

Profiling Time Periods for Literary Works: A Computational Linguistic Study on Fiction in the 19th & the 20th Century

Teslin Roys

teslin.roys@student.uni-tuebingen.de

Connor Kirberger

connor.kirberger@student.uni-tuebingen.de

Fidan Can

fidan.can@student.uni-tuebingen.de

Abstract

This study aims to distinguish between 19th century and 20th novels based on their linguistic features, since the prevailing idea for this research is that languages evolve over time as societies evolve; accordingly, we focus our study on the literary works of the HUM19UK and Project Gutenberg corpora. To distinguish between these two periods, we aimed to extract and use the following features: the frequency of such content words as "king", "queen," and "squire", verb endings for "-ise" and "-ize", use of French words (i.e. diacritics), sentence length, use of "nor" and contractions, the size of PP and nominal structures, and the use of progressive passive. Appropriate subsets of the HUM19UK and Gutenberg datasets were collected to represent the 19th century and the 20th century respectively. A linear SVM was trained to classify samples of 19th century and 20th century novels using the aforementioned features as well as n-gram features. We present the syntactic and content features which were most predictive, i.e. characteristic, of each century.

1 Introduction

The 19th century was a distinctive era for cultural and literary development in the western world, and particularly the United Kingdom, producing some of the most influential writers in history as well as numerous works that remain highly regarded and studied even to this day. Charles Dickens, Oscar Wilde, and Emily Brontë are just a few of the notable names from this time period. This development stemmed from the rise of Queen Victoria's rule, starting from 1837 and preceded by the 1832 Reform Act that drastically changed the political sphere of the country. Under her, the British empire reformed and industrialized, transitioning to a society with newfound political opportunities for

the public and an evolving set of morals and values that reflected these.

The literature of this era, called the Victorian era, is marked by themes such as class consciousness, a byproduct of the rapid industrialization and the critiques of wealth and working conditions that accompanied this. Other important themes that commonly appeared in works were those of religion, gender, morality, and politics. Discourse around social class and daily life were common, and the social status of an individual was typically a reflection of their wealth and education. In some contrast to the eras before it, writing in the Victorian era was generally more to the audience of the masses, thus the topics tended to be more relatable to the average person. This expansion in audience was helped in part with the fruition of the first railroads, which enabled literature to be more widely accessible.

Starting around the end of the 19th century, thematic literature began to evolve again. This time period, which antithetically is not as rigidly defined with an exact starting and ending date, blends together modernism and postmodernism. As society evolved to become more urbanized with the widespread growth and development of cities, cultural perspectives shifted towards that of utopianism, and art and literature became commonly marked by themes of technology, urban life, and war.

Following the second World War, however, this also began to change. It was around this time that a new philosophy emerged: postmodernism. This came largely as a direct response to modernism, where previously idealist perspectives on the world became darkened by the harsh brutality of the war, and the longstanding traumatic effects that this had on society. Influenced also by the rapid growth of consumerism, information technologies, and the media as a conglomerated entity, writers from this time period tended to use a variety of themes and

topics in their writing, but some notable ones include paranoia, technoculture, and hyperreality.

As they are products of the time period and world in which they were created, literature can be distinguished from one another based on linguistic, historical, and other contextual features that set each apart. In this way, we aimed to discern those features that best identified 19th century and 20th century writings. The literature review, methodology, results, discussion, and conclusion of our research are presented as follows.

2 Literature review

Rubino et al. (2016) suggest in their article that time periods classification has been rarely studied until now when compared to other classification tasks.

To start with for the scope of our study, the 20th century differs from the Victorian era fiction with respect to how it reflects life: Yekini et al. (2017) argue that the 20th century novels focus on the intricate thoughts of human beings, and unlike Victorian period novels (the 19th century), there were no pure villains or pure good described. They rather include the characters with a mixture of the good and the bad.

