

Guidelines for the Syntactic Annotation of the Ancient Greek Dependency Treebank (1.1) *

David Bamman and Gregory Crane

The Perseus Project, Tufts University

September 1, 2008

Contents

1	Introduction	2
2	Dependency Grammar	3
3	Annotation Style	3
3.1	PRED (predicate)	4
3.2	SBJ (subject)	5
3.2.1	Nominative nouns	6
3.2.2	Accusative nouns	6
3.2.3	Genitive nouns	7
3.2.4	Infinitive verbs	8
3.2.5	Relative Pronouns	8
3.3	OBJ (object)	8
3.3.1	Accusative nouns	9
3.3.2	Accusative + infinitive constructions	9
3.3.3	Relative clauses	10
3.3.4	Subordinate clauses	10
3.3.5	Infinitive verbs	11
3.3.6	Complements	11
3.3.7	Indirect objects	12
3.3.8	Passive agents	13
3.4	ATR (attribute)	14
3.4.1	Adjectives	14
3.4.2	Determiners	15

*These guidelines are based on those developed for the annotation of Latin syntax in collaboration with the Index Thomisticus [1]. Thanks are due to Meg Luthin for finding many of these example sentences and to Dan Ullucci for illustrating them.

3.4.3	Participles	15
3.4.4	Prepositional phrases	16
3.4.5	Agreeing nouns	16
3.4.6	Non-agreeing nouns	16
3.4.7	Relative clauses	18
3.4.8	Pronouns	19
3.5	ADV (adverbial)	19
3.5.1	Adverbs	19
3.5.2	Prepositional phrases	20
3.5.3	Nouns	20
3.5.4	Participles	20
3.5.5	Subordinate clauses	21
3.6	ATV/AtvV (Non-governed complement)	21
3.7	PNOM (predicate nominal)	23
3.8	OCOMP (object complement)	24
3.9	“Bridge” structures	25
3.9.1	COORD (coordinator)	25
3.9.2	APOS (apposing elements)	27
3.9.3	AuxP (preposition)	27
3.9.4	AuxC (conjunction)	28
3.10	Punctuation	29
3.10.1	AuxX (commas)	29
3.10.2	AuxG (bracketing punctuation)	30
3.10.3	AuxK (terminal punctuation)	31
3.11	AuxY (sentence adverbials)	32
3.12	AuxZ (emphasizing particles)	34
4	How to Annotate Specific Constructions	34
4.1	Ellipsis	34
4.2	Relative Clauses	36
4.3	Particles	38
4.4	The Genitive and Accusative Absolute	39
4.5	Accusative + Infinitive	40
4.6	Tmesis	41
4.7	Direct Speech	42
4.8	Direct Address	42

1 Introduction

Trebanks – 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 [6]), Early Modern English (Kroch et al. [5]), Old English (Taylor et al. [10]), Early

New High German (Demske et al. [2]) and Medieval Portuguese (Rocio et al. [8]).

In what follows we present a preliminary set of annotation guidelines for the Ancient Greek Dependency Treebank adapted from those developed for the Latin Dependency Treebank in collaboration with the Index Thomisticus. The annotation style proposed here is predominantly informed by two sources: the dependency grammar used by the Prague Dependency Treebank [4, 3] (itself based on Sgall et al. [9]), and the Latin grammar of Pinkster [7].

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 Greek, 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*, for instance, would look like Figure 1.

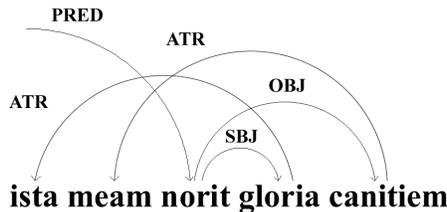


Figure 1: Dependency graph of *ista meam norit gloria canitiem* (Prop. I.8.46). Arcs are directed from heads to their dependents.

Dependency grammar is also appropriate for Greek and 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

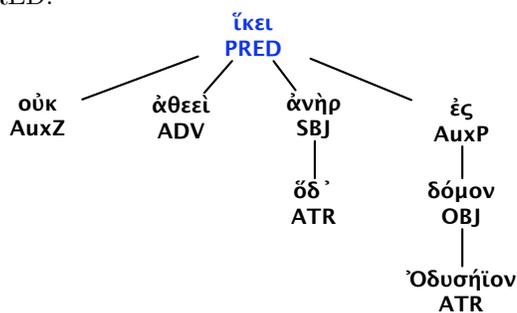
Different treebanks and grammars, however, assign syntactic functions differently. The general model for our style of representation is that used by the Prague Dependency Treebank, with several important departures arising from Pinkster’s [7] Latin grammar. The following table lists all of the tags currently in use; the following subsections further elaborate each.

PRED	predicate
SBJ	subject
OBJ	object
ATR	attributive
ADV	adverbial
ATV/AtvV	complement
PNOM	predicate nominal
OCOMP	object complement
COORD	coordinator
APOS	apposing element
AuxP	preposition
AuxC	conjunction
AuxV	auxiliary verb
AuxX	commas
AuxG	bracketing punctuation
AuxK	terminal punctuation
AuxY	sentence adverbials
AuxZ	emphasizing particles
ExD	ellipsis

Table 1: Complete tagset.

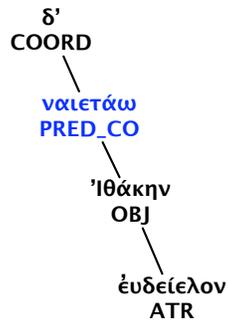
3.1 PRED (predicate)

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.

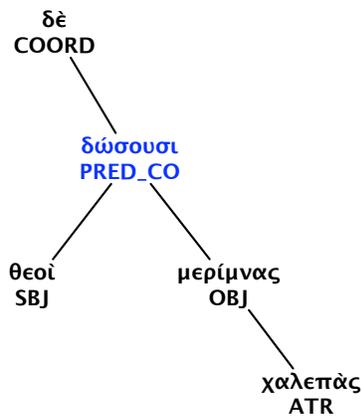


οὐκ ἄθξει ὄδ' ἄνθρωπος Ἰουδοσθηίων ἐς δόμον ἴκει
 ("This man does not come to the Odyssean palace without the will of the gods," Od. 18.353)

If a sentence begins with an initial conjunction, the main verb is dependent on that conjunction and the conjunction then depends on the root.



ναιετάω δ' Ἰθάκην ἐυδείελον
 ("I inhabit clear-seen Ithaca," Od. 9.24)

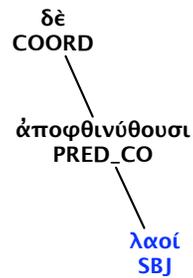


χαλεπὰς δὲ θεοὶ δώσουσι μερίμνας
 ("And the gods will give difficult troubles," Op. 178)

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:

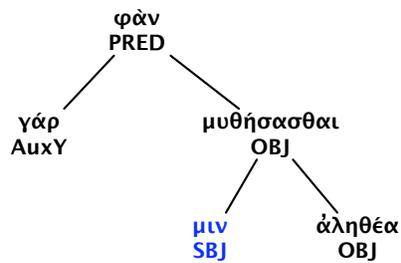
3.2.1 Nominative nouns



ἀποφθινύθουσι δὲ λαοί
("And the men perish," Op. 243)

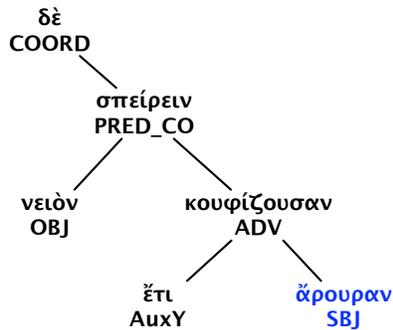
3.2.2 Accusative nouns

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



φάν γάρ μιν ἀληθέα μυθήσασθαι
("For they thought that he spoke the truth," Od. 18.342)

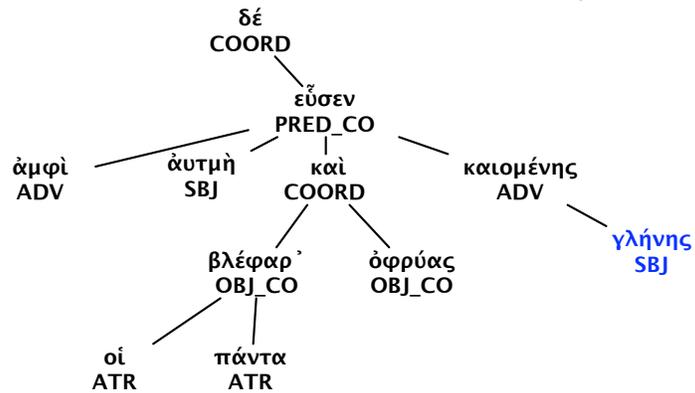
Accusative absolutes, like the more common genitive absolutes below, are treated as a form of embedded predication, with the accusative dependent on the participle as its SBJ.



