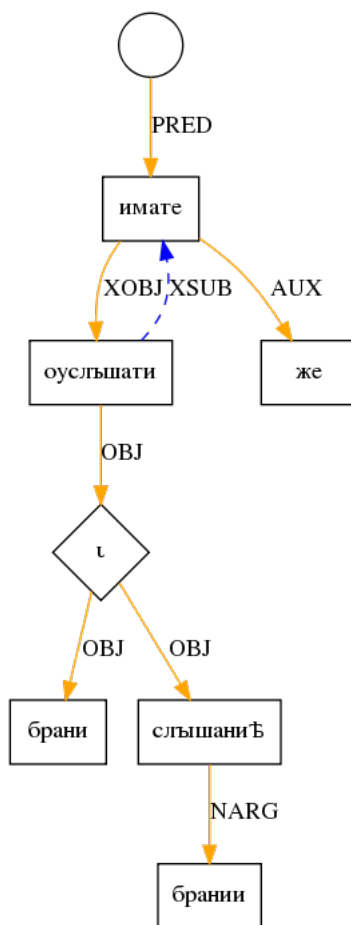


Figure 55: ‘You will hear of wars and rumors of wars.’ (Mar. Matt. 24.6, 39461)

13.5 Arguments of nouns

Nouns, especially deverbal nouns, can take arguments just like verbs.⁶ A clear case are so-called objective genitives (55), but objective datives are also very frequent (56), and NARGs may also have other realisations, they are for instance often headed by prepositions (57).

Infinitives are often NARGs when they are dependent on deverbal nouns (58), but they can also restrict the nouns, see section 19.22.

⁶However, nouns very rarely take direct object arguments, although some exceptions are found e.g. in the PVL.

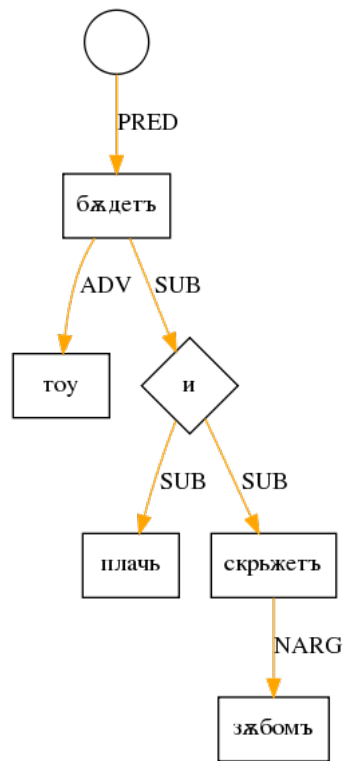


Figure 56: ‘In that place there will be weeping and gnashing of teeth.’ (Mar. Matt. 25.30, 39556)

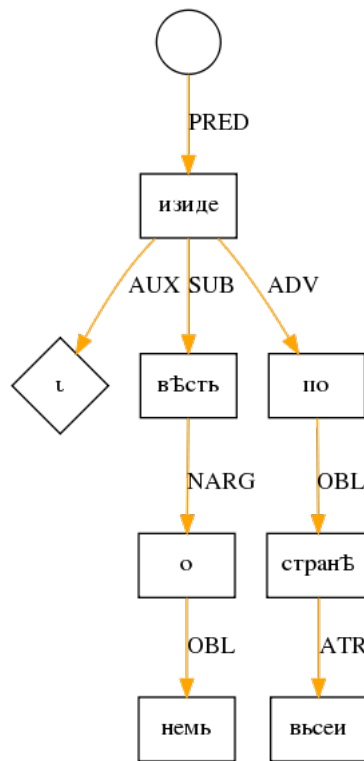


Figure 57: ‘And news about him spread through the whole countryside.’ (Mar. Luke 4.14, 40114)

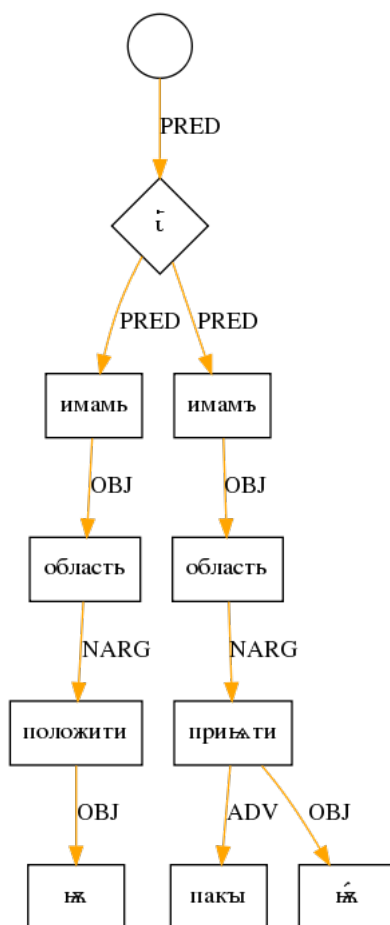


Figure 58: ‘I have authority to lay it down and authority to take it up again.’
(Mar. John 10.18, 42343)

14 Embedded predication

Embedded predications are divided in two major classes according to whether they can have their own, overt subject or not. Predications which cannot have an overt subject are called *open*, because their subject is supplied from outside the construction. They include some infinitives, such as the complement infinitives of *mošti*, and conjunct participles, whose subject is always coreferent with some other element of the sentence, but never expressed as a dependent of the participle. Open predications are further discussed in section 14.8. Sections 14.1 to ?? describe various kinds of closed predications.

Open and closed predications

The distinction between open and closed predications has to do with being a sentence or not. ‘Real’ subordinate sentences, introduced by subjunctions, have all it takes to be a sentence: a subject, a predicate, aspect, voice, etc. etc. Accusative with infinitive structures and absolute constructions also have these characteristics. So we can refer to these as subordinate *sentences*. Conjunct participles are also like sentences in some ways, but differ in a crucial respect, since they cannot not have their own subject. Note that closed predications may have their subject omitted for various reason (most frequently due to topicality), so the distinction does not hinge on the presence of an overt subject, but on the *possibility* of such a subject.

14.1 The basic types of subordinate clauses

Traditional grammar divides subordinate clauses into three groups: substantival clauses, adjectival clauses and adverbial clauses. This distinction reflected in our annotation.

14.1.1 Substantival clauses

Substantival clauses are also called complement clauses because they complement the main verb. They often seem to have the same functions as nouns, most often as objects, but sometimes as subjects. However, we do not attempt to keep such functions apart, but use COMP for all of them. We treat them this way because it is not always easy to tell whether the clause in question is the subject or the object.

In effect COMP ‘overrules’ other relations whenever an argument is sentence-formed. Most of the time it functions as a subject or an object, but sometimes sentences can be arguments of nouns, in which case they are marked as COMP rather than NARG; or as predicatives (in sentences like ‘My will is that’), where they are COMP instead of XOBJ.

Substantival clauses are typically subordinate clauses introduced by subjunctions such as *jako* ‘that’ (59) (not in the meaning ‘because’, which typically introduces adverbial clauses).

Types of complement clauses

Indirect questions (60) are also common, and we also find accusatives with infinitives (mostly Greek-inspired) and datives with infinitives.

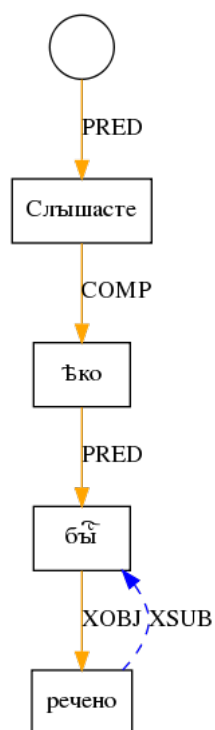


Figure 59: ‘You have heard that it was said’ (Mar. Matt. 5.38, 50673)

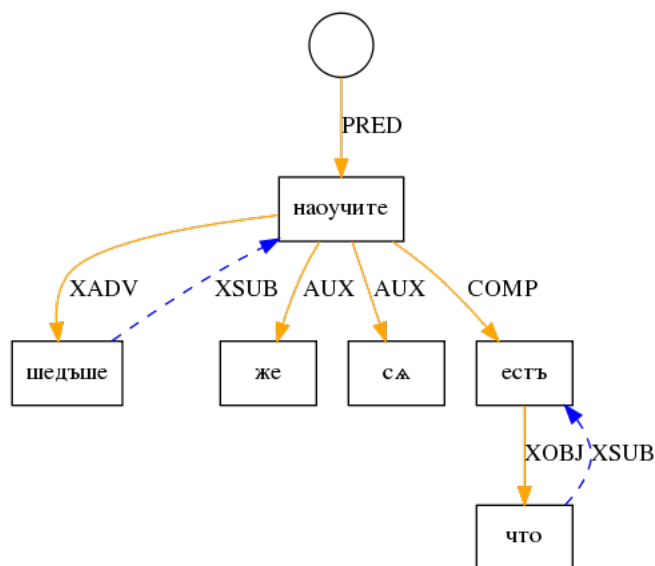


Figure 60: ‘But go and learn what this means’ (Mar. Matt. 9.13, 38557)

The criterion for a clause to serve as a COMP is that it is selected by the verb: in (61), *xotěti* ‘want’ selects a *da* clause, for instance, so this clause is a COMP.

Notice that a *da* clause which depends on *xotěti* differs from the type of *da* clause which can be added to any verb, to express intention.

Sometimes a subordinate clause is not introduced by any subjunction at all, and insofar as such subordinate clauses are required by the main verb, they are COMPs. (Adverbial subordinate clauses without subjunction are also possible.)

For guidelines on how to annotate the various complement sentences, consult sections 14.2 (clauses with subjunctions), 14.3 (finite clauses without subjunctions), 14.6 (accusative/dative with infinitive) and 14.5 (indirect questions).

14.1.2 Adjectival clauses

Relative clauses are adjectival in the sense that they have the same functions as adjectives. They are often adnominal (ATR and APOS), but can also be ‘substantivized’ and used directly for sentence-level functions without having an antecedent.

Relative clauses are the only adjectival clauses, and their annotation is further described in section 14.4.

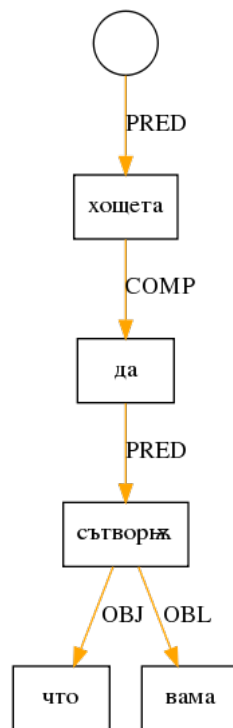


Figure 61: ‘What do you want me to do for you?’ (Mar. Matt. 20.32, 39240)

14.1.3 Adverbial clauses

Adverbial clauses are clauses which express information about time, intention, result, cause, concession, condition etc. Absolute constructions (section 14.7) belong here. Otherwise, they are normally introduced by subjunctions and treated in section 14.2, though asyndetic adverbial clauses also exist. Some subjunctions which introduce adverbial clauses are *ašte*, *jako*, *da*, *donědeže*, *zane*, *eda*, *poneže*, but there are many more.

14.2 Subordinate clauses with subjunctions

In subordinate clauses, except relative clauses, the verb is attached to the subjunction via the same relation PRED that we use in main clauses. The subjunction is related to the matrix clause via COMP if it is a complement clause (59).

Only a restricted group of subjunctions introduce COMPs, see the discussion in section 14.1.1. Many more subjunctions introduce adverbial clauses, consult section 14.1.3.

Verbs of saying such as *glagolati* or *rešti* may take a complement clause introduced by *da*, in which case the speech verb itself is interpreted as expressing the will of the speaker (62).

When a speech verb is overtly present, we assign the relation COMP to the subjunction *da*. When no speech verb is present, however, we take *da* to function as an ‘imperative’- or ‘subjunctive’-marking particle, with the whole clause acting as a main clause equivalent to an imperative clause. In such examples we lemmatise *da* as an adverb rather than a subjunction, and make it an AUX dependent on the main verb (63).

The PRED under a subjunction is usually finite, but note that infinitive PREDs are quite common as well (64).

Notice that in most cases the subjunction will only have a single PRED daughter, the verb, although sometimes a subjunction can be modified, for instance a negation may have scope over the whole dependent clause, as *ne* has over the *zane* clause in (65).

**Infinitive
PRED
Modification
of a subjunction**

14.3 Finite subordinate clauses without subjunctions

It can happen that a finite subordinate clause is not introduced by a subjunction, e.g. when *jako* is left out. In such cases, the function of the subordinate clause is annotated directly on the verb. Both COMPs (67) and ADVs without subjunctions are sometimes found in the Slavic material.

Conditional clauses seem especially prone to occur without subjunction (??).

14.4 Relative clauses

Relative clauses differ from other finite subordinate clauses in that they do not have a subjunction. Instead, they are introduced by an element (the relative

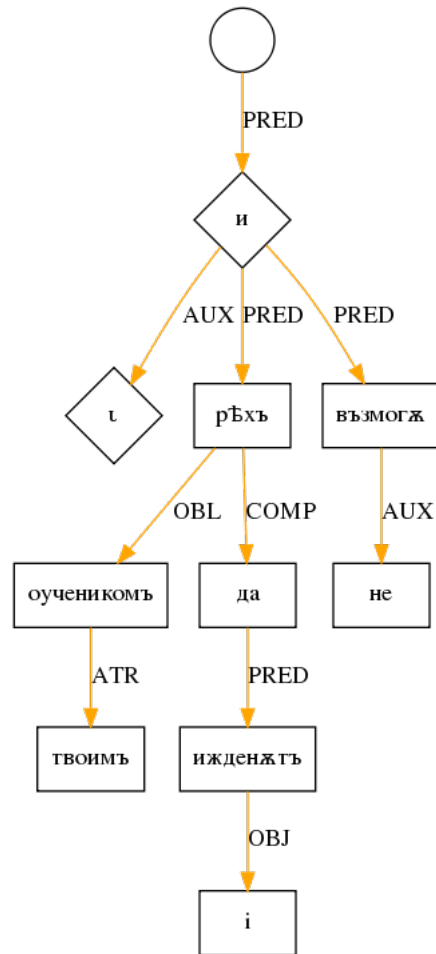


Figure 62: ‘I asked your disciples to drive out the spirit, but they could not.’
(Mar. Mark 9.18, 36831)



Figure 63: ‘Your kingdom come.’ (Mar. Matt. 6.10, 50685)

pronouns or adverbial) which also has a role in the sentence, such as subject, object, oblique or adverbial. Therefore these sentences should be headed by the verb directly, and the relative or interrogative should be annotated with its function within the clause. In other words, it is **never** correct to let an adverbial or relative word head a sentence. This also goes for relative adverbs such as *ideže* ‘where’, *jegda* ‘when’ and similar words.

Annotators are often tempted to let the relative pronoun depend on its antecedent: after all it agrees with its antecedent in gender and number. But we annotate syntactic *function*, and function is indicated by *case*. Thus, the verb of the relative clause must be made dependent on the relative pronoun’s antecedent (if there is one) since that is what the relative clause modifies.

The verb of relative clauses is given the relation corresponding to the function of the relative clause as a whole. There are several possible functions.

The verb of a relative clause with an antecedent is attached to the antecedent via the relation ATR (restrictive relative clauses) (68) or APOS (non-restrictive relative clauses) (69).

