CHAPTER 5

The Ancient Greek Dependency Treebank: Linguistic Annotation in a Teaching Environment

Francesco Mambrini

German Archaeological Institute

Abstract

This chapter argues that manual linguistic annotation of Ancient Greek texts can be effectively employed to teach of Greek literature and languages. Under the supervision of a teacher, students can be engaged into the ongoing creation of the Ancient Greek Dependency Treebank. With the help of one example from Sophocles (*Tr.* 962–3), we will illustrate how the collective work of treebanking in a class environment provides an ideal occasion to discuss the methods of Classical Philology and the history of interpretation of a given passage; more importantly, while producing a treebank annotation, students can learn how to read a complex text in its literary and communicative context following the methods of textual criticism. New and old research questions emerge from the work; at the same time, through the final annotation the students will produce a tangible contribution to a crucial initiative that is likely to change the way Greek grammar will be studied in the future.

1 Introduction

In the fall of 2009, the Perseus Project published the first edition of the *Ancient Greek Dependency Treebank* (AGDT),[1] a digital corpus of Greek literary texts that include a word-by-word morphological and syntactic annotation. At the

moment, the last published version of the treebank (AGDT 1.6) include the complete extant opus of Homer, Hesiod, Aeschylus, five of the seven surviving plays of Sophocles, and smaller selections of Plato (the *Euthyphro*) and Athenaeus (Book 12 of the *Deipnosophistae*).[2]

Treebanks are a powerful resource for data-driven linguistic research which are likely to have a great impact on the way the grammar of the ancient languages is studied. Traditional grammars have often limited themselves to register the existence of certain linguistic facts, providing at best a detailed classification of the constructions and a number of examples from the ancient texts. For example, we learn from grammar that in Greek coordinated subjects can trigger either plural and singular agreement with the verb, and that singular verbs occur more often with or-coordinates than with subjects joined by 'and';[3] we do not find any indication, though, of how frequent each of these constructions is, how the two agreement patterns are distributed among the authors and texts, and what the different meaning of these constructions (if any) might be. A large digitized repertoire of texts that can be searched for specific syntactic constructions will make these information easily retrievable to students and scholars alike.[4] Moreover, the treebank formalism, which allows to represent the sentence as a tree-shaped graph as in Fig. 1, provide a formidable tool to visualize the sentence structure.
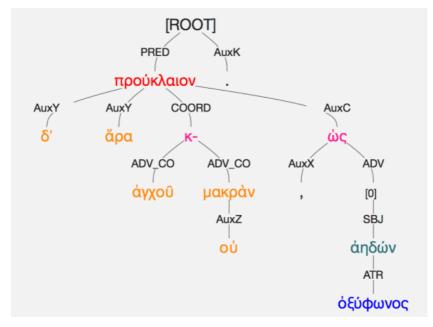

Figure 1: Sophocles, *Trach., 962–3.*

1.1 Treebank annotation

Although some experiments on (semi-)automatic parsing of Latin and Greek have already been carried out,[5] so far, all the information, including part of speech, morphological features (tense, mood, person etc.), and the syntactic relations between each word in the texts, have been entered manually by human annotators. This process of word-by-word enrichment can be facilitated with the help of graphical interfaces and online tools, such as Arethusa (see Section 1.2).

Some of the texts, and the poems of Homer and Hesiod in particular, were annotated by students in the context of graduate programs in Classics.[6] The pedagogical value of such an exercise of close reading cannot be overestimated.[7] In fact, by using the formalism of the AGDT to enrich an ancient text with morpho-syntactic information, students can both practice their language skills and contribute to the advancement of the available resources to an extent that is scarcely matched elsewhere in the Classics. Not only will it be possible for students to 'learn by doing', but the publication in the AGDT corpus can also offer an immediate gratification to their efforts.

Along this line, my paper will draw the attention on what contribution the annotation process can bring to the teaching of Greek language, literature, and civilization. We will present the case for engaging the students in the practice of collective treebanking, using the formalism and the guidelines of the AGDT and under the supervision of a teacher.[8] The advantages in terms of grammar and language acquisition that are inherent in the process of a word-by-word morpho-syntactic annotation are immediately evident. However, instead of focusing only on them, my analysis will attempt to show the wide spectrum of methodological and historical problems that a formalized linguistic annotation entails. Reading even a simple sentence at the level of detail that treebanking requires is a process that goes far beyond grammar and can be leveraged to teach Greek civilization *tous azimuts*.

