*Date: 30 December 2019*

**Selecting, Deriving and Applying Computational Authorship Attribution Methods that are Suitable for Studying Early Modern Dramas:**

**A Preliminary Exploratory Systematic Review**

CN8002 Coursework / Student No. 1720146

Chung Hoi Chiu

**Abstract:** Computational authorship attribution is using computers to determine who wrote a piece of text the author of which is disputed or unknown. This discipline began in 1960s. Thereafter, as computers have become increasingly powerful, more methods in this discipline have been invented. The purpose of the proposed research is to find two or three methods in this discipline that are particularly suitable for studying early modern dramas, to derive a new method that is based on certain special features of early modern dramas, and to apply the methods to one or two early modern dramas or parts thereof the authors of which are in dispute. It is hoped that the research can shed some new light on the disputes. This initial exploratory systematic literature review is produced to guide future literature review, to provide evidence of literature reviews conducted so far, to summarise the findings from literature reviews conducted so far, and to point out the future direction of the research.

**1. Introduction**

Computational authorship attribution is a discipline of study that uses computers to determine who wrote a piece of text the author of which is disputed or unknown. This discipline began in 1964. In that year, with the permission of the United States Government, the two pioneers, Federick Mosteller and David Wallace, reported in their book *Inference and Disputed Authorship: The Federalist* (Mosteller and Wallace, 1964, p. x) that they and their team studied texts of about 210,000 words that were typed on punch cards with computers located at the M.I.T. Computer Center, Cambridge, Massachusetts, and at the Center for Advanced Study in the Behavioral Sciences, Stanford, California respectively and found that the actual author of all of the 12 disputed Federalist Papers is James Madison, the fourth president of the United States (the Federalist Papers consist of 85 papers. The disputed ones are Nos. 49-58 and 62-63). Mosteller and Wallace used Bayes's Rules to make statistical inferences on log-negative binomial distribution probabilities (or on the simpler but less accurate log-poisson distribution probabilities) of the 30 carefully selected marker words (examples of the words are while, whilst, on and upon). The conclusion is convincing. Since 1964, dozens of different methods have been invented to study the 12 disputed papers, and most of them obtained the same or slightly deviated results. *Non-Stop*, one of the songs contained in the musical play *Hamilton* which has been performed in the West End (Victoria Palace Theatre) since 2017, has the following line: “John Jay got sick after writing five. Madison wrote twenty-nine. Hamilton wrote the other fifty-one.” The numbers contained in this line are based on Mosteller and Wallace’s research results.

Since the publication of Mosteller and Wallace's ground-breaking work in 1964, hundreds of works on this discipline have been published. According to Joseph Ruman, until 1998, more than 300 were published (Rudman, 1998); between 1998 and 2010, more than 600 were published comprised of journal articles, books, book chapters, dissertations, newspaper articles, on-line self-published papers and encyclopaedia entries (Rudman, 2010).

The most practical usage of authorship attribution should be finding forensic evidence, such as finding who is the actual author of a suicide note. However, it is not easy to find suitable texts for conducting research on forensic authorship attribution. Even a PhD student at Aston University (this university has a forensic linguistics centre) needed to find 90 volunteers with their gender, education, age, etc. known to fabricate 287 malicious texts. He then used 104 real malicious texts (the authors of which unknown) to gauge the fabricated texts (Nini, 2015). Therefore, I intend to focus my research on early modern dramas, especially those of the Elizabethan era.

I intend to aim my research at:

1. finding two or three existing authorship attribution methods that are the most suitable for studying early modern dramas and, if needed, improving them;
2. devising a new authorship attribution method that is especially suitable for studying early modern dramas;
3. finding or devising a suitable method for optimally synthesising results obtained from applying the authorship attribution methods mentioned in i. and ii. above; and
4. applying the above methods to one or two disputed early modern dramas or parts thereof and then analysing the results.

Therefore, I formulated the following two research questions, the resolution of which will substantially achieve the above research aims:

**Q1.** What are the current common methods for conducting computational authorship attribution studies, especially those that are suitable for studying early modern dramas?

**Q2.** Do commonplace deficiencies exist in the discipline of computational authorship attribution studies, such as lack of rigour and lack of reproducibility?

