Building CorefLat
A linguistic resource for coreference and anaphora resolution in Latin

Eleonora Delfino[1,*,†], Roberta G. Leotta[2,†], Marco Passarotti[2,†] and Giovanni Moretti[2,†]

[1]Università di Udine, Via Palladio 8, 33100 Udine, Italy
[2]CIRCSE Research Centre, Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123 Milano, Italy

Abstract

This paper presents the initial stages of a project focused on coreference and anaphora resolution in Latin texts. By building a corpus enhanced with coreference/anaphora annotation, the project wants to explore empirically a layer of metalinguistic analysis that has not been yet extensively investigated in linguistic resources and natural language processing for Latin. After reviewing the related work on this NLP task, the paper discusses annotation criteria and data analysis, providing examples about a few issues that emerged during the annotation process.

Keywords

Latin, Coreference, Anaphora, Annotation, Corpora

1. Introduction

Over the past decade, research on linguistic resources and natural language processing (NLP) for Latin has seen remarkable growth[1]. However an important layer of metalinguistic annotation such as coreference and anaphora resolution still remains quite neglected. Indeed, except for the (meta)data produced by the FIR-2013 project *Development and Integration of Advanced Linguistic Resources for Latin* [2], there are neither corpora enhanced with coreferential/anaphoric annotations nor NLP tools for automatic coreference/anaphora resolution for Latin. This absence limits the degree of granularity of information extraction from Latin corpora. Such a limitation is particularly compelling, as Latin texts are mainly used for purposes of research in the Humanities, like literary, stylistic and philosophical analysis. To give an example, investigating in Latin texts a philosophical concept conveyed by a word, like *voluntas* 'will', or studying the turns of a certain character in a drama would highly benefit from a textual resource where, for instance, the ana-/cataphoric references of pronouns are resolved.

The PRIN 2022 project *Textual Data and Tools for Coreference Resolution of Latin* was granted funding to overcome such situation. Run jointly by the Università Cattolica of Milan and the University of Udine, the project stems from the FIR-2013 pilot experience, having the short-term objective of developing a large-scale and balanced dataset of Latin texts enhanced with coreference/anaphora resolution (called CorefLat). Based upon this annotated dataset, the project has two long-term objectives.

The first aims to develop and evaluate a set of trained models for automatic coreference/anaphora resolution of Latin.

The second long-term objective wants to publish the metadata pertaining to coreference/anaphora resolution as Linked Data, to make them interoperable with other (meta)data in the Web. To this aim, the texts of the annotated dataset are selected among those published in the LiLa Knowledge Base, a collection of multiple linguistic resources for Latin modelled using the same vocabularies for knowledge description and interconnected according to the principles of the Linked Data paradigm [3][2].

This paper details the initial stages of the creation of the CorefLat annotated dataset.

CLiC-it 2024: Tenth Italian Conference on Computational Linguistics, Dec 04 — 06, 2024, Pisa, Italy
*Corresponding author.
†These authors contributed equally.
eleonora.delfino@uniud.it (E. Delfino); robertagrazia.leotta@unicatt.it (R. G. Leotta); marco.passarotti@unicatt.it (M. Passarotti); giovanni.moretti@unicatt.it (G. Moretti)
https://docenti.unicatt.it/ppd2/en/docenti/102059/roberta-grazia-leotta/profilo (R. G. Leotta); https://docenti.unicatt.it/ppd2/it/docenti/14144/marco-carlo-passarotti/profilo (M. Passarotti)
0009-0002-5947-5011 (E. Delfino); 0009-0004-5631-1032 (R. G. Leotta); 0000-0002-9806-7187 (M. Passarotti); 0000-0001-7188-8172 (G. Moretti)

[1]For an overview of the available linguistic resources for Latin, see [1]. As for NLP tools, see the three editions of the evaluation campaign EvaLatin (last edition: https://circse.github.io/LT4HALA/2024/EvaLatin).

