Stylistic palimpsests: Computational stylistic perspectives on precursory authorship in Aphra Behn's drama

Mel Evans (iD) and Alan Hogarth

School of Arts, University of Leicester, Leicester, UK

Abstract

Computational stylistics has developed various methods for investigating and attributing authorship of collaborative literary texts. This article investigates 'precursory authorship' (Love, 2002): that is, the authorial traces of a source text that inform—to a greater or lesser degree—a subsequent literary output, in order to establish its relevance for our approach to and understanding of the linguistic properties of literary style. Precursory authorship and derivative adaptations are common features of early modern English drama, and the study focusses on two case studies relating to the plays of Restoration playwright, Aphra Behn (c. 1640–89). Using a combination of quantitative methods (Rolling Delta (RD), principal components analysis (PCA), Delta, and Hierarchical Cluster Analysis), the investigation highlights the presence of precursory authorial style in Behn's *The Rover* and an anonymous work associated with Behn, *The Counterfeit Bridegroom*. The results suggest that precursory authorial style is identifiable in both cases, not only through a similarity with the source text but, to a lesser degree, other texts by the precursory author as well. The anonymous play yields complex and non-confirmatory evidence for Behn's authorship. Methodologically, RD is most sensitive to precursory collaboration. Collectively, the findings highlight the importance of stylistic factors when describing and interpreting literary linguistic quantitative data: precursory authorial style is another facet that intersects with properties such as time period and genre. The article urges a more critical and theoretically informed view of authorially aligned linguistic style.

Correspondence:
Mel Evans, School of Arts,
University of Leicester,
Leicester, UK.
E-mail:
mel.evans@le.ac.uk

1 Introduction

One of the substantial challenges for computational authorship attribution is how to pinpoint the contributions and, in some cases, identities of multiple individuals who 'author' a single text. This problem is pervasive in literary studies, whereby collaboration has long been the norm for the creative process, rather than the exception. The topic has provoked especial debate in relation to early modern English dramatic works, particularly those of Shakespeare, in which critique has concentrated on the validity of computational stylistic methods for the identification of collaboration, and the implications of the findings for understandings of authorial style, influence, status, and literary worth (e.g. Freebury-Jones, 2017; Gladwin *et al.*, 2017;

Ilsemann, 2018; García-Reidy, 2019). This article focusses on a lesser studied, but no less significant, kind of collaborative praxis, in which one author adapts, edits, and/or emends a source text from an earlier period: what Love (2002, pp. 40–3) calls 'precursory authorship', and what can be conceived as a kind of asynchronous collaboration. The processes of temporally distant derivation and adaptation sit alongside the synchronic, sustained, and interactive collaborations more commonly studied, in which authors work together to produce a single text for shared financial and cultural reward. However, the derivation of literary works from pre-existing sources is a practice that was pervasive in early modern theatrical writing and therefore warrants more substantial attention from a computational stylistic perspective.

In this article, we offer two case studies that represent exploratory steps in the investigation of precursory authorship. Our objective is to evaluate the extent to which a source text is traceable in the lexical make-up of its derivative successor, based on the stylistic features typically considered in computational stylistic studies of authorship. Three questions shape our discussion: (1) Is precursory authorship identifiable in a given text? (2) What methods best account for this kind of collaboration? and (3) To what extent does precursory authorship impact the stylistic features associated with the (presumed) adaptor? The two case studies focus on the dramatic works of Aphra Behn (c. 1640–89), an English Restoration writer and one of the first women to make a living from her pen. Behn's dramatic canon comprises sixteen plays of confident attribution, plus works of more questionable provenance. Her most famous, securely attributed play, *The Rover*, is in fact a reworking of Thomas Killigrew's unperformed mid-century comedy *Thomaso*—its adaptive status was recognized (and criticized) by Behn's contemporaries. This work therefore offers an explicit test case for investigating the stylistic traces of precursory authorship. The second case study investigates a work with a less secure attribution to Behn, *The Counterfeit Bridegroom (CB)*, published in 1677. The play is closely sourced from Thomas Middleton's play *No Wit, No Help like a Woman's* (first performed in

1638), and thus the analysis must address the potential stylistic influence of the source author, Middleton, alongside the inquiry into Behn's potential authorship of the Restoration play.

