Sense-Based Annotation of Geographical Nouns in Ancient Greek and Latin: A Diachronic Study with LLMs

Andrea Farina[1], Michele Ciletti[2], Barbara McGillivray[1], Andrea Ballatore[1]
[1]Department of Digital Humanities, King's College London, United Kingdom
[2]Department of Humanities, University of Foggia, Italy
Correspondence: andrea.farina@kcl.ac.uk

Abstract

This paper investigates the lexicalisation of geographical nouns in Latin and Ancient Greek using a diachronic, multi-genre corpus (8th cent. BCE – 2nd cent. CE) and Large Language Models for Word Sense Disambiguation. We focus on two main aspects: the onomasiological question of which words encode core geographical concepts, and the semasiological distribution of senses across lemmas. Across both languages, city-related concepts are the most frequently expressed, but Greek shows a stronger focus on maritime terms, whereas Latin favours concepts related to land. Semasiologically, Latin shows clearer evidence of semantic change over time (e.g., *civitas* 'citizenship' → 'city', *aequor* 'flat surface' → 'sea'), while Greek displays more gradual or distributed shifts. These results show that computational annotation enables cross-linguistic and diachronic analysis of geographical semantics, allowing us to compare the frequency of concepts across languages, genres, and periods, and to track when semantic change occurs and how core concepts evolve over time.

1 Introduction

Language technologies have opened new avenues for studying cultural heritage texts, enabling large-scale analyses of phenomena traditionally explored qualitatively. One such phenomenon is the lexicalisation of geographical concepts, which is central to understanding how ancient societies conceptualised space. This paper investigates the distribution and semantic disambiguation of common nouns referring to geographical entities in Latin and Ancient Greek using computational approaches. Our work adopts an onomasiological perspective, asking: given a concept, which words express it, and how does this set of words evolve across time and genres? We leverage Large Language Models (LLMs) to annotate geographical expressions and perform Word Sense Disambiguation (WSD), testing whether temporal metadata improves performance. With a comparison between Latin and Ancient Greek, we aim to gain insights into cross-linguistic tendencies in geographical semantics and contribute to the broader study of diachronic conceptual change.

2 Related Work

Research on historical lexical semantics distinguishes two complementary perspectives: semasiology (from words to concepts) and onomasiology (from concepts to words). Semasiology asks, given a word, which meanings it expresses – for example, Latin *civitas* can mean 'citizenship', 'the citizens', or 'a city'. Onomasiology asks, given a concept, which words encode it – for instance, <CITY> in Latin may be *urbs*, *oppidum*, *civitas*, or *metropolis*. These relations vary across genres, registers, and time, shaped by cognitive, social, and contextual factors (Meillet, 1958; Blank, 2001). Semasiological studies focus on polysemy and sense-frequency trajectories (Haspelmath, 1999; Geeraerts, 1997), while onomasiological analyses track concept lexicalisation and synonym evolution (Bréal, 1897; Lehrer, 1985). While qualitative studies have documented such phenomena for Latin and Greek (Haverling, 2012; Moussy, 1965; Fruyt, 1994; Torrego, 2020; Williams, 1976; Lehrer, 1985; Stolova, 2015; Adams, 2013), with a few exceptions (Farina, 2023) large-scale quantitative analyses are missing due to the lack of sense-annotated diachronic corpora.

Recent NLP research has introduced automatic methods for semantic change detection based on distributional semantics (Hamilton et al., 2016; Kutuzov et al., 2022). These approaches model words as vectors in high-dimensional spaces and track their shifts across time-defined subcorpora. While effective for modern languages, they typically treat lexemes as atomic units, ignoring sense-level vari-

Proceedings of SIGHUM 2026, pages 266–279

ation and genre effects. For historical languages, preliminary experiments on WSD using contextual embeddings (Latin BERT) achieved promising results (Bamman and Burns, 2020; Lendvai and Wick, 2022; Ghinassi et al., 2024; Mercelis et al., 2025), but coverage and temporal sensitivity remain limited. Lendvai and Wick (2022) fine-tuned Latin BERT on data from the *Thesaurus Linguae Latinae*, achieving substantial gains over static embeddings. Ghinassi et al. (2024) proposed a language-pivoting approach that propagates sense annotations from English to Latin via parallel corpora. Kaše et al. (2025) explored transformer-based embeddings and attention mechanisms for WSD using dictionary senses.

On the onomasiological side, resources such as WordNet (WN) (Fellbaum, 1998) and its Latin and Ancient Greek counterparts (Minozzi, 2010) provide structured sense inventories, yet they have not been systematically linked to diachronic corpora. This prevents quantitative analyses of how concepts were lexicalised over time. Existing diachronic corpora annotated with WN senses for other languages (e.g., Bulgarian, Dutch, Slovenian) were generated automatically without manual curation and cover only a few centuries (Martelli et al., 2022).

In this context, WN-style sense inventories provide a principled framework for operationalising an onomasiological perspective by anchoring lexical items to shared conceptual units (Fellbaum, 1998) (Section 3). WN has been successfully employed for – often cross-linguistic – semantic and lexical studies both on modern (Gonzalo et al., 1998; Burgun and Bodenreider, 2001; Apidianaki and Sagot, 2012; Rudnicka et al., 2022; Klimczak et al., 2024) and historical (Biagetti et al., 2023; Khan et al., 2023; Brigada Villa et al., 2025) languages, and it therefore offers a solid basis for constructing comparable sense inventories for historical languages.

Our study focuses on geographical nouns (GNs), intended as common nouns denoting geographical entities (as opposed to toponyms), a lexically and conceptually salient domain that is central to both narrative structure and cultural conceptualisation in Ancient Greek and Latin. Annotating GNs in historical texts is challenging because of diachronic change and sparse data, yet it is essential for linking lexical mentions to conceptual and real-world geographical domains (Erdmann et al., 2016; Beersmans et al., 2023; Kenyon et al., 2023). GNs attract attention in ethnophysiography and cognitive linguistics as evidence of how cultures conceptualise the natural and human-made environments, and are generally characterised by high cross-linguistic and cross-cultural variability (Burenhult and Levinson, 2008; Mark et al., 2011; Falcinelli et al., 2024). Compared to previous studies, this is the first computational study on two historical languages combining both the semasiological and the onomasiological perspectives, and employing LLMs, specifically targeted towards GNs.

3 Research Questions

Our study integrates sense-annotated diachronic resources for historical languages by combining onomasiological and semasiological perspectives with recent advances in LLMs. To operationalise this approach, geographical meanings are anchored in WN sense inventories, where each synset (synonym set) represents a concept shared across lexical items. Linking lexical forms to synsets allows us to analyse both how concepts are lexicalised across words and how words map onto multiple senses in a sense-aware way. This framework also allows for a principled integration of computational annotation: LLMs are applied to automatically disambiguate senses in a diachronically diverse corpus including multiple literary genres (Section 4), providing both linguistic and quantitative data on the lexicalisation of geographical concepts. Building on this framework, we focus on two main research questions (RQs):

RQ1: LLM-based annotation of geographical common nouns How accurately can LLMs annotate geographical senses of common nouns in Ancient Greek and Latin texts? How does model performance vary across languages, centuries, and literary genres?

RQ2: Onomasiological patterns of geographical lexicalisation How are geographical concepts (e.g., <CITY>, <SEA>) lexicalised through common nouns in Ancient Greek and Latin? How do the distributions of lemmas associated with the same geographical sense change diachronically and across literary genres? To what extent do these languages exhibit convergent/divergent onomasiological patterns?

