Hartmut Ilsemann

The Shakespeare Authorship of *Edmund Ironside*

Abstract

The study explores the authorship of *Edmund Ironside*, an anonymous Elizabethan play first printed in 1927. While earlier critics, including Ephraim Everitt and Eric Sams, attributed it to Shakespeare—with Ramon Jiménez later supporting this claim but arguing for Edward de Vere as the true author—the paper turns to modern stylometry using the R Stylo toolset.

Analyses with Rolling Delta, Rolling Classify, and the General Imposters method consistently show strong stylistic proximity between *Ironside* and Shakespeare's early plays, particularly *1* and *2 Henry VI*, though Robert Greene and Anthony Munday also appear as significant influences. Classification methods (nsc and svm) further highlight Greene's role as a stylistic source, reflecting the documented rivalry between Greene and Shakespeare.

The findings suggest *Ironside* is best seen as an early Shakespearean work, shaped by Greene's style and confirming Shakespeare's early emergence as both actor and playwright in the mid-1580s. The evidence also undermines claims for Edward de Vere's authorship.

Introduction

Thanks to the processing of large amounts of data and a technique known as 'distant reading', stylometric analysis has entered a new phase in terms of determining authorship. This is particularly evident in English Renaissance drama, where many texts are either anonymous or apocryphal. One such text, acquired by the British Museum in 1865 as part of the Egerton Collection, plays a special role in this regard. First printed in 1927, Edmund Ironside has no known origin, including the time and place of its performance. Only a few contributions have been made as far as literary criticism is concerned. In 1954, an American scholar named Ephraim Everitt identified *Ironside* as an early Shakespeare play, but this idea received little acclaim. However, it was the musicologist Eric Sams who conducted pioneering research and corroborated the claim that William Shakespeare wrote this early history play. In 2003, Ramón Jiménez confirmed Sams's findings, providing a comprehensive analysis and scholarly interpretation that left no doubt as to its place in the Shakespeare canon. However, Jiménez believed that Edward de Vere was the true author of Shakespeare's works. An exciting question is whether modern, non-traditional stylometry using the R Stylo program suite can provide clarifying data.

Analyses

The main features applied were Rolling Delta (1), Rolling Classify (2), and the General Imposters method (3). These are all explained in detail in the R Stylo paper by Maciei Eder, Jan Rybicki and Mike Kestemont, as well as in a blog post by Eder (30 May 2018). Further information about the chosen parameters can be found in the respective Notes section. Corpus preparation is certainly an important aspect of Rolling Delta analyses, since authors, genres and dates must be taken into account. However, initial analyses of 154 reference texts showed that the sorted delta values of the reference plays already indicate the degree of similarity and exclude unsuitable texts. Thus, a history play shows no similarity to comedies, and time-wise delta values also group plays from the same time period together. When Rolling Delta was carried out using 56 carefully selected reference texts, the list of similar plays was the same as that listed at http://www.shak-stat.engsem.uni-hannover.de/eallironside.html on the Shakespeare Statistics webpage. Conditional formatting in red and blue provides an impressive overview of the range of plays in the vertical dimension and the variations within a single play in the horizontal dimension. The target text, 'anon_ironside.txt', was used as a reference text again to obtain its overall delta value. This is not zero, as one might expect, but is derived from the average standard deviation of the entire corpus. The next plays that are stylistically closest are *1* and *2 Henry VI*, followed by further Shakespeare plays. However, similarity in style does not prove identity of authorship. Other playwrights whose texts appear in the reddish sections include Chettle, Greene, Marston, Munday, and Peele. Table 1 shows the first 20 stylistically similar texts, from which the plays in bold were selected as reference texts, with a maximum of two plays per author.