[2]https://lila-erc.eu

2. Related Work

Coreference (henceforth CR) and anaphora (henceforth AR) resolution are often treated as a single, yet diverse, task in NLP. To understand the difference between CR and AR, it is necessary to distinguish between the concept of "mention" and that of "entity". A mention is defined as an instance of reference to an object, while an entity is the object to which a mention refers in a text. CR consists in finding in a text all mentions of (strictly speaking, real-world) entities such as persons or organisations, regardless of their textual representation. Instead, in AR the interpretation of a mention (known as "anaphora" or "cataphora", e.g., a pronoun) depends on another mention present in the text, whether antecedent or following in the word order. If both mentions refer to the same entity, they are considered to be coreferential, which makes AR and CR closely bound to each other. Since ana-/cataforic relations are present in the text, the need of world knowledge in AR is minimal. In contrast, CR has a much broader scope: co-referential terms can have completely different grammatical properties and/or functions (e.g., different gender and part of speech) and yet, by definition, they can refer to the same entity.

In NLP, the CR task is usually not meant in a strict sense, as it consists in finding all mentions of each entity in a text regardless of their relation to the real world. Accordingly, our project adopts this same interpretation of the CR task [4].

Since the 1960s, coreference and anaphora resolution has been a central topic in NLP studies, but it was considered a difficult task, typically requiring the use of sophisticated knowledge sources and inference procedures. In 1983, Roberto Busa pointed out the absence of resources and tools for pronoun coreference resolution: "[...] avete mai incontrato tavole e concordanze computerizzate nelle quali il programma automaticamente abbia [...] collegato i pronomi alle forme di cui sono vicari?" [5, 7.2][3].

Like for other NLP tasks, during the 1990s research on CR/AR gradually shifted from heuristic approaches to machine learning approaches, thanks to the public availability of annotated corpora produced for the aims of shared tasks dedicated to coreference resolution, such as Message Understanding Conference (MUC) conferences [7], and Automatic Content Evaluation (ACE) Program conferences [8]. These corpora mainly include news article and newswire texts in English. The ACE corpus also features Arabic and Chinese texts from web-blogs and telephone conversations. The tendency to focus coreference and anaphora annotation on newspaper texts is also confirmed by those selected for the CoNLL shared task on modeling unrestricted coreference in OntoNotes [9, 10], as well as by the NXT-format Switchboard Corpus [11]. In addition, some treebanks feature CR/AR, encompassing a wide range of languages, including English and Czech [12], German [13], Japanese [14], Italian [15], Spanish and Catalan [16]. To the best of our knowledge, there is no specific Latin corpus enriched with CR/AR. The only currently available texts that include this layer of annotation come from Latin treebanks. The FIR-2013 project mentioned above built a CR-annotated dataset including works by Sallust, Caesar and Cicero (taken from the Latin Dependency Treebank [17]), and by Thomas Aquinas (from the *Index Thomisticus* Treebank [18]). However, the selection of texts in this dataset is quite unbalanced as for both literary genres and authors. Out of the more than 45,000 total annotated tokens, about 27,000 are taken from Thomas Aquinas' *Summa contra Gentiles*, and more than 10,000 are from Sallust's *In Catilinam*. This given, our project wants to create a more balanced dataset by increasing and differentiating the quantity of annotated texts for both Classical and Late Latin.

3. Building CorefLat

3.1. Annotation Criteria and Data Selection

To create a resource that adheres to the most unified and widely shared annotation criteria for CR/AR, the annotation style of CorefLat resembles the one developed for the GUM corpus and follows the recommendations proposed by the (ongoing) Universal Anaphora (UA) project[4], which aims to create, gather, and distribute harmonized resources for CR/AR.

