Diachronic Parsing of Pre-Standard Irish

Kevin P. Scannell
Department of Computer Science
Saint Louis University
St. Louis, Missouri, USA 63103
kscanne@gmail.com

Abstract
Irish underwent a major spelling standardization in the 1940’s and 1950’s, and as a result it can be challenging to apply language technologies designed for the modern language to older, “pre-standard” texts. Lemmatization, tagging, and parsing of these pre-standard texts play an important role in a number of applications, including the lexicographical work on Foclóir Stairiúil na Gaeilge, a historical dictionary of Irish covering the period from 1600 to the present. We have two main goals in this paper. First, we introduce a small benchmark corpus containing just over 3800 tokens, annotated according to the Universal Dependencies guidelines and covering a range of dialects and time periods since 1600. Second, we establish baselines for lemmatization, tagging, and dependency parsing on this corpus by experimenting with a variety of machine learning approaches.

Keywords: parsing, part-of-speech tagging, diachronic treebank, Irish, lexicography

1. Introduction
Irish is relatively well-resourced in terms of language technologies for grammatical analysis, including a rule-based part-of-speech tagger (ÓDhonnchadha and van Genabith, 2006) and a dependency parser (Lynn, 2016) that both achieve high levels of accuracy. Older texts present a problem for these resources, however, in part because of a significant spelling reform that was undertaken in the 1940’s and 1950’s with the introduction of an official standard for the written language, An Caighdeán Oifigiúil (Rannóg an Aistriúcháin, 1945). The standard resulted in an orthography that was both simpler (e.g. déidheannaighe becomes déanai) and more consistent (e.g. Meirceá, Meiricea, Aimeirice, Meirioca, … and so on all become Meiriceda), and has been embraced widely by the Irish-speaking community. In addition to the challenges presented by this orthographic discontinuity, older texts exhibit a number of grammatical features that have all but disappeared in the modern language, e.g. various synthetic verb forms, wide use of the nominal dative case, etc. The language technologies that exist for the modern language are unable to handle these phenomena in a reliable way.

Lemmatization, tagging, and parsing of these pre-standard texts are all of tremendous importance. First and foremost, these are important enabling technologies for lexicography. There are two significant lexicographical projects underway in Ireland at present: the Royal Irish Academy’s historical dictionary of Irish covering the period from 1600 to the present and new general-purpose monolingual and bilingual dictionaries funded by Foras na Gaeilge. Both projects make use of large corpora that include millions of words of pre-standard text. Effective searching of these corpora for lexicographical purposes is impossible without, at minimum, indexing them by standardized lemmas and parts of speech.

Grammatical analysis of older texts has other potential applications, for example as an aid to historians or linguistic scholars who are engaging with Early Modern Irish source texts, a challenging task even for those with a fluent command of modern Irish. The Léamh project was established with precisely this audience in mind; the project website provides a grammar and glossaries for Early Modern Irish, as well as several carefully annotated texts to help scholars learn the nuances of the language. At present, these texts are produced through time-consuming manual annotation; with suitable language technologies tailored to this time period, additional texts could be prepared much more quickly. Currently, there are no resources for direct tagging or parsing of pre-standard texts. Instead, the general strategy has been to start with a best-effort automatic standardization (Scannell, 2014), and then to make use of modern taggers and parsers. Good results have been obtained with this approach, although there are some inherent limitations. First, given the absence of tools for direct analysis of the source texts, the standardizer must do its job without part-of-speech tags or other linguistic annotations. Instead, it relies only on “shallow” techniques: a set of rule-based spelling changes, a large lexicon of pre-standard/standard word mappings, and a language model on the target (modern Irish) side. Second, the standardization task generally becomes more difficult for older texts, and errors introduced by the standardizer, along with the frequent occurrence of out-of-vocabulary words, negatively impact the quality of tagging and parsing. Third, by definition, this approach is unable to handle grammatical phenomena that do not

---

1See https://www.ria.ie/research-projects/focloir-stairiul-na-gaeilge
2See https://léamh.org/about-the-project/
occurs in the modern language.
The goal of this paper is two-fold. First, we present a
new reference corpus of pre-standard texts published
between 1602 and 1936, representing various time
periods and dialects, and annotated according to the
Universal Dependencies (UD) guidelines (Nivre et al.,
2016). Second, we experiment with a number of tag-
gging and parsing models and evaluate them on this
reference corpus, establishing baseline scores for lemmat-
ization, part-of-speech tagging, and dependency parsing
on pre-standard Irish.

