Greek Dependency Syntax Labels in Perseids

SunoikisisDC 20 Feb 2020

Vanessa B. Gorman

Professor of Ancient History

University of Nebraska-Lincoln

vgorman1@unl.edu

Dependency Syntax

- A system of showing **relationships** between words in a sentence.

- Simpler than constituency grammar [Chomsky]

- It is used internationally on various languages
  - Suitable for comparison between languages
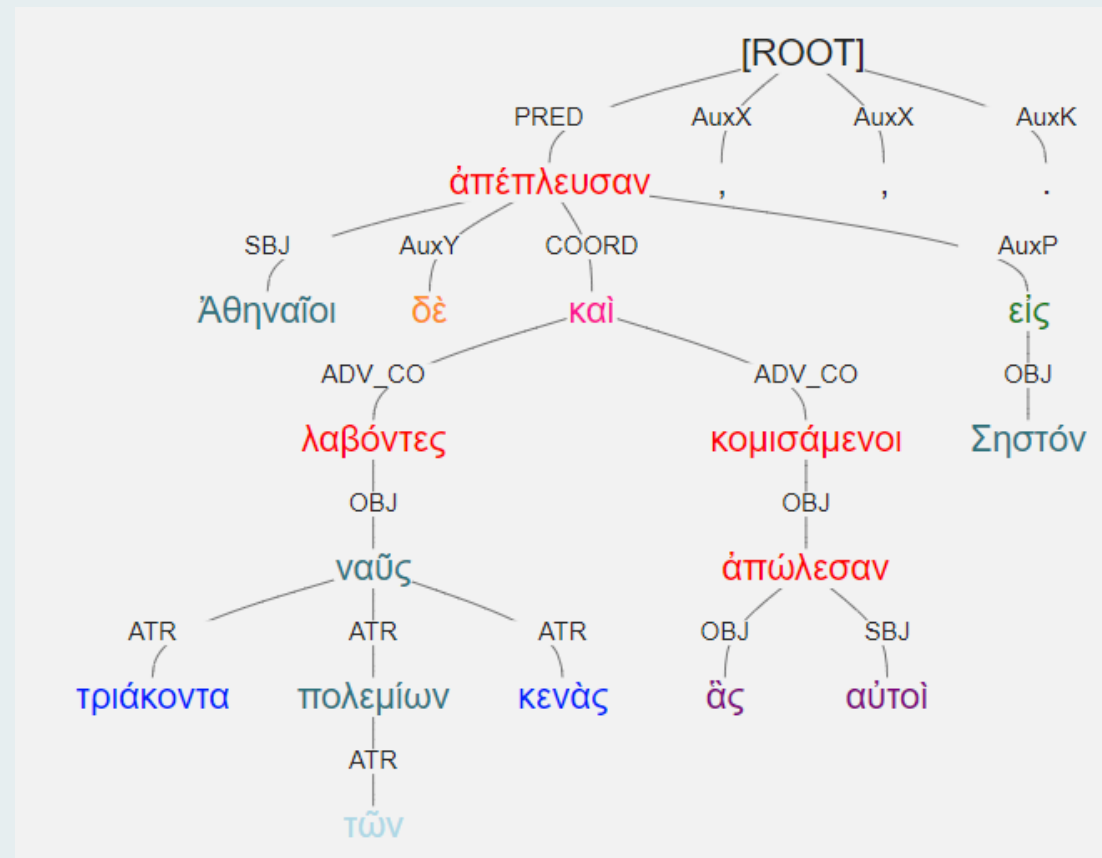 - Works well for sentences in which the word order is not linear

A single word anchors the entire sentence:
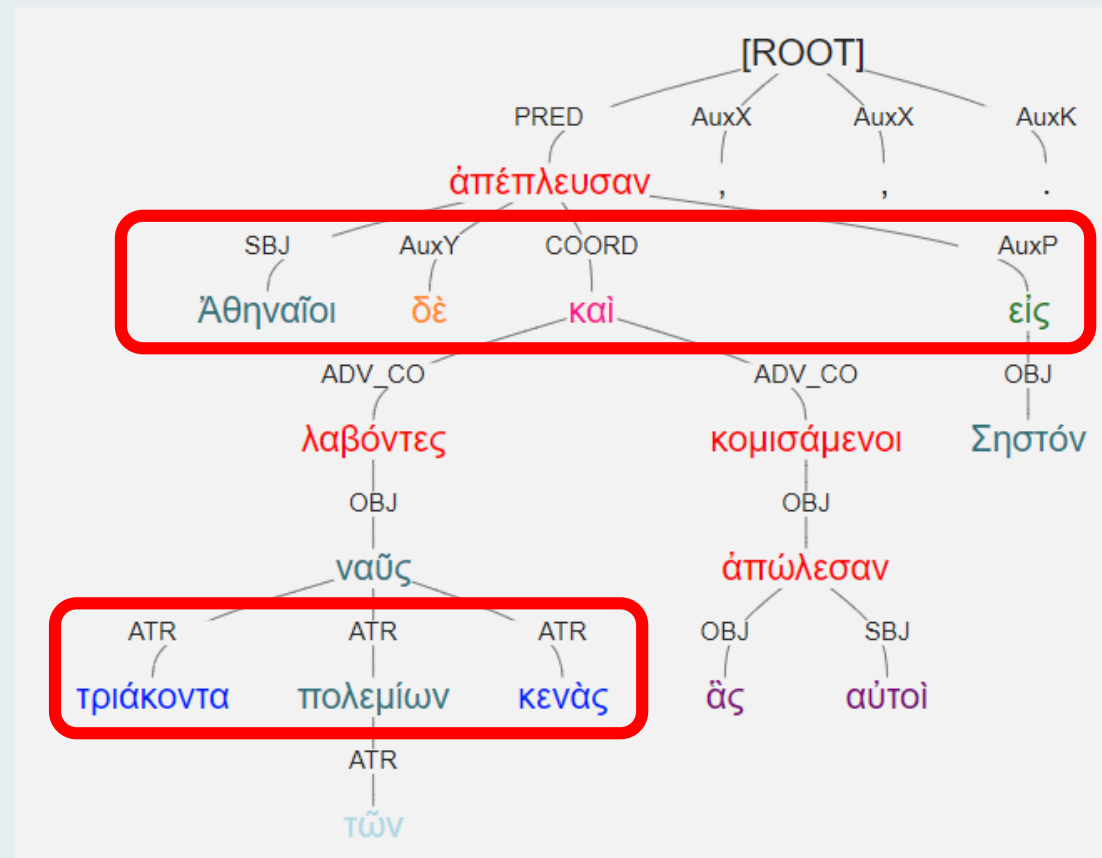## a <u>main verb</u> or a <u>coordinating conjunction</u>
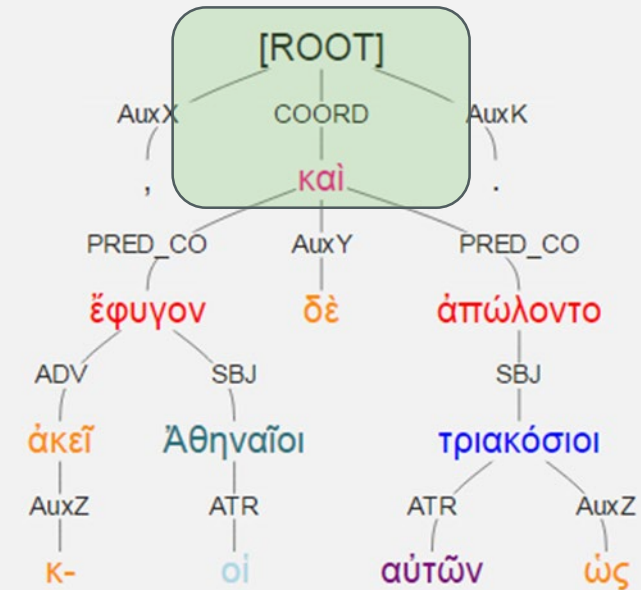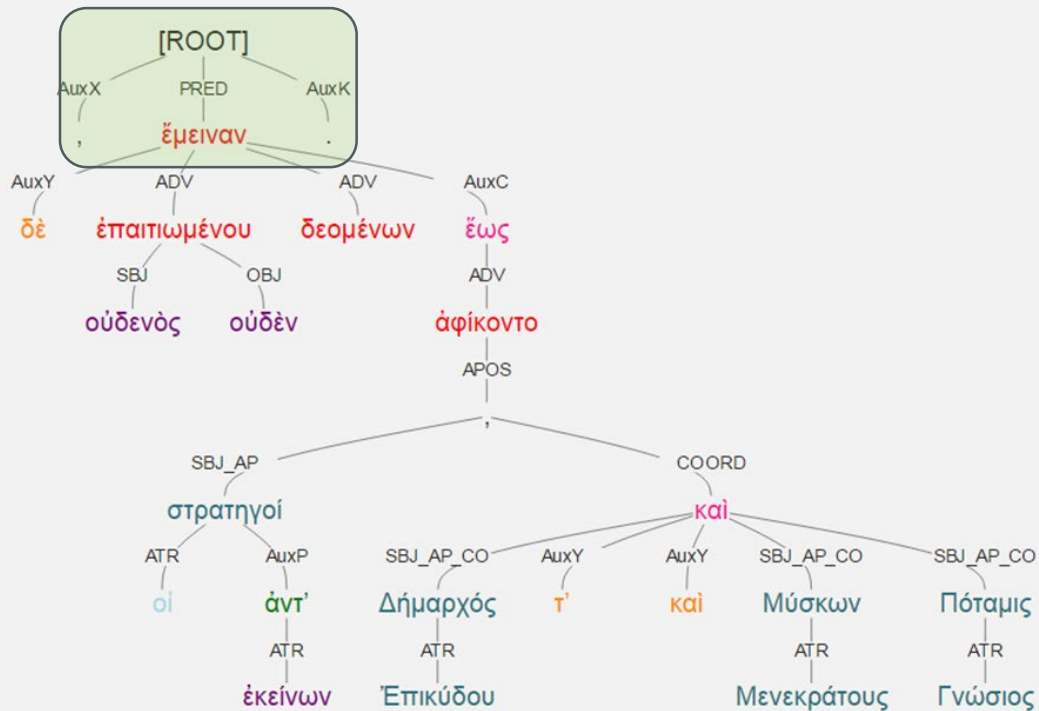
Each word in the tree is located within the syntax of the whole sentence
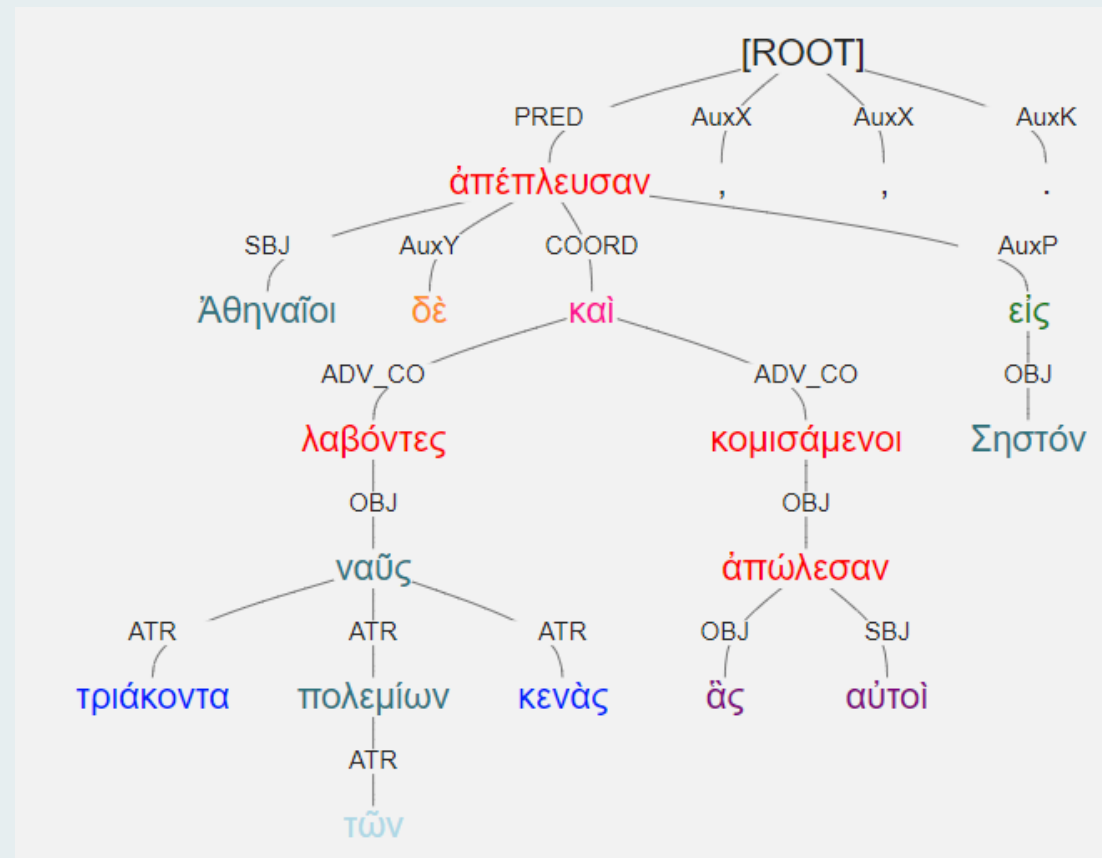
Each word is hung on a single "parent" word