Like the medieval manuscript palimpsest, in which the 'original text has been effaced or partially erased, and then overwritten by another' (OED Online), 'precursory' collaborative play texts are problematic, yet of immense interest for computational stylistic approaches to authorship, because they comprise a series of stylistic layers representing the characteristics of multiple authors writing in different periods. Our case studies of these Restoration theatrical palimpsests (following Harbage, 1940) show that precursory authorship complicates (further) our understanding of the authorial signal. Our investigation suggests that different lexical measures and statistical tests show different sensitivities to the stylistic contributions of a source text in its derivation, and that in some circumstances, the stylistic preferences of the source text's author are evident, and even prominent, in the newer text under analysis, and that this needs to be taken into account when attributing the authorial identity of the later adaptor.

2 Computational Perspectives on Authorship and Collaboration

Computational stylistics, or stylometry, identifies linguistic patterns across (primarily) literary texts, using a range of computational techniques that enable 'distant reading' and 'macroanalysis' (Jockers, 2013; Eder *et al.*, 2016, p. 108) to answer broad authorial, thematic, and cultural questions. The approach expands the scope of conventional literary textual analysis, which traditionally uses 'close reading' techniques focussed on a limited set of texts for salient features, through the analysis of large textual datasets for patterns of similarity and difference that are invisible to the eye of the reader (e.g. Underwood, 2019). The two approaches are not, of course, mutually exclusive, but rather offer a set of complementary tools for the interrogation of literature and its contexts.

Authorship attribution has been a primary focus of computational stylistics. Underlying this approach

is the idea that each individual has a unique 'linguistic fingerprint' (in linguistics, known as an idiolect), although the distinctiveness of this fingerprint is not equal across individuals, weakening the validity of the metaphor somewhat (see van Halteren *et al.*, 2005). Nevertheless, the term captures the persistent findings that linguistic features are often authorially aligned, demarcating the works of one writer from another. The analysis of authorial style needs also to account for the other situational and contextual factors that shape how an individual uses language (Kestemont, 2014, p. 64). No authorial signal is an island; it operates within, and in relation to, macro-level factors, such as genre, time, and gender (Rybicki, 2016; Weidman and O'Sullivan, 2018). Computational stylistic methods typically seek to cancel out these confounding variables to provide a honed measure of a text that, by the best means possible, reflects the idiolectal distinctiveness of a text's author. Quantitative 'statistically-focused' (Auerbach, 2018, p. 2) authorship attribution conventionally concentrates analysis around high-frequency linguistic features such as the most frequent function words (MFFW) and the most frequent words (MFWs; comprising function and content words) (see summary in Stamatatos, 2009). Such items have the benefit of high frequencies and a balanced distribution, and a relative independence from content and topic (cf Vickers, 2018). Word frequencies in a reference corpus of works by known authors and the questioned text are processed using either descriptive statistical methods such as hierarchical cluster analysis and principal components analysis (PCA), or machine-learning classifiers, such as support vector machines (see Stamatatos, 2013).

Quantitative authorship attribution has attracted a steady stream of criticism, constructive and otherwise, arguably because the approach poses ideological challenges to the qualitative, external, and internal evidential processes of attribution traditionally used in literary studies. Yet lexical frequencies have proven to be a reliable delimiter of style (in carefully controlled settings, at least). What is perhaps needed, which has also been recognized in the parallel discipline of forensic linguistics, is for proponents of computational approaches to authorship to start to broaden the discussion: not only in terms

of the features of analysis, but also their theoretical framing and our understanding of their effectiveness (see discussions in Wright, 2017; Underwood, 2019). The present discussion sits within this emerging context, seeking to open up the interrogation of authorship and style to explore the confluence of factors, from the local (micro-level) to the social (e.g. genre), the temporal and the authorial (macro-level), that shape the linguistic and stylistic properties of a text and texts.