Table 1 Rolling Delta analysis with mf3c and 4,000-word windows

	A	B	C
1	mf3c,4000		
2		Δ	dist.
3	anon_ironside	15,4	
4	**shak_2henry6**	24,0	8,6
5	anon_weakwall	24,3	8,9
6	**shak_1henry6**	24,3	8,9
7	shak_richiii1592	24,3	8,9
8	shak_2henryiv	24,5	9,1
9	shak_richii1595	24,7	9,3
10	shak_hamlet1600	24,9	9,5
11	anon_edwardiii	24,9	9,5
12	shak_1henryiv1596	25,0	9,5
13	**mun_deathh1601**	25,0	9,5
14	shak_h5	25,2	9,8
15	**greene_jamesiv1591**	25,2	9,8
16	shak_john1596	25,4	10,0
17	**chettle_hoffman**	25,5	10,1
18	shak_iiihenryvi	25,7	10,3
19	**greene_selimus**	25,7	10,3
20	anon_moregut	25,8	10,4
21	**greene_lookingglass**	25,9	10,5
22	anon_kingleir1594	25,9	10,5
23	**peele_edward1**	26,1	10,7

Their plays were used as reference texts in Table 2, which displays the window centroid deltas in column A, arranged vertically. The lowest deltas in columns B to J have a dark background; the second-lowest deltas have a grey background; and the third-lowest deltas have a lighter grey background. When Burrows introduced the concept of 'delta' in 2002, he deliberately referred to 'guide to likely authorship' in the title of the paper and included the next lowest deltas as well.

Table 2 Rolling Delta analysis of *Edmund Ironside*

	A	B	C	D	E	F	G	H	I	J	K	L
1		**Rolling Delta** Analysis of *Edmund Ironside* with 9 selected										
2		reference texts due to their vicinity to the target text										
3	**window centroids**	chettle_hoffman	greene_jamesiv	greene_selimus	kyd_soliman	mars_malcontent	mun_deathh1601	peele_edward1	shak_1henry6	shak_2henry6	**SCENES**	**WORDS**
4	2000	25,7	24,2	26,7	27,8	27,3	24,6	26,6	25,1	24,5	I,1	1999
5	2250	25,5	24,2	26,6	27,4	26,7	24,3	26,7	24,6	24,2		
6	2500	25,5	24,3	26,6	27,6	27,0	24,4	26,6	25,1	24,4	I,2	2423
7	2750	25,8	24,7	26,7	27,7	26,5	24,6	26,8	24,9	23,9	I,3	2810
8	3000	26,3	25,3	26,8	27,9	26,9	25,0	27,3	25,3	24,3		
9	3250	26,7	25,5	27,0	27,8	27,1	25,4	27,9	25,9	25,0		
10	3500	25,9	25,6	26,9	28,0	27,0	24,8	27,7	25,6	25,0	II,1	3415
11	3750	25,6	25,5	26,2	27,4	26,7	24,3	27,1	25,5	24,3		
12	4000	24,4	24,7	25,4	26,3	25,5	23,3	25,9	24,5	23,3		
13	4250	24,9	25,0	25,9	26,5	25,9	24,0	26,4	24,8	23,9	II,2	4272
14	4500	25,2	25,0	25,5	26,8	26,0	24,4	26,0	24,4	23,7		
15	4750	25,3	24,7	25,6	26,9	26,5	24,2	25,8	24,3	23,5		
16	5000	25,3	24,7	25,5	27,1	26,6	24,9	26,2	24,2	23,3		
17	5250	25,5	24,6	25,2	27,5	26,3	25,2	25,9	23,9	23,2		
18	5500	26,5	26,2	25,6	28,3	27,1	26,7	26,8	24,6	23,3		
19	5750	26,8	26,8	25,7	28,4	27,2	27,3	27,2	23,9	23,5		
20	6000	27,4	27,8	26,1	28,9	27,6	28,1	27,3	23,8	24,1	II,3	6101
21	6250	27,4	27,5	26,7	29,3	27,6	28,2	27,7	24,2	24,6		
22	6500	26,4	26,7	26,0	28,7	26,7	27,5	26,8	23,7	24,0	III,1	6456
23	6750	25,2	25,3	24,9	27,4	25,5	26,6	25,3	22,6	23,0		
24	7000	24,8	25,1	24,4	26,8	24,6	26,0	25,0	22,7	22,9	III,2	7155
25	7250	24,4	24,7	24,2	26,2	24,4	25,5	24,6	22,6	22,3		
26	7500	24,6	24,7	25,0	26,4	24,3	25,4	24,8	23,2	22,8	III,3	7472
27	7750	24,7	24,7	26,0	27,2	25,1	25,6	25,3	23,7	23,5	III,4	7602
28	8000	25,2	24,7	26,4	27,3	25,8	25,7	25,6	24,2	23,8		
29	8250	25,5	25,2	26,8	27,7	25,9	26,1	25,9	24,3	24,5		
30	8500	26,0	25,4	27,4	28,5	26,4	26,2	26,6	25,0	25,3		
31	8750	26,1	25,7	27,5	28,3	26,5	26,1	26,8	25,3	25,6		
32	9000	25,4	25,5	27,0	27,0	25,7	25,5	26,4	24,8	25,4		
33	9250	25,2	25,1	26,9	26,3	25,8	25,1	26,2	24,9	25,0	III,5	9174
34	9500	25,2	24,9	27,2	26,1	25,9	24,5	26,2	25,1	25,5	III,6	9453
35	9750	24,9	24,1	26,9	25,9	25,4	23,8	25,5	24,9	24,9		
36	10000	25,3	24,3	27,0	25,9	25,7	23,7	25,8	25,2	25,1		
37	10250	25,4	24,7	26,7	25,5	25,7	23,8	25,9	25,7	24,9		
38	10500	25,8	25,2	26,6	25,2	26,0	23,9	26,0	25,5	24,9	IV,1	10602
39	10750	26,4	25,9	27,1	26,1	27,1	24,1	26,4	25,9	25,5		
40	11000	26,0	25,0	27,2	26,2	26,9	23,7	25,9	25,7	25,3		