**Adnominal
relative
clauses**

Restrictive relative clauses restrict the reference of the head noun. In (68), the relative clause picks out exactly the place where the body of Jesus was lying before he was resurrected. Non-restrictive relative clauses don’t: in (69), the referent is already uniquely identified by deixis, and the relative clause just contributes additional information. If it is not possible to decide whether a

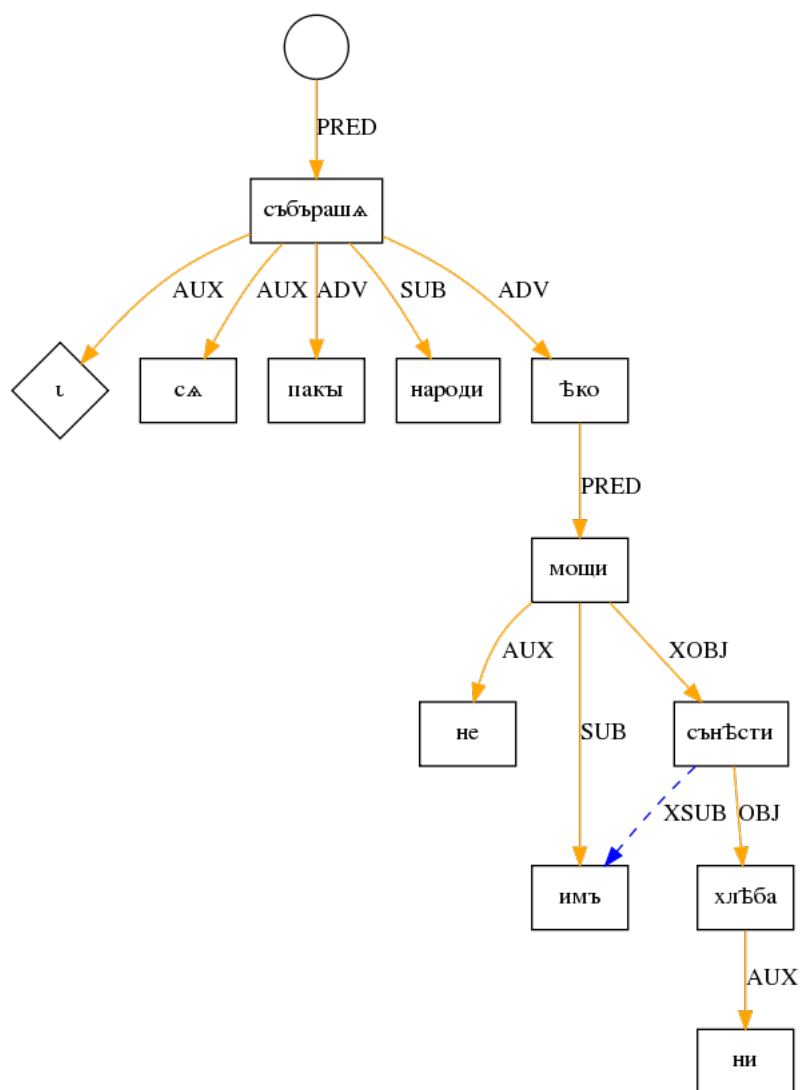


Figure 64: ‘And the crowd gathered again, to such an extent that they could not even eat a meal.’ (Mar. Mark 3.20, 36487)

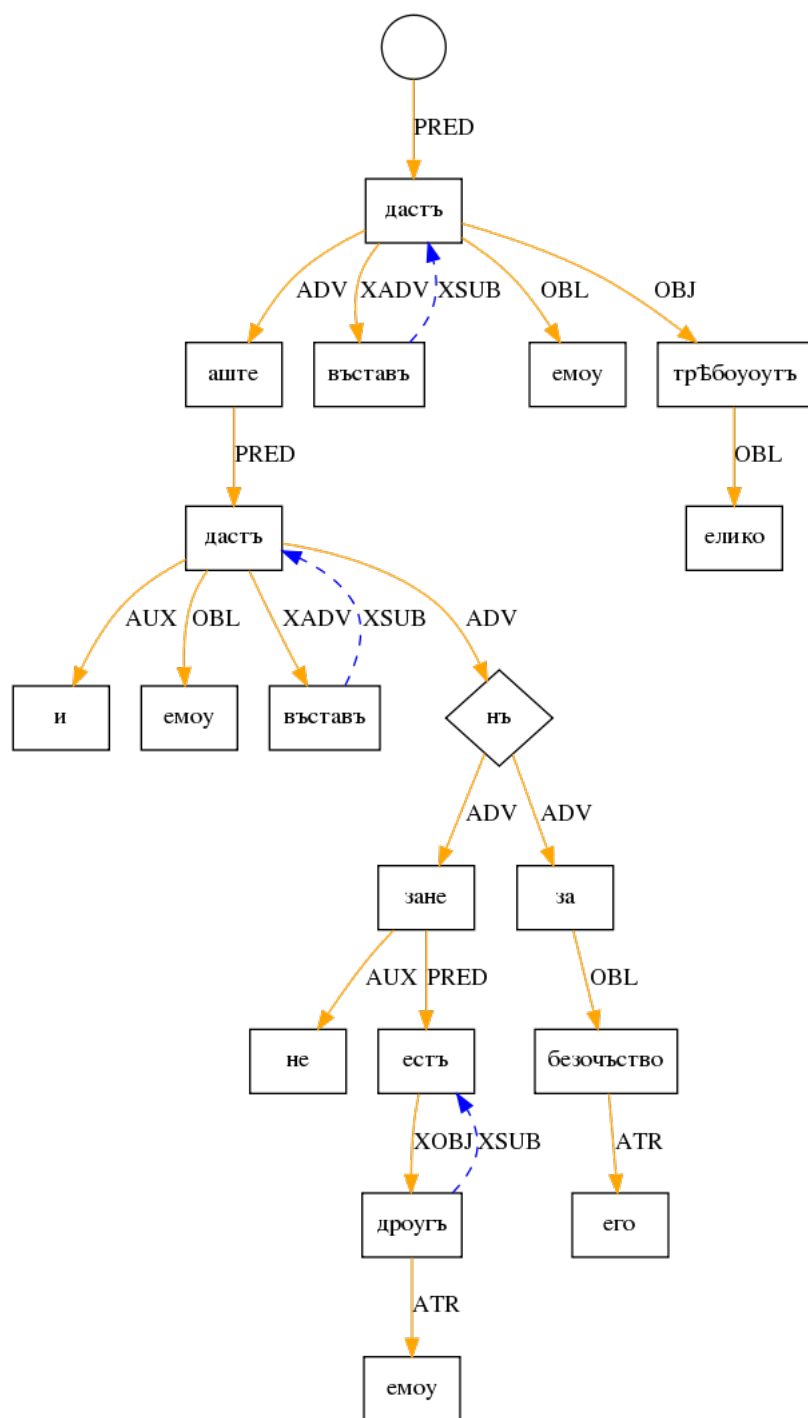


Figure 65: 'Though he will not get up and give him anything because he is his friend, yet because of his impudence he will rise and give him whatever he needs.' (Mar. Luke 11.8, 51456) 72

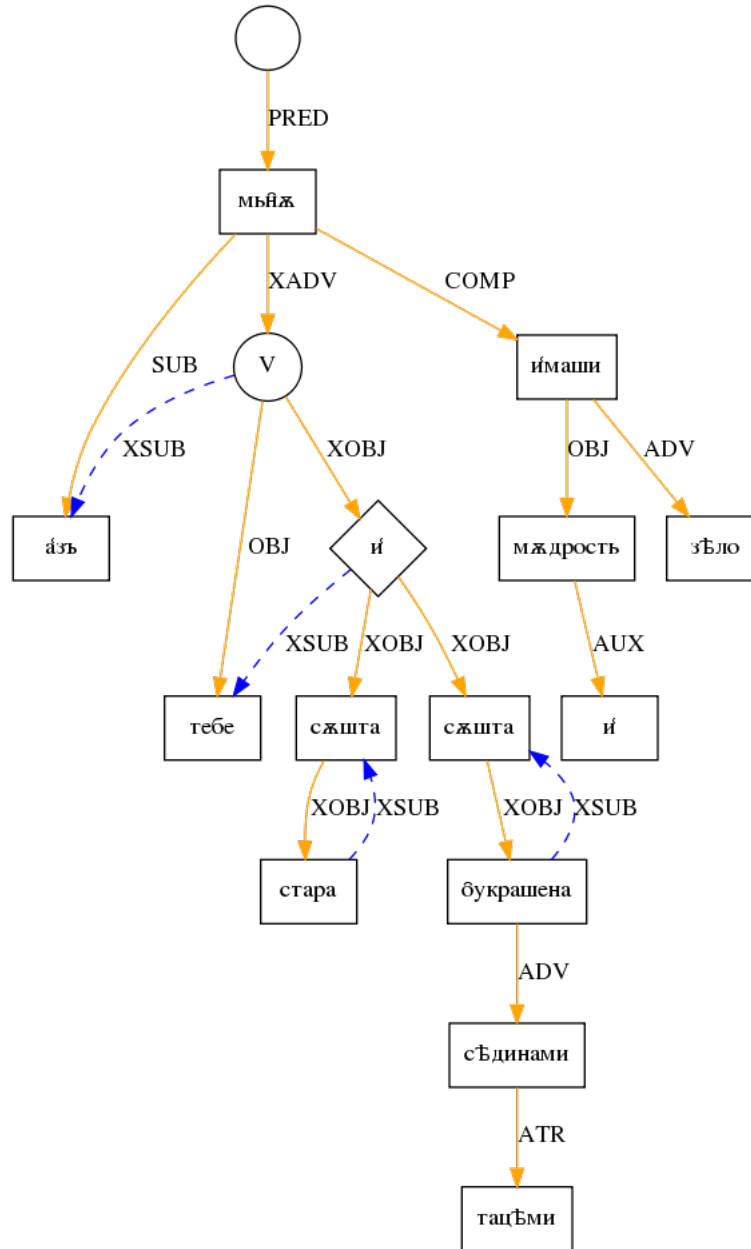


Figure 66: ‘Seeing that you are old and adorned with such gray hair, I think you also possess wisdom to a great extent’ (Supr. 3, 83099)

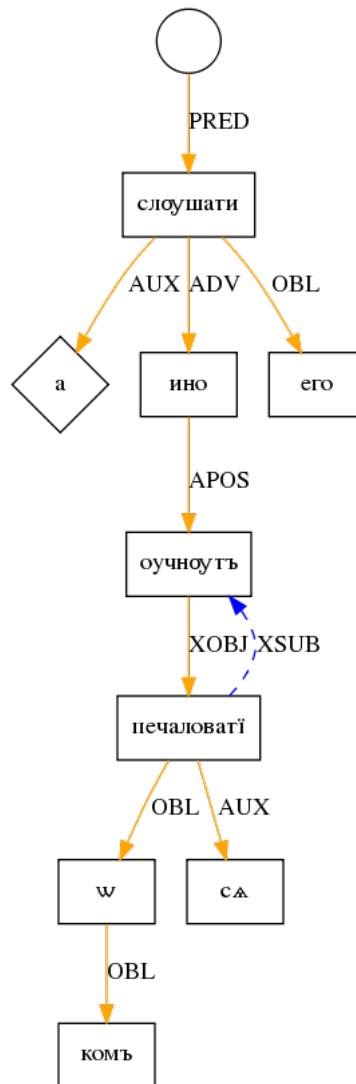


Figure 67: ‘And if they begin begging on behalf of someone, one must listen to it’ (Domostroj 14, 160531)

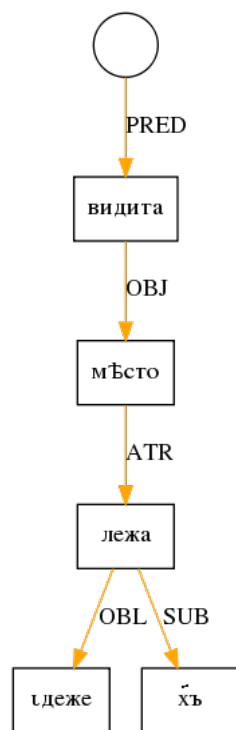


Figure 68: ‘See the place where his body was lying.’ (Mar. Matt. 28.6, 39820)

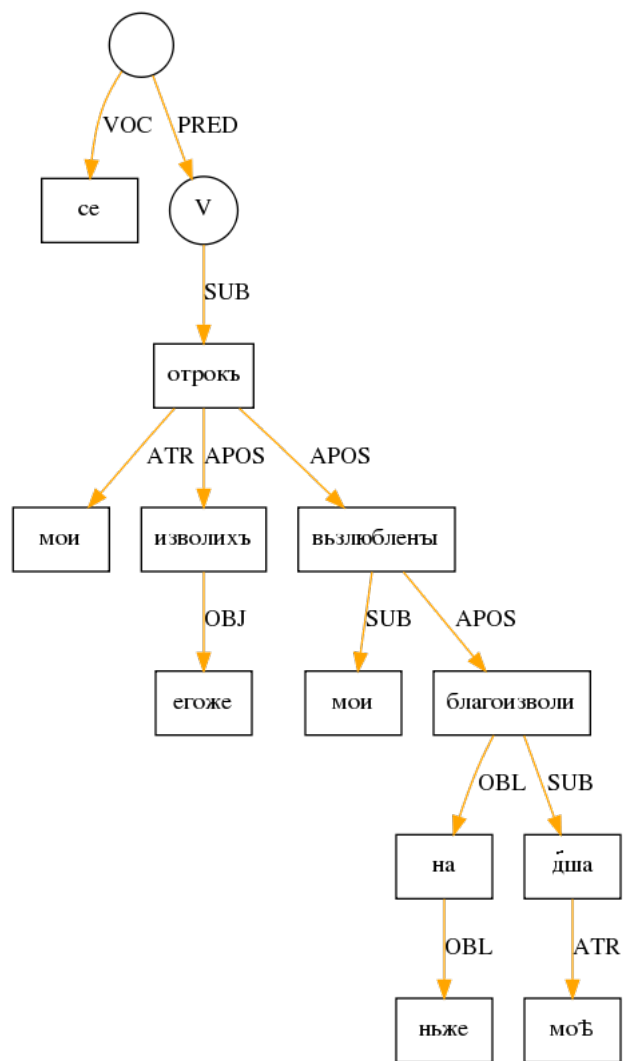


Figure 69: ‘Behold, my servant whom I have chosen, my beloved with whom my soul is well pleased.’ (Mar. Matt. 12.18, 38725)

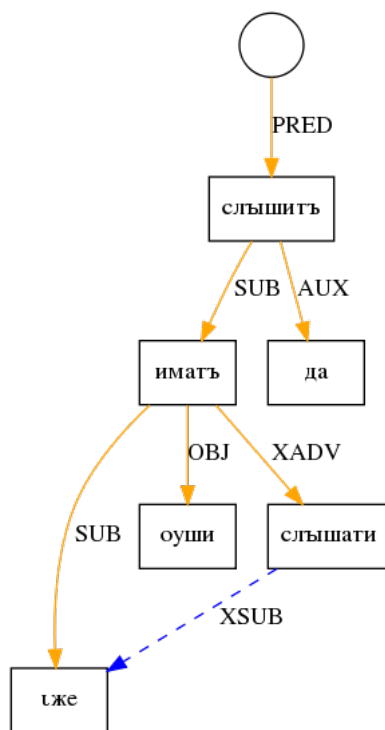


Figure 70: ‘He who has ears to hear, let him hear.’ (Mar. Mark 4.23, 36534)

relative clause is restrictive or not, use the supertag REL.

Whenever there are two verbs which share one relative word, we make the relative word dependent on the closest verb, but it is important to add a shared argument slash from the second verb to show that we are dealing with a relative clause. However, this is only possible when the relative pronoun has the same function in both clauses, see below.

It is important that not only clauses with relative pronouns, but also ones with relative adverbs should be annotated this way, since relative adverbs also have a function within their own clause. In (68), *ideže* serves as an OBL to the positional verb in the relative clause.

**Relative
adverbs**

If there is no antecedent for the relative pronoun, the verb of the relative clause will itself bear a sentence-level function within the matrix clause (70).

**Relatives
without
antecedent**

Although this implies that we take the relative clause to be a noun phrase, we avoid adding nominal modifiers to it, so quantifiers such as *всь* and *всьjakъ* are taken as heads of the relative clauses (71), see also ??.

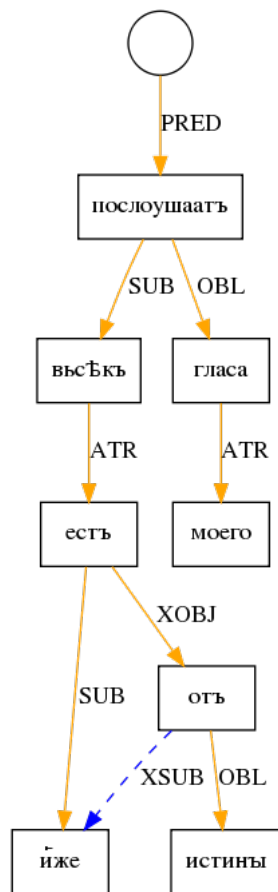


Figure 71: ‘Everyone who is of the truth listens to my voice.’ (Mar. John 18.37, 42838)

Notice that this style of annotation implies that the relation of the subordinate verb changes if the correlate is left out. Leaving out the correlate changes the function of the relative clause, which is annotated on the relative clause verb, from an attribute which restricts the correlate (as in example 71) to a direct argument of the matrix verb (as in example 70).

A variant of headless relative clauses are those which have an internal head. In this case the noun is given its function inside the relative clause, normally with the relative pronoun as an ATR (72).

**Internally
headed
relative
clauses**

Sometimes the correct position of the relative pronoun can be deep inside the clause it introduces: in (73) it is found inside a PP OBL dependent on an XOBJ to the relative clause head verb.

Sometimes a sentence where the relative pronoun does not have a role is coordinated with the real relative clause. We still coordinate such clauses, and use shared argument slashes where we can. In (74), the relative pronoun arguably does have a function in both sentences, but not the same function (it could be an ADV in the second clause), so we cannot express this in our notation.

In correlative constructions the relative sentence is made a dependent on the correlative: if the relative sentence precedes it is in most cases an APOS (as this is a topicalizing construction, cf. 19.20). Else it is normally, but not always, restrictive.

Finally, notice that infinitives sometimes can have a function similar to that of relative clauses, see section 19.22.

14.5 Indirect questions

Indirect questions are like relative clauses in that there is no element which introduces the sentence. The interrogative pronoun/adverb has a function inside the clause and must be annotated accordingly. The verb of the indirect question is given the function of the whole subordinate clause, which in the case of indirect questions is almost always COMP, as in (60).

However, in rare cases a dependent question can have adverbial (e.g. conditional) force (75).