3 Collaboration

Elizabethan and Jacobean plays were produced in a fundamentally collaborative culture (see discussions in Stallybrass, 1992; Orgel, 1992). Orgel (1992) proposes that the authority of a play resided with theatre companies, rather than those who were commissioned to write them. Companies 'usually stipulated the subject, often provided the plot, [and] often parcelled it out, scene by scene, to several playwrights. The text thus produced was a working model, which the company then revised as seemed appropriate' (Orgel 1992: 84). Brown, in his 2017 PhD thesis, offers a narrower view, whereby authors had greater autonomy, and thus collaboration should be considered in the terms of co-authorship undertaken for financial reward. Brown's survey of early modern play databases indicates that a quarter of plays performed between 1567 and 1642 (so earlier than those considered here) are recorded as collaborative; a proportion lower than the 50% traditionally accepted in earlier, twentieth-century scholarship. Interestingly, the proportion is higher in lost plays (33.2%), possibly suggesting that the prestige of a work, and therefore its likelihood of preservation, was enhanced through the 'respectability' of authorship (Brown, 2017: 177).

In terms of the mechanics of collaboration, multiple authors could work together in diverse ways, including 'prior agreement on outline, vetting of successive drafts by a partner, composition in concert', and 'brief and possibly infrequent intervention' (Zitner, 1984, p. 10). Potter (2008, p. 5) identifies differing co-authorship practices,

sequential and concurrent, to describe the probable 'workflow' process of theatrical collaboration in the period. These practices of collaboration are broadly synchronous, focussed upon the creation of a text for a specific time and place. Alongside this kind of textual authoring process is what Love (2002, pp. 40–3) calls 'precursory authorship', in which a text by one author informs, through direct replication, paraphrase, or with a more abstract, narratological or topical derivation, the creation of another. This derivative kind of authorship, the precursory collaboration, is usually asynchronous, in that the authors of each text are not working in the same time and place on the same textual output.

While Shakespeare and his contemporaries drew without reservation on the works of their predecessors, the Restoration saw a change in how precursory collaboration was evaluated. Plays were typically published anonymously in the decades leading up to 1600, but by the end of the seventeenth century an author's name on the title page, and a stable system of financial reimbursement for their authorial efforts, was the norm (Kewes, 1998, pp. 225–6). By the 1690s, a play was understood to be 'a product of an individual imagination' (Kewes, 1998, p. 228). Consequently, synchronous collaboration between two or more playwrights working on the same theatrical output became a more overt and demarcated affair, with care taken to specify the various authors' involvement (Kewes, 1998, p. 228; see discussion in Erne, 2013, pp. 60–8). Precursory authorship also became more contentious. This can be seen in the fact that Behn was lambasted for the perceived plagiarism in her best-known play *The Rover*, which is heavily indebted to Thomas Killigrew's ten-act play *Thomaso* (see Hobby, forthcoming). Behn (1677) challenged these accusations in the published edition of the play, asking the reader to make up their own mind about the work's originality.

Several computational stylistics methods have been developed for cases of (synchronous) collaboration (e.g. Rybicki *et al.*, 2014; Gladwin *et al.*, 2017). One specific method is Rolling Delta (RD) (Burrows, 2010; Eder, 2016), which was developed as an alternative to the 'bag of words' comparison of texts. RD uses Burrows's (2007) Delta method to

identify the MFWs in a dataset, but unlike the bag of words model, the results map the degree of similarity between candidate authors and the dubious text over its course, using overlapping sequential slices e.g. 5,000-word segments created every 1,000 words. According to Eder, the method allows for the detection of 'periodic regularities in the time series on the one hand, and possible disturbances or local idiosyncrasies on the other' (Eder, 2016, p. 458). RD has been enthusiastically taken up to investigate collaboration (e.g. Ilsemann, 2016, 2018; Schöberlein, 2016), although caution has been advised. For example, Ilsemann (2018) claims that the technique can reliably attribute to Christopher Marlowe only two out of seven plays that traditionally form his dramatic canon; this has been rigorously contested (Barber, 2019). The caution reflects the fact that, as with any lexical measure, the Delta statistical method 'is genre sensitive as well as author sensitive and unsupported results should not be taken either as conclusive or as purely authorial' (Craig and Burrows, 2012, p. 36). When run through the package Stylo for R, RD is also less explicit in its workings than other statistical analyses. For this reason, the analyses conducted here utilize David Hoover's (2019) Excel macros for Delta, configured to provide a rolling analysis using the Java software Intelligent Archive (Craig, 2019), and following Burrows's (2010) original model for the approach. As is discussed in more detail below, the results obtained are not entirely dissimilar to those produced using Stylo for R, but the greater control over the calibration of the wordlist and culling criteria produces more consistent results.

