Using modern languages to parse ancient ones: a test on Old English

Luca Brigada Villa
University of Bergamo/Pavia
luca.brigadavilla@unibg.it

Martina Giarda
University of Bergamo/Pavia
martina.giarda@unibg.it

Abstract

In this paper we test the parsing performances of a multilingual parser on Old English data using different sets of languages, alone and combined with the target language, to train the models. We compare the results obtained by the models and we analyze more in deep the annotation of some peculiar syntactic constructions of the target language, providing plausible linguistic explanations of the errors made even by the best performing models.

1 Introduction

The performance of dependency parsing models for high-resource languages (HRLs) has improved significantly in recent years due to the availability of large annotated corpora and the advancement of deep learning techniques. Among others, models such as Stanza (Qi et al., 2020) and UDPipe (Straka, 2018) can achieve very high accuracy, with F1 scores approaching or even exceeding 0.90 on some treebanks datasets. This is true for some models for parsing data of (both modern and ancient) languages that have plenty of annotated resources, upon which is possible to train the models, while dependency parsing of low-resource languages (LRLs) is more problematic. The challenges that dependency parsing for LRLs has to face can be summarized as follows: a) data scarcity: LRLs often have limited annotated text corpora, which makes it difficult to train high-quality models and b) transfer learning limitations: transfer learning approaches that rely on models pre-trained on HRLs may not work well for LRLs due to the language-specificity of syntactic constructions.

As thoroughly discussed in Section 2.1, for what concerns dependency parsing, Old English (henceforth OE) can be considered a LRL, since the amount of annotated data available for this historical variety is scarce. Given these premises, we attempted an automatic parsing of OE, starting from the automatic conversion of the *York-Toronto-Helsinki Parsed Corpus of Old English Prose*[1] (henceforth YCOE) into a CoNLLU file, in which, however, the annotation is restricted to the sole morphological features retrievable from the YCOE annotation. Taking this as a starting point, we manually annotated 292 sentences, following the standards of Universal Dependencies (de Marneffe et al., 2021). Then we tested the results obtained training UUParser v2.4 (de Lhoneux et al., 2017b; Kiperwasser and Goldberg, 2016) on data coming from our set of annotated sentences in OE and a set of treebanks of three related languages, following Meechan-Maddon and Nivre's (2019) methodology, also followed by Karamolegkou and Stymne (2021) to test the performances of cross-lingual transfer learning for parsing Latin.

The paper is structured as follows: in Section 2 we introduce Old English providing a brief description of its history, developments, and typological features. In addition, we provide a brief survey of the main available resources for this language and introduce some issues that an automatic parsing of OE may face. In Section 3 we present our data and methodology. In Section 4 we overview the results of the parsing of OE data and discuss them. Finally, Section 5 concludes the paper and summarizes our findings.

2 Old English

Old English is a West-Germanic language, classified with Old Frisian and Old Saxon among the so-called Ingvaeonic languages. It was the language spoken in England after Angles, Saxons, Jutes and Frisians came to Britain and settled in the island in the 5th century. It is attested from the 7th century, except for some older brief runic inscriptions, whereas its ending point is conventionally established in 1066, date of the Norman Conquest of England (von Mengden, 2017b).

[1] https://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm

Typologically, OE shows a nominative-accusative alignment. Like other Indo-European languages, OE is a fusional language with inflectional word classes. Nouns are inflected by number and case, and follow three inflectional classes, depending on their original Proto-Germanic stem. After some merging processes, only four of the eight original Indo-European cases are found in OE: nominative, accusative, genitive, and dative. Some traces of the instrumental are present, but residual. Depending on the class, different cases can show syncretism. As other Germanic languages, OE has two main conjugational system: the so-called strong and weak verbs, the former building the preterit by means of apophony, i.e. the vowel alternation found in Present-Day English (PDE) irregular verbs, the latter with a dental suffix, just as PDE regular verb, whose past form is constructed with the *-ed* suffix. Finite OE verbs inflect for mood (indicative, subjunctive, imperative), tense (present and past), number, and person. Some forms show syncretism, in particular the plural in all moods and tenses, and the first and third person singular in the subjunctive (von Mengden, 2017a). Although some regularities may be found, word order in OE is not as rigid as in PDE (Mitchell and Robinson, 2012: 63-65), and it is still debated whether the basic word order was (S)VO or (S)OV. Like other ancient and modern Germanic languages, OE also exhibits V2, i.e. the tendency of the finite verb to follow the first constituent, regardless of its type. Concerning the order of other constituents, nouns are generally preceded by modifiers, e.g. demonstratives, adjectives, genitive complements. However the latter can follow the noun if another preceding modifier is present. In PPs, adpositions tend to precede a noun, but generally follow a pronoun; however, the opposite is also attested (Molencki, 2017). Contrary to PDE, OE allowed discontinuous constituents, above all in relative constructions.

2.1 Annotated resources for OE

Differently from other ancient languages, such as Latin or Ancient Greek,[2] and its contemporary counterpart, scholars have devoted little attention to the creation of resources to study Old English. The sole syntactically annotated resources for this language are the constituency treebank YCOE and its poetry counterpart, the *York-Helsinki Parsed Corpus of Old English Poetry*[3] (henceforth YCOEP), which follow the Penn style. Despite their value in size, these treebanks are hardly machine- nor user-friendly, have no interface and can only be investigated through their tool CorpusSearch2,[4] which require an intensive training in order to write even simple queries. There have been several attempts to convert constituency treebanks (particularly, Penn-style treebanks) into dependency-formats as the Estonian-EDT (Muischnek et al., 2014) and the Indonesian CSUI (Alfina et al., 2020), whereas, to our knowledge, no attempts in the opposite direction have been made.

