Deep Syntax and Semantics of Latin with UMR Thesis Proposal

Federica Gamba
Charles University
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
gamba@ufal.mff.cuni.cz

Abstract

The Latin language, evolving over two millennia across Europe, exhibits significant syntactic changes that challenge the accuracy of parsing models when applied to diverse datasets. This phenomenon, rooted in the language's diachronic and diatopic development, as well as in genre variations, has been observed in various studies and emphasized in the EvaLatin campaigns. Traditional surface-syntactic parsing struggles with these variations, prompting the need for deeper linguistic analysis.
Our research aims to transcend these limitations by investigating Latin through the lens of deep syntax and semantics, leveraging the Uniform Meaning Representation (UMR) framework. This approach aligns with recent NLP research trends focused on semantic representations, as representation of meaning can prove helpful for Natural Language Understanding. By annotating Latin data according to UMR, we seek to explore the potential of semantic analysis to provide a more consistent understanding of Latin texts and mitigate the impact of variability issues inherent in syntactic parsing. This investigation will also extend to a comparative analysis with Romance languages, enhancing the comprehension of the diachronic evolution of Latin.

1 Introduction

Over the course of more than two millennia and across a vast area that roughly corresponds to today's Europe, the Latin language has undergone numerous changes which have affected various linguistic layers, including syntax. Consequently, parsing accuracy scores on Latin texts tend to drop significantly when a model is applied to data that differ from those it was trained on, more so than is typically observed with out-of-domain data. This phenomenon is strongly tied to the diachronic and diatopic development of the language itself, as well as to differences due to genre (e.g. poetry/prose). For instance, Passarotti and Ruffolo (2010) and Ponti and Passarotti (2016) observed how performances drop when a model is employed to parse out-of-domain data, and Passarotti and Dell'Orletta (2010) addressed the need to adapt an existing parser to the specific processing of Medieval Latin. The issue of Latin variability has been a central focus of the EvaLatin campaigns (Sprugnoli et al., 2020; Sprugnoli et al., 2022, Sprugnoli et al., 2024), devoted to the evaluation of NLP tools for Latin. While the first two editions concentrated on lemmatization, morphological analysis and POS tagging, the third edition focused on dependency parsing, and highlighted the ongoing challenge of model portability across different literary genres.[1]

Besides addressing the variability issue at the syntactic level, it is natural to question whether shifting from a surface-level to a deeper representation would impact such challenge. Therefore, it is interesting to delve deeper into the syntax-semantics interface and examine how Latin behaves when it is processed at the semantic level. As we expect the linguistic variability observed with respect to syntax to decrease substantially when the language is addressed from the semantic perspective, we extend our investigation by taking the diachronic variability to the extreme and comparing the behavior of Romance languages to Latin. Despite being closely related, these languages display relevant differences, such as the presence or absence of a case system.

The proposal is structured as follows. In Section 2 an overview of the syntactic resources available for Latin is provided, and then the harmonization process undertaken to unify their annotations and rule out extrinsic levels of variability (2.1) is discussed. Subsection 2.2 delves into the EvaLatin 2024 Dependency Parsing Shared Task, explain-

[1] Moreover, one of the participating teams explicitly targeted model performance across different epochs (Behr, 2024).

ing the rationale behind the transition to a level of analysis deeper than surface syntax. Section 3 thus presents the decision to address semantics, and namely to adopt the Uniform Meaning Representation (UMR) formalism (3.1). Subsection 3.2 details on the need of lexical resources in order to annotate data according to UMR, with a particular focus on the resource available for Latin. In Subsections 3.3 and 3.4, an overview of the ongoing work on Latin is presented, as well as the plan and initial steps for the expansion of the corpus in a multilingual direction. Finally, Section 4 outlines future research directions and anticipated outcomes.

2 From Syntax

A significant number of resources is available for Latin. With respect to syntax, notable are the six treebanks in Universal Dependencies[2] (de Marneffe et al., 2021):

- **Index Thomisticus Treebank** (ITTB) (Passarotti, 2019): the largest of the Latin treebanks, it encompasses texts by Thomas Aquinas (1225–1274) and other authors related to Thomas, representing an example of philosophical Medieval Latin.
- **Late Latin Charter Treebank** (LLCT) (Cecchini et al., 2020b): it consists of Early Medieval (VIII-IX century) Latin charters written in Tuscany, Italy, all representing the legal/documentary genre.
- **Perseus** (Bamman and Crane, 2011): it includes some of the most representative Classical Latin texts (e.g., by Augustus, Cicero, Propertius, Sallust, Tacitus, Vergil) of different genres.
- **PROIEL** (Haug and Jøhndal, 2008): it contains most of the Vulgate New Testament translations, and selections from Caesar's *De bello Gallico*, Cicero's *Epistulae ad Atticum*, Palladius' *Opus Agriculturae* and the first book of Cicero's *De officiis*. Such texts are examples of Classical Latin, yet they represent different genres.
- **UDante** (Cecchini et al., 2020a): it includes literary texts (letters, treatises, poetry) by Dante Alighieri, representing an example of literary Medieval Latin (XIV century).
- **CIRCSE**: it contains poetry (tragedies *Agamemnon* and *Hercules Furens* by Seneca) and prose (Tacitus's *Germania*) texts dating back to the classical period (I-II centuries CE).

As it emerges from the provided overview, the treebanks highly differ in terms of included texts (e.g., genre, period) and size. However, despite the treebanks all following the UD annotation guidelines, some differences in the annotation scheme have also been observed. Specifically, the treebanks have been annotated by different teams and in different moments of the development of UD guidelines, resulting in different annotation choices. Therefore, all annotation levels – from word segmentation to lemmatization, POS tags, morphology, and syntactic relations – present divergences. The issue is not specific to Latin; for instance, it has been observed with respect to English by Zeldes and Schneider (2023) and with respect to Turkish by Akkurt et al. (2024) too.

2.1 Harmonization

In order to investigate genuine syntactic diversity, we first have to assess how much the observed drop in parsing performances is due to such divergences in annotation style. A deeper understanding, and possibly levelling of such divergences allows to isolate the impact of annotation choices and highlight intra-linguistic syntactic variability.

For the harmonization process we decided to model our interventions on the UDante v2.10 treebank. This choice is motivated by several factors: a) at the time when the harmonization was carried out, UDante was the only Latin treebank to have been annotated directly in UD,[3] rather than being converted from another framework, which allowed us to rule out conversion errors; b) it was the newest Latin treebank in UD, thus following the latest version of the UD guidelines; c) it is developed by the same team maintaining the other valid[4] Latin treebanks and defining the UD guidelines for Latin. In light of these reasons, we selected UDante as the Latin treebank most conforming to the current UD guidelines, and based on it the whole harmonization process.

