A Computational Analysis of the Voices of Shakespeare's Characters

Liviu P. Dinu, Ana Sabina Uban
University of Bucharest, Romania
ldinu@fmi.unibuc.ro, auban@fmi.unibuc.ro

Abstract

In this paper we propose a study of a relatively novel problem in authorship attribution research: that of classifying the stylome of characters in a literary work. We choose as a case study the plays of William Shakespeare, presumably the most renowned and respected dramatist in the history of literature. Previous research in the field of authorship attribution has shown that the writing style of an author can be characterized and distinguished from that of other authors automatically. The question we propose to answer is a related but different one: can the styles of different characters be distinguished? We aim to verify in this way if an author managed to create believable characters with individual styles, and focus on Shakespeare's iconic characters. We present our experiments using various features and models, including an SVM and a neural network, show that characters in Shakespeare's plays can be classified with up to 50% accuracy.

1 Introduction

The problem of authorship identification is based on the assumption that there exist stylistic features that can help distinguish the real author of a text from any other theoretical author, and that these can be computationally measured and exploited in order to automatically identify the true author of a text. Automated authorship attribution has a long and rich history (starting from the early 20th century (Mendenhall, 1901)) and has since then been extensively studied and elaborated upon.

One of the most influential studies in authorship attribution is the study of (Mosteller and Wallace, 1963) on the Federalist Papers, in which the authors try to determine the real author of a few of these papers which have disputed paternity. In this work, they both introduce a standard dataset and propose an effective method for distinguishing between the author's styles, based on function words frequencies, that is still relevant and used to this day. Many types of features have been proposed and successfully used in subsequent studies to determine the author of a text. These types of features generally contrast with the content words commonly used in text categorization by topic, and are said to be used unconsciously and harder to control by the author. Such features are, for example, function words (Mosteller and Wallace, 1963; Dinu et al., 2012), grammatical structures (Baayen et al., 1996), part-of-speech n-grams (Koppel and Schler, 2003), lexical richness (Tweedie and Baayen, 1998), or even the more general feature of character n-grams (Kešelj et al., 2003; Dinu et al., 2008). Recent studies focusing on stylistic variation within the writings of a single author combine traditional function word features with stylistic markers such as lexical richness and readability, as well as topic modelling, to compare the importance of the the stylome and the topics discussed in in the evolution of an author's writing (Dinu et al., 2017; Dinu and Uban, 2018).

A related problem that has been approached much less in computational linguistics and even in digital humanities scientific literature is that of distinguishing between the writing styles of fictional people, namely literary characters. This problem may be interesting to study from the point of view of analyzing whether an author managed to create characters that are believable as separate people with individual styles, especially since style is a feature of speech that is hard to consciously control. Shakespeare, as arguably the most renowned dramatist in the history of literature, is the ideal case study for understanding whether it is possible to create characters that are as individualized as humans are.

One of the first authors to study literary characters stylistically is John Burrows, who (Bur-

Proceedings of Recent Advances in Natural Language Processing, pages 295–300
Varna, Sep 4–6, 2023
https://doi.org/10.26615/978-954-452-092-2_033

rows, 1987) shows that Jane Austen's characters show strong individual styles, then later Burrows and Craig (2012) look at a corpus of seventeenth-century plays and tries to cluster them by character and by playwright. Another recent study (van Dalen-Oskam, 2014) analyzes the works of two epistolary novels authors, who are known to have written their books together, and tries to distinguish automatically between passages written by each author, and between styles of each character in the novel. Dinu and Uban (2017) propose an experiment on classifying the characters in the epistolary novel *Les Liaisons Dangereuses*, showing that the characters can be automatically distinguished stylistically even using simple models and features. Muzny et al. (2017) propose a metric for characterizing spoken dialogue in the novel, which they call "dialogism", and Vishnubhotla et al. (2019) publish a study reporting automatic measures of dialogism in plays from the nineteenth and twentieth centuries by automatically classifying their characters.

