Modal markers in co-occurrence: a study on Terence's comedies

Paola MARONGIU
Institut des sciences du langage (ISLa), University of Neuchâtel
paola.marongiu@unine.ch

This paper presents a corpus-based study on the co-occurrence of lexical modal markers expressing possibility or necessity in the six comedies written by Terence. Although the co-occurrence of modal markers has long been discussed in the literature, a systematic and comprehensive study of this phenomenon is still lacking, and studies on Latin are almost absent. In this paper I will first provide a quantitative overview of the phenomenon, discussing my methodological approach and challenges linked to it. Then I will delve into the qualitative analysis of the passages with co-occurrence, focusing on both the interactions observed between types of modality and the use of the co-occurrence for argumentative purposes in specific contexts.

1. Introduction

1.1 A definition of modality

Providing a clear-cut definition of modality is not a trivial task. This notion has been and still is subject to discussion within the research community. Nuyts and van der Auwera (2016) provide an extensive overview on the history of the study of modality and of the different systematisations produced by various theoretical approaches. For the sake of this study, I will define modality as the association of the notional domains of possibility, necessity, probability and volition with a representation, called the State of Affairs (SoA).[1] This association describes the attitude of the speaker towards the representation itself. In the sentence *Mary may be home,* the speaker expresses his/her belief that there is a possibility that Mary is at home. The speaker encodes this judgment using the marker *may* to qualify the representation of Mary being home as possible (Dell'Oro 2023: 6). The portion of text to which a marker refers, and which (usually) describes a SoA, is called scope. The definition of modality and the theoretical framework adopted for this study are borrowed from the WoPoss project (Dell'Oro et al. 2020; Dell'Oro 2023; Bermúdez-Sabel et al. 2024), which draws on the functional approach described by Nuyts (2016). The project aims at studying the diachrony of modality in Latin by annotating modal passages in a diachronic Latin corpus, and in order to meet this goal, it has developed a fine-grained

---

[1] This work is a by-product of my doctoral thesis, in which I investigate the co-occurrence of modal markers at the sentence level in a corpus of classical Latin texts. The method outlined in section 3 and a substantial portion of the qualitative analysis are, therefore, adapted from a broader study on this phenomenon in a later stage of the Latin language. My PhD was part of the project *WoPoss: A world of possibilities. Modal pathways over an extra-long period of time: the diachrony of modality in the Latin language*. WoPoss was funded by the Swiss National Science Foundation (SNSF project n° 176778) for the period from February 2019 to January 2023.

annotation schema which includes 23 subtypes of modality. In this study I adopt a simplified version of the WoPoss annotation schema for the analysis of the modal passages. The main types of modality that I annotated are dynamic, deontic authority, deontic acceptability and epistemic.

Dynamic modality is defined as the expression of a necessity or a possibility enabling or compelling the main participant who is engaged in the SoA.

- Dynamic: *He **can** stand on his head without using his hands.*[2]

Deontic modality is associated to two subtypes: deontic authority and deontic acceptability. Deontic authority can be defined as the expression of a permission or an obligation addressed to the participant by some source of authority. Deontic acceptability expresses an evaluation on the degree of moral desirability of the SoA.

- Deontic authority: *You **may** go now.*

- Deontic acceptability: *We **cannot** fire him just like that, he's been our best employee for years.*

Finally, epistemic modality expresses an evaluation of the SoA made by the speaker in terms of likelihood.

- Epistemic modality: *That's **probably** the postman bringing today's newspaper.*

1.2   The co-occurrence of modal markers

The co-occurrence of modal markers is a field of study which is still under-explored and, to my knowledge, there is no systematic analysis of the co-occurrence of modal markers in Latin at the present time. One of the first mentions of this phenomenon in literature is found in Halliday (1970: 331), who (marginally) presents the case of co-occurrence of modal adverbs scoping over a modal verb, e.g. ***Perhaps** he **might** have built it.* After Halliday, the theoretical discussion on the topic was developed by Lyons (1977), Sueur (1978) and Palmer (2001), but it still only focused on epistemic modality. On the other hand, there have been some empirical corpus-based studies that considered other types of modality, such as Huot (1974), Coates (1983), Nuyts (2004), Narrog (2009), Kratochvílová (2018) and Hütsch (2020).[3]

Prior to performing this study, I established some parameters to define the co-occurrence of modal markers. Co-occurrence is observed when at least two lexical modal markers expressing the notions of possibility or necessity are

---

[2]   All the examples presented for the types and main subtypes of modality are taken from Nuyts (2016). Throughout the paper, the markers relevant for the analysis will be indicated in bold.

[3]   It is worth also mentioning Sueur (1975) and Chu (2008), who worked on the co-occurrence of modal markers (mainly adverbs and verbs) in French, although their work does not fall into the category of corpus-based studies as they worked with acceptability tests.

found within a portion of text delimited by a strong punctuation sign (see examples 1 and 2):

1) I **have to** go now, but we **can** meet on Thursday.

2) He **probably will be able to** come home for Christmas.

As it appears from examples 1 and 2, the markers can co-occur in different syntactic structures: in example 1 the markers co-occur in two coordinate clauses; in example 2 the adverb *probably* modifies *be able to* in a simple clause. The type of syntactic structure can impact the scopal relation that is established between the markers (in example 2, the adverb *probably* includes in its scope the co-occurring marker *be able to*,[4] whereas the two markers have separate scopes in example 1). The aim of this study is to determine I) how extensive the co-occurrence of modal markers is in Terence II) how many and which lexical markers appear in co-occurrence in Terence III) what the relation between the co-occurring markers is, from both a syntactic and a semantic viewpoint.

1.3   Modality in Latin and in Terence

The investigation of the expression of modality in Latin counts some important contributions. Bolkestein (1980), Núñez (1991), and Fruyt and Moussy (2002) have published relevant volumes on this topic. Bertocchi and Orlandini (2001), Fruyt and Orlandini (2003), Fruyt (2004), Magni (2005, 2010), among others have contributed to the study of specific aspects of Latin modal markers (both lexical and morphological, including verbal mood), both in synchrony and diachrony.

However, the description of the co-occurrence of modal markers in Latin is not as developed. If the literature on this phenomenon in other languages leaves room for improvement, there is absolutely no dedicated study in Latin to my knowledge (marginal observations can be found in Urbanova and Cuzzolin 2016: 333, for example, on the co-occurrence of deontic *possum* with verbs such as *permitto* 'to allow').

Similarly, specific studies on the expression of modality in the production of Terence are missing. Terence has composed in the 2nd century BC six comedies: *Andria* (166 BC), *Hecyra* (165 BC), *Heauton Timoroumenos* (163 BC), *Eunuchus* (161 BC), *Phormio* (161 BC), *Adelphoe* (160 BC).[5] The studies on the language of Terence are numerous: we can recall Bléry (1909), Vairel (1980), Karakasis (2005), Smith (2005), de Melo (2007), Pezzini (2015)*,* Halla-

---

[4]   I will refer to this type of interaction as 'scope hierarchisation', see section 4.

[5]   I propose in this paper the widest accepted chronology for Terence's production (see Conte 1994; Hanchey 2013: 29 among others). However, it should be specified that the dating of Terence's comedies is debated (Hanchey 2013: 127-130).

Aho (2018) among others.[6] However, with the exception of Vairel (1980), Terence is only mentioned in works with a broader scope on the expression of modality throughout Latin, and often together with Plautus (e.g. Núñez 1991; Fruyt and Moussy 2002).

In this paper I will present the co-occurrence of lexical modal markers in the language of Terence, so as to try to fill the gap by proposing a method of analysis and some preliminary observations. I will give an overview of the type of data with their curation and processing in section 2; in section 3 I will present the results of the extraction of the co-occurrences and describe them in a quantitative perspective, also discussing the limitations encountered during the analysis; finally, in section 4 I will provide an in-depth qualitative analysis of the co-occurrences, resorting to the results presented in section 3 and discussing the argumentative and/or semantic patterns that underlie specific types of co-occurrence.

2.   Data description

The retrieval of the co-occurrences was based on a list of targeted modal markers to be searched in a corpus containing all of Terence's comedies.

The list of modal markers is largely based on the one used within the framework of the WoPoss project (Dell'Oro 2023: 8-9), and contains the following lexical modal markers (or constructions): *aequus* 'right, appropriate', *aptus* 'apt to', *certe* 'certainly', *debeo* 'I have to', *decet* 'It is appropriate', *dubie* 'doubtfully', *dubium* 'doubtfully', *dubius* 'doubtful' (or *dubium est* 'it is doubtful'), *facultas* 'possibility, ability', *forsitan* 'maybe', *fortasse* 'maybe', *forte* 'maybe', *(il)licitus* '(not) allowed', *(in)certus* '(un)certain', *indubitanter* 'undoubtedly', *indubitate* 'undoubtedly', *ineptus* 'not apt to', *iniquus* 'not right, not appropriate', *ius est* 'it is right, it should', *licet* 'It is allowed, possible', *meum est* 'I can/it is my duty', *necessario* 'necessarily', *necessarium* 'necessarily', *necessarius* 'necessary', *necessarie* 'necessarily', *necesse (est)* '(it is) necessary', *necessitas* 'necessity', *necessitudo* 'necessity', *(ne)queo* 'I can(not)', *oportet* 'It ought', *opus est* 'it is necessary', *possiblilis* 'possible', *possibilitas* 'possibility', *possibiliter* 'possibly', *possum* 'I can, I am able', *potestas* 'power, possibility', *probabilitas* 'probability', *probabiliter* 'probably', *valet* 'to have the ability or power to', *usus est* 'it is necessary'. Many of these markers are polyfunctional (see for example Magni 2010: 212) i.e., they can express more than one type of modality. For instance, *debeo* or *possum* can express dynamic, deontic (authority or acceptability) or epistemic modality depending on the context in which they occur. Other markers e.g. *forte, dubius* (with *dubium est* and *dubie*), *certum est* (with *certe*), are typically associated with epistemic modality (see the modal maps by Marongiu

---

[6]   See also Augoustakis & Traill (2013) for an overview of Terence's production.

and Dell'Oro 2022 for an overview of the markers and their modal meanings with reference to lexicographic sources).