4 Corpus Overview

The PREMOVE Base Corpus (Farina, 2026) is a diachronic and multi-genre collection of Ancient Greek and Latin texts originally developed for a

manually-curated comparative analysis of prefixed verbs of motion. The corpus was explicitly designed to enable cross-linguistic comparison by ensuring balanced coverage across time periods, genres, and authors in both languages. While originally constructed for the study of motion verbs, its comparable design and systematic sampling make it equally suitable for investigating GNs and other geographical expressions (Farina et al., 2025). As is the case for all historical languages, the fragmentary survival of texts and their uneven transmission prevent the construction of a fully representative corpus (McEnery and Wilson, 2001). The corpus is therefore conceived as a controlled and purpose-oriented sample, structured to support comparative investigations of lexical and semantic phenomena.

The corpus builds on a subset of the Ancient Greek and Latin Dependency Treebank (AGLDT) (Bamman and Crane, 2007, 2011), supplemented with additional works to improve chronological continuity and genre coverage. It comprises 35 texts (19 Ancient Greek, 16 Latin) from the 8th cent. BCE to the 2nd cent. CE, covering major linguistic phases – Archaic, Classical, Hellenistic, and Imperial Greek (Charalambakis, 1996; Babiniotis, 2002; Ralli, 2012) and Early, Classical, and Post-Classical Latin (Cuzzolin and Haverling, 2009). Texts span historiography, poetry, philosophy, theatre, oratory, and the novel (labels from Biagetti et al., 2021), allowing us to evaluate whether the semantic behavior of place nouns varies across registers, including figurative or freer-word-order contexts.

The full corpus contains 541,620 tokens (punctuation excluded), divided almost evenly between Ancient Greek (275,913 tokens; 50.9%) and Latin (265,707 tokens; 49.1%). Historiography is the most represented genre (approximately 44% of Greek tokens and 36% of Latin tokens), followed by poetry (around 29% in both languages). For a more detailed breakdown of authors, works, centuries, and genres we refer to Appendix A and Farina (2026).

5 Annotation Pipeline

The selection of target GNs is grounded in the place vocabulary by Ballatore (2015). This is a structured collection of nouns describing both natural (e.g., rivers, mountains) and human-made (e.g., cities, roads) places. In addition to lexical coverage, this resource associates each entry with GeoNames[1], DBpedia[2] and WN (Bond and Bond, 2019). Although the resource is English-based, it offers a theoretically informed and systematically curated starting point for defining a cross-linguistic set of place nouns and for grounding their semantic interpretation.

Starting from this English inventory, candidate Latin equivalents were obtained through a combination of dictionary-based translation for well-attested and semantically central terms, and automated generation for less frequent or contextually complex ones. Primary lexicographic support for direct translation was provided by a suite of classic English-Latin dictionaries,[3] including Smith and Hall (1871), Döderlein (1874), Lewis and Short (1879), and Ogilvie (1901), which together cover a broad range of core vocabulary and synonymic nuances in Latin. Candidate lemmas were then manually reviewed, discarding entries that were semantically inappropriate, exhibited excessive ambiguity with respect to the intended geographical meanings, or were anachronistic (e.g., Eng. *airport*).

This procedure resulted in an initial list of 227 Latin lemmas. To ensure cross-linguistic comparability, these lemmas were then mapped to their Ancient Greek counterparts using the digitised version of Schrevel (1832). Latin entries for which no suitable Ancient Greek equivalent could be identified were excluded. The final inventory comprises 227 place-noun lemmas attested in Latin and 264 lemmas attested in Ancient Greek, which form the basis for the automatic annotation of this study.

Our data was obtained by extracting each token corresponding to an identified Latin or Ancient Greek place-noun lemma from the PREMOVE Base Corpus. The source texts were lemmatised with latinCy (Burns, 2023) and greCy (Myerston, 2024). Then, for each matching token, we stored its lemma, its context (the sentence it occurs in), additional metadata (author, language, passage reference), and gave it a unique identifier. The lemmatisation process accounted for dialect variation (e.g., *thálassa* vs *thálatta*) by linking variants to a stable form.

[1] https://www.geonames.org/.
[2] https://www.dbpedia.org/.
[3] https://latinitium.com/latin-dictionaries/.

6 GNs in the PREMOVE Base Corpus

Lemma Frequencies Before considering the LLM-based annotation, it is informative to examine the raw distribution of place-noun lemmas in the PREMOVE Base Corpus. The extraction procedure (Section 5) yielded 16,681 occurrences (Table 1), of which 252 had already been manually annotated and serve as our validation set (Section 8). Of the remaining 16,429 occurrences, 8,555 are from Greek and 7,874 from Latin. This roughly balanced distribution provides a solid foundation for cross-linguistic comparisons of geographical terminology and contextualises the frequency patterns of lemmas prior to any automated sense annotation.

Table 1: Dataset distribution

	Latin	Ancient Greek	Total
Unlabelled data	7,874	8,555	16,429
Validation set	106	146	252
Total	7,980	8,701	16,681

Table 2 presents the ten most frequent lemmas for each language. The frequency data for Greek and Latin lemmas shows a strong focus on urban, territorial, and natural features, though with some language-specific emphases. In Greek, the most frequent lemmas are *pólis* 'city' (1,537 occurrences), *thálassa* 'sea' (384), *archḗ* '(ruled) territory'[4] (373), *nêsos* 'island' (338), and *teîkhos* 'wall' (304). The least frequent Greek lemmas include specialised or context-specific terms such as *bórboros* 'mud', *domḗ* 'building', *kamára* 'vaulted chamber', *kapēleîon* 'tavern', and *xenṓn* 'guest chamber'. Latin shows a broadly similar pattern with the most frequent lemmas being *urbs* 'city' (721), *domus* 'house' (482), *terra* 'land' (310), *civitas* 'community, city' (271), and *ager* 'field' (261), while the least frequent terms include *Oceanus* 'Ocean', *ager publicus* 'public land', *apotheca* 'storehouse', *canalis* 'channel', and *zona* 'zone', reflecting more specialised contexts.

Groupings Examining groups of GNs reveals interesting parallels and differences between Greek and Latin (Table 3). In both languages, urban centres (Gr. *pólis, ástu*, Lat. *urbs, civitas*) dominate frequency counts, followed by natural features such as seas and rivers (Gr. *thálassa, pélagos, póntos, háls*, *potamós*, Lat. *mare, pelagus*, *flumen*), mountains (Gr. *óros*, Lat. *mons*), and fields or land (Gr. *pedíon, khthôn*, Lat. *ager*). Greek texts show a relatively higher occurrence of maritime terms (*thálassa, pélagos, háls, póntos* sea', *nêsos* island'), which is consistent with the historically central role of the sea in Greek civilisation, both for trade (Reed, 2003), colonisation (Nash, 2018), and cultural exchange (Lindenlauf, 2004; Beaulieu, 2016). Latin, by contrast, emphasises inland features, administrative regions, and cultivated land (*terra* 'earth, land', *ager* 'land', *provincia* 'province', *regnum* 'kingdom'), reflecting the Roman focus on land (Roselaar, 2010; Zuiderhoek, 2015).

Greek		Latin	
Lemma	**#**	**Lemma**	**#**
pólis 'city'	1,537	*urbs* 'city'	721
thálassa 'sea'	384	*domus* 'house'	482
archḗ '(ruled) area'	373	*terra* 'land'	310
nêsos 'island'	338	*civitas* 'city'	271
teîkhos 'wall'	304	*ager* 'land'	261
khôra 'region'	299	*regnum* 'kingdom'	252
nomós 'district'	259	*via* 'street'	210
potamós 'river'	239	*finis* 'border'	206
khōríon 'region'	213	*aedes* 'house, temple'	200
oîkos 'house'	202	*provincia* 'province'	181

Table 2: Most frequent GNs in Greek and Latin.