2. Related Work

Text analysis tools for standard Irish

As noted above, modern Irish is relatively well-
resourced among minority languages in terms of lan-
guage technology. There is an rule-based part-of-
speech tagger and lemmatizer going back to Elaine
Uí Dhomhnaill’s Ph.D. thesis in the early 2000s
(Uí Dhomhnaill and van Genabith, 2006) [Uí Dhom-
hnaill, 2008]. Teresa Lynn produced a large depend-
tency treebank for Irish (Lynn et al., 2021) as part of
her Ph.D. work (Lynn, 2016), and has used that to train
dependency parsers that achieve very good results on
a range of domains and text types (Lynn et al., 2012]
Lynn et al., 2014] Lynn and Foster, 2016] Barry et al.,
2021). The present author has developed a standardiza-
tion tool (Scannell, 2014) that grew out of earlier work
on spelling and grammar correction, and which plays
an important role in this research.

Old and Middle Irish

Although outside of the scope of this paper, it is worth
mentioning some important work on grammatical analysis
for Old and Middle Irish, given that Early Modern
Irish texts exhibit linguistic phenomena that sur-
vive from these older varieties. In the future it might
be desirable to unify some of these efforts to produce
diachronic corpora ranging from the earliest Old Irish
texts to the modern Irish of present-day speakers.

Plain text corpora for Old and Middle Irish exist in
abundance and there are even some annotated corpora,
including the Parsed Old and Middle Irish Corpus
(Lash, 2014) and the St. Gall Priscian Glosses
(Bauer et al., 2018), the latter having been converted
into Universal Dependencies format by Adrian Doyle,
although with part-of-speech tags and morphological
features only.

Tools for lemmatization, tagging, and parsing of Old
and Middle Irish are still at an early stage of develop-
ment, although there has been significant progress in
recent years; see [Dereza, 2016] Dereza, 2019] Doyle

Parsed corpora in other languages

Finally, we would like to situate this work among oth-
ers that involve the development of treebanks, taggers,
and parsers for historical language varieties, and the in-
teresting linguistic work on diachronic syntax enabled
by these efforts (Eckhoff et al., 2020).

In addition to the work on Old and Middle Irish cited
above, we are aware of constituency or dependency
treebanks for Medieval French (Prévost and Stein,
2013), Middle and Early-modern English (Kroch,
2020), Old High German (Petrova et al., 2009), and
historical varieties of Portuguese (Galves, 2018), Ice-
lantic (Wallenberg et al., 2011), Basque (Estarrona
et al., 2020), and Russian (Berdichevskis and Eckhoff,
2020).

3. Datasets

Motivation

As noted above, our strategy for analyzing pre-standard
Irish texts has traditionally been to pass them through
the standardizer and then use tools designed for the
modern language. Tagged corpora created with this ap-
proach have been used in lexicographical projects, and
have been incorporated into the search functionality on
the corporas.ria.ie site.

Evaluations of the individual components in this
pipeline have been performed and reported in the lit-
erature. See [Uí Dhomhnaill et al., 2014] and [Scann-
ell, 2014] for the standardizer, [Uí Dhomhnaill and
van Genabith, 2006] for the lemmatizer and tagger, and
[Lynn et al., 2012] Lynn and Foster, 2016] Barry et al.,
2021] for the dependency parser. Nevertheless, no for-
mal evaluation of the effectiveness of the full pipeline
has been performed on pre-standard texts, and so we
have no objective measure of how well it is working,
and no way to decide if modifications to the process
result in significant improvements.

Our primary aim is therefore to put this research on
a more solid foundation by establishing an annotated
test corpus consisting of texts from the period 1600
to 1936, annotated according to the Universal Depen-
dencies guidelines. The resulting treebank (Scannell,
2022) is freely available for others to use in their own
experiments on tagging and parsing of pre-standard
Irish; our aim is to have it included in the 2.11 release
of the Universal Dependencies treebanks.

The Texts

With limited time for manual annotation, we decided
to keep the test corpus quite small, while at the same
time endeavoring to include texts that represent a range
of time periods and dialects.