In addition, Langland (1992) claims that there is a distinction between the classes in societies in the 19th century: although men and women can marry, classes do not, and servants of this era were not allowed to marry somebody from the upper classes. She (1992) also adds that there is “the active deployment of class power” (p.291). In line with the idea of socially structured classes formed in the 19th century and language representation of the people from the lower classes such as servants, Hodson (2016) shows in his research that out of 100, 68 novels in the 19th century included non-standard language used by the servants. He also points out in his paper that contractions can be used as a linguistic feature to show the social status of the characters and exemplifies it through Bronte’s novel *Wuthering Heights* where a servant’s speech included contractions and some variations in pronunciation.

However, Fries (1940) claims that in the 20th century, especially after the WWII, not much attention was given to descriptive grammar in Britain, and it was noticed in this period that there was not necessarily a direct connection between grammar and good use of English, which was emphasised in

the 19th century. (as cited in Hudson and Walmsley, 2005)

Apart from grammaticality and the importance given on the standard English in the 19th century, another linguistic characteristic is the use of loan words from French. Bailey (1996) states that English was the language of the era in the 19th century (as cited in Schäfer, 2013) as Schäfer (2013) claims that this is due to the power of England in industrialization and it becoming the leading colonizer. However, according to Görlach (1999) when the English equivalents of some French words were shallow or simple, using English was ridiculed. Furthermore, he also puts forward that in Britain, French was still widely taught despite the political criticism against the French Revolution. One of the interesting points made is that when the French loanwords were used often, it was ridiculed and appeared pretentious; however, he (1999) also claims that “French loans retained their prestige value” and the number of them rose later on in the 19th century as well (p.110). In a parallel view, (Schäfer, 2013) states that French language was not useful for the lower class in this period.

Moreover, in the article *Changing Grammar and New Words in the Nineteenth Century*, it was seen that progressive passive, which was a new structure, started to rise in the 19th century, and became dominant only towards the end; nonetheless, resistance to progressive passive was ongoing across the period on grounds of the debate on the correct use of English (n.d.).

According to *Twentieth century English – an overview*, during the 20th century, the strict understanding of using standard English was relaxed, and variety in accent and grammar was tolerated alongside globalization. In other words, the adherence to standard English was less rigid. Globalization also led to an increased use of Americanized words due to mass media and marketing (n.d.).

Finally, to get an understanding of the 19th c. and the 20th c. distinction from another aspect in terms of their linguistic features – although not in a literary context – Rubino et al. (2016) found in their study that the nominal structures of the 20th century scientific abstracts were more complex (p.758).

(i.e. consisting of compounds with a at least two nouns, e.g. DT NN NN such as the heat transfer)” while in the 19th century, nominal structures were shorter consisting of POS sequences with one

noun, e.g. followed by a comma (DT NN , such as the heat,) and prepositional phrases (e.g. IN DT NN such as on the eye) were discriminative for the 19th c. abstracts.

3 Methodology

3.1 Dataset

To compare the 19th and the 20th century fiction, two data sets were chosen: HUM19UK provided by [University of Huddersfield \(2019\)](#) and Project Gutenberg dataset provided by [Lahiri \(2014\)](#). While the former is a 19th century British fiction corpus which includes one hundred novels from one hundred authors over the span of one hundred years, the latter corpus is composed of a mix of 19th and 20th century fiction and nonfiction. The authors in this corpus are American, Canadian, British, and more. Due to the fact that the Project Gutenberg data set also includes texts from the 19th century, a number of 20th century novels were extracted by hand. Twelve 20th century novels (plays or poetry collections were discarded) were extracted from the Gutenberg corpora, primarily British authors from the first quarter of the 20th century, and only one novel per author.

Due to the nature of the original dataset, copyrights on more recent texts prevented their inclusion, so the availability of works from the latter three quarters of the 20th century was low. Preference was made for British authors in the Gutenberg-derived dataset, but in the interest of including additional female authors fitting our time period criteria, as well as particularly well-known works, some non-British English-speaking authors were also included. Unfortunately even with this loosening of criteria, still only four female authors of the specified time were available in the source Gutenberg dataset. The genres vary and include drama, adventure, science fiction, etcetera. For the Gutenberg-derived set, in general preference was also given to authors who wrote a greater proportion of their works in the 20th century than the 19th.