Diachrony When normalising for total tokens per century, the distribution of place-related vocabulary in Greek and Latin texts becomes clearer. In Greek, the most evident difference lies in the chronological gap between the Homeric lexicon (8th cent. BCE) and Classical Greek (5th cent. BCE). The most frequent place name in Homer is *dôma* 'house', dominating with roughly 22.8 occurrences per 10,000 tokens, followed by *póntos* 'sea' (17.4) and *oîkos* 'house' (13.7). This distribution also reflects the composition of the corpus: the Homeric material for the 8th cent. BCE includes a portion of the *Odyssey* (Appendix A), which features many scenes set around houses and numerous references to the sea due to Odysseus' journey.

In the 5th cent. BCE, *pólis* 'city' becomes extremely frequent at 98.7 per 10,000 tokens, followed by *teîkhos* 'wall' and *thálassa* 'sea' at only 23.3 and 19.8 per 10,000 tokens. Notably, the terminology for 'sea' changes by this period: *thálassa* is now preferred over *póntos* (Section 9.2). As for *pólis*, it remains consistently among the top three most frequent words up until the 2nd cent. CE.

[4]The main meanings of *archḗ* are 'origin' and 'power', which accounts for its high overall frequency. The semantic annotation (Section 8) was used to isolate those instances in which *archḗ* functions as a GN.

Semantic Group	Greek (occurrences)	Latin (occurrences)
City / Town	*pólis* 'city' (1537), *ástu* 'town' (111), *agorá* 'marketplace' (108), *oikía* 'house' (80), *hierón* 'temple' (91)	*urbs* 'city' (721), *civitas* 'community, city' (271), *oppidum* 'town' (131), *vicus* 'village' (50), *colonia* 'colony' (46)
Sea / River	*thálassa* 'sea' (384), *potamós* 'river' (239), *pélagos* 'sea' (75), *póntos* 'sea', *háls* 'sea' (10), *ōkeanós* 'ocean' (24)	*mare* 'sea' (178), *flumen* 'river' (154) *pelagus* 'sea' (49), *pontus* 'sea' (66), *oceanus* 'ocean' (31)
Mountain / Hill	*óros* 'mountain' (140), *hóros* 'mountain' (140), *lóphos* 'hill' (53), *líthos* 'rock' (52), *skópelos* 'peak' (13)	*mons* 'mountain' (160), *saxum* 'rock' (116), *collis* 'hill' (44), *rupes* 'peak' (31), *scopulus* 'peak' (45)
Land / Territory	*khthṓn* 'land' (102), *pedíon* 'plain' (87), *agrós* 'field' (37), *erēmía* 'desert' (64), *khorós* 'region' (299)	*terra* 'earth,land' (310), *ager* 'field' (261), *campus* 'plain' (133), *regio* 'region' (83), *patria* 'territory' (141)
Wall / Fort	*teîkhos* 'wall' (304), *phulakḗ* 'garrison' (128), *púrgos* 'tower' (49)	*murus* 'wall' (148), *moenia* 'wall' (144), *praesidium* 'garrison' (103), *castellum* 'fort' (16)

Table 3: Frequency of major semantic groups for GNs in Greek and Latin.

Latin shows a similar pattern. In the 3rd cent. BCE, *aedes* 'house, temple' and *domus* 'house' are the most frequent, with 44.4 and 23.4 per 10,000 tokens respectively. Across all centuries, *urbs* 'city' and *domus* 'house' remain extremely frequent. These distributions are also shaped by genre (see below): for instance, the early Latin texts in our corpus are dominated by comedies, which often emphasise domestic settings, explaining the prominence of house-related vocabulary.

Genre In Greek, *pólis* 'city' largely dominates historiography with 101.2 occurrences per 10,000 tokens, far surpassing any Latin counterpart, where the top word *aedes* 'house, temple' reaches only 44.4 per 10,000 tokens in the 3rd cent. BCE. This prominence of *pólis* may reflect the centrality of the concept of <CITY> in the Greek history: city-states (e.g., Athens, Sparta) were fundamental for history of Greece, so historiographical narratives often revolve around individual cities, their governance, and their interactions (Rood, 2001; Azevedo, 2010; Stefanovski and Čavoški, 2023; Harris and Lewis, 2024). By contrast, in Roman historiography, *urbs* primarily refers to Rome itself (Witcher, 2005), while other words such as *oppidum*, *castellum*, or *vicus* denote more specific types of settlements (Becker, 2008), resulting in a more distributed lexical pattern for urban terms.

Other Greek words show strong genre specialisation: *nomós* 'district' is highly frequent in oratory (78.9 per 10,000 tokens), *thálassa* 'sea' and *teîkhos* 'wall' in historiography (22.7–23.1 per 10,000 tokens), and *dôma* 'house' and *póntos* 'sea' in poetry (14.4–15.3 per 10,000 tokens), whereas philosophy prefers terms like *arkhḗ* (41.8 per 10,000 tokens) – often employed with its primary sense of 'beginning, authority' – and *pólis* (39.0 per 10,000 tokens). In Latin, the picture is broadly similar but less extreme: *aedes* and *domus* dominate across genres, *urbs* 'city' appears frequently in oratory (66.2 per 10,000 tokens) and historiography (42.2 per 10,000 tokens), and *civitas* 'community' is most frequent in historiography (22.9 per 10,000 tokens).

Some correspondences emerge: both Greek and Latin historiography privileges words for 'city' and 'walls' (Gr. *pólis*, Lat. *urbs*; Gr. *teîkhos*, Lat. *murus*), while poetry in both languages seems to favour natural and domestic space (Gr. *dôma, póntos, nêsos*; Lat. *domus, terra, saxum*). Greek exhibits higher lexical peaks and more pronounced genre-specific distributions, suggesting greater stylistic and lexical flexibility, whereas Latin shows more even frequencies across genres.

Specialised terms, such as Gr. *stratópedon* 'camp', *diéxodos* 'passage', or *múlē* and Lat. *praesidium* 'camp', *palus* 'marsh', or *balneum* 'bath', are concentrated in historiography or poetry. This pattern can be explained on two grounds: in historiography, the more technical or administrative nature of the discourse naturally calls for precise terminology, so specialised terms appear frequently (Ampolo et al., 2004; Cardinali, 2017); in poetry, the use of rare words may reflect particular imagery or expressive nuance, meter constraints, or a desire to avoid repetition (Ferri, 2011).

7 Experimental Setup

During our automatic annotation procedure, we tasked an LLM with assigning a WN synset to each token (Section 3). We selected GPT-5.2 (with medium reasoning effort) as our annotator. This is the most advanced iteration of the GPT-5 family and demonstrated strong potential for Latin and An-

cient Greek semantic annotation among 13 model families in Farina and Ciletti (2026). Sample tests confirmed its competence in handling GNs, and we avoided models marked as deprecated by their creators (such as GPT-5) to ensure reproducibility.