Methodological questions on how to reconstruct, interrogate, and interpret an ancient text are, as we shall see, especially prominent in treebank annotation. Moreover, we will stress that students must be encouraged to question the meanings and interpretations of a text that each of the possible reconstructions of a sentence imply. Finally, by searching the collection of the already annotated texts, students can also be asked to reflect on the relation of the text they are reading in the larger context of the ancient Greek literature.

1.2 Reading through treebanking: Sophocles, Trachiniae, 962–3

In what follows, we try to articulate this program with one example, limited to a short sentence from Sophocles. The discussion will touch only a minimal part of the potential benefits of treebanking in a classroom environment.

Others application (e.g. interdisciplinary projects involving students in computer science and linguistics in cooperation to improve the efficiency of the research tools) will be left out of the present work. The use of corpora to generate drills and exercises, that can also be applied to measure each student's familiarity with single grammatical aspects and assess personalized training sessions on the weakest point, is also a potentially crucial use that we will have to leave aside.[9]

We will consider one sentence taken from the fourth choral ode (*stasimon*) of Sophocles' *Women of Trachis*.[10] Fig. 1 shows how the sentence is annotated in 'Arethusa', the new annotation framework that has been recently made available as part of the Perseids editing environment and can be freely accessed on the Internet.[11] The Greek text of Sophocles, along with a minimal paraphrase, is reported below; this starting point should mirror the situation in a class: students should be confronted immediately with the original, and no translation (except for the basic meaning of some of the most unusual words) should be provided. A more articulate translation will emerge while we progress in the annotation:

> ἀγχοῦ δ᾽ ἄρα κοὐ μακρὰν προύκλαιον,
> ὀξύφωνος ὡς ἀηδών
> near and not far off then [I was?] weeping beforehand, like the shrill-voiced nightingale.

The *Women of Trachis* is probably not one of the most popular tragedies in school curricula. Moreover, the short passage that we selected does not belong to the most memorable passages of the play; these words are likely to be overlooked as a moment of transition between two important scenes. Yet precisely these reasons convey interest to our choice: one of the aim of the paper is to illustrate how even a short and apparently uninteresting sentence can in fact, when considered through the lenses of treebank annotation, raise complex literary and historical questions to engage students in fruitful discussions.

2 The Sentence in its Context

It is often customary to remind beginners in Greek and Latin that every fresh analysis of a sentence should start by the identification of the main verb. This approach is certainly sound; since the prototypical dependency tree is generally rooted to the main predicate, which in turns governs a bunch of satellites (as in Fig. 1), the indication to start there is also well suited to the theoretical frame our treebank is built upon.[12]

Yet one of the first lessons that can be learned while reading a sentence like this is that knowledge about the context constitutes an even more fundamental premise. Context (intended both as the 'intra-textual' net of references

and presuppositions to other passages of the work, and as the communicative situation a text is inserted in) is a primary linguistic element, which is often crucial in disambiguating syntactic and semantic problems. As we will see, our sentence offers a good illustration of this point.

The *Women of Trachis* dramatizes the agony and death of Herakles, which is involuntarily caused by the gift sent by the hero's wife Deianeira on the occasion of his return to Greece. After the narrative of the lines 899–946, where the Nurse told how Deianeira killed herself after she discovered the real effects of her actions, the Chorus awaits the second and final evil; the agonizing Herakles will be eventually brought to the scene and displayed to the audience.

As it is typical of Greek tragedy, it is a song by the Chorus, which in the play impersonates a group of young maidens from the town of Trachis, which builds the dramatic tension and bridges the two sections. Our fourth song is dominated by the opening questions: 'which evil shall we bewail first, which of the two is more grievous' (947–9)? The sight of the escort that brings the bier of Herakles reveals that the evil that the Chorus has already anticipated is almost at hand.