νειὸν δὲ σπείρειν ἔτι κουφίζουσαν ἄρουραν
 (“Sow fallow-land with the soil still getting light,” Op. 463)

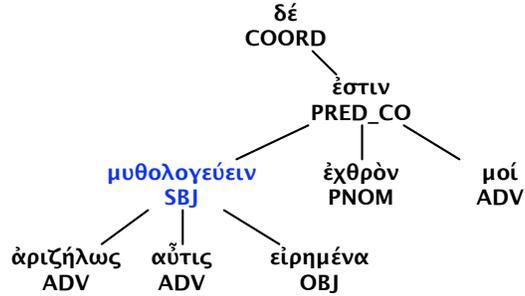
3.2.3 Genitive nouns

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



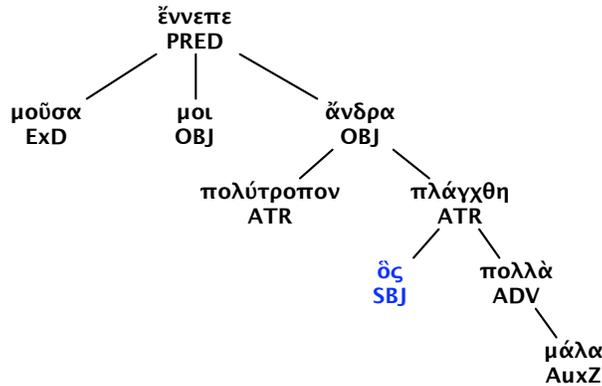
πάντα δὲ οἱ βλέφαρ' ἀμφὶ καὶ ὄφρῦας εὔσεν ἀυτμὴ γλῆνης καιομένης ...
 (“And the flame singed his eyelids and brow all around, his eyeball burning,”
 Od. 9.389-90)

3.2.4 Infinitive verbs



ἐχθρὸν δέ μοί ἐστιν αὐτίς ἀριζήλως εἰρημένα μυθολογεύειν
 (“To tell a tale clearly again is irksome to me,” Od. 12.452-3)

3.2.5 Relative Pronouns

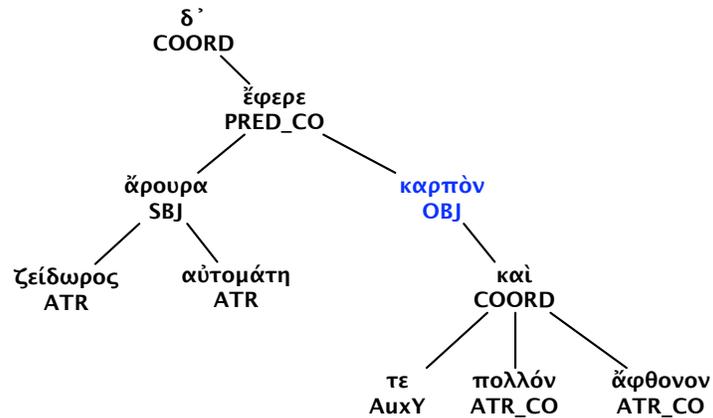


ἄνδρα μοι ἔννεπε, μοῦσα, πολύτροπον, ὃς μάλα πολλά πλάγχθη
 (“Tell me, muse, of the many-devised man, who wandered very many ways,”
 Od. 1-2.)

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:

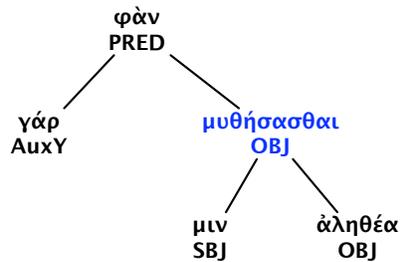
3.3.1 Accusative nouns



καρπὸν δ' ἔφερε ζείδωρος ἄρουρα αὐτομάτη πολλὸν τε καὶ ἄφθονον
 ("The fruitful earth unforced bare them many and abundant fruit," Op.
 117-18.)

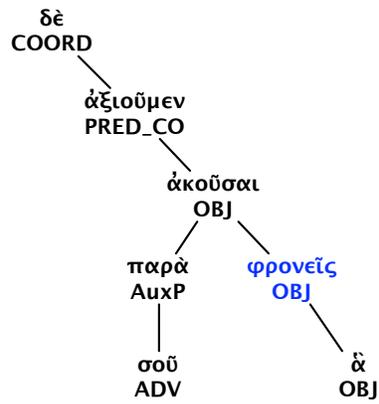
3.3.2 Accusative + infinitive constructions

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



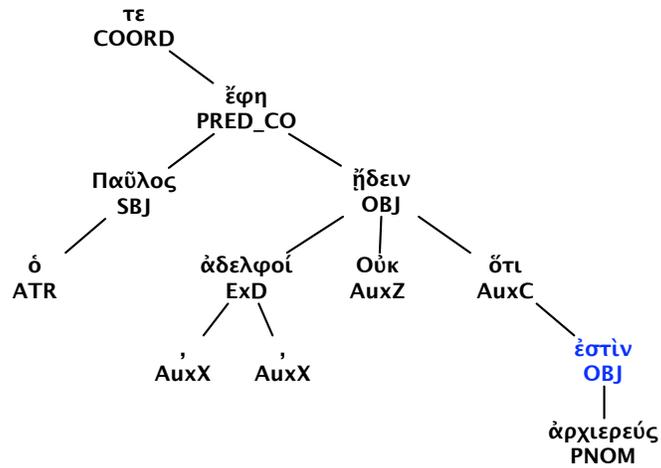
φάν γάρ μιν ἀληθέα μυθήσασθαι
 ("For they thought that he spoke the truth," Od. 18.342)

3.3.3 Relative clauses



ἀξιοῦμεν δὲ παρὰ σοῦ ἀκοῦσαι ἃ φρονεῖς
 ("We want to hear from you what you think," Acts 28.22)

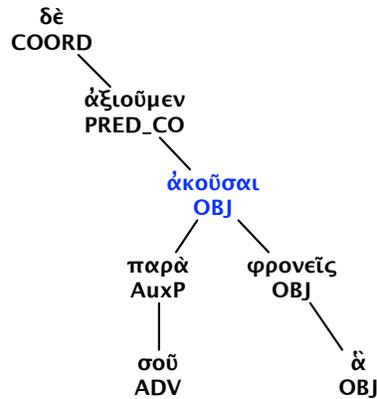
3.3.4 Subordinate clauses



ἔφη τε ὁ Παῦλος Οὐκ ᾔδειν, ἀδελφοί, ὅτι ἐστὶν ἀρχιερεὺς
 ("Paul said, 'I didn't know, brothers, that he was high priest,'" Acts 23.5)

3.3.5 Infinitive verbs

This include both verbs that function as traditional direct objects as well as those that complete verbs like ἐθέλω (to be willing), μέλλω (to intend), δύναμαι (to be able) and βούλομαι (to want).



ἀξιοῦμεν δὲ παρὰ σοῦ ἀκοῦσαι ἃ φρονεῖς
 (“We want to hear from you what you think,” Acts 28.22)

3.3.6 Complements

Our notion of object 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). These are arguments that pertain specifically to the verb in question and cannot readily be applied to other verbs as well. The following English examples all have one argument that is typically considered the “direct object”: this would generally appear in the accusative case in Greek or Latin and should be annotated OBJ. However, they also each contain one phrase that completes the action of the verb (*with wine*, *onto his right hand* and *out of pork*) that should be annotated with OBJ as well.

- They sprinkled the lamp *with wine*
- He threw the ring *onto his right hand*
- The cook fashioned birds and fish *out of pork*

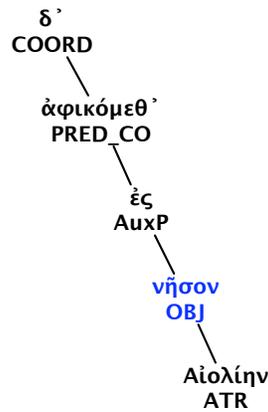
In practice, complements like these are often confused with ADV. Objects are obligatory arguments of verbs, while adverbials are always optional. If an object is left out of the sentence, one of two things happens: the sentence

becomes ungrammatical or a different sense of the verb is implied (one with a reduced valency). If we were to add any of the complements above to another verb (like “runs”) it wouldn’t make sense (e.g., “runs” in “he runs *out of pork*” does not mean the same thing as “he runs”).

Adverbs, however, can apply to almost any verb, so we could easily add “yesterday” to any of the examples and they would still be grammatical:

- They sprinkled the lamp with wine *yesterday*
- He threw the ring onto his right hand *yesterday*
- The cook fashioned birds and fish out of pork *yesterday*

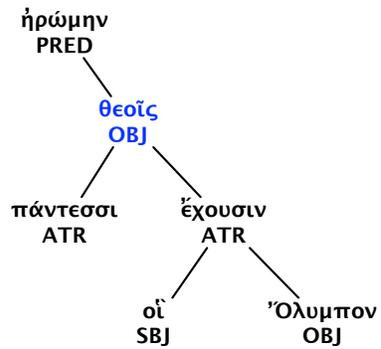
One category in particular that should always be annotated with OBJ consists of words that specify “motion toward” something.



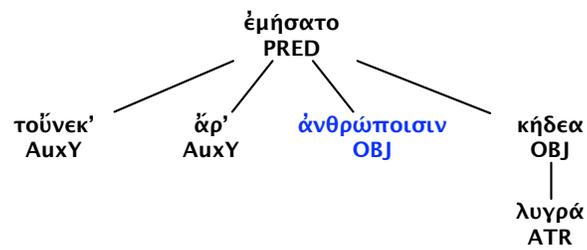
Αἰολίην δ' ἐς νῆσον ἀφικόμεθ
 (“We came to the Aeolian island,” Od. 10.1)

3.3.7 Indirect objects

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



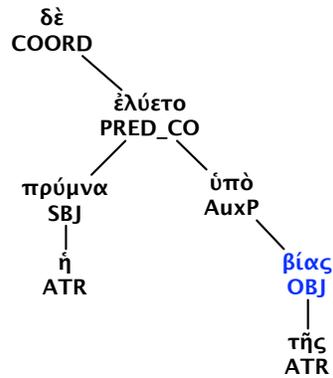
ἠρώμην πάντεσσι θεοῖς οἱ Ὀλυμπον ἔχουσιν
 (“I prayed to all the gods who hold Olympus,” Od. 12.337)



τοῦνεκ' ἄρ' ἀνθρώποισιν ἐμήσατο κήδεα λυγρά
 (“Therefore he planned mischievous troubles against men,” Op. 49)

3.3.8 Passive agents

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

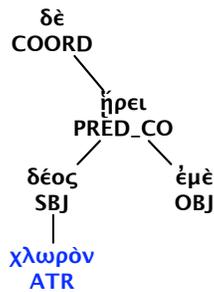


ἡ δὲ πρύμνα ἐλύετο ὑπὸ τῆς βίας
 (“The stern began to be broken up by the waves,” Acts. 27.41)

3.4 ATR (attribute)

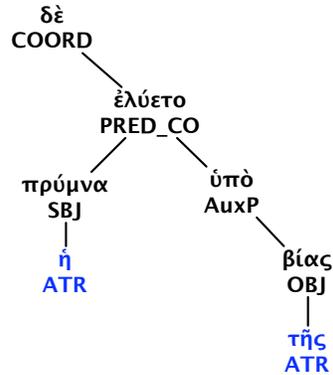
Attributes are those phrases that attributively specify (or delimit) the meaning of their head. Most commonly these are adjectives, but can include other classes as well, such as nouns, relative clauses and prepositional phrases.