**2. Literature search strategy**

**2.1 Basic search terms**

I derived basic search terms which will be used for searching literature to answer the two research questions mentioned at the end of Section 1 above. They are:

(all commas used to separate search terms below indicate "OR")

**Basic search terms for Q1**: common, authorship attribution, computational methods, early modern, drama

**Basic search terms for Q2**: deficient, authorship attribution, rigour, reproducibility

I then derived synonyms, and alternative and related search terms from the above basic search terms:

Synonyms, and alternative and related terms derived from the basic search terms for **Q1**:

**common**: accepted, acclaimed, classic, common, famous, long-established, popular, renowned, prominent, well-known, seminal, successful, widely used

**authorship attribution**: authorship analysis, authorship attribution, authorship identification, authorship recognition, authorship verification, disputed authorship, forensic linguistics, stylometric analysis, stylometric studies, stylometry, text classification

**computational methods**: back-propagation algorithm, bayes theorem, bayesian statistics, chi square, clustering, computational methods, cosine similarity, cross-validation, decision trees, deep learning, delta, functional grammar, function word adjacency networks, function word proximities, information entropy, k-nearest neighbors, KNN, Kullback-Leibler Divergence, linear kernel, logistic regression, machine learning, markov chain, MLP, monte carlo attribution, multi-layer perceptron, multivariate analysis, n-gram, naive bayes, natural language processing, nearest shrunken centroids, negative binomial distribution, neural networks, PCA, NLP, POS, normal distribution, part of speech tag, poisson distribution, principal component analysis, random forest, rolling delta, shannon entropy, singular value decomposition, soft voting, student's t, SVD, support vector machines, SVM, syntactic level frequency, term-frequency inverse-document-frequency, tf-idf, vector-space model, vocabulary richness, WANs, word embedding, z-score

**early modern**: elizabethan, golden age, jacobian, shakespearean era

**drama**: act, play, blank verse, iambic pentameter, scene, stage, theatre

Synonyms, and alternative and related terms of the basic search terms for **Q2**:

**deficient**: bad, faulty, flawed, insufficient, poor, meager, sloopy, unsatisfactory

**authorship attribution**: authorship analysis, authorship attribution, authorship identification, authorship recognition, authorship verification, disputed authorship, forensic linguistics, stylometric analysis, stylometric studies, stylometry, text classification

**rigour**: accuracy, elaborate, precise, provable, rigorous, robust, sober, verifiable

**reproducibility**: duplicable, repeatable, replicable, reproducible

**2.2. Search engines and databases to be used for searching literature**

The following search engines and databases will be used for searching literature:

i. Cambridge Journals Online

ii. Early English Books Online

iii. Engineering Village, which includes Compendex and Inspec, which in turn include IEEE Xplore and Springer Link

iv. Explore the British Library

v. Oxford Scholarly Editions Online

vi. ProQuest

vii. Scopus, which includes the ACM Digital Library and ScienceDirect

viii. Wiley Online Library

Except the engine Explore the British Library, all other engines and databases mentioned above can be indirectly searched through a university library's search engine, provided that the library has subscribed to the service of the engine. However, for journal articles and ebooks published in recent months, one should still need to login to the search engines and databases to access them.

**2.3 Literature inclusion and exclusion criteria**

**2.3.1 Inclusion criteria**

I will not set up rigid inclusion criteria. However, I will pay particular attention to:

i. articles published by peer-reviewed journals,

ii. articles contained in article theses or coherent monographs published by esteemed publishers, and

iii. textbooks published by esteemed publishers.

However, in the field of computational authorship attribution, it is not uncommon for a simple method mentioned in an article published in a less important journal later turns out to be a very successful method. On the other hand, a well-funded research project conducted by well-known scholars was later found to contain fundamental errors. Furthermore, it is a quite common phenomenon in this field that critiques would be posted only on the Web. Therefore, I also pay attention to open access articles or even articles posted only on the Web, provided that the articles have been frequently cited or mentioned.

**2.3.2 Exclusion criteria**

I will exclude articles or books that are related to a similar field but not exactly related to computational authorship attribution. Such similar fields include text sentiment analysis, spam email detection, automatic information extraction, etc.

