From Static to Interactive Maps: Drawing Diachronic Maps of (Latin) Modality with *Pygmalion*

PAOLA MARONGIU (iD)

FRANCESCA DELL'ORO (iD)

Author affiliations can be found in the back matter of this article

]u[ubiquity press

ABSTRACT

In this paper, we present the diachronic maps of a selection of 75 Latin modal markers designed through the tool *Pygmalion*. Both the maps and *Pygmalion* were conceived in the framework of the *WoPoss* project, which aims at analysing the diachronic pathways of modality in Latin. While the description of the tool and its functionalities is beyond the scope of this paper, we focus here on the description of our diachronic modal maps.

Using visualisations to represent semantic shifts is a well-known practice in some linguistic fields such as typology and lexicography, and they have already been applied to modality. Though the situation is rapidly evolving, typological semantic maps as well as lexicographic maps are still for the most part static and usually not-interactive visualisations. Our modal maps stand out not only for their interactivity, but also for the richness of the information conveyed: chronology, etymology, semantics, syntax, first attestation and diachronic relationships between the meanings.

After presenting our conceptual framework for modality, we illustrate the process of conceptualisation and development of our diachronic maps of modality. More specifically, we explain how we gathered and organised the data in order to transpose it into a visual representation. We then showcase the map of *possum* as an example of our results. Subsequently, we discuss the results with respect to previous literature concerning both visualisation of modal evolution from a general point of view and the investigation of modality in Latin. Finally, we outline possible applications within and beyond the *WoPoss* project.

(1) CONTEXT AND MOTIVATION

In this paper we show how modality and some related semantic, syntactic and textual aspects can be visualised along a diachronic axis by using digital, interactive maps. We focus on the case of modal meanings in the Latin language and we showcase how we transposed the gathered data from a discursive to a visual form. In this introductory section, after having defined modality, we briefly present the research project from which the Latin maps stem. Finally, we outline the types and subtypes of modality that the user will find in our diachronic maps.

Modality is a fascinating, though sometimes elusive, concept. Indeed, in the relevant literature, the term is used to refer to different notions. In our work (cf. Dell'Oro, 2019 drawing on Nuyts, 2016) as well as in this paper we intend it as the association of the notional domains of necessity, possibility and volition to any state of affairs (SoA): an example is the sentence *I must go now*, in which *must* is the modal marker of necessity and *I … go now* is the modalised SoA. The diachronic maps of modality presented here were mainly conceived as a synthesis of lexicographic works on Latin modal markers to be variously used in the framework of the *WoPoss*[1] project (cf. Dell'Oro, 2020; Dell'Oro, Bermúdez Sabel & Marongiu, 2020). *WoPoss* aims at reconstructing the pathways of modality in the diachrony of the Latin language on a corpus-based approach. In order to achieve this goal we are currently setting up and annotating a corpus of Latin texts spanning from the 3rd-century BCE to the 7th-century CE. In order to obtain visualisations of the continuity or of the changes of the modal meanings along the dimension of time, we set up *Pygmalion* (2020b, cf. also 2020a and 2021), a user-friendly tool to create highly informative, interactive, diachronic visualisations of words, functions or concepts and their relationships (cf. Dell'Oro, Bermúdez Sabel & Marongiu, forthcoming). Though initially conceived to describe Latin lemmas and modal concepts, *Pygmalion* works for any language, with any script and can be applied to represent any notion in its diachronic evolution. Thanks to this tool we designed diachronic visualisations for each of the 75 Latin modal markers—lemmas or constructions—selected to be analysed in the context of the *WoPoss* project. These markers were selected as they were considered to be the most representative for the expression of modality in Latin. Note that we did not limit our choice to verbs, but we also included other parts of speech. Though we focus on Latin, our methodology can be easily applied to other languages.

For the annotation of the *WoPoss* corpus we adopt a fine-grained analysis (see Dell'Oro, 2019) based on 23 modal subtypes[2] and on the possibility of multiple annotations of the same modal passage. However, in the case of the maps we decided to simplify the annotation schema, while preserving the main types and subtypes of modality (see *Figure 1*; in the examples, modal markers are in bold). We distinguished two basic types for dynamic modality—necessity or possibility[3]—, five subtypes for deontic modality—i.e., permission and obligation (both depending on the presence of a source of authority), acceptability (evaluative modality), volition and intention[4]—and one basic type for epistemic modality.

The modal maps are mainly based on the *Thesaurus Linguae Latinae* (1900–)—from now on *ThLL*[5]—, complemented with other reference works and selected literature (see 2.1 below). The resulting descriptions are adapted and enriched according to the *WoPoss* theoretical framework that we have outlined above.

The paper is organised as follows. In the next section we outline the steps followed to draw the interactive diachronic maps (see 2 below). Then we show the main results that we obtained, taking the map of *possum* as a reference example and highlighting our contribution to the previous discussion (see 3 below). Finally, we present some implications and applications of the maps, from research to didactics (see 4 below).

---

1   *https://woposs.unine.ch/about.php* (last accessed: 10 December 2021).

2   The *WoPoss* theoretical framework is mainly based on the systematisation by Nuyts (2016).

3   For the subtypes of dynamic modality i.e., participant-inherent, participant-imposed and situational see Nuyts (2016, pp. 34–35).

4   The inclusion of volition and intention in the domain of deontic modality is a matter of debate, see Nuyts (2016, p. 37).