Regardless of the method of implementation, it is imperative that RD results are corroborated with the findings of other tests (as recently demonstrated by Gladwin *et al.*, 2017; see also Burrows, 2010, pp. 30–3). Our work, therefore, builds on the triangulation principles of previous studies of collaboration, through a comparison of the same dataset using three established, descriptive statistical treatments: RD, PCA, and Delta (whole texts). The results highlight the different sensitivities of these tests to precursory authorship, and the different behaviours of the individual texts included in an authorial sample (despite macro-level variables, such as 'dramatic

comedy' being held constant). In the two case studies, the source text and the authorial style of that source are evident to different degrees across the three test results (Gladwin *et al.*'s study of twentieth-century prose found a similar variability in their study of synchronic collaboration). Although the results exemplify Burrows' admission that computational stylistics 'cannot offer certainty' (2010, p. 35), we suggest that such evidence offers a rich basis to further develop the methodological and theoretical frameworks for the description and interpretation of authorial style, and thus provide a stronger gauge of exactly how confident we can be in interpreting computational stylistic findings.

4 Aphra Behn and Restoration Drama: Datasets

Aphra Behn (c. 1640–89) is recognized as one of the first professional female English writers, producing works of drama, poetry, and prose fiction over two decades (1670–89). She is perhaps best known for her work of short prose fiction, *Oroonoko* (1688). Sixteen plays are securely attributed to Behn. *The Rover* (1677), which follows the romantic liaisons of Willmore, the eponymous hero, and his acquaintances around Naples at carnival time was the most popular work during her lifetime, and continues to be staged today (e.g. RSC 2016-7). *The Rover* draws on Thomas Killigrew's unperformed comedy *Thomaso*, published 15 years earlier. As our first case study shows, Killigrew's play, and other examples of his work, shows similarities with Behn's *The Rover* at a quantitative lexical level.

There are also five other plays from the period that have a more speculative association with Behn, lacking external verifiable evidence such as her name on the title page, and/or a lifetime date of publication. One such work, *The Counterfeit Bridegroom* (1677) (henceforth CB), represents a particular challenge because of the combination of its precursory authorship and source text, and the uncertain identity of its Restoration adaptor. The play was not attributed to Behn in print until the nineteenth century. Its plot and dialogue are derived (to varying degrees over the course of the text) from the

Thomas Middleton play *No Wit, No Help Like a Woman's* (henceforth *No Wit*) (perf. 1638). Our stylistic analysis does indeed identify Middletonian traits (the precursory author), elements that affect the stylistic profile of the play and thus our ability to assess its similarity to Behn's dramatic style.

The investigation uses a corpus of dramatic texts published between 1660 and 1710, taken from Early English Books Online - Text Creation Partnership (EEBO-TCP) and Visualizing English Print. These include Behn's dramatic works, the dubia, and a corpus of Restoration plays (comprising comedies and tragedies). All texts are prepared using Text Encoding Initiative - Extensible Markup Language (TEI-XML), demarcating genre features, such as speech prefixes and stage directions, from the dialogue, with spelling regularized using VARD 2.4 (Baron, 2017). The texts were checked and proofed manually. Further information about the corpus can be found on the project website (www.aphrabehn. online).

5 Case 1: *The Rover* and Killigrew's *Thomaso*

The Rover was first published anonymously in 1677, although editions the following year appeared with Behn's name on the title page. Thomas Killigrew's *Thomaso* was probably written around 1654 (Hobby, forthcoming). Behn and Killigrew were acquaintances of the London theatrical world, as well as having a previous professional relationship during Behn's 1666 activities as a Royalist spy (Todd, 2017); Todd suggests that Behn may have acted as amanuensis for Killigrew's original play, although it seems more likely that Behn read the work in the 1664 publication of Killigrew's collected drama (Hobby, forthcoming). Behn's contemporaries certainly recognized the similarities between the plays—similarities that, as well as parallels in plot, character, and dialogue, included unchanged street names, despite Behn relocating Killigrew's play from Madrid to Naples (Hobby, forthcoming). As noted above, changing attitudes towards authorship and appropriation meant that Behn's reputation was attacked, prompting a vigorous defence by

the author in the play's published postscript: 'the Plot and Bus'ness (not to boast on't) is my own: as for the Words and Characters, I leave the Reader to judge and compare 'em with Thomaso' (1677, p. 85). Perceptions of female professionalism may have informed the accusations; Killigrew himself drew on other plays in his work, which went without comment (Hobby, forthcoming).