2.2 Issues in automatically parsing OE data

An automatic parsing of such a free-ordered language can meet several problems. Regarding syntax, some problems may arise, given the freedom of word order and case syncretism, which may lead to a confusion, for instance, between subject and object constituents. Moreover, the use of both pre- and postpositions may result in erroneous annotation of oblique phrases. Another problematic issue is the parsing of relative clauses, which can be marked by a variety of means or even left unmarked, and often show non-projectivity.

3 Data and methods

3.1 Starting point and initial issues

Our data consist of two prose OE texts, *Adrian and Ritheus* and the first homily of Ælfric's *Supplemental Homilies*,[5] for a total of 292 sentences.[6] Both texts are written in the West-Saxon dialect and have religious content. First, the texts were coverted from the YCOE-format to a CoNLLU-file containing the POS and the morphological features retrievable from the YCOE annotation, i.e. case for nouns and adjectives, and mood and tense for verbs (when not ambiguous). Second, we manually annotated the remaining morphological fea-

[2]The latest release of UD (v2.11) includes 5 treebanks for Latin and 2 for Ancient Greek.

[3]https://www-users.york.ac.uk/~lang18/pcorpus.html

[4]https://corpussearch.sourceforge.net/CS.html.

[5]These are the first two texts in the YCOE treebank. *Adrian and Ritheus* is dialogue on several biblical issues (Cross and Hill, 1982 : 3-4). On the other hand, Ælfric's homily, *Nativitas Domini*, is a Christmas homily, with several expansions, consisting in scriptural elaborations (Pope, 1968 : 191-195).

[6]Data and scripts can be found at https://github.com/unipv-larl/wundorsmitha-geweorc/tree/main/paper_projects/parsing_oe_modern

tures, lemmatization and syntactic dependencies, following Universal Dependencies guidelines. This choice is due to these reasons: UD is the *de facto* standard for the annotation dependency treebanks; moreover, it allows for comparison, which is useful for both typological and historical analyses.

Some problematic issues derive from the conversion of texts itself: the YCOE tags as `P` both adpositions and subordinating conjunctions, which would be tagged, respectively, as `ADP` and `SCONJ` in Universal Dependencies. In the conversion, both options have been kept, to manually disambiguate them. Moreover, the verbs *beon* and *wesan* 'to be' and *weorþan* 'to become' have their specific tag in the YCOE annotation, i.e. `BE*`. Given the frequency of copular and passive constructions in which they appear, they have been all converted to `AUX`. However, this tagging disregards their occurrences as existential verbs, which should be tagged as `VERB`. As a general tendency, we chose not to include subtypes of the syntactic lables, except for the following cases:

- `advcl:relcl`, indicating a relative clause;
- the subtypes indicating a passive construction, i.e. `nsubj:pass`, `aux:pass` and `obl:agent`;
- `advmod:neg` for the negative particle and adverb *ne* and *na*,
- the specific `advmod:tmod` and `advmod:lmod` only when they were single-word adverbs, tagged in the YCOE as `ADV^L` and `ADV^T`;
- `obl:tmod` and `obl:lmod` have only been used when there was a unambiguous, not metaphorical interpretation.

3.2 Support languages

We used UUParser v2.4 (de Lhoneux et al., 2017b), a transition-based parser which is able to train multilingual models. Given the small amount of annotated sentences, we chose a multilingual parser, in order to test whether the inclusion of support languages in the training phase could have a beneficial impact on the parsing of OE sentences or not. To do so, we selected three languages related to OE since the addition of related languages has shown to be effective in the tests described in de Lhoneux et al. (2017a) and Meechan-Maddon and Nivre (2019).

While Meechan-Maddon and Nivre (2019) had three modern languages (Faroese, Upper Sorbian and North Saami) as target languages for the experiment, which resulted in an easier choice of languages to be used as support to train the models, our choice to focus on OE brings some issues in selecting the support languages. PDE has been excluded, due to its diachronic evolution: English has lost both nominal and verbal inflection, has developed a rigid SVO order, and its lexicon has been enriched by many French loanwords. Even though not part of the same sub-branch, i.e. Ingvaeonic, other modern Germanic languages present features that are closer to OE morphosyntax. In particular, we selected Modern Icelandic, Modern Swedish, and Modern German. The former two are part of the North-Germanic branch, whereas the latter is part of the West-Germanic branch, to which OE, too, belongs. Icelandic is considered the most archaic of Germanic languages, since it has retained many morphological and syntactical characteristics of Old Norse (Bandle et al., 2005: 1872). Some of its features compatible with OE are: a) prenominal definite determiners; b) pre- and post-nominal attributive genitive; c) the so-called "oblique objects" (i.e. impersonal constructions); d) the presence of verb-auxiliary constructions. The last feature is lost in Swedish, which has also undergone a process of morphological simplification. However Swedish features, as OE, prenominal possessive determiners, while Icelandic has mainly postnominal possessives (Bandle et al., 2005: 1874). Nonetheless, both Scandinavian languages show a fixed SVO order, which contrasts with the free OE word order. The Scandinavian languages, as well as German, are V2 languages, like OE. Regarding the West-Germanic branch, German is similar to OE in that it retains, at least in subordinate clauses, a verb-final order. Similarly to OE, it has both prepositions and postpositions. Both German and OE have prenominal definite determiners and attributive genitive both pre- and postnominal positions (Haider, 2010). Of the three support languages, Swedish shows the major innovations, whereas Icelandic and German may give better results.