**2.4 Evidence matrices**

Evidence matrices of the literature found so far are shown below:

**2.4.1.** **Evidence matrix of literature found which are concerned with one or more methods**

(See next page)

|  |  |  |  |
| --- | --- | --- | --- |
| Article | Text type | Method(s) | Comments and (< 3 words opinion) |
| Argamon and Koppel, 2013 (17 pages) | Blogs | Systemic functional grammar | Only in theoretical discussion stage (qa: high) |
| Burrows, 2002 (22 pages) | Verses by 25 poets of the English Restoration Period | Delta | John Burrows is the inventor of the Delta method (qa: high) |
| Custodio and Parabono, 2018 (7 pages) | Fanfictions (closed-set) | Character n-gram, non-diacritic n-gram, word n-gram, multinomial logistic regression and support vector machines | This article is the report of the winning team of the 2018 Pan-Clef authorship attribution competition |
| Eisen et al., 2017 (29 pages) | Early modern dramas | Function work adjacency networks | -- |
| Feng, 2003 (9 pages) | The Federalist Papers | Selected function words, concave minimization and support vector machines | Results obtained the same as those of Mosteller and Wallace, 1964 |
| Ilsemann, 2018a (33 pages) | Early modern literature | Character n-gram, word n-gram and moving Delta | -- |
| Ilsemann, 2018b (40 pages) | Early modern literature | Character n-gram, word n-gram and moving Delta | Responding to criticism from Barber, 2019b |
| Keselj et al., 2003 (10 pages) | 19th Century English novels and Greek newspaper columns | Character n-gram | A simple formula is used as the classifier. But the results are still very satisfaction |
| Kjell et al., 1993 (10 pages) | The Federalist Papers | Character n-gram, cosine similarity, visualisation | Results obtained the same as those of Mosteller and Wallace, 1964 |
| Kocher and Savoy, 2018 (17 pages) | The Federalist Papers, the State of the Union addresses and the columns of Glasgow Herald and La Stampa newspapers | Distributed language representation (deep learning), cosine similarity, k-nearest neighbours | -- |
| Koppel et al., 2012 (10 pages) | Blogs | Character n-gram, Delta and Unmarking | Accuracy rates claimed to be above 80%. However, the article does not provide information about the texts used |
| Merriam, 2019 (8 pages) | Early modern literature | Word n-gram | Corpora built by Martin Mueller and Pervez Rizvi for helping people to learn how to apply word n-gram method to study early modern dramas were used |
| Miao et. al, 2005 (38 pages) | Texts obtained from web scraping | Character n-gram, word n-gram and Student's t-test | -- |
| Miranda-Garcia and Calle-Martin, 2012 (20 pages) | The Federalist Papers | Lemma-based Delta and part-of-speech-based Delta | Results obtained the same as those of Mosteller and Wallace, 1964 |
| Muttenthaler et al., 2019 (9 pages) | Fanfictions (open-set) | Character n-gram, distorted character n-gram, word n-gram, tf-idf, singular value decomposition and support vector machines | This is the report of the winning team of the 2019 Pan-Clef authorship attribution competition |
| Nance, 2017 (22 pages) | Early modern dramas | Microattribution | -- |
| Rosso et al., 2008 (11 pages) | Early modern literature | Word frequency, Shannon entropy and statistical complexity | -- |
| Segarra et. al, 2017 (30 pages) | Early modern dramas | Function word adjacency networks and relative Shannon entropy | Average accuracy rate claimed to be 92.6%. Deviation in spelling ignored |
| Taylor and Nance, 2017 (16 pages) | Early modern dramas | Microattribution | -- |
| Willams, 2014 (25 pages) | Early modern dramas | Visual representation of metrical patterns | Very litter mathematics involved |

(Table: Evidence matrix of literature found which are concerned with one or more methods)