While building CorefLat, we decided to focus on a subset of the different types of coreference and ana-/cataphora prescribed by the GUM and UA recommendations. The types that we selected are listed below:

- anaphoric pronouns referring back to something: ***domine qui*** *et semper vivis* (Aug. *Conf.* 1.6.8) '**Lord** (you) **who** live for ever';
- cataphoric pronouns referring forward to something: *invocat* ***te***, ***domine*** (Aug. *Conf.* 1.1.1) 'invokes **you**, **Lord**';
- content-rich lexical item - coreferring the same lexical mention: ***laudes*** *tuae, domine,* ***laudes*** *tuae per scripturas tuas suspenderent palmitem cordis mei* (Aug. *Conf.* 1.17.27) 'Your **praises**, Lord, your **praises** throughout your Scriptures would have supported the vine shoot of my heart';

[3] "[...] have you ever come across computerized tables and concordances in which the programme automatically [...] connects pronouns with the nouns that they represent?". Translation taken from [6, 137-138].

[4] https://universalanaphora.github.io/UniversalAnaphora/

- split antecedents - the referred items are more than one: *an vero **caelum** et **terra**, **quae** fecisti et in **quibus** me fecisti, capiunt te?* (Aug. *Conf.* 1.2.2) '**heaven** and **earth**, **which** you made, and in **which** you made me, encompass you?'.

Such a limited set of types of coreference was selected to address the fundamental aim of the two-year long funded project, namely building and distributing a Latin corpus enhanced with coreferential annotation, which is not yet available for this language.
Texts are annotated manually by two independent annotators, using the Content Annotation Tool (CAT)[19], formerly known as the CELCT Annotation Tool, which was created specifically for textual coreference annotation. The tool is highly customizable, making it possible, for instance, to distinguish between annotations of mentions and those of entities. (Meta)data are saved in XML and are then converted in CoNLL-U Plus following the recommendations of the UA initiative[5].
In CorefLat, coreferences are not annotated as chains, but rather as relations. In a coreference relation two elements are involved: the one referring (mention) and the one referred (entity). In our annotation, each mention points directly to the one entity it refers to, rather than to any previous mention of the same entity. Consider the example in (1).

(1) *Magnus es, **Domine**, et laudabilis valde. Magna virtus **tua** et sapientiae **tuae** non est numerus.* (Aug. *Conf.* 1.1.1)
'Great are you, O **Lord**, and surpassingly worthy of praise. Great is **your** goodness, and **your** wisdom is incalculable'[6].

In sentence (1), we identify two coreference relations: the first one involves the mention *tua* and the entity *Domine*, and the second one involves the mention *tuae* and the same entity *Domine*. Typically, the referred element is a noun, nevertheless it happens to get through cases where the referred entity is represented by a function word, such a pronoun, like in example (2):

(2) *nec valerem **quae** volebam **omnia** nec **quibus** volebam **omnibus**.* (Aug. *Conf.* 1.8.13)
'I was incapable of achieving **all that** I wanted, and by **all that** I wanted.'

In (2), the relative pronoun *quae* refers to the quantifying pronoun *omnia*, like *quibus* refers to *omnibus* in the reminder of the sentence. Since *omnis* 'all' (lemma of both *omnia* and *omnibus*) is a function word, no content-rich entity is concerned in this coreference relation. Moreover, it should be noted that sometimes the entity is not explicitly expressed in the text. To address this issue, we create external entities to which the respective mentions are linked by tagging. For instance, in example (3), the pronoun *nos* 'we' refers to the two lovers in Plautus' comedy *Curculio*, namely the girl *Planesium* and the boy *Phaedromus*, whose names are not explicitly mentioned in the sentence for economy's sake, as the two characters are present on stage and pronounce these lines themselves.

(3) *quo usque, quaeso, ad hunc modum / inter **nos** amore utemur semper surrupticio?* (Pl. *Curc.* 1, 204-205)
'How much longer, please, will we always conduct **our** love affair in secret?'

In such a case, we tag the mention *nos* as linked to the entities "Planesium" and "Phaedromus" that are created external to the text.

The annotation task is performed on a collection of Latin texts already enriched with lemmatization and Part-of-Speech (PoS) tagging and linked to the LiLa Knowledge Base. The following texts were chosen according to selection criteria aimed to ensure a sufficiently representative and balanced corpus as for both literary genre and era.