41	11250	26,2	25,6	26,9	26,4	26,8	**24,2**	26,1	**25,6**	**25,3**	IV,2	11318
42	11500	26,4	25,3	25,8	26,2	27,0	**24,7**	26,2	**24,7**	**24,2**	IV,3	11585
43	11750	26,4	25,5	25,5	25,8	27,1	**24,9**	25,9	**24,2**	**23,6**		
44	12000	25,9	25,3	25,0	25,4	26,5	**24,9**	25,5	**23,1**	**23,2**	IV,4	11892
45	12250	26,0	25,1	**24,5**	25,3	26,6	24,8	25,2	**23,2**	**23,1**		
46	12500	25,7	25,0	24,5	25,0	26,7	**24,5**	24,9	**23,4**	**22,5**		
47	12750	25,8	25,3	24,4	24,9	26,8	**24,3**	24,9	**23,6**	**22,6**	V,1	12823
48	13000	26,3	25,5	24,6	25,5	27,4	**24,5**	25,1	**23,9**	**23,2**		
49	no.		3				10		11	21		
50	%		6,7				22		24	47		
51									Σ	71		
52	Window size: 4.000 words											
53	Step size: 250 words											
54	Culling value: 70 %											
55	Extracted from 56 reference texts											
56											V,2	15246

Of the forty-five window measurements, thirty-two belong to early Shakespeare plays, amounting to 71%. However, the second- and third-lowest measurements are also noteworthy, particularly in regard to Robert Greene and Anthony Munday. Munday can be seen as a borrower, as his play was written long after those of Greene and Shakespeare. Greene records 24 second- or third-lowest delta values here, putting him ahead of Munday, who records 17. Interestingly, the classification results also favour Greene, with both the NSC (nearest shrunken centroid) and SVM (support vector machine) classifiers.