It is typical of choral odes, and of Sophocles in particular, that the first strophes of a *stasimon* are concerned with general questions or with mythical paradigms, while the last stanzas bring the focus back to the stage events and introduces the scene to come.[13] Our sentence operates precisely this shift. The meaning of the words (with an emphasis of 'near' and 'to mourn in advance'[14]) points to the dramatical function of introducing the new characters that are about to enter the scene and the theme of the episode; if the general meaning is clear, the exact grammatical interpretation of the words proves to be more challenging.

3 Morphology

Identifying the main verb of the sentence requires students to define the part of speech of each of the words and then to concentrate on the full morphological analysis (mood, tense, person) of the verbs. Often, students will meet ambiguous words, where more than one analysis is possible. In such cases, disambiguation will have to rely on the syntax or on the general knowledge of meaning and context.

In our sentence, only one word is liable to two different interpretations, as it is shown in the interface for morphological annotation of Arethusa (Fig. 2).[15] The main verb προύκλαιον can be interpreted as:

1. indicative imperfect 1st person singular of προκλαίω (*we mourned/were mourning*);
2. indicative imperfect 3rd person plural of προκλαίω (*they mourned/were mourning*).

Figure 2: Arethusa: annotation interface for morphology.

Since the sentence lacks an explicit subject, both are theoretically possible. With n.1 the implied subject is the Chorus, who can, as usual, shift between *I* and *we* for self-reference.[16] With n.2, the subjects are the men who carry Herakles on the litter.

Advanced readers of Greek will be in no doubt about the correct answer, but it is interesting to note that both interpretations are attested in the history of scholarship. N.2 is adopted by an ancient commentator whose opinion is preserved in the medieval manuscripts, in a marginal note (*scholium*) to the line:

> ἀντὶ τοῦ προκλαίουσιν· ὁ χορὸς αἰσθάνεται τοῦ Ἡρακλέους πλησίον φερομένου καὶ πλήθους θρηνούντων ἐπακολουθούντων αὐτῷ
>
> [*they were lamenting*] instead of *they lament*: the Chorus perceives that Herakles is brought near and [perceives] the crowd of mourners that is escorting him.[17]

At this point, students should be encouraged to discuss: are the two interpretations equally admissible? Do we have arguments to choose between them? The context that we described above provides several strong arguments to reject the interpretation of the scholium. The closing of the stanza, in which the Chorus asks why the escort is advancing in such an ominous silence (cf. 965–7), speaks strongly against it, as it was already noted by an eminent scholar.[18] Another argument is grounded in grammar: the equivalence between imperfect and present that the scholiast invokes cannot be seriously considered.

On the contrary, the imperfect makes a perfect sense if it is referred, as it is, to the laments that the Chorus was uttering in the preceding stanzas. The meaning that we chose for the verb ('lament in advance', 'weep beforehand') is perfectly at home in reference to the first part of the ode where maidens were lamenting the sort of Herakles even before seeing it. Now, with the approach of the litter, the time of foreboding is over.

4 Syntax

4.1 ὀξύφωνος ὡς ἀηδών

The easiest syntactic structure of the sentence is that formed by the last three words. This phrase introduces a simile in which the lament of the Chorus is compared with the wailing of the nightingale. 'Nightingale' (ἀηδών) and 'shrill-voiced' (ὀξύφωνος) thus make a noun–epithet pair, and the similitude has to be connected with the main verb (προύκλαιον). In the case of comparisons introduced by 'like', the guidelines of the AGDT require annotators to take the term of comparison (ἀηδών) as the argument of an implied circumstantial of the compared verb (as shown in Fig. 3); the phrase is therefore annotated as if it were: '[we mourned] (Predicate) as (Conjunction) a shrill-voiced nightingale (Subject, SBJ) [does/mourns] (Implied circumstantial, ADV)'.