Words may have none, one, or several "children"

The <u>relationship</u> [dependency] between the "child" and the "parent" is labeled

PRED	Predicative
SBJ	Subject
OBJ	Object
ATR	Attributive (Adjectival)
ADV	Adverbial
AuxP	Preposition
AuxC	Subordinating conjunction
AuxY	Sentence adverb
AuxZ	Emphasizing particle
AuxV	Auxiliary verb
AuxR	Reflexive passive
AuxG	Bracketing punctuation
AuxX	Commas
AuxK	Terminal punctuation
COORD	Coordinating conjunction
ATV/AtvV	Verbal attributive
PNOM	Predicative nominative
OCOMP	Object complement
APOS	Apposing element [usually a comma]
ExD	External dependency

The Prague Tagset [1.1]

Only 20 possible relationship tags

PRED	Predicative
SBJ	Subject
OBJ	Object
ATR	Attributive (Adjectival)
ADV	Adverbial
AuxP	Preposition
AuxC	Subordinating conjunction
AuxY	Sentence adverb
AuxZ	Emphasizing particle
AuxV	Auxiliary verb
AuxR	Reflexive passive
AuxG	Bracketing punctuation
AuxX	Commas
AuxK	Terminal punctuation
COORD	Coordinating conjunction
ATV/AtvV	Verbal attributive
PNOM	Predicative nominative
OCOMP	Object complement
APOS	Apposing element [usually a comma]
ExD	External dependency

> Three are used for plain punctuation

PRED	Predicative
SBJ	Subject
OBJ	Object
ATR	Attributive (Adjectival)
ADV	Adverbial
AuxP	Preposition
AuxC	Subordinating conjunction
AuxY	Sentence adverb
AuxZ	Emphasizing particle
AuxV	Auxiliary verb
AuxR	Reflexive passive
AuxG	Bracketing punctuation
AuxX	Commas
AuxK	Terminal punctuation
COORD	Coordinating conjunction
ATV/AtvV	Verbal attributive
PNOM	Predicative nominative
OCOMP	Object complement
APOS	Apposing element [usually a comma]
ExD	External dependency

> Three are used for plain punctuation

>Two rarely or never occur in Greek

PRED	Predicative
SBJ	Subject
OBJ	Object
ATR	Attributive (Adjectival)
ADV	Adverbial
AuxP	Preposition
AuxC	Subordinating conjunction
AuxY	Sentence adverb
AuxZ	Emphasizing particle
AuxV	Auxiliary verb
AuxR	Reflexive passive
AuxG	Bracketing punctuation
AuxX	Commas
AuxK	Terminal punctuation
COORD	Coordinating conjunction
ATV/AtvV	Verbal attributive
PNOM	Predicative nominative
OCOMP	Object complement
APOS	Apposing element [usually a comma]
ExD	External dependency

> Three are used for plain punctuation

>Two rarely or never occur in Greek

> **Two are always labeled the same way = "functional bridges"**

PRED	Predicative
SBJ	Subject
OBJ	Object
ATR	Attributive (Adjectival)
ADV	Adverbial
AuxP	Preposition
AuxC	Subordinating conjunction
AuxY	Sentence adverb
AuxZ	Emphasizing particle
AuxV	Auxiliary verb
AuxR	Reflexive passive
AuxG	Bracketing punctuation
AuxX	Commas
AuxK	Terminal punctuation
COORD	Coordinating conjunction
ATV/AtvV	Verbal attributive
PNOM	Predicative nominative
OCOMP	Object complement
APOS	Apposing element [usually a comma]
ExD	External dependency

> Three are used for plain punctuation

>Two rarely or never occur in Greek

> Two are always labeled the same way = "functional bridges"

>Two are used to indicate coordination or apposition

PRED	Predicative
SBJ	Subject
OBJ	Object
ATR	Attributive (Adjectival)
ADV	Adverbial
AuxP	Preposition
AuxC	Subordinating conjunction
AuxY	Sentence adverb
AuxZ	Emphasizing particle
AuxV	Auxiliary verb
AuxR	Reflexive passive
AuxG	Bracketing punctuation
AuxX	Commas
AuxK	Terminal punctuation
COORD	Coordinating conjunction
ATV/AtvV	Verbal attributive
PNOM	Predicative nominative
OCOMP	Object complement
APOS	Apposing element [usually a comma]
ExD	External dependency

> Three are used for plain punctuation

> Two rarely or never occur in Greek

> Two are always labeled the same way = "functional bridges"

> Two are used to indicate coordination or apposition

> **Eleven real choices**

Predicative (PRED) = Main verb of the sentence
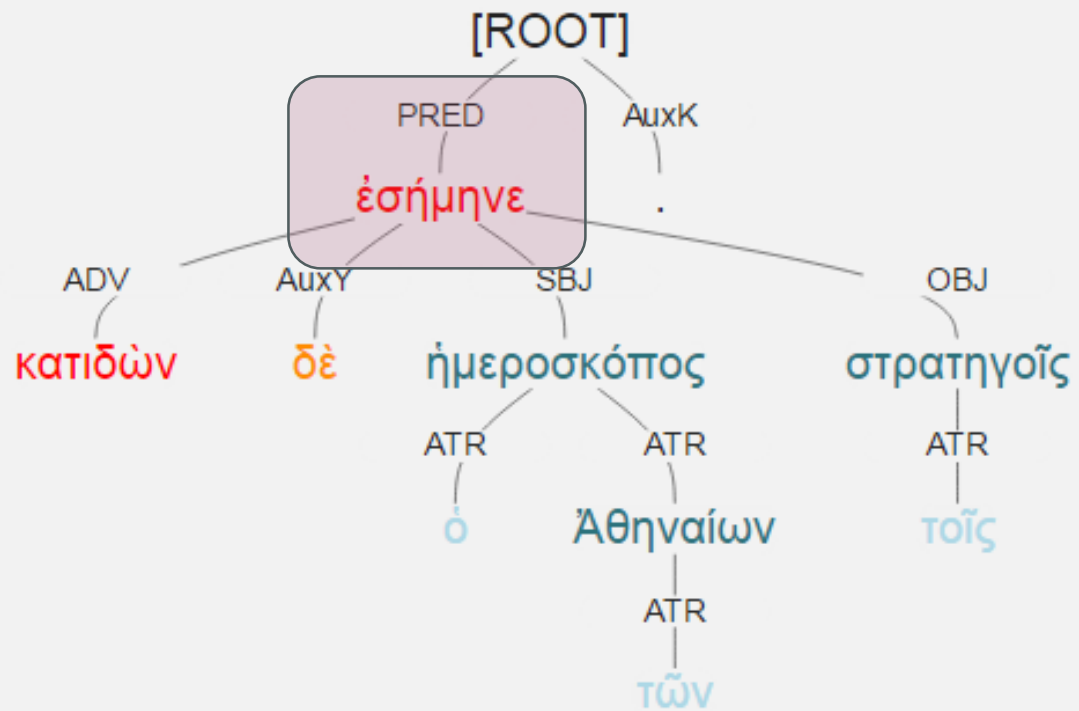
"The watcher was signaling the generals…"

Dependencies of verbs

- Dependency grammar is based on idea of verbal **valency**

- Valency = number of elements "required" by given verb

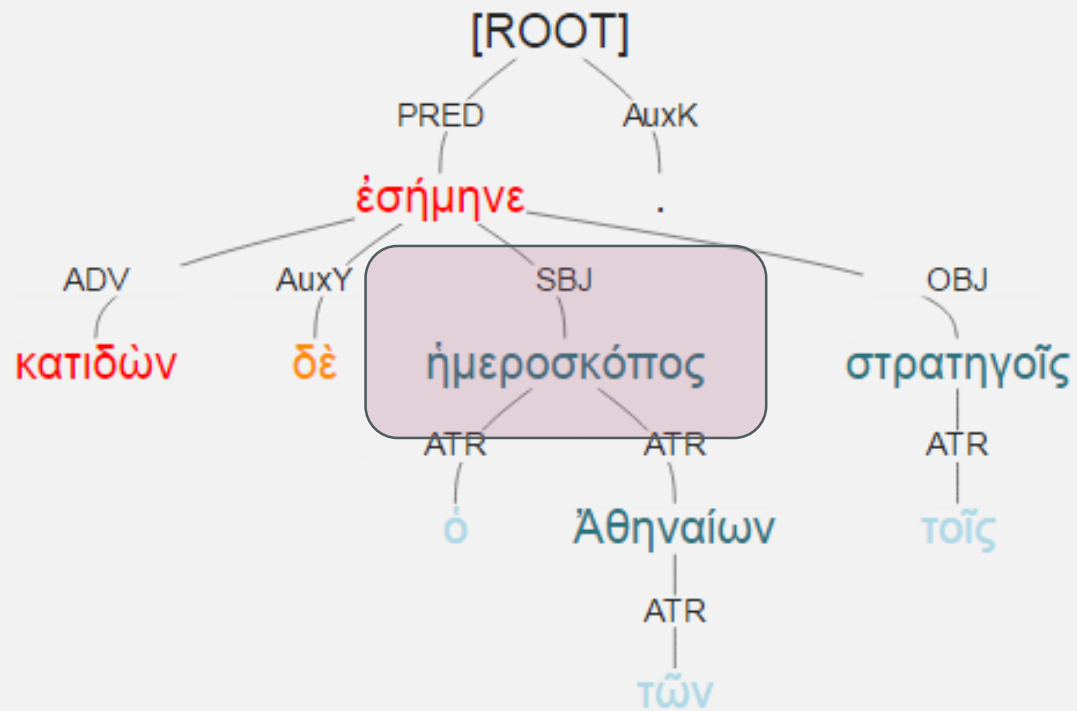- Required elements are called **arguments** of the verb

Valency: Verbs commonly "require" either one, two, or three arguments (expressed or assumed).

- The lone argument of a mono-valent verb or the "first" argument of a bi- or tri-valent verb is labeled as the **subject** (**SBJ**)

- The King died.

- The army was victorious.

Subject (SBJ) = Subject of the sentence, clause, infinitive, or participle

"The watcher was signaling the generals…"

Additional arguments of bi- and tri-valent verbs are called one of three things

- Direct or indirect object [OBJ]
  - Often expressed as a prepositional phrase
- Predicate nominative [PNOM]
- Object complement [OCOMP]

Two arguments

- ## SBJ, OBJ
  - Pericles governed the empire.
  - The Ionians decided to obey the king.

- ## SBJ, PNOM
  - The woman was Pericles' wife.
  - The army was in danger.