From the HUM19UK dataset, we selected twelve novels as well for a fair comparison, primarily from the early 19th century. For this selection, it was simpler to achieve gender parity - of the twelve included authors in this selection, six authors were male, six were female. Exactly matching the proportions of the Gutenberg-derived set, even if desirable, was not possible without includ-

ing a wider time range (due to insufficient male authors within the time period).

Finally, with these constraints in mind, an effort was made to avoid extremely drastic differences in novel lengths (e.g. short stories versus full length novels). To this end, for the 19th century dataset, selected novels are between 54,288 and 166,636 words in size. For the 20th century data selection, the novels are between 46,282 and 160,450 words in size. Because the HUM19UK dataset contained XML metadata and the Gutenberg dataset did not, this was removed from the 19th century data. Title, table of contents, and author were likewise removed from the 20th century dataset. Overall, the 18th century novel data spans the years 1804-1826 and the 19th century novel data spans 1901-1922.

To provide an adequate number of distinct samples to our model, we broke each novel into approximately 20 equal-sized chunks, resulting in over 250 instances per class. Initially, we attempted to break each novel into chapters, however inconsistent metadata and format of the novels meant the size of each sample was extremely varied. To minimize slicing into sentence boundaries, we applied the Punkt sentence tokenizer implemented in the NLTK framework, and divided each text by sentence rather than token. This means the samples are not all the same size, but they vary in the worst case according to the greatest difference in average sentence lengths between any two samples.

3.2 Pattern-based features

To extract linguistic features from these two data sets, different kinds of methodologies were made use of. The features are as follows:

The first feature is the use of some content words such as "king", "queen", and "squire", as the rationale behind this is our expectation that the Victorian era fiction will have more words related to social classes when compared to the 20th century, which hosts modernism and postmodernism. By making use of the regular expressions for the specified words, which are "king", "queen", "squire", "castle", "sir", "ma'am", "hither", "England", "London", "scrooge", "pounds", "shall", "lord", "grey", "dollar", "new/syork", "gray", "America", "city", "parlor", "Boston", and "Washington", and making the regular expression case insensitive, we checked whether these words occur and/or how often they occur in alignment with the time periods that the texts were written in. The regular expres-

sions are as following:

```
(king|queen|squire|castle|sir|  
ma'am|hither|england|london|  
scrooge|pounds|shall|lord|grey)  
(dollar|new\sYork|gray|america|  
city|parlor|boston|washington)
```

This second regular expression is included as well since there are some American authors in the Project Gutenberg data set.

The second feature is the use of "-ize" and "-ise" to differentiate between the British authors and the American authors and to see whether more Americanized words (i.e. occurrences of "-ize") are seen in the later periods alongside globalization. This feature can be one of the indicators to profile the different eras, or at least to see how the language has evolved in time due to societal changes such as globalization. Via a regular expression, we aimed to investigate how predictive American-style verbs are of 20th century text. The case insensitive regular expression for this feature is:

```
(\w*(i|y)z(e|i|a)\b)
```

The third linguistic feature is the use of French words, and instead of looking for any possible French words, we narrowed down the scope by having the purpose of looking for the words including diacritics and some characters that are special to French that do not exist in English. These special characters are "ç", "é", "â", "ê", "î", "ô", "û", "à", "è", "ï", "ò", "ù", "ë", "ï", and "ü". The reason for having this feature is to check whether it would be a distinguishing linguistic characteristic to differentiate the 19th century from the 20th century as it is expected that due to the highlighting differences in social classes and elitism, the former period would include more French words than the 20th century. As a result, the case insensitive regular expression formulated for this linguistic feature is:

```
(ç|é|â|ê|î|ô|û|  
à|è|ï|ò|ù|ë|ï|ü)
```

The fourth feature extracted is the sentence length to have an understanding of whether the way of communication in the later periods was abbreviated, and how the language in the literary works was influenced by the changing time with regard to syntactic level. For this measure, we once again leveraged the Punkt sentence tokenizer, and took the average sentence length in characters in each sample. We use characters rather than tokens because this also includes word length implicitly as a factor.