5   The *ThLL* is available in open access at *https://thesaurus.badw.de/en/tll-digital/tll-open-access.html*.

type of modality	subtype	definition (mainly with reference to the selected subtypes)	example
dynamic possibility or necessity	not specified	Roughly speaking, dynamic modality can be defined as the expression of necessities and possibilities compelling or enabling the (main) participant in the SoA. Unlike in the case of deontic modality, there is not a source of authority from which such necessities and possibilities derive.	(a) "I'll **be able** to help you in a few minutes." (b) "I'll come down for dinner soon, darling, but I **need to** finish this letter first." (from Nuyts, 2016, p. 35)
deontic	permission and obligation (authority)	We speak of (deontic) authority when, in the context in which the sentence is produced, there is an implicit or explicit source which allows (permission) or compels (obligation) someone to do something.	(c) "You **may** go now." (from Nuyts, 2016, p. 37) (d) "You **must** go now." (personal knowledge of the authors)
	acceptability (evaluative modality)	Evaluative modality conveys "the degree of moral desirability of the SoA expressed in the utterance" (Nuyts, 2016, p. 36).	(e) "This initiative of the federal government is highly **deplorable**." (from Nuyts, 2016, p. 36)
	volition and intention	In the case of volition, the sentence expresses the desire of someone for a SoA to take place. In the case of intention, the person who desires the event expressed in the SoA is also primarily involved as an agent in its realisation.	(f) "I **want** to hear the whole story." (from Nuyts, 2016, p. 37) (g) "I **will** never do it again, I promise." (from Nuyts, 2016, p. 37)
epistemic	not specified	Epistemic modality conveys an estimation about "the chances or the likelihood that the state of affairs expressed in the clause applies in the world" (Nuyts, 2016, p. 38).	(h) "That's **probably** the postman bringing today's newspaper." (from Nuyts, 2016, pp. 38–39)

Figure 1 Main types and subtypes of modality used in the diachronic maps of modality.

(2) METHOD

We usually drew a single map for each lemma annotated in the *WoPoss* project, though sometimes different lemmas (e.g., an adjective and the corresponding adverbs) are outlined in the same map. The choice of drawing one or more maps was determined by the relative complexity of the description, as we did not want to charge a map with an excessive amount of information, though technically this would have been possible. Each lemma is described under different aspects: etymology, semantics and syntax (in particular collocations) with respect to their diachrony.

Our reference model is a map that is able to convey knowledge about the following aspects:

1. a precise chronology (within the limits imposed by data)
2. the etymology
3. the (dis-)continuity of a meaning over time
4. the relations between the meanings
5. the directionality of the relations
6. some examples and relevant textual references

7. the presence of constructions, including collocations
8. the presence of groups of meanings
9. the presence of incertitude in defining some of the previous aspects
10. the presence of analytical categories, such as modal types

Most of such issues (1 to 4, 6, 7 and 10) were already solved in our prototype map of the modal lemma *potestas* 'power, possibility' (Bermúdez Sabel, Dell'Oro, & Marongiu, 2020), in which modal types were color-coded.

In the following we will focus on the process underlying the development of a diachronic modal map.

(2.1) GATHERING AND ORGANISING THE INFORMATION

The first type of information displayed in the maps is the etymology. We mainly relied on de Vaan (2008) and Ernout and Meillet (2001), complementing them with Meiser (1998), Walde and Hofmann (1938–1956), and Meyer-Lübke (1935).

The core of the maps is the list of meanings (or functions), together with their first attestations, positioned with respect to a chronological axis. The description of the semantics of each modal marker is based on the *ThLL*, a monolingual monumental dictionary of Latin, written in Latin. The *ThLL* ideally reconstructs all the meanings attested for a lemma, listing the majority of the attestations of each meaning and organising them in chronological order. However, while the *ThLL* is the best and most complete dictionary of Latin (though the project is still ongoing), its description of modal markers is not compliant with any specific framework of a theory of modality. Therefore, building on the information given by the *ThLL,* we add our semantic analysis, which is oriented at identifying the modal meanings of each marker in a more fine-grained way. When relevant, we also add premodal and postmodal meanings, as defined in typological studies, in particular the one by van der Auwera and Plungian (1998). Sometimes the analysis of the attestations brings to light meanings, or, more often, specific nuances that were not mentioned by the *ThLL*. When this is the case, we include them in the description. For instance, one of the meanings illustrated by the *ThLL* for *debeo* is "to be obliged by necessity". However, a closer analysis of the attestations revealed that this meaning could have three different modal readings depending on the context: dynamic necessity, deontic authority and deontic acceptability. We decided to represent in the map three instances of this meaning, one for each modal reading.[6] At times, in particular in the case of collocations, we also needed to add one or more meanings to the ones already suggested by the *ThLL*. An example is the collocation *necessitatem habeo* in the map of *necessitas* 'necessity'. The *ThLL* specifies all possible constructions (i.e., with the accusative *necessitatem* standalone or followed by a gerund(ive) in genitive or by an infinitive), but it does not associate it with any specific meaning. Building on the semantic analysis of the attestations, we assigned *necessitatem habeo* two meanings: 'to show/imply a state of necessity' and 'to be under the necessity of'. The former is not modal, while the latter shows a basic value of necessity further specified into dynamic or deontic modality depending on the context.[7]

The changes or adjustments that we made to the semantic pieces of information given by the *ThLL* mainly resulted from a deeper analysis of the modal meanings of the relevant lemma. We rarely added or modified meanings that did not show any possible modal reading, unless it seemed necessary. As an example, we can mention the case of *possum* 'to be strong, to be able'. For this verb, the *ThLL* presents a specific context of use, in which *possum* can be referred to words, money and similar. In this context, *possum* indicates the value of something: the economic value of money, or the semantic value of a word, i.e., its meaning. However, the *ThLL* does not provide a specific meaning description for these cases. Therefore, we decided to introduce the meaning 'having some meaning or value', as we did not want to discard this use of the verb, although not modal.[8]

---

6   Map of *debeo: https://woposs.unine.ch/maps/map-debeo.php*.

7   Map of *necessitas: https://woposs.unine.ch/maps/map-necessitas.php*.

8   Map of *possum: https://woposs.unine.ch/maps/map-possum.php*.

The *ThLL* offers extremely detailed descriptions that are often organised in different sections depending on various criteria, such as the animacy of the participants, the context of use, or specific constructions of the lemma (e.g., with dative, with gerund(ive) in genitive or with the subordinating conjunction *ut*). As the maps need to be informative and legible at the same time, we partially discarded this type of information. However, we made some exceptions, based on our interest in the study of modality. For the modal marker *licet* 'it is allowed' the *ThLL* lists, among the others, the meaning 'it is necessary'. This is in contrast with the basic value of possibility expressed by the marker, and with the other listed modal meanings (e.g., 'it is allowed', 'it is permitted', 'to have the chance/possibility/opportunity of doing/undergoing something'). In fact, the *ThLL* adds that this meaning is attested in Plautus with an ironic nuance, where it is used as a mockery. In this case, the indication "ironically, hapax in Plautus" appears in the map next to the meaning 'it is necessary' in order to avoid possible confusion on the part of the user.[9] As mentioned above, the maps also inform about syntactic structures or collocations in which the modal marker participates and that convey a specific meaning altogether. For instance, the lemma *certus* 'certain' can be used in collocations such as *certum habeo* 'I consider it certain', *certo certius* 'certainly', *certior sum* 'to be informed'.[10] Therefore, in the map of *certus* the user will also find these three collocations and their respective meanings.