Hobby notes that Behn must have worked closely with Thomaso, observing that '[Behn] combines – sometimes within a single speech – materials drawn from different sections of this ten-act play'. She expands 'the size and complexity of the women's parts' and 'cuts and tightens plots and dialogue' (Hobby, forthcoming). Behn also draws on Brome's Novella, which was itself a source for Killigrew's play, although the relationship to this work is less direct. Behn's borrowing of 'many snatches of dialogue' (Hobby, forthcoming) means that Thomaso is the most immediate source text, with Killigrew having a potential role as a precursory author. In principle, the precursory role of Thomaso in Behn's The Rover should mean that the two plays show greater stylistic similarities than a comparison of Behn's play with Killigrew's other works and, more generally, with works by any author other than Behn. The impact of literary derivation and adaptation on lexical features, such as MFWs, typically considered in computational stylistic authorship attribution, has not been extensively studied, and our findings here provide some evidence for how the most frequent, typically function, words distribute and align with different authorial contributors in cases of precursory authorship.

5.1 Rolling Delta

One theme emerging out of the two case studies is how the identification of the precursory author and the source text differs according to the descriptive statistical test applied. Rolling Delta (RD) appears to be the most sensitive to precursory style; this is a finding that makes sense, in light of the test's focus on style identification in collaborative texts at a fine-grained level. The visualization of the results plots each candidate authors' texts across the x-axis, which represents the derivative text (e.g. The Rover) in words. The y-axis displays the Delta z-score values. The lower the score, the greater the stylistic similarity with the derivative text, which may suggest an author's involvement with the work in question. Three authors are included in the tests: Behn and Killigrew, as author and precursory author, respectively, and another Restoration author, Thomas D'Urfey, who provides a point of comparison with Behn's dramatic style, and therefore may highlight possible temporally linked (rather than authorially linked) differences with the precursory author and source text.

The tests use 5,000-word segments, overlapping every 2,500 words, providing ten segments of The Rover for analysis. The iterations of the RD tests follow those preprogrammed in Hoover's (2019) workbook macros, which tests the MFWs in the given text set. The tests analyse 100–1,000 MFW at 100-word intervals, using different percentages of culling and the inclusion/exclusion of pronouns. Culling is of particular interest when investigating precursory authorship. It considers the distribution of words in the texts in the corpus as a factor in their inclusion in the wordlist, rather than selecting words on the basis of frequency alone; frequency-based selection can potentially include words over-represented in one text, therefore skewing results. This is relevant for the analysis of precursory authorship, in which the test text is derived from one particular source. Hoover's workbooks define culling as a threshold for which the number of occurrences of a word occurs proportionally in a single text: e.g. if more than 60% of instances of 'Willmore' occur in one text (i.e. The Rover) then it will not be included in the wordlist. Hoover (2019) notes that a threshold of 60–80% culling appears to provide the most accurate results. In the present case study, the impact of culling in the exploration of precursory authorship appears to be one of degree rather than difference: tests with no culling have a more exaggerated depiction of similarity between a given play and The Rover than those with 60 or 80% culling.[1] In the straight full-text Delta analyses, reported in the subsequent section, the culled iterations are less likely to identify the precursory author as the main stylistic likeness: a finding that (albeit logical) highlights the impact

of the statistical approach used when exploring facets of authorial style in a text.

To interpret and visualize the RD results, several strategies are available. One approach is to look at the ranked list of the comparison plays' segments per segment of *The Rover* and identify the lowest Delta score for each author included in the test, with no attention paid to the specific text. This gives a general impression of the maximum authorial likeness between each author and each segment of the test text. Another option, which highlights the variable stylistic profiles of play texts relevant to questions of precursory authorship, is to map the lowest score achieved by any segment representing each play included in the test: in effect, providing the maximum text-specific likeness with *The Rover*. Other options include averaging the results for all segments of the authorial corpus or a play, or running a strict comparison of scores between Segment 1 of play text and test text, Segment 2 of play text and test text, and so on. Given that the process of precursory authorship may involve a more diffuse or selective acquisition of a source text's linguistic content, the 'maximum likeness' measure (by author and by play segments) is used in the following visualizations.

