Guidelines for the Syntactic Annotation of Latin Treebanks (v. 1.3)

David Bamman¹, Marco Passarotti², Gregory Crane¹, and Savina Raynaud²

¹Latin Dependency Treebank, The Perseus Project, Tufts University
²Index Thomisticus, Catholic University of the Sacred Heart, Milan

March 1, 2007

Contents

1 Introduction 3
2 Dependency Grammar 3
3 Annotation Style 4
  3.1 PRED (predicate) 4
  3.2 SBJ (subject) 6
    3.2.1 Nominative nouns 6
    3.2.2 Accusative nouns 6
    3.2.3 Ablative nouns 6
    3.2.4 Infinitive verbs 7
    3.2.5 Subordinate clauses 8
    3.2.6 Relative clauses 8
    3.2.7 Pronouns 9
  3.3 OBJ (object) 9
    3.3.1 Accusative nouns 9
    3.3.2 Accusative + infinitive constructions 10
    3.3.3 Relative clauses 10
    3.3.4 Subordinate clauses 11
    3.3.5 Gerunds 11
    3.3.6 Infinitive verbs 11
    3.3.7 Gerundives 13
    3.3.8 Indirect objects 13
    3.3.9 Passive agents 14
  3.4 ATR (attribute) 14
    3.4.1 Adjectives 15
1 Introduction

Treebanks - large collections of syntactically parsed sentences - have recently emerged as a valuable resource not only for computational tasks such as grammar induction and automatic parsing, but for traditional linguistic and philological pursuits as well. This trend has been encouraged by the creation of several historical treebanks, such as that for Middle English (Kroch and Taylor [5]), Early Modern English (Kroch et al. [4]), Old English (Taylor et al. [9]), Early New High German (Demiske et al. [1]) and Medieval Portuguese (Rocio et al. [7]).

The Perseus Project and the Index Thomisticus are currently in the process of developing treebanks for Latin (Perseus on works from the Classical period and IT on the works of Thomas Aquinas). In order for our separate endeavors to be most useful for the community, we must come to an agreement on a common standard for the syntactic annotation of Latin - and not only for the two projects under development now, but for any other Latin treebanks focussing on different eras that may arise in the future.

In what follows we present a preliminary set of annotation guidelines. The annotation style proposed here is predominantly informed by two sources: the dependency grammar used by the Prague Dependency Treebank [3, 2] (itself based on Sgall et al. [8]), and the Latin grammar of Pinkster [6].

2 Dependency Grammar

Dependency Grammar (DG) differs from constituent-based grammars by foregoing non-terminal phrasal categories and instead linking words themselves to their immediate head. This is an especially appropriate manner of representation for languages with a moderately free word order (such as Latin and Czech), where the linear order of constituents is broken up with elements of other constituents. A DG representation of *ista meam norit gloria canitiem* canitiem, for instance, would look like Figure 1.

![Diagram](image)

Figure 1: Dependency graph of *ista meam norit gloria canitiem* (Prop. I.8.46). Arcs are directed from heads to their dependents.

Dependency grammar is also appropriate for Latin since it is not too theoretically distant from Classical pedagogical grammars, where the highly inflected nature of the language leads to discussions of, for example, which adjective "modifies" which noun in a sentence. A dependency grammar simply assigns one such "modification" to every word.
3 Annotation Style

Different treebanks and grammars, however, assign syntactic functions differently. The general model for our style of representation is that used by the Prague Dependency Treebank, with several important departures arising from Pinkster’s [6] Latin grammar. The following table lists all of the tags currently in use; the following subsections further elaborate each.