The fifth one is the use of the functional word "nor" to compare the two different time periods and distinguish them via looking at the occurrence and/or the frequency of this functional word since we expect a more frequent use of this word in the earlier periods. For this task, we made use of regular expressions again to exclude some exceptions such as "nor" occurring in different positions in a word (e.g. donor). Thus, specifying the word boundary seemed a plausible solution for it. Furthermore, as it is possible to use "nor" in different positions in a text, one of the essential points is that the regular expression should be case insensitive too, and it is as follows:

```
\bnor\b
```

The sixth linguistic feature is the use of contractions, and specifically we looked for occurrences of the "n't" sequence. Due to the fact that contractions can give us another insight into language use with respect to what contractions can reflect (e.g. casual use/non-strict use and social status) and insight into how changing times have an impact on the language structures, making analysis on contractions was one of the goals of this study as well.

The seventh feature we included is the size of the PP and nominal structures to see if the result of the [Rubino et al. \(2016\)](#) study would apply to the literary context we have or whether the distinction between POS sequences would differ from genre to genre.

Lastly, use of progressive passive is another feature we wanted to look at since we expect that 20th century fiction should include progressive passive more than that of the 19th century in the light of the background reading. For this reason, a regular expression was used. The cases of reduction such as "being implemented" without copula was excluded. We only looked for the full forms for which the rule is "be + being + V3". However, there are critical points to consider while forming the regular expression. To exemplify, for the V3 form, both regular and irregular verbs are used to form a passive structure in English. Therefore, we included the most common 50 irregular verbs in our regular expression. In addition, contractions can also be used with the verb be (e.g. "it's being improved"), and it is important that the regular expression is case insensitive. As a result of bearing these points in mind, the following regular expression comes out:

```
((\s*(was|were|am|is|are|'s|'re|m)
\sbeing\s[a-zA-Z](ed|ied|d)\s)|\s
(was|were|am|is|are|'s|'re|m)
\sbeing\s(said|made|gone|taken|
come|seen|known|got|gotten|given|
found|thought|told|become|shown|
left|felt|put|brought|begun|kept|
held|written|stood|heard|let|meant|
set|met|run|paid|sat|spoken|lain|
led|read|grown|lost|fallen|sent|
built|understood|drawn|broken|
spent|cut|risen|driven|bought|worn|
chosen)\s*)
```

3.3 Model

In order to learn what features may distinguish the two centuries under investigation, it was necessary to choose a model where the feature weights or coefficients impacting the prediction were relatively transparent. Two of the most promising models under that criteria were either a random decision forest or a support vector machine. Although random forests may have some advantages in a multi-classification setting, because in our case this was merely a linear classification, the coefficients of an SVM are particularly interpretable (with negative coefficients pointing toward one class and positive toward the other).

In addition to evaluating the linguistic features described above, we also trained a linear SVM on the same data, but using unigrams and bigrams as features. Our expectation was that these n-gram features may predictively outperform the handpicked features, but may be more difficult to draw human generalizations from.

4 Results & Discussion

4.1 General

To train and evaluate the SVM model, we used 70 percent of the data for training and reserved 30 percent for testing. We did not use stopwords, because our analysis is specifically interested in changes in use of high frequency words between eras. We also experimented with training and testing the model on the full-length novels (without chunking each document into many), with the same test/training split. In general, classification accuracy of the data in that form was very poor.

Class	Precision	Recall	F1 Score
19th Century	0.6000	1.0000	0.7500
20th Century	1.0000	0.6000	0.7500
Overall Accuracy	-	-	0.7500

Table 1: N-gram-based classification results with TF-IDF weighting, on full-length novel data.

Class	Precision	Recall	F1 Score
19th Century	0.9241	0.9733	0.9481
20th Century	0.9726	0.9221	0.9467
Overall Accuracy	-	-	0.9474

Table 2: N-gram-based classification results without TF-IDF weighting, on chunked data.

