WoPoss - A *Wo*rld of *Poss*ibilities. An open access annotated corpus to study the diachrony of modality in the Latin language

Francesca Dell'Oro - Université de Lausanne

Until now the study of lexical modality – i.e. the expression of the notion of 'necessity' and 'possibility'[1] – in the Latin language has focused almost exclusively on the synchronic level and mainly on verbs[2]. The investigated corpus has also been limited to a small group of classical authors. Later sources or documentary sources, such as inscriptions, have been neglected. Moreover, no fine-grained annotation of modal types and subtypes has yet been attempted. The SNSF-funded project WoPoss - *A Wo*rld of Poss*ibilities. Modal pathways over an extra-long period of time: the diachrony of modality in the Latin language* (agreement n° 176778) aims to fill these gaps and to make available to the scientific community an open access annotated corpus which will be useful for research on modality or modality-related issues. As annotation progresses, data will be made available during the project lifespan (February 2019–January 2023).

In this paper I want to briefly present the main innovative features of the WoPoss corpus (A1-3) and of the WoPoss annotation (B1-5, see Dell'Oro 2019, Dell'Oro - Bermúdez Sabel – Marongiu submitted)[3]:

A1) WoPoss adopts a diachronic approach: the corpus spans from the 3rd century BCE to the 7th century CE;

A2) the list of markers to be annotated is not limited to modal verbs (e.g. *debeo*), but includes nouns (e.g. *necessitas*), adjectives (e.g. *necessarius*), adverbs (e.g. *necessarie*), modal phrases (e.g. *necesse est*/*habeo*) and adjectival suffixes (such as *-bilis*, *-ndus*, *-turus*);

A3) the WoPoss corpus aims to be representative of any type of (socio-)linguistic variation:

- as it emerges from different kinds of transmission support (codex, papyrus, inscription);

- at the geographical level;

- in terms of textual and literary genres.

B1) the WoPoss annotation is mixed, partially automatic (e.g., for morpho-syntax and metadata), partially manual (fine-grained semantic annotation)

B2) in contrast to the few projects annotating at least one type of modality, the main goal of the WoPoss annotation is not to label different types of modality, but **to annotate in parallel a**

---

[1] There is no consensual definition of modality. The theoretical frame of the project was largely inspired by Nuyts' attempt to formulate a broad, ideally consensual definition of the modal domain (Nuyts 2016).

[2] Cf. the monographs by Bolkestein 1980 and Núñez 1991. There are of course welcome exceptions among shorter publications, e.g. Fruyt - Moussy 2002.

[3] All team members are annotators : http://woposs.unil.ch/team.php

series of parameters which will allow any user of the corpus to understand both why the annotators recognised a certain (sub-)type of modality and what the correlations are between these parameters and a certain (sub-)type of modality;

B3) the WoPoss annotation allows to annotate more than one modal reading, if ambiguity is inherent in the use of a modal marker (polysemy) in the context;

B4) it takes into consideration also pragmatic and rhetoric uses of modal markers, by distinguishing a literal level of analysis and a contextual one;

B5) gradable concepts are evaluated on wide-range scales (e.g. epistemic modality is evaluated on a five-degree scale).

The WoPoss fine-grained semantic analysis will enable the reconstruction of modal pathways as well as to study co-occurrence and competition mechanisms among modal forms in the history of the Latin language. It will also enhance dialogue among researchers using other frameworks or working on other languages.

Selected references

BYBEE, J., R.PERKINS & W. PAGLIUCA (1994), *The evolution of grammar: tense, aspect and modality in the languages of the world*, Chicago.

BOLKESTEIN, A.M. (1980), *Problems in the description of modal verbs: an investigation of Latin*, Assen.

DELL'ORO, F. (2015), 'What role for inscriptions in the study of syntax and syntactic change in the old Indo-European languages? The pros and cons of an integration of epigraphic corpora', in Viti, C., ed., *Perspectives on Historical Syntax*, Amsterdam, 271–290.

DELL'ORO, F. (2019), *WoPoss guidelines for annotation*, 10.5281/zenodo.3560951

DELL'ORO, F - BERMÚDEZ SABEL, H. – MARONGIU, P. (submitted), 'Implemented to be shared: the WoPoss annotation of semantic modality in a Latin diachronic corpus'.

FRUYT, M. - MOUSSY, CL. (2002), eds, *Les modalités en latin. Colloque du Centre Alfred Ernout, Université de Paris IV 3, 4 et 5 juin 1998*, Paris.

MAGNI, E. (2010), 'Mood and modality', in Ph. Baldi - P. Cuzzolin (2009-2011), eds, *New perspectives on historical Latin syntax*, 4 vol., Berlin, vol. 2, 193–275.

MENDES, A., I. HENDRICKX, L. ÁVILA, P. QUARESMA, T. GONÇALVES & J. SEQUEIRA (2016), 'Modality annotation for Portuguese: from manual annotation to automatic labeling', *LiLT* 14.5, 1–38.

NÚÑEZ, S. (1991), *Semántica de la modalidad en latín*, Granada.

NUYTS, J. (2016), 'Analysis of the modal meaning', in J. Nuyts - J. van der Auwera, eds, *The Oxford handbook of modality and mood*, Oxford, 31–49.

Prague Dependency Treebank 3.0: https://ufal.mff.cuni.cz/pdt3.0/documentation

ZIEGELER, D. (2016), 'The diachrony of modality and mood', in J. Nuyts - J. van der Auwera, eds, *The Oxford handbook of modality and mood*, Oxford, 387–405.