[2] See https://universaldependencies.org/.

[3] In the current release (v2.14) this holds true for the CIRCSE treebank as well.

[4] Intended as a technical label of the UD infrastructure. See the UD Validation Report at http://quest.ms.mff.cuni.cz/udvalidator/cgi-bin/unidep/validation-report.pl.

The harmonization consisted in two main phases: first, we mainly focused on the syntactic layer (Gamba and Zeman, 2023b); however, some interventions were done also with respect to tokenization, lemmatization and POS tags. The second step focused only on morphological features (Gamba and Zeman, 2023a). Udapi (Popel et al., 2017), a framework providing an application programming interface for UD data, was exploited to manipulate data.

As an example of the first phase of the harmonization process, Figure 1 and 2 illustrate the sentence *Successio autem propter motum aliquem est.* 'Succession results from change of some kind.' (ITTB test-s2084). The original annotation considers the copula *est* to be the root of the tree, instead of attaching it to the correct head of the nominal predicate *motum*, annotated instead as an `oblique`. After the harmonization, the resulting tree (Figure 2) correctly follows the UD guidelines.

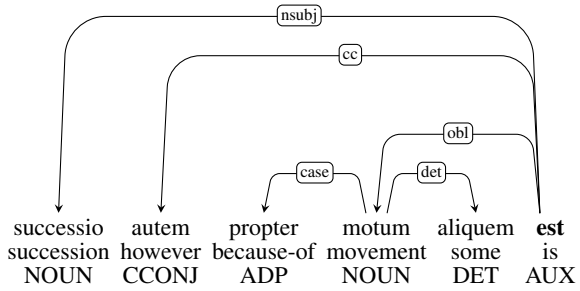

Figure 1: Annotation in UD 2.10.

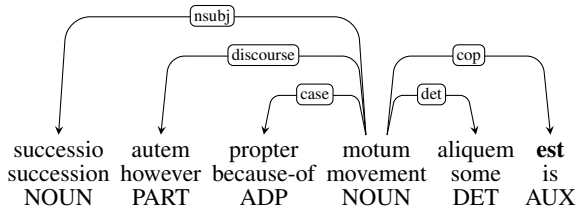

Figure 2: Harmonized annotation.

Results of the harmonization were two-fold. The assessment of the impact of the harmonization on parsing results proved that models trained on harmonized treebanks obtain better scores than the pre-harmonization ones. The improvements are notable especially with respect to results obtained on Perseus and PROIEL, which initially were the treebanks diverging the most. We observed all post-harmonization models to perform better on the two treebanks. The improvement is substantial (up to +9% in several cases), and is mirrored by parsing performances of models trained on harmonized Perseus and PROIEL, which achieve better scores on all the five treebanks, with peaks of around +17% both in LAS and UAS (LLCT parsed with a Perseus model). The parsing experiments confirmed once more the absolute relevance of a truly universal and shared annotation. Additionally, the harmonization process resulted in the contribution of enhanced resources. Indeed, the harmonized version of Perseus was directly contributed to UD official release (since v2.12), and a close collaboration with the team maintaining ITTB, LLCT and UDante led to the release of harmonized versions of such treebanks by the team itself.

2.2 EvaLatin 2024 Dependency Parsing Shared Task

The harmonization process allowed to rule out the effects of the additional layer of variability constituted by divergences in annotation, and thus set the conditions for a fair assessment of the variability of the language itself. An important milestone in this respect is represented by the EvaLatin 2024 Dependency Parsing shared task (Sprugnoli et al., 2024), which contributed to draw attention on Latin variability at the syntactic level.

The EvaLatin 2024 Dependency Parsing shared task worked as a test-bed for the harmonization efforts, which proved to be beneficial for the results obtained by the ÚFAL LatinPipe team (Straka et al., 2024). A significant difference was observed depending on whether models were trained on harmonized or non-harmonized treebanks. Specifically, test data did not contain any punctuation; the only UD treebank lacking punctuation at the moment of the shared task[5] was PROIEL. However, PROIEL is also the only treebank in the UD release which has not been consistently and thoroughly harmonized (unlike Perseus, ITTB, LLCT, and UDante, whose harmonized versions had already been incorporated in the official release). As a consequence of the lack of punctuation in test data, shared by PROIEL alone, the multi-treebank model trained by ÚFAL LatinPipe would pick PROIEL as the most similar, and thus relevant, treebank. Exploiting not the officially released PROIEL version but its harmonized counterpart for training thus proved beneficial, leading to around +3% improvements both in LAS and UAS.

[5]Currently, the CIRCSE treebank (released in v2.14) does not contain punctuation either. The two texts employed as test data in the shared task are now included in this treebank.

As far as the model architecture is concerned, ÚFAL LatinPipe – the winning submission to the shared task – consisted of a fine-tuned concatenation of base and large pre-trained LMs, with a dot-product attention head for parsing and softmax classification heads for morphology to jointly learn both dependency parsing and morphological analysis. The system was trained for a few initial epochs with frozen weights before fine-tuning. Local relative contextualization was added by stacking the BiLSTM layers on top of the Transformers, and output probability distributions from seven randomly instantiated networks were then ensembled.

Nevertheless, the shared task revealed that even with the incorporation of cutting-edge deep learning methodologies and access to high-quality harmonized data, the ÚFAL LatinPipe model still failed to surpass a 80% threshold, reaching a LAS of 75.75% on poetry and of 77.41% on prose. These findings underscore the intrinsic challenges posed by the linguistic variability of Latin, and suggest that models tested on data different from those leveraged for training are likely to keep posing comparable challenges. Consequently, the prospects for significant improvements in parsing scores appear limited, rather calling for a shift of perspective.

3 To Semantics

In light of what has been discussed, it thus seems hard to further improve the surface-syntactic parsing. However, morphosyntax can be viewed as an intermediate step to actual language understanding; we hence chose to delve further into exploring the potential and boundaries of deep syntax and semantics.

This choice reflects current directions of NLP research, as recent years have seen an enduring interest in annotation schemes that allow natural language texts to be parsed into semantic representations exploitable for tasks such as information extraction, machine translation, and other downstream purposes where understanding the meaning of text is crucial.

In line with the latest developments in the field, we have made the decision to explore the Latin language according to the Uniform Meaning Representation (UMR) framework, with the goal of releasing Latin data annotated according to UMR and investigating the implications and observations that can be derived from the annotation process.

3.1 Uniform Meaning Representation

Uniform Meaning Representation (UMR) (Van Gysel et al., 2021b) is a meaning representation framework designed to annotate the semantic content of a text. UMR is primarily based on Abstract Meaning Representation (AMR) (Banarescu et al., 2013), an annotation framework initially designed to representing the meaning of English sentences in a structured and abstract manner, independently of their surface syntax. Subsequent attempts to extend the formalism and adapt it to other languages have taken place, although without resulting in a truly cross-lingual framework. Cross-lingual AMR adaptations have been developed, for instance, for Czech (Urešová et al., 2014) and Chinese (Xue et al., 2014; Li et al., 2016), but overall most efforts have been focused on a restricted set of well-represented languages (notably English), as observed for instance by Vigus et al. (2020).