4.2 N-gram features

First, we evaluated classification performance and examined the most predictive features under different training conditions. In Tables 1 and 3, we can see that training and testing on the full-length (unsubdivided) novel data results in a drastic reduction in accuracy from employing the same methodology on the chunked (subdivided) texts. In general, with so few samples per class, even if the accuracy of the model were more comparable under both training scenarios, we would also expect the results on new, unknown data to be less stable.

The choice to employ TF-IDF weights on the n-gram features or not was one of the most significant in terms of results. In terms of accuracy, precision, and recall, it appears that TF-IDF weighting does improve the quality of the model’s results by several percent.

However, this does not tell the whole story. In this setting, we are also interested in discovering interpretable, generalizable features – classification performance is relevant but not the only criteria in evaluating the model. In Figure 1, we see the top 20 most predictive n-grams of each class include a large number of character or place names. We suggest this is because if (for example) a character

Class	Precision	Recall	F1 Score
19 th Century	1.0000	0.9867	0.9933
20 th Century	0.9872	1.0000	0.9935
Overall Accuracy	-	-	0.9934

Table 3: N-gram-based classification results with TF-IDF weighting, on chunked data.

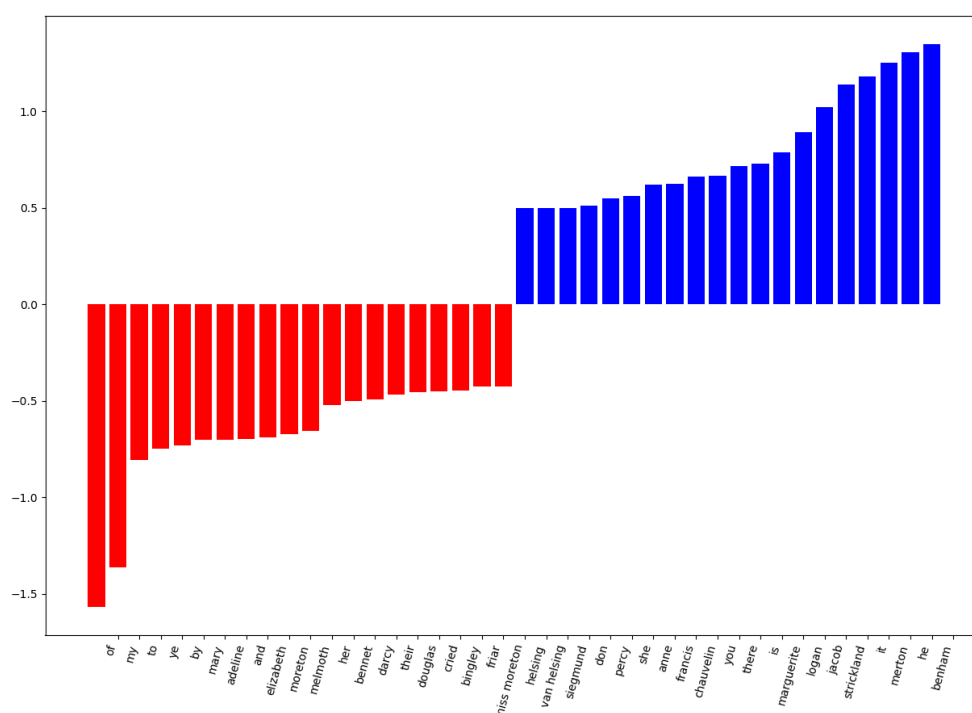


Figure 1: Top 20 most predictive n-gram features, with TF-IDF weighting, on chunked data. The y-axis represents the SVM feature weight – negative values are predictive of 19th century text while positive values predict 20th century text.

name uniquely appears in only one of the two novel time periods (which is more likely because of the small number of novels in the dataset overall), it becomes disproportionately predictive than it would necessarily be on unseen data.

Of course, the popularity of names may indeed also change over a century – however, even if this were true, the size of the available dataset wholly precludes us from making any such inferences. This highlights a problem of the dataset: though we have sampled the novels into smaller parts, these chunks are not wholly independent documents and will share at least some commonalities with other chunks from the same source text. In Figure 3, however, we can see that names will still occur as highly predictive features even without applying the chunking strategy.