The corpus contains the six comedies written by Terence, for a total of 56k tokens (see Table 1). The text used to build the corpus was retrieved from the Perseus Digital Library (Smith et al. 2000) and corresponds to the edition by Parry (1857). The texts were automatically annotated using the model IT-TB of the parser Stanza (Qi et al. 2020), developed by the Stanford NLP group.[7] In order to retrieve the co-occurrences I relied on the results of the lemmatisation and sentence segmentation tasks, looking for at least two of the lemmas listed above in every sentence of the corpus. The retrieval of the co-occurrences and part of the quantitative analysis, including the visualisations, were performed using a combination of Python scripts, a workflow in Orange Textable (Xanthos 2014) and the visualisation software Cytoscape (Shannon et al. 2003).[8]

Title of the comedy	N° of tokens
Adelphi	8941
Andria	9049
Eunuchus	9869
Heauton Timorumenos	9708
Hecyra	9041
Phormio	9491
Total	**56099**

Table 1: Number of tokens per comedy and in total in the corpus

The sentences extracted from the corpus were then manually annotated for the following features:

- the polarity of the clause in which the markers appear;[9]

- if the marker is modal or not, and the type of modality expressed by the marker in context;

- the comedy from which the sentence was extracted;

- the type of syntactic relation between the co-occurring markers (distinguishing between complex clauses, related by coordination or subordination, and simple clauses where an adverb modifies another

---

[7]   The IT-TB is not the only model available in Stanza for the automatic annotation of Latin texts. As the retrieval of the co-occurrences mainly relies on their lemmas, I chose the model with the highest performance on the lemmatisation task (99%).

[8]   The code is available on GitHub at https://github.com/paoma370/Co-occurrences-in-Terence/tree/main

[9]   The annotation of polarity is based on the annotation performed in the WoPoss project: the clause in which a marker appears can have a positive or negative polarity, and negation must be expressed by an independent lexical unit e.g. *non* 'not' (Dell'Oro 2023: 13).

constituent, usually a main verb, see section 3.4 for examples from the corpus);

- the coordinating or subordinating particle that participates in the syntactic structure, if any (e.g. coordinating particle *et* 'and', subordinating particle *quia* 'since, because');

- if one of the co-occurring markers scopes over the other (see example 2) i.e., if there is scope hierarchisation;

- the semantic scope of each co-occurring marker.

3.   A quantitative overview of the co-occurrences

Table 2 shows the number of sentences which host the co-occurrence of modal markers, against the total number of sentences counted in the corpus. The sentences are determined based on the results of the sentence segmentation performed by the parser, and correspond to a portion of text delimited by a strong punctuation sign.

Number of sentences with 0 modal markers	9760	94.40%
Number of sentences with 1 modal marker	550	5.32%
Number of sentences with at least 2 modal markers	29	0.28%
Total number of sentences	**10339**	**100%**

Table 2: Sentences in the Terence corpus with respect to the presence of modal markers.

In the Terence corpus the number of sentences with the co-occurrence of at least two (lexical) modal markers is only 29. This is not entirely surprising: this phenomenon was also proven to be quite limited in Latin in another similar study performed on a bigger corpus-150k tokens-of Classical Latin texts (see Marongiu 2021). However, while the rate of co-occurrence of lexical modal markers reached the value of 2.4% in the Classical Latin corpus, this percentage lowers to 0.28% in Terence. This phenomenon is therefore extremely rare in Terence's productions, but this is not completely unexpected if we look at the sentence length and syntactic complexity values shown and discussed by de Melo (2007: 103) and Halla-Aho (2018: 78). As I will show in section 4, the co-occurrence of modal markers is often found in combination with argumentative patterns that require a certain sentence length and syntactic complexity. De Melo and Halla-Aho show that the average sentence length in Terence's plays is only 6.34 words per sentence, whereas all the sentences with co-occurrence retrieved in the corpus exceed this value. Therefore, the rarity of this phenomenon seems to be related to the rarity of long, complex sentences throughout Terence's production.

3.1   Co-occurring markers

The markers that co-occur at least once in the corpus are listed in Table 3. This Table shows both markers that are modal, and markers that turned out to be not modal after a qualitative analysis. As explained in Section 2, the co-occurrences extracted from the corpus were manually analysed to determine, among others, if the markers were modal and what type of modality they show in the passage. In Table 4 I filtered out all cases in which one or both markers were not modal, and kept only the markers showing a modal value in the co-occurrences. Consider example 3.

3) CH. *Mea quidem hercle* **certe** <u>**in dubio**</u> *vita est* (Andria, verse 347).

  CH. But my life is <u>at risk</u> **for sure**.

This line is uttered by the character Charinus, overtly devastated by the idea that he will never be able to marry his beloved Philumena, as her father has already promised her in marriage to his friend Pamphilus. In this sentence the markers *certe* and *dubius* co-occur. *Certe* is modal in this context, and it expresses epistemic certainty. However, *dubius,* in the locution *in dubio*, simply means 'at risk' (i.e., 'in doubt' because it is at risk). This co-occurrence is therefore excluded from Table 4.

Markers (both modal and not modal)	**Count in co-occurrence**	**Count in corpus**	**Co-occurrence rate**
dubius	4	16	25,00%
licet	10	42	23,81%
iniquus	3	14	21,43%
aequus	10	63	15,87%
incertus	2	14	14,29%
certe	1	10	10,00%
forte	2	22	9,09%
potestas	1	11	9,09%
possum	19	218	8,72%
certus	3	38	7,89%
opus	4	59	6,78%
decet	1	15	6,67%
queo	1	27	3,70%
oportet	1	28	3,57%

Table 3: Markers in co-occurrence (both not modal and modal).

Markers (only modal)	Count in co-occurrence	Count in corpus	Co-occurrence rate
licet	10	42	23,81%
dubius	2	16	12,50%
potestas	1	11	9,09%
opus	4	59	6,78%
decet	1	15	6,67%
possum	14	218	6,42%
aequus	3	63	4,76%
forte	1	22	4,55%

Table 4: Markers in co-occurrence (only modal co-occurrences).

The column 'Markers' in Tables 3 and 4 lists the markers that occur together with another marker in the corpus. The column 'Count in co-occurrence' shows the marker count in co-occurrence i.e., how many times the marker is found in co-occurrence with another one. The column 'Count in corpus' shows the marker count in the corpus i.e., how many times the marker occurs in the corpus regardless of its co-occurrence with other markers. The column 'Co-occurrence rate' shows the marker co-occurrence rate i.e., the ratio of the marker count in co-occurrence to the marker count in the corpus. The relevance of these measures will be discussed later in this section. For now, let us focus on the column 'Marker' in Tables 3 and 4.

Table 3 shows that, among the 40 markers in the list, only 14 appear at least once in co-occurrence (before the modal analysis of the passages). The absence of some of the markers from Table 4 was expected. The diachronic maps of modal markers produced within the WoPoss project (Marongiu and Dell'Oro 2022) show that *(il)licitus*, *possibilis*, *possibilitas* and *probabilis* all emerged, according to lexicographic sources, after the 2nd century BC. Therefore these four modal markers (and their negative counterparts) do not occur at all in the comedies. Others—*certe, facultas, forsitan, ius est, necessario, nequeo*—occur in the corpus, and they can be modal, but they never co-occur with another marker. Consider example 4.

4) THA. *Ego impetrare **nequeo** hoc abs te, biduum saltem ut concedas solum?* (Eunuchus, verses 181-182)

 THA. I **cannot** obtain this from you, not even that you grant me just two days?

The speaker Thais is expressing her despair, as she is unable (*nequeo*) to convince her lover Phaedria to leave her alone for a couple of days. She wants to accomplish her plan of recovering her friend Pamphila from an army officer who keeps her in his household. Here *nequeo* is modal, and expresses dynamic modality, but it does not co-occur with any other lexical modal marker.

Finally, the markers *debeo, valet, aptus* and *necessitas* also occur in the corpus but never show a modal value. An illustration is given in example 5 for *debeo* (I include the left context for clarity)*.*

5) DAV. *Accipe, hem, lectum est: conveniet numerus quantum **debui.*** (Phormio, verses 52-53)

   DAV. Here, take it, it has been examined: the number (of coins) will correspond to the amount **I owed** you.

At the beginning of the comedy *Phormio*, the slave Davus has collected and is giving his friend Geta the money that he owed him. In this context, *debeo* means 'to owe something (to someone)' and therefore it is not modal (but more specifically premodal following van der Auwera and Plungian 1998).

Let us now compare the column 'Marker' in Tables 3 and 4. Firstly, the markers *certe, certus, incertus, iniquus* and *queo* disappear from Table 3 to Table 4. The case of *certe* was already illustrated in example 3, in which *certe* is modal but the marker with which it co-occurs is not. Another example is the use of *certus, incertus* and *iniquus* as adjectives. In such cases, they are always annotated as not modal, because their scope does not describe a SoA. Consider example 6.

6) *Quia scibam **dubiam** fortunam esse scenicam, spe **incerta certum** mihi laborem sustuli.* (Hecyra, verses 16-17)

   As I knew that fortune in theatre is **unsteady**, with **hesitating** hope I endured **certain** toil.

This sentence is part of the second prologue of the *Hecyra*: an actor introduces the play on behalf of its author, and while addressing the audience he refers to the acting career as challenging and unpredictable. In this context, the three potential markers *dubius, incertus* and *certus* do not modalise a SoA: they are rather used as adjectives that qualify a noun, respectively *fortunam* 'fortune'*, spe* 'hope' and *laborem* 'toil, work'. As such, they are annotated as not modal.

To delve deeper into the comparison between Table 3 and Table 4, we can now proceed to the analysis of the data in the columns 'Count in co-occurrence', 'Count in corpus', and 'Co-occurrence rate'. In order to show the relevance of the co-occurrence rate measure in this type of study let us consider the example of *possum.* In Table 3, this is the marker with the highest count in co-occurrence (19). However, it is also the marker with the highest count in corpus (218 occurrences). Its overwhelming presence in the corpus noticeably lowers the co-occurrence rate, assigning it the same value as *potestas* (always in Table 3), which only appears once in co-occurrence. The opposite case is shown for example by *licet,* a relatively rare marker with respect to *possum,* which occurs in the corpus only 42 times, compared to the 218 occurrences of *possum,* but has a much higher co-occurrence rate. The examples of *possum* and *licet* show the importance of considering both co-occurrence count and co-occurrence rate as relevant measures of co-occurrence, especially when dealing with such

evident cases of data sparsity.[10] The count in co-occurrence accounts for how many times a certain marker will be found in the co-occurrences retrieved in the corpus (the higher the number of occurrences, the higher the chances for the marker to appear in co-occurrence); the co-occurrence rate, however, gives an indication of how likely it is to find a specific marker in co-occurrence (see Narrog 2009: 172-175). By relying on the count in co-occurrence alone, one knows that *possum* is the most present marker in the co-occurrences, although this is due to the extensive presence of the marker in the corpus in general. But considering the number of occurrences of *possum* in the corpus, the number of co-occurrences in which this marker appears is not extremely high. Indeed, other markers in Table 3 (e.g. *licet* and *aequus)*, have much lower counts in co-occurrence and in the corpus but show a co-occurrence rate more than twice as important as that of *possum* (*licet* co-occurs with another marker for roughly a fourth of its occurrences in the corpus)*.