<table>
<thead>
<tr>
<th>Tag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRED</td>
<td>predicate</td>
</tr>
<tr>
<td>SBJ</td>
<td>subject</td>
</tr>
<tr>
<td>OBJ</td>
<td>object</td>
</tr>
<tr>
<td>ATR</td>
<td>attributive</td>
</tr>
<tr>
<td>ADV</td>
<td>adverbial</td>
</tr>
<tr>
<td>ATV/AtvV</td>
<td>complement</td>
</tr>
<tr>
<td>PNOM</td>
<td>predicate nominal</td>
</tr>
<tr>
<td>OCOMP</td>
<td>object complement</td>
</tr>
<tr>
<td>COORD</td>
<td>coordinator</td>
</tr>
<tr>
<td>APOS</td>
<td>apposing element</td>
</tr>
<tr>
<td>AuxP</td>
<td>preposition</td>
</tr>
<tr>
<td>AuxC</td>
<td>conjunction</td>
</tr>
<tr>
<td>AuxR</td>
<td>reflexive passive</td>
</tr>
<tr>
<td>AuxV</td>
<td>auxiliary verb</td>
</tr>
<tr>
<td>AuxX</td>
<td>commas</td>
</tr>
<tr>
<td>AuxG</td>
<td>bracketing punctuation</td>
</tr>
<tr>
<td>AuxK</td>
<td>terminal punctuation</td>
</tr>
<tr>
<td>AuxY</td>
<td>sentence adverbials</td>
</tr>
<tr>
<td>AuxZ</td>
<td>emphasizing particles</td>
</tr>
<tr>
<td>ExD</td>
<td>ellipsis</td>
</tr>
</tbody>
</table>

Table 1: Complete Latin tagset.

3.1 PRED (predicate)

Every complete sentence (i.e., non-elliptical with at least one predicate) has one word unattached to any other; this is attached to the root of the sentence with the relation PRED.
spectat inter occasum solis et septentriones

If a sentence begins with an initial conjunction (either coordinating or subordinating), the main verb is dependent on that conjunction and the conjunction then depends on the root.

at illum ingens cura atque laetitia simul occupavere

et stetit super harenam maris
3.2 SBJ (subject)

Subjects are dependent on their verb (which is the predicate of either a main or subordinate clause), and come in a variety of parts of speech and phrases, including:

3.2.1 Nominative nouns

Belgae ab extremis Galliae finibus oriuntur

3.2.2 Accusative nouns

These are typically found in indirect discourse and other accusative + infinitive constructions.1

contentum te esse dicebas

3.2.3 Ablative nouns

Since ablative absolutes are treated as an embedded predication, the ablative noun in such constructions should be annotated as the subject of the participle.2

---

1 For more information on this construction, see section 4.7, "Accusative + Infinitive."
2 See section 4.4, "The Ablative Absolute."
His rebus cognitis Caesar Gallorum animos verbis confirmavit
This is also the case with *ab urbe condita* constructions.

3.2.4 Infinitive verbs

dulce et decorum est pro patria mori
3.2.5 Subordinate clauses

datum est illis ne occiderent eos

3.2.6 Relative clauses

qui habet aurem audiat
3.2.7 Pronouns

ea autem quae dicuntur per modum formae inhaerentis, non recipiunt dictorum pronominum additionem

3.3 OBJ (object)

Likewise, objects are also dependent on their verb, and come in as large a variety of phrase types as subjects, including:

3.3.1 Accusative nouns

Caesar Gallorum animos verbis confirmavit
3.3.2 Accusative + infinitive constructions

The infinitive verb is the head of the accusative + infinitive construction, and depends on the verb introducing the construction via OBJ.

contentum te esse dicebas

3.3.3 Relative clauses

quod vides scribe
3.3.4 Subordinate clauses

dicit
PRED

quod
AuxC

significat
OBJ

usia
SBJ
compositum
OBJ

ex
AuxP

et
COORD

materia
ATR_CO
forma
ATR_CO

dicit ... quod usia significat compositum ex materia et forma

3.3.5 Gerunds

disponitur

... 

ad
AuxP

receptandum
OBJ

formam
OBJ

disponitur ad receptandum formam

3.3.6 Infinitive verbs

This include both verbs that function as traditional direct objects (as in dabo ei edere de ligno vitae) as well as those that complete verbs like possum, volo or incipio.
dabo ei edere de ligno vitae

post haec vidi turbam magnam quam dinumerare nemo poterat
3.3.7 Gerundives

Our notion of object, however, follows that used by the PDT, and includes a wider range of phrases than traditional direct objects. OBJ should also be used to annotate the complements of a verb (i.e., those required arguments that cannot become subjects if the verb is made passive). Examples (all taken from Pinkster [6]) include:

- miseretur pater filii (filii -> miseretur)
- nupsit Sempronia Scipioni (Scipioni -> nupsit)
- pater gladio utitur (gladio -> utitur)
- pater litteris studet (litteris -> studet)
- abundat Germania fluminibus (fluminibus -> abundat)
- contendunt Romani cum Germanis (OBJ, Germanis -> cum; AuxP, cum -> contendunt)
- accusavit pater me avaritiae (avaritiae -> accusavit)
- docet magister me linguam Latinam (linguam -> docet)

3.3.8 Indirect objects

Traditional indirect objects are also included in this category, and can show up either as prepositional phrases or as dative nouns:
3.3.9 Passive agents

And as in the PDT, actors in passive constructions should also be annotated as OBJ.

\[
\text{mittuntur}
\]

\[
\text{litterae}
\]

\[
\text{Caesarem}
\]

\[
\text{Cicerone}
\]

\[
\text{ex urbe}
\]

\[
\text{periculis}
\]

3.4 ATR (attribute)

Attributes are those phrases that attributively specify (or delimit) the meaning of their head. Most commonly these are adjectives, but can include other classes as well, such as nouns, relative clauses and prepositional phrases.
3.4.1 Adjectives

Belgae ab extremis Galliae finibus oriuntur

3.4.2 Participles

et post haec vidi alium angelum descendentem de caelo
3.4.3 Prepositional phrases

et
COORD

/\

vocem

vocem

/\

magnam

magnam

/\

in

in

/\

magnum

magnum

/\

in

in

/\

caelo

caelo

et audivi vocem magnam in caelo

3.4.4 Agreeing nouns

/\

Rhodano

Rhodano

/\

a

a

/\

flumine

flumine

... a flumine Rhodano

3.4.5 Non-agreeing nouns

Possessive genitives, objective genitives and partitives all fall in this category.
Belgae ab extremis Galliae finibus oriuntur

faciam initium narrandi

Ex eo numero compluris ... in vincula coniecerat
3.4.6 Relative clauses

mercatores ... ea quae ad effeminandos animos pertinent important

3.4.7 Relative pronouns

quam ob rem discede atque hunc mihi timorem eripe
3.4.8 Pronouns

\[
\text{PRED} \quad \text{continebat} \\
\text{in} \quad \text{AuxP} \quad \text{occulto} \\
\text{hostes} \quad \text{SBJ} \quad \text{intra} \quad \text{AuxP} \\
\text{silvas} \quad \text{OBJ} \quad \text{eas} \quad \text{ATR} \\
\]

intra eas silvas hostes in occulto sese continebat

3.5 ADV (adverbial)

Similarly, adverbials further specify the circumstances under which a verb, adjective or adverb takes place. These include adverbs, prepositional phrases, nouns in oblique cases, participles and subordinate clauses.

3.5.1 Adverbs

\[
\text{PRED} \quad \text{mittuntur} \\
\text{litterae} \quad \text{SBJ} \quad \text{ad} \quad \text{AuxP} \\
\text{Caesarem} \quad \text{OBJ} \quad \text{confestim} \quad \text{ADV} \\
\text{ab} \quad \text{AuxP} \\
\text{Cicerone} \quad \text{OBJ} \\
\]

mittuntur ad Caesarem confestim ab Cicerone litterae

3.5.2 Prepositional phrases

\[
\text{PRED} \quad \text{veniunt} \\
\text{ceteri} \quad \text{SBJ} \quad \text{sine} \quad \text{AuxP} \\
\text{mora} \quad \text{ADV} \\
\]

ceteri sine mora veniunt
3.5.3 Nouns

Caesar Gallorum animos verbis confirmavit

3.5.4 Participles

Just as prepositional phrases can either modify a noun (with ATR) or a verb (with ADV), so can participial phrases as well. When a participial phrase delimits the possible reference of a noun phrase (as in section 3.4.2 above), it should depend on that noun via ATR. When it further specifies the action of a verb, it should depend on the verb via ADV.

et cantant novum canticum dicentes dignus es accipere librum
3.5.5 Subordinate clauses

Helvetii quoque reliquos Gallos virtute praecedunt, quod fere cotidianis proeliis cum Germanis contendunt.