**2.4.2** **Evidence matrix of literature found which are surveys of the field or critiques**

|  |  |  |
| --- | --- | --- |
| Article | Brief summary | Comments and (< 3 words opinion) |
| Athira and Thampi, 2015 (22 pages) | A general survey of the present situation of the field | -- |
| Auerbach, 2018 (5 pages) | Criticising the book New Oxford Shakespeare Authorship Companion as being lack of vigour and containing numerous fundamental statistical errors | -- |
| Barber, 2019a (11 pages) | Criticising the Function Word Adjacency Networks (WANs) method as being based on wrong assumptions and having not been independently tested | -- |
| Barber, 2019b (12 pages) | Criticising the article "Christopher Marlowe: Hype and Hoax" as not sufficiently rigorous (for example, some texts using their modern editions, some not) | -- |
| Daelemans, 2013 (12 pages) | A general survey of existing problems the field encountered. It also suggests putting more efforts on explaining existing methods | -- |
| Freebury-Jones, 2019a (10 pages) | A general survey of current situation of the field in the area of early modern literature | -- |
| Freebury-Jones, 2019b (10 pages) | Criticising the microattribution method | -- |
| Kestemont, 2014 (9 pages) | Acknowledging character n-gram is the best method / Suggesting directions for improving those methods which using function words as the features | -- |
| Kestemont et al., 2018 (24 pages) | An overview of the Pan-Clef 2018 authorship attribution competition. Closed-class. Texts used: fanfictions | Most of the participants used n-gram (different types) as features (character, word, POS, etc.), tf-idf to add weights and support vector machines as the classifier |
| Kestemont et al., 2019 (15 pages) | An overview of the Pan-Clef 2019 authorship attribution competition. Open-class. Texts used: fanfictions | All used n-gram (different types) as features (character n-gram, word n-gram, POS tags n-gram, punctuation n-gram, etc.). Most of the participants used tf-idf to add weights and support vector machines as the classifier |
| Mueller, 2019 (3 pages) | Describing how to apply a common computational authorship attribution method to early modern dramas / Suggesting a test: using statistics dept. students who have no knowledge on early modern literature to conduct the work | -- |
| Rudman, 1998 (15 pages) | A survey of the field from its starting in 1964 to 1998 | -- |
| Rudman, 2010 (16 pages) | A survey of the field from 1998 to 2010 | -- |
| Rizvi, 2018 (10 pages) | Early modern dramas | Criticising the microattribution method |

(Table: Evidence matrix of literature found which are surveys of the field or critiques)

**2.5 Literature quality assessment and data extraction**

I will not produce a form to guide me in performing the caption mentioned assessment and extraction. I am concerned that doing so would rigidise my thoughts. Instead, I will create a Google docs file, so that I can jot down my thoughts and notes at any time and place while reading the literature. Each literature will occupy at least one page. I will print out very important articles.

**3. Synthesis of the evidence obtained so far through reading or skimming the literature found**

**3.1 Common features and algorithms**

**3.1.1 Common features**

To conduct a computational authorship attribution study, one should first study a certain feature of the target text (a target text is a text the author this study aims to determine), and the test texts (test texts are texts the author of each of which are known and one or none of the authors is the author of the target text). The pattern of the feature of the target text is then compared with that of the test texts to find out which author of the test texts is the most likely author of the target text, or to conclude that it is unlikely that any one of the authors of the test texts is the author of the target text.

Common features used are:

i. The simplest two:

Sentence length and word size are the two simplest features. They have been used for more than a century. However, numerous experiments have already confirmed that they are not very useful (Mosteller and Wallace, 1964, pp. 6-7 and 259-262).

ii. Character n-gram, the most successful feature:

Character n-gram is the most successful feature so far. This feature was first suggested by a textbook published in 1976 (Bennett, 1976) and was first implemented in 1993 (Kjell et al., 1993). To tally 1-gram frequencies of a text just means to count the frequency of occurrence of each and every character of the text. To tally 2-gram frequencies of, for example, the text "Shall I compare thee to a summer’s day?" just means first to form the following two-character strings: sh | ha | al | ll | l\_ | \_i | i\_ | \_c | co … | r\_ | \_s | s\_ | \_d | da |ay|, and then count the frequency of occurrence of each and every such two-character strings (some people ignore character case, some people do not. The same also applies to punctuation, space, line feed character, etc.). The most likely author of a target text can be found by comparing the frequencies of occurrence of 1-gram, 2-gram, 3-gram, ... (usually only up to 5-gram) of the target text with those of the test texts. There are so many algorithms for performing the comparison, from as simple as basic algebra formulae, to as complicated as deep learning neural network models. However, no matter which algorithm is used, surprisingly, "character n-grams often outperform more complex feature sets" (Daelemans, 2013).