Having clarified the relevance of these measures for this study, let us return to the comparison between Tables 3 and 4. In Table 3 the first markers in terms of co-occurrence rate are *dubius*, *licet, iniquus* and *aequus*. In Table 4, after the qualitative analysis of modal passages, *licet* and *dubius* still appear on top of the ranking, but their behaviour is completely different. While the co-occurrence rate of *licet* remains exactly the same, the co-occurrence rate of *dubius* is halved, leaving the top position to *licet*. A similar behaviour is displayed by *aequus* and *forte* with respect to *dubius.* As shown in example 6*, the drop in co-occurrence rate for *dubius* and *aequus* is due to their frequent use as modifiers to a noun, and therefore to the absence of a modal meaning in such cases. The co-occurrence rate of *possum* slightly lowers, but not to the extent of the previous markers. All the remaining markers *licet, potestas, opus* and *decet* maintain their co-occurrence rates stable. These changes impact the final ranking of co-occurrences with exclusively modal markers in the following way. *Licet* becomes the marker with the highest co-occurrence rate in the corpus in Table 4. It records a co-occurrence rate of almost the double compared to the second-ranked marker, *dubius* (23,81% against 12,50%), and more than the double compared to all the other markers that follow. This is especially interesting because in Table 3 *licet* co-occurs 10 times, while the other markers co-occur from 1 to 4 times, with the exception of *aequus* and *possum* (10 and 19 respectively). *Licet* is the only marker among those with a high co-occurrence count to maintain this value steadily after the manual analysis.

---

[10]   Data sparsity describes a phenomenon by which the data sample, in this case the corpus, provides a limited amount of data to carry out a specific analysis, therefore making it highly difficult or impossible to identify true distributions and patterns within the data (Allison et al. 2006: 327). In this specific case, this translates to a low number of observations and a high number of zero values among the co-occurrences.

3.2   The types of modality

The counts in co-occurrence for each type of modality are shown in Table 5. All the types of modality described in section 1.1 are present in at least one co-occurrence. Dynamic modality stands out with the highest count in co-occurrence, outnumbering by far all the other types of modality. Deontic acceptability, deontic authority and epistemic modality show a similar behaviour, with counts in co-occurrence between 3 and 4.

Type of modality	Count in co-occurrence
deontic acceptability	4
deontic authority	3
dynamic	26
epistemic	3

Table 5: Counts in co-occurrence for each type of modality.

Let us now look at the combinations between the modal markers and the types of modality i.e., the type of modality expressed by each marker in the co-occurrences. This is shown in Table 6, and the corresponding heatmap is given in Figure 1.

	Deo. Accept.	Deo. Auth.	Dyn.	Epist.
aequus	3	0	0	0
decet	1	0	0	0
dubius	0	0	0	2
forte	0	0	0	1
licet	0	2	8	0
opus	0	0	4	0
possum	0	1	13	0
potestas	0	0	1	0

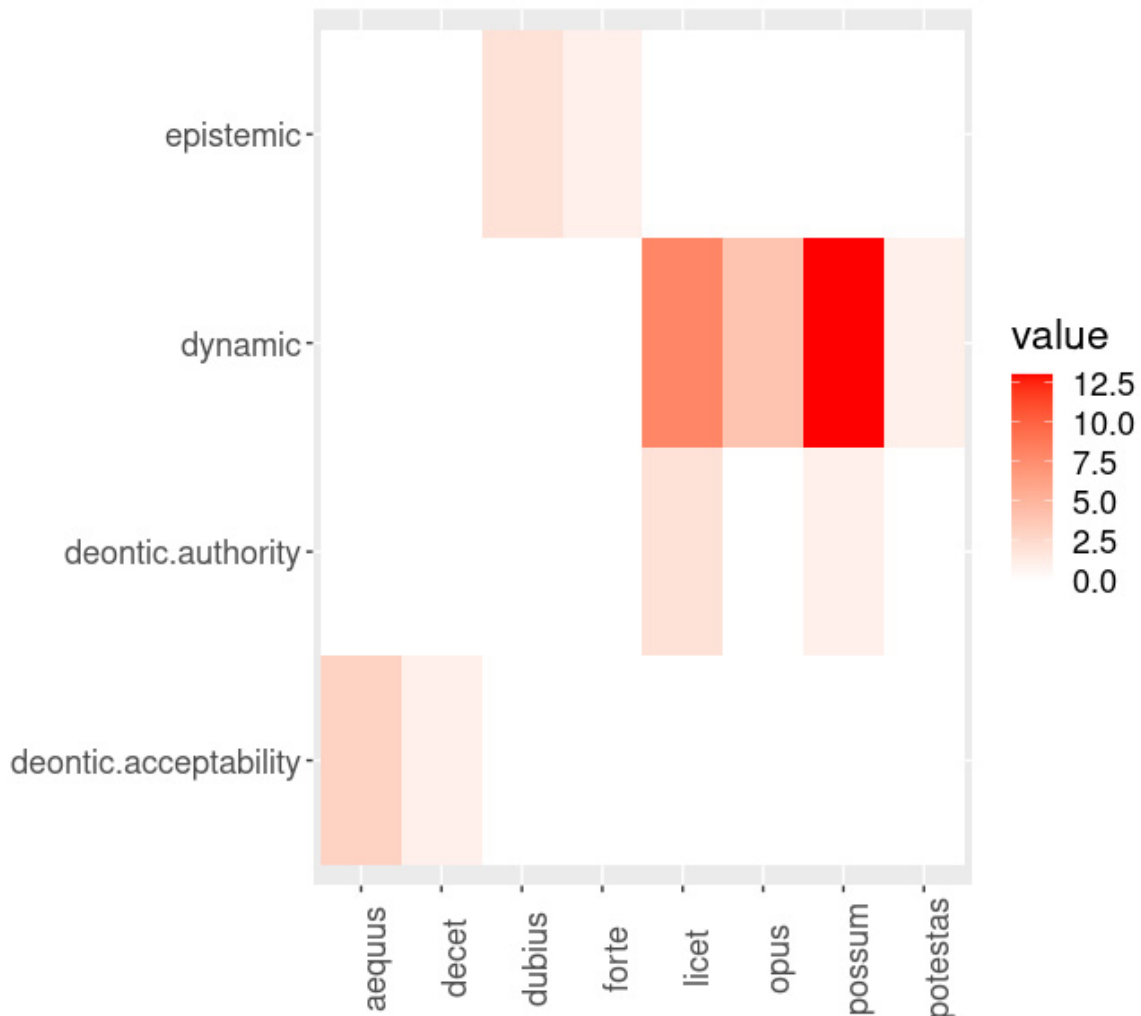Table 6: Markers in co-occurrence and types of modality expressed.

Figure 1: Heatmap representing the types of modality expressed by the markers found in co-occurrence in the corpus.

As shown by the scale in the figure, darker shades correspond to higher counts of associations between a marker and a type of modality. Some of these associations were expected. As mentioned in section 2, *forte* and *dubius* are two of those markers that are not considered to be polyfunctional, as they are only associated with epistemic modality. Therefore, they appear as epistemic markers in the co-occurrences, as illustrated in example 7 (I include the previous line to provide some context).

7) SYR. *et quidem tuo nepoti huius filio hodie prima mammam dedit haec.* DEM. *Hercle vero serio, si quidem prima dedit, <u>haud</u> **dubium** <u>quin</u> emitti **aequum siet**.* (Adelphoe, verses 974-976)

SYR. And today she breastfed (lit. gave the breast) your grandson, his son, first. DEM. For god's sake, truly and seriously, if she really did it first, **there's <u>no</u> doubt** that it **should be right** to set her free.

The phrase *haud dubium (est),* which co-occurs in this line with *aequum est,* is used by the character Demea to convey his certainty that it should be right to set Syrius' wife free, given her primary role in raising his grandson. *Dubius*

appears twice in the co-occurrences, and always in the phrase *haud dubium est*, to express certainty ('it is <u>not</u> doubtful') about a (modalised) SoA.

On the other hand, although most of the other markers found in co-occurrence exhibit varying degrees of polyfunctionality *per se*, they consistently associate with a single type of modality (e.g. *aequus,* in the phrase *aequum est,* with deontic acceptability, *potestas* with dynamic). This is for instance illustrated by the occurrence of *aequum est* in example 7, where it expresses deontic acceptability, in particular the moral desirability of freeing Syrius' wife, given her commendable behaviour.

An exception to this are *possum* and *licet,* which are both associated with the deontic authority (example 8) and dynamic (example 9) types of modality, depending on the context. Consider example 8 first.

8) *Quod si personis iisdem uti aliis <u>non</u> **licet**, qui magis **licet** currentes servos scribere, bonas matronas facere, meretrices malas, parasitum edacem, gloriosum militem, puerum supponi, falli per servum senem, amare, odisse, suspicari?* (Eunuchus, verses 35-40)

But if it is <u>not</u> **allowed** to use the same characters as others, how is it any more **allowed** to describe running slaves, to make up good matrons, a greedy parasite, a boastful soldier, a child who's been substituted, an old man who's been tricked by a slave, love, hatred, suspicion?

These verses belong to the prologue of the *Eunuchus.* As he often does, Terence uses the prologue to declare his influences and to present a critical defense of his own work. This specific comedy is a rework of Menander's homonimous play, with elements from Menander's *Kolax* ('The flatterer'). More specifically, Terence is here responding to an hypothetical accusation of plagiarism mentioned in the preceding lines, launched in presence of a magistrate (the *aedilis*) attending the preview of the play, and is asserting his right to use other authors' narrative material, as it was customary. Given the context, where there is a reference to an accusation of plagiarism and a specific figure tasked with verifying the veracity of this accusation, the deontic interpretation appears to be the most relevant for both occurrences of *licet* in this passage.