- Classical Latin: data are excerpted from the *Opera Latina* corpus by LASLA[7], an extensive collection of approximately 1.7 million words from over 130 lemmatized and morphologically tagged Classical and Late Latin texts[8].
- Late Latin: data are taken from the text of Augustine's *Confessiones* provided by The Latin Library[9].

At present, no annotation of Medieval Latin texts was performed, as data from this era are largely provided, albeit in unbalanced fashion, by the results of the FIR project.

3.2. Results

So far, we annotated the following excerpts: the first book from Augustine's *Confessiones*, a philosophical prose text, and a comedy of Plautus: *Curculio*. The workload was split equally between the two annotators; however, the last 50 sentences of the first book of Augustine's *Confessiones* were annotated by both annotators to measure

[5] https://github.com/UniversalAnaphora/UniversalAnaphora/blob/main/documents/UA_CONLL_U_Plus_proposal_v1.0.md
[6] English translations of Latin examples are taken, with minor changes, from [20] (Augustine) and [21] (Plautus).
[7] https://lasladb.uliege.be/OperaLatina/
[8] The *Opera Latina* corpus in the LiLa Knowledge Base is available at https://lila-erc.eu/data/corpora/Lasla/id/corpus.
[9] http://www.thelatinlibrary.com. The text is available in LiLa at https://lila-erc.eu/lodview/data/corpora/CIRCSELatinLibrary/id/corpus/Confessiones

their agreement. Inter-annotator agreement was calculated through the Dice coefficient similarity metric, which is widely adopted in NLP [22, 23]. Its value ranges from 0 to 1, with 1 indicating that two sets are identical and 0 meaning that they have no overlap. Once evaluated that the annotated markables span the same tokens for the two annotators in all cases, we calculated the similarity values as for entities (0.817) and mentions (0.824), which are comparatively highly acceptable for this task [24, 25, 26]. Additionally, the Cohen's Kappa coefficient was measured, yielding the following agreement values for each markable class: for the markable class 'mention' the resulting value is 0.8139902, whereas for the markable class 'entity', the value obtained is 0.8118851.
Table 1 presents the data derived from the analysis of the two texts. To highlight the quantitative significance of the coreference phenomenon, it shows the total number of tokens in the texts analyzed, along with the number of tokens involved in coreference relations. Additionally, the table shows the total number of coreference relations, and their respective entities and mentions. The

Table 1
Data obtained from the analysis of the corpus

Category	*Confessiones*	*Curculio*
Tot. token	6,133	5,853
Token in coref.	746	976
Coref. relation	521	796
Entity	202	577
Mention	542	569

tokens involved in a coreference relation account for the 12.16 percent of the total in *Confessiones*, while in *Curculio* they represent the 16.7 percent of the total. In both cases the percentages exceed the data produced by the FIR project, where the phenomenon concerns approximately the 8 percent of the tokens of the Latin texts annotated therein. The table clearly indicates that *Curculio* exhibits a greater number of coreferences despite having a lower total number of tokens. This difference is statistically significant: the chi-squared test performed on these data yielded a chi-squared statistic of 49.18 and a p-value lower than 0.00001. Given that the p-value is lower than the conventional alpha level of 0.05, coreference relations vary significantly from a statistic point of view in *Confessiones* and in *Curculio*. The coreference phenomenon is indeed widespread in the language of Plautus's theatre. This may be due to the fact that Plautus's language mimics, to some extent, everyday spoken language. Furthermore, the presence of numerous dialogues, where speakers often interrupt each other's turns, implies frequent references to the recipients with whom the characters interact. The text structure, characterized by numerous allocutions, also contributes to the high number of coreferences.