In some cases, the *ThLL* organises the materials according to a semantic criterion, which we tried to preserve by gathering the meanings around the suggested groups. An example is the lemma *incertus* 'uncertain', whose description is divided into two sections: active sense, describing someone who is uncertain, and passive sense, relative to something doubtful, of which one should doubt. In the digital map, the two sections are represented as semantic groups, and each meaning is assigned to the relevant one. For instance, the meaning 'lacking knowledge, unaware' refers to the semantic group labelled 'who is in doubt', whereas 'not determined, delimited, defined' is assigned to the semantic group 'doubtful'.

For some lemmas a description of the *ThLL* is not available yet. Therefore, we decided to rely on the *Oxford Latin Dictionary—OLD* (1968). This is the case for *nolo* 'to want not', *(ne)queo* '(not) to be able', *usus* 'use' (with the modal meaning of 'need' arising in collocations such as *usus venit* and *usus est), valeo* 'to have strength, to be able', *volo* 'to want' and *voluntas* 'will'. Although the *OLD* is undoubtedly less rich in information, especially about semantic groups and contexts, we applied the same procedure, by extracting the list of meanings with their first attestations and providing additional modal analysis. Specific constructions indicated by the dictionary were also listed as constructions/collocations, when pertinent.

Once all relevant pieces of information are gathered, the meanings are organised in space according to a chronological criterion. In this phase we provide an additional layer of analysis that aims at individuating the pathways of the semantic evolution of the lemma throughout time. This is done by detecting the relations between the different meanings of a lemma, based on their semantics and on previous studies. An important reference for this operation is represented by the models illustrated in van der Auwera and Plungian (1998), in a cross-linguistic perspective, and Magni (2005, 2010), for an application to Latin. The former elaborated a ground-breaking model of a cross-linguistic modal map, illustrating semantic relations involving premodal, modal and postmodal concepts in both synchrony and diachrony.[11] Building on this model, Magni (2005, 2010) developed a comprehensive map specific to Latin. Her map is based on the analyses of the modal pathways of *debeo, licet, necesse (est)* '(it is) necessary', *oportet* 'it is right, proper', and *possum* (for a further discussion of the interplay between our maps and previous studies, see 3.2 below).

(2.2) FROM THE DRAFT TO THE DIGITAL VERSION

The digital map is created by means of the tool *Pygmalion*. The information is encoded by filling out two forms. In the first form the user enters the lemma for which the map will be created and the etymological, chronological and semantic information, together with the attestations.

---

9   Map of *licet*: *https://woposs.unine.ch/maps/map-licet.php*.

10   Map of *certus*: *https://woposs.unine.ch/maps/map-certus.php*.

11   The model is based on the single pathways for possibility and necessity elaborated by Bybee, Perkins and Pagliuca (1994).

The second form is dedicated to determining the relations between the meanings and their directions.

In the first form, the number of etymological steps can be set by the user. For the modal maps we encoded the Proto-Indo-European root and the Proto-Italic form that gave rise to the relevant Latin lemma. The user can also choose the width of the chronological subdivision to appear in the bar above the map: it can be set by centuries, decades or specific years. Our diachronic modal maps are based on centuries. Then, the meanings selected as pertinent are encoded one by one. For each meaning of the lemma we indicate: if it is conveyed by a collocation and which one (as for *certus*); if it is part of a larger semantic group (as for *incertus*); if it is modal (and what type of modality); the passage in which it is first attested; the century of the first attestation.

In the second form, the relations are drawn by selecting the two semantically related meanings and the direction of the link between them. It is also possible to leave the direction of the relation unspecified. We chose this last option for cases of two related meanings that are both first attested in the same century or, more generally speaking, when the directionality of the relation could not be established for some reason.

The etymological steps, the modal readings, or the relations between meanings can all be flagged as uncertain in the *Pygmalion* forms.

Once the two forms have been filled out, the map is automatically saved and can be downloaded as a JSON file. It can be visualised by uploading it on the *Pygmalion* platform at *https://woposs. unine.ch/map.php*.

(2.3) SUPPLEMENTARY MATERIAL: LOCATION AND FORMAT OF THE DATA

The supplementary material accompanying this article contains the data (i.e., the maps) and additional instructions on how to use them. The maps are stored in a repository in Zenodo.[12] The dataset contains: a README file which provides information about the contents of the dataset and the context in which the dataset was produced, and instructions on how to use the data; a JSON file for each modal map. The reason for using JSON format is that the tool employs the JavaScript library D3 for rendering the information in a visual form. *Figure 2* shows how information about the meaning 'ought, should (for logical or similar reasons)' of the lemma *debeo* is encoded in a JSON file, and *Figure 3* shows how it is rendered visually. The structure of a JSON file consists of pairs of name and corresponding value. In this case, the meaning

```
{
  "definition": "Ought, should (for logical or similar reasons)",
  "construct": "_",
  "group": "_",
  "analysis": [
    {
      "id": "mkny",
      "category": "Modal: epistemic",
      "emergence": -1,
      "disparition": "None",
      "attestation": "CIC. Verr. 4, 65 erat candelabrum eo splendore qui ex... pulcherrimis gemmis esse debebat.",
      "certainty": true,
      "relationships": {
        "origins": [
          {
            "rel": "jbsmy",
            "cert": true
          }
        ],
        "destinations": [],
        "unspecified": []
      }
    }
  ]
}
```

Figure 2 Section of the JSON file for the map of *debeo*.
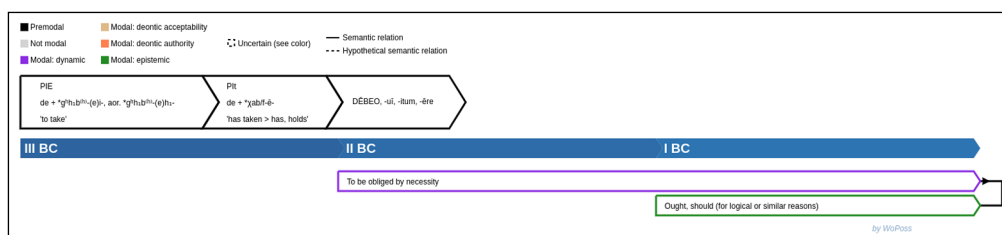
Figure 3 Map of *debeo*, with a focus on the meaning 'ought, should (for logical or similar reasons)'.