In this paper we take a look at one of the most interesting authors in literary history: William Shakespeare. Shakespeare is seen by scholars and readers alike as one of the greatest dramatists in the history of literature. His characters are iconic, with strong well defined personalities. The question we propose to answer in this study is whether a computational analysis would lead to the same conclusion – did Shakespeare manage to write distinct characters with unique speaking styles, and can we measure that? Moreover, are the features that distinguish characters the same as the features that distinguish between different authors?

Shakespeare's characters have also been the subject of a few previous studies, such as Nalisnick and Baird (2013), where the authors try to map the relationships between characters. Culpeper (2009) study keyness, and use Shakespeare's *Romeo and Juliet* as a case study. In Vogel and Lynch (2008), the authors investigate the interesting problem of strength of characterization of a character, using plays of four authors, including Shakespeare. They use text similarity methods to measure how similar a character's utterances are to the lines of the other characters in the same play and in other plays, proposing that stronger characters are most self-similar compared to other characters and plays. We are also interested in how individualized and realistic the characters are in their construction, but we assume that the strength of a character relies in how belivable it is as a unique person, and that this can be measured by the ability to distinguish characters the same as we do humans, from the perspective of their writing style.

Character	**Nr lines**
King Lear	190
Timon	220
Cleopatra	180
Duke Vicentio	210
King Henry V	200
Hamlet	370
Iago	280
Mark Anthony	220
Othello	240
Brutus	190

Table 1: Number of lines per character for top 10 characters

2 Data and Methodology

We constructed our set of labeled texts by first splitting each of Shakespeare's plays into individual lines, labeled with the characters that speak them, and excluding characters with less than 500 lines, and were left with a total of 50 characters. Since it can be difficult to extract meaningful information from the short individual lines, we further concatenated them in groups of 10 lines (spoken by the same character) and used the resulted texts as our data points.

We artificially balanced the number of datapoints pertaining to each class during training, using oversampling. Table 1 includes the number of lines per character before rebalancing.

3 Classification Experiments

We formulated the problem as a supervised learning problem, and trained several models using various features to try and understand how well a machine learning model can predict a character based on its utterances within a play, and what are the features that help shape characters the most.

We start by tokenizing the texts in our dataset and encoding them using a bag-of-words representation, which we further use to extract features for our classifiers. We perform different kinds of feature selection in order to then compare their performance and conclude on which are the most

Character	Precision
King Lear	10%
Timon	68%
Cleopatra	22%
Duke Vicentio	66%
King Henry V	20%
Hamlet	40%
Iago	50%
Mark Anthony	59%
Othello	62%
Brutus	31%

Table 2: Precision for top 10 characters

Feature Set	Accuracy
SVM with all words	30%
SVM with K-best (100)	13%
SVM with content words	30%
SVM with function words	6%
SVM with character n-grams (2-10)	18%
MLP with all words	50%

Table 3: Overall accuracy for each feature set

helpful features for predicting characters. The various features extracted from text are:

All words. We first experiment with using all the words in the text as features, encoded as bag-of-words. We obtain a vocabulary of 13,559 words.

Function words. Function words have been traditionally successfully used as features for authorship attribution, and are considered to be the aspects of the text that can encode a writer's style. In some of our experiments, we try to limit our features to only function words, in order to understand whether these are as useful in distinguishing between characters as they are for distinguishing between different authors.

Content words. In a separate experiment, we try to limit our features to only content words, ignoring function words. In this way, we hope to understand how important the content or topic of the text matters for distinguishing a character. We represent a list of content words using a bag-of-words model, but each word is represented by its *tf-idf* score instead of its frequency.

K-best. We attempt to use statistical methods to extract features that contribute most to separating between our classes. We use chi^2 feature selection to limit our vocabulary to the k-best features, then use only these words as features in classification.

Character n-grams. We finally experiment with character n-grams instead of words. These are a more versatile kind of feature, able to capture sub-word and multi-word content as well as individual words. We consider all character n-grams from 2-grams to 10-grams and encode them with a bag-of-words representation.

We experimented with different classifiers:

SVM. SVMs have shown to be successful in authorship attribution, since the features are usually predictive enough in this task without the need for an overly-complex model.