3.3. Annotation Issues

In this section, we present and discuss three examples of annotation issues. On one hand, we address a problematic case regarding the application of our annotation scheme on the data, which was the primary reason for disagreement between the two annotators (example 4). On the other hand, we present two cases that highlight the fundamental role of context (example 5) and of the literary genre (example 6) for the coreference resolution task. The limited number of cases presented below is consistent with our prior decision to restrict the scope of annotation to only a subset of coreferential phenomena. We hypothesize that expanding the range of annotated coreference types or enlarging the corpus of annotated texts (in terms of quantity and literary genre) would lead to greater annotation challenges.
Starting from the first annotation issue, the most relevant disagreement between the two annotators concerns how to link mentions that are distant in the text from the entity they refer to. Example (4) shows a representative case of this type of disagreement.

(4) *Bonus ergo est* ***qui*** *fecit me, et* ***ipse*** *est bonum meum, et* ***illi*** *exulto bonis omnibus quibus etiam puer eram. Hoc enim peccabam, quod non in* ***ipso*** *sed in creaturis* ***eius*** *me atque ceteris voluptates, sublimitates, veritates quaerebam, atque ita inruebam in dolores, confusiones, errores.* (Aug. *Conf.* 1.20.31)
'Therefore the one who made me is good, and he himself is my good, and I rejoice in him for all the good things of which I consisted even in childhood. This was my sin: I sought pleasures, exaltations, truths not in he himself but in his creations, which is to say, in myself and other things'.

The pronouns in (4) are references to the entity God, which is explicitly expressed six sentences above in the text. The reader has no difficulty decoding these pronouns because the first-person narrator is discussing his relationship with God, to whom he is constantly referring. Therefore, it is not necessary to explicitly state the entity in every sentence.

The sentence in (4) can be annotated in two distinct ways: each pronoun can either be directly linked to the entity 'God' within the text, or be linked to the first pronoun concerned in (4) (*qui*), which gets then linked to the external entity 'God'. During the annotation process, the two annotators diverged: one selected the former method, while the other opted for the latter. There is no upper limit to the number of sentences after which a mention cannot be associated with the entity to which it refers [27]. When CR and AR first emerged as NLP tasks, there were concerns that machines could not yield acceptable results if the mention and the entity were too distant

from each other [28]. However, contemporary methods achieve satisfactory results even with long-distance coreference, exceeding 200 sentences [29]. Additionally, given that we focus on literary texts, which feature long-distance coreferences more frequently than other textual types [30], it is imperative that we devote particular attention to this specific type of coreference. The two options chosen by the annotators are both equally valid. To harmonize the annotation process, we decided to link the mention to the external entity beyond a certain threshold, which was set at five sentences[10].

Sentence (5) from Plautus' *Curculio* exemplifies another challenging case of ambiguity, which further complicates the annotation process:

(5) Pal.: *Quid? tu te pones Veneri ieientaculo?* Phaed.: *Me, te atque* ***hosce omnis****.* (Pl. Curc. 1, 73-74)
Pal.: 'What? You'll offer yourself a breakfast to Venus?' Phaed.: 'Yes, myself, yourself, and all these here.'

As is typical in theatrical texts, much is left to the audience' inference. In this instance, the actor's gestures serve to disambiguate the phrase *hosce omnis*, which could refer either to the group of slaves accompanying the character Phaedromus or to the audience itself [31, 32, 33]. The annotators decided to follow the interpretation provided by Paratore [34], according to whom, *hosce omnis* refers to the audience. In this example, an agreement in gender and number between the mentions and the potential antecedents inferred from the context can be observed. Disambiguating the antecedent not only requires understanding the text but also knowing the specific characteristics of the literary genre concerned.

Another case in which the importance of literary genre and knowledge of context becomes evident is as follows.

(6) Cvrc.: [...] *Lyconem quaero tarpezitam.* Lyc.: *Dic mihi, quid eum nunc quaeris?* (Pl. Curc. 3, 406- 407): Cvurc.: 'I'm looking for the banker Lyco.' Lyc.: 'Tell me, why are you looking for him now?'