For each Latin token, we retrieved the WN synsets associated with its lemma via the Latin WordNet (LWN) API[5]. Additionally, we retrieved all synsets associated with its English counterpart(s) in the Open English WordNet (OEWN) (McCrae et al., 2020) based on our initial mapping, to ensure maximum semantic coverage. For Ancient Greek, whose WN is still under construction (Marchesi et al., 2025), we followed a similar procedure: each token was mapped to its corresponding Latin and English lemma(s), and candidate synsets were stored from both the LWN and OEWN. The model input included the target token, its lemma, context, and the list of candidate synsets. The synsets were presented with glosses and customised IDs in a randomised order, stripped of any information regarding their source language. The prompt also included detailed instructions (Appendix B) and up to five few-shot examples of the same lemma from the validation set. The LLM was instructed to return the most appropriate synset, or to explicitly reject all candidates if none were accurate, along with a binary classification on whether the meaning was literal (`yes`) or metaphorical/metonymic (`no`), and a confidence score (0–1).

8 Results

We evaluated the model against our gold standard of 252 annotated tokens (Section 6). Across all tokens, including cases where the model abstained (`NA`), it achieved a precision of 0.583, recall of 0.966, and F1 of 0.727, reflecting the high frequency of predictions but the challenge of fine-grained sense disambiguation (see below). The model predicted a sense for 98.0% of tokens, and on these committed predictions, the precision was 0.583. As this metric considers only tokens with predictions, precision, recall, and F1 are identical, reflecting accuracy when the model commits. The mean confidence score given by the model itself was 0.85 – high even for incorrect predictions, but proportionally higher for correct annotations –, and 94.7% of tokens were tagged as literal, reflecting their status as GNs.

[5] https://latinwordnet.exeter.ac.uk/api.

Languages Breaking down performance by language (Table 4), Latin tokens achieved slightly higher precision on committed predictions (0.594), while Greek tokens scored lower (0.574). Inter-annotator agreement between human and the model's annotation, measured with unweighted Cohen's Kappa, also varied slightly: Latin achieved 0.562 and Greek 0.542, indicating moderate agreement in both languages.

These results suggest that the model is slightly more conservative in Latin, avoiding some false positives, while Greek tokens – despite higher coverage and more examples per token – show marginally lower precision. Overall, literal GNs are reliably annotated, whereas metaphorical or less frequent senses remain challenging, highlighting the limits of current LLM-based WSD for fine-grained WN synsets.

Language	Precision
Greek	0.574
Latin	0.594
All	0.583

Table 4: Model precision by language and overall.

Previous Latin WSD studies report higher F1 scores on standard benchmarks (Ghinassi et al., 2024: 0.618; Lendvai and Wick, 2022: 0.794; Kaše et al., 2025: 0.68). We compare our precision on committed predictions to their F1 because, in the committed subset, precision, recall, and F1 are identical. Comparable LLM-based approaches to semantic analysis using English meanings rather than WordNet synsets achieved F1 scores of 0.596 for Latin and 0.666 for Greek (Farina and Ciletti, 2026).

Centuries and Genres We also examined model performance across centuries and macro-genres (prose/poetry) to assess temporal and stylistic variation (Table 5). Overall, Latin achieves higher precision values across most periods and shows a relatively homogeneous pattern, although data for the 3rd century BCE and 2nd century CE are very limited. Greek is also fairly homogeneous, with peaks in the 8th and 5th centuries BCE; these likely reflect the higher volume of texts from these periods rather than systematic differences in model performance.

Macro-genres were used because coverage is highly uneven across more specific genres. In Latin, poetry exhibits higher precision than prose (0.692 vs. 0.563), whereas Greek is more homogeneous

across poetry and prose (0.588 vs. 0.570). This suggests that the model handles Greek texts relatively consistently across genres, while in Latin the differences between poetry and prose may reflect the particular types of GNs present in poetry within the PREMOVE Base Corpus.

Lang.	Genre	Precision
Greek	Poetry	0.588
Greek	Prose	0.570
Latin	Poetry	0.692
Latin	Prose	0.563

Table 5: Model performance by macro-genre.

Qualitative Insights An analysis of the synsets with the highest precision reveals clearer patterns. After excluding single-occurrence synsets, 12 achieved a perfect precision (1.0), while five others fell within a moderate range of 0.625–0.928. All other synsets failed to score higher than 0.375. The highest-performing synsets generally benefit from distinct glosses and low semantic ambiguity. For instance, `02901994-n` "a structure that allows people or vehicles to cross an obstacle such as a river or canal or railway etc.", was unequivocally assigned to tokens linked to Eng. *bridge* (Lat. *pons*). In general, high-scoring synsets are related to the following GNs: human-made places ('camp', 'temple', 'house'), water masses ('lake', 'river'), and land masses ('land').

Most discrepancies appeared to stem from issues of granularity (Burgun and Bodenreider, 2001; Bond and Piasecki, 2017; Kafe, 2017; McGillivray et al., 2023). WN synsets can be highly specific, and multiple similar senses may exist for a single lemma. Consequently, many annotation divergences reflected slightly different interpretations rather than outright errors. Consider, for instance, synsets `08542298-n` "a large and densely populated urban area; may include several independent administrative districts" (Eng. *city*) and `08683242-n` "an urban area with a fixed boundary that is smaller than a city" (Eng. *town*). Precision for `08542298-n` was high (0.815), while for `08683242-n` it dropped to 0.375. A closer examination shows that in all misclassified cases the model predicted `08542298-n` instead of `08683242-n`. This reflects the fact that, in some passages, Lat. *urbs* and Gr. *pólis* – normally denoting larger urban centres – actually refer to smaller settlements, but the model defaulted to the more common, prototypical sense for these lemmas.

9 Discussion

In this section, we outline general tendencies in the semantic development of place-related lemmas in Latin and Greek, before turning to two case studies in Sections 9.1 and 9.2. Although the overall precision of the automatic annotation is moderate (Section 8), the case studies focus on high-frequency concepts for which the model achieves a precision of 0.660 (i.e., <CITY> and <SEA>). A closer examination of the 18 errors shows that in 13 cases, apparent mistakes stem solely from the fine granularity of WN synsets, with only 5 representing genuine misannotations. Some of these apparent errors involve, for example, synsets for towns being annotated as cities and vice versa (Section 8); since these still denote closely related concepts, they do not affect our analysis, which examines how these concepts are lexicalised and whether some lemmas undergo genuine semantic shifts across distinct meanings.

In Latin, the data show that certain lexical items undergo noticeable shifts in dominant meaning over time, often reflecting cultural and functional changes in the environment. For instance, *aedes* initially denotes a dwelling (42 per 10,000 occurrences in the 3rd cent. BCE) but gradually acquires a prominent secondary meaning as a temple or place of worship (0.94 per 10,000 occurrences in the 3rd cent. BCE; 4.8 per 10,000 occurrences in the 1st cent. BCE; 15.4 per 10,000 occurrences in the 2nd cent. CE), illustrating a narrowing of sense toward religious buildings. Similarly, *moenia*, originally a general term for architectural partitions, increasingly refers specifically to defensive city walls, while *villa* shows a shift from extensive rural estates to domestic dwellings, with residual usage for large properties.

Greek place-related lemmas, by contrast, tend to display more gradual and distributed changes across coexisting senses, rather than abrupt reorganisations of dominant meaning. Semantic evolution often manifests as moderate redistributions of lexical prominence, suggesting slower or more incremental shifts in sense hierarchy. While the Greek data do not show the sharp transitions observed in Latin, this apparent stability may partly result from the temporal resolution of the available corpus (8th cent. BCE – 2nd cent. CE): longer diachronic windows might reveal more substantial reorganisation of sense dominance (Section 10).