Multi-layer perceptron (MLP). We use a simple feed-forward neural network (multi-layer perceptron) that takes as input our features encoded as bag-of-words, passes it through one hidden layer of 1000 units, and finally predicts the most probable class using Softmax on the final layer. The vocabulary size is approximately 13K words, equal to the number of input units.

Classification accuracy was measured for each character separately, in a series of experiments where the model was trained on 80% of the texts, and tested on the remaining 20%. The overall accuracy was obtained by averaging the per-character accuracy scores.

4 Results and Analysis

The overall accuracy for each of the experiments is shown in Table 3. The results show that we were able to distinguish between characters with an accuracy superior to a random guess (which would yield an average accuracy of 2%, given there are a total of 50 classes, assuming balanced a class distribution). Precision of classification per character for the top 10 characters is shown in Table 2. The most successful feature were the content words, by far outperforming function words, which are usually successful in authorship problems. This shows that even though characters are indeed distinguishable, it may not be their style that differentiates them, at least not in the same way as it does for authors.

We take a closer look at the landscape of Shakespeare's characters as represented by our model, by reducing our bag-of-words representation to two dimensions using principal component analysis. Figure 1 illustrates this, showing that, even in this lower dimensional space, lines of the same character cluster together for some of the characters. Furthermore, it is interesting to see which characters are more similar by looking at their relative

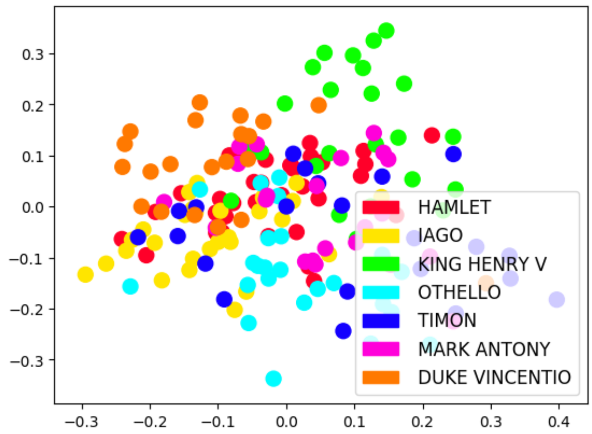

Figure 1: 2D view of character's lines (most frequent 7 characters)

Feature Set	Accuracy
SVM with all words	30%
SVM with K-best (100)	20%
SVM with content words	39%
SVM with stopwords	8%
SVM with character n-grams (2-10)	24%
Neural network with all words	20%

Table 4: Overall accuracy for each feature set for classifying plays

positions in this space.

Our results suggest lines spoken by different characters can be distinguished, especially through the content words used in them. A question raised by this is whether we are truly capturing features specific to the characters, or predicting something else, such as the play they belong to. To tackle this problem, we perform a second experiment where we try to predict the play a book belongs to, using the same models, features and experimental settings as in our character classification experiment.

The results for classifying texts by play are shown in Table 4. There are 32 plays in our dataset (32 classes), so the expected accuracy for a random classifier in the case of plays is around 3%. We can then conclude that results are comparable between the first and second classification experiments. Useful features tend to be the same as for the previous experiments as well. Content words perform best, and removing function words even adds an improvement to the results in the case of plays.

We also replicate the visualization experiment, plotting in 2 dimensions lines belonging to top 10 (most prolific) characters in the top 5 plays (longest), shown in Figure 2. Here too the distinction between lines in different plays is visible even in 2D, though less apparent than in the case of characters, which suggest characters may be more separable than plays.

The results of the classification experiments do suggest that identifying the play it belongs to is a factor in determining which character utters a line. Nevertheless, the classifier can still distinguish between characters of the same play, so other factors may contribute as well. We further try to understand how easy it is to classify between the characters to belonging the same play. Only 4 of the 32 plays have more than 2 characters in our class set: *The Tragedy of Othello, the Moor of Venice*, *The First Part of Henry the Fourth The Tragedy of Antony and Cleopatra* and *The History of Troilus and Cressida*. For each of the mentioned plays, we perform an experiment to classify between its characters, using the setting that performed best at both character and play classification: an SVM with content words as features. We average the accuracy per character for each play, then average the obtained accuracy per play, and get an average accuracy of 58.5% (almost double compared to the 30% accuracy that would be obtained by a random choice classifier). Table 6 shows the results per play, which seem to confirm that characters can be distinguished within plays as well.