3.3 Experimental setup

We split our sample of manually annotated OE sentences in three sets (see Table 1) and from Universal Dependencies v2.11 (de Marneffe et al., 2021), we selected one treebank for each of the support languages, namely UD Swedish-Talbanken (Nivre and Megyesi, 2007), UD Icelandic-Modern and

UD German-GSD (McDonald et al., 2013).

	train	dev	test	total
tokens	2673	1308	1334	5315
sentences	149	73	70	292

Table 1: The sets resulted from splitting OE data.

We reduced the treebanks of the support languages to 60k tokens to avoid the effect on the results that the size of the treebanks might have,[7] and we converted the characters which were not in the target language as shown in Table 6 in Appendix A.

Then, for each one of the combinations of the four languages (the target language and the three support languages), we performed the training of the models and, after the training phase, we used the best model to parse the OE test set. Our workflow followed these steps:

1. we used UUParser to train the model (30 epochs)

2. the epoch that had the best LAS on OE dev data was selected as the best model

3. we parsed the OE test data using the best model

The training phase did not take into account the part-of-speech tags, even if the parser is able to learn embeddings of POS tags if a specific option is given. We decided not to use that option since we wanted to test how well the model performed in a common situation when it comes to work with OE data, that is not having POS annotated texts.

In Section 4 we show the results achieved by each model and discuss them.

4 Results and discussion

Table 2 shows the accuracy reached by each model measured on the parsing of OE test data. At first glance, we can see that the model trained using only OE data significantly outperforms each of the models trained without OE data in the training set. This applies for both the monolingual and multilingual models and seems to confirm what was found by Meechan-Maddon and Nivre (2019).

[7]This is the main reason why we did not consider the Gothic PROIEL treebank (Haug and Jøhndal, 2008), even if its inclusion could have improved the parsing scores. In this work we decided to restrict the set of languages to be considered as support data for our models to the modern ones.

Considering the metrics, the best-performing models were the ones trained with Icelandic and OE data, which achieved the best Unlabeled Attachment Score (UAS) and Labeled Attachment Score (LAS), together with the one trained with Icelandic, German and OE data, which achieved the best Label Accuracy (LA).[8]

Considering the UAS and the LAS achieved by the models, it is surprising to notice that the model that performed best was the one trained upon only Icelandic and target language data, since the Icelandic monolingual model was the one which obtained significantly worse results than the other monolingual models. For what concerns the LA, it seems reasonable to see a model trained on German data performing better that the others considering that the monolingual model trained upon German was the one that achieved the best scores among the monolingual models trained without OE data, even though such multilingual model was trained also upon Icelandic data. Finally, all models trained including the target language data achieved better results than their counterparts trained without having the target language data in the training set, even though the best performances are achieved combining Icelandic and German with OE data. This seems reasonable in light of what discussed in Section 3.2.

In the following sections we will analyze more in detail the output of the parsing phase of the two models which scored the highest metrics (is+target and de+is+target) and the monolingual model trained only upon OE data. We will focus on the deprels `advmod` and `obl` for the following reasons: the former showed unexpectedly low results for the OE model (as shown in Table 3); the latter allows investigating whether postpositions have been recognized and correctly annotated. We will also concentrate on `advcl:rel`, as relative clauses can be marked by different pronouns and can show non-projectivity. We will discuss and exemplify the output of the models for these constructions, using four erroneously annotated sentences. Finally, in Section 4.4, we will show some recurrent errors made by the models tagging the dependency relations and the impact of a rule we designed to correct the output of the parsing process.

[8]The LA was measured dividing the number of token whose deprel was tagged correctly by the number of tokens in the test set.

	-Target			+Target		
	UAS	**LA**	**LAS**	**UAS**	**LA**	**LAS**
Old English				60.79	64.39	47.23
sv	27.06	24.44	9.45	65.07	73.61	57.20
de	32.91	25.34	10.12	65.82	72.19	56.45
is	20.31	22.64	4.57	**68.44**	73.76	**58.70**
sv+de	32.16	25.56	10.42	65.82	72.19	57.42
sv+is	26.39	23.76	9.45	64.62	70.09	54.42
de+is	30.73	27.74	11.17	66.34	**74.29**	57.42
sv+de+is	32.46	24.96	11.02	65.97	71.66	57.57

Table 2: UAS, LA and LAS of each model measured on the parsing of OE test data. -Target = cross-lingual models trained without target language data. +Target = models trained including target language data.

		`advmod`	`obl`	`acl:relcl`
oe	**P**	41.67	61.26	62.50
	R	35.71	80.95	52.63
oe-is	**P**	65.45	65.42	72.22
	R	51.43	83.33	86.67
oe-de-is	**P**	58.21	70.93	51.85
	R	55.71	72.62	63.64

Table 3: Precision (P) and Recall (R) for the dependency relations `advmod`, `obl` and `acl:relcl`.

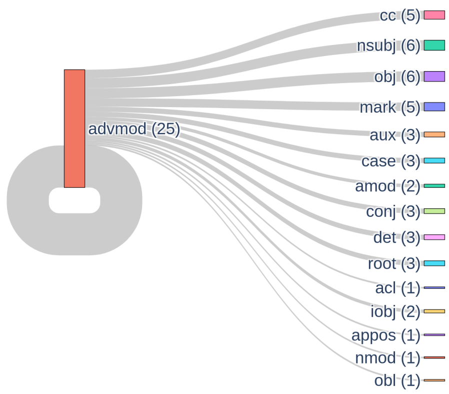

Figure 1: How oe model tagged tokens which had to be tagged as `advmod` (see Figures 9 and 10 in Appendix A for the other models).