The pre-standard texts published in the late 19th cen-
tury and early 20th century (from roughly the founding
of Conradh na Gaeilge in 1882 through the introduction
of the Official Standard in the 1940’s) are, generally
speaking, much easier to process than older texts. Even
though the orthography is still quite different from the
standardized orthography, there is much more consistency and the grammatical differences are relatively minor. We selected three texts from this period, one from each of the major dialects: Deoraidheacht by Pádraic Ó Conaire (Connacht Irish, first published in 1910), Peig by Peig Sayers (Munster Irish, first published in 1936), and Scáirt an Díithceáis, a translation of Jack London’s Call of the Wild by Niall Ó Domhnaill (Ulster Irish, first published in 1932).

We then selected three older and much more challenging texts to round out the corpus: Foras Feasa ar Éirinn by Seathrún Céitinn (1634), the 1602 translation of the Gospel of John by Uilliam Ó Domhnaill, and Cín Lae Amhlaibh, a hand-written diary kept by Amhlaoibh Ó Súilleabháin between 1827 and 1835. This diary is perhaps the most challenging text for computational processing despite being written in the 19th century, because of the informal nature of the writing and tremendous variation in spelling.

All six source texts are included in the Royal Irish Academy’s Historical Corpus of Irish (Dillon, 2017).

Annotation Guidelines

There are two existing Universal Dependencies treebanks for modern Irish that use the same annotation guidelines: the Irish Universal Dependencies Treebank (IUDT) (Lynn et al., 2021) and the TwittIrish treebanks for modern Irish that use the same annotation guidelines (Cassidy et al., 2021).

Generally speaking, we followed these guidelines very closely; the details are provided on the Universal Dependencies website. Here we will make note of a few consequences of this design choice that arose when annotating the pre-standard corpus, and a couple of ways that we diverged from the existing guidelines.

First, the modern Irish treebanks perform some gentle standardization in the lemmatization field. For example, a misspelling like neamhspleách is corrected in the lemma field to neamhspéácht, and a pre-standard or dialect spelling like thaisbeáint is lemmatized to taíspeáint. We followed this convention in the pre-standard treebank as well, but in our case it applies to a large proportion of the words in the corpus vs. the occasional misspelling or dialect spelling. We believe this is the correct design choice for the lexicographical applications we have in mind, where indexing by a standard spelling is sure to be useful. That said, this also makes the task of “lemmatization” much more difficult from a machine learning perspective, since the task now really amounts to both lemmatization and standardization, and there is no easy way for a machine learning algorithm to tease apart strictly morphological phenomena from changes that come from standardization of the lemma (e.g. when we lemmatize inneosad to inis vs. innís).

Nouns with explicit marking for the dative case are much more common in the pre-standard corpus than in modern Irish. The modern Irish treebanks only include the feature Case=Dat in the few set phrases where the noun has a distinct dative form in standard Irish, e.g.: ar leith, i gcrích, in Éirinn, os cionn, etc. We followed this convention in the pre-standard treebank, even though explicitly-marked datives are common enough that an argument could be made for annotating all nouns that appear in a dative context with Case=Dat, in much the same way that all genitives in the modern treebanks are annotated with the feature Case=Gen, even when the surface form agrees with the nominative (e.g. uisce in acmhaimní uisce).

We leave this point for future discussion with the other Irish treebank maintainers.

Some care was needed in dealing with noun genders, since some nouns have changed genders over time, and there is some variation across dialects as well. We reviewed all cases where internal evidence (usually an initial mutation) suggested that a noun might be of an unexpected gender, and determined whether these were actual variations or mere performance errors, the latter being exceedingly common in Cín Lae Amhlaibh, e.g. Do sheid an gaoth go ciuin . . . or . . . an smolach, an fúiseog, agus gac einín bin eile. Even the well-edited texts from the 20th century contain some examples like this; the first edition of Peig contains the phrase Is beag an beann a bheadh agansa . . ., where beann would normally be feminine and therefore leントed in this context (and indeed, later editions of the book “correct” this to an bheann). In cases like these, we referred to existing dictionaries as well as the wider corpus for evidence of gender variation of the given noun before deciding on the best annotation.

Tokenization was the one place where we diverged significantly from the annotation guidelines for modern Irish. The general UD guidelines allow for so-called “multiword tokens”; these are orthographic tokens that are decomposed into multiple words for the purpose of syntactic analysis (e.g. the French treebanks decompose the token du into two syntactic words, the preposition de and the determiner le). The modern Irish treebanks do not use multiword tokens at all. For the pre-standard treebank, we decided to make use of them in cases where a single token would be normally be written as two or more words in the modern orthography. For example, ar anadhbhárnsain is common in the 17th century Bible translations (usually corresponding to therefore in English translations), but would standardize to ar an abhar sin. Here we would annotate anadhbhárnsain as a multiword token. As another example, in older texts it was common to fuse the verbal particle do with the verb: dochuáidh, dorinne, etc., whereas these would be written separately in the standard orthography.