Character n-gram is also a language independent feature, which means that it can be applied to different languages. It also means that a researcher using this feature does not need to know the language of the texts.

iii. Word n-gram and function words:

After reading or skimming dozens of pieces of literature, I believe that the second most successful feature is word n-gram (the meaning of word n-gram can be inferred from ii. above), and the third most successful feature is the frequency of occurrence or distribution of selected function words.

Function words are those words that have low semantic meaning and are "linguistic glue" for connecting content words to form sentences. It is difficult to coin a new function word. For example, although we have longed for decades for the emergence of a gender neutral singular third-person personal pronoun (i.e. a gender neural form of he or she), it is unlikely that it will come into being in the foreseeable future. Furthermore, interestingly, in other types of text classification, for example, in spam email detection, function words are classified as stop words and will be filtered out from texts before they are studied. In authorship attribution, how to use function words betrays an author’s unconscious habit (use while or whilst, on or upon).

iv. Higher level features

The features mentioned in ii. and iii. above are at character level and word level. Features beyond these two levels are less successful. The most mentioned high level (sentence level) feature extraction tool is systemic functional grammar which was invented by Michael Halliday (Halliday and Matthiessen, 2004). However, using this tool to extract features from texts is now still in theoretical discussion stage (Argamon and Koppel, 2013).

**3.1.2 Common Algorithm**

The following passage extracted from the article summarising the 2019 Pan-Clef competition can be used to show that, currently, the most successful algorithms are support vector machines (SVM) and logistic regression:

As concerns the classiﬁers, the most popular choices are SVMs and ensembles of classiﬁers, usually exploiting SVM base models followed by Logistic Regression (LR) models.

(Kestemont et al., 2019)

The first computational authorship attribution project mentioned at the beginning of this essay used Bayesian log-logistic regression because the dispute is confined to two persons, namely, whether the 12 disputed papers were written by Alexander Hamilton or James Madison. For projects involving more than two candidate authors, SVM usually performs better than multilevel logistic regression.

Other common algorithms are chi square, k-nearest neighbours, Delta, Markov Chain, naive Bayes, and random forests.

**3.2 Methods used to study early modern dramas**

Currently, there are three methods being widely discussed, which are briefly described below:

**3.2.1 Word n-gram**

Using the word n-gram feature to perform authorship attribution studies is very successful. It attracted a Shakespearean scholar, Pervez Rizvi (He always calls himself an "independent student" and is now living in Croydon, London) to build an online corpus that contains 527 early modern plays (a total of 9,577,660 words) with the spelling of all the words in the corpus modernised and lemmatised. He also wrote code to help people to perform the word n-gram analysis on those plays. The corpus is located at www.shakespearestext.com, which is accessible by the public.

**3.2.2 Microattribution**

This method was invented by Gary Taylor, the general editor of the four-volume set *The New Oxford Shakespeare: The Complete Works*, which was published in 2017. One of the four volumes is *The New Oxford Shakespeare: Authorship Companion* (776 pages). This method is simple. It first divides the target text (in one article, the text used is a disputed chunk of an early modern drama of 63 words; in another article, 173 words) into strings of mostly two-word or three-word length (the method used is similar but not the same as n-gram. See 3.3 below). These strings will then be used to compare with strings obtained from dozens of plays the authors of which are known. The author who has the highest number of strings in the author’s texts the same as those of the target text will be regarded by this method as the most likely author of the target text. This method attracted many criticisms, one of which is from Pervez Rizvi, the person mentioned in 3.2.1 above (Rizvi, 2018).

**3.2.3 Function word adjacency networks**

This method first selects a number (for example: 100) of very frequently used function words and then uses the Markov Chain to study the preference of each and every candidate author in putting those function words adjacent to each other. This method first appeared in *IEEE Transaction on Signal Processing*, an engineering journal in 2015 (Segarra, et al., 2015). It then appeared in two humanities journals. Gary Taylor, the general editor, and the editorial board of *New Oxford Shakespeare: Authorship Companion* (mentioned in 3.2.2 above) regarded this method as convincing (Pollack-Pelzer, 2019). However, this method is under quite heavy criticism, one example of which is Auerbach, 2018. (Please also see the last two paragraphs of 3.3 below)