4.1 The deprel `advmod`

As shown in Figure 1, no relevant patterns of error seem to be present. However, it is remarkable that many of the errors are found with the word *ne* 'not' and *swa* 'so'. Both can have several functions in the sentences: *ne* can either be a negative adverb or a negative conjunction, whereas *swa* can introduce a subordinate clause or function as an adverb. The different usages are distinguished in the UPOS, which however is not considered by the models, causing confusion in the syntactic annotation as well.

An interesting example is shown in Figure 2, where the adverb *swutelicor* 'more clearly' has been annotated by the three models as `obj` of the verb *cweðað* '(we) talk' in this context, but generally 'say'. This can be accounted for in light of the absence of a true direct object depending on the verb.

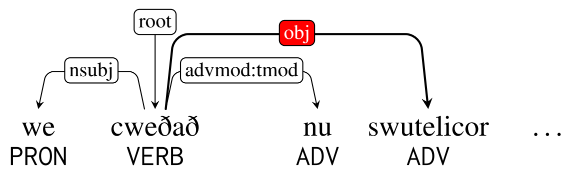

Figure 2: Dependecy tree of part of the sentence 'we cweðað nu swutelicor , on þam Godes wisdome , þe is witodlice lif , & cann wyrcan his weorc be his dihte' ('we now talk more clearly about God's wisdom, which truly is life, and can make his actions by his command'). Correct annotation in Figure 15 in Appendix B.

4.2 The deprel `obl`

As in 4.1, no significant patterns of error can be identified for the deprel `obl`. Remarkably, Figure 5 compared to Figures 11 and 12 in Appendix A shows that the best model in this respect is the one trained only on OE data.

One example of incorrect annotation is worth discussing: as touched upon in Section 2.2, the annotation of postpositions has been problematic. None of the three models could correctly recognize that the adposition *ongean* 'against, towards' depended on the preceding pronoun *hiom* 'them'. The OE model considered *hiom* as `case`, directly

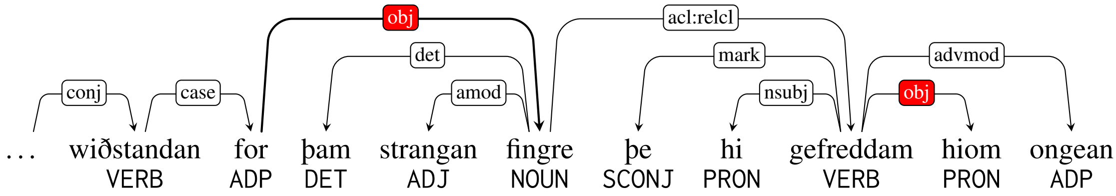

Figure 3: Dependency tree of '[... & hi ne mihton na leng Moyse] wiðstandan for þam strangan fingre þe hi gefreddan hiom ongean.' ('[and they could no longer] withstand [Moyses] for that strong finger that they felt against them '). This is the output of the oe-de-is model, see Figure 16 in Appendix B for the correct tree.

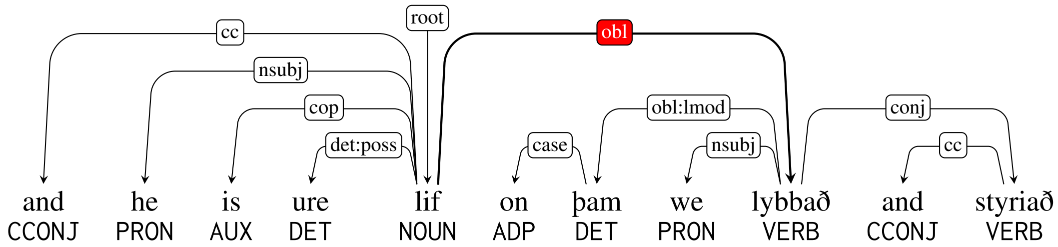

Figure 4: Dependency tree of part of the sentence '& he is ure lif on þam we lybbað & styriað, & on þam we syndon, swa swa us sæde Paulus.' ('and he is our life, in which we live and move, in which we are, so as Paul said to us'). This is the output of the oe-is model, see Figure 17 in Appendix B for the correct tree.

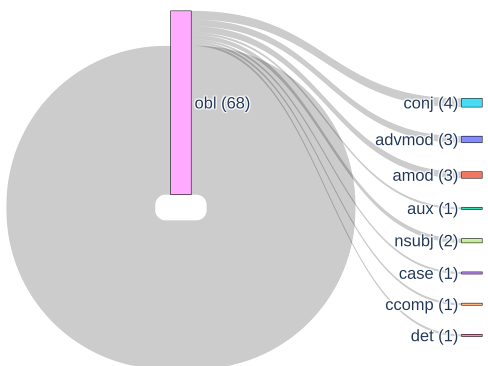

Figure 5: How oe model tagged tokens which had to be tagged as `obl` (see Figures 11 and 12 in Appendix A for the other models).

depending on the preceding verb *gefreddan* 'feel, perceive'. On the other hand, both multilingual models considered *hiom* as the object of *gefreddan* and *ongean* an adverb modifying the verb, as shown in Figure 3.

4.3 The deprel `acl:relcl`

Most problems in the annotation of relative clauses are: a) the great variability in the relative pronouns marking them, and b) non-projectivity. Concerning the point in a), OE has an invariable complementizer *þe*, which generally functions as a relative marker, at times accompanied by the determiner *se, seo, þæt*. However, the determiner can be found without the complementizer to mark relative clauses, above all when part of PPs, or relative clauses can simply be left unmarked. Other POS, e.g. locative adverbs, can function as relative pronouns. All three models tended to make the same errors, generally recognizing and annotating correctly only the sentences with *þe*, and making mistakes when this element did not occur.