3.4.1 Adjectives



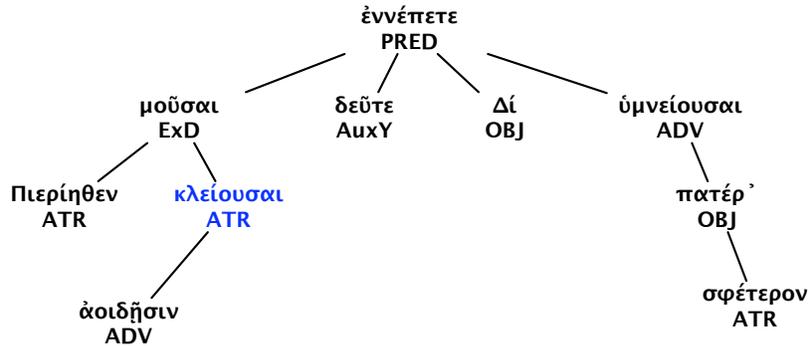
ἐμὲ δὲ χλωρὸν δέος ἦρει
 (“And pale fear seized me,” Od. 11.43)

3.4.2 Determiners



ἡ δὲ πρύμνα ἐλύετο ὑπὸ τῆς βίας
 (“The stern began to be broken up by the waves,” Acts. 27.41)

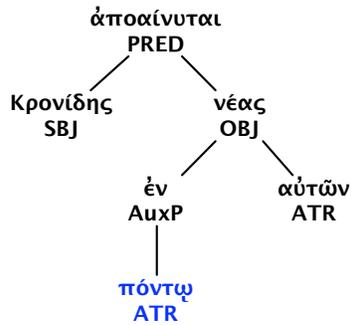
3.4.3 Participles



μοῦσαι Πιερίηθεν ἀοιδῆσιν κλείουσαι δεῦτε, Δί' ἐννέπετε, σφέτερον πατέρ'
 ὑμνείουσαι
 (“Muses of Pieria giving glory by song, here!, tell of Zeus, singing of your
 father,” Op. 1.1)

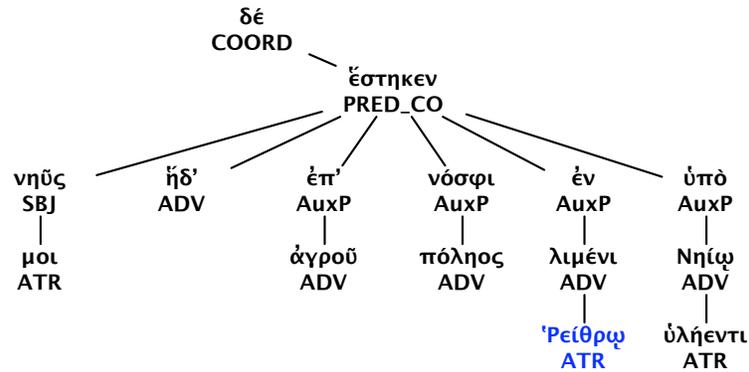
Note that there are two participles in this sentence, one attributive (κλείουσαι) and one adverbial (ὑμνείουσαι). The first restricts the reference of the noun (“Muses *who give glory* ...”) while the second modifies the verb because it gives more information on *how* that action is undertaken (e.g., “tell of Zeus by singing”).

3.4.4 Prepositional phrases



νέας ἐν πόντῳ Κρονίδης ἀποαίνυται αὐτῶν
 (“The son of Cronos takes away their ship at sea,” Op. 247)

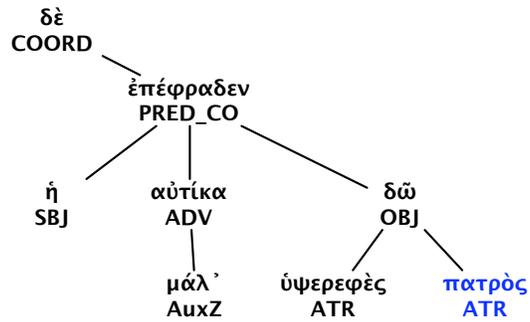
3.4.5 Agreeing nouns



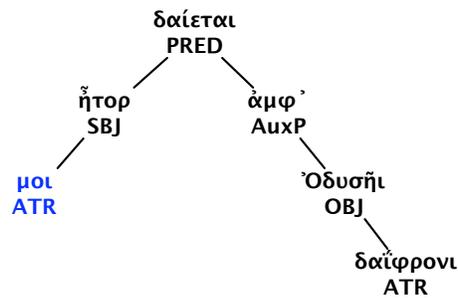
νηῦς δέ μοι ἧδ' ἔστηκεν ἐπ' ἀγροῦ νόσφι πύλῃος, ἐν λιμένι ῥεῖθρων ὑπὸ Νηίῳ ὕληεντι
 (“My ship lies there beside the field, away from the city, in the harbor of Rheithron, under woody Neion,” Od. 1.185-6)

3.4.6 Non-agreeing nouns

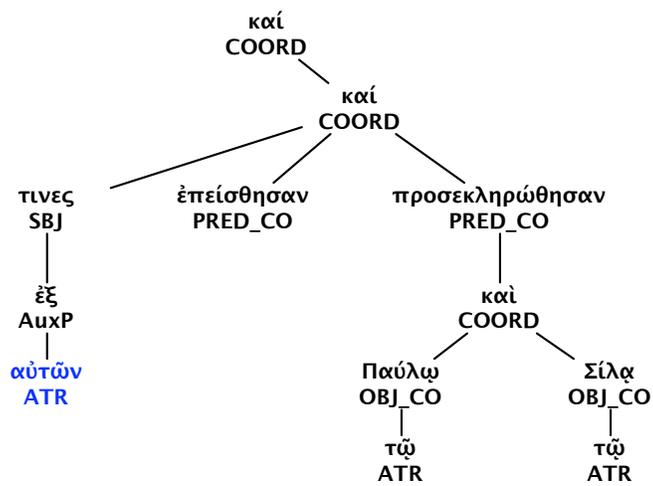
Possessive genitives and datives, objective genitives and partitives all fall in this category.



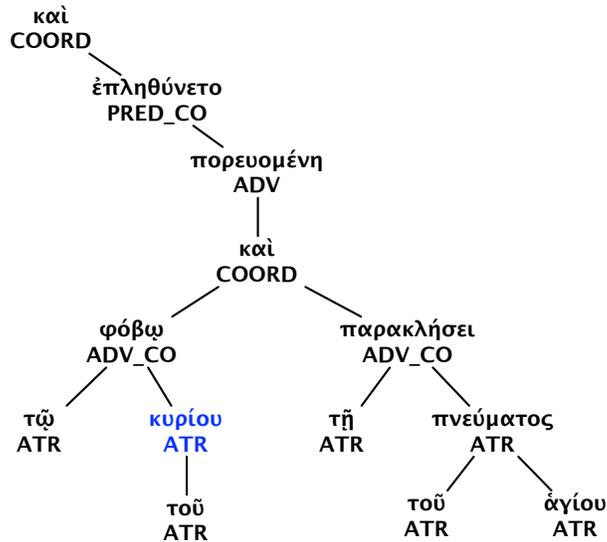
ἡ δὲ μάλ' αὐτίκα πατρὸς ἐπέφραδεν ὑψερεφές δῶ
 (“And she quite immediately pointed out the high-roofed homes of her father,” Od. 10.111)



καρπὸν δ' ἔφερε ζείδωρος ἄρουρα αὐτομάτη πολλὸν τε καὶ ἄφθονον
 (“My heart is torn over wise Odysseus,” Od. 1.48.)

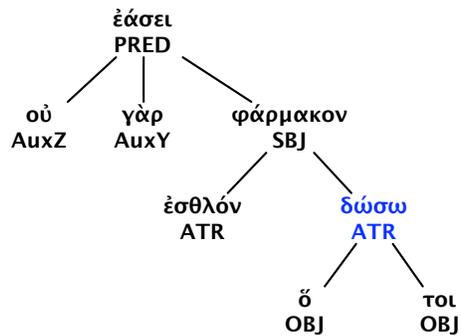


καὶ τινες ἐξ αὐτῶν ἐπίσθησαν καὶ προσεκληρώθησαν τῷ Παύλῳ καὶ τῷ Σίλῳ
 (“And some of them were persuaded and joined Paul and Silas,” Acts 17.4)



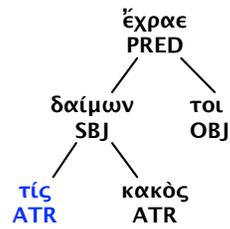
καὶ πορευομένη τῷ φόβῳ τοῦ κυρίου καὶ τῇ παρακλήσει τοῦ ἁγίου πνεύματος
 ἐπληθύνετο
 (“They multiplied, walking in the fear of the Lord and the comfort of the Holy Spirit,” Acts 9.31)

3.4.7 Relative clauses



οὐ γὰρ ἔασει φάρμακον ἔσθλόν, ὃ τοὶ δώσω
 (“For the potent drug which I will give to you will not permit it,” Od. 10.291-2)

3.4.8 Pronouns

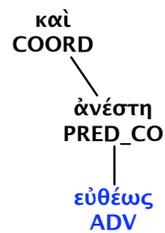


τίς τοι κακός ἔχραε δαίμων
("What evil spirit assailed you?" Od. 10.64)

3.5 ADV (adverbial)

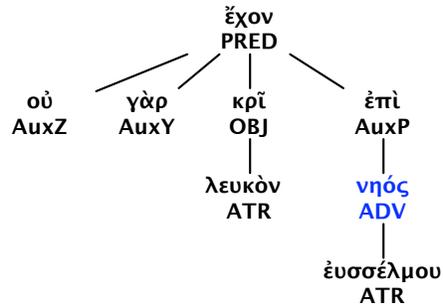
Similarly, adverbials further specify the circumstances under which a verb, adjective or adverb takes place. These include adverbs, prepositional phrases, nouns in oblique cases, participles and subordinate clauses.