Object (OBJ) = Obligatory arguments of the verb
[no more than 2 per verb]
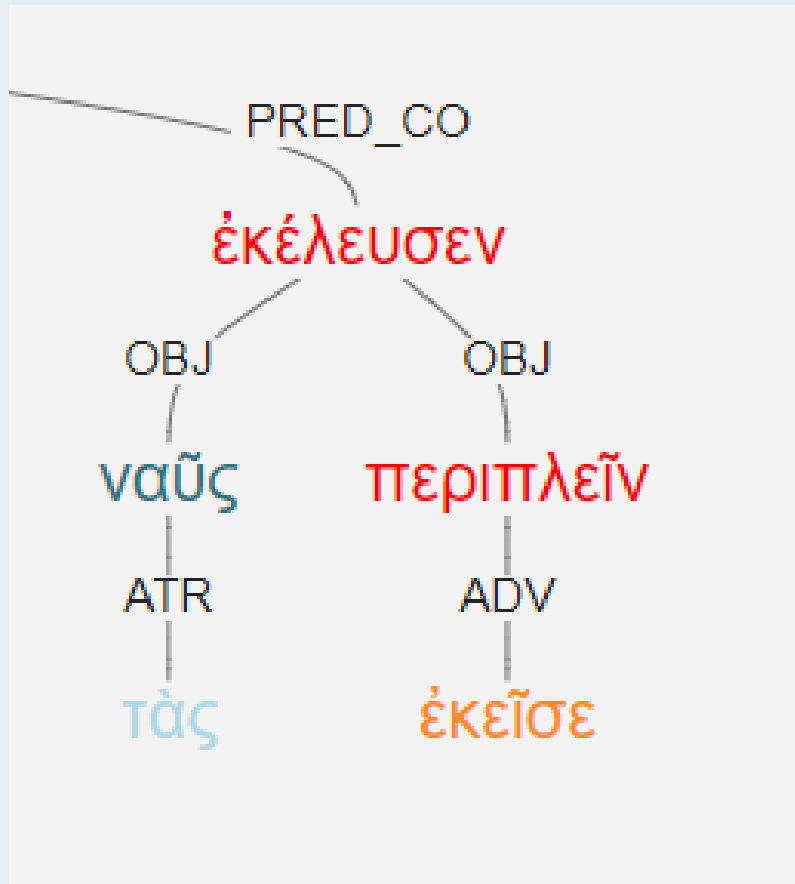
"…and he ordered the ships to sail around to there."

Predicate Nominal (PNOM) = subject complements

Usually with verbs of being/becoming/seeming, or <u>with some passives</u>
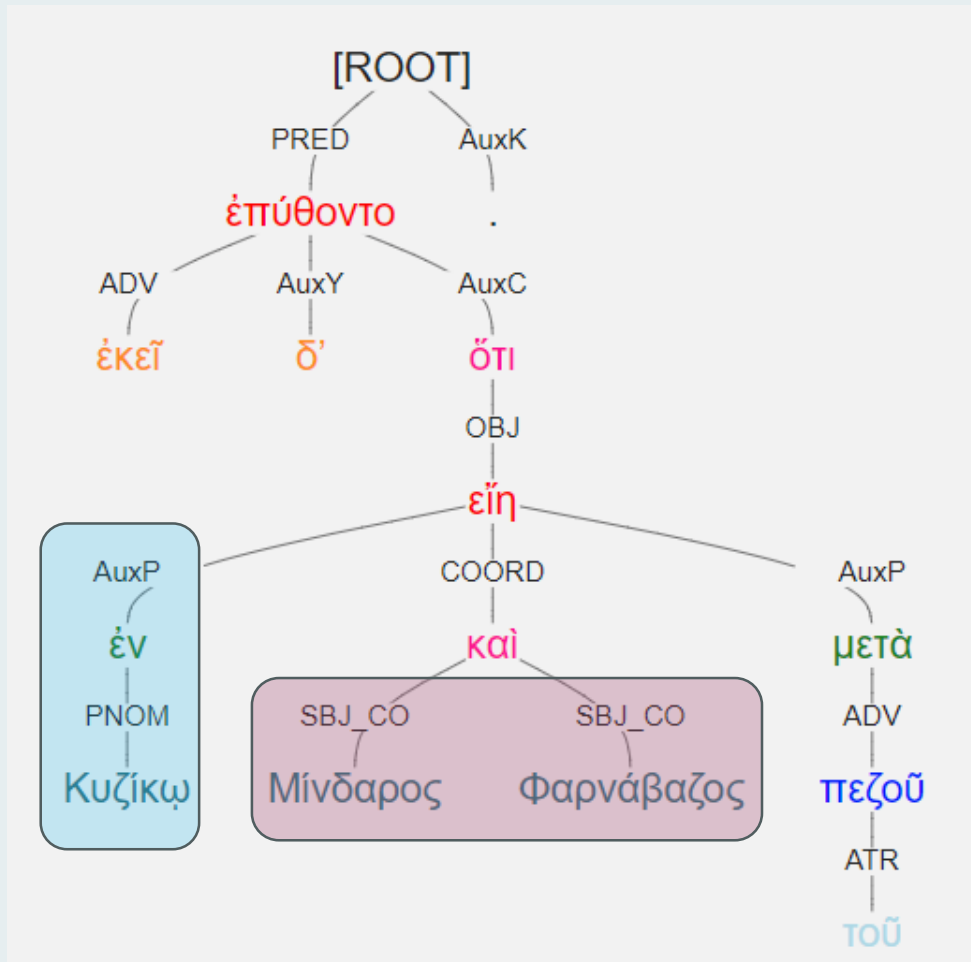
"And there they learned that Mindaros and Pharnabazos were in Kyzikos with the infantry."

Three arguments

- **SBJ, OBJ, OBJ**
  - The god gave them victory.
  - The merchant loaded his ship with many jars.

- **SBJ, OBJ, OCOMP**
  - The Athenians selected him general.
  - Pericles made Athens the supreme power in the Aegean Sea.

Object Complement (OCOMP)
To make [elect/choose] x y.

"And they offered as witnesses of these things the pilots of those ships and many others of those who had sailed with them."
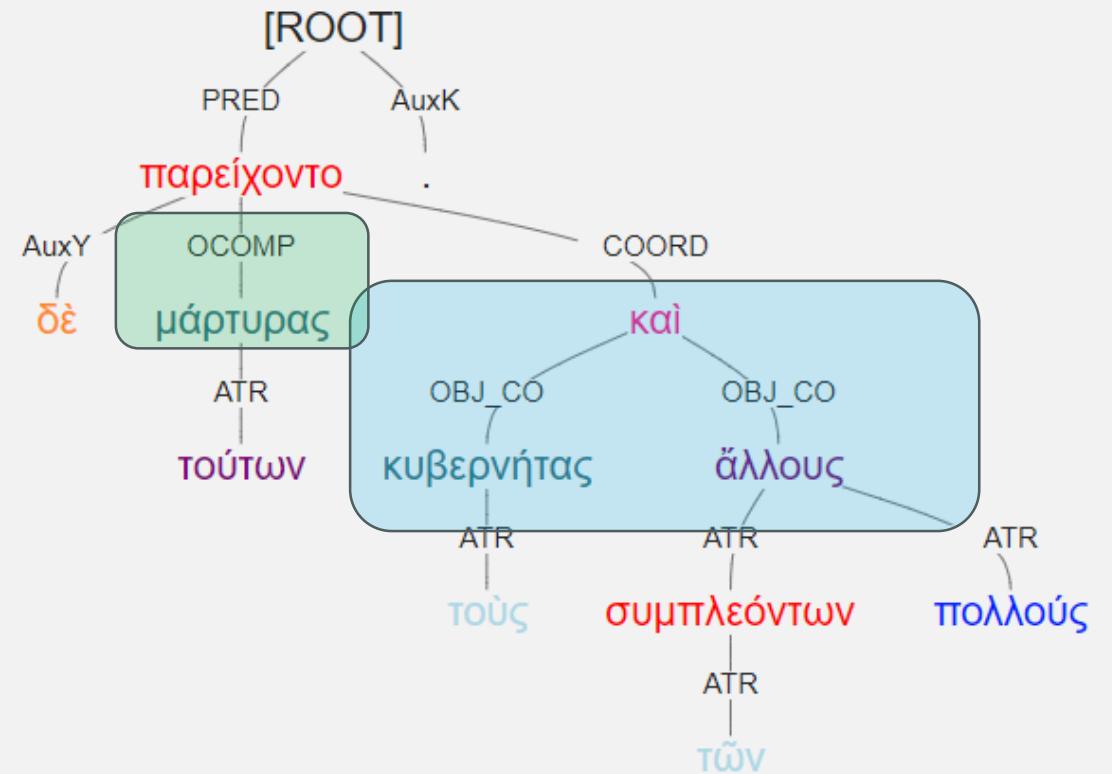
Zero-Valency Verbs

- It is raining.
- It is thundering.
- *There is no "it".*

Subordinate verbs may also have subjects and objects

"But when he [Alkibiades] arrived to him [Tissaphernes] with one trireme, bringing both friendship offerings and gifts, arresting [him], he shut Alkibiades up in Sardis, **saying that the king ordered [him] to make war on [the] Athenians**."
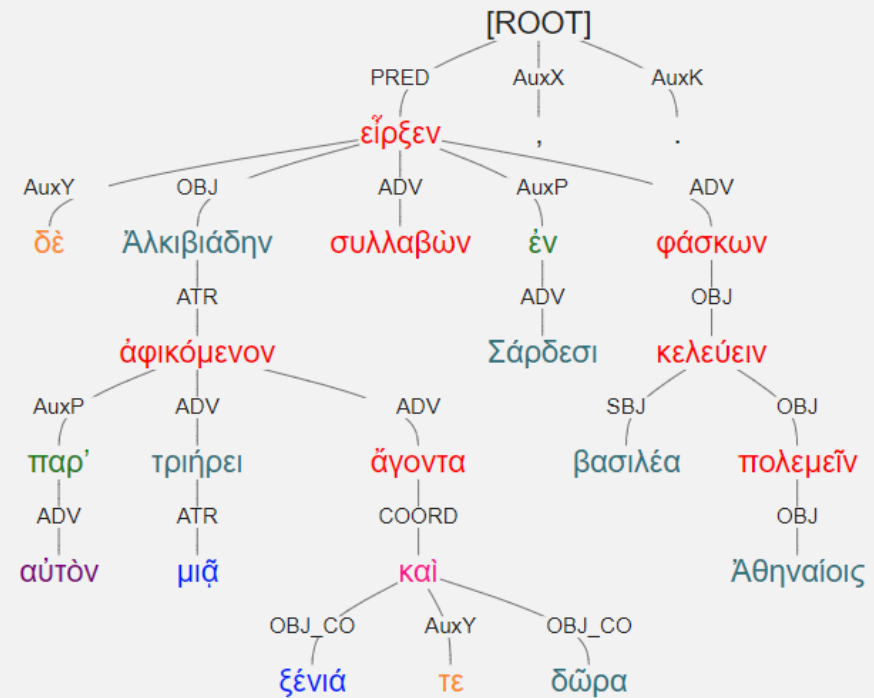
Adverbial [ADV] = optional information that is not a <u>necessary</u> argument of a verb

Then, at night, nearby, there, afterwards, very, etc.