Results also show that overall, content seems to be more predictive of the character, and that function words don't seem to capture a character's style in the same way they do an author's, in the case of Shakespeare. Nevertheless, the accuracy above chance obtained with function word features show they are not entirely unhelpful, confirming previous results in literary character classification (Dinu and Uban, 2017).

Finally, we perform a last experiment where we select only the 4 plays with more than 3 prolific characters and group them together into a set of 12 total characters that we try to classify. Looking at the errors the algorithm makes, whether or not it tends to mistake characters with other characters of the same play, should help us understand to what degree it learns to classify characters versus plays. Table 5 shows for each of the 12 characters, how many datapoints were classified correctly, how many were misclassified to a character in the same play, and how many were predicted to belong to a character in a different play.

Character	Same character	Diff character, same play	Diff character, diff play
Iago	20	6	2
Othello	18	2	4
Desdemona	3	5	4
Marc Antony	12	3	7
Cleopatra	5	6	7
Octavius Caesar	1	1	9
Falstaff	8	2	6
Prince Henry	3	5	7
Hotspur	5	4	5
Troilus	6	0	8
Ulysses	5	1	7
Pandarus	1	0	10

Table 5: Correct and mistaken classifications

Play	Accuracy
Othello	64%
Henry IV	62%
Antony and Cleopatra	51%
Trolius and Cressida	57%

Table 6: Average accuracies for character classification within plays

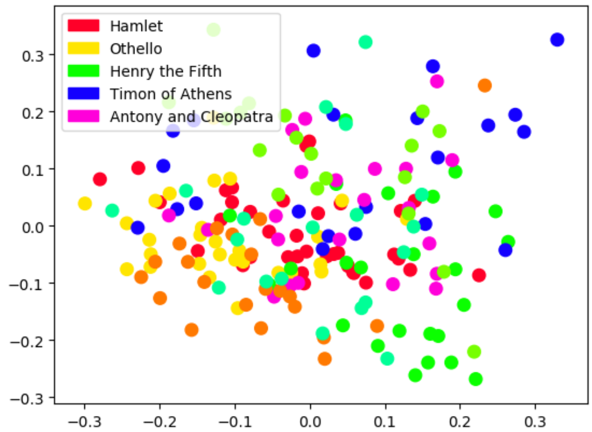

Figure 2: 2D view of character's lines grouped by play (most frequent 5 plays)

5 Conclusions and Future Directions

Our experiments have shown that it is possible to automatically distinguish between the characters of Shakespeare's plays using a machine learning model. The texts were most successfully classified using content words, not function words, that are known to capture the stylistic dimension of a text. This suggests the question Shakespeare's characters mostly differ in the topics they approach, and less in style, as defined in authorship attribution. We have also compared character classification to play classification, and have shown that, while the play a character belongs to is a useful indicator to its identity in classification, it is not the only factor which helps tell characters apart. It might be interesting to further explore other features such as sentiment or emotion features, or to use a more powerful classifier (such as a convolutional/recurrent neural network). Many of the challenges of this analysis stemmed from the scarceness of data (many characters were discarded, lines were grouped together), so a learning algorithm that would be able to better handle small data might help expand the set of possible experiments and give more insight into the issue.

In the future it may also be interesting to look at how various authors pertaining to different periods and literary currents compare in terms of their ability (and desire) to create individual, stylistically independent characters. Literary theory (Wellek et al., 1956) tells us that the practice of giving characters strongly individual voices is a rather modern idea, and that characters evolved with time and literary current from the classical figures, who represented a typology, to the realist characters, who are pictured with strong individualities. This would be interesting to confirm experimentally, by extending the study to perform a diachronic analysis of characters in literary works.