Instead of mechanically applying those rules to similar easily identifiable constructions, students should be encouraged to reflect about the meanings that each of the elements in the graphs introduces. The edges that connect the verb of mourning to the noun and the noun to the adjective are both laden
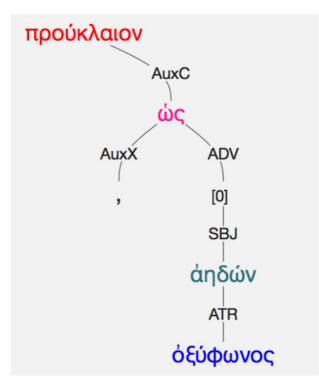

Figure 3: Like the shrill-voiced nightingale.

with a rich cultural history. The piercing voice of the wailing bird is one of the most traditional images of Greek literature, and so is the connection of the nightingale with the poetical representation of mourners. The adjective ὀξύς ('sharp', 'shrill') and derived, whether they point to the high pitch of the sound or to his piercing emotional effects, are often used for the characterization of sounds.[19] The nightingale, via the mythical paradigm of Procne, is the model for the everlasting mourn of a woman (Penelope) already in the *Odyssey*.[20] And especially in Attic tragedy, the bird is often invoked as a paradigm for the performers of dirges.[21]

The value of a treebank goes also beyond the process of annotation, even in a discussion about such questions of literary history. The AGDT includes the whole text of the *Iliad* and *Odyssey*, which notoriously provided a vast cultural repertoire of models for similes. Using the same formalism as in our passage (ὡς + implied ADV + noun), students may interrogate the treebank to extract, classify and discuss the similes in a given text.

4.2 A Polar Expression

One of the clearest features of this sentence is the coordination of a positive affirmation with the negation of its contrary ('near and not far'). This kind of antithesis is generally referred to as *polar expression*; it is typically a solemn way for a speaker to stress the reality of his utterance.[22]

Once again, a treebank provides a formalism to describe the syntax of the construction (visualized in Fig. 4), and a general corpus where comparable phrases can be searched. A polar expression, however, requires a supplementary semantic level (the two terms must have opposite meaning) which is not captured by the
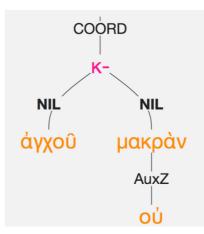

Figure 4: The polar expression.

current annotation. A query on the treebank for a coordinated phrase where one of the term in negated will probably return pertinent results along with a number of false positives. Students may be asked to search for relevant construction and identify and discuss the authentic polar expressions in their context.

4.3 How to Mourn Near and Far

If it is clear that the two adverbs are coordinated in a polar expression, it remains to be seen what their construction and exact meaning is in this particular sentence. Many readers have noted the conciseness, if not the oddity, of the expression 'to mourn near and not far off'; not surprisingly, a majority of them have tried to explain the syntax by looking for a word that is left implied and would, if supplied mentally, clarify the syntax.

The object of προκλαίω is left unexpressed: unless we suppose that the verb is used absolutely ('lament in advance'), we have to reconstruct it from the context. Logically, it is precisely the object of the Chorus' lamentation that is drawing near to the scene. Thus, some commentators take the two adverbs to refer to the (implied) object, thinking either of Herakles or of a more generic evil of the woeful situation. The former solution is argued by Hermann (1848), the latter by Jebb (1892). This reconstruction (which is reflected in the tree of Fig. 5) is by far the most widely accepted interpretation of the sentence and is
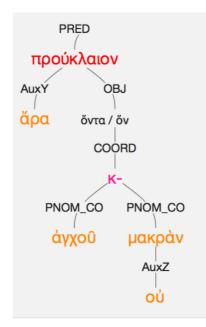
Figure 5: Hermann's interpretation.

reflected in most of the translations of the play.[23] The construction, however, is rather bold. The two passages from Sophocles that are brought as parallels by the commentators seem, as noted by Jebb, easier to understand.[24]

The plainest alternative to Hermann's construction is to attach the adverbs directly to the verb, as complements that specify the location where the main action is performed.[25] The sentence can be then paraphrased as: '(being) near and not far off (to the object of my lamentations), I mourned in advanced' and can be annotated as in Fig. 6.