The dialogue cited here between the two characters, Curculio and Lyco, plays on a comedic ambiguity: Curculio knows he is speaking to Lyco, while Lyco believes that Curculio is unaware of his identity. When Curculio asks to speak with Lyco, Lyco responds by speaking about himself in the third person, thereby concealing his identity. For this reason, both the first-person pronoun 'mihi' and the third-person pronoun 'eum' refer to the same entity. This case clearly demonstrates the importance of understanding both the context and the specific narrative techniques of the textual genre in order to effectively resolve coreferences.

[10]The threshold is sentence-based rather than token-based as sentence is the usual relevant unit adopted in CR/AR, where indeed it is regular distinguishing between, for instance, intra- and inter-sentential anaphora.

4. Conclusion and Future Work

In this paper, we provide an overview of the current state of a project aimed to build a Latin corpus enhanced with coreference and anaphora resolution. We detailed the annotation criteria and discussed a few annotation challenges, highlighting how this annotation layer necessitates a profound interaction among various fields of expertise, including linguistics, textual criticism, and literature.

In the near future, our aim is to expand the annotated corpus and to further extend the evaluation of inter-annotator agreement by incorporating the metrics as those proposed by Kopeć and Ogrodniczuk [35], such as the MUC score [36]. Once a sufficiently large dataset will be available, NLP will be concerned too, as we plan to exploit the annotated dataset to train and evaluate a stochastic model in supervised fashion to perform automatic CR/AR of Latin, usable also in NLP pipelines like, for instance, UDPipe [37] and Stanza [38]. We expect such a model to prove helpful to provide the Latin treebanks currently available in the Universal Dependencies (UD) initiative [39] with a layer of so-called enhanced dependencies, which also includes coreference and anaphora resolution. This would position Latin on an equal footing with other contemporary languages for which CR/AR annotations are also publicly accessible in treebanks [40] [11]. Given that one of the UD Latin treebanks, the *Index Thomisticus* Treebank, is already published as Linked Data in the LiLa Knowledge Base [41], having the treebank enriched with enhanced dependencies will require to model and publish therein the metadata about CR/AR.

The contribution of our project can also be considered within the broader context of NLP task on Latin. For instance, the corpus enriched with coreference annotations could enhance a task such as Emotion Polarity Detection, which was one of the shared tasks at the last edition of the evaluation campaign EvaLatin 2024. In the long term, a follow-up of the project will consist in building further textual datasets that feature other layers of coreferential annotation recognized by the GUM framework, such as appositive, attributive, and predicative coreferences, along with discourse deixis, and non-proper coreferences. Finally, given the current spread of Large Language Models and their highly promising accuracy rates on a wide range of NLP tasks, our data could be used to fine-tune

[11]https://universaldependencies.org/u/overview/enhanced-syntax.html

already models for Latin, such as the Latin BERT [42].

5. Acknowledgements

This contribution is funded by the PRIN 2022 project "Textual Data and Tools for Coreference Resolution in Latin" (CUP J53D23013680008), a project carried out jointly by the Università Cattolica del Sacro Cuore in Milan and by the University of Udine.

References

[1] M. Passarotti, F. Mambrini, G. Franzini, F. M. Cecchini, E. Litta, G. Moretti, P. Ruffolo, R. Sprugnoli, Interlinking through lemmas. the lexical collection of the lila knowledge base of linguistic resources for latin, Studi e Saggi Linguistici 58 (2020) 177–212.

[2] M. Passarotti, From syntax to semantics. first steps towards tectogrammatical annotation of latin, in: Proceedings of the 8th Workshop on Language Technology for Cultural Heritage, Social Sciences, and humanities (LaTeCH), 2014, pp. 100–109.

[3] T. Berners-Lee, J. Hendler, O. Lassila, The semantic web, Scientific american 284 (2001) 34–43.

[4] R. Sukthanker, S. Poria, E. Cambria, R. Thirunavukarasu, Anaphora and coreference resolution: A review, Information Fusion 59 (2020) 139–162.