Further, the analogous problem to author profiling could be tackled with regard to literary characters. Separately of whether characters are easy to distinguish stylistically from one another, it may be interesting to see if an author managed to belivably build a character's style that is consistent with features of the character's personality: such as age or gender.

Acknowledgments

This research is supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI—UEFISCDI, CoToHiLi project, project number 108, within PNCDI III.

References

Harald Baayen, Hans Van Halteren, and Fiona Tweedie. 1996. Outside the cave of shadows: Using syntactic annotation to enhance authorship attribution. *Literary and Linguistic Computing*, 11(3):121–132.

JF Burrows. 1987. Computation into criticism: a study of jane austen's novels andan experiment in method.

John Burrows and Hugh Craig. 2012. Authors and characters. *English studies*, 93(3):292–309.

Jonathan Culpeper. 2009. Keyness: Words, parts-of-speech and semantic categories in the character-talk of shakespeare's romeo and juliet. *International Journal of Corpus Linguistics*, 14(1):29–59.

Karina van Dalen-Oskam. 2014. Epistolary voices. the case of elisabeth wolff and agatha deken. *Literary and Linguistic Computing*, 29(3):443–451.

Anca Dinu, Liviu P Dinu, and Bogdan Dumitru. 2017. On the stylistic evolution from communism to democracy: Solomon marcus study case. In *RANLP*, pages 201–207.

Liviu P Dinu, Vlad Niculae, and Octavia-Maria Şulea. 2012. Pastiche detection based on stopword rankings. exposing impersonators of a romanian writer. In *Proceedings of the Workshop on Computational Approaches to Deception Detection*, pages 72–77.

Liviu P Dinu and Ana Sabina Uban. 2017. Finding a character's voice: Stylome classification on literary characters. In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 78–82.

Liviu P Dinu and Ana Sabina Uban. 2018. Analyzing stylistic variation across different political regimes. In *International Conference on Computational Linguistics and Intelligent Text Processing*, pages 110–123. Springer.

Liviu Petrisor Dinu, Marius Popescu, and Anca Dinu. 2008. Authorship identification of romanian texts with controversial paternity. In *LREC*.

Vlado Kešelj, Fuchun Peng, Nick Cercone, and Calvin Thomas. 2003. N-gram-based author profiles for authorship attribution. In *Proceedings of the conference pacific association for computational linguistics, PACLING*, volume 3, pages 255–264. sn.

Moshe Koppel and Jonathan Schler. 2003. Exploiting stylistic idiosyncrasies for authorship attribution. In *Proceedings of IJCAI'03 Workshop on Computational Approaches to Style Analysis and Synthesis*, volume 69, pages 72–80.

Thomas Corwin Mendenhall. 1901. A menchanical solution of a literary problem. *Popular Science Monthly*, 60(2).

Frederick Mosteller and David L Wallace. 1963. Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed federalist papers. *Journal of the American Statistical Association*, 58(302):275–309.

Grace Muzny, Mark Algee-Hewitt, and Dan Jurafsky. 2017. Dialogism in the novel: A computational model of the dialogic nature of narration and quotations. *Digital Scholarship in the Humanities*, 32(suppl_2):ii31–ii52.

Eric T Nalisnick and Henry S Baird. 2013. Character-to-character sentiment analysis in shakespeare's plays. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, volume 2, pages 479–483.

Fiona J Tweedie and R Harald Baayen. 1998. How variable may a constant be? measures of lexical richness in perspective. *Computers and the Humanities*, 32(5):323–352.

Krishnapriya Vishnubhotla, Adam Hammond, and Graeme Hirst. 2019. Are fictional voices distinguishable? classifying character voices in modern drama. In *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pages 29–34, Minneapolis, USA. Association for Computational Linguistics.

Carl Vogel and Gerard Lynch. 2008. Computational stylometry: Who's in a play? In *Verbal and Nonverbal Features of Human-Human and Human-Machine Interaction*, pages 169–186. Springer.

Rene Wellek, Austin Warren, et al. 1956. *Theory of literature*. Harcourt, Brace New York.