Table 3 nsc classification with various window sizes and variables

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
	Rolling Classify analysis with nsc and window sizes between 1.000 and 8.000 words																								
		1000	2000	3000	4000	5000	6000	7000	8000	1000	2000	3000	4000	5000	6000	7000	8000	1000	2000	3000	4000	5000	6000	7000	8000
1	0																								
2	250																								
3	500	G				mf1w				S				mf2c				S				mf3c			
4	750	G								S								S							
5	1000	G	G							G	S							S	S						
6	1250	G	G							S	S							S	S						
7	1500	G	G	G						S	S	S						S	S	G					
8	1750	G	G	G						S	S	G						G	G	S					
9	2000	G	G	G	G					S	S	G	G					G	G	G	G				
10	2250	G	G	G	G					S	G	G	G					G	G	G	G				
11	2500	G	G	G	G	G				G	G	G	G	S				G	G	G	G	G			
12	2750	G	G	G	G	G				G	G	G	G	G				S	G	G	S	G			
13	3000	G	G	G	G	G	G			G	G	G	G	G	S			P	G	G	S	G	S		
14	3250	G	G	G	G	G	G			M	G	G	G	G	S			M	M	G	G	G	G		
15	3500	G	G	G	G	G	G	G		M	M	G	G	G	S	S		M	M	G	G	G	G	G	
16	3750	C	G	G	G	G	G	G		M	G	G	G	G	S	S		M	S	G	G	G	S	S	
17	4000	C	G	G	G	G	G	G	G	M	G	G	G	G	S	S	S	C	S	G	G	G	S	S	S
18	4250	G	G	G	G	G	G	G	G	S	G	G	G	G	S	S	S	S	G	G	G	G	G	S	S
19	4500	G	G	G	G	G	G	G	G	S	S	G	G	G	S	S	S	S	G	G	G	G	S	S	S
20	4750	G	G	G	G	G	G	G	G	S	S	G	G	G	S	S	S	S	G	G	S	G	S	S	S
21	5000	G	G	G	G	G	G	G	G	S	S	S	S	S	S	S	S	S	S	G	S	S	S	S	S
22	5250	G	G	G	G	G	G	G	G	S	S	S	S	G	S	S	S	G	S	S	G	S	S	S	S
23	5500	G	G	G	G	G	G	G	G	S	S	S	S	S	S	S	S	G	S	S	S	S	S	G	S
24	5750	G	G	G	G	G	G	G	G	G	S	S	S	S	S	S	S	G	G	S	S	S	S	G	S
25	6000	G	G	G	G	G	G	G	G	G	S	S	S	S	S	S	S	G	G	S	S	S	S	G	S
26	6250	G	G	G	G	G	G	G	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	G	S
27	6500	S	S	G	G	G	G	G	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	G	S
28	6750	S	S	S	G	G	G	G	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	G	S
29	7000	S	S	G	G	G	G	G	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	G	S
30	7250	S	S	S	G	G	G	G	G	S	S	S	S	S	S	S	S	S	S	S	S	G	S	G	S
31	7500	S	S	G	G	G	G	G	G	S	S	G	G	G	S	S	S	S	S	S	S	G	G	S	S
32	7750	S	S	S	G	G	G	G	G	S	S	G	G	G	S	S	S	S	S	S	S	G	G	G	S
33	8000	S	S	G	G	G	G	G	G	G	S	P	G	G	G	S	S	S	S	S	S	G	G	G	S
34	8250	G	S	G	G	G	G	G	G	G	G	G	G	G	G	S	S	S	S	G	S	G	G	G	S
35	8500	G	G	G	G	G	G	G	G	G	G	G	G	G	G	S	S	G	A	G	G	G	G	G	S
36	8750	G	G	G	G	G	G	G	G	G	M	G	G	G	S	S	S	A	M	G	G	G	S	G	S
37	9000	M	G	G	G	G	G	G	G	M	M	G	G	G	G	S	S	M	M	G	G	G	S	G	S
38	9250	G	G	G	G	G	G	G	G	M	M	G	G	G	S	S	S	M	M	G	G	G	S	S	S
39	9500	M	G	G	G	G	G	G	G	M	G	G	G	G	S	S	S	M	M	G	G	G	S	G	G
40	9750	G	G	G	G	G	G	G	G	M	G	G	G	G	G	S	S	M	M	G	G	G	G	G	S
41	10000	G	G	G	G	G	G	G	G	G	G	G	G	G	G	S	S	G	M	G	G	G	G	G	S
42	10250	G	G	G	G	G	G	G	G	G	G	G	G	G	G	S	S	G	G	G	G	G	G	G	S
43	10500	G	G	G	G	G	G	G	G	S	G	G	G	G	G	S	S	G	G	M	G	G	G	G	S