"And the things around Mytilene were fought in this way."
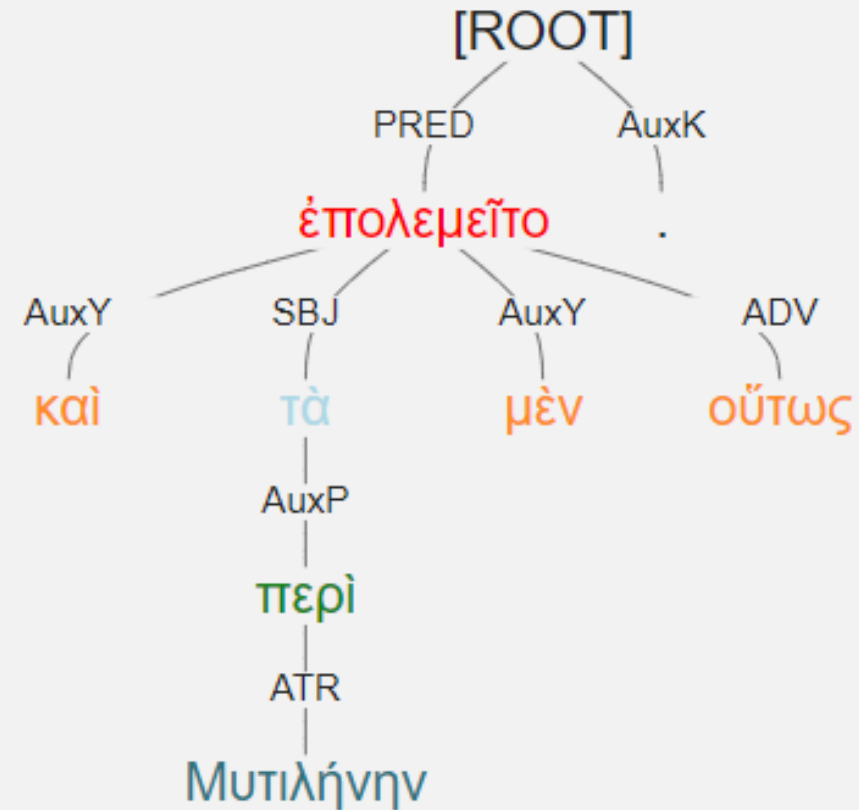
Adverbial [ADV]
May be participles or prepositional phrases

"The Athenians, setting out, fought [them] along the beach near Abydos from morning until afternoon."
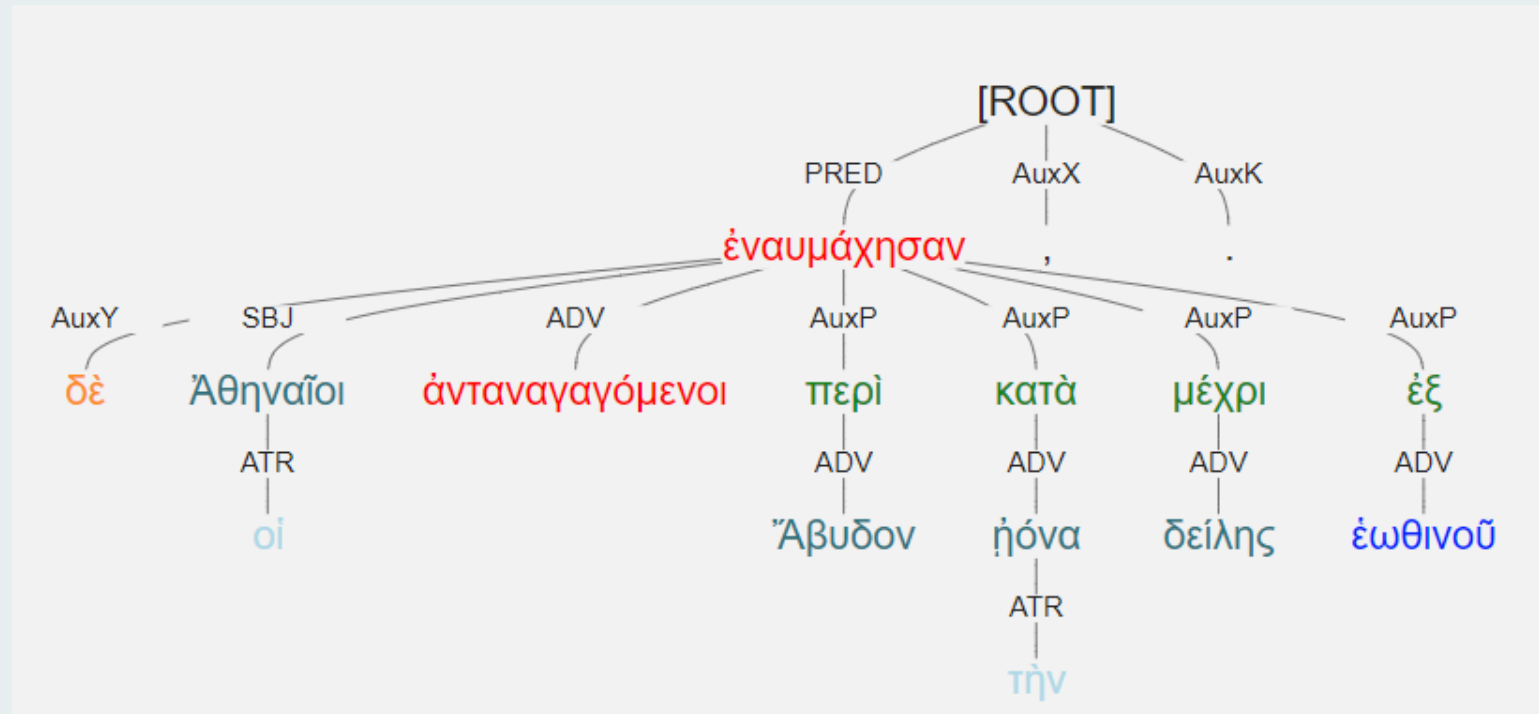
The noun phrase

- We have seen SBJ, OBJ, OCOMP, PNOM, and ADV (usually) that are dependencies from verbs
- Nominal forms may also have dependencies
  - Adjectives
  - Adverbs
  - Prepositional phrases
  - Participles
- Labeled **ATR** [attributive]

Attributive (ATR) = Words that specify or delimit or describe their head, including the article [*adjectival elements*]
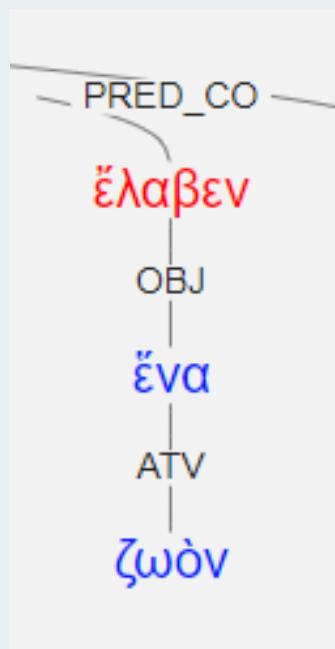
"And when the Athenian watcher perceived him, he signaled to the generals."

Literally: "the watcher of the Athenians"
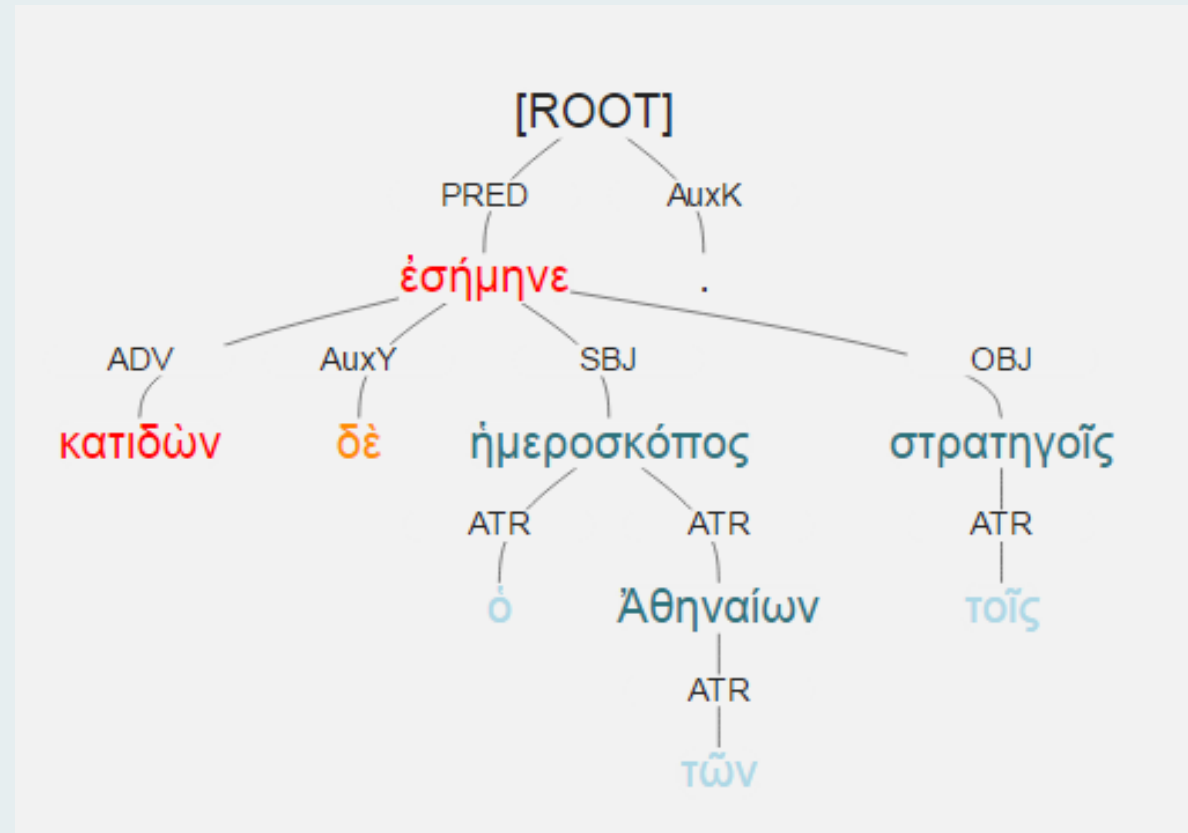
Verbal Attribute= ATV/ATvV

Agrees with its head noun <u>morphologically</u>, but also <u>qualifies</u> the function of the verb necessarily.

- The neighbors left **<u>unarmed</u>**. [NOT: The unarmed neighbors left.]

- I drink my coffee **<u>hot</u>.** [Not: I drink my hot coffee.]

- I am going to sand the floor **<u>smooth</u>**. [NOT: Sand the smooth floor.]

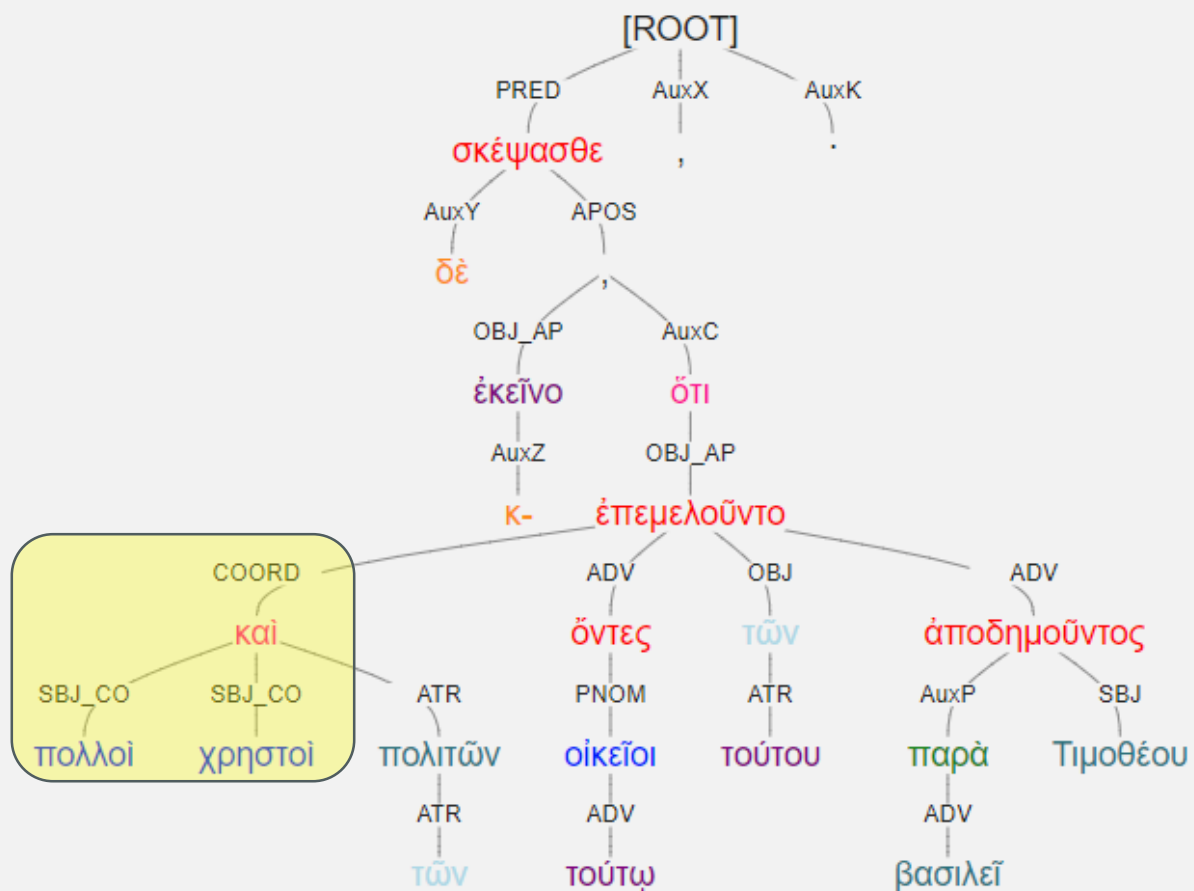- He captured one man **alive**. [NOT: He captured one live man.]