Conversely, UMR was explicitly developed with cross-linguistic scope in mind. Its main goal is to extend AMR to other languages, and in particular to morphologically complex, possibly low-resource languages, in a solid typological perspective. UMR improves on AMR in two major ways (Bonn et al., 2023b): first, it adjusts the AMR schema to make it more cross-linguistically applicable; secondly, AMR only includes sentence representation. On the one hand, UMR adds new semantic coverage to the schema by providing representation (i.e., graph elements) for tense, aspect, modality, and scope; on the other hand, it enhances the representation by designing document-level dependency structures for linguistic phenomena such as temporal and modal relations, as well as coreference, which may extend beyond sentence boundaries (Vigus et al., 2019; Pustejovsky et al., 2019). Overall, UMR is intended to be scalable, learnable, and cross-linguistically plausible, and it is designed to support both lexical and logical inference. The current release (UMR 1.0) (Bonn et al., 2023a, 2024) includes data for six languages: Arapaho, Chinese, Kukama, English, Navajo, Sanapaná.

Attempts to expand the framework to other languages are emerging, bringing along new (possibly language-specific) challenges to tackle. For instance, the UMR annotation of Chinese required addressing the issue of Chinese verb compounds, as compounding is a productive process in Chinese (Sun et al., 2023). The expansion of the UMR framework to other languages has also raised addi-

tional issues, which do not apply to English. Notably, a non-negligible challenge is represented by the lack of native speakers of a given language. Semantic annotation is typically done by native speakers, since native intuitions are assumed to be necessary to make judgments required for semantic annotation. This issue may be circumvented with low-resource languages that yet have millions of speakers, but it represents a core challenge when native speakers are not available (Van Gysel et al., 2021a). This will raise a problem for historical languages as well, for whose annotation no native speakers can be exploited. As of now, however, UMR annotation is not available for any historical language, and the issue has not been addressed yet.

Practically, the UMR annotation of each sentence is composed of three blocks: 1) the sentence-level graph, representing the meaning of the sentence itself; 2) the document-level graph, annotating temporal and modal relations as well as coreference beyond sentence boundaries; 3) alignments, clarifying the sentence tokens to which nodes in the sentence-level graph are aligned.

3.2 Need for Lexical Resources

The UMR framework is built upon predicate-argument structure annotation. Nodes in a UMR graph represent semantic concepts; semantic concepts are defined as word senses (if available), while PropBank (Kingsbury and Palmer, 2002) participant roles[6] associated to each predicate are included in the graph if realized in the sentence. For this reason, language resources encompassing information about semantic roles are essential to UMR annotation. Latin has a fair amount of data available, although far from what is available for example for English or Czech, both in quality and in quantity. However, those resources cannot be readily exploited for annotation purposes in light of their present condition.

More specifically, two such resources exist for Latin, i.e. the first and the second version of the Latin valency lexicon (Vallex). Despite the name, the two resources appear to be independent and do not really present points of contact - if not some of the encompassed lemmas.

The first version of Vallex (Passarotti et al., 2016) contains around 2,500 valency frames for more than 1,000 lexical entries and is stored as a single XML file. It is built upon the tectogrammatical layer of Latin texts annotated in the style of the Prague Dependency Treebank (PDT), which is based on Functional Generative Description (FGD) (Sgall, 1967). The tectogrammatical layer in PDT-like resources focuses on the syntactic-semantic properties of language; while keeping the dependency structure used at the surface-syntactic level, it also specifies semantic properties such as argument (valency) structure, predicate senses, semantic attributes of nodes, like tense, aspect, number, modality. In addition, elided arguments are restored as separate nodes, with the possibility of linking them e.g. by coreference, function words are removed and replaced with semantic relations similar to UMR roles. Besides coreference, other discourse relations that extend beyond sentence boundaries are annotated, as well as paratactic relations within sentences. Such information could be leveraged to facilitate the production of UMR annotated data through the conversion of already annotated information. Indeed, semantic representation is more difficult to obtain than for instance the syntactic one, so obtaining data (semi-)automatically would represent a significant contribution to our efforts. Tectogrammatical annotation for Latin is available only for a selected set of texts, which can yet prove very useful in this perspective. Namely, texts annotated in PDT style are the Index Thomisticus Treebank (ITTB), encompassing text by Thomas Aquinas, and a portion of the Latin Dependency Treebank (LDT) comprising works by Caesar, Cicero, and Sallust. We observe here a partial overlap of these resources with the Latin treebanks available in Universal Dependencies: namely, the whole ITTB is also annotated according to the UD formalism, and the same holds true for some of the texts included in the LDT (e.g., for portions of *De Coniuratione Catilinae* by Sallust, available in the Perseus UD treebank). The availability of different syntactic annotations for the same texts could provide us with yet another way to explore, and another source of information to exploit.

The first version of Vallex presents several limitations: first of all, the lack of definitions, which makes it hard to fully understand the intended meanings of the different frames, even more so in the case of a language with no native speakers. Secondly, there appears to be a redundancy of entries, which are unnecessarily distinguished even when they clearly refer to the same frame and meaning of a verb. Nevertheless, the resource cannot be completely overlooked, since it is directly linked to

[6] E.g., ARG0, ARG1.

(and built upon) the tectogrammatical layer of the treebanks which we could exploit to automatically obtain annotated data.

Latin Vallex 2.0 (Mambrini et al., 2021a) is a revision of the first version, but it adopts a different approach: it is intuition-based, which means that for each sense listed for a lemma or hypolemma, there is a valency frame, established on the basis of the dictionary meaning listed for that lemma. It contains about four times the entries of the first version, and through the link with WordNet (Franzini et al. 2019; Mambrini et al. 2021b) synsets it provides definitions.[7] However, the second version does not include examples, which would be very useful in understanding the definitions and distinguishing frames. Indeed, we already observed that many of the entries present extremely similar definitions as well as identical frames, which makes the resource not practical in terms of usability. For instance, we can observe two senses of *porto*, both with frame ACT (Actor), PAT (Patient), defined respectively as

definition	**synset_id**
have on one's person	v#00047745
have with oneself; have on one's person	v#02717102

Although not infrequent, such extreme cases are not the majority. *Metior* can serve as a more moderate example, yet still informative about Vallex granularity; see a list of its 9 synsets, all with frame ACT, PAT:

1. measure (distances) by pacing
2. determine the measurements of something or somebody, take measurements of
3. judge tentatively or form an estimate of (quantities or time)
4. evaluate or estimate the nature, quality, ability, extent, or significance of
5. set, mark, or draw the boundaries of something
6. determine the capacity, volume, or contents of by measurement and calculation
7. travel across or pass over
8. give out as one's portion or share
9. administer or bestow, as in small portions

Although with different nuances, synsets 1-6 all revolve around the concept of *measuring*, and are possibly too fine-grained for automatic detection. *Metior* does not represent an isolated occurrence, but a standard entry in Vallex.

