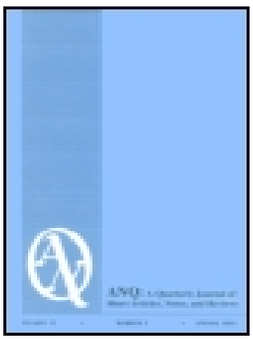

ANQ: A Quarterly Journal of Short Articles, Notes and Reviews

ISSN: 0895-769X (Print) 1940-3364 (Online) Journal homepage: https://www.tandfonline.com/loi/vanq20

Authorship Attribution for Early Modern Plays Using Function Word Adjacency Networks: A Critical View

Pervez Rizvi

To cite this article: Pervez Rizvi (2018): Authorship Attribution for Early Modern Plays Using Function Word Adjacency Networks: A Critical View, ANQ: A Quarterly Journal of Short Articles, Notes and Reviews, DOI: 10.1080/0895769X.2018.1554473

To link to this article: https://doi.org/10.1080/0895769X.2018.1554473

ANQ: A QUARTERLY JOURNAL OF SHORT ARTICLES, NOTES, AND REVIEWS
https://doi.org/10.1080/0895769X.2018.1554473

Authorship Attribution for Early Modern Plays Using Function Word Adjacency Networks: A Critical View

Pervez Rizvi

Authorship attribution scholars must deal nowadays with the rapid growth of the infant discipline known as computational stylistics. Scholars whose expertise is in English literature face considerable difficulty with the results published by researchers in this discipline, since the methods used involve large-scale data processing and, in some cases, advanced mathematics. In this note I hope to show that, even without expertise in mathematics or computing, we may subject such methods to scrutiny and decide if they are appropriate for the uses to which they are now being put. I shall do this by critiquing a new method known as function word adjacency networks, which has been used to assign parts of the *Henry VI* trilogy from William Shakespeare to Christopher Marlowe.

This method was introduced in an engineering journal a few years ago; however, it came to the attention of Shakespeare scholars when it appeared in an article in *Shakespeare Quarterly* in 2016 (Segarra et al.) and was thereafter used in press releases to gain publicity for the New Oxford Shakespeare complete works edition of 2017, and it appeared again in 2018 in an article by the same authors (Eisen et al). It is simple to state in non-technical terms. Any text may be regarded as containing two kinds of words: function words and non-function words. Function words are, as the name suggests, words such as *the, and, of, about*, and so on. The method asserts that, when deciding who wrote a text, it is not necessary to know the non-function words it contains. It is enough to know only the function words, the order in which they occur, and how far apart they are from each other. As the inventors of the method admit: "We cannot explain in psychological, neurological, or artistic terms just why various words cluster with other words" (Segarra et al. 235). The method takes no account of an author's choice of non-function words, still less their meaning. It regards as indistinguishable the phrases "O, for a muse of fire" (*Henry V*) and "O, for a couple of faggots" (Fletcher and Massinger's *The Little French Lawyer*), since they use the same function words and in the same places.

The example I have just given was a little tongue-in-cheek, since Fletcher and Massinger might have been parodying Shakespeare. More seriously, we should ask if, *contra* the premise of the method I am considering, an author's choice of a function word can be influenced by a non-function word he has just used. This is intuitively obvious and should not need examples, but examples can be found. To find them, I consulted Pervez Rizvi's self-published concordances of N-grams—that is, phrases of N words—from more than five hundred plays from the early modern period.[1] The simplest search is to look for cases where the same function word always follows a non-function word. For example, *upside* is listed in the concordance as occurring 23 times, and it is followed every time by *down*. Or consider the word *looker* and its plural *lookers*, which occur 48 times and are followed by *on* every time. Similarly, *devoid* occurs 24 times and is followed every time by *of*. The obvious inference is that the non-function word influenced the choice of the function word that followed it. To be thorough, let us check this explanation. Rizvi has provided a separate concordance of function word skip N-grams. These are N-grams found by skipping over non-function words, as if they were not present in the text. If, for example, we look at all function words preceding *of* in that concordance, we discover that *of* is preceded by many different function words within the same texts. This is exactly as we should expect, given how common *of* is in the English language. If we had found that *of* was always preceded by the same function word, we might have had some reason to believe that function words alone determine

CONTACT Pervez Rizvi pervez.rizvi@gmail.com 15 Blake Road, Croydon CR0 6UH, UK

the author's choice of the next function word, but the evidence tells us the opposite. The reader may verify from the concordance that, unsurprisingly, the same is true of the other function words I exampled above, *down* and *on*. The obvious and natural view, consistent with the evidence, is that, for example, the use of *devoid* influenced the author to use *of* immediately afterward, or that he chose *devoid of* as a phrase rather than choosing the non-function word and the function word separately. If it is true, as both our intuition and the textual evidence suggest, that non-function words exercise an influence over the function words that follow them, then the premise of the method I am considering becomes highly suspect.