Both interpretations, as phrased in most of the modern commentaries, have one point in common: they try to identify a precise lexical word that is left implied and should be supplied to govern the adverbs. Yet according to the guidelines of the AGDT this step is not mandatory. Even in case of ellipsis, all that annotators are requested to do is to mark that a word performing a certain syntactic function is missing in the text, without any obligation to explicitly identify the lexical element that is left out. Rather than being a limitation, this simpler notation can even be thought to carry a more radical interpretation of our sentence. In its turn, this interpretation has interesting implications on the way we reconstruct the communication between the actors and the audience of the play. In the original design of the Prague Dependency Treebank, from which the guidelines of the AGDT were derived, the tag for elements governed by elided words (ExD) is used as a signpost that marks the absence of an explicitly realized construction.[26] It can be liable to two different readings:
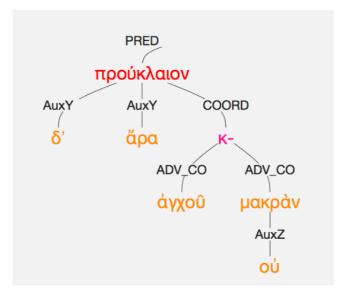

Figure 6: Circumstantials with the verb.

1. the governing element is implied from the context;
2. the element is lacking any proper construction; for example, its construction is suspended as the sentence, which started according to a certain pattern, moves abruptly toward a different organization.

Read from the standpoint of interpretation n.2, the annotation represented in Fig. 7 implies that ἀγχοῦ is left hanging, because with the addition of οὐ μακράν προύκλαιον the sentence shifts to a different structure where the first adverb has no proper place.

As a matter of fact, even this reading has already been anticipated in the critical literature. Some commentators (especially Campbell 1881) have noted that the two adverbs are not quite as well coordinated as they appear at first sight. μακράν ('for a long stretch') constructed with a verb of speaking points normally to duration in time, rather than to distance in space; so κοὐ μακρὰν προύκλαιον could mean: 'not for long did we mourn in anticipation'. According to Longo (1968), this passage should be seen as an example of 'blurring' of different constructions, where competing tendencies operate in the same sentence. Two ideas are fused together, namely: 1. that Herakles is nearer than it was imagined 2. that 'lamentation in forebode' is over and did not last for long; the transition is reached in the polar expression. The adverb 'near' brings about the formulation of the former idea. At the same time the adverb μακράν places
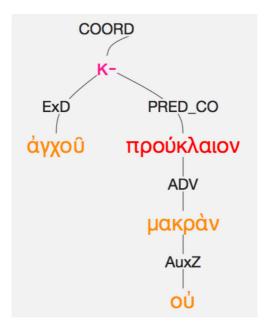

Figure 7: A broken edge?

the main accent on the latter. As a result, the first adverb is loosely attached to the main verb, to which only the second properly belongs.

Such a progressive restructuring of the syntax is often observed in spoken language, since there is no other way in which uttered sentences or sentence fragments can be corrected.[27] Yet this phenomenon is by no means exclusively a mark of hesitation or error. It can also be a strategy to unfold the meaning of a performed utterance progressively. We should certainly not forget that tragedy was a performance-oriented genre; some of the phenomena which are more typical of the spoken rather than written language should not surprise us even in the carefully polished poetical language of the Athenian dramatists. Other cases of sentences whose syntax is structured progressively, and yet very carefully, are indeed often found in Sophocles.[28] This interpretation points to a tension between the syntax of the coordination and the semantics of the two adverbs that is certainly operative in the sentence. And yet it is equally undeniable that, however 'perturbed' by competing structures and in spite of any 'false start' that the first adverb suggests, the words do reach a coherent syntactic construction centred around the coordination of ἀγχοῦ and μακράν. This is a crucial difference with the model of sentence restructuring typical of oral conversation (and diagrammed in the tree of Fig. 7), where the starting elements are obliterated or left unrealized. We would certainly go too far if we posited that ἀγχοῦ is left without a construction. My personal preference, therefore, goes to the interpretation of Fig. 6, even if the sort of zeugma that can be seen in the construction of the verb with the two adverbs cannot be properly captured in it.[29]

4.4 Complex Syntax in a Class Environment

The previous discussion involves a level of complexity and subtlety that might be suitable only for advanced students in Classical philology. However, two crucial points must be stressed.

Firstly, annotators must be encouraged to notice how even the smallest change in the collocation of the words within the sentence tree or in the use of the AGDT labels for the syntactic relations is going to affect dramatically the general interpretation of the sentence.