In conclusion, the existing resources cannot be exploited as they are, and additional work is needed to obtain the kind of resource required for UMR annotation. Therefore, we envisaged a task of combination of the two resources, which however cannot be done automatically. The output of this task is expected to be Vallex4UMR, a new resource that does not disregard information contained in the available resources, but improves it in terms of completeness and efficacy, at the same time by making it exploitable for UMR purposes. Its coverage will be sufficient for the needs of UMR annotation, although it will not exhaustively cover the language. The current state of Vallex4UMR will be presented in the following subsection.

3.3 Current state of UMR for Latin

The first text that has been selected for UMR annotation is *De Coniuratione Catilinae* 'Conspiracy of Catiline', a historical monograph in 61 chapters written by Sallust in the I century BC. This choice is motivated by the availability of a version of the text annotated in the PDT format,[8] which could represent a valuable source of information to annotate the data according to the UMR formalism, and could possibly be leveraged for automatic conversion. However, for the time being the current state of the guidelines as well as the quality of released data renders manual annotation the only viable option. In fact, the annotation process itself represents a valid occasion to refine the guidelines, which have been observed to be often incomplete, unclear, and under-specified, as well as currently skewed towards English despite the stated cross-linguistic approach.

As of now, a sample of 50 sentences has been annotated, corresponding to the first five chapters of the text. The annotation of the first 50 sentences required a couple of months, primarily due to the need for thorough comprehension and discussion of the guidelines. We now anticipate the annotation process to proceed more quickly. An example of annotated sentence can be found in Appendix A. The annotation of the text has required to address

[7] Entries in Latin Vallex v2 are linked to WordNet synsets through the LiLa Knowledge Base (Passarotti et al., 2020).

[8] This also implies that all the predicates of the text are associated with a valency frame in Latin Vallex v1, which is indeed built upon the tectogrammatical annotation of LDT texts (3.2).

the need of an exploitable valency lexicon, with the two versions of Latin Vallex having proved inadequate in their current state. For this reason, we started to build Vallex4UMR, which was compiled as presented hereafter. Its current stage is exclusively built on *De Coniuratione Catilinae*; it may be subject to extension as new texts are to be annotated.

To begin with, we retrieved all predicates of the text by extracting from the tectogrammatical layer all the nodes with a valency frame, and then restricting to verbs thanks to their part-of-speech. We also mapped all frame ids of Vallex v2 to a newly generated set of labels that conform to the UMR requirements (e.g., *verb-01*), which resulted in the set of Vallex4UMR entries inheriting definition and roles from Vallex v2. We then manually paired each extracted predicate, which is already associated to a Vallex v1 valency frame, with one of the newly created Vallex4UMR frames; as already discussed, a 1:1 mapping of v1 and v2 frames has proven to be unrealistic in light of the above-mentioned issues, thus demanding a manual mapping case by case. Consequently, each predicate is associated with a v1 frame and a v2 one, with the latter implying a definition and semantic roles. Moreover, to each Vallex4UMR entry we associate an example as well as possible additional grammatical notes (e.g., the verb voice). Choosing the most appropriate frame from v2 has, however, turned out to be non-trivial in several cases, due to the already discussed issue of identical v2 frames (*porto*, *metior*). We decided to merge such entries into a single one.

Overall, most verbs received a frame from Latin Vallex v2. When no appropriate frame could be found in the resource, a new one was defined. Its definition could consist of an existing WordNet synset which was not assigned to the verb so far, or a brand-new definition could be created. E.g., for *vivo* there is no entry in Latin WordNet; to its occurrence in *alii alio more viventes* 'living with different customs' we assigned a new frame with synset `v#02614387` "lead a certain kind of life; live in a certain style". Finally, we chose not to assign any frame to some predicates, as they can be treated as UMR abstract predicates; it is often the case of the verb *sum* 'to be', which can be treated e.g. as *identity-91*, *belong-91*, *have-mod[ification]-91*. We thus excluded such predicates from Vallex4UMR.

As a result, more than one frame can be paired with the same Vallex4UMR entry, i.e. to the same v2 frame. This holds true the other way round as well, with several v2 frames that can correspond to a single v1 frame.

The current version of the corpus encompasses 2,046 predicates.[9] At present, the resource is not yet organized as a lexicon; currently each predicate extracted from Sallust's texts is annotated with the above-mentioned information. We will need to transition from such predicate-based structure to a new format centered around UMR lexical entries. We will thus need to devise a conversion process to generate the appropriate Vallex4UMR frame dictionary. This conversion process will entail handling a complex mapping of entries from Vallex v1, Vallex v2 and Vallex4uMR.

It emerges clearly from the presented picture how manual annotation represents a time-consuming task. This comes as no surprise, given that semantic annotation typically relies on extensive lexical resources such as frame dictionaries, and that a deeper annotation correlates with greater annotation effort. For this reason, possible strategies to obtain annotation (semi-)automatically will have to be investigated.

3.4 Towards a Parallel Corpus

While our primary focus revolves around Latin, we also intend to broaden our scope to encompass a diverse set of languages, thereby integrating a multilingual dimension into our research. Specifically, we aim to expand our annotation efforts to include some Romance languages – potentially Italian, French, and/or Spanish. As a result, a subset of the UMR corpus derived from this work will incorporate the selected Romance languages as well. Ideally, we aspire to develop a prototype of parallel corpus, leveraging the choice of *De Coniuratione Catilinae*, as Sallust represents one of the most renowned and extensively translated authors from classical antiquity. A parallel corpus would constitute a valuable resource for comparative analysis. Our objective is to conduct comparative assessments between texts that share the same content, which is however expressed in different languages. Such investigations will elicit linguistic observations, for instance with respect to the isomorphism of graphs representing identical sentences across different languages. Noteworthy remarks within a

[9] For a total of more then 1,100 unique Vallex4UMR entries. In other words, one Vallex4UMR entry is one frame of a verbal lemma, and predicates are instances of these frames in the sentences. We will then group Vallex4UMR entries by lemma.

comparative perspective may emerge when graphs exhibit non-isomorphism, indicating structural disparities. Exploring the underlying reasons for such divergences may provide interesting insights about the languages themselves.