9.1 The concept of <CITY>

The concept of <CITY> provides a productive case study for examining semasiological organisation and diachronic distribution in Greek and Latin. Quantitatively, Greek exhibits a higher overall lexical dispersion for the concept, with 1,585 tokens distributed across several lemmas, compared to 1,057 tokens in Latin. In Greek, the lexicalisation of <CITY> is strongly dominated by *pólis* (Section 6), which alone accounts for 1,457 tokens, while *ástu* (111 tokens) and *políteuma* (17 tokens) play marginal and more specialised roles. Latin, by contrast, shows a more balanced lexical system: although *urbs* is the most frequent term (693 tokens), it is complemented by *civitas* (222 tokens) and *oppidum* (121 tokens), with minor contributions from *patria* (14 tokens) and *municipium* (7 tokens).

This distribution reflects semantic differentiation rather than lexical competition. In Greek, *pólis* is the unmarked term for the city, *ástu* denotes the citadel or built core, and *políteuma* refers to the civic body rather than urban space. Latin shows a similar but less polarised pattern: *urbs* is the general urban term, *civitas* overlaps with 'city' and 'citizenship', and *oppidum*, *patria*, and *municipium* mark more specific urban notions (McGillivray and Nowak, 2025).

From a semantic perspective, both languages cover comparable conceptual domains, but distribute them differently across their lexicons. In Greek, `08542298-n` "large and densely populated urban area" (897 tokens) and `08185877-n` "politically organized body of people under a single government" (593 tokens) are both largely subsumed under *pólis*. Conversely, Latin displays a sharper semantic partitioning: while `08542298-n` is predominantly encoded by *urbs* (731 tokens), the political sense is primarily associated with *civitas* (202 tokens), and smaller settlements (`08683242-n` "an urban area with a fixed boundary that is smaller than a city") are lexicalised with *oppidum* (124 tokens). Genre distribution reveals a contrast: in both languages historiography dominates quantitatively, but Greek <CITY> terminology remains overwhelmingly centred on a single lemma across genres, while Latin exhibits greater lexical and semantic diversification.

Semasiologically, the Latin lemma *civitas* emerges as a particularly revealing example of semantic shift within a system that explicitly differentiates between civic community ('citizenship') and urban space ('city'). In the present corpus, this semantic shift becomes more evident from the 1st cent. BCE onward, when the initially dominant sense of *civitas* as a politically organised body of citizens progressively redistributes toward the sense of a densely populated urban area (Figure 1). This development is a case of metonymic extension, whereby the collective entity of citizens comes to stand for the physical space they inhabit.

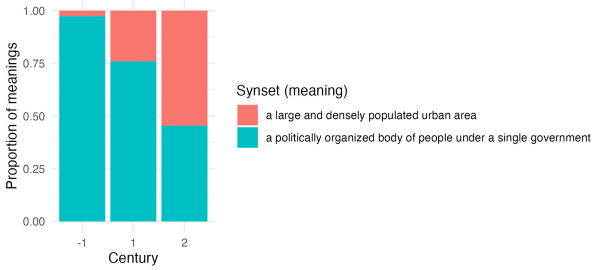

Figure 1: Semantic shift of Lat. *civitas*.

9.2 The concept of <SEA>

The concept of <SEA> is lexicalised in both Greek and Latin by four etymologically distinct lemmas. In Greek, it is expressed through *thálassa*, *póntos*, *pélagos*, and *háls*,[6] collectively covering two similar synsets: `09368829-n` "any very large body of (salt) water" (288 tokens) and `09449666-n` "a division of an ocean or a large body of salt water partially enclosed by land" (225 tokens). *Thálassa* dominates the corpus with 307 tokens (ca. 60% of all occurrences), followed by *póntos* (139 tokens), *pélagos* (60), and *háls* (7).

At the genre level, *thálassa* is mostly employed in historiography (273 tokens), while *póntos* in epic poetry, probably reflecting their etymologies. This preference is statistically confirmed by a Chi-squared test ($\chi^2 = 727.83$, $p < .001$). The preference for *póntos* in epic does not imply that *thálassa* is absent from Homeric poetry. Rather, qualitative inspection of the contexts shows a systematic semantic contrast. *Thálassa* tends to denote the sea as an unmarked geographical setting, without an explicit focus on movement or traversal, as in (1), where it simply designates the maritime landscape. By contrast, *póntos* typically occurs in contexts of motion, journeying, or crossing, foregrounding the sea as a route or medium of transport, as in (2).

[6]The lemmas differ etymologically (Beekes, 2009): *thálassa* is the unmarked term for the sea; *póntos* denotes the sea as a route or crossing; *pélagos* refers to the open or high sea; *háls*, etymologically 'salt', denotes the sea by its substance (metonymy).

(1) *Tēlémakhos d' apáneuthen iõn epì* ***thîna thalássēs***
'Telemachus went far off to **the shore of the sea**.' (Od. 2.260)

(2) ***Pléōn epì oínopa pónton*** *ep' allothróous anthrṓpous*
'**Sailing on the wine-dark sea** to men of another language.' (Od. 1.183)

In Latin, <SEA> is lexicalised through *mare*, *pontus*, *aequor*, and *pelagus*,[7] covering the same two synsets as Greek: `09368829-n` (214 tokens) and `09449666-n` (73 tokens). Here, *mare* and *aequor* are the dominant forms, with poetry accounting for the majority of occurrences (190 tokens), followed by historiography (56 tokens) and theatre (32 tokens). *Aequor* also provides an example of semantic change within the corpus: its primary sense of 'flat surface' diminishes over time, while the sense of 'sea' rises in the 1st cent. BCE and becomes dominant, reflecting a metonymic shift from a general flat expanse to the specific surface of the sea (Figure 2).

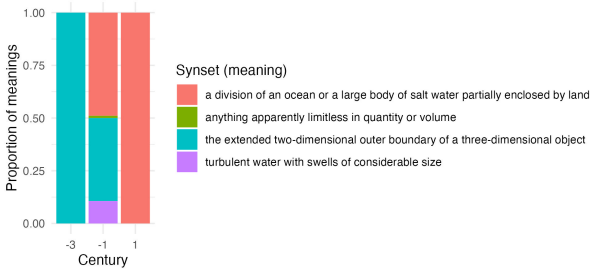

Figure 2: Semantic shift of Lat. *aequor*.

10 Conclusion and Future Work

This study demonstrates that combining LLM-based WSD with an onomasiological and semasiological framework may provide new insights into the lexicalisation of geographical concepts in Latin and Ancient Greek. Greek shows higher lexical dispersion and semantic conflation, while Latin exhibits sharper semantic partitioning and gradual sense shifts. Genre and diachrony modulate lexical patterns, highlighting the interaction between linguistic, cultural, and textual factors. Due to space constraints, we focused on two case studies, while the full annotated dataset is made available on Zenodo (Ciletti et al., 2026).

Although the model proved useful for comparative onomasiological analyses, its overall accuracy was moderate, reflecting challenges in disambiguating fine-grained WN synsets. Future work should explore fine-tuning alternative models and systematically investigating the best strategies for conducting onomasiological and semasiological analyses. Preliminary studies (Farina and Ciletti, 2026) show that fine-tuned open models can reach F1 scores up to 0.85 when assigning English meanings to Latin and Ancient Greek verbal prefixes, doubling the scores of their base versions. It still remains to be tested whether this strategy is feasible for assigning meanings via WN synsets. Moreover, expanding the corpus diachronically and by adding further texts may provide a clearer picture of semantic change, enabling more reliable identification of when meanings shift and when concepts come to be expressed by different lexical items, and contributing to the development of historical lexical resources.