[5] R. Busa, Trent'anni d'informatica su testi: a che punto siamo? quali spazi aperti alla ricerca?, in: Convegno su L'Università e l'evoluzione delle Tecnologie Informatiche, volume 1, CILEA, Milano, Italy, 1983, pp. 7.1–7.4.

[6] J. Nyhan, M. Passarotti, One Origin of Digital Humanities: Fr Roberto Busa in His Own Words, Springer, 2019.

[7] N. A. Chinchor, Overview of MUC-7, in: Seventh Message Understanding Conference (MUC-7): Proceedings of a Conference Held in Fairfax, Virginia, April 29 - May 1, 1998, 1998. URL: https://aclanthology.org/M98-1001.

[8] G. R. Doddington, A. Mitchell, M. A. Przybocki, L. A. Ramshaw, S. M. Strassel, R. M. Weischedel, The automatic content extraction (ace) program-tasks, data, and evaluation., in: Lrec, volume 2, Lisbon, 2004, pp. 837–840.

[9] S. Pradhan, L. Ramshaw, M. Marcus, M. Palmer, R. Weischedel, N. Xue, Conll-2011 shared task: Modeling unrestricted coreference in ontonotes, in: Proceedings of the fifteenth conference on computational natural language learning: shared task, 2011, pp. 1–27.

[10] S. Pradhan, A. Moschitti, N. Xue, O. Uryupina, Y. Zhang, Conll-2012 shared task: Modeling multilingual unrestricted coreference in ontonotes, in: Joint conference on EMNLP and CoNLL-shared task, 2012, pp. 1–40.

[11] S. Calhoun, J. Carletta, J. M. Brenier, N. Mayo, D. Jurafsky, M. Steedman, D. Beaver, The nxt-format switchboard corpus: a rich resource for investigating the syntax, semantics, pragmatics and prosody of dialogue, Language resources and evaluation 44 (2010) 387–419.

[12] A. Nedoluzhko, M. Novák, S. Cinková, M. Mikulová, J. Mírovský, Coreference in Prague Czech-English Dependency Treebank, in: N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis (Eds.), Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), European Language Resources Association (ELRA), Portorož, Slovenia, 2016, pp. 169–176. URL: https://aclanthology.org/L16-1026.

[13] E. Hinrichs, S. Kübler, K. Naumann, H. Telljohann, J. Trushkina, et al., Recent developments in linguistic annotations of the TüBa-D/Z treebank, Universitätsbibliothek Johann Christian Senckenberg, 2004.

[14] R. Iida, M. Komachi, K. Inui, Y. Matsumoto, Annotating a japanese text corpus with predicate-argument and coreference relations, in: Proceedings of the linguistic annotation workshop, 2007, pp. 132–139.

[15] A. Minutolo, R. Guarasci, E. Damiano, G. De Pietro, H. Fujita, M. Esposito, A multi-level methodology for the automated translation of a coreference resolution dataset: an application to the italian language, Neural Computing and Applications 34 (2022) 22493–22518.

[16] M. Recasens, M. A. Martí, Ancora-co: Coreferentially annotated corpora for spanish and catalan, Language resources and evaluation 44 (2010) 315–345.

[17] D. Bamman, G. Crane, The design and use of a latin dependency treebank, in: Proceedings of the Fifth Workshop on Treebanks and Linguistic Theories (TLT2006), Citeseer, 2006, pp. 67–78.

[18] M. Passarotti, The project of the index thomisticus treebank, in: Digital Classical Philology, De Gruyter Saur, 2019, pp. 299–320.

[19] V. B. Lenzi, G. Moretti, R. Sprugnoli, Cat: the celct annotation tool., in: LREC, 2012, pp. 333–338.

[20] W. W. Augustine, Confessions, Vol. 2: Books 9-13 (Loeb Classical Library, No. 27), 1912.

[21] P. Nixon, et al., Plautus, Vol. II: Casina. The Casket Comedy. Curculio. Epidicus. The Two Menaechmuses (Loeb Classical Library), William Heinemann; GP Putnam's Sons, 1917.

[22] L. R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (1945) 297–

302.