'ought, should (for logical or similar reasons)' is the value associated with the name 'definition'. Constructions (including collocations) or semantic groups are specified by the two corresponding names, respectively 'construct' (an abbreviation for 'construction') and 'group'. As in this case no constructions or semantic groups were associated with the meaning, their value is null, represented here with '_'. As more than one type of analysis is possible for a meaning, the information about semantic analysis is encoded in the array of objects 'analysis'. Within this array, the name 'category' illustrates the modal (or not modal) reading ('Modal: epistemic'), while the other names are associated with the date of emergence and disappearance of the meaning,[13] the first attestation, the certainty of the modal analysis (defined by the boolean values 'true' or 'false') and the relationships between the concerned meaning and the other meanings in the map. The information about the relationships is stored in three different arrays depending on their type: 'origins' for directed relationships (ingoing arrow), 'destinations' for directed relationships (outgoing arrow), and 'unspecified' for undirected relationships (simple line). Each meaning (or semantic analysis, in case there are more modal readings for the same meaning) is assigned a random ID (in this case 'mkny'). The IDs are then used as values to create the relations between meanings.

(3) RESULTS AND DISCUSSION

(3.1) MAIN RESULTS: AN OVERVIEW OF AN INTERACTIVE MAP AND ITS FUNCTIONALITIES

In this section we will give an overview of an interactive map, taking *possum* as a reference example (*Figure 4*).

The etymology of the lemma is displayed in the arrow above the map, which is divided in as many sections as the number of etymological steps set by the user. In the case of *possum*, the arrow shows first the Indo-European root **pót-i-*, followed by the Proto-Italic forms **poti-* and **pot-ē-* and their meanings, respectively 'master, in control of' and 'to be master', ending with the Latin paradigm *possum, potuī, posse*.

The chronological bar below, set by centuries, indicates the time span in which the attested meanings of the lemma appear.

All the meanings are displayed inside arrows. Each arrow begins in correspondence with the century in which the meaning was first attested. As the *ThLL* does not usually provide information about the date of (possible) disappearance of a meaning, the arrows continue until the end of the chronological bar. One exception is represented by the already mentioned meaning
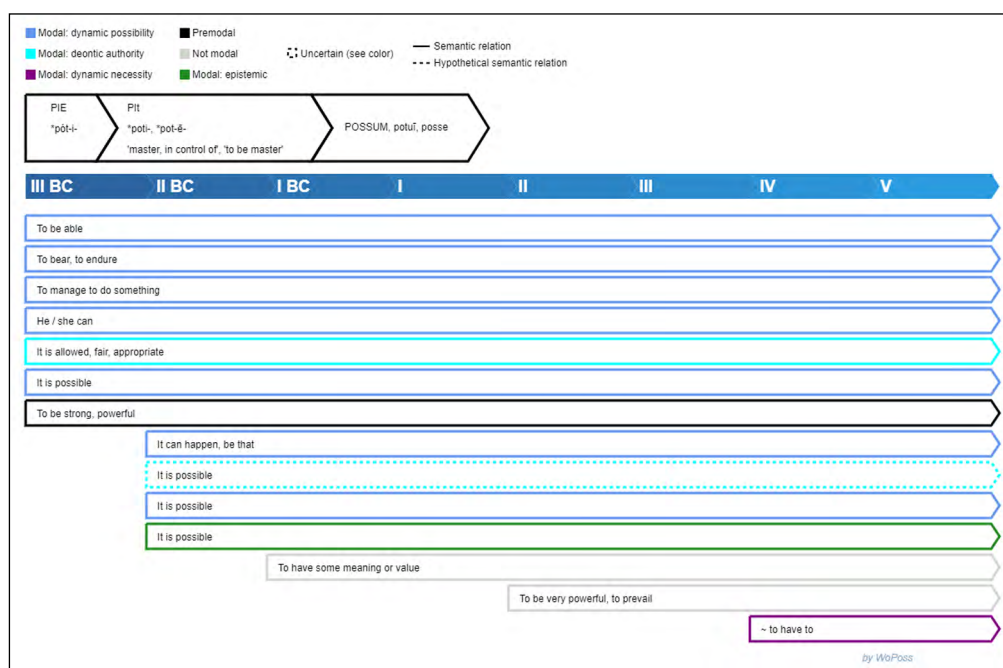

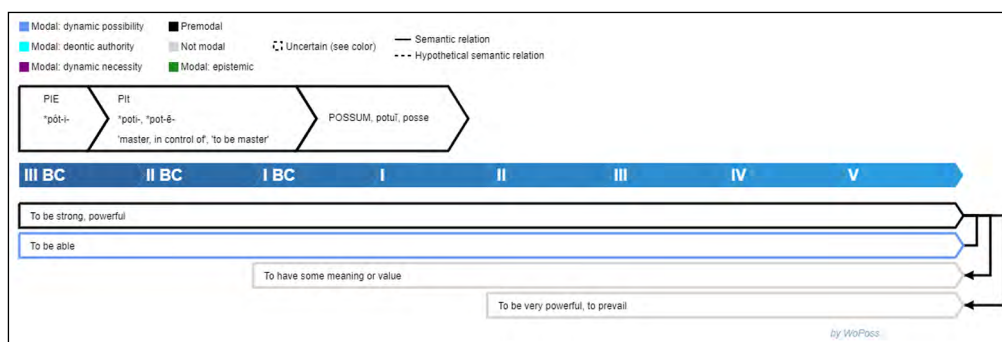**Figure 4** Visualisation of the meanings of *possum*.

---

13   The date of disappearance is never specified ("disparition": "None"), as the lexicographic resources we used do not provide this type of information.