Acknowledgments

This work is supported by the UKRI under the Horizon Europe Guarantee (grant number UKRI947) for the project COALA (Computational Corpus Annotation for Quantitative Analysis of Latin Lexical Semantics) successfully evaluated by the ERC, and by King's College London's AHRS Research Grant (Research & Scholarship Development Stream) for the project "Mapping meaning with Large Language Models".

Author Contributions

AF: Conceptualisation; Data curation; Formal analysis; Investigation; Methodology; Visualisation; Writing – original draft (Sections 3–6, 8–10); Writing – review & editing.
MC: Investigation; Software; Writing – original draft (Sections 5, 7, 8); Writing – review & editing.
BMcG: Conceptualization; Funding acquisition; Supervision; Writing – original draft (Sections 1, 2); Writing – review & editing.
AB: Conceptualization; Data curation; Writing – original draft (Section 2).

[7] The lemmas differ etymologically (de Vaan, 2008): *mare* is the general term for the sea; *pontus* and *pelagus* are borrowings from Greek (*póntos*, *pélagos*) and retain their original nuances; *aequor* derives from Latin *aequus* 'level, even' and originally denotes a flat expanse, later extending metonymically to bodies of water.

References

J. N. Adams. 2013. *Social Variation and the Latin Language*. Cambridge University Press, Cambridge.

Carmine Ampolo, Ugo Fantasia, and Leone Porciani. 2004. *Lexicon historiographicum Graecum et Latinum (LHG & L)*. Number 1 in Lexicon historiographicum Graecum et Latinum (LHG & L). Ed. della Normale, Pisa.

Marianna Apidianaki and Benoît Sagot. 2012. Applying cross-lingual WSD to wordnet development. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 833–840, Istanbul, Turkey. European Language Resources Association (ELRA).

Flávia Lemos Mota de Azevedo. 2010. The polis and the marvelous in the narrative of herodotus. *Revista Archai*, (1):33–42.

Georgios Babiniotis. 2002. *Lexikó Tis Néas Ellinikís Glóssas [Dictionary of Modern Greek]*. Athens: Kentro Lexicologias.

Andrea Ballatore. 2015. *Place Vocabulary*. GitHub.

David Bamman and Patrick Burns. 2020. Latin bert: A contextual language model for classical philology. arXiv preprint arXiv:2009.10053.

David Bamman and Gregory Crane. 2007. The Latin Dependency Treebank in a cultural heritage digital library. *Proceedings of the Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2007). Prague (Czech Republic)*, pages 33–40.

David Bamman and Gregory Crane. 2011. The Ancient Greek and Latin Dependency Treebanks. In *Language Technology for Cultural Heritage*, pages 79–98, Berlin/Heidelberg. Springer.

Marie-Claire Beaulieu. 2016. *The sea in the Greek imagination*. PENN/University of Pennsylvania Press, Philadelphia.

Hilary Becker. 2008. Urbs, oppidum, castellum, vicus: settlement differentiation and landscape nomenclature in Etruria. *La Città Murata in Etruria. Atti del XXV Convegno di Studi Etruschi ed Italici*, pages 73–80.

Robert Stephen Paul Beekes. 2009. *Etymological dictionary of Greek*. Number 10 in Leiden Indo-European etymological dictionary series. Brill, Leiden.

Marijke Beersmans, Evelien de Graaf, Tim Van de Cruys, and Margherita Fantoli. 2023. Training and evaluation of named entity recognition models for classical Latin. In *Proceedings of the Ancient Language Processing Workshop*, pages 1–12.

Erica Biagetti, Chiara Zanchi, and Silvia Luraghi. 2023. Linking the Sanskrit WordNet to the Vedic dependency treebank: a pilot study. In *Proceedings of the 12th Global Wordnet Conference*, pages 77–83, University of the Basque Country, Donostia - San Sebastian, Basque Country. Global Wordnet Association.

Erica Biagetti, Chiara Zanchi, and William M. Short. 2021. Toward the creation of wordnets for ancient indo-european languages. In *Proceedings of the 11th Global Wordnet Conference*, pages 258–266. Global Wordnet Association.

A. Blank. 2001. Words and concepts in time: Towards diachronic cognitive onomasiology. In *Words in Time: Diachronic Semantics from Different Points of View*, pages 37–65.

Francis Bond and Arthur Bond. 2019. GeoNames Wordnet (geown): extracting wordnets from GeoNames. In *Proceedings of the 10th Global Wordnet Conference*, pages 387–393, Wroclaw, Poland. Global Wordnet Association.

Francis Bond and Maciej Piasecki. 2017. Proceedings of the Workshop on Challenges for Wordnets. *Language, Data and Knowledge Conference. Galway, Ireland, 18 June 2017*.

Michel Bréal. 1897. *Essai de sémantique*. Paris: Hachette.

Luca Brigada Villa, Andrea Farina, and Chiara Zanchi. 2025. Formulaic Networks as Prototypical Categories: Combining the Ancient Greek Dependency Treebank with the Ancient Greek WordNet for a Pilot Study on the Iliad. In Jesús De La Villa, Araceli Striano, and Rodrigo Verano, editors, *Advances in Ancient Greek Linguistics*, pages 737–758. De Gruyter.

Niclas Burenhult and Stephen C Levinson. 2008. Language and Landscape: A Cross-linguistic Perspective. *Language sciences*, 30(2-3):135–150.

Anita Burgun and Olivier Bodenreider. 2001. Comparing terms, concepts and semantic classes in WordNet and the Unified Medical Language System. *Proceedings of NAACL'2001 Workshop, "WordNet and Other Lexical Resources: Applications, Extensions and Customizations"*, pages 77–82.

Patrick J Burns. 2023. Latincy: Synthetic trained pipelines for latin nlp. *arXiv preprint arXiv:2305.04365*.

Luca Cardinali. 2017. Alcune osservazioni sulla lingua degli storiografi arcaici. Pallas. Études de linguistique latine II(103):171–179.

Christophoros Charalambakis. 1996. Istoría tis ellinikís glóssas [history of the greek language]. *Encyclopedia Papyros-Larousse-Britannica: Greece*.

Michele Ciletti, Andrea Farina, Barbara McGillivray, and Andrea Ballatore. 2026. A diachronic dataset of semantically-annotated geographical nouns in ancient greek and latin.

Pierluigi Cuzzolin and Gerd V. M. Haverling. 2009. Syntax, sociolinguistics, and literary genres. In *New perspectives on historical Latin syntax*, pages 16–63.

Michiel de Vaan. 2008. *Etymological Dictionary of Latin*. Brill, Leiden / Boston.

Ludwig Döderlein. 1874. *Handbook of Latin Synonyms*. W. F. Draper, Andover.

Alexander Erdmann, Christopher Brown, Brian Joseph, Mark Janse, Petra Ajaka, Micha Elsner, and Marie-Catherine de Marneffe. 2016. Challenges and solutions for Latin named entity recognition. In *Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH)*, pages 85–93.

Ilenia Falcinelli, Chiara Fini, Claudia Mazzuca, and Anna M Borghi. 2024. The geo domain: a review on the conceptualization of geographical and geopolitical entities. *Frontiers in Psychology*, 15:1389581.

Andrea Farina. 2023. Lost at sea: A dataset of 25+ sea words morpho-semantically annotated in ancient greek and latin. *Journal of Open Humanities Data*, 9(1):24.

Andrea Farina. 2026. *A comparative study on Ancient Greek and Latin preverbed motion verbs. The creation of PrevNet as an expandable online resource for preverbs and preverbation*. King's College London: PhD Thesis.