3.5.1 Adverbs



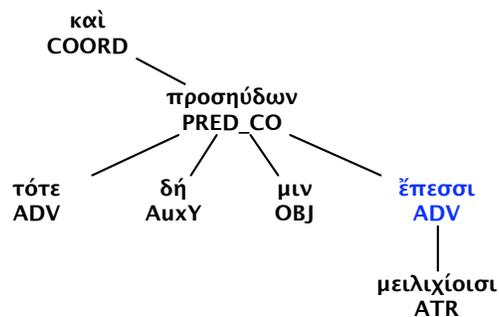
καὶ εὐθέως ἀνέστη
("And he immediately rose," Acts. 9.34)

3.5.2 Prepositional phrases



οὐ γὰρ ἔχον κρῖ λευκὸν ἐυσσέλμου ἐπὶ νηὸς
 (“For they did not have white grain on the well-benched ship,” Od. 12.358)

3.5.3 Nouns

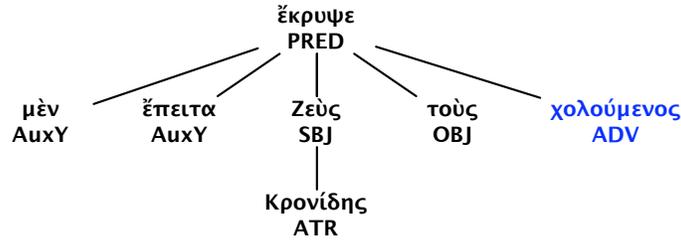


καὶ τότε δὴ μιν ἔπεσσι προσηύδων μελιχίοισι
 (“Then, at that time I addressed him with simple words,” Od. 9.363)

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. Note that even if a participle shares the same case, number and gender as another noun in the sentence (as *χολούμενος* matches *Ζεὺς* in the example below), it should still depend on the verb via ADV unless it further *restricts* the reference of the noun - e.g., it should depend on Zeus only if it means “the Zeus who was angry” (as opposed to some other Zeus who was not). If it describes *how* the verb was completed (e.g., with

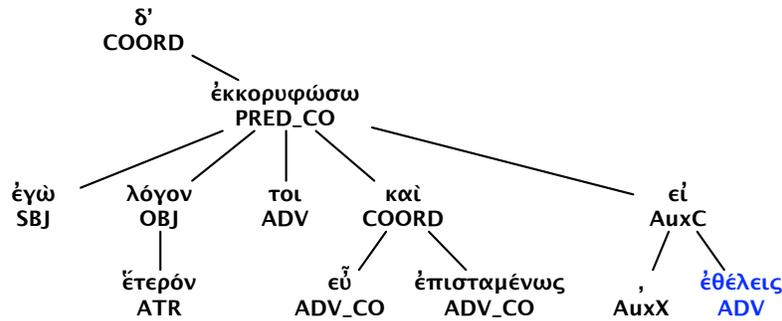
the subject momentarily being angry), as it does here, it should depend on that verb via ADV.



τοὺς μὲν ἔπειτα Ζεὺς Κρονίδης ἔκρυψε χολούμενος
 (“Then Zeus the son of Cronos covered them, being angry,” Op. 137-8)

3.5.5 Subordinate clauses

Subordinate clauses that can be left off the sentence without it becoming ungrammatical generally express optional information about the circumstances surrounding the verb. Clauses beginning with “if” or “because” almost always fall into this category.



εἰ δ' ἐθέλεις , ἕτερόν τοι ἐγὼ λόγον ἐκκορυφώσω εὖ καὶ ἐπισταμένως
 (“If you are willing, I will tell another story for you well and skillfully,” Op. 106-7)

3.6 ATV/AtvV (Non-governed 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 – but not optionally, as ADVs do. The PDT

use of ATV is largely similar to the account of praedicativa given in Pinkster [7], as for example:

- Cicero *consul* coniurationem Catilinae detexit (“Cicero *as consul* uncovered the conspiracy of Catiline”).

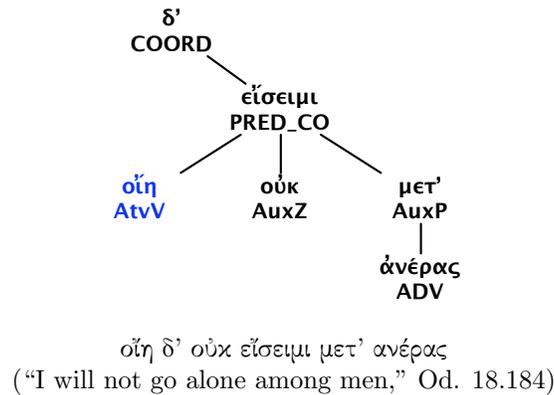
Here *consul* cannot be left out without changing the meaning of the verb, since what is being stressed is the *state* of the actor. Similarly in Greek:

- γείτονες ἄζωστοι ἔχιον (“The neighbors went unarmed,” Op. 345)

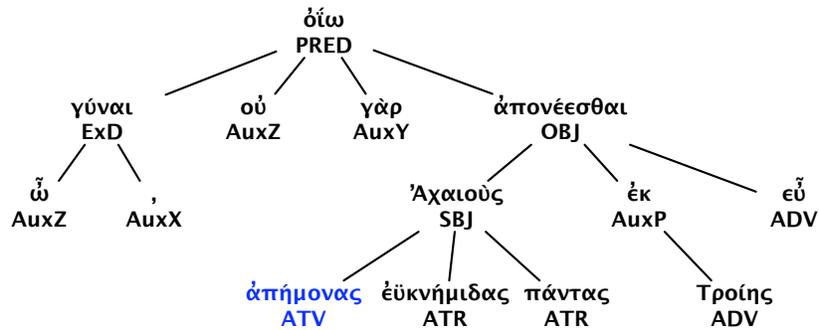
Here, ἄζωστοι cannot be optionally left out without changing the fundamental meaning of the sentence. Similarly, even though it matches the noun γείτονες morphologically, it shouldn’t depend on it via ATR since this analysis leads to the translation “The unarmed neighbors went.” The solution is to have it depend on the noun via ATV.



If the head noun phrase in such constructions is implied, the praedicativum should depend on the main verb via AtvV.



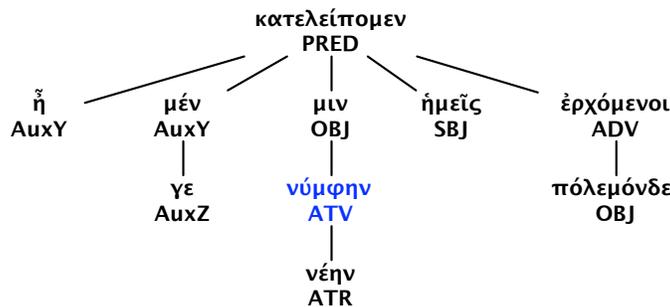
ATV and AtvV are both relatively rare constructions – before using it, be careful to consider that the word should not be annotated as an optional ADV or a restrictive ATR. The verbs you find used with ATV/AtvV are typically confined (as Pinkster notes) to a limited number of groups, mainly verbs involving motion (as the example above and this below) and several that behave like copulas.



ὦ γύναι, οὐ γὰρ οἶω εὐκνήμιδας Ἀχαιοὺς ἐκ Τροίης εὖ πάντας ἀπήμονας ἀπονέεσθαι

(“O wife, I do not foresee that all the well-greaved Achaeans will well return safe from Troy,” Od. 18.259-60)

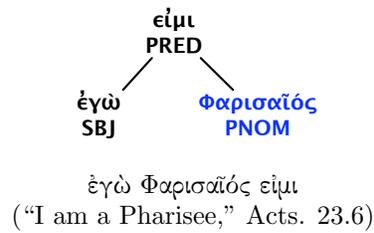
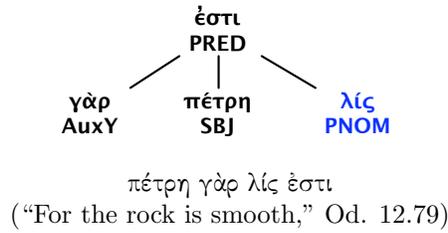
Most phrases that involve x doing something *as y* should be annotated with ATV. In the example below, the direct object (μιν) has been left *as a young bride*.