Figure 1 shows the authorial profiles for Behn, Killigrew, and D'Urfey based on the RD comparison with Behn's play *The Rover*. For clarity, clustered bars are used to represent the scores for each author per segment. The results suggest that there is a relatively strong similarity between Killigrew and *The Rover* in the first four segments of the play, and particularly in the third segment. This stylistic closeness is far less evident in the latter half of the play. Behn's profile, however, shows a relatively consistent likeness throughout.

No culling provides the greatest distinction between the three authors, and the strongest pull between Killigrew and *The Rover*; this likely reflects topic-related lexis as well as features of less conscious authorial preference (i.e. function words). In Figure 2, which includes words for which no more than 60% examples occur in a single text, the difference is reduced. However, there is still a strong similarity between Killigrew and *The Rover*, with Killigrew showing maximal scores comparable

with those of Behn for the first part of the test text. In the latter half of *The Rover* the likeness diminishes, with Killigrew's scores now higher, that is, less similar to *The Rover*, than those of Behn's contemporary, D'Urfey.

A finer-grained perspective reveals the role of the source text, *Thomaso*, in the authorial likeness, when compared with other texts by Killigrew (Fig. 3): segments from *Thomaso* are consistently more similar to *The Rover* than other works by Killigrew in the test. The distinction between Killigrew and the other authors is perhaps not as pronounced as might be expected given the temporal gap between works: Killigrew's comedy *Parson's Wedding* is no less like *The Rover* than D'Urfey's comedies for Segments 1–5, for example. The decrease in similarity between *Parson's* and *The Rover* for the latter half mirrors the trend for *Thomaso* which suggests that the RD test may, in fact, be picking up on precursory authorial similarities outside those appertaining to the source text. By comparison, the third Killigrew play, *Bellamira* is least like *The Rover* of all of Killigrew's works analysed; the play is a tragedy and this is likely a strong factor in its stylistic profile.

Of course, the lexical frequencies are derived only from the texts included in each test, and thus can only indicate the relative likenesses within that dataset. This point is made by Burrows (2010, p. 30) but is one that bears repeating. This is especially pertinent in the second case study, as will be seen. However, before turning to this, it is useful to consider how the 'bag of words' tests, PCA and Delta, compare with RD in identifying a stylistic relationship between *The Rover* and its source text and precursory author.

5.2 PCA

PCA offers an indicative, descriptive statistical profile of the main stylistic dimensions of a dataset, with a long tradition in authorial analyses (e.g. Binongo, 2003). One limitation of PCA in this context is that the analyst must identify what the components represent. Da (2019) suggests that the absence of contextual information may lead to a tendency for scholars to 'overfit' their interpretation in terms of preconceived stylistic variables, such as
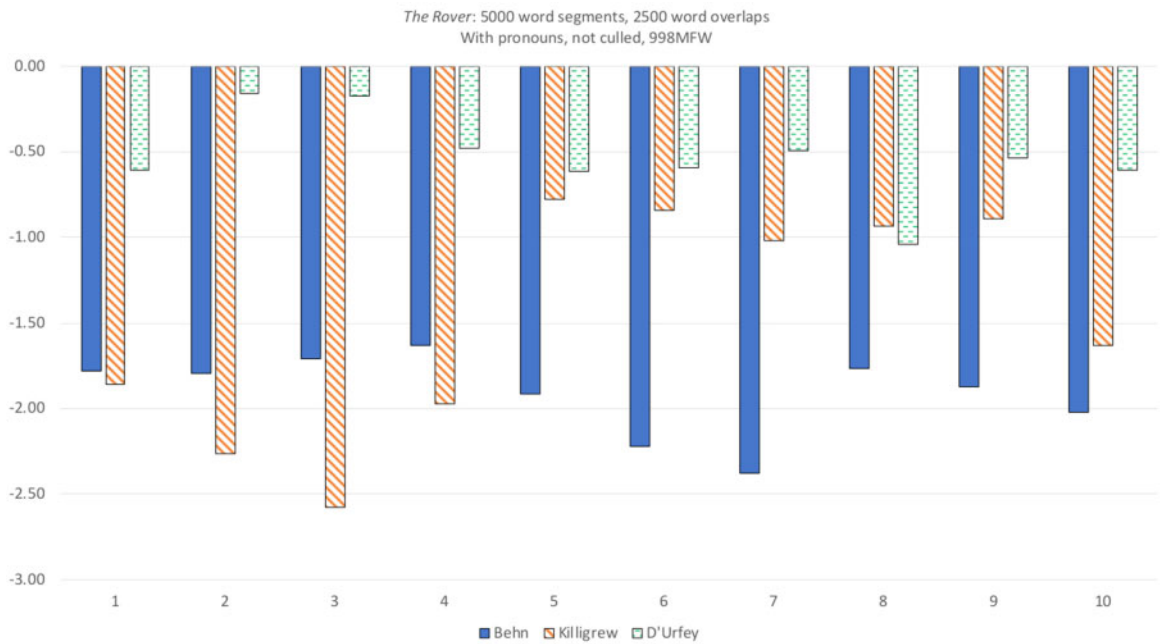
The Rover: 5000 word segments, 2500 word overlaps
With pronouns, not culled, 998MFW