There are further subtleties to take into account when annotating these multiword tokens. In the examples above, the decomposed words all appear explicitly as part of the surface token (do + rinne, etc.). When they...
do not appear explicitly in this way, we choose not to annotate as a multiword token. For example, the standardizer converts the synthetic verb form *thóigéubh-tháoi* to *thógfaidh sibh* but this is treated as a single token in the treebank, with features **Number=Plur** and **Person=2**, the same way synthetic verbs in the modern language would be handled.

### Building the treebank

The Irish standardizer outputs word-aligned standardizations; these alignments are critical in what follows, because our goal is to build the pre-standard treebank using **cross-lingual projection** via these word alignments (Yarowsky and Ngai, 2001). Our six chosen books were run through the standardizer, and then the resulting standardized texts were annotated using a parser trained on the IUDT corpus (see §4.1 below for details), with the goal of projecting these annotations back to the original, pre-standard source. Across the six texts, 97.5% of tokens are aligned one-to-one with their standardizations, and in these cases the annotations were projected directly. Of the remaining 2.5% of tokens, the majority involve one-to-many standardizations, of the type discussed in the previous subsection (*anadhbhársain*, *dorinne*, etc.). These are trivial to annotate given our decision to treat them as multiword tokens; the annotations on the individual standardized words are simply projected back to the individual source words comprising the multiword token.

The remaining cases involve many-to-one standardizations; these require a bit more care and some manual intervention. Typical examples include:

- *ana mhaith* (standard *an-mhaith*)
- *deagh* *Ghaedheal* (standard *dea-Ghael*)
- *ró naomhtha* (standard *rónaofa*)
- *cé ‘r bh’* (standard *cérbh*)
- *dh’á ríribh* (standard *dáiríre*)
- *le n’ár* (standard *lenár*)
- *ní fhilim* (standard *nílim*)

The most common 700 of these many-to-one mappings were surveyed, and the correct annotation of the individual words was determined manually and stored in a database for the projecting parser to use. These rules include the part-of-speech tags for each token, an indication of the head of the phrase, and internal dependency relations so these can easily be incorporated into the annotation of the full sentence. In the remaining (rare) cases of many-to-one mappings, we default to assigning the part-of-speech tag *X* to each pre-standard token, and assign the root of the sentence as the head. We call this process, starting with a pre-standard source text and ending with a valid CoNLL-U file, the **projecting parser**. We applied the projecting parser to each of our six texts, shuffled the sentences, and then split into training, development, and test sets. The test sets were chosen to be balanced across the six books, with 25 sentences taken from each, resulting in a treebank containing 150 sentences and 3804 tokens. This treebank was then manually corrected, resulting in the gold-standard corpus used in our evaluations below.

### 4. Parsing Models

In this section, we will introduce the seven parsing models that we evaluated on the test set described in the previous section. All models were trained using version 1.2.1 of UDPipe (Straka and Straková, 2017) using the “swap” transition system. UDPipe also allows the incorporation of pre-trained **word2vec** word embeddings into the parsing models. We did this for each of the models below, using the skip-gram model, a window size of 10, and 50-dimensional word vectors (following the recommendations of the UDpipe maintainers). The details of the corpora that we used to train the word embeddings varied from model to model; these details are given in the subsections that follow.

#### Modern Irish parser

Our first baseline involved looking at the performance of the unmodified standard Irish parser on pre-standard texts, as a kind of “zero-shot” evaluation. For this, we trained a model using the IUDT training set distributed with version 2.9 of the Universal Dependencies treebanks. This corpus contains 95881 tokens across 4005 sentences. We incorporated pre-trained word vectors using **word2vec**, trained on a large web-crawled corpus of modern Irish containing about 127 million words. The results for this model are labeled “UD” in Table 2 below.

#### Projecting parser

This model is precisely the projecting parser described above in §3.4. In short, it involves standardizing a given input text, parsing the standardized text with the modern Irish parser, and then projecting those annotations back to the original text using the word alignments output by the standardizer. Again, most of the care is needed to handle the cases of many-to-one standardization. The results for this model are labeled “Projecting” in Table 2 below.