The dynamic value of *licet* is illustrated in example 9.

9) *date **potestatem** mihi statariam agere ut **liceat** per silentium*; (Heaut. 35-36)

   Give me the **opportunity** so that **I can** stage the play in silence;

This passage belongs to the prologue of the *Heauton Timoroumenos*. The actor assigned to the prologue is addressing the audience, asking for the opportunity of performing the comedy in silence. In this sentence, the co-occurring markers are *potestas* and *licet.* Both markers in this sentence show a dynamic possibility value. We will get back to this passage when discussing cases of scope hierarchisation (see section 4).

3.3   The co-occurrences

Having provided an overview of the markers and types of modality that were found in co-occurrence, I will show in this section how they combine with each other.

The co-occurrence count[11] values for only modal markers (Table 4) are shown in Table 7. This Table is a symmetric matrix: the variables in the rows and columns are the same, and the values in the cells are symmetric with respect to the main diagonal. This means for example that the co-occurrence *decet-licet* and *licet-decet* holds the same value. The "real" co-occurrence count for the pair *decet* and *licet* is obtained by summing the values of the two pairs. The reason for this lies in the fact that the order in which the markers appear in the sentence was not considered as a parameter for the representation of the co-occurrences in the matrix. Several studies have shown that Latin is a language with a relatively free word order (see for example Spevak 2010), and that the position of specific constituents in the clause can be due to factors not directly related with the co-occurrence itself, e.g. pragmatic ordering of the constituents (Spevak 2010: 47-59), or in this particular case, metric constraints. Since the Terentian corpus is of particularly small size, I have decided to discard this type of information in the formal representation of the co-occurrences to avoid further issues related to data sparsity. However, the order of appearance of the markers will be addressed in section 4, as it will turn out to be relevant for certain types of co-occurrence in particular.

	aequus	decet	dubius	forte	licet	opus	possum	potestas
aequus	0	0	0.5	0	0	0.5	0.5	0
decet	0	0	0	0	0.5	0	0	0
dubius	0.5	0	0	0	0	0	0.5	0
forte	0	0	0	0	0	0.5	0	0
licet	0	0.5	0	0	4	0	0	0.5
opus	0.5	0	0	0.5	0	0	1	0
possum	0.5	0	0.5	0	0	1	5	0
potestas	0	0	0	0	0.5	0	0	0

Table 7: Co-occurrence counts for modal markers in the corpus.

The values shown in Table 7 can be visualised as a network in Figure 2.

---

[11]   The co-occurrence count is the number of co-occurrences of two specific markers in the corpus.
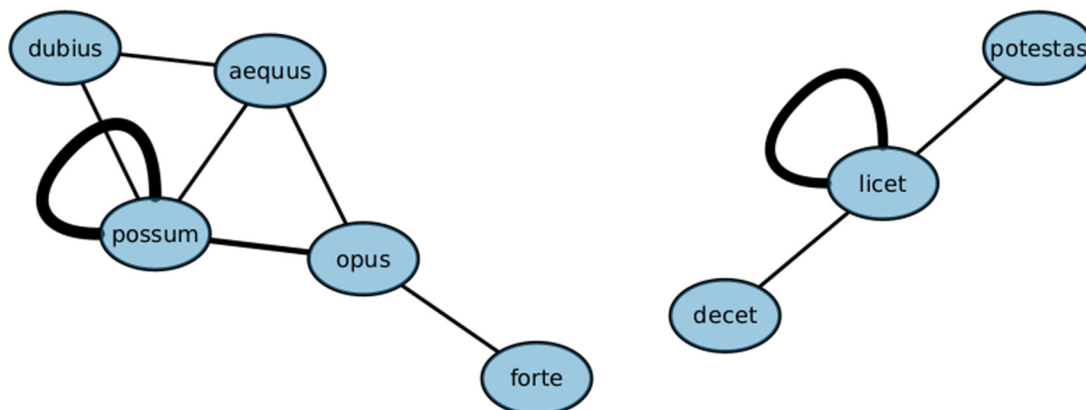
Figure 2: Co-occurrence of the modal markers in Terence's corpus as a network (co-occurrence counts).

The representation of language as a network is a well-known practice (see Ferrer I Cancho 2010 for an overview). In this type of network, the nodes represent a certain linguistic item, e.g. word or part of speech, and the links (or edges) represent the type of relation between them. In our case, the nodes are the modal markers and the links represent the presence of at least one co-occurrence between them. By studying this network, we can highlight some of the features of the co-occurrences. First of all, two clusters of co-occurrence emerge: the first one including *decet, licet* and *potestas,* and the second one including *aequus, possum, dubius, opus* and *forte.* Moreover, *licet* and *possum* show what is called a self-loop, that is a link to the same node. This means that they are the only markers for which we can find two occurrences in the same sentence (e.g. *possum* co-occurring with *possum*). The width of the links between the markers represents the value of the co-occurrence count: the wider the link, the higher the value of co-occurrence count for two markers. Figure 2 shows that the highest values of co-occurrence correspond to the two self-loops in the network.

In general, the co-occurrence count alone is not enough to account for the degree of attraction between two markers or two types of modality. This measure does not consider the count in co-occurrence for each marker, i.e., the total number of co-occurrences in which the marker appears (shown in Table 4) nor the total number of co-occurrences in the sample. To address this issue, I attempted to calculate the independence quotient for each pair. The independence quotient establishes the degree of attraction between two items of a pair, under the hypothesis of their independence (see Egloff & Bavaud 2018: 17). This type of measure takes into account both the total number of co-occurrences in the sample and the number of co-occurrences for each marker

in the pair: if the value of the independence quotient is higher than 1, there is a relation of mutual attraction between the two markers. Table 8 shows the results for the independence quotient.

	aequus	decet	dubius	forte	licet	opus	possum	potestas
aequus	0	0	6	0	0	3	0.86	0
decet	0	0	0	0	3.6	0	0	0
dubius	6	0	0	0	0	0	1.29	0
forte	0	0	0	0	0	9	0	0
licet	0	3.6	0	0	2.88	0	0	3.6
opus	3	0	0	9	0	0	1.29	0
possum	0.86	0	1.29	0	0	1.29	1.84	0
potestas	0	0	0	0	3.6	0	0	0

Table 8: Independence quotient values for the modal markers in co-occurrence.

The problem with using such a measure on this data sample is data sparsity. There are not enough observations (in this case, co-occurrences) for the results to be reliable. The measure tends to overestimate the attraction for markers that have a very low count in co-occurrence. An exemplary case is the co-occurrence between *forte* and *opus,* which obtains an extremely high independence quotient (9), simply because *forte* appears only once in co-occurrence, and that is with *opus*. On the other hand, the independence quotient for *possum* and *licet,* which co-occur approximately ten times more than *forte,* is quite low. The same type of problem arises with the co-occurrence of the types of modality (see Table 9 and the corresponding visualisation in Figure 3): the only co-occurrence of deontic authority with epistemic modality gives an extremely high value to this specific co-occurrence. Because of this issue, and given the low number of observations, I will show from now on only the co-occurrence counts, keeping in mind that the dynamic possibility type, as observed for *possum, is* over-represented in the co-occurrences.

	Deo. Accept.	Deo. Auth.	Dyn.	Epist.
Deo. Accept.	0	0	1.5	0.5
Deo. Auth.	0	1	0	0.5
Dyn.	1.5	0	11	0.5
Epist.	0.5	0.5	0.5	0

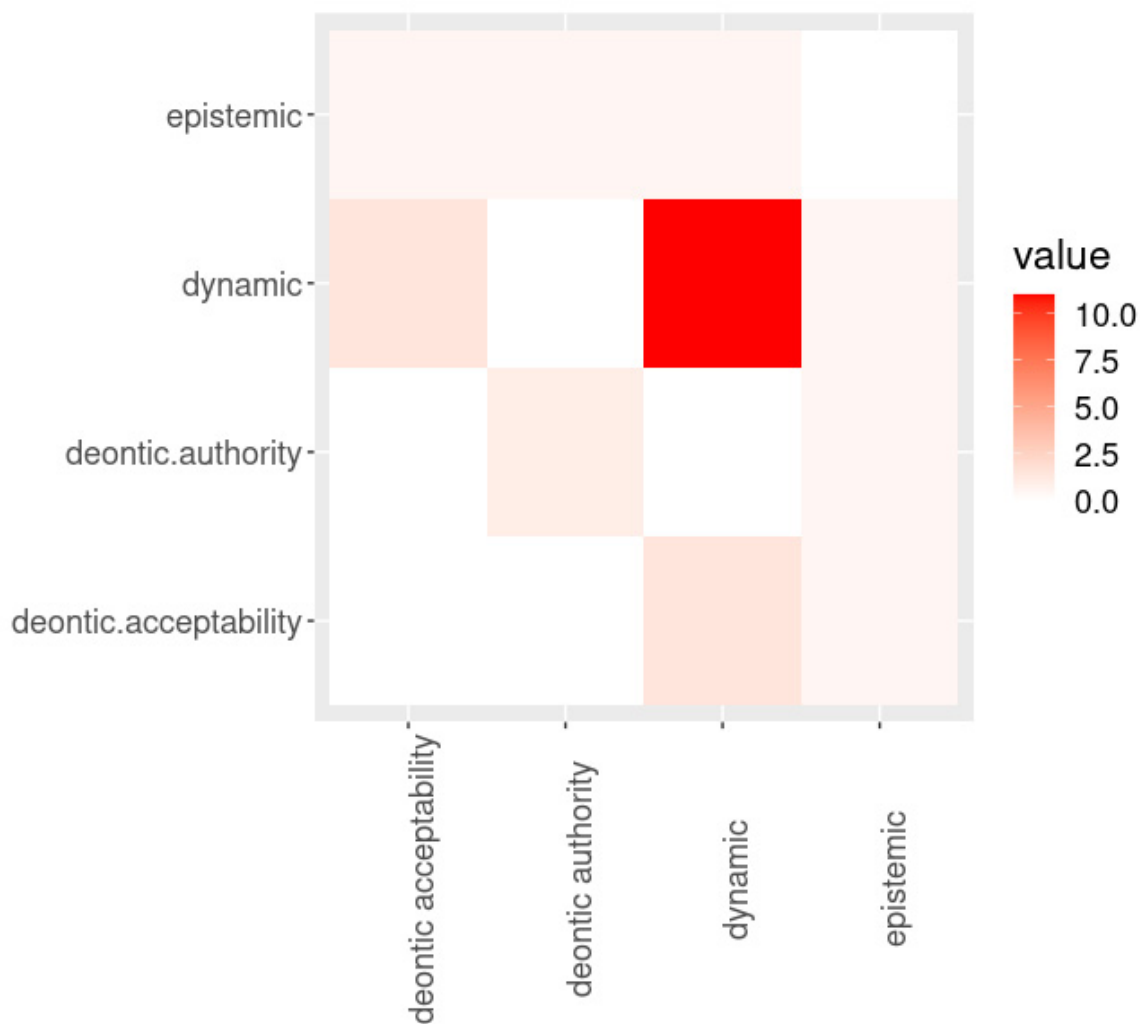Table 9: Co-occurrence count for the types of modality.