44	10750	G	G	G	G	G	G	G	G	G	S	G	G	G	G	S	S	G	S	G	G	G	G	G	S
45	11000	S	G	G	G	G	G	G	G	S	S	G	G	G	G	S	S	S	S	G	G	G	G	G	G
46	11250	G	G	G	G	G	G	G		S	S	S	G	G	G	S		S	S	S	G	G	G	G	
47	11500	G	G	G	G	G	G	G		S	S	S	G	G	S	S		S	S	G	G	G	G	G	
48	11750	G	G	G	G	G	G			S	S	S	S	S	K			S	S	G	G	G	G		
49	12000	G	G	G	G	G	G			K	S	S	S	S	S			S	S	G	G	G	G		
50	12250	G	G	G	G	G				K	S	S	S	S				S	G	G	G	G			
51	12500	G	G	G	G	G				G	S	S	S	S				G	G	G	S	G			
52	12750	G	G	G	G					G	S	S	S					G	G	S	G				
53	13000	G	G	G	G					S	S	S	S					G	G	G	G				
54	13250	G	G	G						S	S	K						G	G	G					
55	13500	S	G	G						S	S	K						S	G	G					
56	13750	S	G							S	K							S	G						
57	14000	G	G							S	S							S	G						
58	14250	G								K								S							
59	14500	G								G								G							
60	14750																								
61	15000																								
62	15246																								

Eder described NSC as 'classification-friendly', whereas SVM required a high decision level before any judgement could be made. Table 4 records the SVM findings. Once again, classification depended on vocabulary (MF1W) and the frequencies of character bigrams (MF2C) and trigrams (MF3C). All of the larger windows in the vocabulary section of Table 3 opted for Greene, and it is typical that references to Shakespeare only appeared with MF2C and MF3C. Table 3 now has a clear message.

Table 4 svm classification with various window sizes and variables

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
		Rolling Classify analysis with *svm* and window sizes between 1000 and 8000 words																							
		Window sizes																							
1	0	1000	2000	3000	4000	5000	6000	7000	8000	1000	2000	3000	4000	5000	6000	7000	8000	1000	2000	3000	4000	5000	6000	7000	8000
2	250																								
3	500	S				**mf1w**				G				**mf2c**				G				**mf3c**			
4	750	G								S								G							
5	1000	G	S							G	S							G	S						
6	1250	G	G							G	S							S	S						
7	1500	G	G	G						S	S	G						S	S	S					
8	1750	G	S	G						G	S	S						S	S	S					
9	2000	G	G	G	S					S	S	S	G					S	S	G	S				
10	2250	G	S	G	S					S	S	G	G					S	S	G	S				
11	2500	G	S	G	S	S				G	S	G	G	S				G	S	G	S	S			
12	2750	G	G	G	S	S				S	S	G	S	S				G	G	G	S	S			
13	3000	G	G	G	S	S	G			S	G	G	S	S	S			G	G	G	S	S	S		
14	3250	G	S	G	S	S	G			G	G	S	S	S	S			G	G	G	S	S	S		
15	3500	M	S	G	S	S	S	S		G	G	S	S	S	S	S		G	G	G	S	S	S	S	