Fig. 1 Comparison of Behn, Killigrew, and D'Urfey with *The Rover*: based on lowest Delta score in each test. With pronouns, no culling, 998 MFW
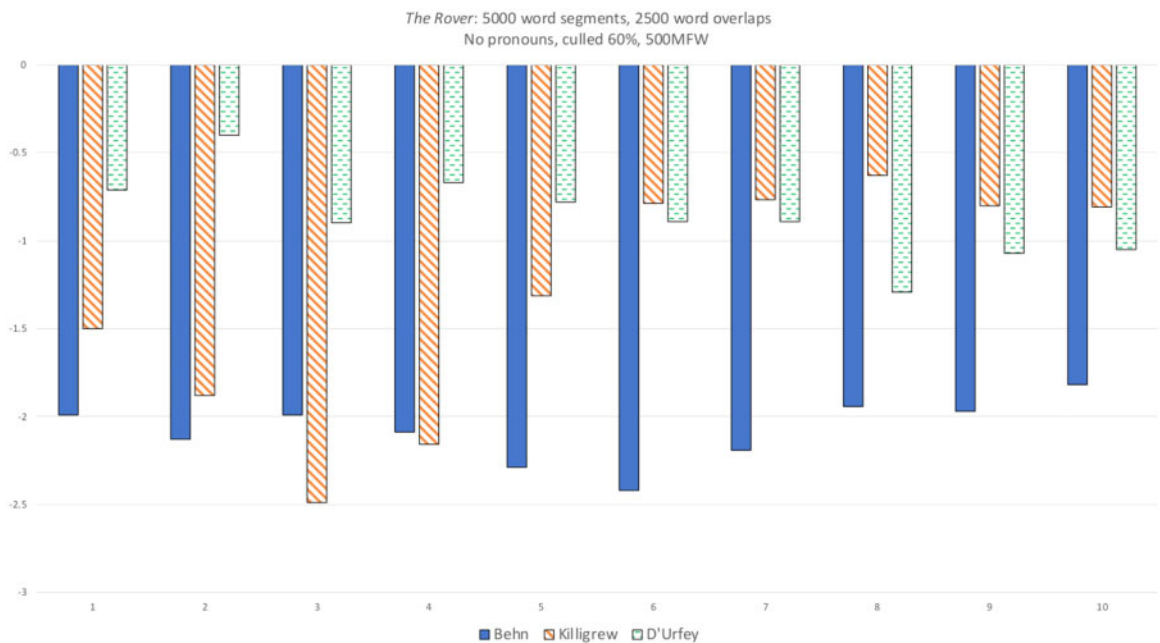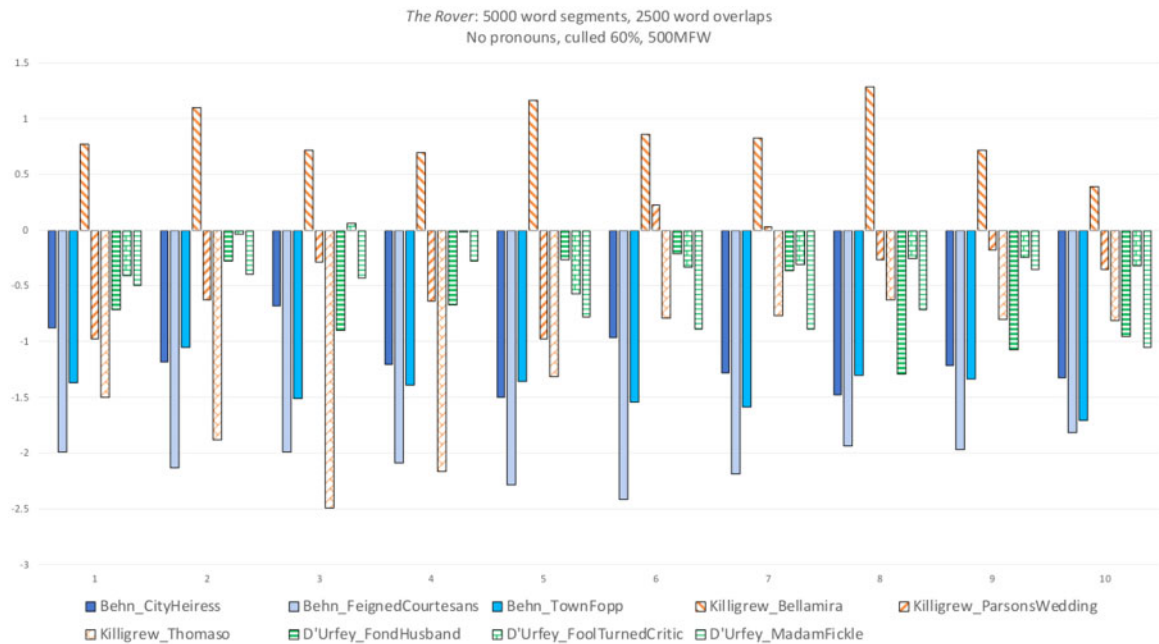