'it is necessary' for *licet* (see 2.1 above). As the *ThLL* specifies that it is only attested in the *Aulularia* by Plautus, the arrow starts and ends in the 2nd-century BCE. The additional semantic analysis is color-coded: a legend at the beginning of the map explains the correspondences. For what concerns the modal readings, we have found attestations for the expression of dynamic possibility (dark blue), e.g., 'to be able'; deontic authority (light blue), e.g., 'it is allowed, fair, appropriate'; dynamic necessity (purple), '~ to have to', where the symbol '~' indicates an approximation of the use of *possum* in that context to the meaning 'to have to'; and epistemic modality (green), 'it is possible'. The meaning 'it is possible' with a deontic authority value is enclosed in dotted lines, which indicate that this type of reading is doubtful. Indeed, the deontic authority value seems to be attested only in this specific fragmentary passage. Since we do not have access to the context of this fragment, we cannot be sure about the modal reading. Non-modal meanings (grey) are 'to have some meaning or value' and 'to be very powerful, to prevail'. We could not find any postmodal value for *possum*. On the other hand, the meaning 'to be strong, powerful' was analysed as premodal, leading to dynamic possibility values.[14]

One innovative feature of the maps is interactivity: the users can set the visualisation according to their needs. The first attestations are not displayed by default, as this would weigh down the visualisation. However, granting access to the first attestations is important for two reasons: each meaning is located on the chronological axis based on this information; moreover, the semantic context in which the marker appears is fundamental for determining a modal reading. In addition, in some cases different arrows can display meanings with the same wording but different modal readings. As shown in *Figure 4*, in the map of *possum* the meaning 'it is possible' can have a dynamic possibility, an epistemic or a (somewhat doubtful) deontic authority modal reading. The user can retrieve the different contexts by hovering over the arrow containing the selected meaning: the first attestation appears showing details about the author, the work and the line or verse corresponding to the attestation. We kept the *ThLL* abbreviations, so that the user can directly refer back to the index available on the *ThLL* website.[15]

Another interactive function is the visualisation of the semantic links between the meanings (*Figure 5*). If the directionality of the relation is specified, the link will appear as a line beginning or ending with an arrow. If the direction is not specified, the two meanings will be linked by a simple line. The user can display the relations by clicking on the arrow containing the selected meaning. This will show all the meanings that were at its origin (outcoming arrow) or were derived from it (incoming arrow). In order to visualise all the meanings again, the user just needs to double click on one of them.

The user can also choose the type of visualisation according to three sorting methods: chronological order, semantic groups and constructions/collocations. If the user selects the first option, the meanings are displayed from the most ancient to the most recent (*Figure 4*). The presence of semantic groups or constructions/collocations, if any, is not displayed with this option. The method 'Groups' shows the meanings gathered by a criterion defined by the user in the first *Pygmalion* form. The map of *possum* does not have any groups, but that of *incertus*, as mentioned in (2.1), shows two groups, labelled 'doubtful' and 'someone who doubts'. The method 'Constructions/Collocations' offers a global visualisation of the meanings where those associated with a construction or collocation appear above (*Figure 6*).

Figure 5 Visualisation of the links between some of the meanings of *possum*.

---

14   As mentioned above, we use "premodal" and "postmodal" in the sense of van der Auwera and Plungian (1998).

15   https://www.thesaurus.badw.de/tll-digital/index/a.html (last accessed: 10 December 2021).

Possum has two collocations: *fieri potest ut* 'it can happen, be that' (dynamic possibility) and *potestur* with passive infinitive 'it is possible', which receives two modal analyses (dynamic possibility and deontic authority).

The option 'Constructions/Collocations' is also used for forms of the same lemma associated with a specific meaning. We exploited this method to illustrate the Latin developments of Proto-Italic **lic-e-*: alongside the impersonal modal verb *licet*, we find the active form *liceo* 'to be for sale, salable', 'to have for sale' or 'to have the possibility/chance/opportunity of doing/undergoing something' (the latter with a dynamic modal reading), and the middle *liceor* 'to bid, make an offer' or 'to bargain, negotiate'.

Each lemma is also illustrated via a network graph shown below the map, where the nodes are the meanings and the edges are the semantic relations between the meanings (*Figure 7*).
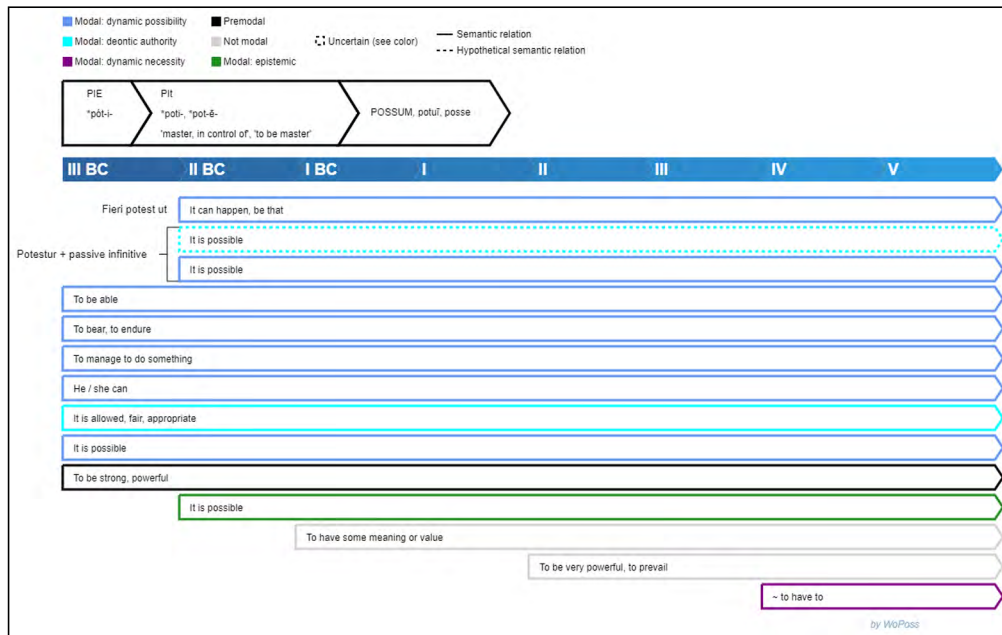