Suffixes with Conjunctions and Appositives

- Elements conjoined receive the suffix _CO
- Elements in apposition receive the suffix _AP
- Elements with both receive _AP_CO

Usually use a comma to indicate apposition [APOS]

"And notice this, that many worthy citizens were looking after his affairs…"

# Functional bridges: some words are always labeled in the same way

- Prepositions (AuxP): *of, from, to, by, across, over,* etc.

- Coordinating conjunctions (COORD): *and, or, yet, but,* etc.

- Subordinating conjunctions (AuxC): *when, since, because, although,* etc.

- Sentence Adverbs [AuxY] and emphasizing particle [AuxZ]

- These words are never given suffixes [_CO or _AP]
 - Put the suffixes on the object of the preposition or the coordinated elements themselves, etc.

Prepositions (AuxP) are functional bridges

"The Athenians, setting out, fought them along the beach near Abydos from morning until afternoon."
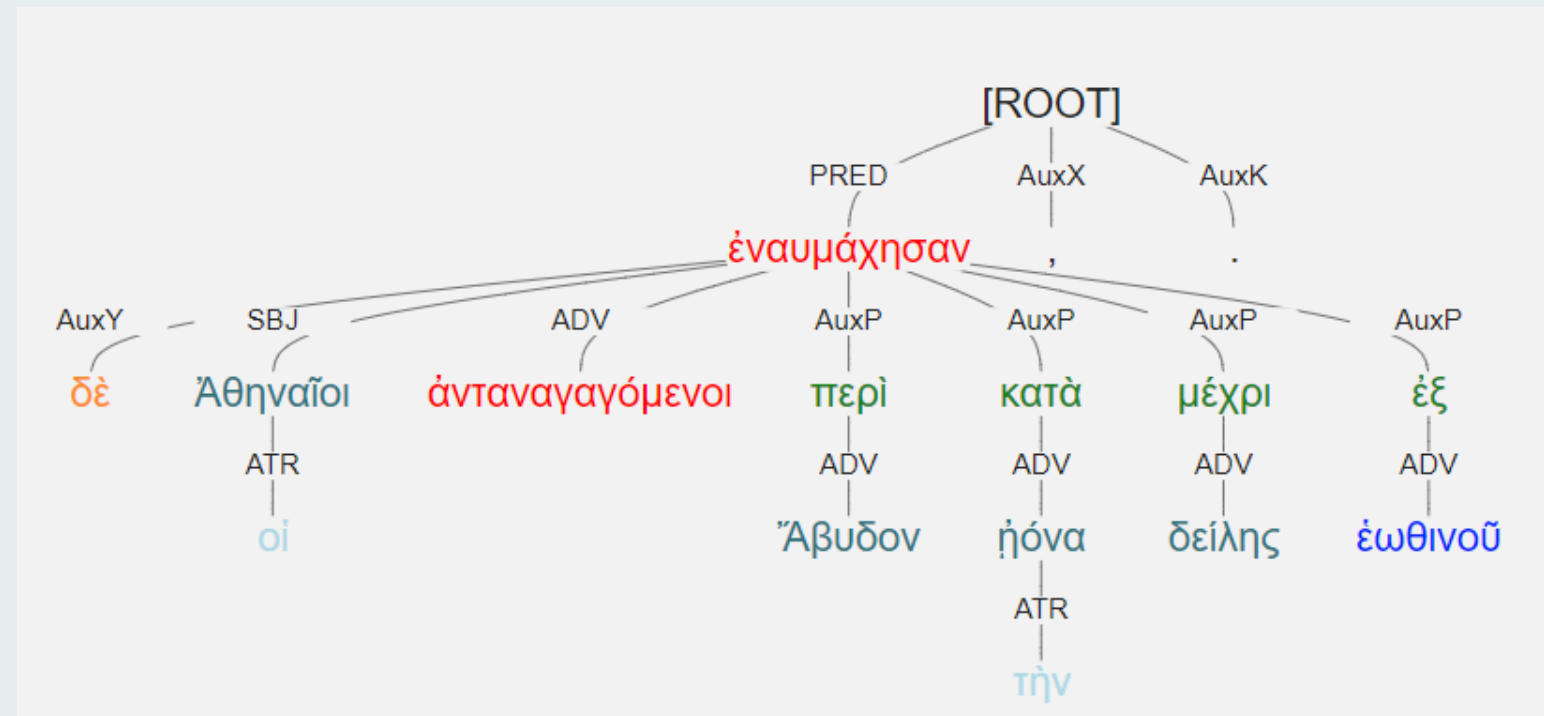
COORDs are functional bridges

"And the Athenians and Lakedaimonians immediately again fought a sea battle, and the Lakedaimonians won, with Agesandridas leading them."
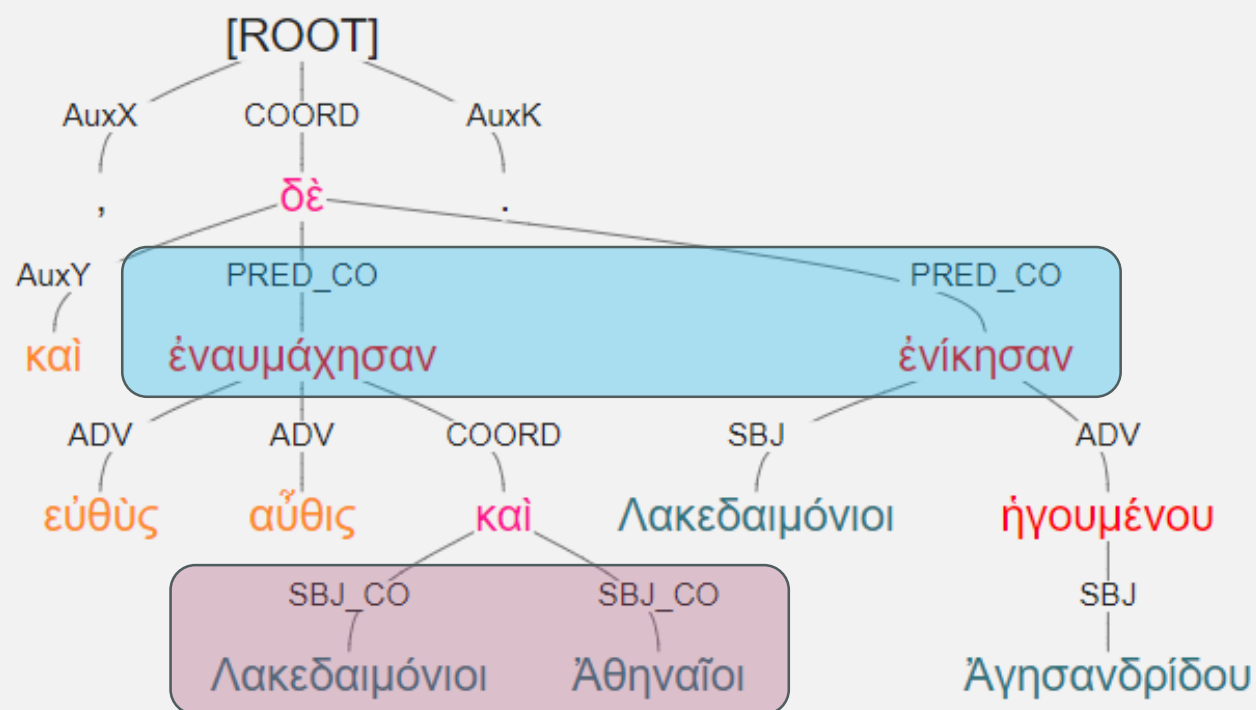
Sentence Adverbials = AuxY
[never _CO or _AP]

- Considered **particles**, which are a subcategory of adverbs
  - Label them morphologically as **adverbs**

- Usually apply to the whole sentence
  - Includes connectives at the beginning of a sentence
  - Includes coordinators in a series
  - Clauses expressing author's attitude ["As they say"] or organizing the discourse ["to make a long story short"]

Sentence Adverbials [Aux Y]

"But when he [Alkibiades] arrived to him [Tissaphernes] with one trireme, bringing both friendship offerings and gifts, arresting [him], he shut Alkibiades up in Sardis, saying that the king had ordered [him] to make war on [the] Athenians."
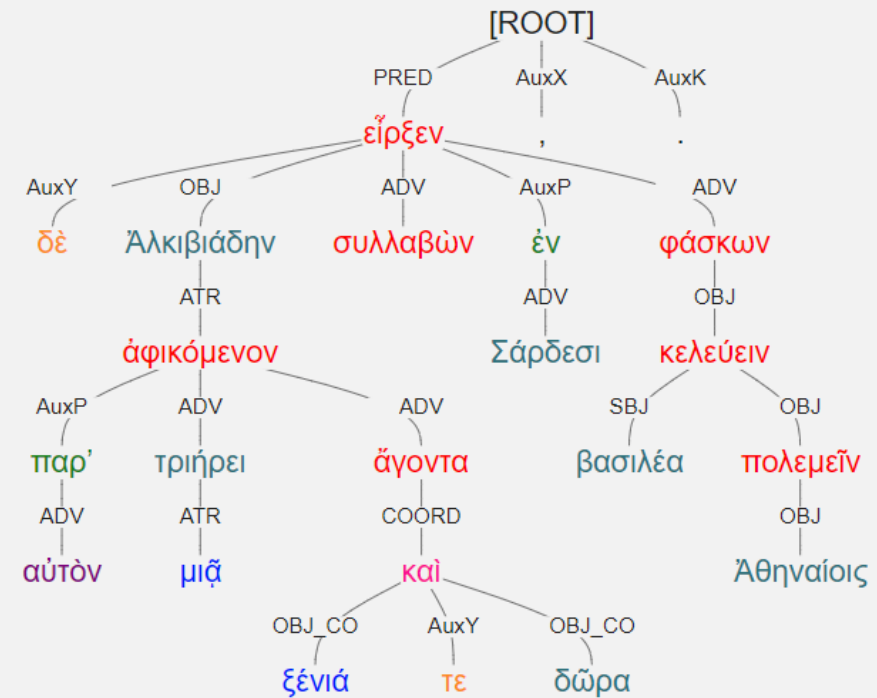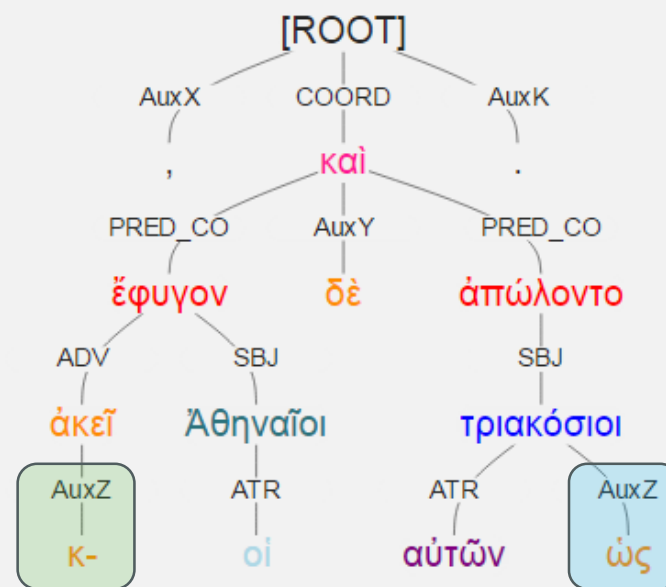
Emphasizing Particles = AuxZ
[never _CO or _AP]