Undoubtedly, however, a significant challenge in this endeavor is represented by the availability of frame dictionaries for the Romance languages to be selected for annotation. What has been discussed about Latin (see 3.2, 3.3) has highlighted the inherent limitations arising from linguistic resources that may not be optimally suited for this task, or lack the requisite structure for straightforward, seamless exploitation.

An initial exploration of available resources that could be exploited for annotation demands caution before finalizing the selection of languages. Spanish seems suitable for our research purposes, given the quality and comprehensiveness of SynSemClass[10] (Alcaina et al., 2023). However, the same does not hold true for French.

While we identified two possible resources for French, they present immediate challenges. Dicoinfo[11] shows an adequate structure, offering examples and frames for each entry. Yet, its scope is confined to fundamental terms within the realm of computer science, posing a significant limitation. The second resource, French FrameNet,[12] presents interpretational hurdles. For instance, the frame defined as 'becoming aware' under the entry *observer* 'to observe' is associated with the following roles: *Cognizer*, *Phenomenon*, *Evidence*, *Topic*, *Instrument*. Although the first two seem to roughly correspond to ACT and PAT in FGD valency theory, it is hard to infer a general mapping for all roles.

The initial findings for Italian appear discouraging. Although Brambilla et al. (2020) allude to a forthcoming FrameNet-like resource for Italian, our attempts to locate this resource have proven futile. PaVeDa - Pavia Verbs Database (Zanchi et al., 2022) consists of a relational database for exploring verb argument structures across various languages, including Italian; nevertheless, it appears to still be in progress as far as the specification of roles is concerned. Indeed, this database defines verb meanings along with their participants; for instance, the Italian verb *caricare* 'to load' is associated with a loader, a loaded object, and a loading location; it is also provided with a basic coding frame.[13] However, only few participants, which are otherwise desribed as the example of *caricare* showed, are linked to more readily interpretable labels such as A (Agent) or P (Patient). The development of role labels within this database appears to be a work in progress. Alternatively, T-PAS (Typed Predicate Argument Structures for Italian verbs) (Jezek et al., 2014) offers an inventory of Typed Predicate Argument Structures (T-PASs) for Italian, structured as *[[Human]] partecipa a 'takes part in' [[Event]]*. However, the absence of a clear correlation to roles that can be mapped to PropBank ones, as required by UMR, implies an additional level of effort that may be unfeasible to undertake.

The Universal PropBank project (Jindal et al., 2022) provides PropBank-like argument labels for a number of Romance languages – including Italian, Spanish and French. However, these are corpora and not dictionaries, with annotations being atop UD data and not gold standard. As our project aims to create a parallel resource centered on Sallust's work, such resource does not meet our needs and cannot be leveraged. In light of such considerations, maintaining flexibility in the selection of languages for the multilingual expansion of the corpus is crucial, as the choice of languages heavily relies on the accessibility of suitable resources. Further examination of these resources may unveil structural challenges beyond our current capabilities. Potential alternatives could involve considering Czech and English, both known to have suitable and accessible frame dictionaries. Alternatively, a possible strategy to address the lack of valency lexicons might involve searching for cognates (Dinu et al., 2024) – for instance between Italian and Spanish – and then leveraging existing Spanish frames to annotate the Italian cognate.

4 Future Work

In summary, we anticipate the future research directions, as partially introduced in the previous sections, to revolve around the main themes outlined hereafter.

The first branch of research will focus on the extension and the release of a new resource, namely Vallex4UMR. As discussed in Section 3.2

[10] https://lindat.mff.cuni.cz/services/SynSemClass50/.

[11] https://olst.ling.umontreal.ca/dicoinfo/moteur/search.cgi.

[12] http://asfalda.linguist.univ-paris-diderot.fr/luIndex.xml.

[13] 1 > V.subj[1] > 2 (su+3).

and 3.3, currently the lexicon is still structured as a corpus of annotated predicates. It will be necessary to convert it to a lexicon-like format, structured in such a way that lexical entries are associated with information about the corresponding identifier in Vallex v1, the corresponding URI in Vallex v2, their definition, an example, the associated roles, and possibly additional notes. This process will involve more than just a simple conversion of formats – from a spreadsheet to a new lexicon-like structure – and will present some challenges demanding non-trivial decisions to be taken.[14] Most of the issues to be handled derive from the lack of a one-to-one correspondence between Vallex v1 and v2 entries, which results in the fact that one Vallex4UMR entry can point to several identifiers from v1 and URIs from v2, and vice versa multiple Vallex4UMR entries can point to a same, shared identifier and/or URI in v1 and v2 respectively. Careful handling of these non-univocal mappings will be necessary to ensure that no information is lost. For instance, the Vallex v1 synset `v#v-w76_MPf24_MP` of the verb *facio* 'to do, to make' has been associated to different v2 synsets, and thus Vallex4UMR entries (*facio-18*, *facio-23*, *facio-25*, ...). At the same time, these entries have been mapped to several v1 synsets: for instance, *facio-23* has been aligned to `v#v-w76_MPf47_MP`, `v#v-w76_MPf43_MP`, `v#v-w76_MPf45_MP`, `v#v-w76_MPf30_MP`, etc. The conversion and refining of Vallex4UMR is expected to take about a month; its extension, dependent on the annotation and/or conversion of texts other than *De coniuratione Catilinae*, would presumably require a few months and would proceed in parallel with the annotation itself.

Our second objective involves the expansion of the corpus of Latin sentences annotated in UMR. This endeavor will demand significant time and require comprehensive understanding of the guidelines, which may need language-specific refinement as they are currently under-specified and not exhaustive. The number of annotated sentences per language and included in the current release is shown in Table 1.[15] *Document* refers to sentences annotated with document level information, while *Sentence* refers to sentences annotated by within-sentence relations only.

[14]We have already reviewed possible technical strategies to encode Vallex4UMR, and decided to replicate the same procedure as for PDT-Vallex.

[15]`https://umr4nlp.github.io/web/data.html`.

Language	Sentence	Document
English	209	202
Chinese	358	358
Arapaho	406	109
Navajo	522	168
Sanapaná	602	602
Kukama	105	86

Table 1: Number of sentences for each language in UMR 1.0 dataset.

Given these figures, it is realistic to anticipate that the annotated sample will encompass a few hundred sentences, featuring both sentence-level and document-level annotations. If we can leverage PDT-like data via a conversion process, obtaining annotated sentences will be significantly more efficient, allowing us to expand the sample size further. We expect the annotation phase to require about half a year. As a result, we plan to publicly release the corpus as part of the UMR collection, accessible via the LINDAT/CLARIAH-CZ repository. Latin will represent the first historical language to be annotated according to UMR, and together with the selected Romance language(s) will contribute to the diversity of the languages represented in the UMR collection, since as of now no Romance languages are included.