Figure 3: The co-occurrence counts for the types of modality in a heatmap.

The most striking feature of the co-occurrence between types of modality is that dynamic modality outshines all the other types of modality in co-occurrence. Again, we observe a self-loop type of co-occurrence for dynamic modality. An illustration for this type of combination was presented in examples 8 and 9, where the two *licet* in self-loop both express, respectively, deontic authority and dynamic modality. Comparing these numbers with the networks in Figure 2, it is fair to hypothesise that the self-loops on *possum* and *licet* might play a role on the high number of self-loops on dynamic possibility, as both of them have a modal base of possibility. Figure 1 previously showed that *licet* and *possum* are strongly associated to dynamic possibility, and to a lesser extent to deontic authority, but additional examples will be provided during the in-depth analysis of the sentences (section 4). An opposite situation can be observed for epistemic modality, which co-occurs with all the other types of modality. An illustration was given in example 7 with *haud dubium est* and *aequum est*. In the next section I will discuss in particular the three cases of epistemic modality combining with deontic acceptability, deontic authority, and dynamic modality. These types of combinations, along with the example of *potestatem dare*, are

particularly important for the cases of scope hierarchisation, mentioned in the introduction (see examples 7, 9 and 13 in this section).

3.4   The types of syntactic structures

Following the qualitative analysis of the modal passages, it turned out that only 18 out of the 29 sentences extracted from the corpus host co-occurrences in which both markers are modal. To add another level of analysis, I integrated information regarding the syntactic relationship between the co-occurring markers. To do so, I categorised them into co-occurrences in simple clauses (involving only one predicate) and in complex clauses (involving more than one predicate). The first category includes cases of modal adverbs modifying another marker, while the second category covers cases of coordination and default subordination (see Table 10[12]). Regarding the last two subcategories, two clarifications are necessary. The first one pertains to coordination: only one co-occurrence belongs to this category and it is an asyndetic juxtaposition of the two predicates (this case will be presented in example 12). That is why it is marked with an asterisk in Table 10. The second clarification concerns the subordination category. This category is intended as rather broad, including both argument and satellite subordinate clauses, as defined by Pinkster (2021: 11), i.e., respectively clauses functioning as arguments to the main verb of the superordinate clause, and clauses functioning as optional constituents to the same verb (adverbial clauses). In this regard, it is important to mention that the modal markers that appear in co-occurrence in the corpus (and specifically those listed in Table 4) are quite diverse concerning their syntactic construction. The construction patterns of the impersonal phrases *aequum est*, *dubium est, opus est*, and the impersonal verbs *decet* and *licet* all require an argument clause functioning as a subject. The verb *possum* allows for a one-slot pattern with an infinitive complement, or a two-slot pattern where it is used as an auxiliary, with a nominative subject and an infinitive complement (Pinkster 2015: 210-215; Magni 2010 for an overview of the construction patterns of Latin modal verbs; cf. also Bolkestein 1980). This type of information can indeed have an influence on the types of syntactic structure that link the co-occurring markers, specifically when one of them is part of the argument structure of the other. It also certainly plays a role in the qualitative analysis of the co-occurrences, which will be discussed in the next section.

---

[12]   In two sentences there were three co-occurring (modal) markers. In the first sentence, I treated the co-occurrences as separate couples of two markers, annotating the type of relation for each pair. The second case will be illustrated in example 12, where *possis dicere* introduces direct speech containing a self-loop for *licet*. In this case I did not consider *possim* for the analysis of the co-occurrence.

Simple clause	Complex clause	Total number of co-occurrences
Adv + V		1
	Coordination*	1
	Subordination	16

Table 10: Types of syntactic relations between co-occurring markers.

As shown in Table 10, the vast majority of the co-occurrences in the Terence corpus appears in a relation of subordination, as just defined. See example 10.

10) SI. *Nam gnatus quod pollicitust, haud **dubiumst** mihi, id si nolit, quin eum merito **possim** cogere.* (Andria, verse 530-531)

SI. If my son refuses to keep his promise, **there's no doubt** I **would** justly **have the right** to compel him.

In this passage, Simo is talking to himself, convinced (*haud dubiumst*) that since his son Pamphilus has officially promised to marry Philumena, he will have every right to compel him (*possim cogere*). *Haud dubium est* expresses epistemic certainty and *quin* introduces the argument clause *possim cogere,* where *possim* functions as a modal auxiliary and conveys a value of deontic authority onto the predicate *cogere* 'to compel'.

Example 11 shows the other case of subordination, where one of the markers is the main predicate of a satellite clause, and the other one is the main predicate of the superordinate clause.

11) SYR. *immo aliis si **licet**, tibi <u>non</u> **licet**. Omnes te in lauta et bene parata re putant.* (Heaut. verses 797-798)

SYR. But if others **can** do that, you **can<u>not</u>**. Everyone thinks you are in a state of wealth and comfort.

In these verses, Syrus, Chremes' slave, is pressuring his master to repay the debt contracted by the slave with the courtesan Bacchis. Syrus points out to Chremes that, although he did not contract it himself, avoiding to pay a debt would tarnish his image as a wealthy and well-to-do man. Therefore, the circumstances prevent him from adopting a behaviour that anyone else in the same situation would adopt, i.e., not repaying the debt. The subordinate clause is introduced by a concessive *si.* As it will be shown in the following section, certain types of co-occurrence found in Terence, such as this one, serve the purpose of building the argumentative scheme of the character.

The overwhelming majority of cases of subordination is an interesting result. Previous studies (Halla-Aho 2018: 79, building on de Melo 2007: 104) have shown how sentence complexity, i.e., the ratio of subordinate clauses to all clauses in Terence's plays, reaches only 31.89%, compared to 58.60% of subordinate clauses in Lucretius, for example. It is also worth reminding that

most of the considered markers allow for a second predicate as an argument in their argument structure. This predicate being another modal marker would thus fall into the category of subordination as defined above. However, of all cases of subordination in Table 10, only 3 of them display a structure where one of the markers occurs in an argument clause depending on the co-occurring marker: in 2 sentences the main predicate is *dubium est* (see example 10) and in one sentence it is the phrase *potestatem dare* (example 9 above). In all the other cases, the subordinate is a satellite clause, like in example 11 above.

Example 12 represents the only case of non-subordination in a complex clause: the marker *licet* occurs in two different clauses, juxtaposed in an asyndetic relation.

12) MIC. [preceding context: *Multa in homine, Demea, signa insunt ex quibus coniectura facile fit, duo quum idem faciunt, saepe ut possis dicere:*] *Hoc **licet** impune facere huic, illi non **licet**;* (Adelphoe, verses 821-824)

     MIC. [preceding context: Demea, there are many signs in men based on which a conjecture is easily made, so that when two people do the same thing, you could often say:] this person **can** safely do this without punishment, that one **can**'t.[13]

This passage belongs to the comedy Adelphoe. Micion, the speaker, has two sons, Ctesipho and Aeschinus, but unable to support both, he gives Aeschinus to his brother Demea, who is the addressee of the speech. Ctesipho and Aeschinus grow up very differently: Micion is very permissive, while Demea is very strict. It is revealed that Aeschinus, the son entrusted to Micion, carries off a girl. It is later discovered that it was actually his brother Ctesipho to commit the crime, and Aeschinus only covered for him. So, Micion and Demea start arguing to understand which is the better educational method. Micion is here trying to excuse Ctesipho to justify Aeschinus who covered for him, and thus, himself, defending the educational method he adopted with Aeschinus. I will refer back to this example in section 4 concerning the use of the repetition of the marker to characterize the character's mood and/or argumentation.

The case in which a modal adverb scopes over a modal verb 'Adv+V' also appears in only one co-occurrence. In example 13, the adverb *forte* (epistemic uncertainty) scopes over *opus sit* (dynamic necessity).

13) DAV. *Quia si **forte opus sit** ad herum iurandum mihi non apposuisse ut liquido possim.* (Andria, verse 728-729)

     DAV. Because, if **maybe I need to** swear to my master that I didn't put it (i.e., the baby) here, I want to be able do so with a clear conscience.

---

[13]   In this sentence there is a third marker, *possis,* which introduces the direct speech containing the two co-occurring instances of *licet*. As it is not semantically related to the two *licet*, it is not considered in the analysis. The same reasoning was applied for the analysis of example 13, where *possim* does not have a role in the co-occurrence between *forte* and *opus sit.*

In this passage, Davus, Pamphilus' slave, is trying to sabotage the marriage between Pamphilus and Philumena. His plan is to expose Glycerium's and Pamphilus' child to Philumena's father, Chremes, and lead him to withdraw his consent to the marriage. Davus is speaking with Mysia, Glycerium's slave, telling her that she should be placing the child in front of the master's door, so that he will be able to swear that he did not do it with a clear conscience. Therefore, *forte* points to a hypothetical situation in which Davus is asked to swear that he was not involved, and *opus sit* to the necessity of perjuring himself.[14]

From the quantitative overview, some preliminary conclusions can be drawn about the phenomenon of co-occurrences in Terence's works. Among the 40 markers, only 8 have been identified as modal markers in co-occurrence after qualitative analysis. *Licet* and *dubius* stand out with the highest co-occurrence rate, while *possum*, despite having the highest count in co-occurrence, has a relatively low co-occurrence rate due to its overwhelming presence in the corpus. The 8 markers in co-occurrence are mostly associated to a single type of modality (epistemic for *dubius,* deontic acceptability for *aequus* and *decet*, dynamic for *potestas* and *opus*). On the other hand, *licet* and *possum*, being highly polyfunctional and having the highest count in co-occurrence, are found to express both dynamic and deontic authority depending on the context.