Andrea Farina, Andrea Ballatore, and Barbara McGillivray. 2025. Mapping meaning in latin with large language models: A multi-task evaluation of preverbed motion verbs and spatial relation detection in llms. In *Proceedings of the Eleventh Italian Conference on Computational Linguistics (CLiC-it 2025)*.

Andrea Farina and Michele Ciletti. 2026. Probing preverbs: Evaluating large language models on Latin and Ancient Greek preverbed motion verbs. Forthcoming. In press.

Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Rolando Ferri. 2011. The Language of Latin Epic and Lyric Poetry. In James Clackson, editor, *A companion to the Latin language*, pages 344–366. Wiley-Blackwell, Malden, MA.

Michèle Fruyt. 1994. Typologie des cas de synonymie. In Claude Moussy, editor, *Les problèmes de la synonymie en latin*, pages 25–46. PUPS.

Dirk Geeraerts. 1997. *Diachronic Prototype Semantics: A Contribution to Historical Lexicography*. Oxford: Clarendon Press.

Iacopo Ghinassi, Simone Tedeschi, Paola Marongiu, Roberto Navigli, and Barbara McGillivray. 2024. Language pivoting from parallel corpora for word sense disambiguation of historical languages: A case study on Latin. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 10073–10084, Torino, Italia. ELRA and ICCL.

Julio Gonzalo, Felisa Verdejo, Carol Peters, and Nicoletta Calzolari. 1998. Applying EuroWordNet to Cross-Language Text Retrieval. *Computers and the Humanities*, 32(2-3):185–207.

William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1489–1501.

Edward M. Harris and David M. Lewis. 2024. City and country in the Greek world. In Miko Flohr and Arjan Zuiderhoek, editors, *A Companion to Cities in the Greco-Roman World*, 1 edition, pages 306–320. Wiley.

Martin Haspelmath. 1999. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft*, 18(2):180–205.

Gerd Haverling. 2012. Literary late latin and the development of the spoken language. In Martti Leiwo, Hilla Halla-aho, and Marja Vierros, editors, *Variation and Change in Greek and Latin*. Foundation of the Finnish Institute at Athens.

Eric Kafe. 2017. How Stable are WordNet Synsets? *Proceedings of the Workshop on Challenges for Wordnets*, pages 1–12.

Vojtěch Kaše, Sarah Lang, and Petr Pavlas. 2025. Embedded in the labyrinth: Investigating latin word senses through transformer-based contextual embeddings and attention. *Anthology of Computers and the Humanities*, 3:498–512.

Jeremy Kenyon, Jason W. Karl, and Bruce Godfrey. 2023. Evaluation of placename geoparsers. *Journal of Map & Geography Libraries*, 19(3):185–197.

Fahad Khan, John P. McCrae, Francisco Javier Minaya Gómez, Rafael Cruz González, and Javier E. Díaz-Vera. 2023. Some considerations in the construction of a historical language WordNet. In *Proceedings of the 12th Global Wordnet Conference*, pages 101–105, University of the Basque Country, Donostia - San Sebastian, Basque Country. Global Wordnet Association.

Karol Marek Klimczak, Jan Makary Fryczak, Dominika Hadro, and Justyna Fijałkowska. 2024. Equivalent sentiment measures for cross-language analysis of corporate communications. *MethodsX*, 12:102745.

A. Kutuzov, E. Velldal, and L. Øvrelid. 2022. Contextualized embeddings for semantic change detection: Lessons learned. *Northern European Journal of Language Technology*, 8.

A. Lehrer. 1985. The influence of semantic fields on semantic change. In Jacek Fisiak, editor, *Historical Semantics–Historical Word Formation*, pages 283–296. Mouton de Gruyter.

Piroska Lendvai and Claudia Wick. 2022. Finetuning latin bert for word sense disambiguation on the thesaurus linguae latinae. In *Proceedings of the Workshop on Cognitive Aspects of the Lexicon (CogALex)*, pages 37–41.

Charlton T. Lewis and Charles Short. 1879. *A Latin Dictionary. Founded on Andrews' edition of Freund's Latin dictionary. revised, enlarged, and in great part rewritten*. Clarendon Press, Oxford: Clarendon Press.

Astrid Lindenlauf. 2004. The sea as a place of no return in ancient Greece. *World Archaeology*, 35(3):416–433.

Beatrice Marchesi, Annachiara Clementelli, Andrea Maurizio Mammarella, Silvia Zampetta, Erica Biagetti, Luca Brigada Villa, Virginia Mastellari, Riccardo Ginevra, Claudia Roberta Combei, and Chiara Zanchi. 2025. Towards the semi-automated population of the ancient greek wordnet. In *Proceedings of the Eleventh Italian Conference on Computational Linguistics (CLiC-it 2025)*, pages 647–658.

David M Mark, Andrew G Turk, Niclas Burenhult, and David Stea. 2011. *Landscape in Language: Transdisciplinary Perspectives*, volume 4. John Benjamins Publishing.

F. Martelli, M. Maru, C. Campagnano, R. Navigli, P. Velardi, R.J. Ureña-Ruiz, F. Frontini, V. Quochi, J. Kallas, K. Koppel, and M. Langemets. 2022. D3.8 lexical-semantic analytics for nlp. Technical Report.

John P. McCrae, Alexandre Rademaker, Ewa Rudnicka, and Francis Bond. 2020. English WordNet 2020: Improving and extending a WordNet for English using an open-source methodology. In *Proceedings of the LREC 2020 Workshop on Multimodal Wordnets (MMW2020)*, pages 14–19, Marseille, France. The European Language Resources Association (ELRA).

Tony McEnery and Andrew Wilson. 2001. *Corpus Linguistics. An Introduction. Second edition*. Edinburgh: Edinburgh University Press.

Barbara McGillivray, Iacopo Ghinassi, and Paola Marongiu. 2023. The challenges of sense granularity in word sense disambiguation for latin (poster).

Barbara McGillivray and Krzysztof Nowak. 2025. Tracing the semantic change of socio-political terms from classical to early medieval latin with computational methods. In *Varietate delectamur: Multifarious Approaches to Synchronic and Diachronic Variation in Latin*, pages 357–570. Brepols Online.

Antoine Meillet. 1958. Comment les mots changent de sens. In *Linguistique historique et linguistique générale*, pages 230–280. Champion.

W. Mercelis, T. Van Hal, A. Keersmaekers, J. de la Villa, A. Striano, and R. Verano. 2025. Tongue, language or noise? word sense disambiguation in ancient greek with corpus-based methods. In *Advances in Ancient Greek Linguistics: Vol. 5.1*, pages 813–828. De Gruyter, Berlin/Boston.

Stefano Minozzi. 2010. The latin wordnet project. In *Latin Linguistics Today: Akten des 15. Internationalen Kolloquiums zur Lateinischen Linguistik*, pages 707–716, Innsbruck.

Claude Moussy. 1965. Gratus et iucundus. *Revue des Études latines*, 42:389–400.

Jacobo Myerston. 2024. greCy: Ancient Greek language models for spaCy.

John M. Nash. 2018. *Rulers of the Sea – Maritime Strategy and Sea Power in Ancient Greece 550-321 BC*. The Australian National University: PhD Thesis.

Robert Ogilvie. 1901. *Horae Latinae: Studies in Synonyms and Syntax*. Longmans, Green, and Co., London.

Angela Ralli. 2012. Greek. *Revue Belge de Philologie et d'Histoire*, 90(3):939–966.