Notwithstanding that, let us continue and observe that the reason to use the method is that its inventors claim that the function words found in a text, the order in which they occur, and how far apart they occur constitute a way of telling authors apart. Indeed, they used this evidence alone to purport to distinguish between Shakespeare and Marlowe in the *Henry VI* trilogy. If this is a sound procedure, then one of the most important kinds of evidence must be when one author sometimes follows a given function word by another given function word, but a different author never does. *Ex hypothesi*, this is very much the kind of evidence that should enable us to distinguish between those authors. It was therefore essential for the method to take such pairs of words into account. Remarkably, it does not. On the contrary, it excludes them. This astonishing decision was not made for any reason to do with authorship or texts. It was made simply because one of the formulae in the method involves division by some numbers calculated from pairs of function words found in texts (Eisen et al. formula 7). When some function words never follow others in an author's work, these numbers are zero. Mathematics does not allow division by zero, so the inventors of this method faced an obstacle. The conclusion they should have drawn was that they were using a method that is not suitable for what they were trying to do. They should have invented a better method. Instead, they chose simply to exclude the inconvenient evidence from consideration (Eisen et al. 503). This means that if one author uses a pair of function words a thousand times, but another uses it just once, the method duly counts it as a difference between them; however, if one author uses it a thousand times and the other never, then the method disregards that evidence. That is utterly irrational.

To bring this point home, I did a further search. For this, I took the twenty-eight Shakespeare plays and the six Marlowe plays that the Segarra et al. article regards as securely attributed (Segarra et al. 255). I looked at all function words in those plays, using the same list of one hundred function words that the article used (Segarra et al. 254). Now, consider that a set of one hundred words generates 10,000 pairs of words, since each of the hundred can follow each of the others, including itself, possibly with some other words between them. Using computerized texts, I searched for each of these 10,000 pairs in each play. I found 2,670 pairs of function words in which the first word is followed by the second in a Shakespeare play but never in a Marlowe play. For example, Shakespeare follows *about* by *at* (e.g., "they may have their throats about them at that time" in *Henry V*) but Marlowe never does. Conversely, I found 116 pairs of such words in Marlowe but not Shakespeare; for example, Marlowe follows *after* by *any* (e.g., "to imagine that after this life there is any pain" in *Doctor Faustus*), but Shakespeare never does.[2] It is difficult to see how we can have confidence in a method that disregards highly relevant evidence such as this, simply because it does not know what to do with it.

It might be argued that, notwithstanding the problems I have demonstrated, the method happens to work, since its inventors claim some high success rates in the tests they did on plays of known authorship. However, the success is illusory, as I shall explain. The claimed successes were obtained by "training" the method. The Eisen et al. article tells us: "The number of function words chosen from the full set of 100 varies for each experiment and is determined by a training process in which we measure the power of each word in helping to discriminate between the particular authors under consideration" (Eisen et al. 505–06). This means that the function words were carefully chosen to ensure that the method attributed specific sets of plays to their already-known authors, the words having to be varied for each set of plays. The article tells us that it achieved success rates as high as 100 percent in some of its tests (Eisen et al. 508). It does not tell us how many function words it used for every one of its tests, still less what they were, but it mentions at one point that it selected 76 out

of its 100 function words for use in some experiments, and 55 in some others (Eisen et al. 516). Now, the number of ways in which we can select several dozen words from a hundred words is astronomically large. The number of ways of selecting 76 words out of 100 is approximately 8 followed by 22 zeroes; and the number of ways of selecting 55 out of 100 is almost a million times greater than that.[3] Therefore, it is no great surprise that if we pick many different subsets from a set of 100 function words, we will eventually find one that gives us the result we are training our method for. Tellingly, despite having literally trillions of subsets of the 100 function words to choose from, the software used in Eisen et al. failed to find one that works well for all authors. Instead, it was obliged to use different subsets for different authors. This means that the claimed success rates are of no great significance. Imagine if you asked someone to think of a number and then allowed yourself trillions of attempts to guess what it is. Eventually, your guess will be equal to or close to the number you are trying to guess. To claim then that you have successfully guessed the number sounds impressive until you disclose how you did it. The value of an authorship attribution method is that, once we have satisfied ourselves of its accuracy by testing it with texts of known authorship, we can apply it, *unchanged*, to texts of unknown authorship and have some confidence in the results. That is not possible with the function word adjacency method.