Secondly, students can be fruitfully reminded that most of the different reconstructions that they can obtain by moving around words in Arethusa and attaching them to different parts of the trees are likely to be already attested in the history of the interpretation of the text they are annotating. Students should be always invited to investigate the commentaries in search to alternative ways of structuring a sentence, and of different arguments to argue either for or against some of the possible reconstructions.

These two steps can be attempted in both direction: starting from the original interpretations of the students to find the precursors in the previous criticism,

or from the history of criticism to an original reading. Yet they both form the indispensable steps toward a fully informed critical annotation.

5 Conclusions

Soph. *Trach*. 962–3 has confronted us with simple linguistic tasks (such as identifying the correct morphological interpretation of προὔκλαιον) and more complex interpretative problems; in cases as such, and in most cases when reading Greek tragedies, the construction of a syntactic annotation of a sentence should be seen more as an open process than a mere application of a series of grammatical rules. Interpretations like those reflected in the trees of Figs. 5, 6, and 7 can be (and in fact, as we saw, have been) defended with good arguments. This situation, which is certainly peculiar of treebanks of ancient literary texts, seems to defy the notion itself of a reference treebank: how could a corpus that allows so much space for conflictual interpretation be used as a research tool to investigate linguistic phenomena?

Several answers can be addressed to these sceptical remarks. On the one hand, we can observe that, for one very controversial point in the reading of the sentence, our treebank annotation records several indisputable facts that contribute positively to the advancement of the resources available for the study of Greek. Such facts include the morphology of the words, the lemmatization, or the syntactic annotation of certain syntactic structures, like the similitude introduced by ὡς; other sentences, no matter how controversial in decisive details they might be, would also include subjects, direct objects or other words whose construction would not pose the minimal problem to readers. If similar pieces of information seem trivial within an eleven word sentence, at the scale of the whole corpus of Sophoclean tragedy (let alone the 5th-century poetry or the whole Greek literature) the impact increases exponentially. Thanks to the students that have annotated every words in the Greek texts they were reading, the AGDT already provides enough evidence to conduct comprehensive studies on e.g. the usage of the nominative in Aeschylus, Sophocles, and Homer.[30]

But in parallel to the 'distant reading' that the massive quantitative evidence of the treebank allows, I hope that my discussion has shown that linguistic annotation encourages the work of critical 'close reading' of ancient texts in their original language.[31] The problems that the annotators will face are indeed the same that Gottfried Hermann or even the ancient *scholia* speculated about. The application of treebank annotation in the class is a crucial opportunity to discuss the methods that constitute the most vital legacy of Classical Philology. Linguistic annotation challenges us to find a solution for passages that are often problematic and then to encode it in a well-defined formalism that can be read, compared, and criticized by all that are familiar with the same annotation schema, across every barrier of language or culture.

Notes

1   AGDT: <http://nlp.perseus.tufts.edu/syntax/treebank/>.
2   A version 2.0, with some revisions and new features, has been announced. More annotated texts are constantly published in the GitHub repository of the Perseus Digital Library: <https://github.com/PerseusDL/treebank_data>.
3   See e.g. Smyth 1920: 265; Kühner, Raphael & Bernard Gerth 1898: 77–82.
4   On the subject, see now the treebank-based analysis of Mambrini & Passarotti 2016.
5   Mambrini & Passarotti 2012.
6   Bamman et al. 2009: 7.
7   For a discussion of a case study on the use of treebank annotation in language and linguistics classes see Gerdes 2013.
8   The guidelines for syntactic annotation of the AGDT can be accessed online at: <http://nlp.perseus.tufts.edu/syntax/treebank/agdt/1.7/docs/guidelines.pdf>.
9   See Gerdes 2013 and the paper of Rydberg-Cox in this volume for a remarkable example of similar applications.
10   The work of annotation starts from the digital text published in the Perseus Digital Library. In this format, the seven extant tragedies of Sophocles consist of 5,973 sentences, which is 71,690 words. Thanks to the integration of Perseus' annotation framework within the Perseids editing environment (see the paper by Beaulieu and Almas in this volume), annotators are free to make modification to the Greek texts they are working with.
11   Perseids: <http://sosol.perseids.org/tools/arethusa/app/#>.
12   A fundamental point of departure for the dependency grammar is the work of Tesnière 1959. In particular, the AGDT is closely inspired by the analytic annotation of the *Prague Dependency Grammar* of Czech, whose theoretical foundations are laid in Sgall et al. 1986.
13   Burton 1980.
14   Others (e.g. Ellendt 1872 *s.v.*) take προκλαίω to mean 'mourn openly, publicly'. This translation is indeed entirely possible; the other interpretation of the pre-verb πρό, however, fits much better in the dramatic context that we are highlighting.
15   In Fig. 2, the data from the morphological analyser Morpheus (Crane 1991) are preloaded in the editor for the annotators to choose. If one prefers, this configuration of Arethusa can be deactivated, so that the identification of all the admissible morphological interpretations can be entirely left to the students. In a class, students that are already familiar with the conjugation of the Greek imperfect will eventually come to (or be guided to) the same conclusions that are output here by Morpheus.
16   Kaminio 1970.
17   Cf. Xenis 2010: 214; note that πλήθους is Nauck's easy correction for πλῆθος transmitted by the manuscripts.