3.5.6 Gerunds

querebatur ... illos dubitando et dies prolatando magnas opportunitates corrumpere

3.6 ATV/AtvV (Complement)

Following the PDT, we use the tag ATV for all complements not participating in government (complements that are governed by their verb are assigned the tag OBJ). These are typically noun phrases and adjectives that agree with their head noun morphologically, but differ from typical attributes in that they also qualify the function of the verb.
The PDT use of ATV is largely similar to the account of praedicativa given in Pinkster [6], and can be simplified to the following two examples contained therein:

- Galli laeti in castra pergunt
- Cicero consul coniurationem Catilinae detexit

In the first example, an attributive reading of laeti would lead to the translation "The happy Gauls entered the camp." As an ATV, it would be rendered "The Gauls happily entered the camp" - while laeti agrees morphologically with the subject Galli, it simultaneously specifies the nature of the predicate. Since it is an inflected adjective (and not the adverb laete), it still bears a syntactic relationship to the noun phrase and should therefore depend on it (and not simply on the verb via ADV). This results in the following tree:

```
pergunt
    PRED
    Galli SBJ
    in AuxP
    laeti ATV
    castra OBJ
Galli laeti in castra pergunt
```

If the head noun phrase in such constructions is implied, the praedicativum should depend on the main verb via AtvV. (If laeti here were a SBJ depending on pergunt, the sentence would mean "The happy ones entered the camp.")

```
pergunt
    PRED
    laeti AtvV
    in AuxP
    castra OBJ
laeti in castra pergunt
```

The main verbs in this variety of ATV construction are typically confined (as Pinkster notes) to a limited number of groups, mainly verbs involving motion and several that behave like copulas, such as lateo below.
In the second example above, *consul* is not a simple attribute (or appositive) of Cicero since it qualifies the nature of the verb: Cicero uncovered Catiline’s conspiracy *as consul* (i.e., *when he was consul*). Since *consul* agrees with Cicero morphologically while also modifying the main predicate, it should depend on the noun via ATV.

Most phrases that involve x doing something *as y* should be annotated with ATV. In the example below, the subject (an implied Misenus), gives himself *as a friend* to Aeneas.
postquam illum vita victor spoliavit Achilles, Dardanio Aeneae sese fortissimus heros addiderat socium

This use of ATV is also important for the annotation of several absolute constructions. When ablative absolutes do not contain a participle, as in Cicero consule below, we can treat one member as simultaneously modifying both its head noun and the main predicate (with Cicero as consul) and therefore assign it the tag ATV.

Cicerone consule, senatus coniurationem Catilinae detexit

3.7 PNOM (predicate nominal)

Predicate nominals (subject complements) depend on a verbal head.

Predicate nominals are not limited, however, to noun phrases and adjectives in the
same case as the sentence subject. They can also appear in a variety of other constructions, such as genitives.

PNOMs most often appear with inflections of *sum*, but can appear with any "linking" verb, including *videor* and *fio*.

### 3.8 OCOMP (Object Complement)

Like predicate nominals (subject complements), object complements depend on their verbal head. Object complements are generally of the form *to make x y*, and most often appear with verbs such as *facio*.
3.9 "Bridge" structures

In the annotation style adopted by the Prague Dependency Treebank, coordinators (including punctuation), "apposing" words, prepositions, and subordinate conjunctions (all described below) function as "bridges" between their children and their own heads. In the phrase *contendunt cum Germanis*, for example, the noun *Germanis* depends on the preposition, but with the relation it would bear to *contendunt* - namely, OBJ. The preposition is assigned a sort of "dummy" relation AuxP, meant to signify that the true relationship is that between *Germanis* and *contendunt*, and that the preposition simply acts a mediator between the two. The advantage of this method is that it preserves the lexical association between pairs of words, regardless of intermediate structure, as between *differunt* and the words *lingua*, *institutis* and *legibus* in the two examples below.

- *hi omnes lingua, institutis, legibus inter se differunt*
- *hi omnes in lingua, institutis, legibus inter se differunt*

This method therefore recognizes the "omissibility" of prepositions in certain contexts: given Latin's rich inflection, a noun in the ablative case can function as if a preposition were included in its morphology.

In the following subsections, we delineate the different methods by which this approach annotates coordination, apposition, prepositional phrases and subordinate clauses.