In general, TF-IDF weights in favor of n-grams which may occur relatively rarely but relatively uniquely in a document. If we are trying to identify generalizable features of 19th or 20th century texts, we may prefer to see of the n-grams which occur most frequently, which are most predictive. From Figure 2, without this weighting, we can see far

fewer character or place names occur in the top 20 most predictive features.

Apart from names, just from these n-gram results we can see that a wide variety of syntactic categories appear to be among the most predictive of the eras: verbs, prepositions, pronouns, and even some adjectives. A few themes can be observed, however. In Figure 2, we see that the conjunctives (but, and, or) rank as highly predictive of 19th century texts but none are predictive of 20th century texts. We can also see that common verbs used with human subjects appear more predictive of 20th century texts – e.g., think, see, put, said – and none of these type of verbs appear predictive of 19th century texts.

TF-IDF weighting may emphasize character names as features, but it also brings to the fore some less common word usage which nonetheless distinguishes the two time periods. In Figure 1, we see that use of the English second person plural "ye" versus "you" diverges greatly between the two eras. In everyday modern English, use of "ye" is extremely rare, but at least in a literary context the pronoun form still had a regular, if vestigial,

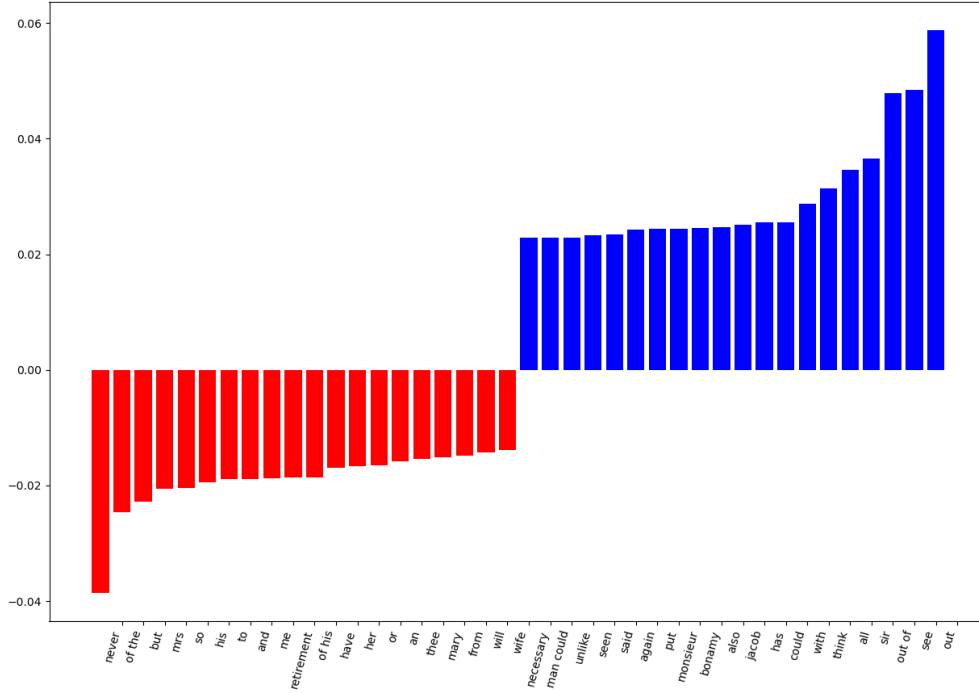


Figure 2: Top 20 most predictive n-gram features, without TF-IDF weighting, on chunked data. The y-axis represents the SVM feature weight – negative values are predictive of 19th century text while positive values predict 20th century text.

presence in the 19th century.

4.3 Pattern-based features

A separate linear SVM was trained with the same input, but different featurization. We tested each of our proposed linguistic features alone to identify which were useful heuristics for classification and which were unproductive.