Regarding the co-occurrences between the markers, the network representation highlights some peculiarities, including the presence of self-loops in correspondence with *possum* and *licet*. This seems to match with a strong value of co-occurrence for the dynamic type of modality. *Licet* and *possum* also seem to form the central node of two different clusters within the network, co-occurring with *potestas* and *decet* for *licet*, and with *dubius*, *aequus*, and *opus* for *possum*. The co-occurrence of *forte* and *opus est* is clearly overestimated by the independence quotient value, but it is nevertheless relevant to the study because it is also the only case pertaining to the syntactic type 'Adv + V', where an (epistemic) adverb scopes over a modal verb. The vast majority of the co-occurrences appears in complex clauses, mainly in subordination (either in argument or satellite clauses), while there is only one example of asyndetic coordination (in a self-loop for *licet*).

The next section will delve into the qualitative analysis of the sentences to investigate how all these elements (the markers, types of modality, and syntactic structure) interrelate. I will explore how they interact with other semantic-syntactic elements such as negation and specific adverbs. When relevant, the focus will be on the rhetorical and argumentative outcomes of these combinations.

---

[14]   As it will be explained in the next section, *si forte* seems to behave almost like a collocation here.

4.   Qualitative analysis of the co-occurrences

Building on a closer qualitative analysis of the passages, and guided by the results shown in section 3, I will now discuss the semantic and syntactic patterns that emerge in the co-occurrence of modal markers in Terence.

As shown in Table 10 (section 3.4), subordination is the most frequent type of syntactic relation in the co-occurrences. 50% of the co-occurrences with subordination host a self-loop either for *licet* or *possum* (see Figure 2 in section 3.3). Another self-loop for *licet* coincides with the only case of asyndetic coordination, presented earlier in example 12. Table 6 also suggests a strong association between the markers involved in self-loops and dynamic modality. On a closer analysis of the co-occurrences, I found that whenever the co-occurrence is a self-loop, the two markers express the same type of modality. Consider example 14.

14) [preceding context: PHAE. *Si non tangendi copiast, eho, ne videndi quidem erit?*] *Si illud <u>non</u> **licet**, saltem hoc **licebit**.* (Eunuchus, verse 639-640)

[preceding context: PHAE. If there is no chance to touch her, hey, it won't even be possible to look at her?] If **it is** <u>not</u> **possible** to do that, it **will** at least **be possible** to do this.[15]

The context of this passage is the same as example 4, section 3.1. Phaedria is complaining about not having the opportunity to see his beloved courtesan Thais, as she is spending a couple of days in another household. In this sentence, both occurrences of *licet* express dynamic modality. The dynamic value of *licet* is anticipated by the expression *non tangendi copia est* 'there is no chance to touch her', where *non […] copia est* conveys the impossibility for Phaedria to touch Thais. The dynamic value is reiterated in the self-loop. Two elements are crucial for the analysis of this co-occurrence. One of them is the opposite polarity of the two co-occurring markers: the first *licet* is negated by *non* 'not', whereas the second *licet* has positive polarity. The second element is the presence of the adverb *saltem* 'at least'. This adverb is often used with an argumentative function: it instantiates an argumentative scale where two different conditions, P and P', compete for the achievement of a specific result R (Bortolussi & Sznajder 2001: 35-36). In this sentence, the situation R is explained in the preceding context: given the circumstances, Phaedria wants to find a way to reach out to Thais. The first condition P, which is presented as the best one to achieve R, is introduced by a hypothetical clause where *si* 'if' is

---

[15]   In this sentence, the scopes of the two co-occurring markers *licet* and *licebit* are the two pronouns *illud* 'that' (for *licet*) and *hoc* 'this' (for *licebit*) respectively. In the framework adopted in this study, for a passage to be modal the scope of the marker must necessarily describe a SoA. The scope can be explicit or implied (Dell'Oro 2023: 6; 16-17). In this passage the pronouns *illud* and *hoc* are used anaphorically to refer to two SoAs described in the immediately preceding sentence (introduced by 'preceding context'): *illud* refers to *tangendi*, the genitive form of the gerund from *tango* ('to touch'); *hoc* refers to *videndi*, another gerund from *video* ('to see, look'). Therefore, *licet* and *licebit* are both modal.

followed by the negative particle *non* 'not'. The hypothetical contains the first co-occurring marker of the self-loop for *licet.* The first condition P is negated (*illud non licet,* i.e., *tangere* 'it is not possible to touch her'), and the second condition P', weaker in the scale, is introduced in the main clause (*hoc licebit,* i.e., *videre*) 'it will be possible to look at her'. The difference between the two argumentative plans is also marked by the two different time plans of the self-loop: the first marker is at the present indicative, while the second one is at the future indicative, emphasising the hopeful spirit with which Phaedria delivers this line. *Saltem* is often found in combination with hypothetical subordinates where it correlates with the hypothetical *si*.[16] The condition introduced by the hypothetical can be negated (*si non*) or not.[17] In this case, within a scale oriented toward achieving the goal of establishing some type of contact with the girl, the best option P is negated within the hypothetical construction, and *saltem* focalises the second-best option P' as possible. The self-loop of *licet* participates in this argumentative structure with a repetition that builds up Phaedria's restlessness in not being able to be with his beloved.

Another case of self-loop building on an argumentative scale is illustrated in example 15, where the two co-occurring *possum* both express dynamic modality.

15) [preceding context: GETA. *Quid faceres, si gravius aliud tibi nunc faciendum foret?*] ANT. *Quom hoc <u>non</u> **possum**, illud <u>minus</u> **possem**.* (Phormio, verse 208)

[preceding context: GETA. What would you do if now you had to do something even more unpleasant?] ANT. Since I **can'**<u>t</u> do this, I'd be <u>even less</u> **able** to do that.

In this passage Antipho is talking to Geta, his father's slave. The two are expecting Antipho's father, who will ultimately find out that Antipho has married a poor orphan girl, without his consent. As Antipho begins to show strong signs of anxiety anticipating his father's furious reaction, Geta asks him to try to conceal it in his father's presence. Antipho says that he is not capable of doing that (*hoc non possum*), and that he would even less be capable of doing something harder (*illud minus possem,* i.e., *gravius aliud*), hence the dynamic reading of the co-occurring markers. Here the first *possum* occurs in a causal clause, but, as in example 14, the subordinate precedes the main clause, presenting the first condition as negated *(illud non licet* in example 14*, hoc non possum* in example 15*).* The second *possum* also receives negative polarity by the adverb *minus* 'less', comparative from the adjective *parvus* 'small, little'. The co-occurrence, as it is presented, has thus the effect of emphasising the

---

16   This structure is already frequent in Plautus (Bortolussi & Sznajder 2001: 39).

17   The negative hypothetical introduced by *si non* 'if not' combined with adverbs like *saltem* 'at least' or *quidem* 'indeed' in the following main clause conveys a rhetorical use of the negative hypothetical, where *si non* expresses the assumption that something (i.e., the condition introduced by *si non*) is not the case (Bertocchi et al. 2006: 56; Pinkster 2021: 319).

character's helplessness: Antipho is unable to do anything unpleasant, to any extent. The argumentative scale is prepared in the preceding context by the adjective *gravius* 'more unpleasant, harder', comparative form of *gravis.* The adjective sets the stage for a comparison between the unpleasantness of one action as opposed to another, pivoting around the notion of capability attributed to the character. This comparison, taken up in the following clause by *minus*, translates into a representation in which *hoc non possum* and *illud minus possem* are treated as two different degrees on the same scale, previously introduced by *gravius*. This structure builds on the rhetorical device of the antithesis, and it is used by the character Antipho to emphasise his incapability of pretending to stay calm in his father's presence. The antithesis is highlighted by the contrast between the indicative mood of the first *possum* and the subjunctive mood of *possim*, which mirrors the *faciendum foret* uttered by Geta and prospects the worst-case scenario as a mere hypothesis.

Examples 14 and 15 show a relation of direct subordination between the first marker and the second one. In other cases, there could be multiple degrees of subordination between the two markers, with an additional subordinate clause interposed between the two co-occurring markers, which I named the "intermediary". This is the case for the sentence illustrated in example 16, where again the co-occurring *possum* both convey dynamic modality, with opposite polarity.

16) BY. *Quaeso edepol, Charine, quoniam <u>non</u> **potest** id fieri quod vis, id velis quod **possit**.* (Andria, verses 305-306)

BY. Please, Charinus, for god's sake, since what you want **can**<u>not happen</u>, you should want what **can** happen.

In this passage, Byrrhia, Charinus' slave, is trying to motivate his master, who is suffering because his friend Pamphilus is about to marry the girl that he loves, Philumena. The structure that links the two markers is slightly more complex: the main clause is *velis* 'you should want', and *id quod possit (fieri)* 'what can (happen)' is a subordinate relative clause functioning as an argument to *velis*. The causal clause introduced by *quoniam* 'since, because' with (*non*) *potest fieri* as main predicate is directly subordinate to the main clause (with *velis* 'you should want' as the main verb). The two subordinate clauses are not directly related to each other, but are both subordinate to the same main clause-which represents here the intermediary-although one is an argument clause and the other one is a satellite clause. This type of construction builds again an opposition between the non-possibility for something to happen (*non potest id fieri*) and the possibility for something else to happen (*quod potest,* i.e., *fieri*). However, in this example, the concept of volition conveyed by *vis* and *velis* participates in the argumentative structure*.* This passage features an interplay of foreground/background between the two representations modalised by the two *possum*, and the notion of volition: the two switch roles, with the non-possibility in the foreground of the subordinate clause, and in the background of

the main clause, and the notion of volition in the background of the subordinate and in the foreground of the main clause.[18] This semantic chiasmus, serving the poetic rhythm, is used by Byrrhia to deliver his *reductio ad absurdum* and unleashes the comedic effect of this line: Charinus could never wish for what causes him suffering, namely his friend's marriage to his beloved.

Another type of co-occurrence that builds on oppositional relations other than polarity was observed with the combination of different modal domains and/or types of modality. This is often the case when a marker which belongs to the modal domain of possibility and another one which belongs to the domain of necessity co-occur. Consider example 17.