**3.3 Rigour and reproducibility**

After reading or skimming dozens of articles concerning computational authorship attribution, I formed the impression that the content of many of these articles lacks rigour or reproducibility. Below are two examples:

i. Lack of rigour:

In the article *Christopher Marlowe: Hype and Hoax* (Ilsemann, 2018a), comparison of character n-gram features of several Shakespeare plays and Marlowe plays are conducted. However, some of the plays used by the article are their modern editions (spelling and punctuation are modernised, while wording usually would be kept unchanged) while others are their old editions. Since the article used character n-gram as the feature for analysis, words in some plays are in modern spelling while in other plays in old spelling should substantially reduce the accuracy of the computations.

Furthermore, the author of the article did not check which old edition upon which a modern edition is based. It was found that the modern edition of *Henry V* that the article used is based on Folio 1 (published in 1623) (The article shows that the *Henry V* text contains the famous prologue "o for a muse of fire that would ascend the …". This prologue is contained only in Folio 1 and not the two older editions, Quarto 1 (published in 1600) and Quarto 2 (published in 1602)). However, Cristopher Marlowe was died in 1593, and English evolved rapidly during the Elizabethan period and the Jacobean period.

ii. Lack of reproducibility:

In the article *Imitation or Collaboration? Marlowe and the Early Shakespeare Canon* (Taylor and Nance, 2015), the method for dividing a chunk of text of 173 words into 377 cells is described as follows:

This identity-profile consists of a system of cells; in this case, the cells record, vertically, a decision-tree, a sequence of overlapping linguistic choices: the first and second word of a passage, then the second and third word, then the string of those first three words together.

It then gives an example of how the passage "within the compass of my curse wherein I" is divided. It is divided into: within the | within the compass | within the compass of my curse | the compass | the compass of | compass of | compass of my | of my | of my curse | my curse | my curse wherein | curse where in | curse wherein I

Why is there a five-word cell while other cells contain only two or three words? How to produce 377 cells from the 173 words. The article does not answer these two queries.

Furthermore, the 173 words should be extracted from a modern edition of *Titus Andronicus*, it is because the passage mentioned above should be "within the compasse of my curse, / wherein I" in Quarto 1 (1604), and "within the compasse of my curse, / Wherein I" in Folio 1 (1623) ("/" indicates a new line).

The article says that the 377 cells are used to compare texts of 80 early modern plays. However, it does not disclose the names and editions of the plays used and does not clarify whether words and punctuations of the 80 plays have been modernised, etc.

The article does not provide sufficient information as mentioned above makes reproduction of the experiment impossible, while reproducibility is a basic requirement for every scientific endeavour.

With regard to reproducibility, Darren Freebury-Jones, a guest editor of the journal *ANQ: A Quarterly Journal of Short Articles, Notes and Reviews* even states in an article of the journal published two months ago the following:

We should also note that practitioners of the Word Adjacency Networks method have, at the time of writing, yet to disclose their actual results. Readers, of course, should not accept authorship claims without seeing the actual results.

(Freebury-Jones, 2019b)

**4. Conclusions and further work**

The preliminary findings as described in Section 3 above indicate that although there are ample of methods for conducting computational authorship attribution studies, it is a common phenomenon that the publications introducing the methods are not sufficiently rigorous and do not provide sufficient technical details to facilitate reproduction. Nevertheless, I will test the suitability of some of these methods for studying early modern dramas. I will first test the method of using character n-gram and word n-gram as the features and support vector machines as the algorithm. I will then test two to four other carefully selected methods. Then, I will try to derive a method which is based on certain special features of early modern dramas. I will then implement two or three finally selected methods to one or two disputed early modern dramas or parts thereof. Finally, I will find or derive an algorithm to synthesise the results obtained from the implementations. I hope that my research will shed some new light on the disputes.

If my research can be published, I will produce a detailed technical report and make it available to the public.

**References**

Argamon S. and Koppel M. (2013) ‘A systemic functional approach to automated authorship analysis’, *Journal of Law and Policy*, 21(2), pp. 299-315.

Athira, U. and Thampi, S. M. (2015) ‘Authorship Analysis: Techniques and Challenges’, in Dawson, M. and Omar M. (ed.) *New Threats and Countermeasures in Digital Crime and Cyber Terrorism*. Pennsylvania: IGI Global, pp. 173-194.