ἦ μὲν μιν νύμφη γε νέην κατελείπομεν ἡμεῖς ἐρχόμενοι πόλεμόνδε
 (“Indeed, we left her as a young bride as we went off to war,” Od. 11.447)

3.7 PNOM (predicate nominal)

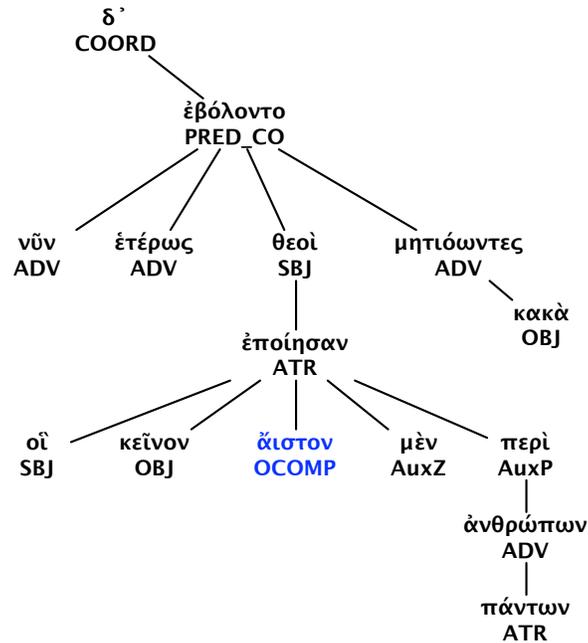
Predicate nominals (subject complements) depend on a verbal head.



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 εἰμι.

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 ποιέω and τίθημι.



νῦν δ' ἑτέρως ἐβόλοντο θεοὶ κακὰ μητιόωντες, οἷ κείνων μὲν ἄιστον ἐποίησαν περὶ πάντων ἀνθρώπων
 (“But now the gods, who had made that man unseen among all men, wish otherwise, planning evil things,” Od. 234-5)

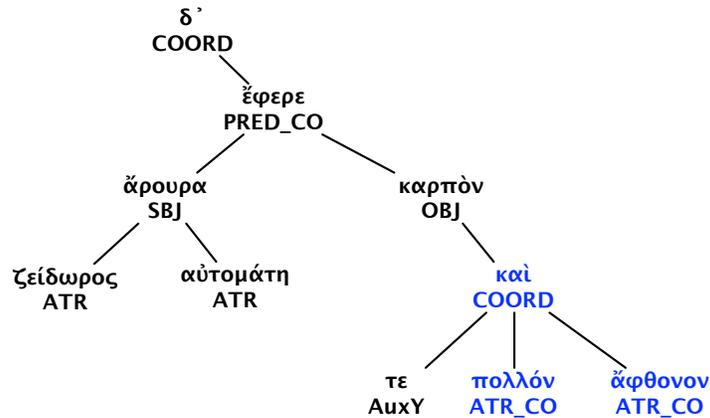
3.9 “Bridge” structures

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 νηῦς ἔστηκεν ἐπ’ ἀγροῦ (“the ship lies beside the field”), for example, the noun ἀγροῦ (“field”) depends on the preposition, but with the relation it would bear to ἔστηκεν (“lies”) - namely, ADV. The preposition is assigned a sort of “dummy” relation AuxP, meant to signify that the true relationship is that between ἀγροῦ and ἔστηκεν, and that the preposition simply acts a mediator between the two.

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

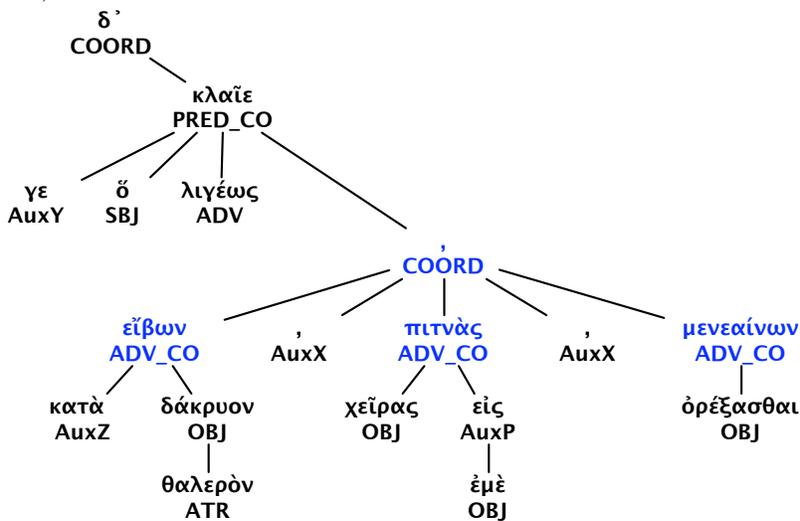
3.9.1 COORD (coordinator)

An example of a coordinated structure is given below.



καρπὸν δ' ἔφερε ζείδωρος ἄρουρα αὐτομάτη πολλὸν τε καὶ ἄφθονον
 ("The fruitful earth unforced bare them many and abundant fruit," Op.
 117-18.)

Here, πολλὸν and ἄφθονον both depend on the single final coordinator that separates them: καὶ. Each of these words depends on that καὶ with a complex tag comprised in part of the relation they bear to the coordinator's head: since each would individually modify καρπὸν as an ATR, each depends on the coordinator with the tag ATR_CO. The coordinator then depends on καρπὸν with the tag COORD. If several coordinating conjunctions are present (e.g. τε here) – and this is especially common in list contexts with three or more coordinated elements – only the *final* conjunction is the head; all others depend on it (via AuxY for actual words like τε and via AuxX for coordinating punctuation like commas).



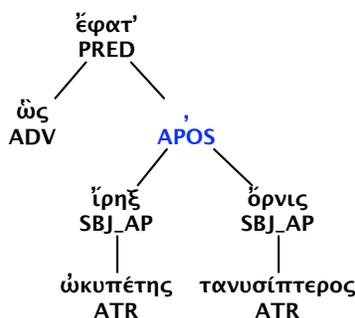
κλαίει δ' ὁ γε λιγέως, θαλερὸν κατὰ ἄκρουον εἵβων, πιτνάς εἰς ἐμὲ χεῖρας ,
 ὀρέξασθαι μενεαίνων
 (“He wept aloud, shedding big tears, stretching his arms toward me, wanting
 to reach me,” Od. 11.391-2)

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).

And even if words are coordinated on different levels, each should only have one `_CO` suffix.

3.9.2 APOS (apposing elements)

An example of apposition is given below.

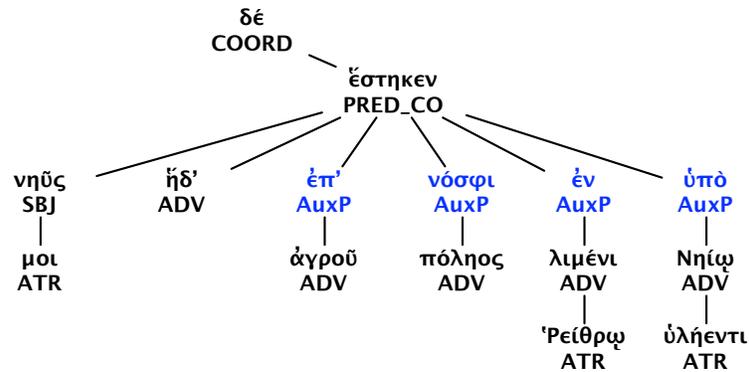


ὡς ἔφατ' ὠκυπέτης ἵρηξ, πανυσιπτερος ὄρνις
 (“So said the swift-flying hawk, the long-winged bird,” Op. 212)

Here the two phrases in apposition to each other are ἵρηξ (“hawk”) and ὄρνις (“bird”); the appositional-coordinating element that separates them is the comma between them. Both ἵρηξ and ὄρνις depend on the apposing comma via the relation they each individually bear to the phrase’s head (ἔφατ’). Since they are both the subjects of ἔφατ’, they modify the apposing word via the complex tag SBJ_AP; the apposing word then modifies ἔφατ’ via APOS.

3.9.3 AuxP (preposition)

An example of several prepositional phrases is given below.

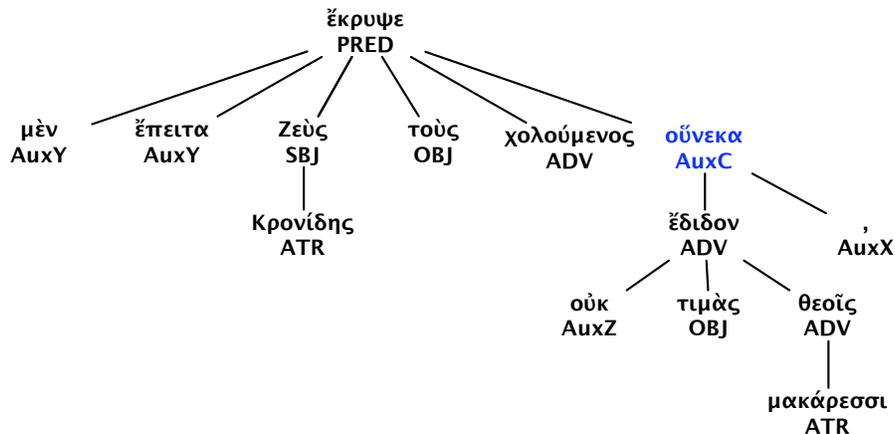


νηῦς δέ μοι ἦδ' ἔστηκεν ἐπ' ἀγροῦ νόσφι πόλης, ἐν λιμένι Ῥεῖθρων ὑπὸ Νηίῳ ὕληεντι
 (“My ship lies there beside the field, away from the city, in the harbor of Rheithron, under woody Neion,” Od. 1.185-6)

Our method of annotation sees prepositions as acting as a functional bridge between their child and head. Here the object of each preposition (ἀγροῦ, ἴπληος, λιμένι and Νηίῳ) would depend on its preposition via the relationship it would hold to the preposition’s head (here, ADV for each). 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.