14.5.1 Indirect questions vs. relative clauses vs. clauses with subjunctions

Since words such as *кто*, *что*, *как*, *которые* and similar words do not only occur in main clauses, but also in dependent clauses, their status is not always easy to determine. The pronominal words of this type typically have an interrogative lemma and an indefinite lemma. In OCS, this is probably enough, but in Old and especially Middle Russian it seems reasonable to posit relative pronoun lemmas on this form as well. Likewise, we may also need relative adverb lemmas for words such as *как*. This is because the system is shifting: whereas we can probably defend an analysis of all occurrences in dependent clauses in OCS as interrogatives in indirect questions (COMP), in Old/Middle Russian there are occurrences where we can no longer defend such an analysis, but must resort to

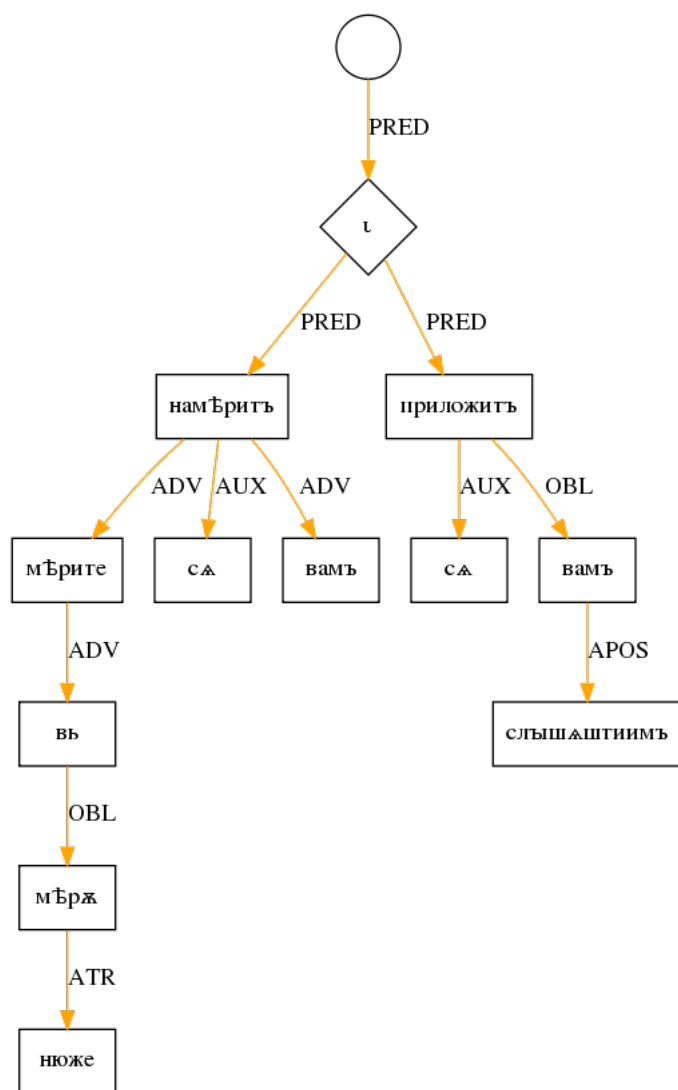


Figure 72: ‘With the measure you use, it will be measured to you, and still more will be added to you who hear.’ (Mar. Mark 4.24, 36537)

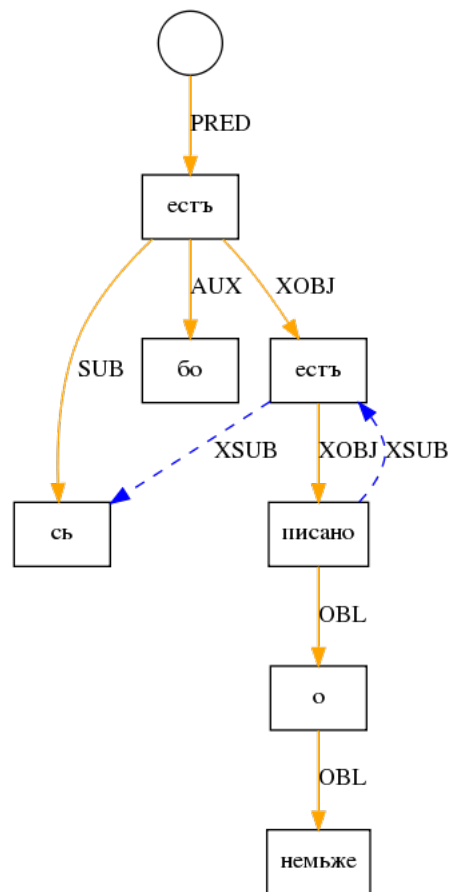


Figure 73: ‘This is the one about whom it is written.’ (Mar. Matt. 11.10, 50798)

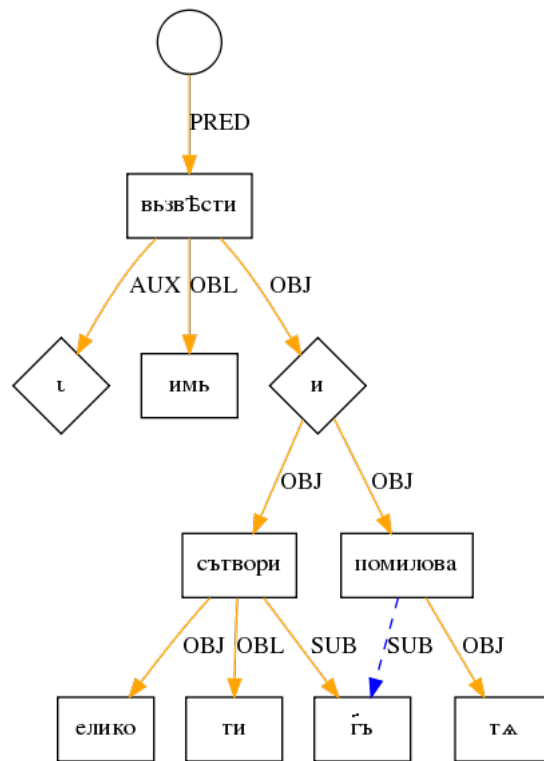


Figure 74: ‘And tell them everything the Lord has done for you and how merciful he has been.’ (Mar. Mark 5.19, 36587)

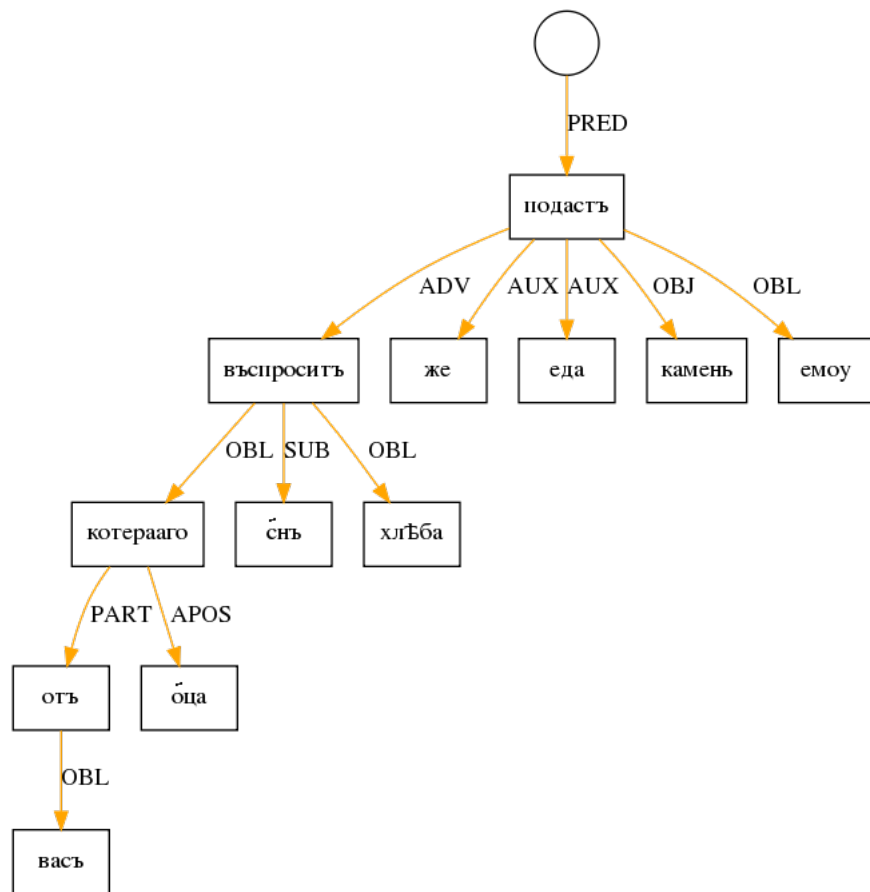


Figure 75: ‘Which of you fathers, if your son asks for bread, will give him a stone instead?’ (Mar. Luke 11.11, 40658)

a relative clause analysis or perhaps even a subjunction analysis (for *čъto*). We should still be cautious with the relative pronoun (adverb) analysis, and only resort to it when the context forces us to. The following is a check list to help deciding:

- If such words occur in direct questions, they should always be lemmatised as interrogative pronouns/adverbs. (*Who is he?*)
- If such words occur in main clauses that are not questions, they must be indefinite pronouns. (*Someone came in.*)
- If such words occur in a dependent clause which does not contain a subjunction and which serves as a subject matter argument to a speech, thought, perception or sometimes even emotion verb, then the dependent clause is an indirect question, and the *къto*, *čъto* etc. should be lemmatised as an interrogative. The whole dependent clause should be a COMP. (*He asked who had come in. He wondered who had come in. He noticed who had come in. He sensed who had come in.*)
- If *čъto* occurs in a dependent clause, but clearly has no syntactic role in the dependent clause, we must assume that it is a subjunction. (*On zametil, čъto oni prišli.*)
- If *къto*, *čъto* etc. occur in a dependent clause that is headed by a subjunction or in an asyndetic adverbial clause, we must assume that it is an indefinite pronoun. (*If someone kills a man, he must pay a fine.*)
- If such words occur in an adnominal dependent clause with no subjunction, then the dependent clause must be a relative clause and *къto* etc. must be lemmatised as a relative pronoun (adverb). (*He saw the man whom he had beaten up.*) The exception is if the dependent clause is an argument of a deverbal noun, in which case we may be dealing with an indirect question again. (*The question who did it.*)
- If a dependent clause contains *къto*, *čъto* etc. and serves as a subject or a non-argument to the main verb, it will typically be a relative clause, and the word in question must be lemmatised as relative. (*Who comes in the middle of the night must pay extra.*)

14.6 Non-finite predication: Accusative with infinitive and dative with infinitive

In Greek and Latin, accusatives with infinitives are very common, and often do the same job as complement clauses with subjunctions. They are found in Slavic as well, especially in OCS translations from Greek. Again we let the infinitive verb stand in for the whole construction. In the normal case this construction consists of a SUBject in the accusative, a predicate which is a verb

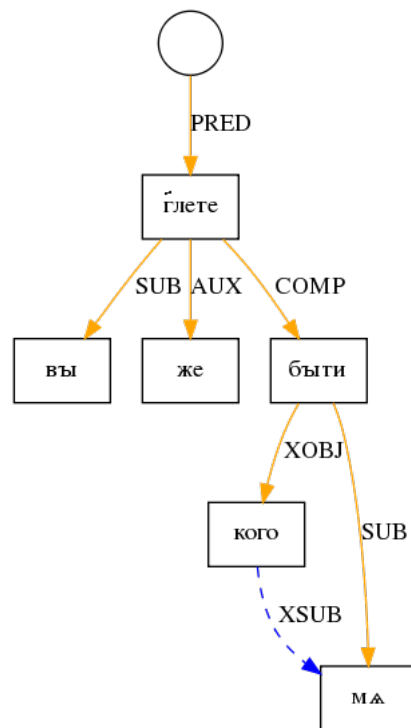


Figure 76: ‘But who do you say I am?’ (Mar. Matt. 16.15, 39009)

in the infinitive and possibly further arguments and adjuncts. The infinitive is given the relation COMP, as in (76).

In Slavic, datives with infinitives are quite frequent in such a function, often behaving very much like Greek accusatives with infinitives. The construction consists of a dative subject, a predicate which is a verb in the infinitive, and possibly other dependents, and the entire construction is a COMP to e.g. a speech verb (77). It should, in other words, be analysed exactly like an accusative with infinitive.

There are structures where one could be in doubt whether an accusative or dative noun is the OBJ/OBL of the main verb or the subject of the subordinate verb. If it is unclear whether the accusative/dative is really part of the main verb's argument structure, the XOBJ solution should be preferred. See also section 19.8.

**COMP vs
OBJ/OBL+XOBJ**

14.7 Infinite predication: absolute constructions

Absolute constructions are embedded predications with an adverbial relation to the rest of the sentence. We consider the participial verb to be the head of the construction. It is therefore related to the matrix clause via the relation ADV (78). The predicate in an absolute construction need not be a verb: it can be a noun or an adjective. In such cases, we assume an empty verb (79). Both in OCS and Old Russian, dative absolutes are quite frequent.

Occasionally, we see constructions that we must analyse as nominative absolutes. These are cases where a participle appears to have a subject in the nominative, and where that subject cannot conceivably be the subject of the main verb. Note that we do not accept word order alone as evidence that a nominative is the subject of the participle and not of the main verb, i.e. if a nominative appears to be a part of the participle constituent because of word order, that is not enough. We only choose the nominative absolute analysis when it's clear for other reasons that it can't be the subject of the main verb, as in (80), where the main verb clearly has a different subject.

**Nominative
absolute**

Nominative absolutes get the same syntactic analysis as dative absolutes: the participle is an ADV on the main verb, and the nominative is a SUB dependent on the participle. Note that there are special rules for 'all', 'self' and similar words, see section 19.1.

14.8 Open predications

Open predications are predications that do not supply their own subject, but get a subject via coreference relations⁷ within the sentence. Their subject is not necessarily a nominative. Finite verb forms never take part in open predication, only non-finite forms. The non-finite verb form heads the whole construction and is related to the matrix verb via the relation corresponding to the function of the embedded predication + a prefix X. We distinguish two kinds of open

⁷These relations are normally analysed in terms of *raising* and *control*.

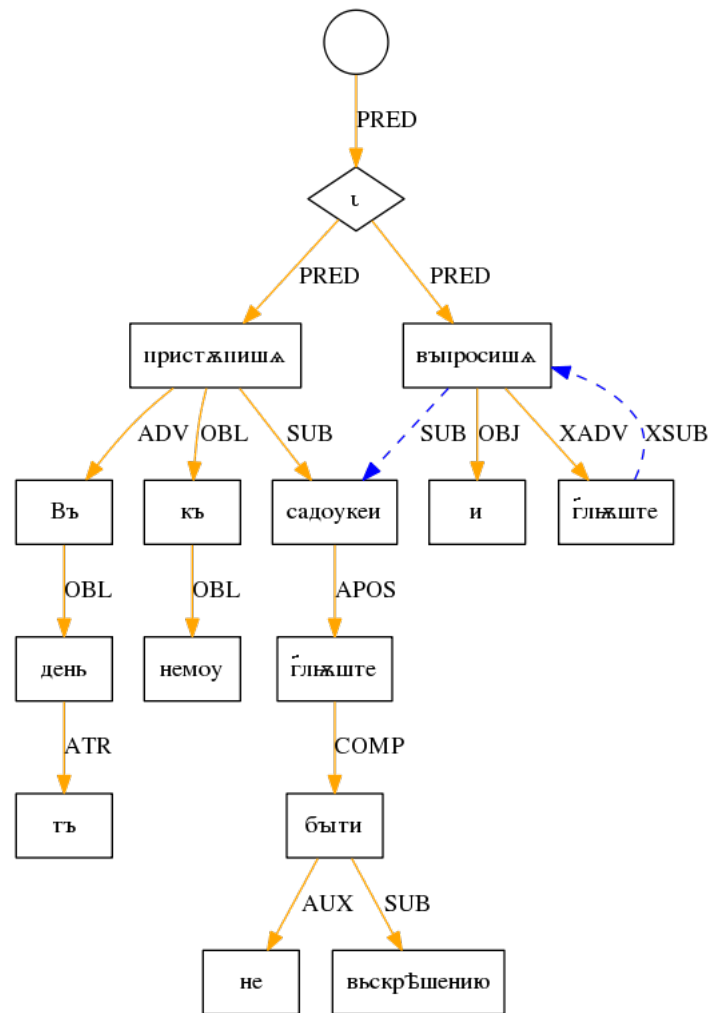


Figure 77: ‘That same day the Sadducees, who say there is no resurrection, came to him with a question, saying’ (Mar. Matt. 22.23, 39372)

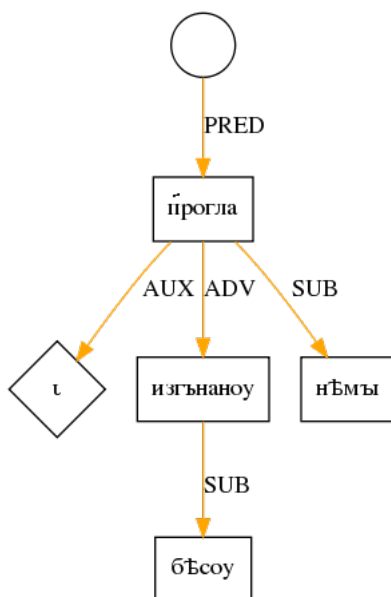


Figure 78: ‘And when the demon had been cast out, the mute man spoke.’
(Mar. Matt. 9.33, 38591)

predications according to whether they are arguments which *complement* a verb or adverbials which *modify* a verb.

In some cases, non-finite verb such as infinitives can depend on nouns. In these cases, they are never given the relations XOBJ or ADV, but are most often NARGs. Consult section 13.5, as well as 19.11 on infinitives.

14.8.1 XOBJ

Infinitives can be arguments of verbs. A fairly broad range of verbs take infinitives with external subjects, typically verbs expressing wanting, needing, ability, as well as phasal verbs (81), but also ones expressing permissions, orders etc. (82).

Such constructions are different from other embedded predications (such as accusatives with infinitives, absolute constructions etc.) in that no subject can be added. Instead the subject is supplied by coreference with an element which is either present in the matrix clause, such as *mi* in (82), or implied by the argument structure of the matrix verb, as the implied subject of *možeši* in (81). This coreference should be marked by means of the slash notation, see section 18.