3.9.1 COORD (coordinator)

An example of a coordinated structure is given below.

```
hi omnes lingua, institutis, legibus inter se differunt
```

Here, *lingua*, *institutis*, and *legibus* all depend on the single final coordinator that separates them: the final comma (that separating *institutis* from *legibus*). Each of these words depends on that comma with a complex tag comprised in part of the relation they bear to the coordinator’s head: since each would individually modify *differunt* as an ADV, each depends on the coordinator with the tag ADV_CO. The coordinator then depends on *differunt* with the tag Coord. See section 3.12.1 (AuxX) on how to annotate the non-final coordinator in lists.

If coordination involves multiple prepositional phrases (AuxP) or subordinate clauses (AuxC), the _CO suffix should be appended to the children of the preposition or subordinating conjunction, respectively (AuxP_CO and AuxC_CO are not valid tags).
cantabant quasi canticum novum ante sedem et ante quattuor animalia et seniores

scio opera tua et laborem et patientiam tuam et quia non potes sustinere malos
And even if words are coordinated on different levels, each should only have one _CO suffix.
et quattuor animalia dicebant amen et seniores ceciderunt et adoraverunt

3.9.2 APOS (apposing elements)

An example of apposition is given below.

In iis erat Fulvius, senatoris filius, quem retractum ex itinere parens necari iussit

Here the two phrases in apposition to each other are Fulvius and senatoris filius; the appositional-coordinating element that separates them is the comma between Fulvius and senatoris. Both Fulvius and senatoris filius depend on the apposing comma via the relation they each individually bear to the phrase’s head (erat). Since they are both the subjects of erat, they modify the apposing word via the complex tag SBJ_AP; the apposing word then modifies erat via Apos.
3.9.3 AuxP (preposition)

An example of a prepositional phrase is given below.

```
contendunt
   ---
       cum
       AuxP
   Germanis
   OBJ
```

contendunt cum Germanis

Our method of annotation sees prepositions as acting as a functional bridge between their child and head. Here the object of the preposition (*Germanis*) would depend on the preposition (*cum*) via the relationship it would hold to the preposition’s head if the preposition were absent (OBJ). The preposition then depends on its head via the relation AuxP.

3.9.4 AuxC (conjunction)

Subordinate (non-relative) clauses are annotated in a manner similar to prepositional phrases, with the subordinating conjunctions acting as a functional bridge between the embedded verb and the parent of the phrase.

```
praecedunt

Helvetii SBJ

quoque AuxZ

Gallos OBJ

virtute ADV

quod AuxC

reliquos ATR

, AuxX

contendunt ADV

proelii ADV

cotidianis ATR

fere ADV

Germanis OBJ
```

Helvetii quoque reliquos Gallos virtute praecedunt, quod fere cotidianis proelii cum Germanis

Here the subordinate verb *contendunt* depends on its head (*quod*) via the relationship ADV. *quod* then depends on *praecedunt* via the "bridge" relationship AuxC.
3.10 AuxR (reflexive passive)

As in the PDT, AuxR should be used for the annotation of reflexive passives (rare in Classical Latin but present in later dialects). Reflexive passives are used to express an action without specifying the agent responsible for it. In the example below, *se habet forma* (literally, "the form holds itself") is equivalent to the agent-less passive "the form is held (i.e., considered) to be." Even though *forma* is the subject of *habet*, it is not the entity that is actually doing the *considering*. Reflexive passives differ from other uses of the reflexive form in that they result in a passive concept, even if the verb form is morphologically active.

```
se habet forma substantialis ad esse naturae
```

AuxR should only be used for these passivizing uses of *se*. Most instances of reflexive *se* in Latin should be annotated with OBJ, as in the example found in section 4.2 (*... paraverunt se ut tuba canerent*).

3.11 AuxV (auxiliary verb)

Auxiliary verbs (mostly seen with passive participles) depend on the tensed verb.

```
e a res est Helvetiis per indicium enuntiata
```

3.12 Punctuation

Our methods of annotating punctuation follow that established by the PDT, which assigns several different functional tags.
3.12.1 AuxX (commas)

If a comma is not the head of a coordinated or appositional phrase, it should be annotated with AuxX and depend on the head of its clause. In coordinated lists, this head is the final comma.

Here, *lingua*, *institutis*, and *legibus* all depend on the final comma (that separating *institutis* from *legibus*) via ADV_Coord. The remaining comma (that separating *lingua* from *institutis*) should then depend on the final comma via AuxX. Note that if the coordinator is a content word (e.g., *et*), it depends on the final coordinator via AuxY; if it is punctuation, it depends via AuxX.