τοὺς μὲν ἔπειτα Ζεὺς Κρονίδης ἔκρυψε χολούμενος , οὖνεκα τιμὰς οὐκ ἔδιδον
μακάρεσσι ἔχουσιν
 (“Then Zeus the son of Cronos covered them, being angry, because they did
 not give honor to the blessed gods” Op. 137-9)

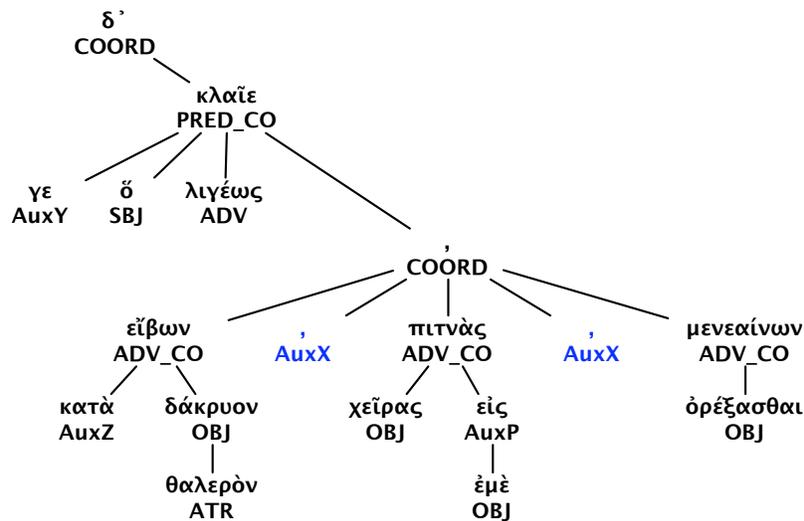
Here the subordinate verb ἔδιδον (“give”) depends on its head (οὖνεκα, “because”) via the relationship ADV. οὖνεκα then depends on ἔκρυψε (“covered”) via the “bridge” relationship AuxC.

3.10 Punctuation

Our methods of annotating punctuation follow that established by the PDT, which assigns several different functional tags.

3.10.1 AuxX (commas)

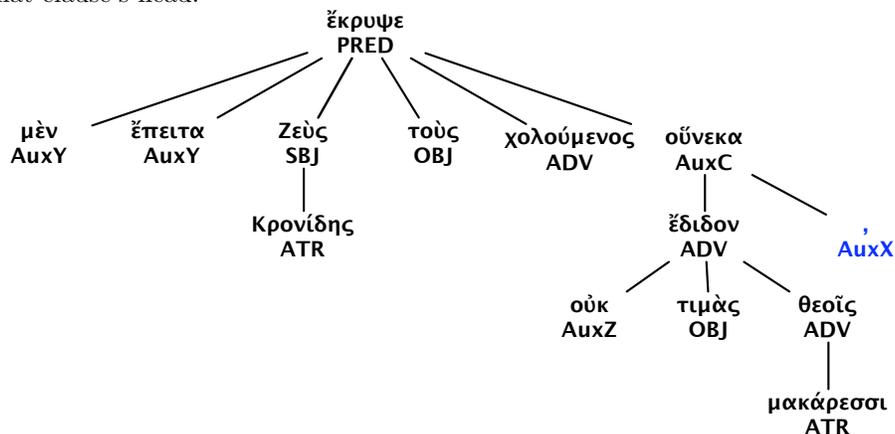
If a comma is not the head of a coordinated or appositional phrase, it should be annotated with AuxX and depend on the head of its clause. In coordinated lists, this head is the final comma.



κλαῖε δ' ὃ γε λιγέως, θαλερόν κατὰ ἄκρουν εἴβων, πιτνάς εἰς ἐμὲ χεῖρας ,
 ὀρέξασθαι μενεαίνων
 (“He wept aloud, shedding big tears, stretching his arms toward me, wanting
 to reach me,” Od. 11.391-2)

Here, εἶβων, πιτνάς, and μενεαίνων all depend on the final comma (that separating χεῖρας from ὀρέξασθαι) via ADV_CO. The remaining commas should then depend on the final comma via AuxX. Note that if the non-head coordinator is a content word (e.g., τε), it depends on the final coordinator via AuxY; if it is punctuation, it depends via AuxX.

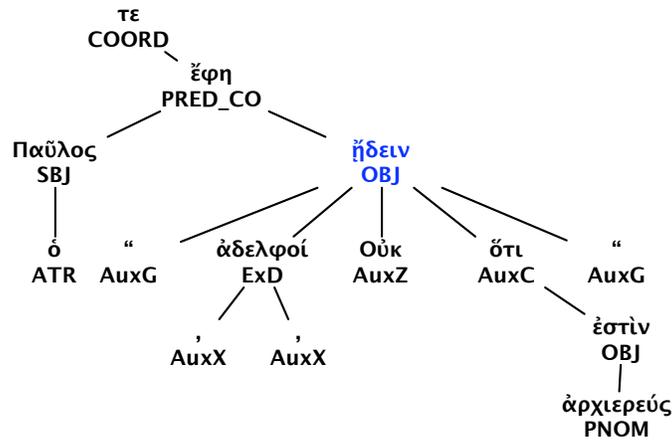
If a comma is used to separate a subordinate clause, it should depend on that clause’s head.



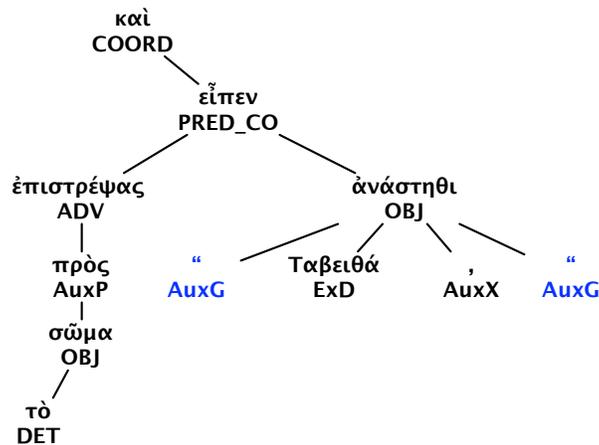
τοὺς μὲν ἔπειτα Ζεὺς Κρονίδης ἔκρυψε χολούμενος, οὖνεκα τιμὰς οὐκ ἔδιδον
μακάρεσσι ἔχουσιν
 (“Then Zeus the son of Cronos covered them, being angry, because they did
 not give honor to the blessed gods” Op. 137-9)

3.10.2 AuxG (bracketing punctuation)

“Bracketing” punctuation surrounds an enclosed phrase, and most frequently appears as quotation marks or parentheses (not commas, which are annotated in such situations with AuxX). These punctuation marks should depend on the head of the bracketed phrase via AuxG.



ἔφη τε ὁ Παῦλος “ Οὐκ ᾔδειν , ἀδελφοί , ὅτι ἐστὶν ἀρχιερεὺς ”
 (“Paul said, ‘I didn’t know, brothers, that he was high priest,’” Acts 23.5)

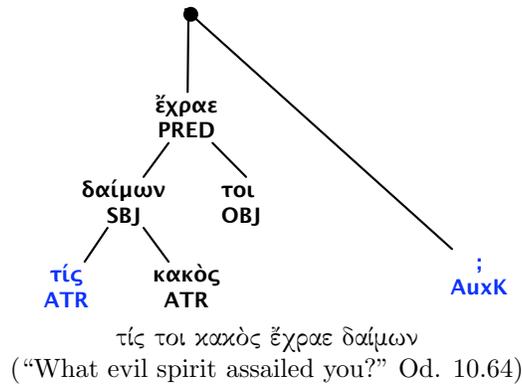


καὶ ἐπιστρέψας πρὸς τὸ σῶμα εἶπεν ἡ Ταβειθά ἀνάστηθι ἡ
 (“Turning toward the body, he said ‘Tabitha, get up,’” Acts 9.40)

AuxG should also be used to annotate abbreviation, with the period depending on the abbreviated word.

3.10.3 AuxK (terminal punctuation)

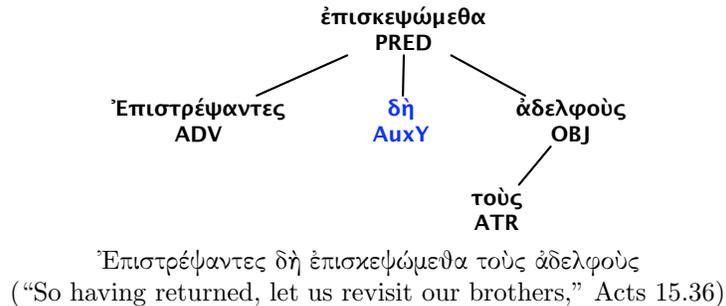
Final punctuation (if present) should depend on the root via AuxK.



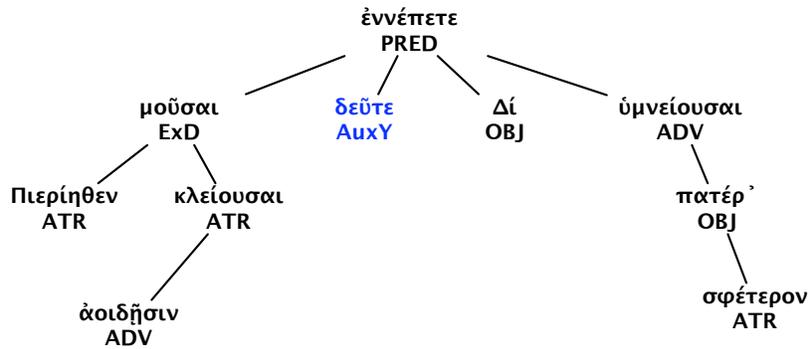
3.11 AuxY (sentence adverbials)

Sentence adverbials (also called disjuncts) are those that pertain to the entire sentence and are used to connect the sentence to the discourse at large (especially to what immediately precedes it). This includes words like “therefore,” “however,” or “regardless” – adverbials that don’t qualify the circumstance of the main verb of the sentence (like ADVs usually do), but rather situate the entire sentence against a context that exists *outside* of it. Similarly, these also often express the author’s opinion (from the perspective of someplace outside the action of the sentence) about the validity of what’s being said (e.g., “truly,” “certainly”).

In Greek this includes words like γάρ, μέν, δέ, δή, ἄρα, ὦ, ἄν, κέ, ἄρα, οὔν, γοῦν, ἦ, ναί, νή, νῦν and ἔτι – though note that several of these words often have other functions (e.g., νῦν as ADV, δέ as a coordinator) – what’s important is to see how they’re used in context.