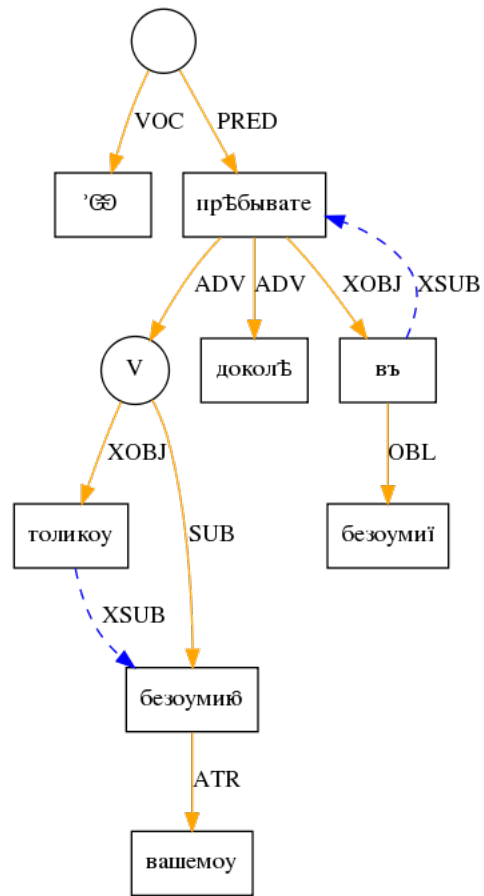


Figure 79: ‘Your madness being this great, how long will you remain in your madness?’ (Supr. 3, 82713)

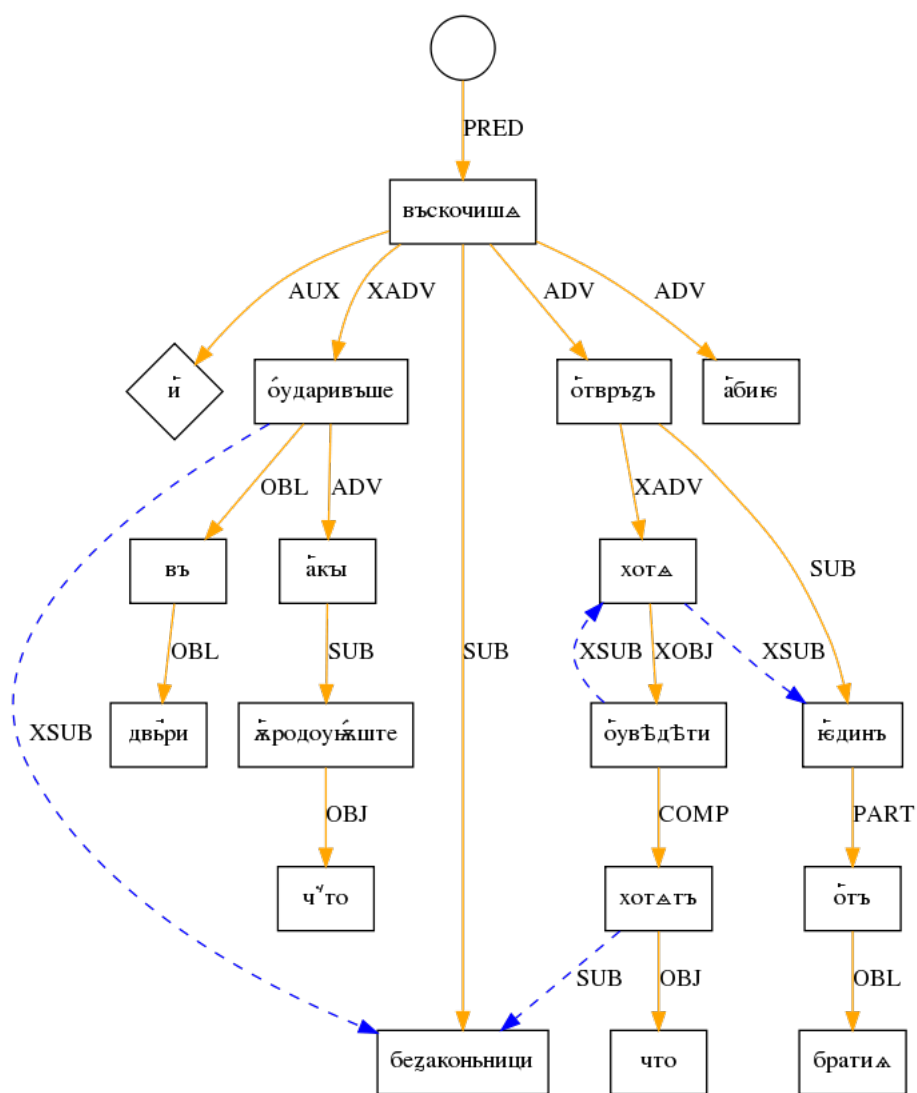


Figure 80: ‘They knocked at the door as if to ask about something and when one of the brothers opened to learn what it was they wanted, the bandits rushed in at once.’ (Supr.11, 122831)

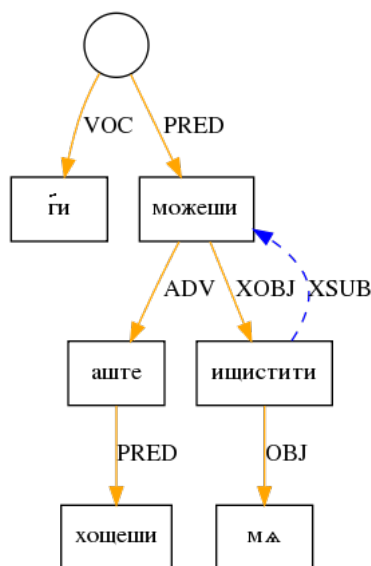


Figure 81: ‘Lord, if you are willing, you can make me clean.’ (Mar. Matt. 8.2, 50700)

This type of XOBJs are not always infinitives; a few verbs, such as *prĕstati* ‘cease’ in (83) may take a participle XOBJ. The coreference between the subject of the participle and the subject of the matrix verb is again marked via the slash notation. Also, perception verbs which take an accusative with participle construction are analysed as taking the accusative as an object and the participle as an XOBJ (84). This also goes for verbs meaning ‘find’ with the same structure.

**Participle
XOBJ**

14.8.2 XADV

Conjunct (or predicative) participles are adverbials which modify the matrix verb, to which they are attached via the relation XADV. They differ from absolute constructions (which are also adverbial participle constructions) in that they cannot have their own subject, but must have the subject supplied by coreference with an element which is either present in the matrix clause (as is the case for *prĭshedъ* and its external subject *isъ* in example 84) or implied by the argument structure of the matrix verb.

Like predicative participles, predicative adjectives are also given the tag XADV, as seen with *nagъ* ‘naked’ in (85).

Participles and adjectives are given the tag XADV when they are predicative,

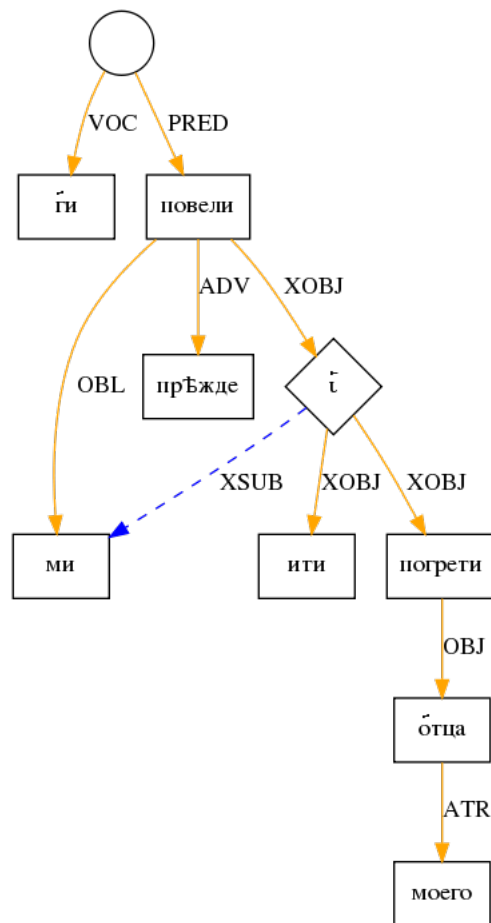


Figure 82: ‘Lord, first let me go and bury my father.’ (Mar. Matt. 8.21, 50729)

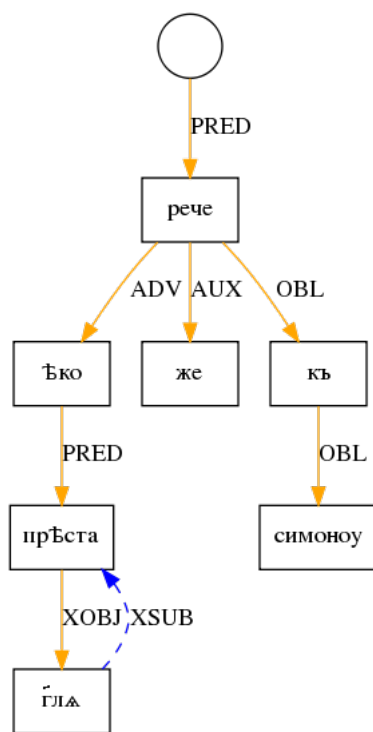


Figure 83: ‘When he had finished speaking, he said to Simon.’ (Mar. Luke 5.4, 40159)

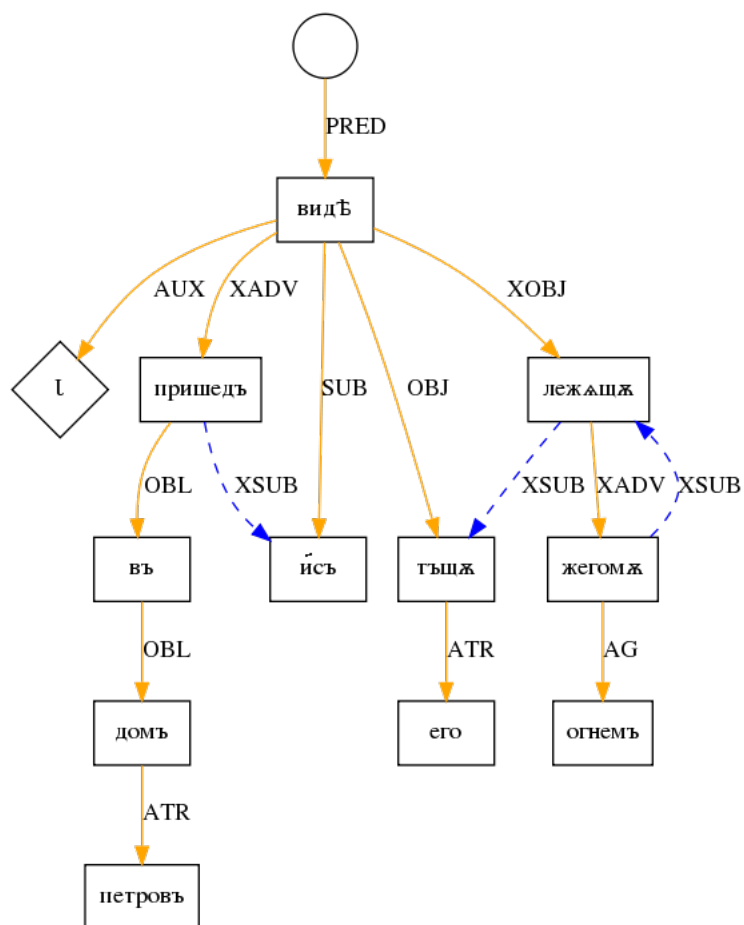


Figure 84: ‘When Jesus came into Peter’s house, he saw Peter’s mother-in-law lying in bed with a fever.’ (Mar. Matt. 8.14, 38512)

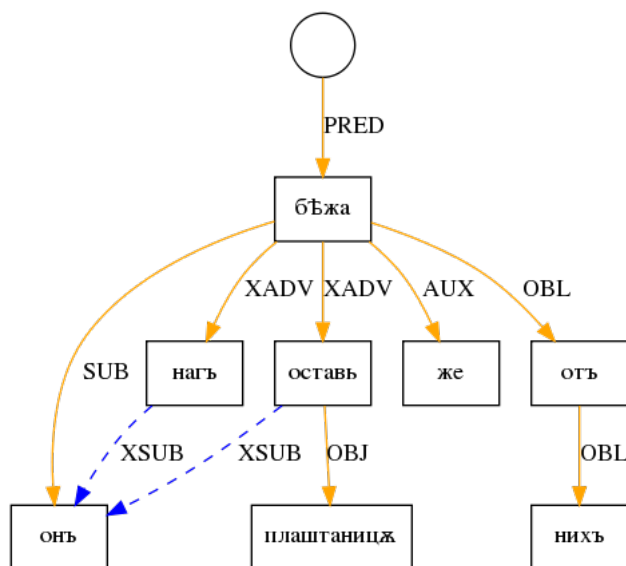


Figure 85: ‘But he left the linen cloth and ran away naked.’ (Mar. Mark 14.52, 37244)

ie. do not restrict the reference of a noun phrase but rather modify the main verb, i.e. there is a difference between ‘The man ran away naked’ (XADV) and ‘The naked man ran away’ (ATR).

OCS and Old/Middle Russian reciprocal pronouns have two components, and prepositions come between them, so they should not be merged. The first component is in the nominative, and should be analysed as an XADV, to capture its apposition-like relationship to the subject (86). The second component, on the other hand, will be an OBJ, OBL or ADV daughter of the verb, as context dictates. Both should be lemmatised as reciprocal pronouns.

Reciprocals

Infinitives of purpose and supines are also adverbial expressions and should be assigned to the relation XADV (87).

**Infinitives
of purpose
and supines**

14.9 Supertags

One supertag is relevant for embedded predications: REL, which is used whenever it is not possible to determine whether a relative clause is restrictive or not.

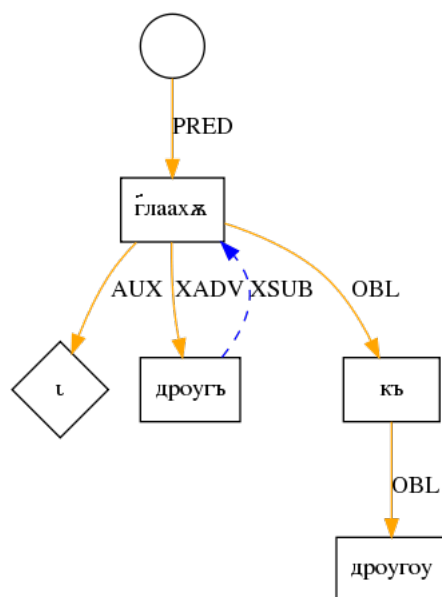


Figure 86: ‘And they said to each other.’ (Zogr. Mark 4.41, 175734)

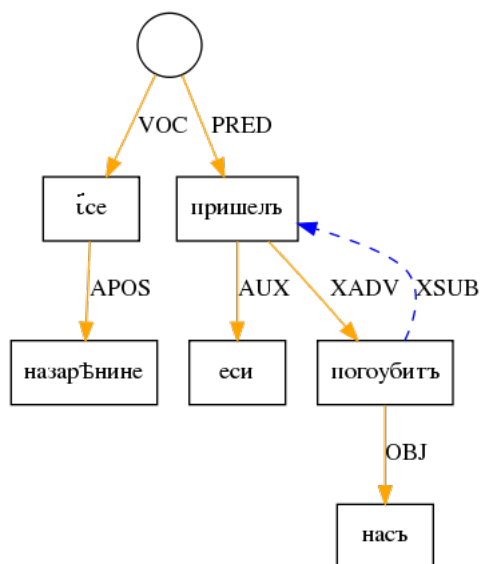


Figure 87: ‘Jesus of Nazareth, have you come to destroy us?’ (Mar. Mark 1.24, 36389)

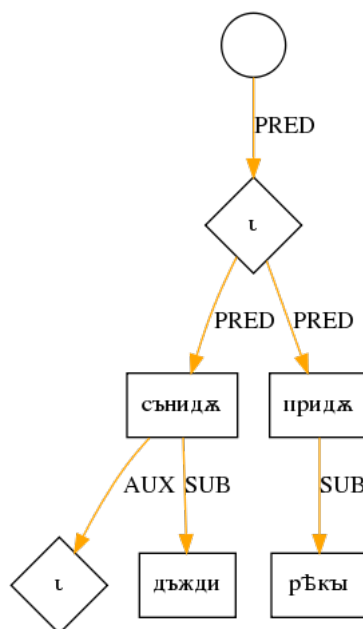


Figure 88: ‘And the rain fell, and the floods came.’ (Zogr. Matt. 7.25, 173141)

15 Coordination

Notice first that we do not recognise monopartite coordination. Thus, in the numerous cases in the TOROT texts where a sentence is introduced by a conjunction (*i* and others), this should not be treated as coordination, but the *i* should be attached to the sentence predicate via the relation AUX. In many sentences there is one introductory conjunction, which should be an AUX under the first verb, and then a coordination of two verbs (88).

We adopt a restrictive definition of conjunctions. Only items which can actually coordinate two words into a phrase are considered conjunctions.

Slash notation of shared dependents is used extensively in the examples this section, and is further explained and discussed in section 18, often with the same or very similar examples.

In analysing coordination, the first step is to ascertain what kind of elements are being conjoined: the conjuncts can be either constituents, multi-rooted elements, or sentences.

**Types of
coordination**

A constituent, in the sense of a dependency grammar, is a node together with all the nodes it dominates: for example, a noun with all its attributes and appositions is a constituent. In (89) *věra tvoě spe tē*, there are exactly

Constituents

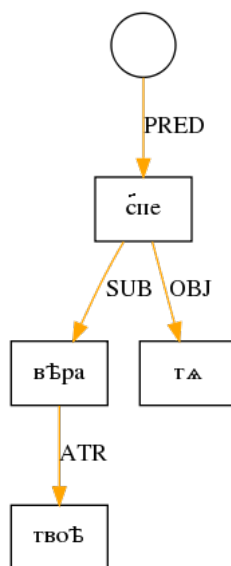


Figure 89: ‘Your faith has saved you.’ (Zogr. Mark 5.34, 191683)

four constituents: *tvoě*, *věra tvoě*, *tě*, and *věra tvoě spe tě*.⁸ However, the latter constituent, which corresponds to a sentence, is treated differently from other constituents in our annotation scheme. By definition a constituent has a single root.

A multi-rooted element is a set of nodes which are not connected, ie. do not have a common mother. Consider the example *prědastъ že bratrъ bratra na smъrtъ* (98) on page 109: the string or node set *bratrъ bratra na smъrtъ* has a common mother node, the verb, but this is not present in the string, so this set of nodes is as such multi-rooted. And still, this set can be coordinated with another element, as in *prědastъ že bratrъ bratra na smъrtъ. i otecъ čedo*.