<table>
<thead>
<tr>
<th>Treebank</th>
<th>Sentences</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>IUDT train</td>
<td>4005</td>
<td>95881</td>
</tr>
<tr>
<td>IUDT test</td>
<td>454</td>
<td>10109</td>
</tr>
<tr>
<td>Silver train</td>
<td>11479</td>
<td>232771</td>
</tr>
<tr>
<td>Older test</td>
<td>75</td>
<td>1530</td>
</tr>
<tr>
<td>Oldest test</td>
<td>75</td>
<td>2274</td>
</tr>
</tbody>
</table>

Table 1: Summary of the treebanks used for training and testing of our parsing models.
Silver parser

Since we do not yet have a gold treebank for pre-standard Irish beyond our small test set, the idea here was to take the output of the projecting parser on the training portion of our six chosen texts, and use those trees to train a new model with no post-editing (hence the name “silver”). In total, there were 323,771 tokens across 11479 sentences in this training set. The resulting model is our first parser trained to act directly on pre-standard Irish without making use of the standardizer as part of the parsing pipeline. We combined it with word2vec embeddings trained on a 30 million word subset of the Royal Irish Academy corpus (Dillon, 2017). The results for this model are labeled “Silver” in Table 2 below.

Bilingual model

We were interested in training a single model that would give good results on both standard and pre-standard Irish. With this in mind, we simply combined the IUDT training set with the silver training data from the previous model. Similarly, we trained word2vec embeddings on the union of the training corpora used for the previous two models. The results for this model are labeled “UD+100%” in Table 2 below.

Cross-lingual word embeddings

This is a small variation on the previous model, again with the aim of getting good results on both standard and pre-standard Irish. We used the same training set, but combined the monolingual word embeddings from the first two models (for standard and pre-standard Irish, respectively) into a single embedding using Facebook’s MUSE [Lample et al., 2018]. MUSE requires “seed” translations in order to build the cross-lingual representation; in our case these were taken from the bilingual lexicon used by the Irish standardizer. The results for this model are labeled “UD+100%+MUSE” in Table 2 below.

Balanced multilingual model

Since we are able to produce virtually unlimited amounts of silver training data, we worried that perhaps the size of the silver corpus would overwhelm the high-quality annotations from the gold IUDT data. We therefore recreated the bilingual model above, but using only 25% of the silver training corpus combined with the full IUDT training corpus. The results for this model are labeled “UD+25%” in Table 2 below.

Modern parser with enhanced lexicon

The syntactic differences between pre-standard and standard Irish are minimal; most of the problems arise from differences in morphology and orthography. We therefore wondered if a modern Irish parser could achieve good results on older texts if it were augmented with a tagged lexicon that provides reasonable coverage of pre-standard Irish. For this, we simply extracted the surface form, lemma, part-of-speech tag, and features for all of the tokens in the silver training corpus and used those as the lexicon with the modern Irish parser (our first model above). In this way we hoped to transfer a good bit of the lexical knowledge embedded in the standardizer to this model without introducing noisy dependency relations.

5. Results

The experimental results are presented in Table 2. Each of the seven models from the previous section was evaluated on three separate test sets. The first test set, corresponding to the columns labeled “Standard” in the table, is the official IUDT test set distributed with version 2.9 of the Universal Dependencies treebanks (Lynn et al., 2021); we included these results to give a sense of how well the models perform on standard Irish. The second test set, labeled “Older” in the table, consists of the 75 gold-standard sentences from the three 20th century texts discussed above (Deoraidheacht, Peig, and Scáirt an Dúthchais). The third test set, labeled “Oldest” in the table, consists of the 75 gold-standard sentences from the three oldest and most challenging texts (Foras Feasa ar Éirinn, the 1602 Gospel of John, and Cín Lae Amhlaoibh).

The “POS” columns refer specifically to the Universal Dependencies (“UPOS”) part-of-speech tags, and “Feat” refers to the UD morphological features. “UAS” and “LAS” are unlabeled and labeled attachment scores, respectively. All scores were computed using the evaluation script from the CoNLL 2017 Shared Task.

The first observation is that, as expected, the IUDT parser performs poorly on the pre-standard test sets, with the worst results on the oldest texts.

Next, we see that the projecting parser achieves the best results across the board for the two pre-standard test sets, although we believe some caution is required when interpreting these results. The Irish standardizer that drives the projecting parser has been under continuous development for almost 15 years, and many improvements have been made based on analysis of its output on various corpus texts, including the six comprising our test set. We expect that similar scores would be obtained on pre-standard texts from the same periods, but verifying this would require expanding the test sets to include a more diverse set of sources, ideally including some that were not available during development of the standardizer.