- Emphasize one particular word
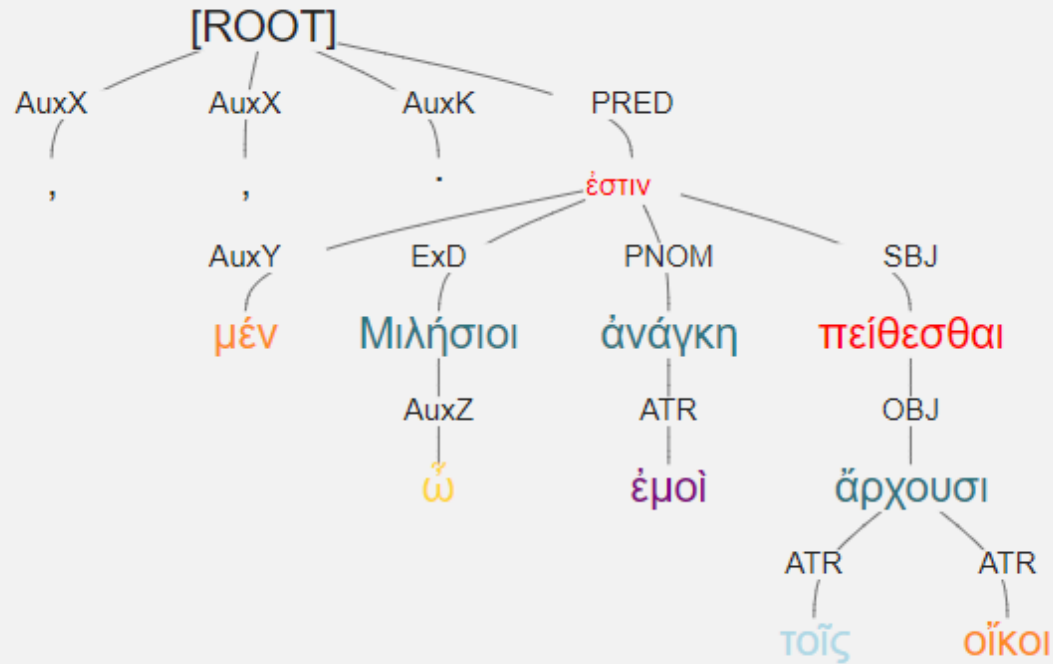- Negations; *only, also, especially*, etc.
- When in doubt, call it an ADV.

Emphasizing Particles = AuxZ

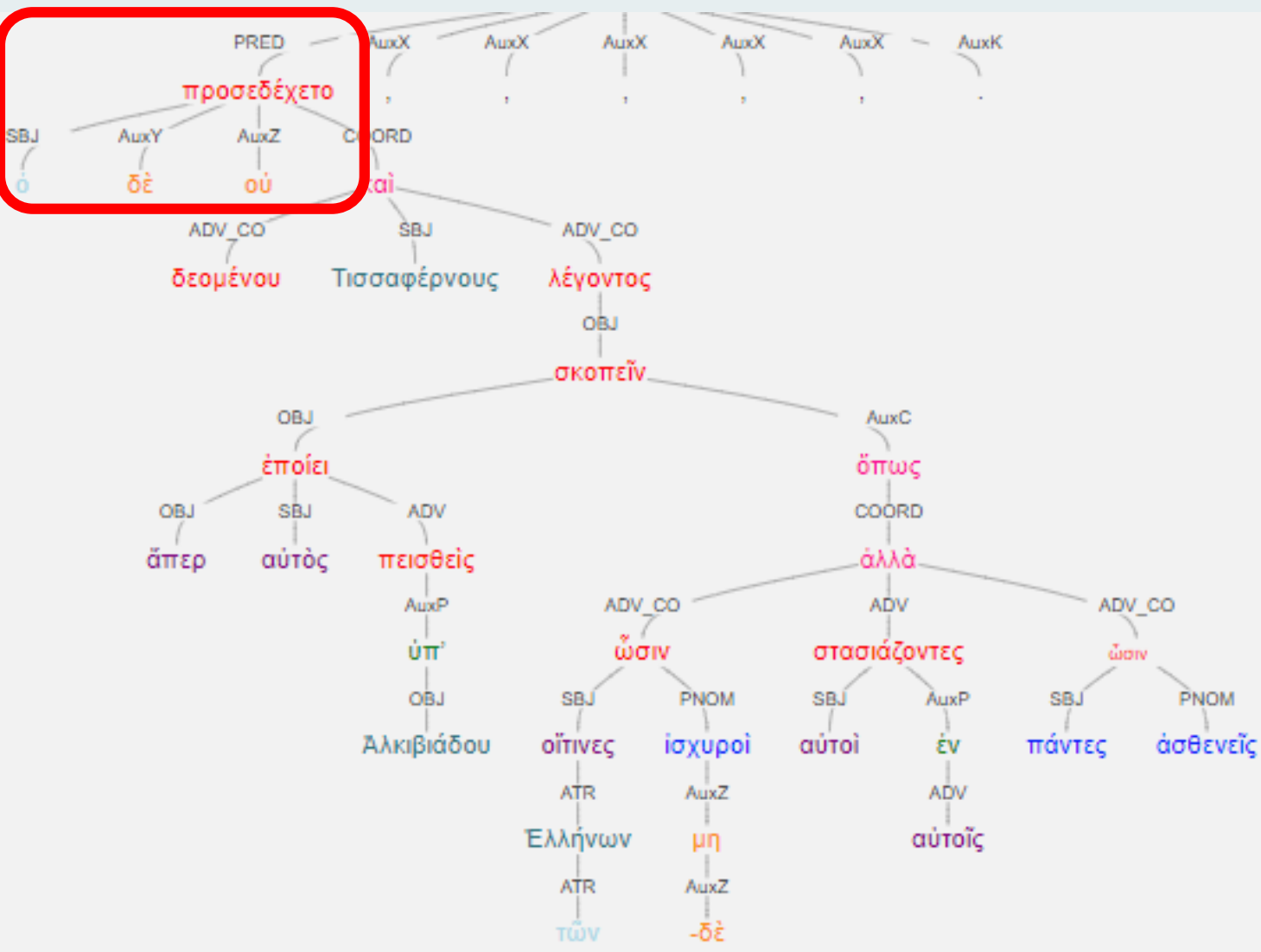"And the Athenians were fleeing also in that place, and about three hundred of them were killed."

External Dependency [ExD]
Asides, vocatives, interjections, etc.

"It is necessary for me, o Milesian men, to obey the archons at home."
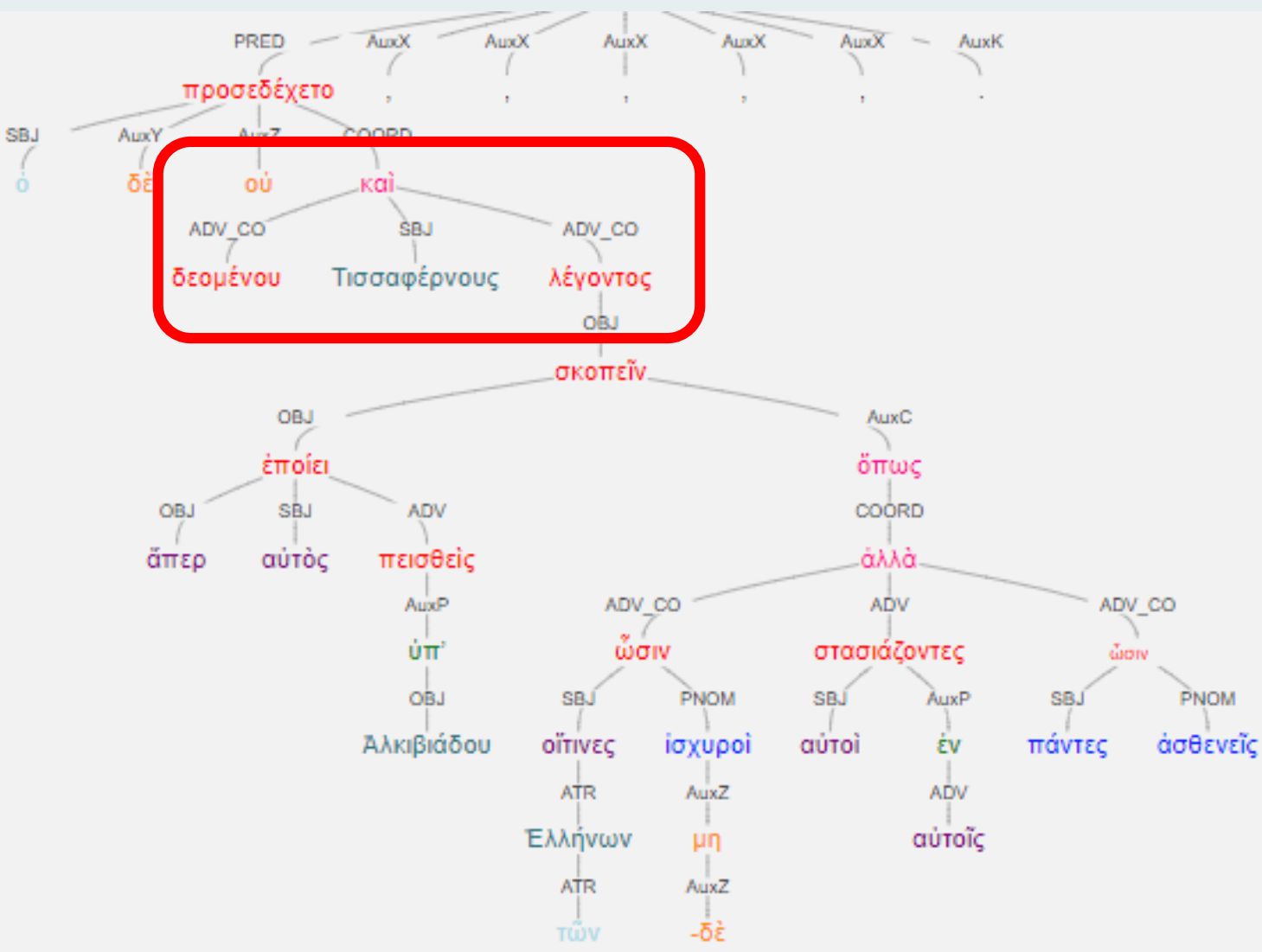
Main Sentence

"**And he would not receive them**, even with Tissapherenes asking him to and saying that he would see the things he himself was doing, persuaded by Alkibiades, so that any of the Greeks would not be strong, but, themselves fighting among themselves, all would be weak."