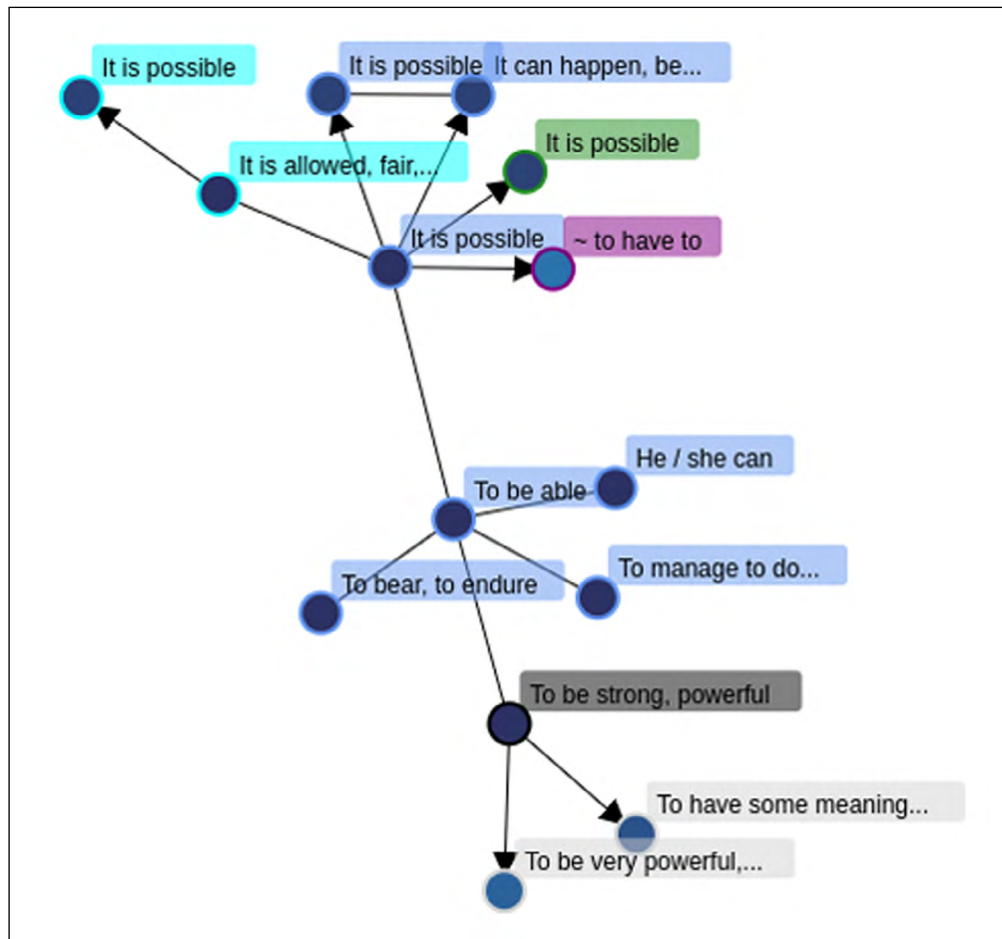**Figure 6** Visualisation of the collocations of *possum*.

Figure 7 Network graph of *possum*.

The edges can be directed or undirected, depending on whether we were able to establish the directionality of the relation. The network graph does not give information about attestations, semantic groups, or collocations. However, the chronological information is color-coded: the most ancient meanings are represented by darker blue nodes, and the shade of blue becomes lighter as the meanings become more recent. The network graph is meant to offer a synthetic overview of the meanings and their relations.

(3.2) FURTHER RESULTS: THE MAPS AS A REFERENCE TOOL FOR STUDYING MODAL MARKERS

The digital modal maps were initially conceived for being used within the project in two main respects. First of all, they represent a support tool for the annotation task. They provide the annotator with a concise and intuitive overview of each modal marker to be annotated in the *WoPoss* corpus. At the same time, the maps are also used as a reference tool for our study of modal markers. At this stage, as explained in (2), they result from the organisation of lexicographic information in the theoretical and more abstract schema of the semantic maps of modality devised by van der Auwera and Plungian (1998) and Magni (2005, 2010). However, on a more advanced stage of the project, this first version of the maps will be compared against the evidence obtained from the annotated *WoPoss* corpus. The aim is to complement or modify the maps basing on the annotated data: this could concern a more ancient attestation for a modal marker or a modal meaning, a semantic nuance that was not captured on the base of the *ThLL* or the *OLD*, a new collocation, a change in the relations between the meanings.

Our maps also raise some interesting issues that we will illustrate by means of the case of *licet*. Magni's work already showed that an attentive review of the Latin attestations can be revealing. For instance, she points out that *possum* is not a case of demodalisation as suggested by van der Auwera and Plungian (1998, p. 106), as the premodal meaning is attested early, already in Plautus (Magni, 2010, p. 217). Building on this, our maps contribute to the discussion in two main respects.

First of all, thanks to a clear chronological visualisation, they make evident when the possibility to establish the directions of diachronic changes cannot be based on the available attestations, but needs to be reconstructed on the basis of comprehensive and cross-linguistic studies. While semantic maps intended in the traditional sense offer a linear visualisation of the semantic changes, our maps, by integrating a philological and lexicographic approach, represent more clearly the actual situation we have to deal with, based on the available evidence. For instance, the map of *licet* shows a very early postmodal meaning conveyed by the construction '*licet* + subjunctive', where *licet* has a concessive value. Notice that this meaning is as anciently attested as the deontic and dynamic ones.[16]

Our maps also allow us to highlight a caveat in the use of lexicographic sources for the study and representation of semantic changes, when these are not complemented with other studies. Although the *ThLL* is extremely rich in attestations, by relying entirely on this type of resource, we risk to end up neglecting rare (modal) meanings or specific semantic shifts. This is even more true for the *OLD*, which provides a smaller amount of attestations. The case of *licet* provides a good example. Magni, building on Núñez (1991, pp. 77–81), classifies *licet* as a modal verb as it can express both dynamic/deontic modality and epistemic modality. However, during our analysis of the attestations provided by lexicographic sources, we did not encounter any instance of epistemic modality for *licet*. The two authors present some evidence for this type of reading. The most striking example is given by Núñez (1991, p. 185):

1. Lucr. 5, 597 *Illud item non est mirandum, qua ratione tantulus ille queat tantum sol mittere lumen, quod maria ac terras omnis caelumque rigando compleat et calido perfundat cuncta vapore. Nam **licet** hinc mundi patefactum totius unum largifluum fontem scatere atque erumpere lumen, ex omni mundo quia sic elementa vaporis undique conveniunt et sic coniectus eorum confluit, ex uno capite hic ut profluat ardor.*

---

16   Without a further corpus-based analysis, it is not possible to date the grammaticalisation of *licet* as the concessive conjunction 'although'. The main issue is that it is often hard to determine if *licet* is used as a verb with a concessive value or as a conjunction.