[18] Hermann 1848.

[19] Kaimio 1977: 174−81, 227−8.

[20] *Od.* 19. 518−23, with full commentary in Nagy 1996: 7−38, 39−58.

[21] See e.g. Soph., *El.* 103−9 and the other examples discussed by Loraux 1999.

[22] See Bruhn 1899, 118−9, and Easterling 1982 *ad* S. *Tr.* 234−5.

[23] Se e.g. Jebb: 'Ah, he [which implies Herakles as the object] was not far off, but close to us, he for whom I cried in advance"; or Lloyd-Jones 1994: 'So when I lamented like the shrill-voiced nightingale, it was for what was near, not what was far [which implies the neuter ὄν as object]'. The same interpretation can also be found in Easterling 1982.

[24] In Soph. *Ph.* 26 μακράν is constructed predicatively with the object of the verb: τοὔργον οὐ μακρὰν λέγεις, i.e. 'not far off is the task that you speak of'. This relation is easier to understand precisely on account of the presence of an object. In Soph. *fr.* 210.38−9 (quoted by Davis 1991), the object is not expressed: ἀγχοῦ προσεῖπας (sc. τὸν δαίμονα)· οὐ γὰρ ἐκτὸς ἑστὼς συρεῖ δὴ φύρδαν. But the προσ- in the verb gives in any case a sense of direction: one can easily be said to 'address near', i.e. to the vicinity (cf. also Soph. *fr.* 380: ἀγχοῦ προσῆψεν).

[25] This is construction is fully compatible with the interpretation of Kamerbeek 1970, which suggests to take the adverbs with the implied οὖσαι.

[26] The guidelines of the PDT for syntactic annotation can be read online at: <http://ufal.mff.cuni.cz/pdt2.0/doc/manuals/en/a-layer/html/index.html>.

[27] This phenomenon is expressed by a famous quotation from Hor. *Ars Poetica* 390: *nescit vox missa reverti*, a word once uttered cannot come back.

[28] This kind of interpretation is privileged in the aforementioned commentary of Longo 1968: for examples in the *Trach.*, see his index on p. 418 under 'sovrapposizione (fusione, contaminazione) di costrutto' ('overlapping, or fusion, contamination between constructions'). To these passages, we may add a sentence in the exchange between Athena and Odysseus in the *Ajax* (42−3); to the question τί δῆτα ποίμναις τήνδ᾽ ἐπεμπίπτει βάσιν; ('why did he fall in this assault upon the flocks?'), the goddess answers: δοκῶν ἐν ὑμῖν χεῖρα χραίνεσθαι φόνῳ ('because upon you he thought that he was staining his hands with murder'). As my translation of line 43 suggests, the phrase ἐν ὑμῖν ('upon you') is not easy to construe with the following χεῖρα χραίνεσθαι φόνῳ, which appear concluded in itself. His initial position in the clause suggests rather that it may be influenced by the common construction of πίπτειν + ἐν ('fall upon') and dative, prompted by the ἐπεμπίπτειν in the question.

[29] This is also the interpretation that I have adopted for my annotation of the *Women of Trachis* for the AGDT, which is represented in Fig. 1.

[30] Seminal works on this direction are already being produced: see e.g. Mambrini & Passarotti 2016 on agreement patterns with coordinated subjects.