16	3750	G	G	G	S	S	S	S		M	S	S	S	S	S	S		G	S	G	S	S	S	S	
17	4000	G	S	S	S	S	S	S	S	G	S	S	S	S	S	S	S	G	S	S	S	S	S	S	S
18	4250	G	S	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
19	4500	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
20	4750	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
21	5000	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
22	5250	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
23	5500	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
24	5750	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
25	6000	S	S	S	S	S	S	S	S	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
26	6250	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
27	6500	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
28	6750	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
29	7000	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
30	7250	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
31	7500	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
32	7750	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
33	8000	S	S	S	S	S	S	S	S	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
34	8250	S	S	S	S	S	S	S	S	S	G	S	S	S	S	S	S	G	S	S	S	S	S	S	S
35	8500	S	G	S	S	S	S	S	S	G	G	G	S	S	S	S	S	G	S	G	S	S	S	S	S
36	8750	G	G	S	S	S	S	S	S	G	G	G	S	S	S	S	S	G	G	G	S	S	S	S	S
37	9000	G	G	S	S	S	S	S	S	K	G	G	G	G	S	S	S	M	G	G	S	S	S	S	S
38	9250	G	G	S	S	S	S	S	S	K	G	G	G	G	S	S	S	M	G	G	S	S	S	S	S
39	9500	G	S	G	S	S	S	S	S	M	G	G	G	G	S	S	S	M	S	G	S	S	S	S	S
40	9750	G	G	G	S	S	S	S	S	G	G	G	G	G	S	S	S	S	S	G	S	S	S	S	S
41	10000	S	G	S	S	S	S	S	S	G	G	G	G	G	S	S	S	S	S	G	G	S	S	S	S
42	10250	S	S	S	S	S	S	S	S	G	G	G	G	G	S	S	S	S	S	G	G	S	S	S	S
43	10500	S	S	S	S	S	S	S	S	G	G	G	G	G	S	G	S	S	S	G	S	G	G	S	S
44	10750	S	S	S	S	S	S	S	S	G	G	G	S	G	S	G	S	G	S	S	S	G	G	S	S
45	11000	S	S	S	S	S	S	S	S	G	S	G	S	G	G	G	S	S	S	G	S	S	G	G	S
46	11250	S	S	S	S	S	S	S		S	S	S	S	G	G	G		S	S	S	S	S	G	G	
47	11500	S	S	S	S	S	S	S		S	S	S	S	G	S	G		S	S	S	S	S	G	G	
48	11750	S	S	S	S	S	S			S	S	S	S	S	G			S	S	S	S	S	G		
49	12000	S	S	S	S	S	S			P	S	S	S	G	S			S	S	S	S	S	G		
50	12250	S	S	S	S	S				S	S	S	S	S				S	S	S	S	S			
51	12500	S	S	S	S	S				S	S	S	S	S				S	S	S	S	S			
52	12750	G	S	S	S					S	S	S	S					S	S	S	S				
53	13000	S	S	S	S					S	S	S	S					S	S	S	S				
54	13250	S	S	S						S	S	S						S	S	S					
55	13500	S	S	S						S	S	S						S	S	S					
56	13750	S	S							S	S							S	S						
57	14000	S	S							S	S							S	S						
58	14250	S								S								G							
59	14500	S								S								S							
60	14750																								
61	15000																								
62	15246																								

The selection of authors and their texts in the window centroids clearly refers to Shakespeare, though Greene is not completely ignored. It may be of interest to view the SVM chart that shows the secondary attribution option.

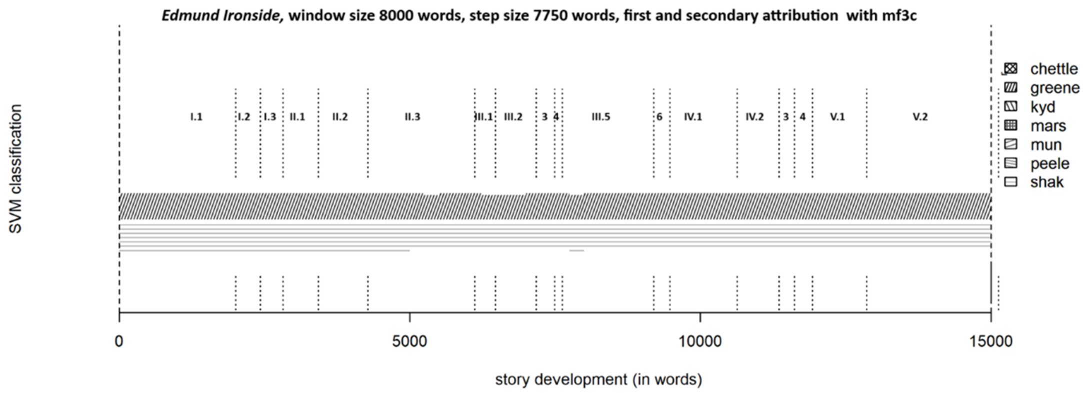

Figure 1 svm classification with mf3c and 8,000-word windows.

The first option in Fig. 1 is Shakespeare, and the second is Greene. Greene's role is significant in two ways. Firstly, it highlights the timeframe in which *Ironside* was created, which coincides with the peak of Greene's activity. It is clear that Shakespeare borrowed from Greene, as evidenced by the famous passage about the upstart crow in *Groatsworth of Wit*. However, this raises the question of what Edward de Vere would have had in common with Robert Greene. No such stylistic connection ever existed.

One would also like to know whether the General Imposters method, newly added to R Stylo, which identifies and explores similarities between two texts, arrives at an equivalent assessment. Table 4 shows the results from Jan Rybicki's unpublished script, which uses delta distances and the Růžička metric. When Kestemont et al. (2016) tested authorship verifications with the Růžička metric, they came to a clear conclusion: "Comparative evaluations across a variety of benchmark corpora show that this metric yields better, as well as more consistent results than previously used metrics" (246).