A sentence is a constituent dominated by a finite verb, or by an empty node which stands in for a verb, or by the infinitive in an accusative with infinitive or dative with infinitive construction (the latter type is found both in main and dependent clauses).

**Multi-rooted
elements**

Sentences

15.1 Coordination of constituents/single-rooted elements

Coordination of constituents is the most straightforward case, especially when the constituents are single words. Consider the example ‘*čto estъ namъ i tebě*’

⁸In a phrase structure grammar, on the other hand, all words would be constituents, as would the VP (verb + object) *spe tě*.

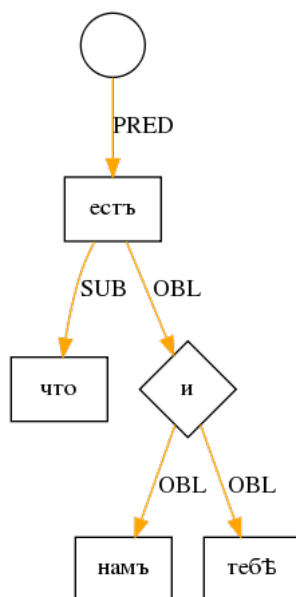


Figure 90: ‘What business do we have with each other?’ (Mar. Mark 1.24, 50225)

(90). *namъ* and *tebѣ* are conjoined by *i*. They are both OBLique arguments of *byti*, but so is the whole string *namъ i tebѣ*. We therefore want to have both *namъ*, *tebѣ* and *namъ i tebѣ* as constituents which bear the OBL relation, and we achieve this by having the conjunction dominating both *namъ* and *tebѣ* via the relation these would have had to the main verb if there was no coordination. The conjunction then inherits this relation and is attached to its head via the same relation that it bears to its daughter(s), as demonstrated in (90).

If there are more than two elements, all conjuncts are attached to the conjunction in the same way. Often there will also be more than one conjunction. In such cases, the first conjunction (not counting conjunctions which occur before the first conjunct!) should serve as a head of the coordinated phrase, and any further conjunctions should depend on the first one via the relation AUX (91).

If there is no conjunction present we insert an empty node which behaves as a conjunction.

We use essentially the same analysis whenever the conjuncts are single rooted but share a subtree. Consider the string *съ страхомъ i радостію велиею*. There are two possible analyses: the intended meaning can be ‘with fear (which is not necessarily great) and great joy’ (92) – in that case we have constituent coordination, since *радостію* and its attribute *велиею* is a constituent which is

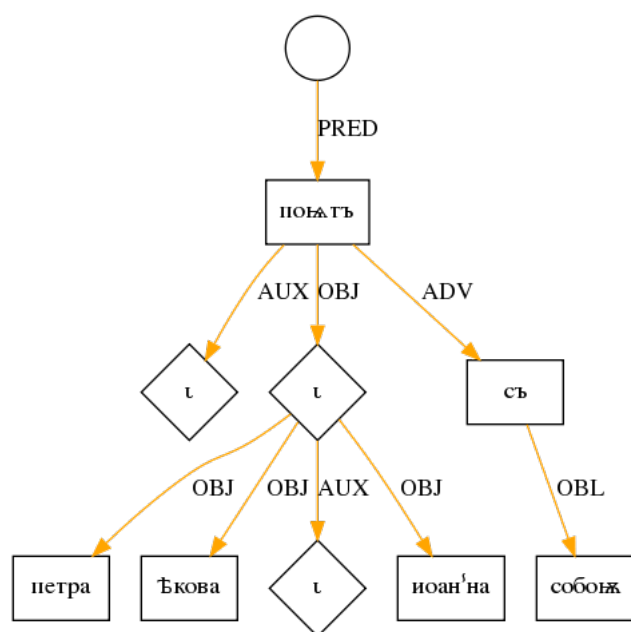


Figure 91: ‘He took Peter, James, and John with him.’ (Zogr. Mark 14.33, 177448)

related to the constituent *страхомъ*. But the meaning can also be ‘with great fear and great joy’ (93), in which case *velijejo* modifies both *radostijo* and *страхомъ*. The difference is captured in (92) and (93). In this way, the adjective has the correct scope, see section 19.7.

Notice that there is an alternative way to express shared attributes like this, namely through slash notation. The practice of putting shared elements on conjunctions is used when the coordinated elements are not finite verbs (or substitute for finite verbs); for finite verbs sharing elements, the slash notation *must* be used.

In analysis of single root conjuncts, it is important to ensure that both conjoined elements have the same function, i.e. they should bear the same grammatical relation. SUBs can only be coordinated with SUBs, OBJs with OBJs etc. Only in this way is it possible to give the conjunction a correct function.

**Identity
of function**

There are some exceptions to this, however. The XADV relation can be coordinated with an ADV element. In such cases the conjunction is given the relation XADV (and not ADV). This makes it possible to slash from the conjunction to the appropriate subject, and have the XADV inherit this slash. The rationale is that XADV and ADV are essentially the same function, the only difference is the presence/absence of an external subject.⁹

**Coordination
of ADV and
XADV**

Also, coordination of COMP and OBJ is allowed, and the resulting coordinator node should get the OBJ function. In a very few cases we also recognize coordination of COMP and OBL, and the coordinator node is labelled OBL in these cases.

**Coordination
of COMP and
OBJ**

If the relations of the two conjuncts are not identical, another form of coordination must be used. Of course, beside the conjuncts themselves, the coordination can have ‘superfluous’ conjunctions bearing the relation AUX, as well as items that are shared between the two conjuncts.

The principles described in this section are only applicable to **constituent** coordination. Whenever the two elements which each have more than one root are coordinated, the construction must be treated as described in section 15.3. One case which is not infrequent is the coordination of two object + object predicative structures. Even though we might like to conceive of an object + an object predicative as a single constituent, they have no common root except the governing verb in our model. We must therefore treat such coordinations as sentence coordinations, see example (97) on p. 108.

15.2 Sentence coordination

Although technically it would be possible to model VP coordination¹⁰ in the same way as other coordinations of single-rooted conjuncts which share a subtree, we do not pursue this option. Instead we follow the spirit of dependency

⁹Note that the relation between OBJ and XOBJ is different, since verbs subcategorise differently for these functions, whereas ADV and XADV are not subcategorised functions, i.e. not part of the verb’s valency.

¹⁰That is, the coordination of two verbs with their arguments, but without the subject.

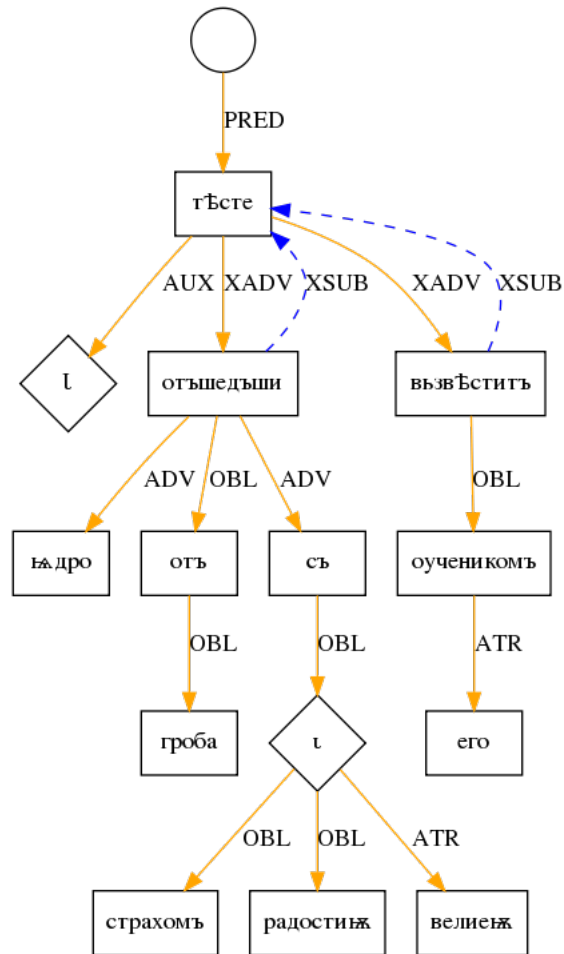


Figure 93: ‘So they departed quickly from the tomb with great fear and joy, and ran to tell his disciples.’ (Mar. Matt. 28.8, 39823)

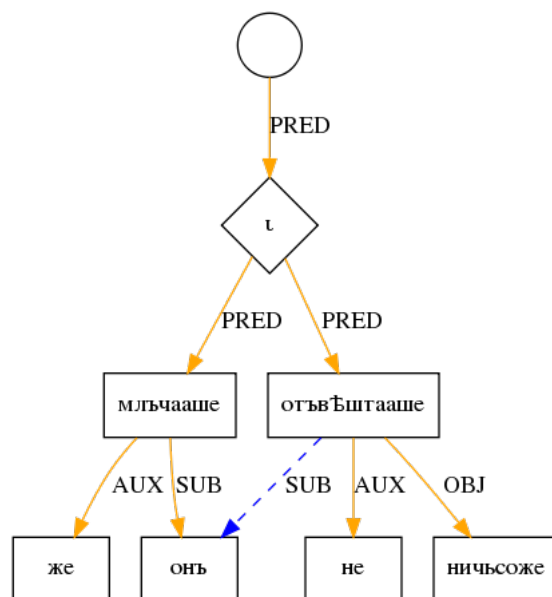


Figure 94: ‘But he remained silent and made no answer.’ (Mar. Mark 14.61, 37253)

grammar and assume that every finite verb, as well as infinitives in accusatives/datives with infinitives, forms a sentence. This also ensures consistency in the analysis, since - given the liberal use of pro-drop subjects in ancient Indo-European languages - it is often not possible to determine whether we have sentence coordination or VP coordination.

Consider *онъ же млъчааше. ꙗ ницьсоже не отгвѣштааше* ‘but he remained silent and made no answer’. There is no way of knowing whether the two conjuncts are *млъчааше* and *ницьсоже не отгвѣштааше*, with the subject belonging to both conjuncts, or rather *онъ млъчааше* and *ницьсоже не отгвѣштааше* with a zero subject in the second conjunct. To avoid making such decisions, we treat all such cases as sentence coordination and mark shared arguments via the slash notation, see section 16. The representation is therefore as in (94).

If there is more than one conjunction present, the first one to occur between two conjuncts will dominate the coordinated elements, whereas subsequent conjunctions will be attached to the first one via the relation AUX. Any conjunction preceding the first conjunct will be an AUX on that first conjunct. This combination is illustrated in (95).

In subordinate clauses without a subjunction, such as relative clauses and indirect questions, the verbs do not bear the relation PRED but a relation

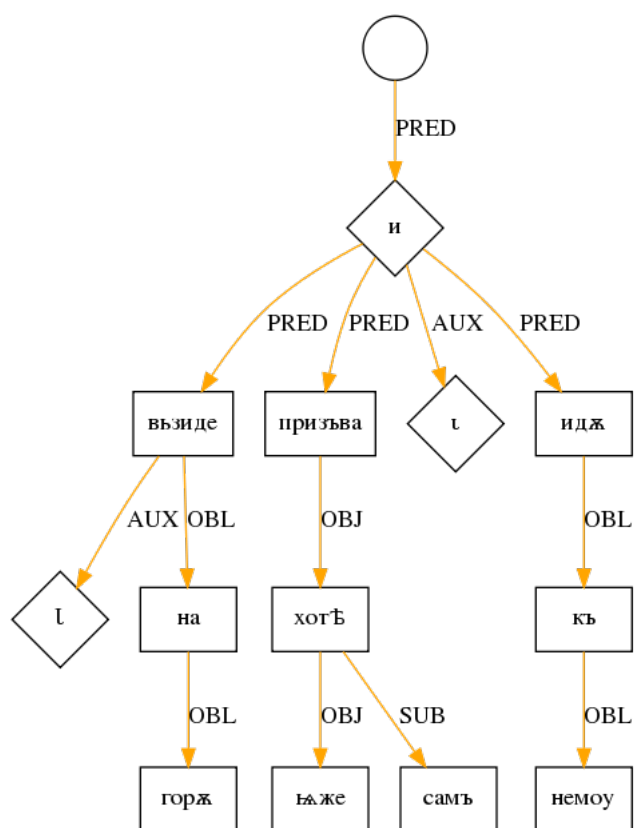


Figure 95: ‘And he went up on the mountain and called to him those whom he desired, and they came to him.’ (Mar. Mark 3.13, 36481)

corresponding to the function of the whole subordinate clause. Conjunct verbs are analysed as sentence coordination here too (96).

15.3 Coordination of non-constituents/multi-rooted elements

The conjuncts do not always have a single root. Consider a sentence such as ‘John drank coffee and Peter beer.’ The second conjunct here is ‘Peter beer’ which is not a constituent and does not have one root, but two unconnected roots, ‘Peter’ bearing a SUB relation and ‘beer’ an OBJ relation. To solve this problem, we must first create a common root for them. This will be an empty node functioning as a conjunction. The empty node and the sentence are then conjoined in a way similar to sentence coordination. Both are given the relation PRED, but the second conjunct lacks the verb and is therefore marked with a slash arrow towards the overt verb, ie. it is a kind of predication which shares the central element, the predicate, with the first conjunct. The solution is illustrated in (97) and (98).

As seen in (98) not only the predicate, but also other elements can be shared between the two conjuncts, in this case an OBL.

15.4 In which conjunct does an element belong?

It is sometimes not clear how the conjoined domains should be delimited and in which conjunct an element belongs. If the element is shared between the two conjuncts, the question is which dependency should be marked directly in the tree, and which one by a slash arrow. This does not affect the semantic interpretation of the sentence in any way, since the element is shared. For such questions, consult section 18.3.5.

It can happen that it is clear that an element belongs to only one conjunct, but it is unclear which one that is. In such cases, the choice will determine the interpretation. If in doubt, the annotator should consult standard translations if available. The principles in section 18.3.5 do not apply to such cases.

15.5 Other issues

Finally, it is important to distinguish the *conjunction* *i* ‘and’ from the *particle* *i* ‘also’. This latter should be attached to the word it emphasises via the relation AUX.

16 Gapping and ellipsis

16.1 Elided copulas

Since we take the verb to be the head of the entire sentence in our model, we always need a verb in the analysis. However, copular verbs are often elided. In such cases we insert an empty copula (99).

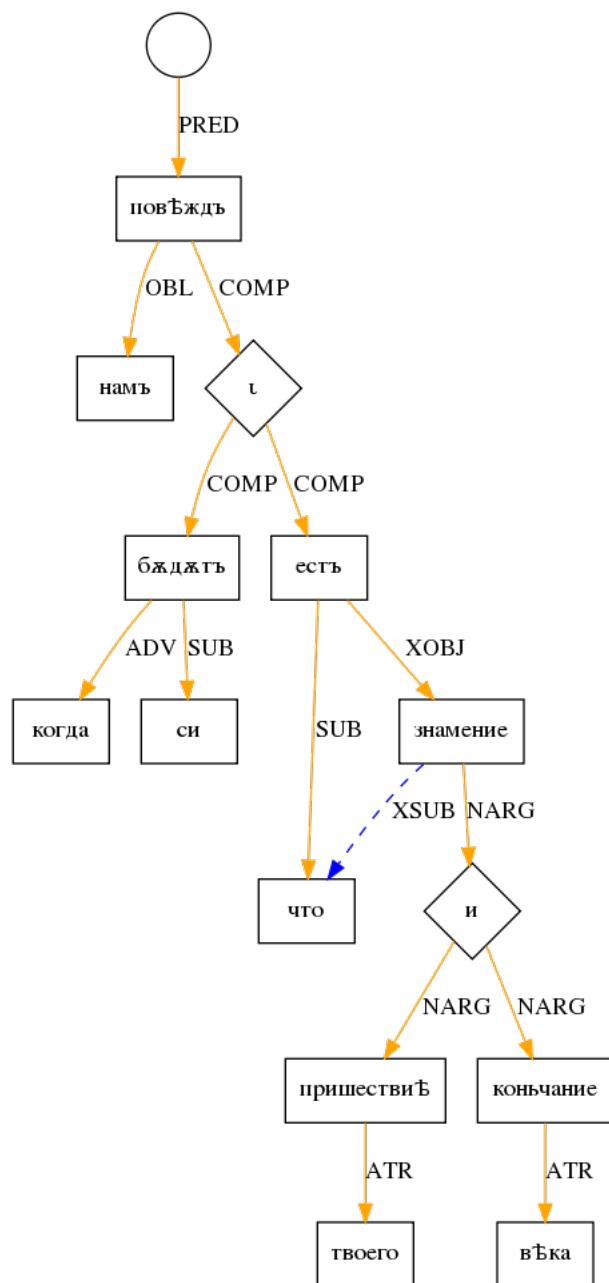


Figure 96: ‘Tell us when these things will be and what will be the sign of your advent and of the end of the world.’ (Mar. Matt 24.3, 51067)