If a comma is used to separate a subordinate clause, it should depend on that clause’s head.
Helvetii quoque reliquos Gallos virtute praecedunt, quod fere cotidianis proeliis cum Germanis contendunt

3.12.2 **AuxG (bracketing punctuation)**

"Bracketing" punctuation surrounds an enclosed phrase, and most frequently appears as quotation marks or parentheses (not commas). These punctuation marks should depend on the head of the bracketed phrase via AuxG.

In this example, the quotation marks surrounding *refer* should both depend on *refer* via AuxG; those surrounding *ad senatum* should depend on *ad*.

AuxG should also be used to annotate abbreviation, with the period depending on the abbreviated word.
libertatem et consulatum L. Brutus instituit

3.12.3 AuxK (terminal punctuation)

Final punctuation (if present) should depend on the root via AuxK.

Vrbem Romam a principio reges habuere;
3.13 AuxY (sentence adverbials)

Sentence adverbials (also called disjuncts) are those that pertain to the entire sentence and often express the author’s opinion about the validity of what’s being said. These include words like *sane, certe, vero,* and also question markers such as *num* and *utrum.* A full list includes *adhuc, alias, consequens, contra, enim, ergo, ideo, igitur, inde, ita, item, postea, praeterea, propterea, sic, tamen, tum, tunc, unde, utrum,* and *vero.* Note that some words are ambiguous between sentence adverbials and subordinating conjunctions: when *enim,* for instance, means “certainly” or “indeed,” it should be annotated with AuxY; when it introduces a subordinate clause (“for ...”) it should be annotated with AuxC.

These also include exclamations.
et
COORD
fui
PRED_CO
statim
AuxZ
in
AuxP
spiritu
ADV
ecce
AuxY
sedis
AuxV
erat
in
AuxP
caelo
OBJ
posita
PRED_CO
statim fui in spiritu et ecce sedis posita erat in caelo

AuxY should also be used to annotate coordinators that are not commas (e.g., \textit{et}, -\textit{que}) when they are not the head of the coordinated phrase. (Non-head commas in these structures should be annotated with AuxX.)

fefellit
PRED
que
COORD
comites
OBJ_CO
que
natum
OBJ_CO
virum
OBJ_CO
comites natumque virumque fefellit

3.14 AuxZ (emphasizing particles)

AuxZ should be assigned to particles with a "poor" meaning content that emphasize one specific word in the sentence (as distinct from AuxY, which emphasizes the sentence as a whole). This occurs most often with words such as \textit{only}, \textit{as well as}, \textit{also}, etc, and also with negation such as \textit{non}. A full list includes \textit{adeo}, \textit{etiam}, \textit{ita}, \textit{item}, \textit{nec}, \textit{non}, \textit{nondum}, \textit{numquam}, \textit{omnino}, \textit{praecipue}, \textit{quanto}, \textit{quantum}, \textit{quasi}, \textit{statim}, \textit{tanto}, \textit{tantum}, and \textit{umquam}. 

35
ab Lentulo, Cethego, Statilio, item Cassio postulant ius iurandum

Here the four names are coordinated on the comma between Statilio and item; item (as well as) should depend on Cassio via AuxZ.

Negative particles should depend on the word that is being negated (whether a verb, adjective, etc.).

4 How to Annotate Specific Constructions

4.1 Ellipsis

Ellipsis - the omission of words in a sentence that are recoverable from contextual cues - is a ubiquitous phenomenon in literary texts. Our method of representing ellipsis attempts to preserve the structure of the tree as much as possible. We accomplish this by assigning a complex tag to orphaned words. This tag preserves the path from the word itself to the elided word’s head. Consider the example of unam incolunt Belgae, aliam Aquitani (Caesar, De Bello Gallico 1.1) given in the figure below. Here, the verb incolunt is missing from the second clause. We can preserve the structure of the
tree by assigning the head of *aliam* and *Aquitani* to be the head that *incolunt* would have if it were in the sentence (the coordinating comma), and by assigning tags to each that preserve the path: *aliam* should be the object (OBJ) of *incolunt*, which should then depend on the coordinating comma via Pred_Co; it therefore receives the tag OBJ_ExD0_PRED_Co (like the PDT, ExD here signifies an external dependency; the following numeral indexes the ellipsis, since in some sentences multiple words are elided). Likewise, *Aquitani* should be the subject (SBJ) of the elided word; it therefore receives the tag SBJ_ExD0_PRED_Co. This method allows us to use the complex tags to reconstruct the tree as necessary.