An example of this is shown Figure 4, where the PP *on þam* 'in which' (lit. 'in the.DAT') was not recognized by the models as marking the relative clause. The sole model which recognized that this was a subordinate clause was the de-is-oe model, which annotated it as `advcl` correctly depending on the noun *lif* 'life', whereas the other models considered it a nominal constituent (either `conj` or `obl`). Another issue is that OE allowed for discontinuity in relative clauses, which could be separated from their antecedent by other constituents. Some of the errors are probably due to this, as shown in the sentence in Figure 6. This sentence shows how the relative clause *þe forlærdon Farao* 'which corrupted the Pharaoh', was not considered as depending on the noun *drymen* 'joys', given that the two constituents are separated by two PPs. What the multilingual model annotated as relative clause (erroneously, as it read the verb *forlærdon* as modifying *Farao*) is dependent on the nearest noun,

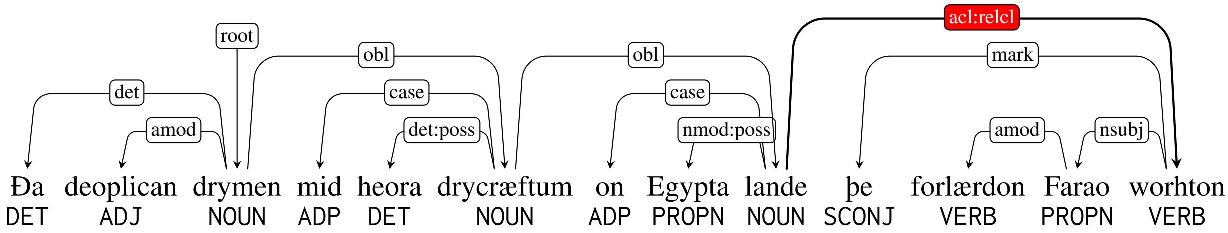

Figure 6: Dependency tree of 'Đa deoplican drymen mid heora drycræftum on Egypta lande þe forlærdon Farao worhton [tacna ongean Moysen of þam ylcan antimbre þe God ær gesceop...]' ('The deep joys, **which corrupted the Pharaoh** with their magical arts in the lands of Egypt, made...'). This is the output of the oe-de-is model, see Figure 18 in Appendix B for the correct tree.

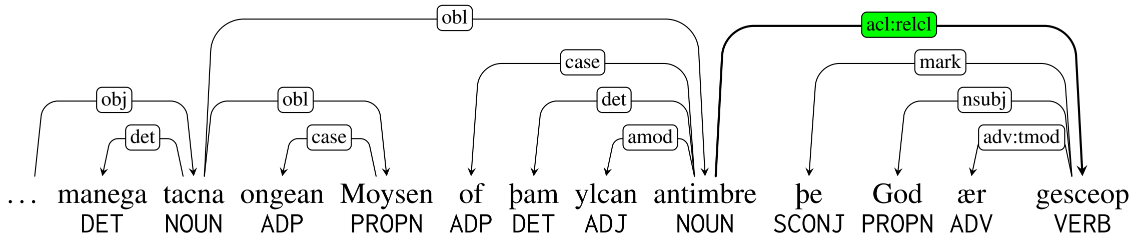

Figure 7: Dependency tree of '[Đa deoplican drymen mid heora drycræftum on Egypta lande þe forlærdon Farao worhton] tacna ongean Moysen of þam ylcan antimbre þe God ær gesceop...' ('...[made] towards Moyses many signs of the same substance, **which God had created before**...'). This is the output of the oe-de-is model, see Figure 19 in Appendix B for the correct tree.

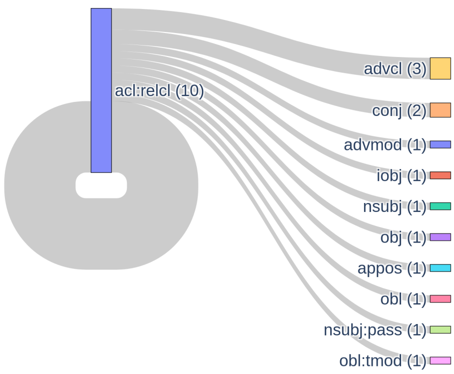

Figure 8: How oe model tagged tokens which had to be tagged as `acl:relcl` (see Figures 13 and 14 in Appendix A for the other models).

i.e. *lande* 'lands'. On the contrary, the following relative clause (Figure 7) *þe God ær gesceop* 'which God had created before', has been annotated correctly by all three models, as it is immediately preceded by its antecedent, *antimbre* 'substance'.

4.4 Recurrent erroneous dependency relations

During the manual check of the output generated by the models, we noticed some recurrent errors that the models could have avoided. These errors are due to the fact that the generated tree and the annotation of dependency relations do not take into account the POS of the tokens.

form	upos	xpos	deprel
ne	CCONJ	any	cc
ne	PART	any	advmod:neg
any	any	starts with MD	aux
any	any	ADV^L	advmod:lmod
any	any	ADV^T	advmod:tmod

Table 4: Deprel correction table (upos=universal part-of-speech; xpos=language-specific part-of-speech).

We decided to assign automatically a dependency relation to tokens which had certain features, as displayed in Table 4. As discussed in Section 4.1, the word *ne* 'not, nor' could function both as negative particle (in which case was assigned PART as universal POS), but also as a negative coordinative conjunction, thus assigned CCONJ as universal POS. Syntactically, the former function can be labeled only as `advmod:neg`, while the latter only as `cc`, whether it conjuncts two NPs/PPs or two clauses. For this reason, we automatically assigned the deprel `advmod:neg` to all the occurrences of *ne*, whose UPOS was PART, and the deprel `cc` to those tagged as `CCONJ`. Together with *ne*, we also mentioned *swa* 'so', as a frequent error in `advmod`. However, we could not proceed to an automatic correction of it, as we did with *ne*, since *swa* tagged as `ADV` can also appear in the fixed expression *swa swa* 'so, in the same way', introducing a subordinate clause. In this case, the first *swa*, whose UPOS is `ADV`, should be annotated as `fixed`, instead of `advmod`.

	before		after	
	LA	**LAS**	**LA**	**LAS**
oe	64.39	47.23	66.79	48.28
oe-is	73.76	58.70	75.34	59.30
oe-de-is	74.29	57.42	75.79	58.17

Table 5: Comparison between the LA and the LAS before and after the correction.