Be that as it may, the success rates claimed in the Segarra et al. article are less impressive than they appear. In the section of the article headed Validation, we are given two success rates, 89.6 percent and 93.6 percent (Segarra et al. 242), but they are misleading. The 93.6 percent success rate is claimed for "the ninety-four plays whose authorship is not in dispute." We are told that the method attributed all but six of those plays to their known authors. However, these attributions were obtained after training the method to recognize the correct authors for those very plays. As the article admits: "Each play is attributed… based on the adjacencies of the one hundred function words that were found in training (based on the full, undisputed sole-authored canon of each dramatist) to be the most discriminating" (Segarra et al. 243). The argument is circular. Since the function words used in the tests were changed until they yielded the claimed success rate, the success rate can hardly be used as evidence of the method's correctness. In passing we may also note that even the claimed success rate of 93.6 percent would cause dozens of early modern plays to be attributed to the wrong authors.

It is easy enough to work out the true success rate of the method. If we exclude the ninety-four plays whose correct attributions were achieved by training, we are left with 60 plays for which the method could be said to be a genuine test. We are not explicitly told for how many of these sixty plays the method succeeded in identifying the correct author. Instead, we are told that it failed for 16 out of all 154 plays, and that it failed for six out of the 94 sole-authored plays among them (Segarra et al. 243). From this information, we can deduce by simple subtraction that the method failed for ten out of the 60 multi-authored plays. We can understand why the article omits to state this explicitly, for it is a shockingly high error rate. If this method were accepted for further use, and if its actual success rate of only 50 out of 60 were repeated across the board, it would attribute more than a hundred early modern plays to the wrong authors. Moreover, if we consider that, for its tests, the inventors of the method picked the low-hanging fruit—that is, the plays most securely attributable to one author—we may feel even less confidence in the method's ability to pick the correct authors of the anonymous plays or the disputed ones, such as the *Henry VI* trilogy.

I end by pointing out a troubling feature of the way in which both the articles I have cited present their results. It is disclosed in what appears to be an innocuous passage:

> We do not report raw relative entropy values between the play being attributed and the author profiles, but instead subtract from these values the relative entropy between the play and a profile containing all available texts. … This is done to make the figures easier to view but does not change the results in any way. (Eisen et al. 507; see Segarra et al. 243, for the equivalent admission)

That is a very significant manipulation of the test results. To understand why, consider that, given any set of numbers, we can choose one of those numbers, or some other number close to them, and

subtract it from them. For example, given the numbers 101, 102, 103, 104, and 105, we might choose to subtract 100 from each of them, giving us 1, 2, 3, 4, and 5. The differences between the numbers have remained unchanged, but their relationship to each other has changed dramatically. Before subtraction the largest number, 105, was only about 4 percent larger than the smallest, 101; after subtraction, it is five times as large. It is obvious that both articles mislead us formally, by claiming that what they did "does not change the results in any way"; what's worse, they mislead us substantively, too. The manipulation of the results disclosed in the quotation I have given above always has the effect of magnifying differences, perhaps by huge amounts, and, in this case, it makes the graphs in the articles look more impressive, perhaps much more impressive, than they would be if the actual test results had been disclosed to us.

I hope I have shown that the function word adjacency network method suffers from defects that should cause us to doubt its efficacy. Moreover, the fact that the method has not been defined in a general form, but instead has an almost unlimited number of variants, depending on our choice of function words, means that it is problematic for use in authorship attribution, because when we are dealing with a text of doubtful authorship, we have no way of knowing which set of function words will give us the right answer. Put simply, this method is not suited for use in authorship attribution.

Notes

1. See www.shakespearestext.com/can (accessed 4 November 2018).
2. Following the method, I considered only pairs of function words that occur within ten words of each other (Eisen et al. 502). The first ten pairs for Shakespeare, in alphabetical order, are: *about after; about at; about before; about every; about hence; about little; about more; about most; about much; about must*; and the first ten pairs for Marlowe are: *about against; about both; about none; about while; after any; against might; against once; against over; another next; another while*.
3. For the formula, see, for example, http://mathworld.wolfram.com/k-Subset.html (accessed 4 November 2018).

Disclosure statement

No potential conflict of interest was reported by the author.

Works cited

Eisen, Mark, et al. "Stylometric Analysis of Early Modern Period English Plays." *Digital Scholarship in the Humanities*, vol. 33, 2018, pp. 500–28. doi:10.1093/llc/fqx059.

Segarra, Santiago, et al. "Attributing the Authorship of the *Henry VI* Plays by Word Adjacency." *Shakespeare Quarterly*, vol. 67, 2016, pp. 232–56. doi:10.1353/shq.2016.0024.