"Another thing also need not excite wonder, how it can be that so small a sun emits so much light, enough to fill with its flood seas and all lands and the heavens, and to suffuse all with warm heat. For it is possible that from this place is opened one single fountain of the whole world, to splash its generous flood and to fling forth light, because the elements of heat gather together from all parts of the world in such a manner, and their assemblage flows together in such a manner, that the heat flows out here from one single source."[17]

In the text that follows this passage, not included for reasons of space, Lucretius offers two other hypotheses that could explain the phenomenon under discussion, i.e., the ability of the sun to emit so much light. Therefore, we can agree with Núñez on a predominant epistemic reading of *licet* in this passage. It is worth mentioning that the attestations offered by the *ThLL* did not include the passages cited by Magni and Núñez; moreover, both authors (Magni, 2010, p. 228; Núñez, 1991, p. 184) underline that epistemic readings of *licet* are rare. Building on this, we decided to add to the map of *licet* the epistemic reading, attested by this passage in Lucretius' *De rerum natura*. This example proves that sometimes relying on lexicographic sources is not enough to provide a representation of the diachronic evolution of a modal marker as complete as possible. However, it also allows us to show the flexibility of our maps, which can be enriched with new information gathered from our sources, such as other researchers' work (as in this case) or more extensive corpus data. In fact, a more comprehensive, corpus-based semantic analysis of the attestations could capture instances of relatively rare modal readings, as it was the case for *licet*, providing new and different insights on the evolution of modal markers in Latin. As already mentioned, we will be able to provide this type of analysis in the next stage of the project, by means of our annotated corpus.

(4) IMPLICATIONS AND APPLICATIONS
(4.1) IMPLICATIONS FOR THE PRACTICE

Thanks to the progress of computer science the field of data visualisation is rapidly evolving. Therefore, it is important to highlight possible implications of our visualisations in designing modality and modality-related issues, as well as their role as linguistic tools of synthesis and discovery.

Our maps represent the digital transposition of traditional non-interactive drawings which are found in many dictionaries and other types of publications. By exploiting computer technologies and in particular the tool *Pygmalion*, we combined this transposition not only with interactivity, but also with the possibility to add layers of information without overloading the design of the map.

From the point of view of more traditional linguistic approaches, our map model sits between that of lexicography—we will use the term 'lexicographic maps'— and that of typology—their maps are usually called 'semantic maps' (for an application of semantic maps to the domain of modality, see van der Auwera & Plungian, 1998).[18] It is beyond the scope of this paper to detail similarities and differences between diachronic lexicographic maps, traditional diachronic semantic maps and our diachronic maps of modality. Moreover, it must be underlined that the domain is rapidly evolving, as shown by the "lexical diachronic semantic maps" and the state of the art outlined by Georgakopoulos and Polis (2021). Notably, their model is based on network graphs and quantitative methods. Considering these recent developments, we can say that, while traditional lexicographic maps represent the etymological history of a lemma by focusing on formal and semantic changes and their relationships, traditional diachronic semantic maps—which often rely on cross-linguistic data—represent the organisation of semantic space with reference to meanings or functions and their relationships. Our diachronic maps of modality combine some of the features of the two traditional types of maps.

As diachronic lexicographic maps and diachronic semantic maps, our modal maps present the diachronic feature too, but they offer a more precise anchoring to the temporal dimension thanks to the chronological bar. This represents a drastic difference in particular with typological

---

17   Translation by Rouse and Smith (1924).

18   One of their maps is reproduced in Dell'Oro (2019, p. 14). Other important references for the semantic map model are Croft (2001) and Haspelmath (2003).

maps: the latter are usually meant to be universally valid generalisations, and therefore are not tied to a specific temporal dimension. Moreover, our maps contain much more data about the represented items, as illustrated in (2) and (3).

Another similarity shared by all the aforementioned types of maps is the representation of the relationships between the meanings or functions (or more generally the represented items). However, in our maps, thanks to their interactivity, the relationships between the meanings only appear when clicking on an arrow. This way, the visualisation is not weighed down by an excessive amount of information, and the user can select specific items and focus on the relations of each one of them individually.

To conclude, while lexicographic and semantic maps are intended to offer synthetic overviews, our maps are rather conceived as tools to interactively explore various types of information. In fact, the major contribution of our map model to the practice of designing maps is probably the fact of allowing the user to explore a big amount of data at the same time.

(4.2) APPLICATIONS

Our maps have a wide range of applications that go beyond the scope of the *WoPoss* project itself.

They could complement the available lexicographic resources for these lemmas, providing a finer analysis of their modal readings. As shown above, e.g., for *possum,* this type of analysis is not present in the *ThLL* nor the *OLD*, even when they mention the basic modal meaning(s) for a lemma.

We also find that the maps can represent a benchmark for scholars interested in different topics, such as modality and evolution of modal pathways, semantic change, history of the Latin language, Latin linguistics. This is especially due to the fact that they summarise in an interactive image a considerable amount of information that is gathered from many different sources. The consistent description provided for each lemma can orient scholars in their research, saving them a time-consuming task, selecting the most relevant information and providing additional analysis on modality that cannot be found in any of these sources. If necessary, the files of our maps can be downloaded and changed according to the requirements of various circumstances.