[23] T. Sorensen, A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on danish commons, Biologiske skrifter 5 (1948) 1–34.

[24] K. B. Cohen, A. Lanfranchi, M. J.-y. Choi, M. Bada, W. A. Baumgartner, N. Panteleyeva, K. Verspoor, M. Palmer, L. E. Hunter, Coreference annotation and resolution in the colorado richly annotated full text (craft) corpus of biomedical journal articles, BMC bioinformatics 18 (2017) 1–14.

[25] I. Hendrickx, G. Bouma, F. Coppens, W. Daelemans, V. Hoste, G. Kloosterman, A.-M. Mineur, J. Van Der Vloet, J.-L. Verschelde, A coreference corpus and resolution system for dutch., in: LREC, 2008.

[26] A. Nedoluzhko, J. Mírovskỳ, P. Pajas, The coding scheme for annotating extended nominal coreference and bridging anaphora in the prague dependency treebank, in: Proceedings of the Third Linguistic Annotation Workshop (LAW III), 2009, pp. 108–111.

[27] R. Simone, Fondamenti di linguistica, volume 9, Laterza Bari, 1990.

[28] T. McEnery, I. Tanaka, S. Botley, Corpus annotation and reference resolution, Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts (1997).

[29] H.-L. Trieu, A.-K. D. Nguyen, N. Nguyen, M. Miwa, H. Takamura, S. Ananiadou, Coreference resolution in full text articles with bert and syntax-based mention filtering, in: Proceedings of the 5th workshop on BioNLP open shared tasks, 2019, pp. 196–205.

[30] R. Thirukovalluru, N. Monath, K. Shridhar, M. Zaheer, M. Sachan, A. McCallum, Scaling within document coreference to long texts, Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021 (2021) 3921–3931.

[31] L. Cappiello, Un commento al curculio di plauto (vv. 1-370) (2015).

[32] G. Monaco, T. M. Plauto, Teatro di Plauto: Il Curculio. I, Istituto editoriale cultura europea, 1963.

[33] T. H. Gellar-Goad, Plautus' curculio and the case of the pious pimp, Roman Drama and its Contexts 34 (2016) 231.

[34] T. M. Plautus, E. Paratore, Il gorgoglione:(Il Gorgoglione). Testo latino con traduzione a fronte, Sansoni, 1958.

[35] M. Kopeć, M. Ogrodniczuk, Inter-annotator agreement in coreference annotation of polish, Advanced Approaches to Intelligent Information and Database Systems (2014) 149–158.

[36] B. Zheng, P. Xia, M. Yarmohammadi, B. V. Durme, Multilingual Coreference Resolution in Multiparty Dialogue, Transactions of the Association for Computational Linguistics 11 (2023) 922–940. URL: https://doi.org/10.1162/tacl_a_00581. doi:10.1162/tacl_a_00581. arXiv:https://direct.mit.edu/tacl/article-pdf/doi/10.116

[37] M. Straka, J. Hajic, J. Straková, Udpipe: trainable pipeline for processing conll-u files performing tokenization, morphological analysis, pos tagging and parsing, in: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), 2016, pp. 4290–4297.

[38] S. N. Group, et al., Stanza–a python nlp package for many human languages, 2018.

[39] M.-C. De Marneffe, C. D. Manning, J. Nivre, D. Zeman, Universal dependencies, Computational linguistics 47 (2021) 255–308.

[40] V. Ng, Supervised noun phrase coreference research: The first fifteen years, in: Proceedings of the 48th annual meeting of the association for computational linguistics, 2010, pp. 1396–1411.

[41] F. Mambrini, M. Passarotti, G. Moretti, M. Pellegrini, The index thomisticus treebank as linked data in the lila knowledge base, in: Proceedings of the Thirteenth Language Resources and Evaluation Conference, 2022, pp. 4022–4029.

[42] D. Bamman, P. J. Burns, Latin bert: A contextual language model for classical philology, arXiv preprint arXiv:2009.10053 (2020).