We also noticed many errors in the annotation of modals, which in the YCOE are all tagged as `MD` (and its variants, which show mood and tense). Following Universal Dependencies guidelines, they should all be annotated as `aux`, making an automatic correction of these errors possible. The original YCOE annotation is useful also with temporal and spatial adverbs. They were originally tagged as `ADV^L` and `ADV^T`, which can easily be automatically converted, respectively, in `advmod:lmod` and `advmod:tmod`, correcting both main deprel and the subtype.

Our correction affected the label accuracy of the treebanks resulting on an increase of 1 or 2 points depending on the model, which had an impact also on the LAS, as shown in Table 5.

5 Conclusion

In this paper we tested the dependency parsing performances of four monolingual models and seven multilingual models on Old English data. We showed that the model trained just using data of the target language achieved far better results than the models (both monolingual and multiliguals) trained without target language data and that, out of the three support languages we selected, Icelandic and German combined better than Swedish according to the scores reached parsing OE test data. As discussed in 3.2, we expected this result given the fact that Modern Icelandic and Modern German retained many morphosyntactic features similar to those of the target language.

Then, we also discussed some cases of problematic annotation: in Sections 4.1, 4.2, 4.3 we gave some linguistic explanations of the errors made by the best models, which include `advmod`, `obl` and `acl:relcl` showing that some poor results might be due to the peculiarity of such constructions in OE. Finally, in 4.4, we discussed the impact which the correction of the dependency relation annotation using some rules based on the word forms, the universal parts-of-speech and the language-specific parts-of-speech had on the results achieved by the best models. This errors might have been avoided if we had used the option to force the models to learn embeddings for the parts-of-speech during the training phase, which would have made the parsing process aware of the already annotated parts-of-speech. The situation in which the POSs are annotated, though, is not so usual for OE, except for the above-mentioned YCOE and YCOEP treebanks.

Our test, following the methodology described in Meechan-Maddon and Nivre (2019), led to the same conclusions in terms of the benefits that support languages have on the parsing scores when combined to OE data during the training phase. In particular this is true when the support languages are related to the target language or, at least, share a significant number of features with the target language.

This approach has proven useful for our broader twofold aim: a) having an alternative to a rule-based conversion of the YCOE(P) treebanks and b) developing a tool to annotate other OE texts, which are not included in the above-mentioned treebanks. Despite the challenges, using this approach to parse historical languages can accelerate the process of creating new resources and produce outputs that, while not perfect, are satisfactory in terms of dependency parsing.

Aknowledgments

We would like to thank the anonymous reviewers who provided great insights and suggestions. We also thank Chiara Zanchi and Luigi Talamo for their precious comments and Andrew Dyer for the suggestions which led to this paper. The paper is the result of close collaboration between the two authors. For academic purposes, Luca Brigada Villa is responsible of training the models and parsing the data, and of Sections 1, 3.3, 4, 5. Martina Giarda is responsible of the manual annotation, and of Sections 2, 3, 3.1, 3.2, 4.1, 4.2, 4.3, 4.4.

References

Ika Alfina, Daniel Zeman, Arawinda Dinakaramani, Indra Budi, and Heru Suhartanto. 2020. Selecting the ud v2 morphological features for indonesian dependency treebank. In *2020 International Conference on Asian Language Processing (IALP)*, pages 104–109.

Oscar Bandle, Kurt Braunmüller, Ernst Hakon Jahr, Allan Karker, Hans-Peter Naumann, Ulf Telemann, Lennart Elmevik, and Gun Widmark. 2005. *The Nordic Languages*, volume 2. De Gruyter, Berlin,Boston.

James E. Cross and Thomas D. Hill. 1982. *The Prose Solomon and Saturn and Adrian and Ritheus*. University of Toronto Press, Toronto.

Miryam de Lhoneux, Yan Shao, Ali Basirat, Eliyahu Kiperwasser, Sara Stymne, Yoav Goldberg, and Joakim Nivre. 2017a. From raw text to Universal Dependencies - look, no tags! In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 207–217, Vancouver, Canada. Association for Computational Linguistics.

Miryam de Lhoneux, Sara Stymne, and Joakim Nivre. 2017b. Arc-hybrid non-projective dependency parsing with a static-dynamic oracle. In *Proceedings of the The 15th International Conference on Parsing Technologies (IWPT).*, Pisa, Italy.

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

Hubert Haider. 2010. *The Syntax of German*. Cambridge University Press, Cambridge.

Dag T. T. Haug and Marius L. 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.

Antonia Karamolegkou and Sara Stymne. 2021. Investigation of transfer languages for parsing Latin: Italic branch vs. Hellenic branch. In *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 315–320, Reykjavik, Iceland (Online). Linköping University Electronic Press, Sweden.

Eliyahu Kiperwasser and Yoav Goldberg. 2016. Simple and accurate dependency parsing using bidirectional LSTM feature representations. *TACL*, 4:313–327.