17) LACH. *Istuc est sapere, qui ubicumque **opus sit** animum **possis** flectere;* (Hecyra, verse 608)

LACH. This is a proof of having good sense, that is, **to be able** to change your mind whenever **it is necessary**;

In this passage, Laches is talking to his wife Sostrata. She has just been accused by their son Pamphilus of having caused his wife to leave their house with her hostile behaviour. In response, Sostrata decides to leave the city and move to the countryside, thus clearing the way for her son and his wife. Laches has overheard the conversation and is praising his wife's conduct and adaptability to a difficult situation. In this context, *possis* expresses dynamic possibility, as it is referred to Sostrata's ability to adapt, and *opus sit* (in the subordinate clause) conveys dynamic necessity, as it refers to any situation in which it is necessary to adapt.

Excluding self-loops, few sentences in the corpus present a co-occurrence of reinforcement, i.e., between markers that express the same type of modality and have the same polarity. One of them was given in example 9, presented here again for clarity.

18 = 9) *date **potestatem** mihi statariam agere ut **liceat** per silentium*; (Heaut. 35-36)

Give me the **opportunity** so that **I can** stage the play in silence;

As mentioned in section 3, in this passage the co-occurring markers *potestas* and *licet* both show a modal value of dynamic possibility. *Liceat* is in a subordinate argument clause to *date potestatem mihi* 'give me the opportunity', introduced by *ut.* This combination has been described in the literature as "modally harmonic" (Lyons 1977: 807): the two markers reinforce each other and the modal notion that they express in context.

Another feature observed in example 9 is that the hierarchy established at the syntactic level by the sentence structure can in some cases lead to a

---

[18]   As mentioned in the introduction, the notion of volition is part of the modal domain in the WoPoss framework. But as this study focuses on the lexical markers of necessity and possibility, I will not discuss the domain of volition here.

phenomenon that I defined as 'scope hierarchisation': a transfer of the hierarchical relations observed at the syntactic level onto the semantic (specifically modal) level. Hierarchical relations between (modal) constituents can be observed both in the complex clause (e.g. one marker in the subordinate clause, and the co-occurring marker in the superordinate clause) and in the simple clause (typically a modal adverb scoping over a modal verb, i.e., the category Adv + V). The two cases are illustrated respectively in examples 7 and 13, both already presented in section 3. I will repeat them here for clarity.

19 = 7) SYR. *et quidem tuo nepoti huius filio hodie prima mammam dedit haec.* DEM *Hercle vero serio, si quidem prima dedit,* <u>*haud*</u> ***dubium*** *quin emitti* ***aequum siet****.* (Adelphoe, verses 974-976)

> SYR. And today she breastfed (lit. gave the breast) your grandson, his son, first. DEM. For god's sake, truly and seriously, if she really did it first, **there's** <u>no</u> **doubt** that **it should be right** to set her free.

20 = 13) DAV. *Quia si* ***forte opus sit*** *ad herum iurandum mihi non apposuisse ut liquido possim.* (Andria, verse 728-729)

> DAV. Because, if **maybe I need to** swear to my master that I didn't put it (i.e., the baby) here, I want to be able do so with a clear conscience.

In order to explain the concept of scope hierarchisation more concretely, I will first briefly describe the syntactic and semantic features of example 13. The context of this specific line in the *Andria* comedy was already explained in the previous section. The co-occurrence is found in a simple clause, which is the hypothetical introduced by *si,* where the adverb *forte* scopes over *opus sit* and its argument clause *ad herum iurandum mihi non apposuisse.* As mentioned before, *forte* expresses epistemic uncertainty and *opus sit* conveys dynamic necessity*.* The syntactic hierarchy between *forte* and *opus sit* is mirrored onto the semantic relation between the two markers: the value of dynamic necessity expressed by *opus sit* is weakened by the epistemic uncertainty inherited from the adverb *forte,* placed on a higher layer of the syntactic-semantic hierarchy. That is, the marker on the lower position of the hierarchy somehow inherits the modal value of the co-occurring marker that governs it. It should be noted that *si forte* appears to behave here almost like a collocation, in which the epistemic value of *forte* seems to bleach in favour of a discursive use of the collocation that is closer to the meaning 'if, by any chance' (see Orlandini 1997: 252-254).

The same hierarchical relation can be observed in example 7. As explained in the previous section, the construction *haud dubium (est)* expresses epistemic certainty, while *aequum siet* relates to the moral desirability of freeing Syrius' wife. In this case *aequum siet* is the main verb of the argument clause introduced by *quin* which depends on *haud dubium* (*est*). Syntactically speaking, there is a hierarchy where *dubium (est)* is located in a higher position than *aequum siet*. From a semantic viewpoint, the epistemic construction modalises the marker of deontic acceptability and the modalised representation *quin emitti aequum siet*.

As in example 13, a semantic hierarchy emerges from the syntax of the clause: the value of epistemic certainty conveyed by *haud dubium (est)* on the top of the hierarchy is inherited by the value of desirability expressed by *aequum siet* in the lower layer. In this case, the desirability conveyed by the modal located in the lower position is reinforced by the epistemic certainty expressed by the higher-ranked modal.

There are four cases of semantic scope hierarchisation observed in the Terence corpus: one of them is in correspondence to the relation 'Adv + V' (example 13), and the other three were found in correspondence of argument clauses. In three cases an epistemic marker (*forte* and the construction *haud dubium est*) scopes over another type of modality, as in examples 7 and 13 (the other construction with *haud dubium est* was presented in the previous section, example 10). There is only one passage where a dynamic marker (*potestas*) scopes over another expression of dynamic modality (example 9).

The qualitative analysis of the passages reveals certain combinatorial trends between markers and types of modality. These combinations can be analysed not only from a purely semantic perspective (the presence or absence of scope hierarchisation) but also from a rhetorical-argumentative standpoint and in terms of characterisation strategies. There seems to be a certain tendency in Terence's comedies to use the co-occurrence in the form of repetition, both in terms of marker and modality types. As shown above, self-loops are often employed for argumentative purposes, in combination with opposite polarity features and adverbs (like *saltem* or *minus*) that establish a scalar structure on which a comparison between the two modalised representations is built. This serves to emphasise, for example, the helplessness and frustration of the character who is speaking (examples 14 and 15 in this section, but see also example 12 in the previous section). Sometimes the co-occurrence seems to be used by subordinate characters like parasites and slaves to reverse their role in a witty statement. In example 11, the slave is trying to persuade his own master to settle a debt that he himself has incurred. In example 16, the slave is advising his master to move his desires towards something that would be more likely to happen, following an absurd reasoning by which his master should hope for something that would distress him. These examples breach the rule of politeness as generally established in the roman comedy, which adheres to the hierarchical organisation of society, and is usually codified at the linguistic level (Unceta Gomez 2022: 301-2; see also Unceta Gomez 2018). In some cases such a hierarchy can be reversed, typically when the master depends on his slave to solve an intricate situation. In these cases, the slave adopts an attitude that they would not normally have, sometimes resulting in comedic or even absurd effects, as seen in examples 15 and 16.

Sometimes, even though there is no self-loop, the co-occurrence is still employed as a reinforcement for the modal notion expressed throughout the character's line (example 9). Cases of scope hierarchisation like the one in

example 9 can be found in two other examples, with structures conveying epistemic certainty (examples 7 and 10). The characters use them to express their absolute certainty regarding a SoA that is already modalised (by an expression of deontic authority in example 10 and by an expression of deontic acceptability in example 7). Epistemic modality appears to be the only type of modality capable of encompassing other types of modality within its scope, consistently with previous studies in other languages (see e.g. Nuyts 2004; Narrog 2009).

5.   Concluding remarks

In this paper I have investigated the co-occurrence of lexical modal markers of necessity and possibility in Terence's comedies, by first providing a quantitative overview of this phenomenon, and then leaning on the quantitative data to deepen the semantic-syntactic analysis of the co-occurrences. This exploratory study has hopefully shed new light on a phenomenon that is generally underexplored, especially in Latin, by characterising it within a specific author and literary genre. By just looking at the interactions between markers, it has been observed that there are different types of combination, among which the main difference is the way in which the co-occurring markers interact, that is, whether they are organised hierarchically or not. Epistemic and to a lesser extent dynamic modality were found to be able to modify other markers hierarchically. Concerning non-hierarchical relations, the use of modals in co-occurrence appears to be an additional device used by Terence for building the dialogues in a dynamic and effective manner. The characters exploit the co-occurrence of modalities, sometimes combined with argumentative structures, to leverage their argumentation, sometimes bordering the absurd and often triggering the comedic effect. It seems worth emphasising that in the study of Terence's language, it is important to keep in mind that his plays were more or less free adaptations of comedies by Menander and Apollodorus (Brown 2013: 19). The model of the Greek New Comedy, however much it may have been contaminated by Terence in terms of both narrative and linguistic structures, certainly exerts an influence on the language of his plays, and therefore probably also on the co-occurrence of modalities. Moreover, as this is a specific study on Terence, it does not claim to make generalisations about the phenomenon in the Latin language as a whole.

Significant attention has been dedicated to the discussion of the measures appropriate for quantitative analysis and to the identification of the most suitable formal representation for co-occurrences. The methodology so described could be applied to other bigger synchronic or diachronic corpora, and extended to other types of modal markers and modality. The limited size of the corpus and data sparsity inevitably entails some limitations, which have been discussed throughout the paper. The number of co-occurrences retrieved from the Terence corpus is particularly low, a result which can be related to various factors. The

Terence corpus consists of theatrical texts heavily characterised by dialogic verbal interactions. This promotes the presence of generally short sentences, with relatively simple syntactic structures (de Melo 2007; Halla-Aho 2018). Given these premises and the rarity of the phenomenon under study in other languages (e.g. Nuyts 2004), a presence of 0.28% of markers in co-occurrence in the Terentian corpus does not seem entirely surprising. Furthermore, as previously mentioned, the study is based on a selection of markers, which, while numerous, do not encompass all the ways of expressing modality in Latin.

Future research perspectives include expanding the list to other types of modal markers, e.g. morphological markers such as the suffix *-bilis* (for adjectives) and *-ndus* (for the gerundive form). In the analysis of example 14, I discussed the influence of certain expressions (in that case, *copia est*) on the modal interpretation of a marker. This is certainly an aspect that I envisage taking into consideration for further work, expanding the list to include all such expressions, e.g. *pace tua* 'with your good leave, with your permission', that co-occur with modal markers and often influence the interpretation of their modal value. Finally, the study will be expanded to include the modal domain of volition, which, as evidenced by example 16, seems to participate significantly in co-occurrences and to contribute to the building of the argumentation of the characters as much as the other types of modality.