The results for the Silver parser are encouraging. They are only a few percentage points worse than the projecting parser, while not making direct use of the standardizer. We do note that its performance on the standard Irish test set is significantly worse than the IUDT model, which is unsurprising since it was trained only on pre-standard texts with noisy annotations.

This defect was fixed in the UD+100% model, which achieves scores comparable to the IUDT model on
standard Irish, and comparable to the Silver parser on the two pre-standard test sets. The next row shows that the addition of the MUSE cross-lingual word embeddings gives a sizable improvement to parsing accuracy on the standard and “older” test sets, while having no significant effect on the “oldest” test set.

As expected, the UD+25% model showed a small improvement in parsing on the standard test set over the UD+100% model, but this was hardly worth it given the steep decline on the two pre-standard test sets. It is clearly important to keep as much of the silver training data as possible to obtain satisfactory performance on these older texts. The results for the UD+Lex model were similar: slight improvements over the UD and UD+100% models on the standard test set, but a large drop-off on the other two, with scores even worse than UD+25%.

6. Conclusion
In this paper, we presented a new dataset for evaluating lemmatization, part-of-speech tagging, and dependency parsing of pre-standard Irish language texts. In addition, we performed a number of experiments to establish baseline scores for these tasks. The results in Table 2 show clearly that a parser trained only on standard Irish performs poorly on pre-standard texts; this observation was the motivation behind this paper. The projecting parser gave very good results, but these may be slightly inflated given that the standardizer achieves very high performance on the six tests comprising the test set. The remaining models show that it is possible to achieve competitive results on both standard and pre-standard Irish without any gold training data, and without making use of the standardizer at all. This suggests that the most promising way forward will be to develop a large gold-standard treebank of pre-standard Irish, most likely by post-editing the output of the projecting parser. This treebank could then be combined with the IUDT training data and MUSE cross-lingual word embeddings to achieve high-quality lemmatization, tagging, and parsing on both standard and pre-standard texts with a single model.

7. Acknowledgements
I would like to acknowledge Teresa Lynn for her many years of work on the Irish treebank; without that resource, none of this research would have been possible. I am grateful to my students Sai Shreyas Bhavanasi and Jianjun Zhang at Saint Louis University for many discussions that helped me understand the mathematics behind cross-lingual word embeddings more deeply. This project originally arose out of conversations with Charlie Dillon at the Royal Irish Academy in early 2020 just before the COVID pandemic; my thanks to Charlie and the RIA for hosting me during that visit, and for inspiring this line of research.

8. Bibliographical References


<table>
<thead>
<tr>
<th>Model</th>
<th>Standard</th>
<th>Older</th>
<th>Oldest</th>
</tr>
</thead>
<tbody>
<tr>
<td>UD</td>
<td>Lem</td>
<td>POS</td>
<td>Feat</td>
</tr>
<tr>
<td></td>
<td>95.8</td>
<td>94.4</td>
<td>82.1</td>
</tr>
<tr>
<td>Projecting</td>
<td>95.0</td>
<td>94.3</td>
<td>81.3</td>
</tr>
<tr>
<td>Silver</td>
<td>90.8</td>
<td>91.0</td>
<td>76.0</td>
</tr>
<tr>
<td>UD+100%</td>
<td>94.6</td>
<td>94.8</td>
<td>83.9</td>
</tr>
<tr>
<td>“+MUSE</td>
<td>94.6</td>
<td>94.8</td>
<td>83.9</td>
</tr>
<tr>
<td>UD+25%</td>
<td>95.3</td>
<td>94.7</td>
<td>83.4</td>
</tr>
<tr>
<td>UD+Lex</td>
<td>95.9</td>
<td>94.9</td>
<td>83.6</td>
</tr>
</tbody>
</table>

Table 2: $F_1$ scores for lemmatization, tagging, and parsing for each model across the three test sets.
historical corpus. In Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects, pages 79–89.


9. Language Resource References


Lash, Elliott. (2014). The Parsed Old and Middle Irish Corpus (POMIC). Dublin Institute for Advanced Studies, 0.1.


Wallenberg, Joel C. and Ingason, Anton Karl and Sigurðsson, Einar Freyr and Rögnvaldsson, Eiríkur. (2011). IcelandicParsed Historical Corpus (IcePaHC), 0.9.