

# Co-occurrence of modal markers in a Latin corpus: from texts to networks

This proposal stems from my PhD project, which focuses on the co-occurrence of modal markers in Latin. Modal markers are linguistic expressions of necessity, possibility and volition. I am specifically interested in studying the syntactic and semantic implications of their co-occurrence in the same syntactic context, as in (1):

(1) *You **can** go to the cinema, but you **have to** finish your homework first.*

In order to perform a structural and quantitative study of this understudied phenomenon, I applied network analysis to a Classical Latin corpus. The networks were built based on a list of 40 lexical Latin modal markers. The nodes are represented by the lemmas of the modal markers in co-occurrence in the corpus. An edge is established if the two modal markers co-occur at least once in the corpus. The weight of an edge is determined by the frequency of co-occurrence. The networks are directed. The direction of the edge is based on the order in which the two modal markers appear in the sentence.

In this poster, I will first showcase the workflow of my project. I will outline the creation of the syntactically annotated corpus, describing: the type of texts included in the corpus; the sources and digital formats with which I have worked; the tools that I used to clean and annotate the texts; the results of the text processing. I will also present the networks that I obtained from my corpus, illustrating the method and describing them by using the metrics that are relevant to my study.

Secondly, I will discuss my results, highlighting the added value of network analysis to this type of research. Based on the properties of the networks, I will be able to delineate the possible pathways of investigation for a further closer analysis of the cases of co-occurrence.

## References

Definition of modality adopted in this study:

Dell'Oro, Francesca. 2019. *WoPoss guidelines for annotation*. Université de Lausanne. doi:10.5281/zenodo.3560951

Nuyts, Jan. 2016. Analyses of the modal meanings. In Nuyts, Jan & Johan van der Auwera (eds.), *The Oxford handbook of modality and mood*. Oxford University Press, Oxford.

Modality in Latin (among others):

Bolkestein, Machtelt A. 1980. Problems in the Description of Modal Verbs: An Investigation of Latin. *Studies in Greek and Latin Linguistics* 1. Assen: Van Gorcum.

Fruyt, Michèle and Claude Moussy (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, Michèle and Anna Orlandini. 2003. L'expression de la modalité et les verbes perfectoprésents en latin. *Revue belge de philologie et d'histoire* 81.3: 693-728.

Núñez, Salvador. 1991. *Semántica de la modalidad en latín*. Granada: Servicio de Publicaciones de la Universidad de Granada.

Co-occurrence of modal markers and interaction with syntactic structures (among others):

Chu, Xiaoquan. 2008. *Les verbes modaux du français*. Paris: Ophrys.

Huot, Hélène. 1974. *Le verbe devoir: étude synchronique et diachronique*. Publications de l'Université de Paris X Nanterre. Lettres et sciences humaines. Paris: Klincksie.

Hütsch, Annalena. 2020. *L'usage des verbes modaux en français et en allemand. Étude contrastive de la combinatoire adverbiale sous l'éclairage quantitatif*. PhD thesis. Université de Neuchâtel. Faculté des lettres et sciences humaines - Institut des sciences du langage et de la communication.

Lyons, John. 1977. *Semantics*. London: Cambridge University Press.

Narrog, Heiko. 2009. *Modality in Japanese. The layered structure of the clause and hierarchies of functional categories*. Amsterdam, Philadelphia: John Benjamins.

Nuyts, Jan. 2004. Over de (beperkte) combineerbaarheid van deontische, epistemische en evidentieële uitdrukkingen in het Nederlands. *Antwerp papers in Linguistics* 108. <http://http://webhost.ua.ac.be/apil/apil108/index.html>.

Palmer, Frank R. 2001. *Mood and modality*. Cambridge: Cambridge University Press.

Rossari, Corinne. (2020). L'exploitation du potentiel argumentatif des formes modales: analyse qualitative et quantitative. In Gaétane Dostie & Pierre Larrivée (eds.) *Représentations du sens linguistique: modalité intra- et extra-phrastique* (Syntaxe et sémantique). Caen: Presses Universitaires de Caen.

Sueur, Jean-Pierre. 1975. *Étude sémantique et syntaxique des verbes devoir et pouvoir*. PhD thesis. Université de Paris X-Nanterre.

Network analysis (among others):

Bollobás, Béla. 2013. *Modern graph theory*. (Graduate texts in Mathematics 184). New York: Springer Science & Business Media.

i Cancho, Ramon Ferrer. 2005. The structure of syntactic dependency networks: insights from recent advances in network theory. In Levickij, Victor, Gabriel Altmann & Valentina Perebejnjs (eds.), *Problems of quantitative linguistics*, 60-75. Černivci: Ruta.

i Cancho, Ramon Ferrer. 2010. Network theory. In Colm Hogan, Patrick (ed.), *The Cambridge encyclopedia of the language sciences*, 555-557. New York: Cambridge University Press.

i Cancho, Ramon Ferrer, Ricard V. Solé, & Reinhard Köhler. 2004. Patterns in syntactic dependency networks. *Physical Review E* 69.5: 051915(8).

Marley, Caitlin A. 2018. Sentiments, Networks, Literary Biography: Towards a Mesoanalysis of Cicero's Corpus. *Theses and Dissertations*. <https://doi.org/10.17077/etd.ot2e09a3>.

Mehler, Alexander. 2008. Large Text Networks as an Object of Corpus Linguistic Studies. *Corpus Linguistics. An International Handbook of the Science of Language and Society*, 328–82.

Passarotti, Marco C. 2014. The importance of being sum. Network analysis of a Latin Dependency Treebank. In Basili Roberto, Lenci Alessandro, Magnini Bernardo (eds.), *Proceedings of the first italian conference on Computational Linguistics. CliC-it 2014. 9-10 December 2014, Pisa, Italy*, 291-295. Pisa: Pisa University Press.

Ponti, Edoardo Maria and Silvia Luraghi, 2018. Non-configurationality in diachrony: correlations in local and global networks of Ancient Greek and Latin. *Diachronica* 35(3): 367-392.

Sudhahar, Saatviga, Gianluca De Fazio, Roberto Franzosi, and Nello Cristianini. 2015. Network Analysis of Narrative Content in Large Corpora. *Natural Language Engineering* 21(1): 81–112. <https://doi.org/10.1017/S1351324913000247>.

#### Text processing:

Stanza = Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton and Christopher D. Manning. 2020. *Stanza: A Python natural language processing toolkit for many human languages*. Association for Computational Linguistics (ACL) system demonstrations.