[31] On the notion of 'distant' and 'close' reading see most recently Jockers 2013.

References

Bamman, D., Mambrini, F. & Crane, G. (2009). An ownership model of annotation: The Ancient Greek Dependency Treebank. In *Proceedings of the Eighth International Workshop on Treebanks and Linguistic Theories (TLT 8)*, Milan: EDUCatt, pp. 5–15.

Bruhn, E. (1899). Anhang. In *Sophokles*, Erklärt von F. W. Schneidewin und A. Nauck. Achter Band: Oedipus auf Kolonus. Berlin: Weidmann.

Burton, R. W. B. (1980). *The Chorus in Sophocles' Tragedies*. Oxford: Clarendon Press.

Campbell, L. (1881). *Sophocles. Edited with English notes and introduction. Vol. II: Ajax. Women of Trachis. Electra. Philoctetes. Fragments*. Oxford: Oxford University Press.

Crane, G. (1991). Generating and parsing classical Greek. *Literary and Linguist Computing* 6(4): 243–245.

Davis, M. (1991). *Sophocles. Trachiniae*. Oxford: Oxford University Press.

Easterling, P. E. (1982). *Sophocles. Trachiniae*. Cambridge: Cambridge University Press.

Ellendt, F. (1872). *Lexicon Sophocleum. Editio Altera. Curavit Hermannus Genthe.* Berolini: Borntraeger.

Gerdes, K. (2013). Collaborative dependency annotation. In *Proceedings of the Second International Conference on Dependency Linguistics (DepLing 2013)*, Prague: Matfyzpress, pp. 88–97.

Hermann, G. (1848). *Sophoclis Trachiniae* (2nd Edn.). Lipsiae: Apud Ernestum Fleischerum.

Jebb, R. C. (1892). *Sophocles. The Plays and Fragments. Part V. The Trachiniae*. Cambridge: Cambridge University Press.

Jockers, M. L. (2013). *Macroanalysis: Digital Methods and Literary History*. Chicago: University of Illinois Press.

Kaimio, M. (1970). *The Chorus of Greek Drama within the Light of the Person and Number Used*. Helsinki: Societas Scientiarum Fennica.

Kaimio, M. (1977). *Characterization of Sound in Early Greek Litterature*. Helsinki: Societas Scientiarum Fennica.

Kamerbeek, J. C. (1970). *The Plays of Sophocles. Commentaries. Part II: The Trachiniae* (2nd ed.). Leiden: Brill.

Kühner, R. & Gerth, B. (1898). *Ausführliche Grammatik der griechischen Sprache. Zweiter Teil: Satzlehre. Erster Band*. Hannover: Hahnsche Buchhandlung.

Lloyd-Jones, H. (1994, January). *Sophocles, Volume II. Antigone. The Women of Trachis. Philoctetes. Oedipus at Colonus*. Loeb Classical Library. Cambridge, MA: Harvard University Press.

Longo, O. (1968). *Commento linguistico alle Trachinie di Sofocle*. Padova: Antenore.

Loraux, N. (1999). *La voix endeuillée*. Paris: Gallimard.

Mambrini, F. & Passarotti M. (2012). Will a parser overtake Achilles? first experiments on parsing the Ancient Greek Dependency Treebank. In *Proceedings of the 11th Workshop on Treebanks and Linguistic Theories (TLT11)*, Lisbon (pp. 133–144). Lisbon: Colibri.

Mambrini, F. & Passarotti M. (2016). Subject-Verb Agreement with Coordinated Subjects in Ancient Greek. A Treebank-based Study. *Journal of Greek Linguistics* 16, forthcoming.

Nagy, G. (1996). *Poetry as Performance. Homer and beyond*. Cambridge: Cambridge University Press.

Sgall, P., Hajičová, E. & Panevová, J. (1986). *The Meaning of the Sentence and Its Semantic and Pragmatic Aspects*. Dodrech: Academia/Reidel Publishing Company.

Smyth, H. W. (1920). *A Greek Grammar for Colleges*. New York: American Book Company.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris: Klinksieck.

Xenis, G. (2010). *Scholia vetera in Sophoclis "Trachinias"*. Berlin: De Gruyter.