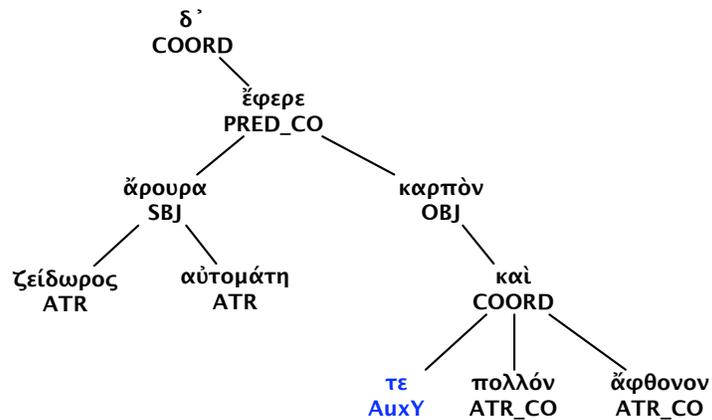


This also includes exclamations.



μούσai Πιερίθην ἀοιδῆσιν κλείουσai δεῦτε, Δί' ἐννέπετε, σφέτερον πατέρ'
 ὑμνεῖουσai
 (“Muses of Pieria giving glory by song, here!, tell of Zeus, singing of your
 father,” Op. 1.1)

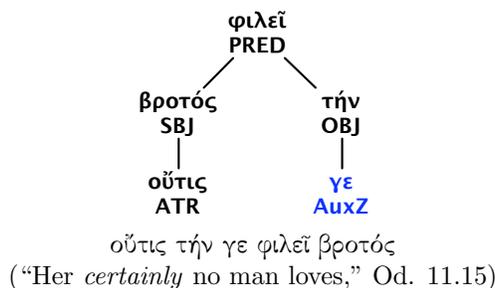
AuxY should also be used to annotate coordinators that are not commas (e.g., τε, καί) when they are not the head of the coordinated phrase. (Non-head commas in these structures should be annotated with AuxX.)



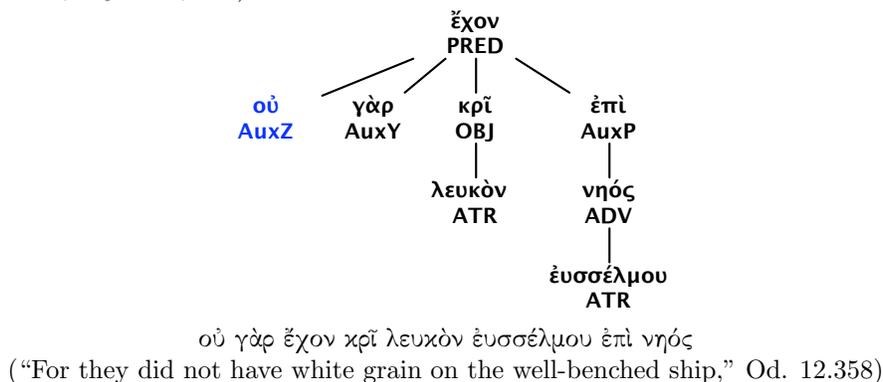
καρπὸν δ' ἔφερε ζείδωρος ἄρουρα αὐτομάτη πολλὸν τε καὶ ἄφθονον
 (“The fruitful earth unforced bare them many and abundant fruit,” Op.
 117-18.)

3.12 AuxZ (emphasizing particles)

AuxZ should be assigned to particles with a relatively poor meaning content that emphasize one specific word in the sentence (as distinct from AuxY, which emphasizes the sentence as a whole). This group consists largely of negators (e.g., οὐ or μή) but also includes words such as γε, ὦ, δὴ, πέρ and τοι.



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

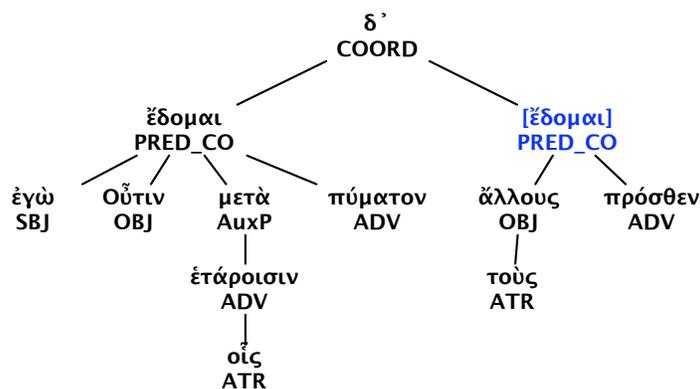


4 How to Annotate Specific Constructions

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. Con-

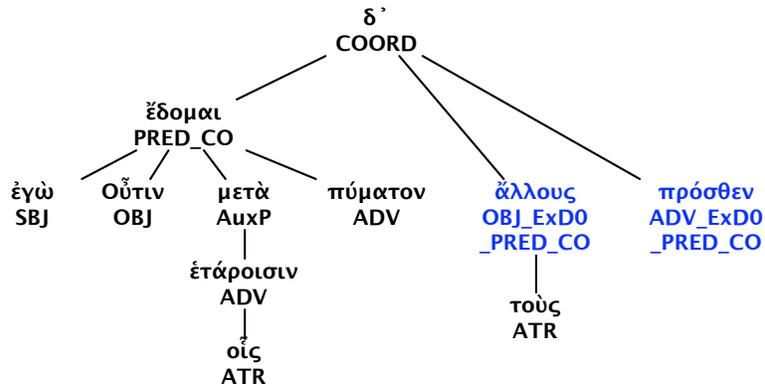
sider the example from Odyssey 9.369-70 given in the figure below.



Οὐτιν ἔγῳ πύματον ἔδομαι μετὰ οἷς ἐτάροισιν , τοὺς δ' ἄλλους πρόσθεν
 (“I will eat Nobody last among his friends, the others beforehand,” Od.
 9.369-70). Abstract structure with elided words reconstructed.

Here, the verb ἔδομαι is missing from the second clause.¹ We can preserve the structure of the tree by assigning the head of ἄλλους and πρόσθεν to be the head that the second ἔδομαι would have if it were in the sentence (δ'), and by assigning tags to each that preserve the path: ἄλλους should be the object (OBJ) of ἔδομαι, which should then depend on δ' via by 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, πρόσθεν should be an adverbial (ADV) dependent on the elided word; it therefore receives the tag ADV_ExD0_PRED_CO. This produces the tree given below (which should be used as a model for annotation). This method allows us to use the complex tags to reconstruct the tree as necessary.

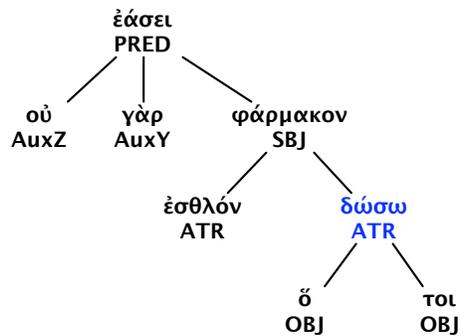
¹Note that we know that the second ἔδομαι is elided because the presence of the two adverbs πύματον and πρόσθεν prevents us from coordinating οὐτιν and ἄλλους on the single ἔδομαι. The two adverbs both must modify the verb, and if all four words depend on the same instance of ἔδομαι, we would not be able to distinguish structurally which person is to be eaten “last” and which “before.”



Οὐτιν ἐγὼ πύματον ἔδομαι μετὰ οἷς ἐτάροισιν , τοὺς δ' ἄλλους πρόσθεν
 (“I will eat Nobody last among his friends, the others beforehand,” Od.
 9.369-70). Tree as annotated with ellipsis.

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 example, are generally attributive, and should modify the antecedent via ATR (this is because the relative clause provides more information to restrict the reference, just like an adjective does – not just “any drug” in the example below, but “the drug that I give you”). The head of a relative clause is the subordinate verb; this is the element that depends on the antecedent.

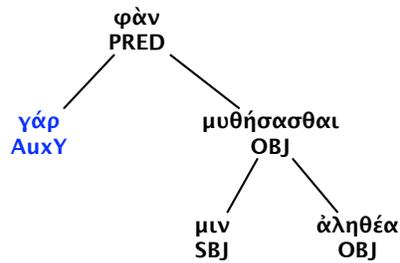


οὐ γὰρ ἐάσει φάρμακον ἐσθλόν, ὃ τοι δώσω
 (“For the potent drug which I will give to you will not permit it,” Od.
 10.291-2)

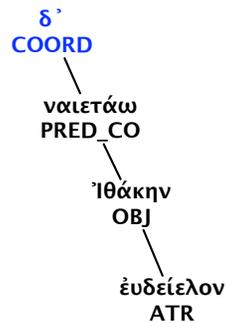
subordinate clauses, in which the subordinate verb depends on the subordinating conjunction, which then depends on a word outside of the clause. See section 3.9.4 (AuxC) for information on annotating subordinate clauses.

4.3 Particles

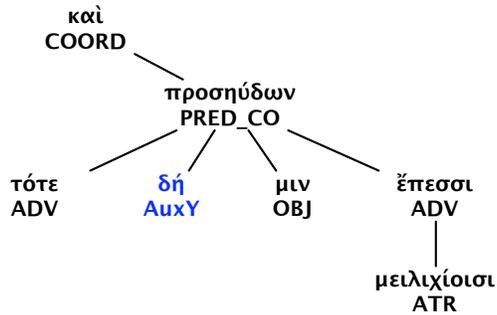
A particle is a morphological category for uninflected function words (like μέν, γάρ and δέ). Just like most other morphological categories (such as nouns and verbs), particles can be annotated in several different ways depending on how they're used in context. Most particles, however, tend to modify either the sentence as a whole (AuxY) or one word in particular (AuxZ).