Table 5 GI table of attributions

	A	B	C	D	E	F	G	H	I	J	K
1	**delta**	low	high	**chettle**	**greene**	**kyd**	**mars**	**mun**	**peele**	**shak**	var.
2	anon_ironside	0.17	0.77	0.24	0.48	0.66	0	0.04	0	0.36	mf1w
3	anon_ironside	0.07	0.72	0.07	0.25	0.04	0.04	**0.94**	0.13	0.4	mf2w
4	anon_ironside	0.15	0.74	0	0.32	0.39	0	0.01	0.08	**0.92**	mf2c
5	anon_ironside	0.2	0.78	0	0.1	0.15	0	0	0.03	**1**	mf3c
6											
7	**wu**	low	high	**chettle**	**greene**	**kyd**	**mars**	**mun**	**peele**	**shak**	var.
8	anon_ironside	0.07	0.71	0	**0.78**	0.36	0.03	0	0	0.29	mf1w
9	anon_ironside	0.04	0.73	0.02	0.6	0.05	0.11	0.48	0.16	0.13	mf2w
10	anon_ironside	0.23	0.59	0	0.45	0.12	0.05	0.03	0.03	**0.82**	mf2c
11	anon_ironside	0	0.73	0	0.35	0.08	0.01	0	0	**0.91**	mf3c
12											

13	ru	low	high	chettle	greene	kyd	mars	mun	peele	shak	var.
14	anon_ironside	0.17	0.81	0.01	0.39	0.05	0	0.19	0.02	**0.9**	mf1w
15	anon_ironside	0.13	0.81	0	0.17	0.01	0	0.23	0.22	**0.92**	mf2w
16	anon_ironside	0.11	0.74	0.01	0.18	0.1	0.01	0.04	0.4	**0.74**	mf2c
17	anon_ironside	0.45	0.54	0.02	0.22	0	0	0.1	0.03	**1**	mf3c
18											

The table shows the cosine delta measurements, with the relevant variables indicated in column K in lines 2–5. The Würzburg Computational Stylistics Group's cosine delta measurements can be found in lines 8 to 11. Lastly, the Růžička metric results (RU) can be found in lines 14 to 17. Columns B and C show the lower and upper limits of the area of insecure attributions. Values above the upper limit are indicative of authorship. These values are printed with a dark background. However, the second-lowest values, printed in light grey, often denote Greene, who does not qualify as an author but can be considered an originator from whom Shakespeare drew inspiration.

Conclusion

In his 1592 pamphlet, 'A Groat's Worth of Wit Bought with a Million of Repentance', Robert Greene warns his fellow playwrights about Shakespeare:

'There is an upstart crow beautified with our feathers, who, with his tiger's heart wrapped in a player's hide, dares to bombard you with blank verse as well as the best of you. And, being an absolute jack-of-all-trades, he imagines he is the only Shake-scene in the country.'

This passage is widely recognised as the earliest surviving reference to Shakespeare as a playwright. If *Edmund Ironside* bears traces of Greene’s style, thus confirming their relationship as lender and borrower, then we have clear proof that *Ironside* is one of Shakespeare's early history plays. Conversely, Greene's status as the first professional writer rules out Edward de Vere as a rival. Instead, 'his tiger's heart wrapped in a player's hide', a phrase taken from *3 Henry VI*, suggests that Shakespeare began his career as an actor. It is this theatrical background that shaped his writing. Regarding the date of *Ironside*, Sams clarifies that, by the time Shakespeare had been in London since the mid-1580s, a fight between clerics on stage would not have been permitted after 1587.