Indeed, we also plan to expand the corpus of annotated sentences to encompass one or more Romance languages – potentially Spanish, French, or Italian.[16] Depending on the outcomes of a more thorough examination of existing resources, we may also consider exploring some machine learning strategies, such as automatic detection of cognates in related languages, to address the lack of resource. Although at present this does not represent our primary strategy due to the fact that potentially sub-optimal results of automatic processing may affect the quality of the annotation, we are prepared to adopt such perspective as a viable solution to an anticipated issue. Moreover, automatic projection of graphs from already annotated Latin sentences to their counterparts in the other languages will be investigated too, as it could help speeding up the annotation. Measuring the accuracy and the increase of efficiency – especially in terms of time – resulting from projecting UMR annotations from Latin to the selected Romance language(s) through

[16]The final decision hinges on the availability and suitability of dictionary frames for our task, as discussed previously.

transfer of graphs and adjustment of alignments, would prove interesting. The creation of a parallel corpus will enable multilingual comparison, allowing for the examination of structural isomorphism and investigation of reasons for non-isomorphic structures. Additionally, we anticipate similar observations to arise during the annotation process, as different surface realizations of the same meaning in different languages are likely to prompt discussion of the guidelines and shed light on linguistic variations across languages. The multilingual phase is expected to extend over a period of approximately half a year.

Additionally, we are also considering the feasibility of automatizing aspects of the annotation process by leveraging existing resources. A first approach would imply exploiting PDT, which indeed conditioned the choice of *De Coniuratione Catilinae*. One possibility involves partially deriving UMR annotation from PDT tectogrammatical layer; if the conversion tool currently under development is ready timely, we will investigate how to exploit it in the most efficient way. A complementary strategy revolves around the extraction of relevant information from Universal Dependencies treebanks. As a valuable starting point, some sentences from Sallust's text are also annotated in UD, enabling us to explore potential points of contact and intersections between the two frameworks.

As far as semantic parsers for UMR are concerned, the amount of data currently available does not allow yet to build and train a parser on UMR data only. For this reason, right now UMR parsing has to be pipelined (Chun and Xue, 2024). The pipeline approach certainly implies a high rate of error propagation, as the model described heavily relies on sub-models trained on different data. As of now, however, this represents the only viable approach. The availability of more data could probably allow for an end-to-end approach, especially for high-resourced languages like English and Chinese. Nevertheless, for less-resourced languages – which, to begin with, are lacking lexical resources for annotation and demand extra challenges to be addressed – a pipelined approach will possibly remain a realistic solution for a longer time. Indeed, the pipeline proposed by Chun and Xue (2024) would not apply to a less-resourced language like Latin, requiring alternative strategies. In light of this, another possible outcome that we envisage would involve integrating available tools into a unified pipeline to streamline manual annotation. For instance, given that coreference forms part of UMR document-level annotation, a tool for coreference prediction could prove beneficial and serve as a valuable component of the pipeline. In general, we will look into the possibility of speeding up the annotation process by investigating possible strategies to replace manual annotation steps through existing resources and machine learning strategies. As of now, we have investigated the feasibility of identifying the correct word sense to define semantic concepts in the UMR graph (e.g., *facio-23*, *porto-01*, ...) by leveraging different Pre-trained Language Models (PLM) (Gamba, 2024). However, such task of Predicate Sense Disambiguation highlighted the challenges that arise from the lexical resources available for Latin (3.2). Indeed, PLMs often fail to retrieve the correct sense, with the high granularity of the resources being a significant contributing factor. We envisage the automatization of (steps of) the annotation process to be carried out in parallel with manual annotation, with an estimated timeline of six months to one year dedicated to it.

To summarize, through the study of Latin the proposed research will examine how the UMR framework can be applied to historical languages, investigating how the lack of native speakers interacts with deep semantic analysis and annotation. Indeed, this is still an unexplored research area, and the challenges that meaning representation of a historical language can pose have not been addressed yet. The study will then adopt a diachronic perspective, examining the syntax-semantics interface in Latin as opposed to Romance languages (most likely Italian, Spanish, or French). This will also result in the expansion of the UMR coverage of languages, by enhancing the already released corpus with data from an Indo-European family – so far represented only by English – and precisely from the Indo-European branch that will develop into Romance languages, absent in UMR as of now. The analysis of UMR parallel sentences will provide insights in terms of isomorphism of graphs as well as highlight the main divergences in how these languages encode meaning. Additionally, the project will address the need to expedite the annotation process by utilizing existing resources and machine learning techniques, thereby highlighting the potential and challenges of these tools in semantic annotation. Leveraging both UD and UMR annotations simultaneously will also allow for the observation of how closely syntactic and semantic trees align. Three new resources (Vallex4UMR,

UMR Latin data, UMR parallel corpus) will be contributed, as described in the previous paragraphs.

References

Furkan Akkurt, Bermet Chontaeva, Çağrı Çöltekin, Mehmet Oguz Derin, Gulnura Dzhumalieva, Soudabeh Eslami, Soudabeh Eslami, Tunga Güngör, Sardana Ivanova, Jumashev Murat, Aida Kasieva, Aslı Kuzgun, Büşra Marşan, Balkız Öztürk, Chihiro Taguchi, Susan Üsküdarlı, Jonathan Washington, and Olcay Taner Yıldız. 2024. Unifying the annotations in Turkic Universal Dependencies treebanks.

Cristina Fernández Alcaina, Eva Fučíková, Jan Hajič, and Zdeňka Urešová. 2023. Spanish synonyms as part of a multilingual event-type ontology. *Jazykovedný časopis / Journal of Linguistics*, 74(1):153–162.

David Bamman and Gregory Crane. 2011. The Ancient Greek and Latin Dependency Treebanks. In *Language technology for cultural heritage*, pages 79–98. Springer.

Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract Meaning Representation for Sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 178–186, Sofia, Bulgaria. Association for Computational Linguistics.

Rufus Behr. 2024. Behr at EvaLatin 2024: Latin dependency parsing using historical sentence embeddings. In *Proceedings of the Third Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA) @ LREC-COLING-2024*, pages 198–202, Torino, Italia. ELRA and ICCL.

Julia Bonn, Matthew J. Buchholz, Jayeol Chun, Andrew Cowell, William Croft, Lukas Denk, Sijia Ge, Jan Hajič, Kenneth Lai, James H. Martin, Skatje Myers, Alexis Palmer, Martha Palmer, Claire Benet Post, James Pustejovsky, Kristine Stenzel, Haibo Sun, Zdeňka Urešová, Rosa Vallejos, Jens E. L. Van Gysel, Meagan Vigus, Nianwen Xue, and Jin Zhao. 2024. Building a broad infrastructure for uniform meaning representations. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 2537–2547, Torino, Italia. ELRA and ICCL.