Genitive Absolute

"And he would not receive them, even with Tissapherenes asking him to and saying that he would see the things he himself was doing, persuaded by Alkibiades, so that any of the Greeks would not be strong, but, themselves fighting among themselves, all would be weak."
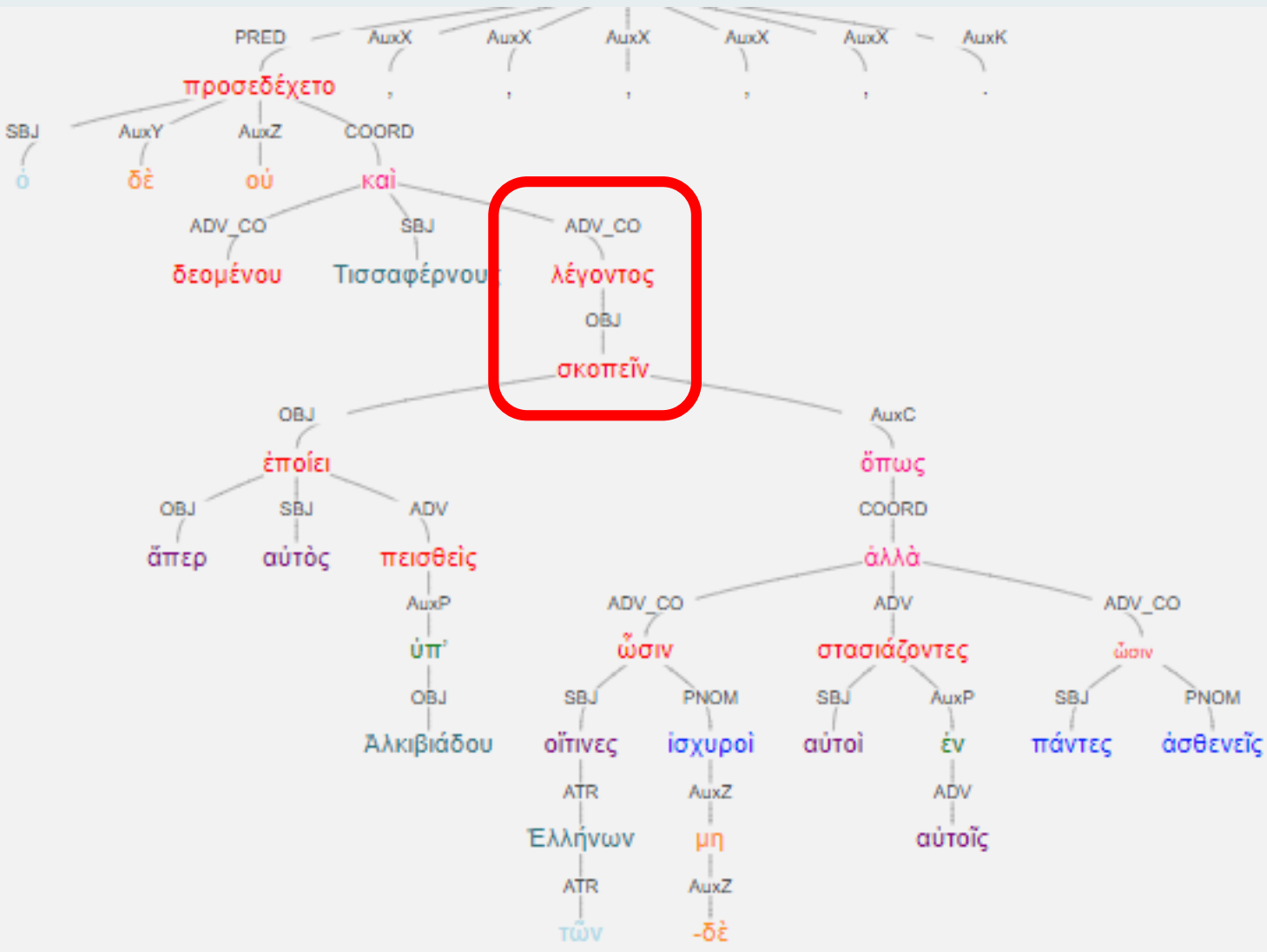
Indirect Speech

"And he would not receive them, even with Tissapherenes asking him to and saying that he would see the things he himself was doing, persuaded by Alkibiades, so that any of the Greeks would not be strong, but, themselves fighting among themselves, all would be weak."
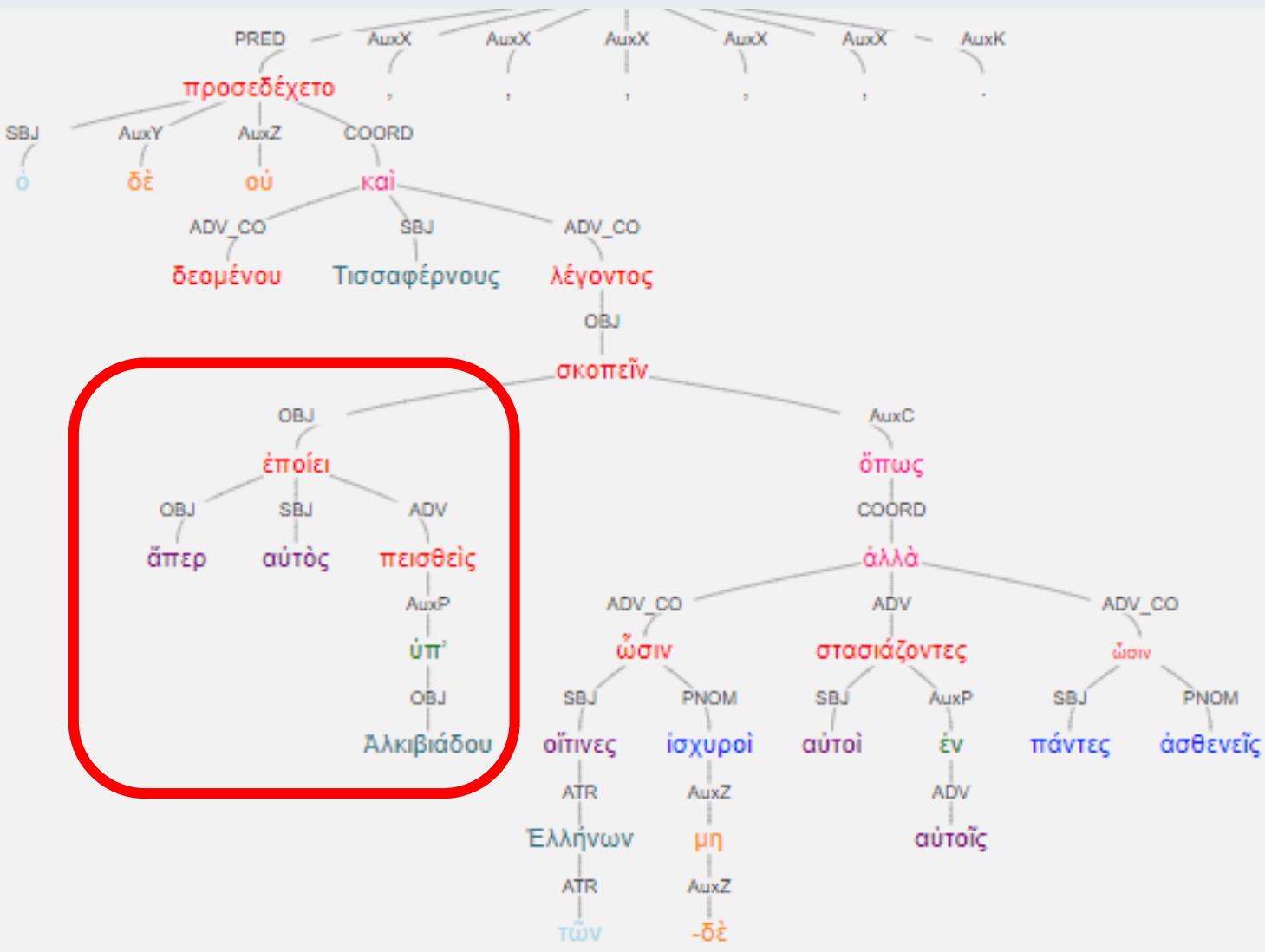
Relative Clause
without an antecedent

"And he would not receive them, even with Tissapherenes asking him to and saying that he would see the things he himself was doing, persuaded by Alkibiades, so that any of the Greeks would not be strong, but, themselves fighting among themselves, all would be weak."

Purpose/Final Clause

"And he would not receive them, even with Tissapherenes asking him to and saying that he would see the things he himself was doing, persuaded by Alkibiades, so that any of the Greeks would not be strong, but, themselves fighting among themselves, all would be weak."
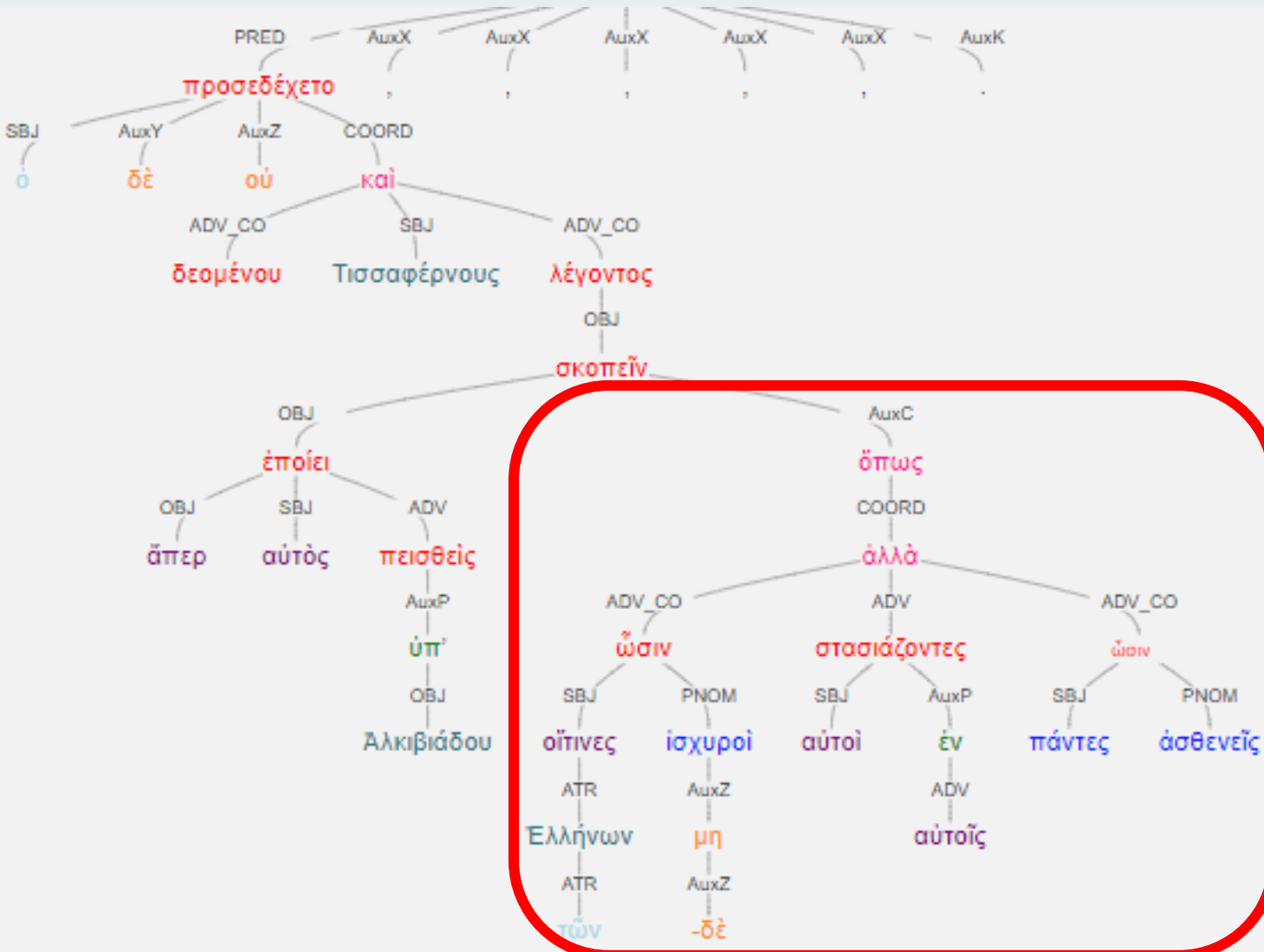
Helpful Links

- *Guidelines for the Syntactic Annotation of the Ancient Greek Dependency Treebank* (1.1), by David Bamman and Gregory Crane [September 1, 2008], http://nlp.perseus.tufts.edu/syntax/treebank/greekguidelines.pdf