Auerbach, D. B. (2018) ‘Statistical infelicities in the New Oxford Shakespeare Authorship Companion’, *ANQ - Quarterly Journal of Short Articles Notes and Reviews*, doi: 10.1080/0895769X.2018.1559023.

Barber, R. (2019a) ‘Function Word Adjacency Networks and Early Modern Plays’, *ANQ - Quarterly Journal of Short Articles Notes and Reviews*,

doi: 10.1080/0895769X.2019.1655631

Barber, R. (2019b) ‘Marlowe and overreaching: A misuse of stylometry’, *Digital Scholarship in the Humanities*, 34(1), pp. 1-12.

‘

Burrows, J. (2002) ‘‘Delta’: a measure of stylistic difference and a guide to likely authorship’, *Literary and Linguistic Computing*, 17(3), pp. 267-287.

Custodio, J. E. and Paraboni, I. (2018) EACH-USP ensemble cross-domain authorship attribution notebook for PAN at CLEF 2018, Available at: https://www.semanticscholar.org/paper/EACH-USP-Ensemble-Cross-domain-Authorship-Notebook-Cust%C3%B3dio-Paraboni/160a0a4f15113cca68288c4e540575e324870190 (accessed: 6 December 2019).

Daelemans, W. (2013) ‘Explanation in Computational Stylometry’, in Gelbukh, A. (ed.) *Computational Linguistics and Intelligent Text Processing. CICLing 2013. Lecture Notes in Computer Science, vol. 7817*, Heidelberg: Springer, pp. 451-462.

Eisen, M. et al. (2017) ‘Stylometric analysis of early modern period English plays’, *Digital Scholarship in the Humanities*, 33(3), pp. 500-528.

Feng, G. (2003) ‘The disputed federalist papers: SVM feature selection via concave minimization’, *2003 Conference on Diversity in Computing*. Atlanta, Georgia, USA, 15 - 18 October 2003. New York: ACM, pp. 42-46.

Freebury-Jones, D. (2019a) ‘Unsound Deductions in Early Modern Attribution: The Case of Thomas Watson’, *ANQ - Quarterly Journal of Short Articles Notes and Reviews*,

doi: 10.1080/0895769X.2019.1612231.

Freebury-Jones, D. (2019b) ‘“When a man hath a familiar style”: An introduction to authorship studies in early modern drama and literature’, *ANQ - Quarterly Journal of Short Articles Notes and Reviews*,

doi: 10.1080/0895769X.2019.1677210.

Halliday, M. A. K. and Matthiessen, C. (2004) *An Introduction to functional grammar*. 3rd edn. London: Arnold.

Ilsemann, H. (2018a) 'Christopher Marlowe: Hype and hoax', *Digital Scholarship in the Humanities*, 33(4), pp. 788-820.

Ilsemann, H. (2018b) *The Marlowe Corpus Revisited*, Available at: http://www.shak-stat.engsem.uni-hannover.de/marlowecorpusrevisited.pdf (Accessed: 6 December 2019).

Keselj, V. et al. (2003) ‘N-gram-based author profile for authorship distribution’, *Pacific Association for Computational Linguistics International Conference 2003*. Dalhousie University, 22-25 August 2003. Computer Science Department at Dalhousie University: Conference published printed book, pp. 255-264.

Kestemont, M. (2014) ‘Function words in authorship attribution from black magic to theory?’, *The 3rd Workshop on Computational Linguistics for Literature*. Gothenburg, Sweden, 27 April 2014. Stroudsburg, Pennsylvania: Association for Computational Linguistics, pp. 59-66.

Kestemont, M. et al. (2018) *Overview of the Author Identification Task at PAN-2018 Cross-domain Authorship Attribution and Style Change Detection*, Available at: https://www.semanticscholar.org/paper/Overview-of-the-Author-Identification-Task-at-and-Kestemont-Tschuggnall/f32c16d632f7100899e18c3a79e68c27f0afe64e (Accessed: 6 December 2019).

Kestemont, M. et al. (2019) *Overview of the cross-domain authorship attribution task at PAN 2019*, Available at: https://www.semanticscholar.org/paper/Overview-of-the-Cross-domain-Authorship-Attribution-Kestemont-Stamatatos/9b78bdd05fdbbedde33db045a3367c08b81339c8 (Accessed: 6 December 2019).