Julia Bonn, Chen Ching-wen, James Andrew Cowell, William Croft, Lukas Denk, Jan Hajič, Kenneth Lai, Martha Palmer, Alexis Palmer, James Pustejovsky, Haibo Sun, Rosa Vallejos Yopán, Jens Van Gysel, Meagan Vigus, Nianwen Xue, and Jin Zhao. 2023a. Uniform meaning representation. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.

Julia Bonn, Skatje Myers, Jens E. L. Van Gysel, Lukas Denk, Meagan Vigus, Jin Zhao, Andrew Cowell, William Croft, Jan Hajič, James H. Martin, Alexis Palmer, Martha Palmer, James Pustejovsky, Zdenka Urešová, Rosa Vallejos, and Nianwen Xue. 2023b. Mapping AMR to UMR: Resources for adapting existing corpora for cross-lingual compatibility. In *Proceedings of the 21st International Workshop on Treebanks and Linguistic Theories (TLT, GURT/SyntaxFest 2023)*, pages 74–95, Washington, D.C. Association for Computational Linguistics.

Silvia Brambilla, Danilo Croce, Fabio Tamburini, Roberto Basili, et al. 2020. Automatic induction of framenet lexical units in italian. In *CEUR WORKSHOP PROCEEDINGS*, volume 2769. CEUR-WS.

Flavio M Cecchini, Rachele Sprugnoli, Giovanni Moretti, and Marco Passarotti. 2020a. UDante: First Steps Towards the Universal Dependencies Treebank of Dante's Latin Works. In *Seventh Italian Conference on Computational Linguistics*, pages 1–7. CEUR Workshop Proceedings.

Flavio Massimiliano Cecchini, Timo Korkiakangas, and Marco Passarotti. 2020b. A New Latin Treebank for Universal Dependencies: Charters between Ancient Latin and Romance Languages. In *Proceedings of the 12th Language Resources and Evaluation Conference*, Marseille, France. European Language Resources Association.

Jayeol Chun and Nianwen Xue. 2024. Baseline Uniform Meaning Representation Parsing as a Pipelined Approach. In *UMR Parsing Workshop*.

Marie-Catherine de Marneffe, Christopher Manning, Joakim Nivre, and Daniel Zeman. 2021. Universal Dependencies. *Computational Linguistics*, 47(2):255–308.

Liviu P. Dinu, Ana Sabina Uban, Ioan-Bogdan Iordache, Alina Maria Cristea, Simona Georgescu, and Laurentiu Zoicas. 2024. Pater incertus? there is a solution: Automatic discrimination between cognates and borrowings for Romance languages. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 12657–12667, Torino, Italia. ELRA and ICCL.

Greta Franzini, Andrea Peverelli, Paolo Ruffolo, Marco Passarotti, Helena Sanna, Edoardo Signoroni, Viviana Ventura, Federica Zampedri, et al. 2019. Nunc Est Aestimandum: Towards an Evaluation of the latin WordNet. In *CLiC-it*.

Federica Gamba. 2024. Predicate Sense Disambiguation for UMR annotation of Latin: Challenges and insights. In *Proceedings of the First Workshop on Machine Learning for Ancient Languages (ML4AL 2024)*, Bangkok, Thailand. Association for Computational Linguistics.

Federica Gamba and Daniel Zeman. 2023a. Latin morphology through the centuries: Ensuring consistency

for better language processing. In *Proceedings of the Ancient Language Processing Workshop*, pages 59–67, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Federica Gamba and Daniel Zeman. 2023b. Universalising Latin Universal Dependencies: a harmonisation of Latin treebanks in UD. In *Proceedings of the Sixth Workshop on Universal Dependencies (UDW, GURT/SyntaxFest 2023)*, pages 7–16, Washington, D.C. Association for Computational Linguistics.

Dag TT Haug and Marius Jøhndal. 2008. Creating a Parallel Treebank of the Old Indo-European Bible Translations. In *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008)*, pages 27–34.

Elisabetta Jezek, Bernardo Magnini, Anna Feltracco, Alessia Bianchini, and Octavian Popescu. 2014. T-PAS; a resource of typed predicate argument structures for linguistic analysis and semantic processing. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 890–895, Reykjavik, Iceland. European Language Resources Association (ELRA).

Ishan Jindal, Alexandre Rademaker, Michał Ulewicz, Ha Linh, Huyen Nguyen, Khoi-Nguyen Tran, Huaiyu Zhu, and Yunyao Li. 2022. Universal Proposition Bank 2.0. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 1700–1711, Marseille, France. European Language Resources Association.

Paul Kingsbury and Martha Palmer. 2002. From TreeBank to PropBank. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02)*, Las Palmas, Canary Islands - Spain. European Language Resources Association (ELRA).

Bin Li, Yuan Wen, Weiguang Qu, Lijun Bu, and Nianwen Xue. 2016. Annotating the little prince with Chinese AMRs. In *Proceedings of the 10th Linguistic Annotation Workshop held in conjunction with ACL 2016 (LAW-X 2016)*, pages 7–15, Berlin, Germany. Association for Computational Linguistics.

Francesco Mambrini, Marco Passarotti, Eleonora Litta, and Giovanni Moretti. 2021a. Interlinking valency frames and wordnet synsets in the lila knowledge base of linguistic resources for latin. In *Further with Knowledge Graphs*, pages 16–28. IOS Press.

Francesco Mambrini, Marco Passarotti, Eleonora Litta, and Giovanni Moretti. 2021b. Interlinking valency frames and wordnet synsets in the LiLa knowledge base of linguistic resources for Latin. In *Further with Knowledge Graphs*, pages 16–28. IOS Press.

Marco Passarotti. 2019. The Project of the Index Thomisticus Treebank. *Digital Classical Philology*, 10:299–320.

Marco Passarotti and Felice Dell'Orletta. 2010. Improvements in parsing the index Thomisticus treebank. revision, combination and a feature model for medieval Latin. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Marco Passarotti, Francesco Mambrini, Greta Franzini, Flavio Massimiliano Cecchini, Eleonora Litta, Giovanni Moretti, Paolo Ruffolo, and Rachele Sprugnoli. 2020. Interlinking through lemmas. The lexical collection of the LiLa knowledge base of linguistic resources for Latin. *Studi e Saggi Linguistici*, 58(1):177–212.

Marco Passarotti and Paolo Ruffolo. 2010. Parsing the Index Thomisticus Treebank. Some Preliminary Results. In *15th International Colloquium on Latin Linguistics*, pages 714—725. Innsbrucker Beiträge zur Sprachwissenschaft.

Marco Passarotti, Berta González Saavedra, and Christophe Onambele. 2016. Latin Vallex. a treebank-based semantic valency lexicon for Latin. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2599–2606, Portorož, Slovenia. European Language Resources Association (ELRA).