Ryan McDonald, Joakim Nivre, Yvonne Quirmbach-Brundage, Yoav Goldberg, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, Claudia Bedini, Núria Bertomeu Castelló, and Jungmee Lee. 2013. Universal Dependency annotation for multilingual parsing. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 92–97, Sofia, Bulgaria. Association for Computational Linguistics.

Ailsa Meechan-Maddon and Joakim Nivre. 2019. How to parse low-resource languages: Cross-lingual parsing, target language annotation, or both? In *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest 2019)*, pages 112–120, Paris, France. Association for Computational Linguistics.

Bruce Mitchell and Fred C. Robinson. 2012. *A guide to Old English. Eighth edition*. John Wiley & Sons, Malden, Oxford.

Rafał Molencki. 2017. Syntax. In Laurel J. Brinton and Alexander Bergs, editors, *The History of English. Old English*, volume 2, chapter 5, pages 100–124. De Gruyter Mouton, Berlin, Boston.

Kadri Muischnek, Kaili Müürisep, Tiina Puolakainen, Eleri Aedmaa, Riin Kirt, and Dage Särg. 2014. Estonian Dependency Treebank and its annotation scheme. In *Proceedings of the 13th Workshop on Treebanks and Linguistic Theories (TLT13)*, pages 285–291, Tübingen, Germany.

Joakim Nivre and Beata Megyesi. 2007. Bootstrapping a swedish treebank using cross-corpus harmonization and annotation projection. In *Proceedings of the 6th international workshop on treebanks and linguistic theories*, pages 97–102. Association for Computational Linguistics Pennsylvania, PA.

John C. Pope. 1968. *Homilies of Ælfric: a Supplementary Collection*. Early English Society, Oxford University Press, London.

Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*.

Milan Straka. 2018. UDPipe 2.0 prototype at CoNLL 2018 UD shared task. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 197–207, Brussels, Belgium. Association for Computational Linguistics.

Ferdinand von Mengden. 2017a. Morphology. In Laurel J. Brinton and Alexander Bergs, editors, *The History of English. Old English*, volume 2, chapter 5, pages 73–99. De Gruyter Mouton, Berlin, Boston.

Ferdinand von Mengden. 2017b. Old english: Overview. In Laurel J. Brinton and Alexander Bergs, editors, *The History of English. Old English*, volume 2, chapter 3, pages 32–49. De Gruyter Mouton, Berlin, Boston.

A Additional tables and figures

character	conversion
ä	æ
ö	o
ü	u
Ä	A
Ö	O
Ü	U
ß	ss
á	a
é	e
í	i
ó	o
ú	u
ý	y
Á	A
É	E
Í	I
Ó	O
Ú	U
Ý	Y
å	a
Å	A

Table 6: The character conversion table.

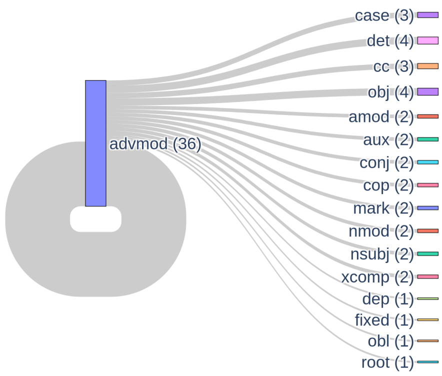

Figure 9: How oe-is model tagged tokens which had to be tagged as `advmod`.

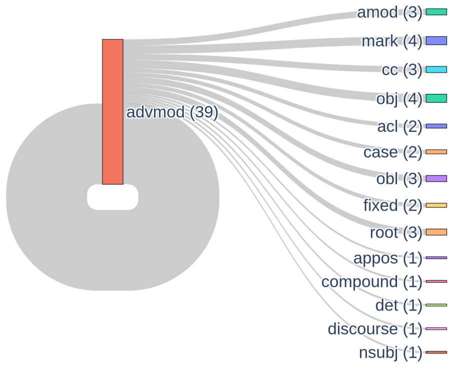

Figure 10: How oe-de-is model tagged tokens which had to be tagged as `advmod`.

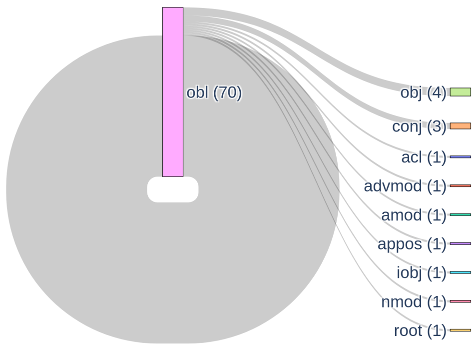

Figure 11: How oe-is model tagged tokens which had to be tagged as `obl`.

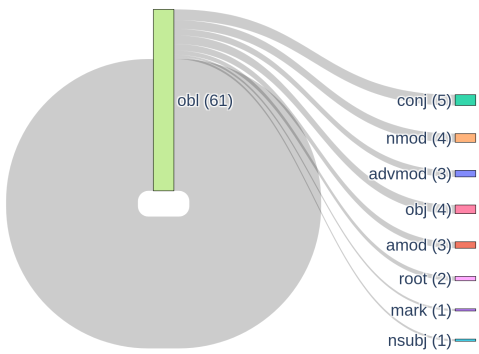

Figure 12: How oe-de-is model tagged tokens which had to be tagged as `obl`.

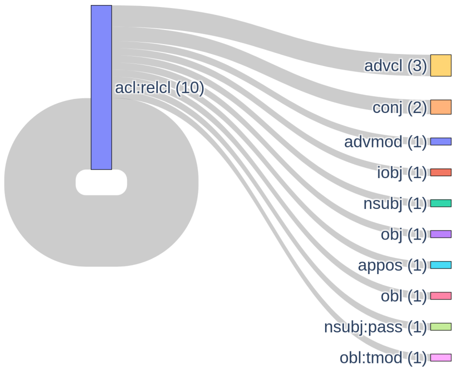

Figure 13: How oe-is model tagged tokens which had to be tagged as `acl:relcl`.