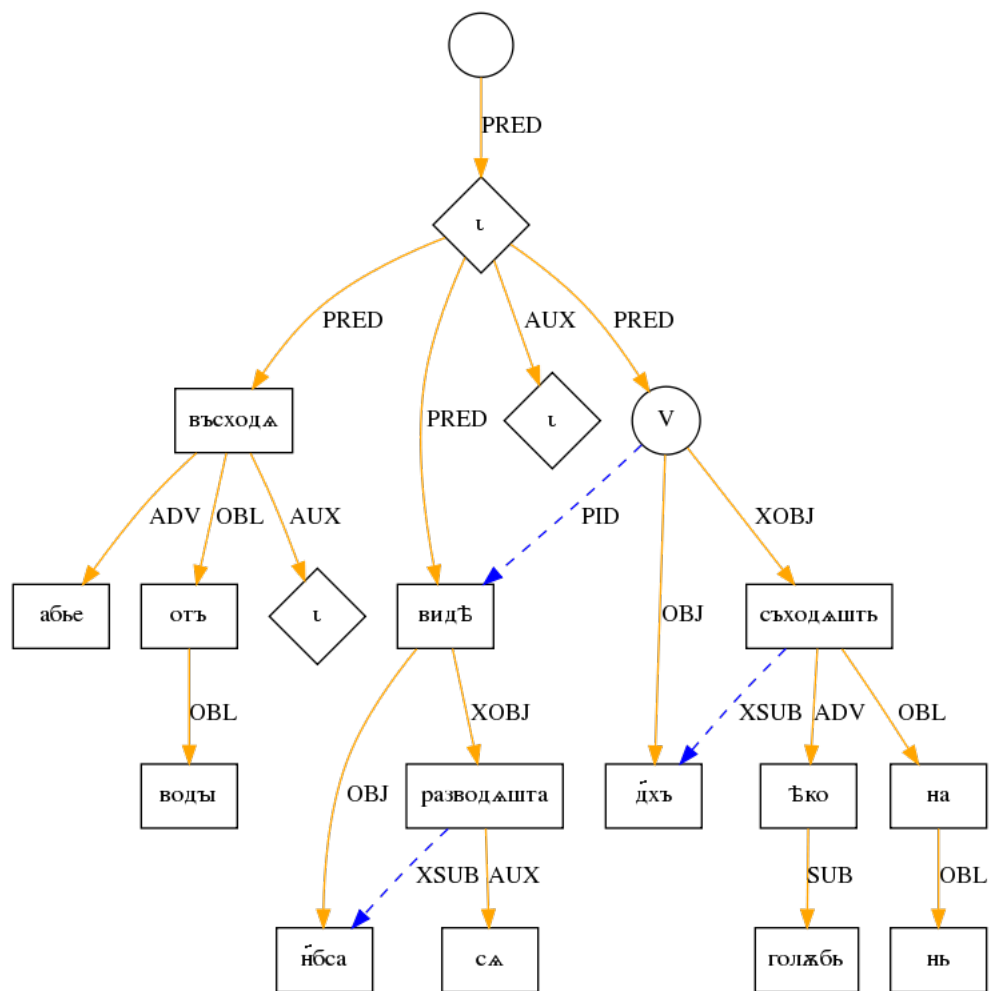


Figure 97: ‘And when he came up out of the water, immediately he saw the heavens being torn open and the Spirit descending on him like a dove.’ (Mar. Mark 1.10, 36370)

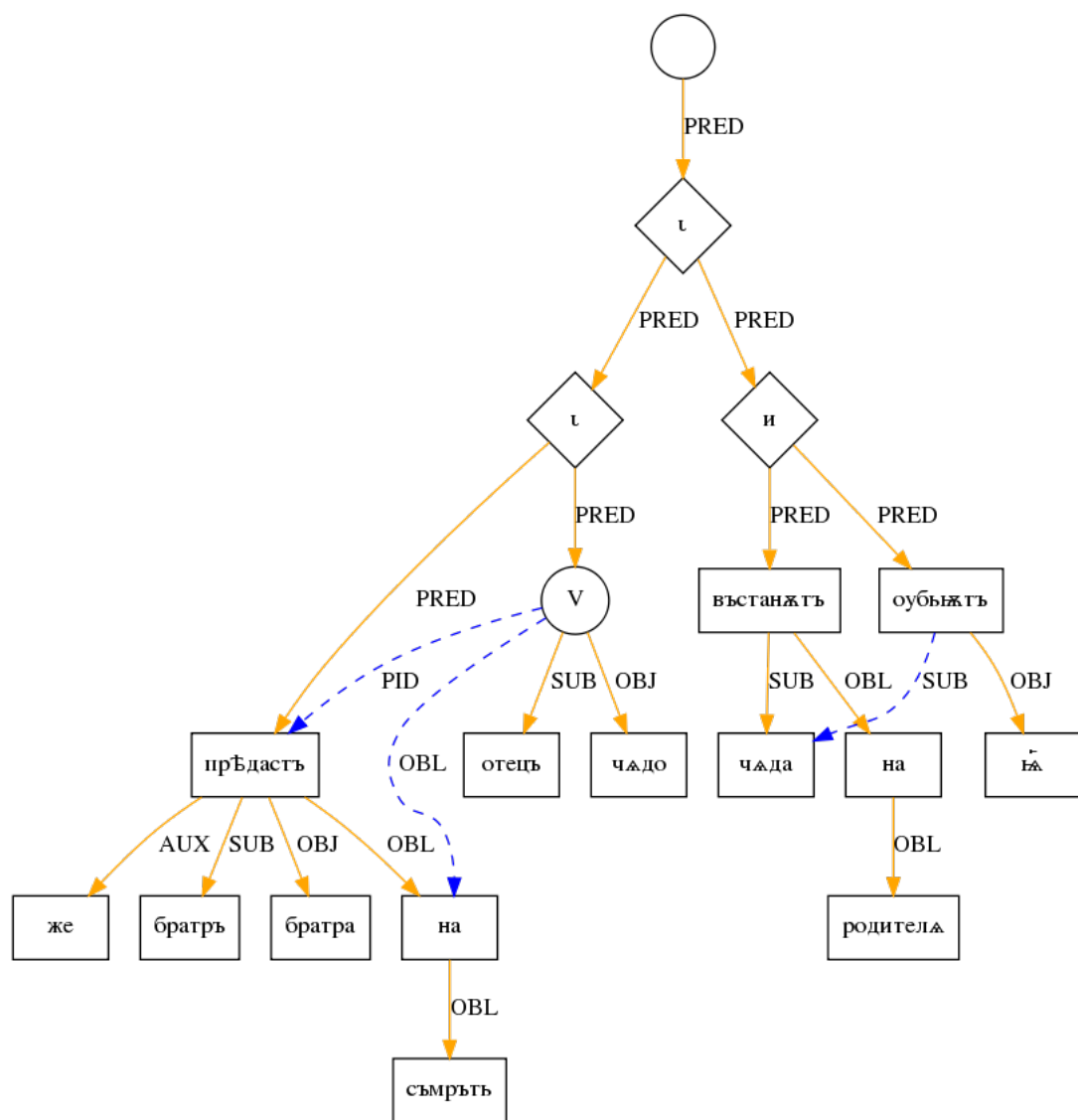


Figure 98: ‘Brother will betray brother to death, and a father his child; and children will rise up against parents and cause them to be put to death.’ (Mar. Matt. 10.21, 38628)

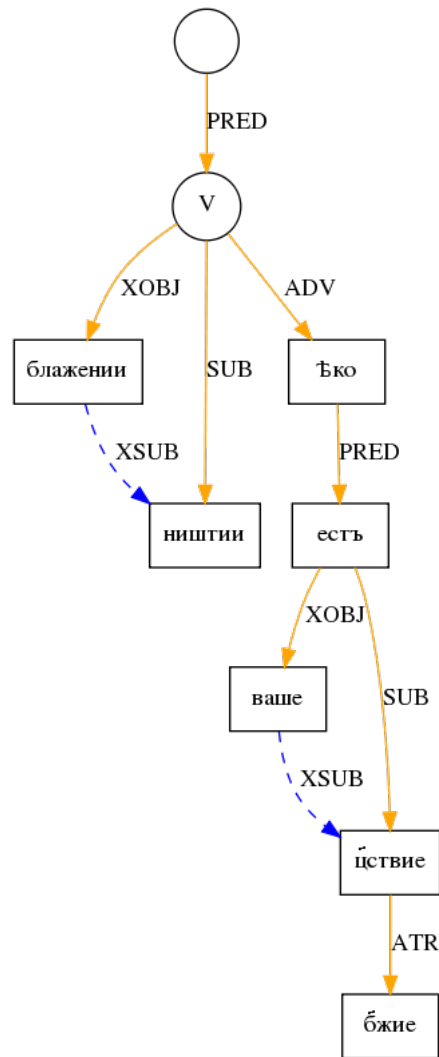


Figure 99: ‘Blessed are you who are poor, for yours is the kingdom of God.’
(Mar. Luke 6.20, 40245)

When the absent verb is not a copula, we will have to analyse this as gapping (see next section), ie. the structure should be attached to the sentence with which it shares a verb.

16.2 Gapping

Note that we do not mark ellipsis of nominal arguments as such, ie. we do not mark absent subjects in sentences such as *umьrětъ* ‘he died’; nor do we mark absent objects in sentences like *ubišę* ‘they killed him’. The only kind of ellipsis that we *do* mark is so-called gapping, i.e. cases where shared material in a coordinated structure is left out in one or more of the conjuncts. Such conjuncts are treated as sentences, ie. we insert an empty verb which is given the relation PRED, as explained in section 15.3, and the shared elements are indicated by slash notation. (100) provides an example with several layers of gapping.

We also mark arguments that are shared by a matrix verb and the verb of an embedded predication, for instance it is allowed to slash from a main verb to the argument of a dependent XADV participle.¹¹

17 The article

None of the TOROT languages have definite articles proper. Ancient Greek, of course, has them in abundance. They therefore have their own part of speech (article), and are analysed as an AUX on their nominal head. For details, we refer the reader to the PROIEL guidelines.

17.1 OCS mock article

Sometimes the translators seem to have used the relative pronoun *iže* to render the Greek article. In particular this happens with infinitives, clearly to mimic the way Greek uses the article to nominalise infinitives. In these cases we lemmatise *iže* as a demonstrative pronoun rather than an article or relative pronoun. In the syntax, we make it an *aux* on the infinitive and let it nominalise the infinitive (102).

This usage is also sometimes found with finite forms (in quotations), see ??.

In addition to the nominalised infinitives and finite verbs, we sometimes see constructions with *eže* that appear to mimic Greek nominalisations of PPs, and which we could potentially have analysed as mock article constructions. However, since it is usually possible to analyse these examples as a relative clause with a null copula, we have opted to do so, and reserve the mock article analysis for cases where there is no such clear alternative.

¹¹The opposite is impossible for technical reasons, since an outgoing slash from an XADV or an XOBJ is automatically interpreted as an XSUB slash by the annotation application. This sometimes causes loss of information.

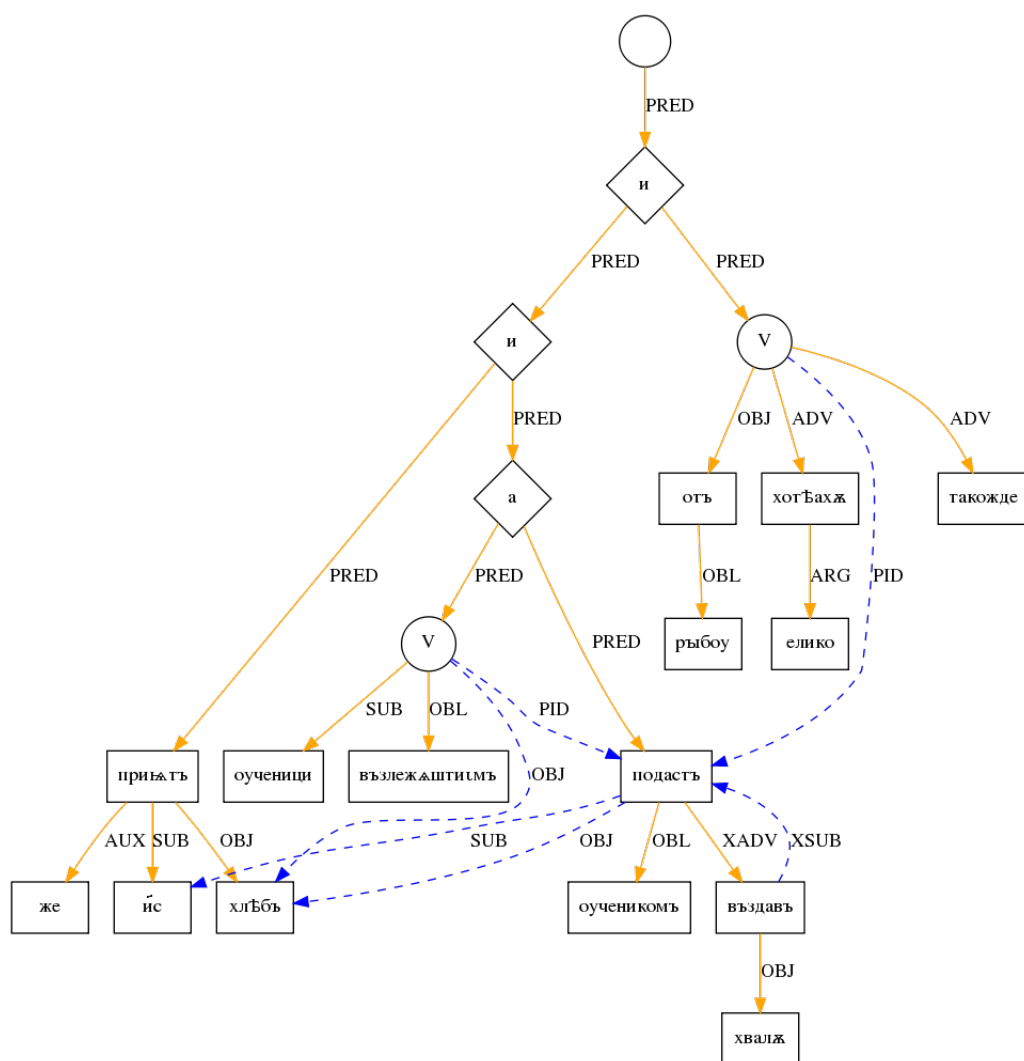


Figure 100: ‘And Jesus took the loaves; and when he had given thanks, he distributed them to the disciples, and the disciples to those who were seated; and likewise of the fishes as much as they wanted.’ (Mar. John 6.11, 41930)

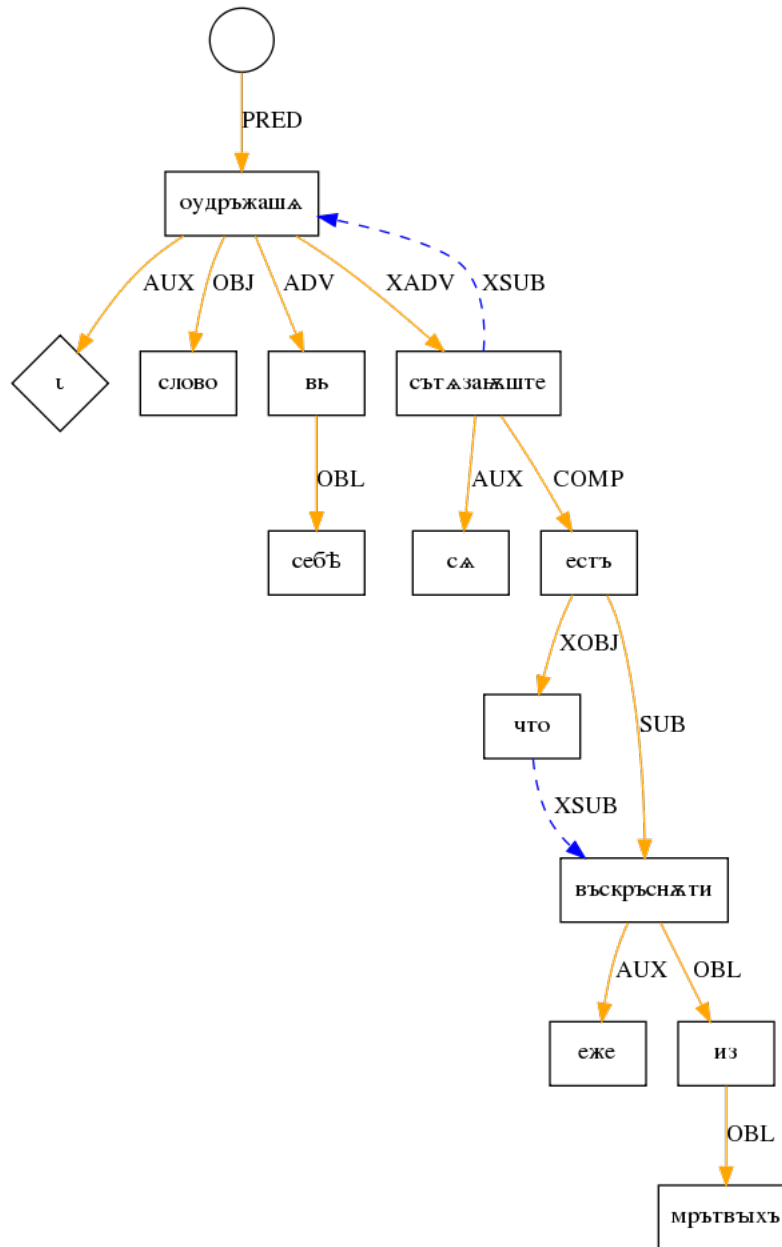


Figure 101: ‘They kept the matter to themselves, discussing what “rising from the dead” meant.’ (Mar. Mark 9.10, 36818)