Due to time constraints and an error with our pre-processing/POS tagging step, our proposed "size of PP/nominal structure" feature was not implemented for this analysis. For the rest of the proposed features, we can see in Table 4 that most of them were useful heuristics for classification. Only the non-English characters and content words features performed worse than a coin toss.

Notably, average sentence length alone as a feature performed within spitting distance of the ngram-based classification. The presence of "nor" and the use of contractions also appear to be strongly indicative of a text's era.

Taking into account all features with greater than 50 percent accuracy alone, we trained a combined model, the results of which can be seen in Table

Feature	Accuracy	Weight
Avg. sentence length	0.9210	-0.01677
Contractions	0.8355	0.1306
Nor	0.7105	-0.2101
"ize" ending	0.5394	0.07602
Progressive passive	0.5131	0.7933
Non-English characters	0.4934	-
Content words	0.4736	-

Table 4: Overview of feature performance. Accuracy scores using single features for classification, for comparison as heuristics. Weights are the SVM coefficients of the features in the combined model, negative being predictive of 19th century text and positive predicting 20th century text.

Class	Precision	Recall	F1 Score
19 th Century	0.9221	0.9467	0.9342
20 th Century	0.9467	0.9221	0.9342
Overall Accuracy	-	-	0.9342

Table 5: Classification results with sentence length + contractions + "ize" endings + "nor" + progressive passive as features, on chunked data.

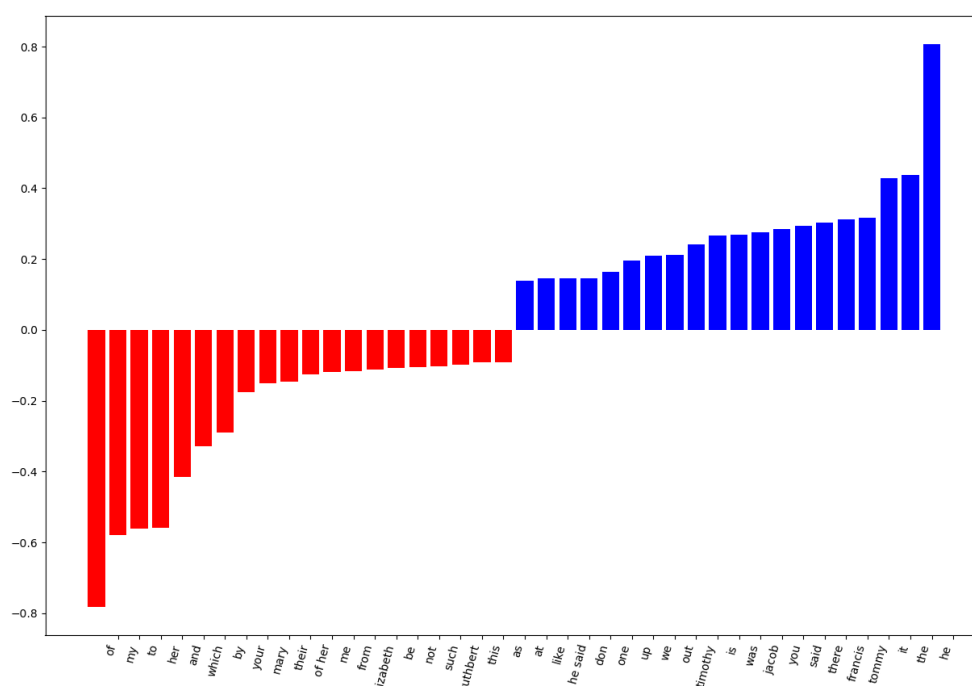


Figure 3: Top 20 most predictive n-gram features, with TF-IDF weighting, on full-length novel data. The y-axis represents the SVM feature weight – negative values are predictive of 19th century text while positive values predict 20th century text.

5. Here we can see that although many of these features may have been less effective alone, including the additional features improved accuracy by around 1.3 percent over sentence length alone.