4.2 Relative Clauses

Different relative clauses must be annotated differently based on their syntactic function in the sentence. Relative clauses with antecedents, as in the following examples, are generally attributive, and should modify the antecedent via ATR. The head of a relative clause is the subordinate verb; this is the element that depends on the antecedent.
collis ab summo aequaliter declivis ad flumen Sabim, quod supra nominavimus, vergebat

Not all relative clauses have antecedents. These should be annotated according to the syntactic function of the entire relative phrase:

qui habet aurem audiat

In the first example, the subject of audiat is the entire phrase *he who has an ear* (*qui habet aurem*). Since habet is the head of this phrase, it depends on audiat as the SBJ (within the phrase, *qui* is the SBJ of habet and *aurem* is its OBJ). In the second example, the object of scribe is the phrase *what you see* (*quod vides*). Since vides is the head of this phrase, it depends on scribe as the OBJ (within the relative phrase, *quod* is the OBJ of vides).

Note that this method of annotation is structurally different from that for subordinate clauses, in which the subordinate verb depends on the subordinating conjunction, which then depends on a word outside of the clause. See section 3.9.4 (AuxC) for
information on annotating subordinate clauses.

Also, not all relative pronouns signal a relative clause; some are simply attributive and modify their head via ATR.

quam ob rem discede atque hunc mihi timorem eripe

quibus rebus confectis omnia propere per nuntios consuli declarantur
4.3 Indirect Questions

Since the interrogative word in indirect questions has a syntactic function within the subordinate clause, it should be annotated in a manner similar to self-contained relative clauses. In the sentence below, *quid* is the direct object of *sentiam* [= "what I think I will explain"] and should therefore depend on it via OBJ.

```
  exponam
    PRED
  sentiam
    OBJ
    ipse
    quid
    SBJ
    OBJ

quid ipse sentiam exponam
```

4.4 The Ablative Absolute

The ablative absolute is a grammatical construction similar to the English nominative absolute, where a noun and (typically) a participle form a phrase that is disjoint from the grammar of the rest of the sentence; in Latin both the noun and participle are inflected in the ablative case, as in the following:

- *his rebus cognitis Caesar Gallorum animos verbis confirmavit*

  Following Pinkster [6], we treat ablative absolutes as an embedded predication that functions as an adjunct. In common absolutes (with a noun + participle), the noun should be annotated as the subject of the participle, with the participle (as the head of the ablative absolute phrase) depending on the main verb as an adverbial. We would annotate the example above in the following way:

```
  confirmavit
    PRED
    cognitis
    ADV
    Caesar
    SBJ
    animos
    OBJ
    verbis
   ADV
    rebus
    SBJ
    Gallorum
    ATR
    his
    ATR

his rebus cognitis Caesar Gallorum animos verbis confirmavit
```

In absolutes involving no participle, the head noun should depend on the main verb via ADV, with its child (the element the head is "functioning as") dependent on it via ATV.
4.5 Direct Speech

Direct speech should be annotated the same way as indirect discourse, by attaching the head of the "spoken" phrase to the predicate that introduces the speaking.

4.6 Direct Address

As in the PDT, vocatives should depend on their verbal heads via ExD.
4.7 Accusative + Infinitive

In indirect discourse and other accusative + infinitive constructions, the infinitive verb is the head of its phrase. This verb represents the entire clause and should depend via OBJ on the word that introduces the discourse. Within the phrase, standard annotation applies (so that the subject, while accusative, still depends on the indirect infinitive via SBJ).