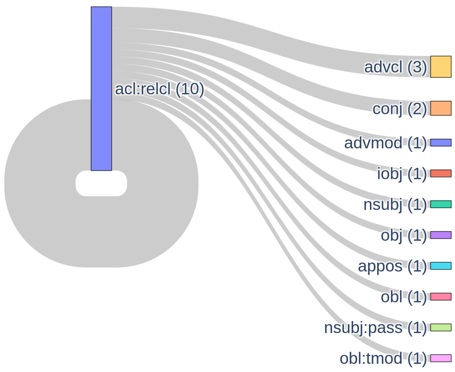

Figure 14: How oe-de-is model tagged tokens which had to be tagged as `acl:relcl`.

B Additional trees

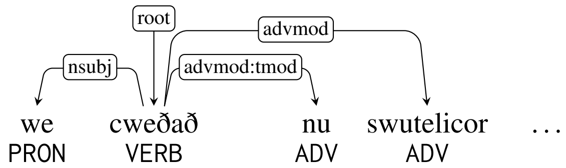

Figure 15: Correct version of the dependency tree in Figure 2.

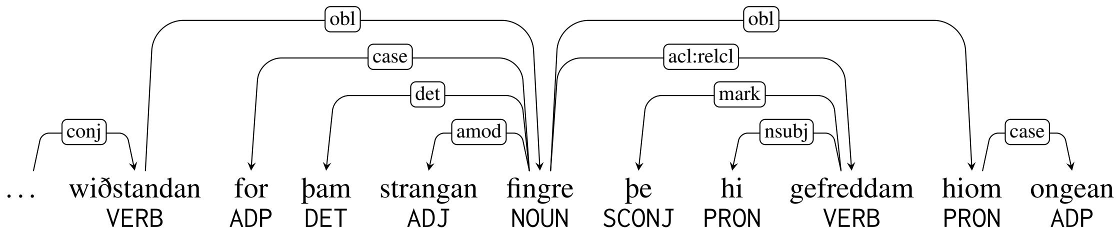

Figure 16: Correct version of the dependency tree in Figure 3.

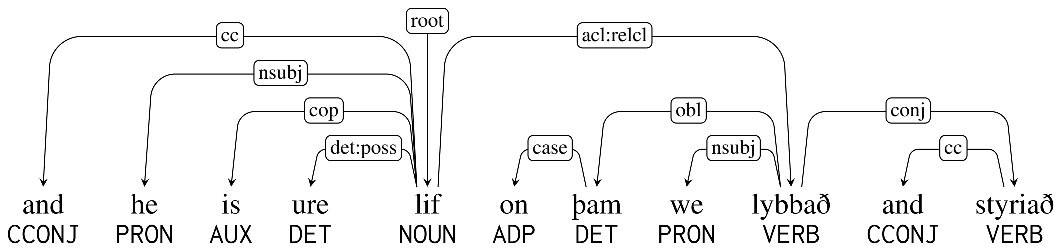

Figure 17: Correct version of the dependency tree in Figure 4.

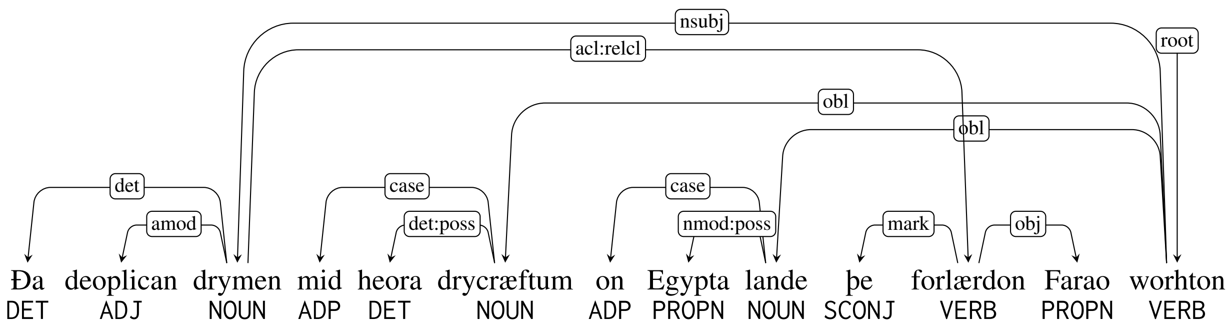

Figure 18: Correct version of the dependency tree in Figure 6.

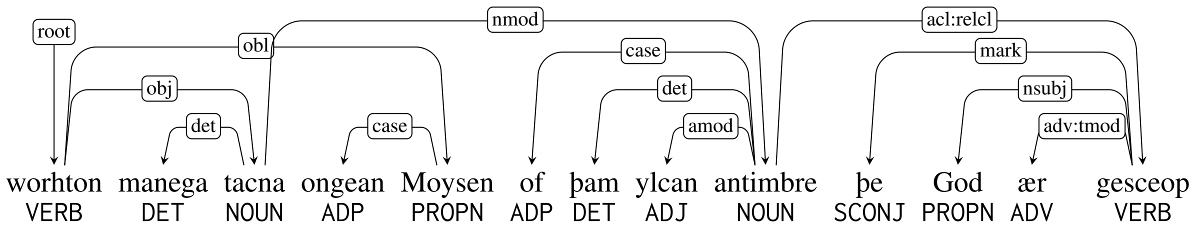

Figure 19: Correct version of the dependency tree in Figure 7.