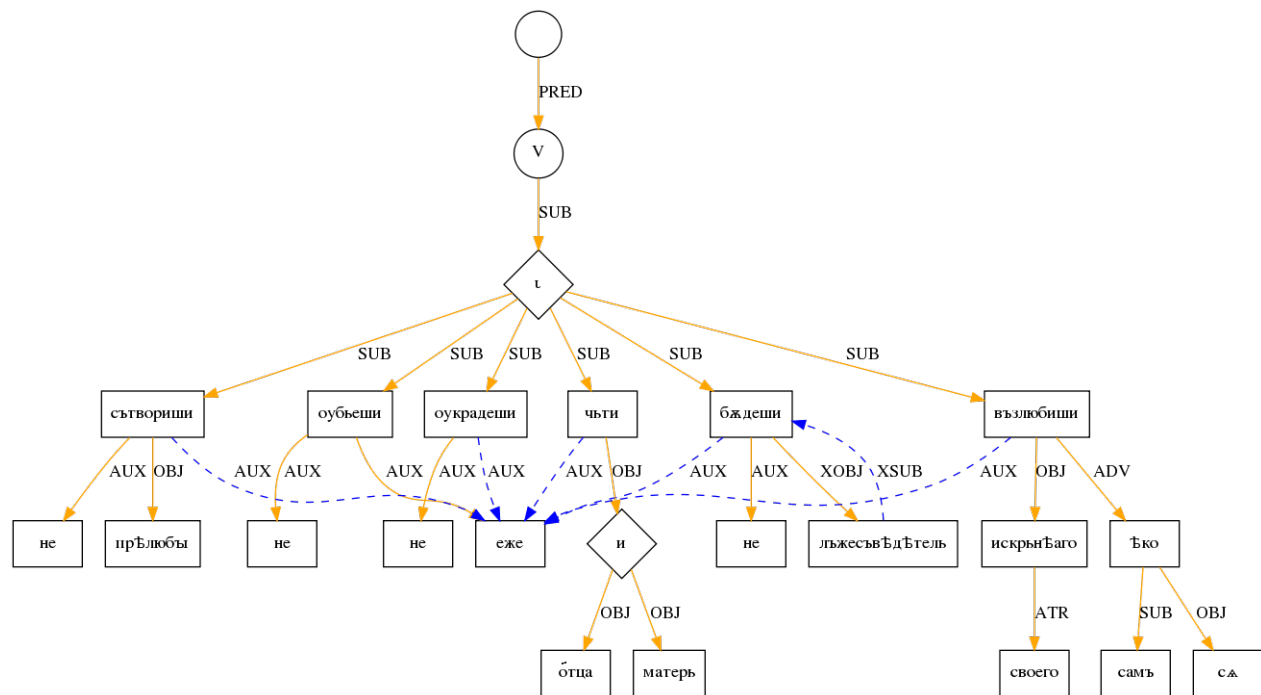


Figure 102: ‘You shall not murder, you shall not commit adultery, you shall not steal, you shall not bear false witness, honor your father and mother, and, you shall love your neighbor as yourself.’ (Mar. Matt. 19.18–19, 39165)

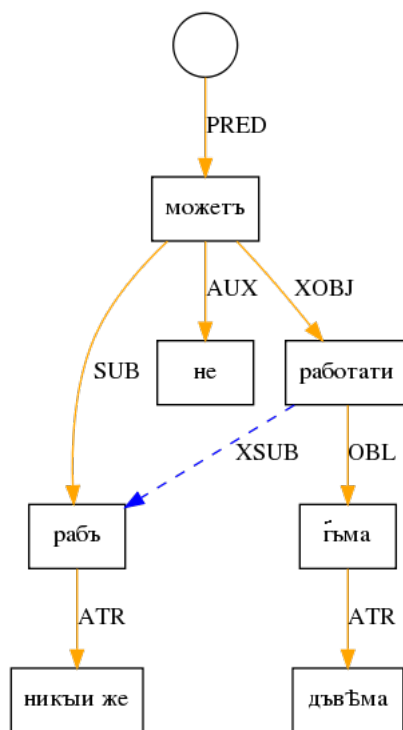


Figure 103: ‘No servant can serve two masters.’ (Mar. Luke 16.13, 40988)

18 Slash notation

18.1 Introduction

Slashes are used in our system to represent coreference relations within one dependency tree (which can of course contain several coordinated main clauses). Informally, the general principle is that a whenever an item ‘needs’ to have a dependent, but this dependent is already ‘used’ (ie. is a dependent of something else), we indicate this relationship by a slash arrow.

In other words, the meaning of a slash is *not* that the slasher node¹² is referentially identical to the slashee node.¹³ Rather, the slash arrows should be conceived of as supplementary dependency relations: the slashee node is a dependent of the slasher. There is one exception to this principle, namely when the slasher is an empty node which bears the PRED-relation, see section 18.3.4.

Interpretation

To understand this general principle, consider the sentence *nikyi že rabъ ne možetъ dѣvѣma gѣma rabotati* ‘no servant can serve two masters’. The depend-

ency tree is seen in (103).

The node *rabъ* is correctly designated as the subject of *možetъ*. However, *rabъ* is also the subject of *rabotati*, and to represent this fact, we put a slash arrow from *rabotati* to *rabъ*. The interpretation of this notation is that the subject of *rabotati* is coreferent with the subject of *možetъ*, which is *rabъ*.

We can distinguish two groups of uses of the slash notation, one associated with the ‘open’ functions XADV and XOBJ, and one associated with gapping and elliptical coordination.

18.2 Open functions: XADV and XOBJ

The open functions XADV and XOBJ were introduced in section 14.8. By definition, they do not supply their own subject, but get a subject via coreference relations within the sentence. From the perspective of traditional grammar, it is perhaps not clear that nominal predicatives behave in the same manner, but consider the example *brěmę moe legъko estъ* ‘my burden is light’, as found in (104). *brěmę* is the subject of *estъ*, and *legъko* is the predicative complement of *estъ*. However, we want to capture the fact that not only is *legъko* an argument of *estъ*, in the sense that *estъ* requires the presence of a predicative complement, *legъko* is also a predication about *brěmę*. In other words, *brěmę* is the subject of *legъko*, and this is represented by a slash arrow from *legъko* to *brěmę*.

The same holds for a traditional ‘object predicative’, as in *i sъtvorjъ vy љkotъ lovъca* ‘and I will make you fishers of men’ (105): In this construction, *sъtvoriti* requires three arguments, a subject, an object and a predicative complement. So, *lovъca* is an XOBJ-dependent of *sъtvorjъ*, and *vy* is an OBJ-dependent of *sъtvorjъ*, but at the same time, *vy* is the subject of *lovъca* and we capture this fact by a slash arrow from *lovъca* to *vy*.

With the relation XADV, the facts are the same. Consider *i otъvѣštavъ isъ reče imъ* ‘and Jesus answered and said to them’ (106): *reče* is the main verb, *isъ* its subject and *imъ* its oblique argument. *otъvѣštavъ* is an adverbial which modifies the main verb, contributing the extra information that the ‘saying’ was ‘answering’. So *otъvѣštavъ* is an XADV-daughter of *reče*, but we also want to express that *isъ* is the subject of *otъvѣštavъ*.

Now what if there was no overt subject in the sentence, ie. if the subject of *reče* was pro-dropped and the sentence read *i otъvѣštavъ reče imъ*? In this case, the subject is supplied by the finite verb itself, and the slash arrow is therefore directed towards the verb (107).

This means that **every node which is the daughter in an XOBJ or XADV relation should have a slash!** The interpretation of the slash arrow is always uniform, since it is always a *subject* dependent that these nodes lack. In other words, the slashes are always the subject of the slasher. Therefore there should **never be more than one** outgoing slash from the daughter in an XOBJ or XADV relation, since this will create an ambiguity. In other words, it is not

Number of
outgoing
slashes
on
XOBJS/XADVs

¹²Ie. the node where the slash is inserted; the origin of the slash arrow in the visual representation of the dependency tree.

¹³Ie. the goal of the slash arrow in the visual representation of the dependency tree.

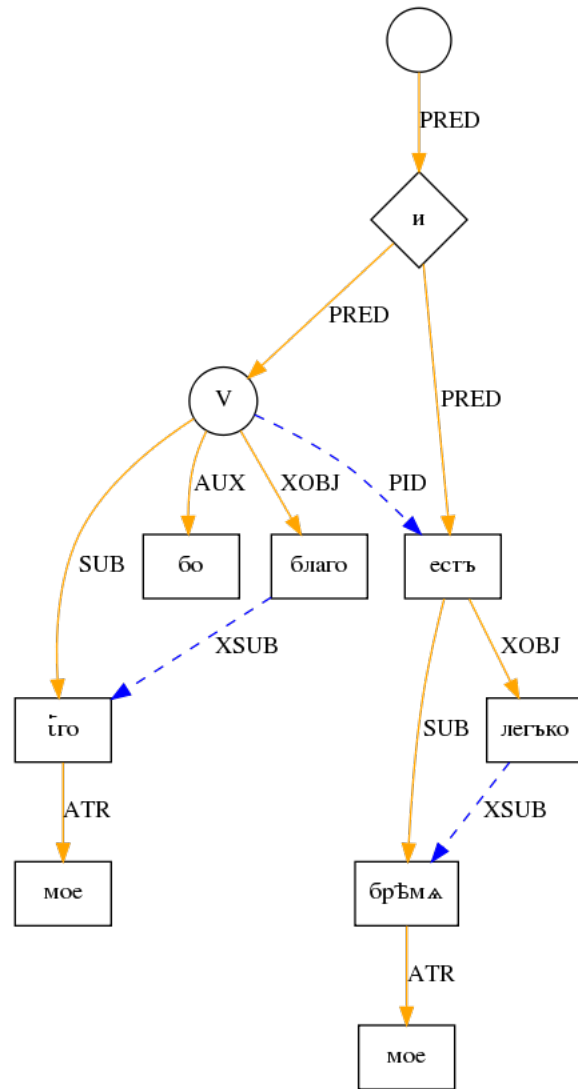


Figure 104: ‘For my yoke is easy, and my burden is light.’ (Mar. Matt. 11.30, 38695)

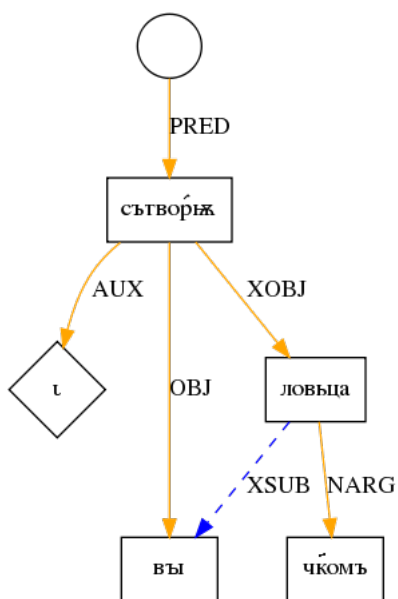


Figure 105: ‘And I will make you fishers of men.’ (Mar. Matt. 4.19, 75193)

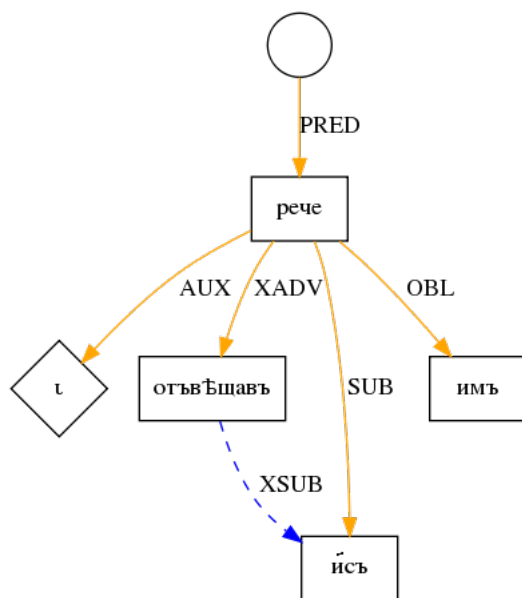


Figure 106: ‘And Jesus answered and said:’ (Mar. Mark 12.17, 37055)

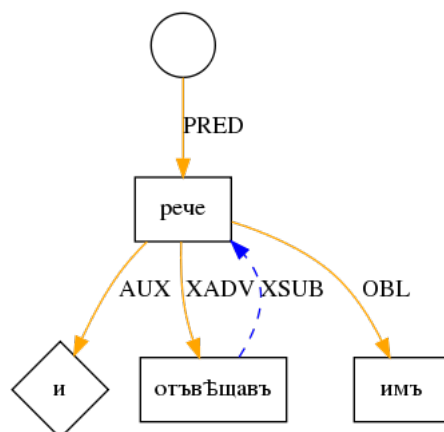


Figure 107: ‘And he answered and said:’ (Mar. Luke 19.40, 41199)

possible to mark that such an element shares an argument with another element of the sentence. Since this does occasionally happen, some information is lost due to this policy.

Notice finally, that the slash arrow from an XOBJ or XADV node should always be local, ie. directed towards either the head verb or an element dominated by the head verb. This restriction is enforced by the annotation interface. However, the annotation interface does allow slashing into a dependent clause under the head verb. This should *only* be done when the argument in question is a shared argument of the main verb.

18.3 Slashes in gapping and coordination

18.3.1 Motivation

Slashes in coordinate *sentences* denote much of the same scoping information that is expressed in other ways in constituent coordination. In section 15.1 we represented adjectives scoping over two coordinated nouns as in (108).

Theoretically, it would be possible to represent *онъ же млѣчааше. и ниѣсоже не отвѣштааше* ‘but he remained silent and made no answer’ by coordinating the two verbs under the conjunction and then attach *онъ* as a SUB-daughter of the conjunction. But consider what would happen if we had a coordination of two SUB relative clauses. This is what we see in (109).

Now, if we were to place their common SUB, the relative pronoun *иже*, on the conjunction *и*, this conjunction would dominate *three* SUB daughters. Although this ambiguity could be resolved through the morphological annotation,

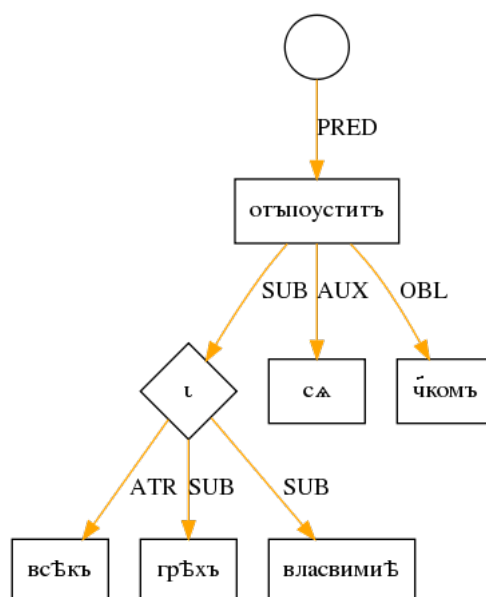


Figure 108: ‘Every sin and blasphemy will be forgiven people.’ (Zogr. Matt. 12.31, 173756)

we prefer to avoid it entirely and represent the sentence as shown in (109) and in (94) on p. 104.

18.3.2 Slashes in sentence coordination

Whenever two coordinated sentences, or a dependent clause or XADV, share an element, be it a SUB, OBJ, XOBJ, OBL, AG or ADV, and this element is overtly represented only once, the element should be attached to the verb to which it belongs in the surface string, and a slash arrow should be set from the other verb to the shared element. Notice that a slash arrow from an XADV to an argument of the verb can only indicate the subject of the XADV, therefore XADVs can only have one outgoing slash arrow. This kind of slash should leave a verbal node and point towards the shared element. One should never use slashes between two verbs to indicate subject identity. For examples, see (98) on p. 109 and (100) on p. 112.

The shared element can also be an auxiliary verb, as seen in (110), which also has a shared negation.

When the slash notation is used in sentence coordination, it is important to assure that the grammatical relation between the slasher node and the slashee

Shared arguments

Shared auxiliaries

Interpretation of slashes in coordination

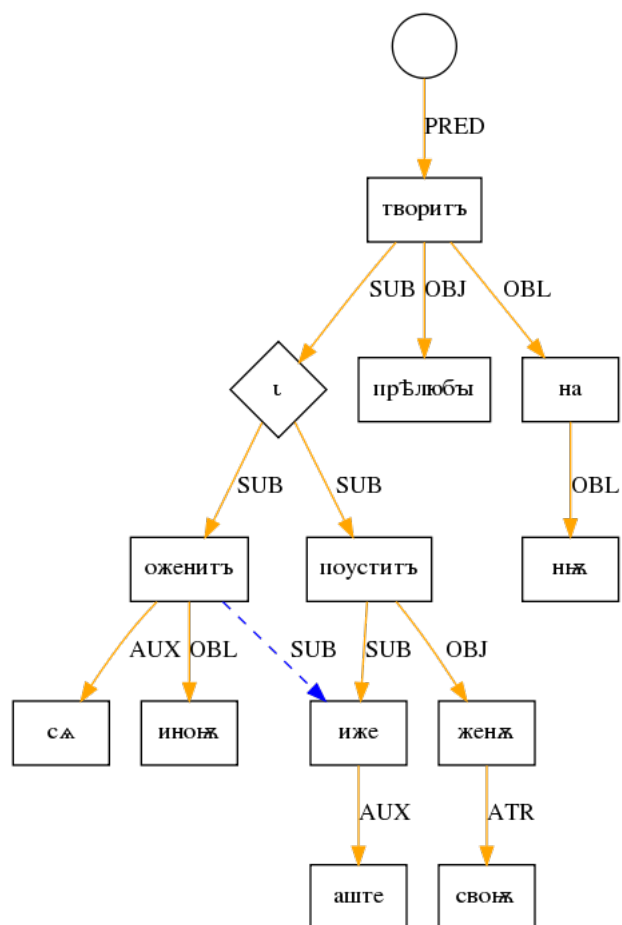


Figure 109: ‘Whoever divorces his wife and marries another commits adultery against her’ (Mar. Mark 10.11, 36897)