REFERENCES

Allison, B., Guthrie, D. & Guthrie, L. (2006): Another look at the data sparsity problem. In P. Sojka, I. Kopeček & K. Pala (eds.), Text, speech and dialogue. TSD 2006. Lecture Notes in Computer Science. Berlin, Heidelberg (Springer), 327-334.

van der Auwera, J. & Plungian, V. A. (1998): Modality's semantic map. Linguistic Typology, 2(1), 79-124. DOI: https://doi.org/10.1515/lity.1998.2.1.79.

Augoustakis, A. & Traill, A. (eds) (2013): A companion to Terence. Chichester, Malden (Wiley-Blackwell).

Bermúdez-Sabel, H., Dell'Oro, F. & Marongiu, P. (2024): Multi-layered semantic annotation and the formalisation of annotation schemas for the investigation of modality in a Latin corpus. Language Resources & Evaluation. DOI: https://doi.org/10.1007/s10579-023-09706-8

Bertocchi, A. & Orlandini, A. (2001). L'expression de la possibilité épistémique en latin. In P. Dendale & J. van der Auwera (eds), Les verbes modaux. Leiden: Brill, 47-65.

Bertocchi, A., Maraldi, M., & Orlandini, A. (2006): Le latin entre coordination et subordination; les cas où la distinction entre co- et sub-ordination est remise en cause: *si* et ses composés. Faits de Langues, 28(1), 47-56. https://doi.org/10.1163/19589514-028-01-900000006

Bléry, H. (1909): Syntaxe de la subordination dans Térence. Paris (Belin Frères).

Bolkestein, A. M. (1980): Problems in the description of modal verbs: an investigation of Latin. Assen (van Gorcum).

Bortolussi, B. & Sznajder, L. (2001): Syntaxe et interprétation de *saltem*. Journal of Latin Linguistics, 7(1), 35-60. DOI: https://doi.org/10.1515/joll.2001.7.1.35

Brown, P. (2013): Terence and Greek New Comedy. In A. Augoustakis & A. Traill (eds), A companion to Terence. Chichester, Malden (Wiley-Blackwell), 17-32.

Chu, X. (2008): Les verbes modaux du français. Paris (Ophrys).

Coates, J. (1983): The semantics of the modal auxiliaries. London, Canberra (Croom Helm).

Conte, G. B. (1994): Latin literature. A history. Translated by J. B. Solodow, revised by D. Fowler and G. W. Most. Baltimore (Johns Hopkins University Press).

Dell'Oro, F., Bermúdez Sabel, H. & Marongiu, P. (2020): Implemented to be shared: the WoPoss annotation of semantic modality in a Latin diachronic corpus. In Sharing the Experience: Workflows for the Digital Humanities. Proceedings of the DARIAH-CH Workshop 2019 (Neuchâtel). DARIAH-CAMPUS. DOI: http://dx.doi.org/10.5281/zenodo.3739440

Dell'Oro, F. (2023): WoPoss guidelines for annotation. Revised version (Version v4). University of Neuchâtel - Swiss National Science Foundation. https://doi.org/10.5281/zenodo.10427053

Egloff, M. & Bavaud, F. (2018): Taking into account semantic similarities in correspondence analysis. Proceedings of the Workshop on Computational Methods in the Humanities 2018 (COMHUM 2018), 45-51.

Ferrer i Cancho, R. (2010): Network theory. In P. C. Hogan (ed.), The Cambridge encyclopedia of the language sciences. Cambridge (Cambridge University Press), 555-557.

Fruyt, M. & Moussy, C. (eds) (2002): Les modalités en latin: colloque du Centre Alfred Ernout, Université de Paris IV, 3, 4, et 5 juin 1998. Paris (Presses de l'Université de Paris-Sorbonne).

Fruyt, M. & Orlandini, A. (2003): L'expression de la modalité et les verbes perfecto-présents en latin. Revue belge de philologie et d'histoire, 81(3), 693-728.

Fruyt, M. (2004): Some cases of grammaticalisation in Latin: subordinating conjunctions, concessivity and modal lexemes. Classica et Mediaevalia, 55, 301-322.

Halla-Aho, H. (2018): Left-dislocation, subordinate clauses and the stylistic difference between Plautus and Terence. Arctos: Acta Philologica Fennica, 52, 73-94.

Halliday, M. A. K. (1970): Functional diversity in language as seen from a consideration of modality and mood in English. Foundations of language, 25(1), 322-361.

Hanchey, D. P. (2013): Terence and the Scipionic *Grex*. In A. Augoustakis & A. Traill (eds), A companion to Terence. Chichester, Malden (Wiley-Blackwell), 113-131.

Huot, H. (1974): Le verbe *devoir*: étude synchronique et diachronique. Publications de l'Université de Paris X Nanterre. Paris (Klincksieck).

Hütsch, A. (2020): L'usage des verbes modaux en français et en allemand. Étude contrastive de la combinatoire adverbiale sous l'éclairage quantitatif. PhD thesis. Neuchâtel (Université de Neuchâtel, Faculté des Lettres et Sciences Humaines, Institut des Sciences du Langage et de la Communication).

Karakasis, E. (2005): Terence and the language of Roman Comedy. Cambridge (Cambridge University Press).

Kratochvílová, D. (2018): Modality in Spanish and combinations of modal meanings. Charles University (Karolinum Press).

Lyons, J. (1977): Semantics. Vol. 2. Cambridge (Cambridge University Press).

Magni, E. (2005): Modality's semantic maps. An investigation of some Latin modal forms. In G. Calboli (ed.), Latina lingua! Proceedings of the twelfth international colloquium on Latin linguistics, Bologna 9-14 June 2003, Papers on Grammar IX. Rome (Herder), 325-336.

Magni, E. (2010): Mood and modality. In P. Baldi & P. Cuzzolin (eds.), New Perspectives on historical Latin syntax. Constituent syntax: Adverbial phrases, adverbs, mood, tense. Vol. 2. Berlin (Mouton de Gruyter).

Marongiu, P. (2021): Co-occurrence of modal markers in a Latin corpus. From texts to networks. Modelling constructional variation and change: agents, networks and vectors. November 17, 2021. Zurich, Switzerland. https://woposs.unine.ch/outputs/zurich_paola.pdf

Marongiu, P. & Dell'Oro, F. (2022): From static to interactive maps: drawing diachronic maps of (Latin) modality with Pygmalion. Journal of Open Humanities Data, 8, 1-14. DOI: 10.5334/johd.58.

de Melo, W. D. C. (2007): Zur Sprache der republikanischen Carmina Latin Epigraphica. In P. Kruschwitz (ed.) Die metrischen Inschriften der römischen Republik. Berlin (Mouton de Gruyter) 97-120.

Narrog, H. (2009): Modality in Japanese: the layered structure of the clause and hierarchies of functional categories. Amsterdam, Philadelphia (John Benjamins Publishing).

Núñez, S. (1991): Semántica de la modalidad en latín. Granada (Universidad de Granada).

Nuyts, J. (2004): Over de (beperkte) combineerbaarheid van deontische, epistemische en evidentiële uitdrukkingen in het Nederlands. Antwerp (Antwerp University). https://repository.uantwerpen.be/docman/irua/cd8cd9/49314.pdf

Nuyts, J. (2016): Analyses of the modal meanings. In J. Nuyts & J. van der Auwera (eds), The Oxford handbook of modality and mood. Oxford (Oxford University Press), 31-49.

Nuyts, J. & van der Auwera, J. (eds) (2016): The Oxford handbook of modality and mood. Chichester (Oxford University Press).

Orlandini, A. (1997): Adverbes d'énoncé et adverbes d'énonciation en latin: *forte, fortasse*. Lalies, 17, 251-260.

Palmer, F. R. (2001): Mood and modality. Cambridge (Cambridge University Press).

Parry, E.S.J. (ed.) (1857): *Publii Terentii Comoediae Sex*. London (Whittaker and Co.; George Bell).

Pezzini, G. (2015): Terence and the verb 'to be' in Latin. Oxford, New York (Oxford University Press).

Pinkster, H. (2015): The Oxford Latin syntax: Volume 1: The simple clause. Vol. 1. New York (Oxford University Press).

Pinkster, H. (2021): The Oxford Latin syntax. Volume 2: The complex sentence and discourse. Vol. 2. New York (Oxford University Press).

Qi, P., Zhang, Y., Zhang, Y., Bolton, J. & Manning, C. D. (2020): Stanza: A python natural language processing toolkit for many human languages. Association for Computational Linguistics (ACL) System Demonstrations, 101-108.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. (2003): Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Research, 13(11), 2498-504.

Smith, D. A., Rydberg-Cox, J. A. & Crane, G. R. (2000): The Perseus Project: A digital library for the humanities. Literary and Linguistic Computing, 15(1), 15-25.

Smith, G. (2005): Sur la différence de niveau de langue entre Plaute et Térence. Journal of Latin linguistics, 943-955.

Spevak, O. (2010): Constituent order in classical Latin prose. Amsterdam, Philadelphia (John Benjamins Publishing). DOI: https://doi.org/10.1075/slcs.117

Sueur, J.-P. (1975): Étude sémantique et syntaxique des verbes *devoir* et *pouvoir*. PhD thesis. Paris (Université de Paris X-Nanterre).

Sueur, J-P. (1978): Adverbes de modalité et verbes modaux épistémiques. Recherches linguistiques 5-6, 235-272.

Unceta Gómez, L. (2018): Gli studi sulla (s)cortesia linguistica in latino. Possibilità di analisi e proposte per il futuro. Studi e Saggi Linguistici, 56(2), 9-37.

Unceta Gómez, L. (2022): Being polite the Roman way: comments about im/politeness in the comedies of Plautus and Terence. In L. Unceta Gómez & Ł Berger (eds), Politeness in Ancient Greek and Latin. Cambridge (Cambridge University Press), 295-316.

Urbanová, D. & Cuzzolin, P. (2016): Some linguistic and pragmatic remarks on the *tabellae defixionum*. Journal of Latin Linguistics, 15(2), 313-345. DOI: https://doi.org/10.1515/joll-2016-0011

Vairel, H. (1980): Le subjonctif parfait de moindre actualisation dans les comédies de Plaute et de Térence. Revue de Philologie, 54, 122-138.

Xanthos, A. (2014): Textable: programmation visuelle pour l'analyse de données textuelles. Actes des 12èmes Journées internationales d'analyse statistique des données textuelles (JADT 2014), 691-703.