Kjell, B. et al. (1993) ‘Discrimination of authorship using visualization’, *Information Processing & Management*, 30(1), pp. 141-150.

Kocher, M. and Savoy J. (2018) ‘Distributed language representation for authorship attribution’, *Digital Scholarship in the Humanities*, 33(2), pp. 425-441.

Koppel, M. et al. (2012) ‘The “Fundamental problem” of authorship attribution’, *English Studies*, 93(3), pp. 284-291.

Merriam, T. (2019) ‘Phrasal Criteria for Authorship, with Consideration of Negative Checks’, *ANQ - Quarterly Journal of Short Articles Notes and Reviews*,

doi: 10.1080/0895769X.2019.1615860.

Miao, Y. et al. (2015) ‘Document clustering using character N-grams: A comparative evaluation with term-based and word-based clustering’, *The 14th ACM international conference on Information and knowledge management*. Bremen, Germany, 31 October – 5 November 2005. New York: ACM, pp. 357-358.

Miranda-García, A. and Calle-Martín, J. (2012) ‘The authorship of the disputed Federalist Papers with an annotated corpus’, *English Studies*, 93(3), pp. 371-390.

Mosteller, F. and Wallace D. L. (1964) *Inference and disputed authorship: The Federalist*. Reading, MA, USA: Addison-Welsey.

Mueller., M. (2019) 'Who wrote what: Is it time for a blind experiment in early modern drama?', *ANQ - Quarterly Journal of Short Articles Notes and Reviews*,

doi: 10.1080/0895769X.2019.1653164.

Muttenthaler, L. et al. (2019) Authorship Attribution in Fan-Fictional Texts given variable length Character and Word N-Grams Notebook for PAN at CLEF 2019, Available at: https://www.semanticscholar.org/paper/Fan-Fictional-Texts-given-variable-length-Character-Muttenthaler-Lucas/2f28d2861efd9f2b90678a02726e7d0ad8936880 (Accessed: 6 December 2019).

Nance, J. V. (2017) *‘“We, John Cade”: Shakespeare, Marlowe, and the authorship of 4.2.33–189 2 Henry VI*’*, Shakespeare*, 13(1), pp. 30-51.

Nini, A. (2015) *Authorship profiling in a forensic context*. Unpublished PhD Thesis, Aston University, Available at: https://ethos.bl.uk/OrderDetails.do?did=1&uin=uk.bl.ethos.642012 (Accessed: 6 December 2019).

Pollack-Pelzner, D. (2017) 'The Radical Argument of the New Oxford Shakespeare', *New Yorker*, 19 February 2017, Available at: https://www.newyorker.com/books/page-turner/the-radical-argument-of-the-new-oxford-shakespeare (Accessed: 6 December 2019).

Rizvi, P. (2018) ‘The problem of microattribution’, *Digital Scholarship in Humanities*, doi:10.1093/digitalsh/fqy066.

Rosso, O. A. et al. (2008) ‘Shakespeare and other English Renaissance authors as characterized by information theory complexity quantifiers’, *Physica A: Statistical Mechanics and its Applications*, 388(6), pp. 916-926.

Rudman, J. (1998) ‘The state of authorship attribution studies: Some problems and solutions’, *Computers and the Humanities*, 31(4), pp. 351–365.

Rudman, J. (2010) ‘The State of NonTraditional Authorship Attribution Studies – 2010: Some Problems and Solutions’, *English Studies*, 93(3), pp. 259-274

Segarra, S. et al. (2015) ‘Authorship attribution through function word adjacency networks’, *IEEE Transactions on Signal Processing*, 63(20), pp. 5464-478.

Taylor, G. and Nance, J. (2017) ‘Imitation or collaboration? Marlowe and the early Shakespeare canon’, *Shakespeare Survey*, 68, pp.32-47.

William, R. B. (1976) *Scientific and engineering problem solving with the computer.* New Jersey: Prentice-Hall.

Williams, R. (2014) *Visual metre as an authorship attribute*, Available at: http://www.shakespeareanauthorshiptrust.org.uk/pdf/VisualMetre (Accessed: 6 December 2019).