Notes

(1) John Burrows introduced delta in 2002 as a key method for measuring stylistic differences between texts, particularly in terms of the frequencies of function words and their z-scores in a target text and a group of reference texts. The idea is that stylistic similarity can be measured as the average difference in standardised word frequencies. Burrows initially used function word frequencies (mf1w), but tests carried out by Grieve (2007) and others proved that character 3-grams (mf3c) were more effective. Additionally, Hoover (2004) found that a culling value of 70% produced the best results by excluding idiosyncratic and specific

vocabulary. The main distinction in the R Stylo culling process is that 0% examines up to 1,000 variables and 100% only considers those present in all texts. This had a disappointing effect in small windows, where only twelve or fifteen variables were expected to determine questions of authorship. Culling at 70% called up variables present in 70% of the reference texts, thus excluding those that would not have played a role anyway. In his extensive study, Eder (2015) warned against using samples that were too small. For this reason, attributions were tested using a proven window size of 4,000 words. The window centroids were moved through the whole text at 250-word intervals. While delta could only measure whole texts, rolling delta recorded stylistic changes and potential collaborations.

(2) To detect potential shifts in authorship, I used the Rolling Classify method implemented in the Stylo R package (Eder, 2015). This is a supervised stylometric technique that segments a text into overlapping windows of various sizes, comparing each window to a reference corpus of known authors. Each segment is then classified based on its proximity to the stylistic profiles of the candidate authors using distance measures such as the nearest shrunken centroid (NSC) and the support vector machine (SVM) method. This method can identify internal stylistic variation and has proven especially effective at detecting collaborative authorship, revisions or later interpolations. Unlike Rolling Delta, which is unsupervised, Rolling Classify depends on a predefined set of authors and provides a sequence of attributions throughout the text.

(3) Unlike traditional classification methods, which assign a text to one of several potential authors, the General Imposters method determines whether two texts are stylistically similar enough to have been written by the same author. Rather than using all available features (e.g. the most frequent words), the method repeatedly selects random subsets of them. This creates many small, slightly different stylistic profiles. For each random subset, the distance between the target text and the known reference texts is measured using Burrows' Delta (delta), the

Cosine Delta metric of the Würzburg Computational Stylistics Group (wu) and the Růžička metric (ru). This process is repeated many times (hundreds or thousands of iterations). Each iteration contributes to a vote on whether the target text is closer to the author in question than to the others. The final score is based on how often the candidate author 'wins' these comparisons across iterations. A high frequency of wins suggests stylistic consistency between the questioned and known texts. Eder emphasises that this method is robust against overfitting and noise because it relies on random sampling and repeated testing rather than a single static feature set. It is seen as a secondary verification system.

References

Burrows, J.F. (2002). 'Delta': a measure of stylistic difference and a guide to likely authorship. Literary and Linguistic Computing, **17**(3): 267–287.

Eder, M. (2015). Does size matter? Authorship attribution, small samples, big problem. Digital Scholarship in the Humanities, **30**(2): 167-182,

Eder, M., Kestemont, M. and Rybicki, J. (2016). Stylometry with R: A package for computational text analysis. *R Journal*, **16**(1): 107-121, https://journal.r-project.org/archive/2016/RJ-2016-007/index.html.

Eder, M. (2018, May 30). Authorship verification with the package 'stylo'. Computational Stylistics Group Blog. https://computationalstylistics.github.io/blog/imposters/

Everitt, E. B. The Young Shakespeare: Studies in Documentary Evidence. Anglistica 2. Copenhagen: Rosenkilde & Bagger, 1954.

Greene, Robert. (1592). A Groatsworth of Wit, bought with a Million of Repentance. In The Dramatic and Poetical Works of Robert Greene and George Peele, edited by Alexander Dyce, 229–247. London: Routledge, 1861.

Grieve, J. (2007). Quantitative Authorship Attribution: An Evaluation of Techniques. Literary and Linguistic Computing, 22(3): 251–270.

Hoover, D.L. (2004). Testing Burrows's Delta. Literary and Linguistic Computing, 19(4): 453–475.

Jiménez, Ramon L. (2003). Edmond Ironside, The English King: Edward de Vere's Anglo-Saxon History Play, The Oxfordian, vol. VI: 7-27.

Kestemont, M. et al. (2016). Authorship verification with the Ruzicka metric. In: Digital Humanities 2016: Conference Abstracts. Kraków: Jagiellonian University & Pedagogical University: 246–249. Available at: [Accessed 27 May 2025].

Sams, Eric. (1986). Shakespeare's "Edmund Ironside": The Lost Play. Wildwood House.

Sams, Eric. (1995). The Real Shakespeare: Retrieving the Early Years, 1564-1594, New Haven, London: Yale.