- Treebanking Tips for Dependency Syntax [VGorman] https://github.com/vgorman1/Greek-Dependency-Trees/blob/master/Treebanking%20Tips%20for%20Dependency%20Syntax.pdf

- "Reading Ancient Greek in the Digital Age." An on-line, open-access course in beginning Greek relying on digital tools rather than rote memorization.
  - https://vgorman1.github.io/Greek-Language-Class/

- Playlist on YouTube: Greek Dependency Treebank Instructions
  - https://www.youtube.com/playlist?list=PLZUQZ95zNNL8S4NrfT7KgdlmtBEgHIq5D
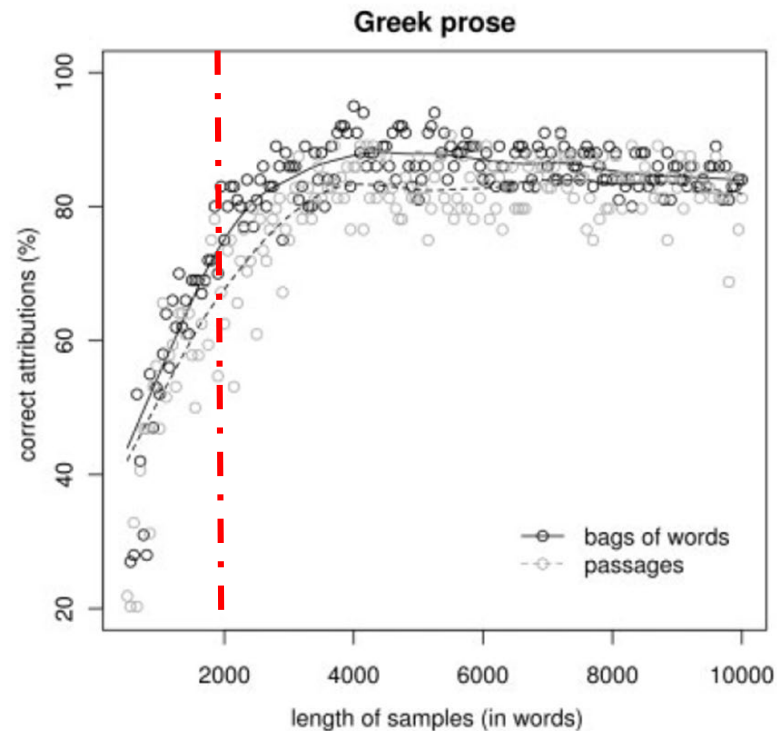
Research directions

- Author attribution studies using computer-parsed trees
  - Gorman, Robert. 2019. "Author Identification Of Short Texts Using Dependency Treebanks without Vocabulary." *Digital Scholarship in the Humanities* 35. academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqz070/5606771.
  - Gorman, Vanessa B. and Robert J. Gorman. 2016. "Approaching Questions of Text Reuse in Ancient Greek Using Computational Syntactic Stylometry." *Open Linguistics* 2: 500-510. https://www.degruyter.com/view/j/opli.2016.2.issue-1/opli-2016-0026/opli-2016-0026.xml?format=INT

N-grams vs stylometry
[Eder 2013 vs. Gorman 2019]

Fig. 6 Dependence of attribution accuracy and length of text samples: 72 Ancient Greek prose texts

R. Gorman

Table 2 Classification of the prose corpus

| Text size	Total 'Guesses'	Total errors	Accuracy (%)
2,000	17,000	0	100
1,900	18,000	0	100
1,800	18,000	0	100
1,700	20,000	2	99.98
1,600	21,000	0	100
1,500	22,000	0	100
1,400	24,000	1	99.99
1,300	25,000	0	100
1,200	28,000	1	99.99
1,100	30,000	1	99.99
1,000	33,000	0	100
900	37,000	4	99.98
800	41,000	0	100
700	47,000	17	99.96
600	55,000	38	99.93
500	66,000	157	99.76
400	83,000	286	99.65
300	110,000	662	99.39
200	165,000	2,705	98.36
100	329,000	18,466	94.38
50	657,000	103,748	84.20

Identify Spurious Passages

Xenophon, *Cyropaedia* 8.8 and 1.2

Red line = similarity to rest of text [Cyr. 1 and 8]
Black line = similarity to 8.8
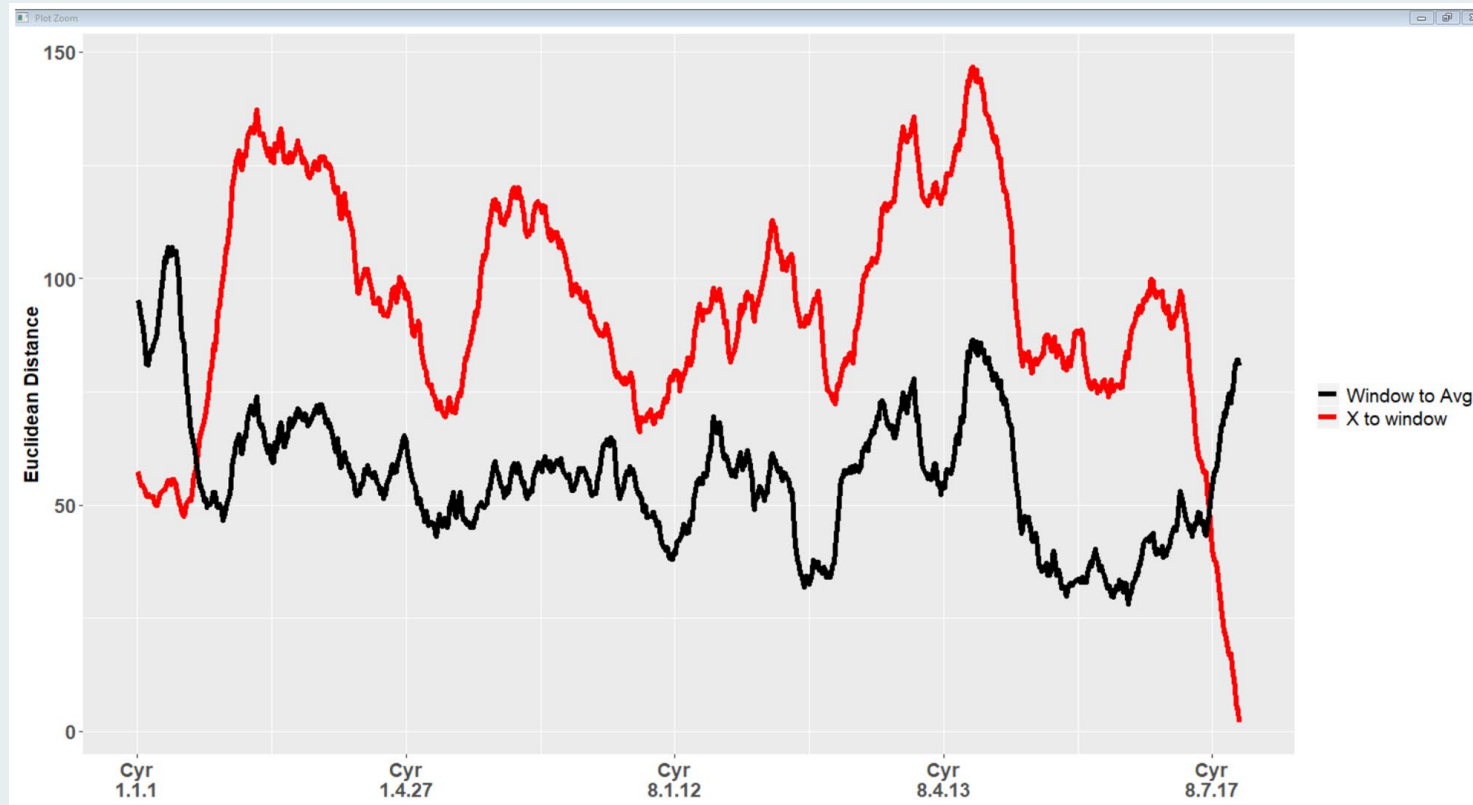
Stylistic variations within one author
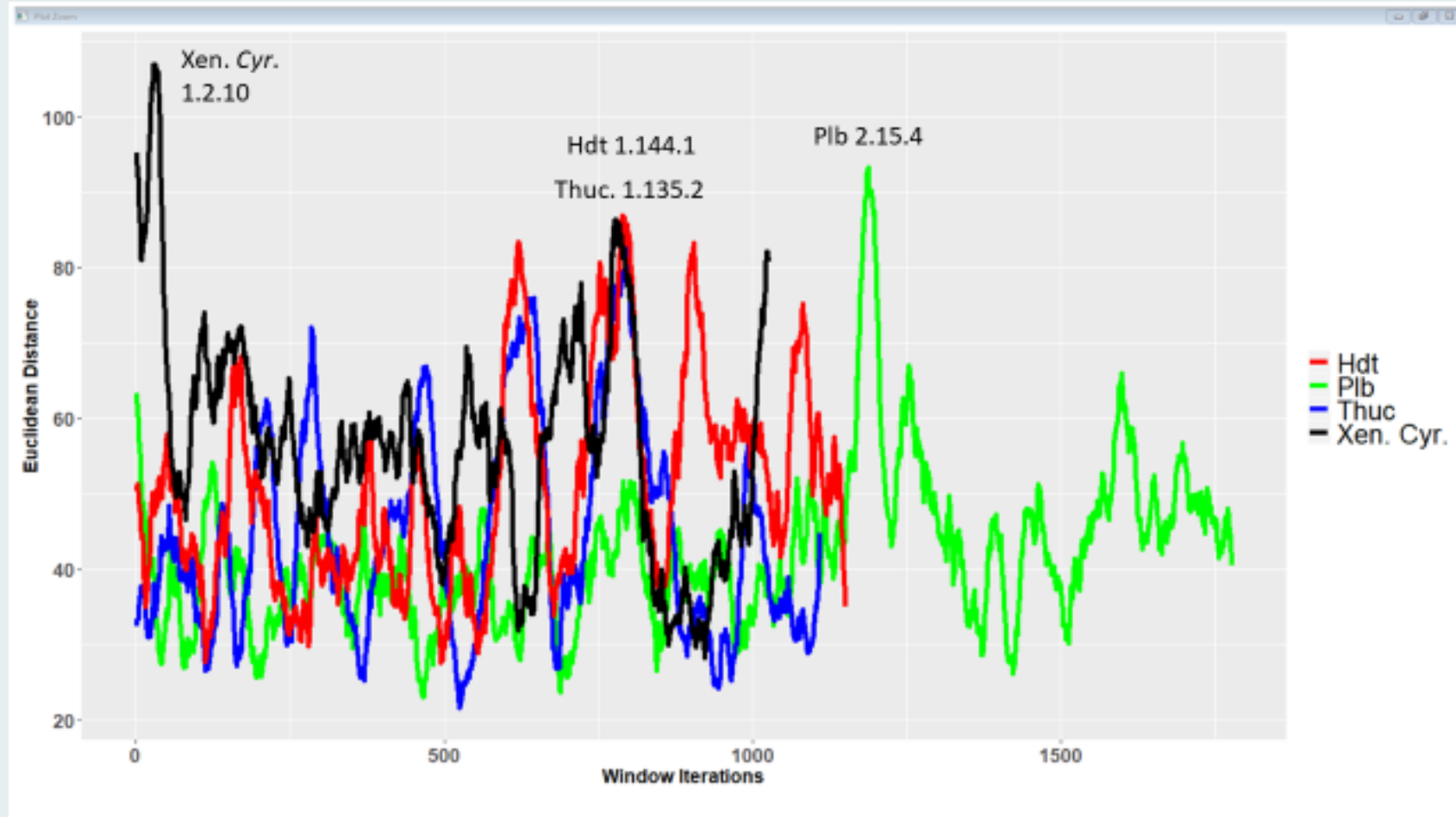
Natural Language Processing

- The Pedalion Project at UK Leuwen (Toon Van Hal)
- http://en.pedalion.org/treebanks