Another field of application for the maps could be the classroom. The benefits of involving visualisation techniques in the teaching and learning process are known in literature (Arcavi, 2003; Jessop, 2008; Rattya, 2013; Klerkx, Verbert & Duval, 2014 among others). The existing maps can be used to teach about modal categories and subcategories and their use in context, but also about more general (historical, semantic and partially syntactic, via the collocations) aspects of the Latin language. For instance, they could be used to introduce pupils and students to the use of impersonal verbs. Our set of modal maps features some impersonal verbs or constructions, e.g., respectively *decet, licet, oportet* and *aequus est, necesse est, meum est* among others. Thanks to the option 'Constructions/Collocations' and to the attestations associated with each meaning, the learners can access the syntactic aspects of these verbs and constructions. Moreover, the interactivity of the maps makes them suitable tools for individual learning: the levels and/or types of information presented in the map (etymology, semantics, chronology, constructions and collocations) are well separated, and learners can easily adapt the visualisation to their learning stage and objectives.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Paola Marongiu: conceptualisation, data curation, investigation, methodology, visualisation, writing – original draft, writing – review & editing

Francesca Dell'Oro: conceptualisation, funding acquisition, investigation, methodology, project administration, supervision, visualisation, writing – original draft, writing – review & editing

AUTHOR AFFILIATIONS

Paola Marongiu   *orcid.org/0000-0002-5060-3307*
Institut des sciences du langage, University of Neuchâtel, Neuchâtel, Switzerland
Francesca Dell'Oro   *orcid.org/0000-0002-8343-356X*
Institut des sciences du langage, University of Neuchâtel, Neuchâtel, Switzerland

REFERENCES

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics, 52*(3), 215–241. DOI: *https://doi.org/10.1023/A:1024312321077*

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*

Bermúdez Sabel, H., Dell'Oro, F., & **Marongiu, P.** (2020, July 22–24). Visualisation of semantic shifts: the case of modal markers [Poster session]. *DH2020*, Ottawa, Canada. DOI: *https://doi.org/10.17613/scy4-br70*

Bybee, J. L., Perkins, R., & **Pagliuca, W.** (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective.* Oxford: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780198299554.001.0001*

Dell'Oro, F. (2019). WoPoss guidelines for annotation. Université de Lausanne. DOI: *https://doi.org/10.5281/zenodo.3560951*

Dell'Oro, F. (2020, July 22–24). WoPoss – A World of Possibilities. An open access annotated corpus to study the diachrony of modality in the Latin language [Conference presentation]. *DH2020*, Ottawa, Canada. DOI: *https://doi.org/10.17613/pp6j-0p86*

Dell'Oro, F., Bermúdez Sabel, H., & **Marongiu, P.** (2020, December 5–6). Implemented to be shared: the WoPoss annotation of semantic modality in a Latin diachronic corpus [Conference paper]. Sharing the Experience: Workflows for the Digital Humanities. *Proceedings of the DARIAH-CH Workshop 2019*. Neuchâtel, Switzerland. DOI: *https://doi.org/10.5281/zenodo.3739440*

Dell'Oro, F., Bermúdez Sabel, H., & **Marongiu, P.** (forthcoming). Pygmalion in the classroom: a tool to draw lexicographic diachronic maps and their application to didactics. In M. Márquez Cruz & V. Ferreira Martins (Eds.), *La Lexicografía didáctica: recursos para el aprendizaje de lenguas*. Madrid: Guillermo Escolar Editor.

Ernout, A., & **Meillet, A.** (2001). *Dictionnaire étymologique de la langue latine: histoire des mots* (reprint of the 4th ed. with additions and corrections by Jacques André). Paris: Klincksieck.

Georgakopoulos, A., & **Polis, S.** (2021). Lexical Diachronic Semantic Maps. Mapping the evolution of time-related lexemes. *Journal of Historical Linguistics, 11*(3), 367–420. DOI: *https://doi.org/10.1075/jhl.19018.geo*

Haspelmath, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. *New Psychology of Language, 2*, 211–242. DOI: *https://doi.org/10.5281/zenodo.831410*

Jessop, M. (2008). Digital Visualization as a Scholarly Activity. *Literary and Linguistic Computing, 23*(3), 281–93. DOI: *https://doi.org/10.1093/llc/fqn016*

Klerkx, J., Verbert, K., & **Duval, E.** (2014). Enhancing Learning with Visualization Techniques. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.) *Handbook of Research on Educational Communications and Technology,* (pp. 791–807). New York, NY: Springer. DOI: *https://doi.org/10.1007/978-1-4614-3185-5_64*

Magni, E. (2005). Modality's Semantic Maps. An Investigation of Some Latin Modal Forms. *Journal of Latin Linguistics, 9*(1), 325–336. DOI: *https://doi.org/10.1515/joll.2005.9.1.325*

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, 2*, (pp. 193–275). Berlin: Mouton de Gruyter. DOI: *https://doi.org/10.1515/9783110215458.193*

Meiser, G. (1998). *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Meyer-Lübke, W. (1935). *Romanisches etymologisches Wörterbuch* (3rd ed.). Heidelberg: C. Winter.

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

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. DOI: *https://doi.org/10.1093/oxfordhb/9780199591435.013.1*

OLD = Glare, P. G. W. (Ed.) (1968). *Oxford Latin Dictionary*. Oxford: Oxford University Press/Clarendon Press.

Pygmalion. (2020a). Dell'Oro, F., Rimaz, L. Bermúdez Sabel, H., & Marongiu, P. (2020a). *Pygmalion-simple 1.0. A tool to draw interactive and diachronic semantic maps*. WoPoss. A World of Possibilities. Swiss National Science Foundation.

Pygmalion. (2020b). Dell'Oro, F., Rimaz, L., Bermúdez Sabel, H., & Marongiu, P. (2020b). *Pygmalion-modal 1.0. A tool to draw interactive and diachronic semantic maps of modality*. WoPoss. A World of Possibilities. Swiss National Science Foundation.

Pygmalion. (2021). Dell'Oro, F., Rimaz, L., Bermúdez Sabel, H., & Marongiu, P. (2021). *Pygmalion-colour 1.0. A tool to draw interactive and diachronic semantic maps of modality*. WoPoss. A World of Possibilities. Swiss National Science Foundation.

Rattya, K. (2013). Languaging and Visualisation Method for Grammar Teaching: A Conceptual Change Theory Perspective. *English Teaching: Practice and Critique, 12*(3), 87–101.

Rouse, W. H. D., & **Smith, M. F.** (Eds.). (1924). *Lucretius. On the Nature of Things*. Cambridge, MA: Harvard University Press.

ThLL = Thesaurusbüro München Internationale Thesaurus-Kommission (Ed.) (1900–). *Thesaurus Linguae Latinae*. Berlin: De Gruyter.

de Vaan, M. (2008). *Etymological Dictionary of Latin*. Leiden: Brill.

Walde, A., & **Hoffmann, J. B.** (1938–1956). *Lateinisches etymologisches Wörterbuch* (3rd ed., 3 vols.). Heidelberg: C. Winter.