Interestingly, although progressive passive performs poorly individually, it contributes nearly half of the accuracy gains in the combined model (92.76 percent without progressive passive, versus 93.42 percent with it). From this we might surmise that progressive passive constructions occur rarely in the data, but are strongly predictive. The high weight (coefficient) value as reported in Table 4 corroborates this, and further shows that these constructions are very indicative of 20th century texts. In this regard, only the contraction feature and the "nor" feature come close in terms of strength of association with a time period ("nor" being predictive of the 19th century and contractions predicting 20th century text).

5 Conclusion

Our research found that between the two approaches tested – using handpicked linguistic features, or n-gram features – n-gram features proved

to be more effective, particularly when weighting by TF-IDF. The improvement with TF-IDF weighting may be less than it appears, because in a small dataset it can emphasize features that are more indicative of particular authors or texts than eras. Nonetheless, even when weighing n-grams using TF-IDF, we found that for the 19th century texts, conjunctions (and, or, etc.) were some of the most predictive features. However, many of the hand-picked features we used also surprisingly showed highly accurate results, suggesting that n-grams may not be the only effective feature to consider in this kind of study. Features such as sentence length were broadly an effective discriminator, while the usage of progressive passive, "nor", and contractions all proved to be even stronger indicators of the writings' respective eras in the instances where they occurred. In contrast, the only two tested linguistic features that showed to be completely ineffective heuristics for classification were the non-English characters and the content word patterns. Iteration on these weaker features may improve results – it is possible a more restricted subset of diacritics, for example, might be more predictive.

In our investigation, we employed the hand-picked linguistic features and the n-gram features in separate models to facilitate comparison. Both approaches led to good results, so in future, it may be productive to combine all of the features in a single model. More linguistic features, including but not limited to PP/nominal phrase size, may also shed light on differences in the eras and improve model accuracy. Most of all, however, the size of our dataset introduced limitations to our approach - we expect applying similar methods to a more robust input would be extremely fruitful.

References

- Richard W. Bailey. 1996. *Nineteenth-Century English*. The University of Michigan Press, Michigan, USA.
- Changing Grammar and New Words in the Nineteenth Century . n.d. [Changing grammar and new words in the nineteenth century](#). Accessed March 29, 2021.
- Charles Carpenter Fries. 1940. *National Council of Teachers of English: English Monograph. American English Grammar: The Grammatical Structure of Present-day American English with Especial Reference to Social Differences or Class Dialects*. D. Appleton-Century Company, New York, USA.
- Manfred Görlach. 1999. *English in Nineteenth-Century England: An Introduction*. Cambridge University Press, Cambridge, United Kingdom.
- Jane Hodson. 2016. [Talking like a servant: What nineteenth century novels can tell us about the social history of the language](#). *Journal of Historical Sociolinguistics*, 2(1):27–46.
- University College Roosevelt University of Huddersfield, Utrecht University. 2019. [HUM19UK, Version 1](#). Middelburg.
- Richard Hudson and John Walmsley. 2005. [The english patient: English grammar and teaching in the twentieth century](#). *Journal of Linguistics*, 43(3):593–622.
- Shibamouli Lahiri. 2014. [Complexity of Word Collocation Networks: A Preliminary Structural Analysis](#). In *Proceedings of the Student Research Workshop at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 96–105, Gothenburg, Sweden. Association for Computational Linguistics.
- Elizabeth Langland. 1992. [Nobody’s angels: Domestic ideology and middle-class women in the victorian novel](#). *Pmla*, 107(2):290–304.
- Raphael Rubino, Stefania Degaetano-Ortlieb, Elke Teich, and Joseph van Genabith. 2016. Modeling diachronic change in scientific writing with information density. *COLING 2016 - 26th International Conference on Computational Linguistics, Proceedings of COLING 2016: Technical Papers*, pages 750–761.
- Sebastian Schäfer. 2013. [A comparison of french loanwords in 19th-century english dictionaries](#).
- Twentieth century English – an overview . n.d. [Twentieth century english – an overview](#). Accessed March 29, 2021.
- Ibrahim Yekini, Franck MOUSTAPHA, and Hergie A. Seguedeme. 2017. [Features of the twentieth-century novel](#). *International Journal on Studies in English Language and Literature (IJSELL)*, 5(3):44–53.