Edoardo Maria Ponti and Marco Passarotti. 2016. Differentia compositionem facit. a slower-paced and reliable parser for Latin. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 683–688, Portorož, Slovenia. European Language Resources Association (ELRA).

Martin Popel, Zdeněk Žabokrtský, and Martin Vojtek. 2017. Udapi: Universal API for Universal Dependencies. In *Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017)*, pages 96–101, Gothenburg, Sweden. Association for Computational Linguistics.

James Pustejovsky, Ken Lai, and Nianwen Xue. 2019. Modeling quantification and scope in Abstract Meaning Representations. In *Proceedings of the First International Workshop on Designing Meaning Representations*, pages 28–33, Florence, Italy. Association for Computational Linguistics.

Petr Sgall. 1967. Functional sentence perspective in a generative description. *Prague studies in mathematical linguistics*, 2(1967):203–225.

Rachele Sprugnoli, Federica Iurescia, and Marco Passarotti. 2024. Overview of the EvaLatin 2024 evaluation campaign. In *Proceedings of the Third Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA) @ LREC-COLING-2024*, pages 190–197, Torino, Italia. ELRA and ICCL.

Rachele Sprugnoli, Marco Passarotti, Flavio Massimiliano Cecchini, Margherita Fantoli, and Giovanni

Moretti. 2022. Overview of the EvaLatin 2022 evaluation campaign. In *Proceedings of the Second Workshop on Language Technologies for Historical and Ancient Languages*, pages 183–188, Marseille, France. European Language Resources Association.

Rachele Sprugnoli, Marco Passarotti, Flavio Massimiliano Cecchini, and Matteo Pellegrini. 2020. Overview of the EvaLatin 2020 evaluation campaign. In *Proceedings of LT4HALA 2020 - 1st Workshop on Language Technologies for Historical and Ancient Languages*, pages 105–110, Marseille, France. European Language Resources Association (ELRA).

Milan Straka, Jana Straková, and Federica Gamba. 2024. ÚFAL LatinPipe at EvaLatin 2024: Morphosyntactic analysis of Latin. In *Proceedings of the Third Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA) @ LREC-COLING-2024*, pages 207–214, Torino, Italia. ELRA and ICCL.

Haibo Sun, Yifan Zhu, Jin Zhao, and Nianwen Xue. 2023. UMR annotation of Chinese verb compounds and related constructions. In *Proceedings of the First International Workshop on Construction Grammars and NLP (CxGs+NLP, GURT/SyntaxFest 2023)*, pages 75–84, Washington, D.C. Association for Computational Linguistics.

Zdeňka Urešová, Jan Hajič, and Ondřej Bojar. 2014. Comparing Czech and English AMRs. In *Proceedings of Workshop on Lexical and Grammatical Resources for Language Processing*, pages 55–64, Dublin, Ireland. Association for Computational Linguistics and Dublin City University.

Jens E. L. Van Gysel, Meagan Vigus, Lukas Denk, Andrew Cowell, Rosa Vallejos, Tim O'Gorman, and William Croft. 2021a. Theoretical and practical issues in the semantic annotation of four indigenous languages. In *Proceedings of the Joint 15th Linguistic Annotation Workshop (LAW) and 3rd Designing Meaning Representations (DMR) Workshop*, pages 12–22, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Jens EL Van Gysel, Meagan Vigus, Jayeol Chun, Kenneth Lai, Sarah Moeller, Jiarui Yao, Tim O'Gorman, Andrew Cowell, William Croft, Chu-Ren Huang, et al. 2021b. Designing a uniform meaning representation for natural language processing. *KI-Künstliche Intelligenz*, 35(3):343–360.

Meagan Vigus, Jens E. L. Van Gysel, and William Croft. 2019. A dependency structure annotation for modality. In *Proceedings of the First International Workshop on Designing Meaning Representations*, pages 182–198, Florence, Italy. Association for Computational Linguistics.

Meagan Vigus, Jens E. L. Van Gysel, Tim O'Gorman, Andrew Cowell, Rosa Vallejos, and William Croft. 2020. Cross-lingual annotation: a road map for low- and no-resource languages. In *Proceedings of the Second International Workshop on Designing Meaning Representations*, pages 30–40, Barcelona Spain (online). Association for Computational Linguistics.

Nianwen Xue, Ondřej Bojar, Jan Hajič, Martha Palmer, Zdeňka Urešová, and Xiuhong Zhang. 2014. Not an interlingua, but close: Comparison of English AMRs to Chinese and Czech. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 1765–1772, Reykjavik, Iceland. European Language Resources Association (ELRA).

Chiara Zanchi, Silvia Luraghi, and Claudia Roberta Combei. 2022. PaVeDa - Pavia verbs database: Challenges and perspectives. In *Proceedings of the 4th Workshop on Research in Computational Linguistic Typology and Multilingual NLP*, pages 99–102, Seattle, Washington. Association for Computational Linguistics.

Amir Zeldes and Nathan Schneider. 2023. Are UD treebanks getting more consistent? a report card for English UD. In *Proceedings of the Sixth Workshop on Universal Dependencies (UDW, GURT/SyntaxFest 2023)*, pages 58–64, Washington, D.C. Association for Computational Linguistics.

A Appendix

```
# sent_id = SlaT-2775
# :: snt3
Index:             1       2      3     4        5      6      7      8
Words:           animi   imperio   ,  corporis servitio magis  utimur   ;
Word Gloss (en):  of.mind government , of.body  service  rather we.employ ;
Sentence:        animi imperio, corporis servitio magis utimur;
Sentence Gloss (en): Of the mind we rather employ the government; of the body, the service.

# sentence level graph:
(s3c / contrast-91
      :ARG1 (s3u / utor-03
                  :ARG0 (s3p/ person
                              :refer-person 1st
                              :refer-number plural)
                  :ARG1 (s3i / imperium
                               :poss (s3a / animus))
                  :aspect habitual
                  :modal-strength full-affirmative)
       :ARG2 (s3u2 / utor-03
                   :ARG0 s3p
                   :ARG1 (s3s / servitium
                                :poss (s3c2 / corpus))
                   :aspect habitual
                   :modal-strength full-affirmative))

# alignment:
s3c: 6-6
s3u: 7-7
s3p: 0-0
s3i: 2-2
s3a: 1-1
s3u2: 6-6
s3s: 5-5
s3c2: 4-4

# document level annotation:
(s3s0 / sentence
   :temporal ((document-creation-time :overlap s3u)
           (document-creation-time :overlap s3u2)
            (s3u :overlap s3u2))
    :modal ((root :modal author)
            (author :full-affirmative s3u)
            (author :full-affirmative s3u2))
    :coref ((s2a :same-entity s3a)
            (s2c :same-entity s3c2)
            (s2p :same-entity s3p)))
```