C. M. Reed. 2003. *Maritime traders in the ancient Greek world*. Cambridge University Press, Cambridge, UK ; New York.

Tim Rood. 2001. Thucydides on the polis. *The Classical Review*, 51(2):241–242.

Saskia T. Roselaar. 2010. *Public Land in the Roman Republic: A Social and Economic History of Ager Publicus in Italy, 396-89 BC*. Oxford Studies in Roman Society and Law Ser. Oxford University Press, Incorporated, Oxford.

Ewa Rudnicka, Łukasz Grabowski, Maciej Piasecki, and Tomasz Naskręt. 2022. In Search of Gaps between Languages and Wordnets: the Case of Polish-English WordNet. *International Journal of Lexicography*, 36(1):68–92. _eprint: https://academic.oup.com/ijl/article-pdf/36/1/68/50049909/ecac005.pdf.

Cornelis Schrevel. 1832. *Lexicon manuale graeco-latinum et latino-graecum*. Collins and Hannay, Philadelphia.

William Smith and Theophilus Hall. 1871. *A Copious and Critical English-Latin Dictionary*. Harper & brothers, New York.

Mirjana Stefanovski and Kosta Čavoški. 2023. Polis, Loimos, Stasis: Thucydides about Disintegration of the Political System. *Conatus*, 8(2):629–656.

N. I. Stolova. 2015. *Cognitive Linguistics and Lexical Change: Motion Verbs from Latin to Romance*. John Benjamins, Amsterdam/Philadelphia.

Esther Torrego. 2020. The expression of knowledge in Latin: *cognosco, nosco, scio, nescio* and *ignoro*. In L. van Gils, C. Kroon, and R. Risselada, editors, *Lemmata Linguistica Latina, Volume II: Clause and Discourse*, pages 20–47. De Gruyter.

Joseph Williams. 1976. Synaesthetic adjectives: a possible law of semantic change. *Language*, 52(2):461–478.

Rob Witcher. 2005. The extended metropolis: *Urbs, suburbium* and population. *Journal of Roman Archaeology*, 18:120–138.

Arjan Zuiderhoek. 2015. Introduction: Land and Natural Resources in the Roman World in Historiographical and Theoretical Perspective. In Paul Erdkamp, Koenraad Verboven, and Arjan Zuiderhoek, editors, *Ownership and Exploitation of Land and Natural Resources in the Roman World*, pages 1–17. Oxford University Press.

A Corpus Composition

This appendix documents the textual composition of the *Anonymysed Corpus*, providing an explicit overview of the Ancient Greek and Latin works included. The selection spans multiple centuries and literary traditions in both languages, allowing for controlled comparison across historical periods and genres.

Table 6 reports the authors and texts included in the corpus, together with their chronological attribution, literary genre, and token counts. Works that are fully or partially annotated in the Ancient Greek and Latin Dependency Treebanks (AGDT/LDT) are indicated by an asterisk (*) or a degree symbol (º), respectively. The corpus combines diachronic breadth with genre diversity while taking advantage of existing syntactic resources, and it is intended to support both its original focus on preverbed motion verbs and its reuse for related studies, such as the analysis outlined in this paper. For a deeper discussion about the creation of the corpus, we refer to Farina (2026).

B System Prompt

This appendix presents the complete prompt template used for the automatic annotation. The complete data of this study are available on Zenodo (Ciletti et al., 2026).

System Prompt:

```
You are an expert in Latin and Ancient
Greek semantics, historical linguistics,
and semantic annotation.

# Your task for each query:

- You are given a TOKEN, its LEMMA,
LANGUAGE, a SENTENCE providing context,
and a list of candidate noun synsets.

- Each candidate synset has:

- a small integer ID (1, 2, 3, ...)

- an English gloss.

You may also be given gold-annotated
example sentences with the same lemma,
each with one gold sense gloss.

Use these examples only as guidance;
they may involve different senses than
the current sentence.

# Your job:

1. Choose at most one synset ID whose
gloss best matches the meaning of the
TOKEN as used in the given SENTENCE.

2. If none of the synsets plausibly
matches, select "n/a" instead of any
numeric ID.

3. Decide whether the
TOKEN is used literally
(compositional, non-figurative) or
metaphorically/figuratively in context.

4. Provide a confidence score from 0 to
1 for your overall decision.

## Important:

- Focus strictly on the actual usage in
the SENTENCE, not all possible meanings
of the lemma.

- Prefer precise matches over loose
associations.

- Treat metonymic or metaphorical
readings as "not literal".

# Output:

Return ONLY a single JSON object
matching the provided schema.
```

Table 6: The PREMOVE Base Corpus: Authors, Texts, Centuries, Literary Genres, and Token Counts. Texts fully or partially annotated in the AGDT/LDT are marked with * and º respectively.

Author	Text	Century	Literary Genre	Tokens
Ancient Greek				
Homer	*Odyssey* 1–12*	8th BCE	Poetry, epic	46,021
Pindar	*Olympians*	5th BCE	Poetry, lyric	5,953
Aristophanes	*Thesmophoriazusae*	5th BCE	Theatre, comedy	7,894
Aeschylus	*Agamemnon**	5th BCE	Theatre, tragedy	8,495
	*Prometheus Bound**	5th BCE	Theatre, tragedy	6,155
Sophocles	*Antigone**	5th BCE	Theatre, tragedy	7,658
Euripides	*Medea*	5th BCE	Theatre, tragedy	8,295
Herodotus	*Historiae* 1	5th BCE	Historiography	29,081
Lysias	*De caede Eratosthenis**	5th BCE	Oratory	2,478
	In Alcibiadem	5th BCE	Oratory	2,591
Plato	*Euthyphro**	5th BCE	Philosophy	5,417
Aristotle	*Athenian Constitution*	4th BCE	Philosophy	16,112
Thucydides	*Histories* 1*	5th BCE	Historiography	22,308
Apollonius Rhodius	*Argonautica* 1–3	3rd BCE	Poetry, epic	27,066
Polybius	*Histories* 1*	2nd BCE	Historiography	25,658
Diodorus Siculus	*Bibliotheca Historica* 11*	1st BCE	Historiography	23,336
Plutarch	*Alcibiades**	2nd CE	Historiography, biography	10,249
	*Lycurgus**	2nd CE	Historiography, biography	9,662
Lucian of Samosata	*Vera Historia*	2nd CE	Novel	11,484
Latin				
Ennius	*Annales*	3rd BCE	Poetry, epic	1,194
Plautus	*Amphitruo*	3rd BCE	Theatre, comedy	9,988
	Mostellaria	3rd BCE	Theatre, comedy	9,780
Caesar	*De bello Gallico* 1–4º	1st BCE	Historiography	20,498
Cicero	*In Catilinam* 1–3º	1st BCE	Oratory	11,625
	*De amicitia*º	1st BCE	Philosophy	9,471
Sallust	*Bellum Catilinae**	1st BCE	Historiography	10,655
Livy	*Ab Urbe condita* 1–2	1st BCE	Historiography	39,913
Vergil	*Aeneid*º	1st BCE	Poetry, epic	63,719
Propertius	*Elegiae* 1.1–1.22*	1st BCE	Poetry, elegy	4,384
Horace	*Satires* 1	1st BCE	Poetry, satire	7,048
Seneca	*De ira*	1st CE	Philosophy	22,614
	Medea	1st CE	Theatre, tragedy	5,693
Tacitus	*Historiae* 1º	1st–2nd CE	Historiography	11,852
Suetonius	*Life of Augustus*º	2nd CE	Historiography, biography	13,915
Apuleius	*Metamorphoses* 1–5	2nd CE	Novel	23,358