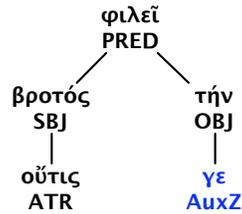
φάν γάρ μιν ἀληθέα μυθήσασθαι
 (“For they thought that he spoke the truth,” Od. 18.342)



ναιετάω δ' Ἰθάκην ἐυδείελον
 (“I inhabit clear-seen Ithaca,” Od. 9.24)



καὶ τότε δὴ μιν ἔπεσσι προσηύδων μελιχίοισι
 (“Then, at that time I addressed him with simple words,” Od. 9.363)



οὐτις τὴν γε φιλεῖ βροτός
 (“Her *certainly* no man loves,” Od. 11.15)

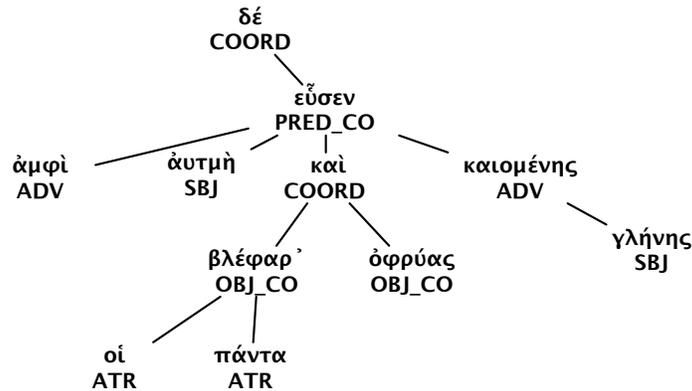
4.4 The Genitive and Accusative Absolute

The genitive and accusative absolute are grammatical constructions similar to the English nominative absolute, where a noun and a participle form a phrase that is disjoint from the grammar of the rest of the sentence; in Greek both the noun and participle are inflected either the genitive or accusative case, as in the following:

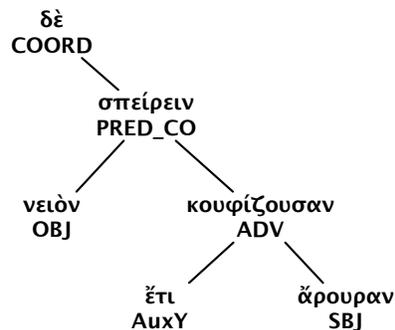
- Genitive: πάντα δὲ οἱ βλέφαρ’ ἀμφὶ καὶ ὀφρύας εὔσεν ἀντιμὴ γλήνης καιομένης
 (“And the flame singed his eyelids and brow all around, his eyeball burning,” Od. 9.389-90)
- Accusative: νειὸν δὲ σπείρειν ἔτι κουφίζουσαν ἄρουραν
 (“Sow fallow-land with the soil still getting light,” Op. 463)

Following Pinkster [7], we treat Greek genitive and accusative absolutes like their Latin ablative cousins: 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 phrase) depending on the main verb as an adverbial. We would annotate the example above in the following way:



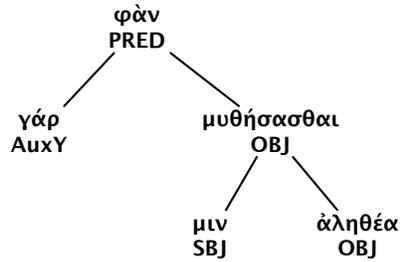
πάντα δέ οἱ βλέφαρ' ἄμφι καὶ ὄφρῶας εὔσεν ἀυτμῆ γλήνης καιομένης ...
 (“And the flame singed his eyelids and brow all around, his eyeball burning,”
 Od. 9.389-90)



νεῖον δὲ σπείρειν ἔτι κουφίζουσαν ἄρουραν
 (“Sow fallow-land with the soil still getting light,” Op. 463)

4.5 Accusative + Infinitive

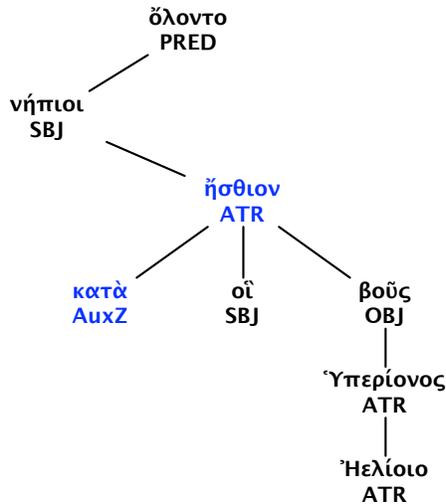
In indirect discourse and other accusative + infinitive constructions, the infinitive verb is the head of its phrase. This verb represents the entire clause and should depend via OBJ on the word that introduces the discourse. Within the phrase, standard annotation applies (so that the subject, while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



φάν γάρ μιν ἀληθέα μυθήσασθαι
 (“For they thought that he spoke the truth,” Od. 18.342)

4.6 Tmesis

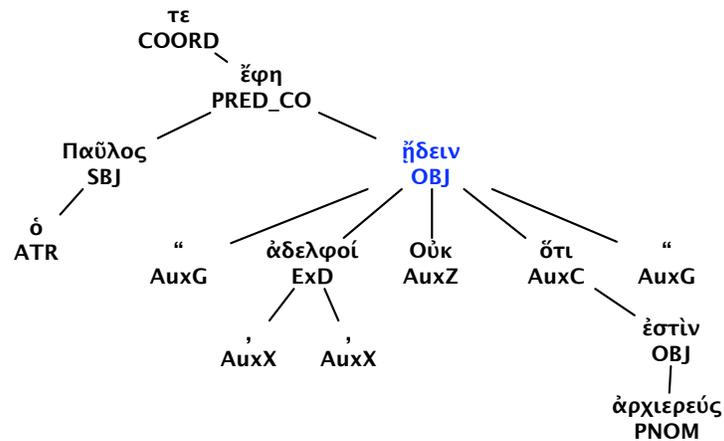
Tmesis is the separation of a verb into two parts, usually involving a prefix that can also function as a preposition or adverb on its own. In the example below, the verb ἤσθιον is not derived from the simple verb ἐσθίω (“to eat”) but rather from the complex verb κατεσθίω (“to devour”). In these cases, the separated prefix should depend on the verb via AuxZ.



ὄλοντο νήπιοι οἱ κατὰ βοῦς Ἰπερίονος Ἡελίοιο ἤσθιον
 (“The fools who devoured the bulls of Helios Hyperion perished,” Od. 1.7-8)

4.7 Direct Speech

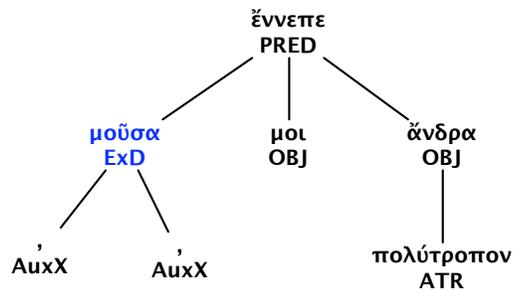
Direct speech should be annotated the same way as indirect discourse, by attaching the head of the “spoken” phrase to the predicate that introduces the speaking.



εφη τε ο Παυλος “ Ουκ ηδεν , αδελφοι , οτι εστιν αρχιερευς ”
 (“Paul said, ‘I didn’t know, brothers, that he was high priest,’” Acts 23.5)

4.8 Direct Address

As in the PDT, vocatives should depend on their verbal heads via ExD.



ανδρα μοι ενηνεπε, μοισα, πολυτροπον
 (“Tell me, muse, of the man of many ways,” Od. 1.1)

References

- [1] David Bamman, Marco Passarotti, Gregory Crane, and Savina Raynaud. Guidelines for the syntactic annotation of Latin treebanks. Technical report, Tufts Digital Library, Medford, 2007.
- [2] U. Demske, N. Frank, S. Laufer, and H. Stiemer. Syntactic interpretation of an Early New High German corpus. In *Proceedings of the Third Workshop on Treebanks and Linguistic Theories*, pages 175–182, 2004.
- [3] Eva Hajičová, Zdeněk Kirschner, and Petr Sgall. A Manual for Analytic Layer Annotation of the Prague Dependency Treebank (English translation). Technical report, ÚFAL MFF UK, Prague, Czech Republic, 1999.
- [4] Jan Hajič. Building a syntactically annotated corpus: The Prague Dependency Treebank. In Eva Hajičová, editor, *Issues of Valency and Meaning. Studies in Honor of Jarmila Panevová*, pages 12–19. Prague Karolinum, Charles University Press, 1998.
- [5] A. Kroch, B. Santorini, and L. Delfs. Penn-Helsinki Parsed Corpus of Early Modern English. <http://www.ling.upenn.edu/hist-corpora/ppceme-release-1>, 2004.
- [6] A. Kroch and A. Taylor. Penn-Helsinki Parsed Corpus of Middle English, second edition. <http://www.ling.upenn.edu/hist-corpora/ppcme2-release-2/>, 2000.
- [7] Harm Pinkster. *Latin Syntax and Semantics*. Routledge, London, 1990.
- [8] Vitor Rocio, Mário Amado Alves, J. Gabriel Lopes, Maria Francisca Xavier, and Graça Vicente. Automated creation of a Medieval Portuguese partial treebank. In Anne Abeillé, editor, *Treebanks: Building and Using Parsed Corpora*, pages 211–227. Kluwer Academic Publishers, 2003.
- [9] Petr Sgall, Eva Hajičová, and Jarmila Panevová. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Dordrecht: Reidel Publishing Company and Prague: Academia, 1986.
- [10] Ann Taylor, Anthony Warner, Susan Pintzuk, and Frank Beths. York-Toronto-Helsinki Parsed Corpus of Old English Prose, 2003.