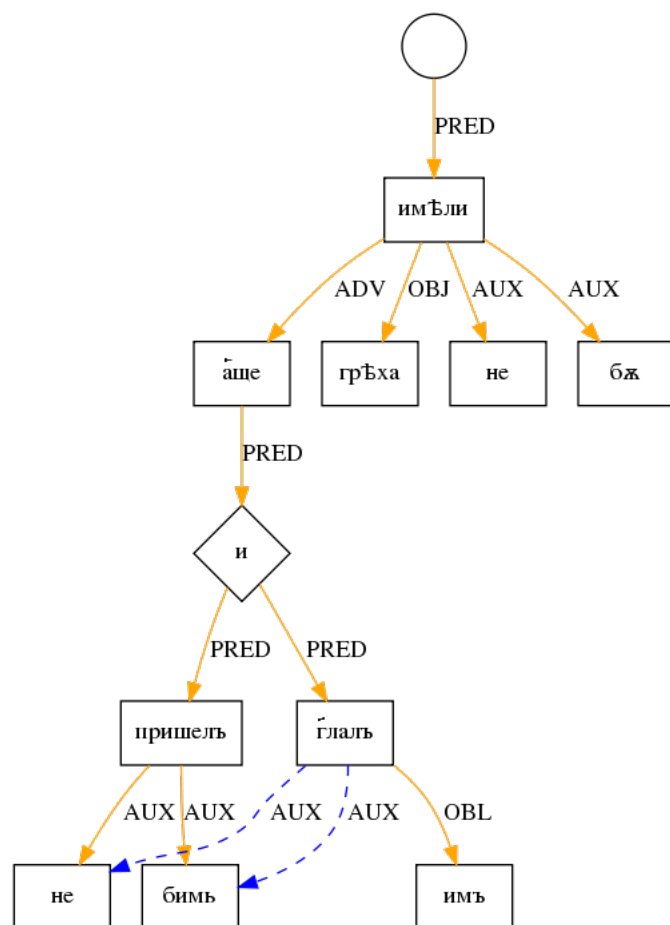


Figure 110: ‘If I had not come and spoken to them, they would not have been guilty of sin.’ (Mar. John 15.22, 42696)

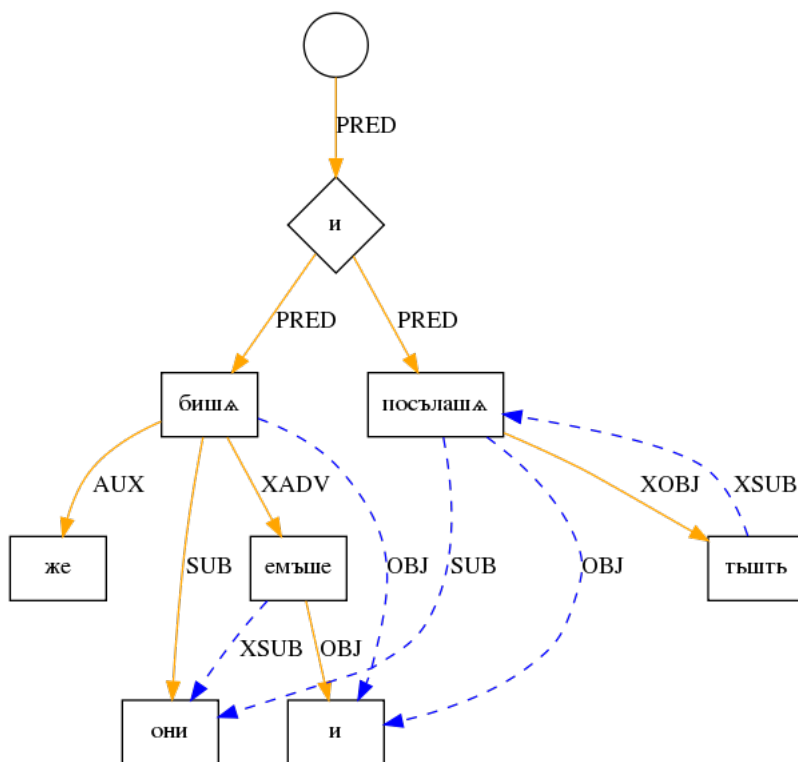


Figure 111: ‘And they took him and beat him and sent him away empty-handed.’
(Mar. Mark 12.3, 37025)

is identical to the relation of the slashes to its mother. In other words, if the slashes are subjects in their own clause, it should also be the subject of the slashes. If the unexpressed object of a coordinated verb is identical to the overt subject of another verb, we cannot express this coreference relation. This is so because, unlike the case of slash notation on open predications, where all slashes are subjects, there is no default relation type with which to interpret the slash relation. Instead, the slash relation ‘inherits’ the function of the slashes. As long as this restriction is respected, there can be several slashes in a tree, as seen in (111).

Here, *i* is taken to be the overt object of *емъше* (for the general principle, see section 18.3.5). The arrows from *бише* and *послаше* indicate that these verbs also have *i* as their object. The arrows from *емъше* and *тыштъ* to their respective dominating verbs *бише* and *послаше* indicate that the subjects of these XADVs

is an unexpressed element which is ‘supplied by the verb’. In the case of *емѣше*, the subject of the participle is the unexpressed subject of the main verb; and in the case of *ѣшѣ*, the subject of the adjective is the object of the main verb. This object is expressed, but via a slash arrow: it is therefore not possible to slash directly from *ѣшѣ* to its subject.

18.3.3 When is an adverbial shared?

While it is generally easy to decide when two coordinated verbs share their subject or another argument (object, oblique), it can be more difficult to decide whether an adverbial actually has scope over both conjuncts. Some are easy: for example, an adverbial conditional clause is shared between two main clauses if they both belong to the apodosis (if X, then Y and Z). Generally, time and space adverbials are the most tricky ones: note in particular that time adverbials should only be shared by two main clauses, if their events are cotemporal and not consecutive.

18.3.4 Slashes in gapping/multi-rooted conjuncts

Slashes in multi-rooted conjuncts behave exactly like slashes in sentence coordination, except that there is also a slash arrow from the empty node representing the verb, to the overt verb.¹⁴

Consider again the example *прѣдасть же братъ братъ на съмъртъ. ѿ оцѣ чѣдо* (98) on page 109. As in this example, it is often the case that other elements than the predicate are shared between the conjuncts. The slash arrow from the empty node to *прѣдасть* indicates predicate identity, whereas the arrow from the empty node to *на* indicates a shared argument. As such it is subject to the same restriction as other slashes in sentence coordination: the slashee must have the same grammatical function in both conjuncts.

18.3.5 In which conjunct does a shared element belong?

Our treatment of shared arguments raises the question where a shared argument belongs in the tree. Sometimes an element can belong in several places. Why do we put *на съмъртъ* in the first conjunct in example 98? Or, in a more difficult case, why does *ѿ* belong to *емѣше* in example 111? The following section provides some guidelines for making such decisions.

For argument sharing between coordinated elements we need to establish the domains of both conjuncts. If there is an overt conjunction, this will indicate the boundary between the two domains. Elements should therefore be made to depend on the verb in their domain. In *прѣдасть же братъ братъ на съмъртъ. ѿ оцѣ чѣдо* we have the domains *прѣдасть же братъ братъ на съмъртъ* and *отѣ чѣдо*, so *на съмъртъ* clearly depends on *прѣдасть* and not on the empty node. If there is no overt conjunction, we appeal to the principle of domain continuity:

**Delimitation
of conjunct
domains**

¹⁴This slash arrow therefore has a different interpretation, meaning that the slasher node has the same meaning as the slashee node.

both domains should, if possible, be continuous. If there are several possibilities, ambiguous elements should belong to the first conjunct. For example, if there were no *i* in example 98, both *prědastъ že bratrъ bratra na sьmьrtъ. // otecъ čędo* and *prědastъ že bratrъ bratra // na sьmьrtъ. otecъ čędo* would yield continuous domains, so we would choose the first option, which attaches *na sьmьrtъ* to the first conjunct.

In practice, then, the principles mean that the first element which belongs only to the second conjunct marks the start of the second conjunct domain.

When there is argument sharing between a clause and a dependent clause or an XADV, the principles remain the same, but the application is slightly different. The subtree dominated by the subordinate verb or the subjunction should correspond to a continuous string in the sentence. In example 111, this means that *i* can be attached to *emъše*; if *i* had appeared to the right of *biše*, such an attachment would not be possible. When several attachments are possible, we again choose to attach elements as early as possible in the surface string: this means that we prefer to attach *i* to *emъše*. However, if the subject of *biše* had intervened between *emъše* and *i*, such an attachment would *not* be possible, since *emъše i* would no longer be a continuous domain.

Notice, finally, that the principles described in this section are only applicable whenever the sentence is ambiguous. Very often, case morphology will make it entirely clear where an element belongs, and in such cases, the morphology should not be overridden by considerations of domain continuity etc.

19 Other constructions

19.1 ‘all’, ‘self’ and similar words

Words like ‘all’, ‘self’ and a few others are special: they are not really pronouns with anaphoric reference, but they behave very similarly when they determine non-overt (‘pro-dropped’) subjects.

19.2 Headedness of compound verb forms

19.3 Comparison

19.4 Verbs of preference

19.5 ‘Approximately’

19.6 Particle coordination

19.7 Scope

Scope is relevant to many constructions. An adjective can modify one or more nouns, and is then said to have scope over these nouns. A genitive can be dependent upon (and thereby for example modify) one or more nouns. The negation can negate one particular element in a sentence, or several, or the

whole sentence - these elements are said to be in the scope of the negation. Emphasizing particles can take scope of one or more words.

We represent scope via the attachment site of scoping items. Scoping items should be placed so as to be dependent on the item in their scope. By default, they also take scope over all elements dominated by their mother node; so a negation dependent on the verb negates the whole sentence. This means that if we want an item to take scope over two coordinated elements, it should be made to depend on the conjunction. However, since the status of the Greek definite article regarding whether it can take scope over two coordinated proper nouns is completely unclear, we currently make it a dependent of the following noun only - in other words, the article is never attached to a conjunction.

If e.g. a negation has scope over only one of two coordinated elements, it must be attached to that element. Note that this holds even in cases of double negation, i.e. a negation may be an AUX on a negative pronoun.

19.8 Impersonal verbs

Impersonal verbs are defined by their inability to take a subject.

19.9 esse + participle

In the New Testament, we fairly often find a construction where the copula combines with a participle to form something which could almost be described as a periphrastic form. However, despite the regularity of the construction, we have chosen to treat it as if the copula had its full force.

19.10 Gerunds

19.11 Infinitives

The two most common functions of dependent infinitives are head of accusative with infinitive/dative (which is a COMP) and complement of auxiliary verb (XOBJ). But infinitives can have other functions. For example, a bare infinitive can express purpose, in which case it is an XADV. It can also be an argument of a noun, in which case it is a NARG.

19.12 Reflexives

Note that the reflexive marker with reflexive verbs is analysed as AUX.

19.13 Datives

19.14 Compound subjunctions

19.15 Appositive names

19.16 Copulae, copula-like verbs and location verbs: *byti* etc.

19.17 Vocatives

It is easy to think that all morphological vocatives will have the relation VOC, but annotators should note that there are other possibilities. First, and most obvious, vocative nouns and adjectives can be ATR or APOS under another vocative.

But there are other possibilities.

19.18 And it came to pass...

19.19 Copula, verb of position, verb of existence

In the old Indo-European languages, the copula ('John is a doctor') is often also a verb of position ('John is in London') and a verb of existence ('There are doctors in London'). We do not attempt to keep the first two apart: both complements ('a doctor') and ('in London') are considered XOBJs. But the existence verb is treated differently: here there is in fact no complement, only an adverbial.

19.20 Resumptive pronouns and topicalizations

19.21 AUX on pronouns – *iže ašte*

19.22 Adnominal infinitives

In several cases, infinitives modify nominal forms like pronouns, adjectives and nouns, as in 'easy to read', 'something to eat' etc. We treat such infinitives as *modifiers* since it tells us *in what respect* something is easy and *what kind* of something we are talking about.

Dependents on pronouns and nouns therefore often become ATRs.

19.23 Impersonal temporal expressions

19.24 It happened to him...

19.25 ‘if not’

19.26 COMP and OBJ in the same sentence

Sometimes a verb seems to govern both a COMP and an OBJ. There are two different cases to distinguish:

1. COMP and OBJ are coordinated
2. the OBJ is co-referent with the subject of the COMP

In the first case, the COMP and the OBJ should simply be coordinated and the coordinating node should be given the OBJ relation.

19.27 Questions with alternatives

Sometimes a question is asked with alternatives to choose between, as in “Who is the better one, me or my brother?”. In these cases, the alternatives are taken as appositions on the question-word.

20 General issues

20.1 ADV or OBL? adjunct or argument?

It can often be hard to decide whether a given element should be an ADV or an OBL, or – in more traditional terms – whether an element is an adjunct or an argument. Therefore we provide a super-tag which should be used whenever the annotator cannot decide, but the following section offers some advice to guide the choice.

Arguments are traditionally defined as ‘elements seen as required by the verb’. In old Indo-European languages, nothing is really ‘required by the verb’, because all elements, even subjects and objects, can be dropped if they are easily inferable from the context. The test is therefore not whether an element can be left out or not, but rather whether it is possible to conceptualize the event expressed by the verb while abstracting from some element in the sentence. If that is possible, the element is an adjunct. Adjuncts (ADV) are elements which elaborate upon an event description and gives extra information about the event.

There are some kinds of adverbials which are almost always adjuncts. These are

- Adverbials of manner
- Adverbials of instrument
- Adverbials of time

- Adverbials of place
- Adverbials of purpose

20.2 ATR or APOS? attribute or apposition?

...

20.3 ATR or XADV?

Adjectives and participles (and other adjective-like categories) can be attributive or predicative/conjunct. In the first case it helps determine the reference of the noun phrase; in the second case, it is a predication which expresses an event or a state which 'accompanies' the main event and which takes the noun as its subject.

20.4 XADV or SUB?

20.5 More on the XOBJ relation

The XADV and XOBJ relations are unfamiliar from traditional school grammar. XADV corresponds quite closely to the concept of predicative/conjunct participle or adjective, so it should not be too hard to grasp. The XOBJ relation, on the other hand, does not correspond closely to one particular category of traditional grammar, so it needs further explanation.

The crucial facts to grasp about the XOBJ relation is that it is used for predications which have external subjects (whence the X-) and which are governed by another verb. This means that nothing which can itself have a subject daughter, is an XOBJ: all finite verbs are excluded.

The prototypical XOBJ is therefore an infinite verbal form, but other word classes can also be predicative: most notably nouns, adjectives and prepositions.

XOBJ is a governed relation, an element which is selected and demanded by the matrix verb.

20.6 Empty nodes

In our model, we only use empty nodes to stand in for missing conjunctions in asyndetic coordination and missing verbs in elliptical constructions. Empty nodes should never be used for any other purpose, ie. to mark a 'null head noun' on which an adjective is dependent.

21 Error messages during syntactic annotation

21.1 'Must have or inherit one outgoing slash edge'

This error message means that a token which is XADV or XOBJ lacks a slash arrow.

Open predications (XADV and XOBJ, see section 14.8) do not have an ‘internal’ subject dominated by the predicate. Instead, the subject is present somewhere else in the same sentence and should be linked to via a slash arrow. If the subject is not overtly present in the sentence, it must nevertheless be present in the argument structure of the governing verb and to represent this we let the slash arrow point towards the verb.

Notice that two XADVs or XOBJs under the same conjunction should not have one slash arrow each: instead the slash should be put on the conjunction. It will be inherited by the daughter nodes.

If you can’t find a suitable target for the slash arrow, the use of XADV or XOBJ is probably wrong.

21.2 ‘May not be a daughter of the root node’

Only tokens bearing the relations PRED, PARPRED or VOC are allowed as daughters of the sentence root. If you let any other element be directly dominated by the root, you will get this error message.

21.3 ‘Subgraphs overlap’

If there are several PRED daughters directly under the root, they must be strictly ordered linearly. If there are two PRED tokens under the root, all daughters of the first one must precede all daughters of the second one. If there are three PRED tokens, all daughters of the first one must precede all daughters of the next two; and all daughters of the second one must precede all daughters of the third one.

The reason for this is that if one sentence root has several PRED daughters, the tree will eventually be ‘sawed up’ in several trees each having only one PRED daughter. To assure that this will be possible, we need to secure that the trees are linearly ordered. If one sentence is contained in another, you should use the PARPRED relation instead.

21.4 ‘Slashes are not contained by subgraph’

If there are two PRED daughters directly under the root, it is not allowed to have a slash arrow between the a daughter of the one of the PREDs and the other. In other words, the slasher and the slashee should both be daughters of the same PRED.

This error message is most likely to occur in the analysis of gapping in asyndetic coordination.

21.5 ‘The head of a PARPRED/VOC relation must be the root node or a valid coordination’

Elements with the relation VOC or PARPRED should always be made daughters of the sentence root.

21.6 ‘Slash must target the node’s head or a node dominated by the head’

The subject of an open predication (XADV or XOBJ, see section 14.8) should always be found in the same sentence. In case the subject of the open predication is not overtly realized in the same sentence, but is found in a conjoined sentence, the slash arrow should point towards the dominating verb, which in turn will have a ‘shared argument’ slash arrow to the overt realization of the subject of the XADV/XOBJ.

21.7 ‘The head of a PRED relation must be the root node, a subjunction or a valid coordination’

There are restrictions on when to use the PRED relation: in fact it is only used under the sentence root or under a subjunction. There are two common situations where annotators get this message.

First, relative sentences and dependent interrogative clauses should *not* be headed by the relative/interrogative word, with the verb as a PRED daughter. Follow the rules in sections 14.4 and 14.5 instead.

Second, many subjunctions are not marked as such in the morphology. Check the morphological annotation and make sure that the subjunction is not annotated as a conjunction, non-comparable adverb or something else.

21.8 ‘A subjunction may only be the dependent in a COMP, ADV or APOS relation’

Subjunctions should always be COMP (if selected for by the verb and necessary for the sentence), ADV (if introducing an adjunct clause) or APOS (if dependent on a noun).

If you don’t understand why you get this message, please go to the morphological annotation and check whether it is caused by a token wrongly marked as a subjunction.

21.9 ‘An infinitive may not be the dependent in an ADV relation’

Infinitives can certainly be adverbials (for example final infinitives), but they always imply a subject and so should be XADVs instead.