- contentum te esse dicebas

This sentence should be annotated in the following way:

```
dicebas
   ↓
  ↓
esse
   ↓
te
   ↓
contentum
   ↓
   ↓
   ↓
P NOM
SBJ
```

4.8 Gerunds and Gerundives

As a verbal noun, gerunds are relatively straightforward to annotate: they should simply be treated as nouns and annotated according to their syntactic function in the sentence:

```
querebatur
   ↓
corrumpere
     ↓
illos
     ↓
   ↓
   ↓
   ↓
SBJ
OBJ
COORD
ADV_CO
OBJ
A TR
```

querebatur ... illos dubitando et dies prolatando magnas opportunitates corrumpere

Gerundives, on the other hand, behave more like participles in that they can function either as an attribute or in a dominating construction. When attributive, gerundives
should be labeled ATR; when dominating, they should be annotated according to their specific role in the sentence. A test for which tag is appropriate is whether or not the gerundive can be omitted: if it can be left out of the sentence without changing the lexical meaning of the predicate, it’s ATR; if not, then it is dominating. In the example below, *effeminandos* cannot be left out of the sentence since *quaer pertinente ad animos* ("which pertain to the minds") doesn’t make sense on its own.

```
pertinent
---
quae SBJ
ad AuxP
effeminandos OBJ
animos SBJ

quaer ad effeminandos animos pertinent
```

Our intuition here may be to treat the noun *animos* as the direct object of the gerundive (since we idiomatically translate the phrase with such a sense: "which pertain to effeminating the mind"), but we should keep in the mind that a gerundive is a passive form, which then makes *animos* a subject.

An attributive use of a gerundive can be seen in the fragment *privatio formae inducendae* ("the privation of the form to be inserted"). Here *inducendae* is omissible and should therefore be labeled with ATR.

```
est PRED
---
vel COORD
privatio PNON

principium SBJ CO
initium SBJ_CO
generationis ATR

principium vel initium generationis est privatio formae inducendae
```

When a gerundive appears in a passive periphrastic construction, it should be treated as a predicate nominal:

```
est PRED
---

Carthago SBJ

delenda PNON
```

43
4.9 Comparison

Adjectives that signify comparison can appear with an overt *quam* or without one (where the compared object appears in the ablative case). These should both be annotated similarly: when *quam* is present, it is the head of an embedded predication and should be annotated with AuxC. The subordinate verb should then modify *quam* via ADV. If *quam* is present but the subordinate verb is not (as in *puella est laetior quam puer*), it should be annotated as an instance of ellipsis. When *quam* is absent and an ablative noun is compared instead, it should modify the comparative adjective via ADV.

```
luce sunt clariora nobis tua consilia omnia
```

```
puella est laetior quam puer est
```

4.10 Participles

Active and passive participles can ambiguously modify both nouns and verbs. When attributively modifying nouns, they restrict the reference of the noun phrase: in the example below, what is seen is not simply any angel, but an “angel descending from the sky.” The possible referent of ”*angelum*” here is restricted from the set of all possible angels to the one satisfying this particular attribute.
et post haec vidi alium angelum descendentem de caelo

Ambiguity arises with nominative participles: are they attributively restricting the reference of the noun phrase (=ATR), or do they further qualify the circumstances of the verb (=ADV)? In the following example, this difference can be captured in two different translations.

- et angeli cantant novum canticum dicentes dignus es accipere librum
  - ATR, modifying angeli: "those angels who were saying 'you are worthy to accept the book' are singing a new song"
  - ADV, modifying cantant: "the angels are singing a new song, saying 'you are worthy to accept the book'"

In this case, the participial phrase further expresses the conditions under which the main verb takes place: it should then modify that verb via ADV (the second option).
et angeli cantant novum canticum dicentes dignus es accipere librum

Active participles are not the only kind that can modify verbs via ADV - passive participles can as well. In the following example, the ATR/ADV ambiguity produces two different translations as well.

- his nuntiis litterisque commotus Caesar duas legiones in citeriore Gallia novas conscripsit
  
  - ATR, modifying Caesar: "the Caesar who was bothered by these announcements and letters drafted a new army in hither Gaul."
  
  - ADV, modifying conscripsit: "Having been bothered by these announcements and letters, Caesar drafted a new army in hither Gaul."

Again, the participial phrase here further expresses the conditions under which the main verb takes place and should depend on the verb via ADV.
4.11 Names

The praenomen and nomen in all names should depend on the cognomen via ATR (if the cognomen is deficient, the praenomen should depend on the nomen). Any abbreviating periods should depend on the word they abbreviate via AuxG.

References