The Rover: 5000 word segments, 2500 word overlaps
No pronouns, culled 60%, 500MFW

Fig. 2 Comparison of Behn, Killigrew, and D'Urfey with *The Rover*: based on lowest Delta score in each test. No pronouns, culled 60%, 500 MFW

Fig. 3 Comparison of Behn, Killigrew, and D'Urfey by play with *The Rover*: based on lowest Delta score in each test. No pronouns, culled 60%, 500 MFW

genre. While this is true to a point, PCA nevertheless provides a useful perspective on the linguistic properties of a dataset; interpretations can be reconciled, or rejected, when other evidence is taken into account (see Burrows, 2010, p. 29).

The PCA comparing Behn and Killigrew's plays uses six Behn plays from the first half of her career (*The Forced Marriage* in 1670 to *The Feigned Courtesans* in 1679) and three plays by Killigrew from the 1650s and 1660s. As in the RD tests, the chosen texts aim to minimize chronological variation within and between the texts analysed (cf Rybicki, 2016). The texts are split into 5,000-word segments to provide a more granular perspective on their stylistic properties, and to provide a rough parallel with the RD segmentation. Figure 4 shows the results for PCA using the 100 MFW. As can be seen, the PCA divides Behn's plays (circles) from those of Killigrew (crosses); a division that could arise from differences in authorial style, and/or temporal variation in dramatic trends. Adding *Rover* and *Thomaso* to the PCA, using this pre-established 100 MFW, means that the lexical

criteria, which are known to distinguish Behn and Killigrew, can be used to profile the two plays and their distribution interpreted accordingly. There is only slight evidence to suggest that Killigrew's authorial signal is evident in the distribution of high-frequency lexis in *The Rover*; one segment of the play (the first 5,000-words of the play, in fact) is positioned closer to the segments from Killigrew's plays than is typical of Behn's other works included in the test. However, there is nothing that indicates that *The Rover* and *Thomaso*, specifically, are stylistically alike to the degree suggested by the RD tests. Comparable distributions are produced using the 100 most frequent function words (MFFWs), and the 100 most frequent 2-grams.

Expanding the MFW list, such as to analyse the 500 most frequent 2-grams in the works of Behn and Killigrew, produces a similar picture (Fig. 5), with no strong evidence for a stylistic overlap between *The Rover* and *Thomaso*. Notably, the genre-related stylistic differences, previously identified in the RD tests for Killigrew's plays, also emerge in this
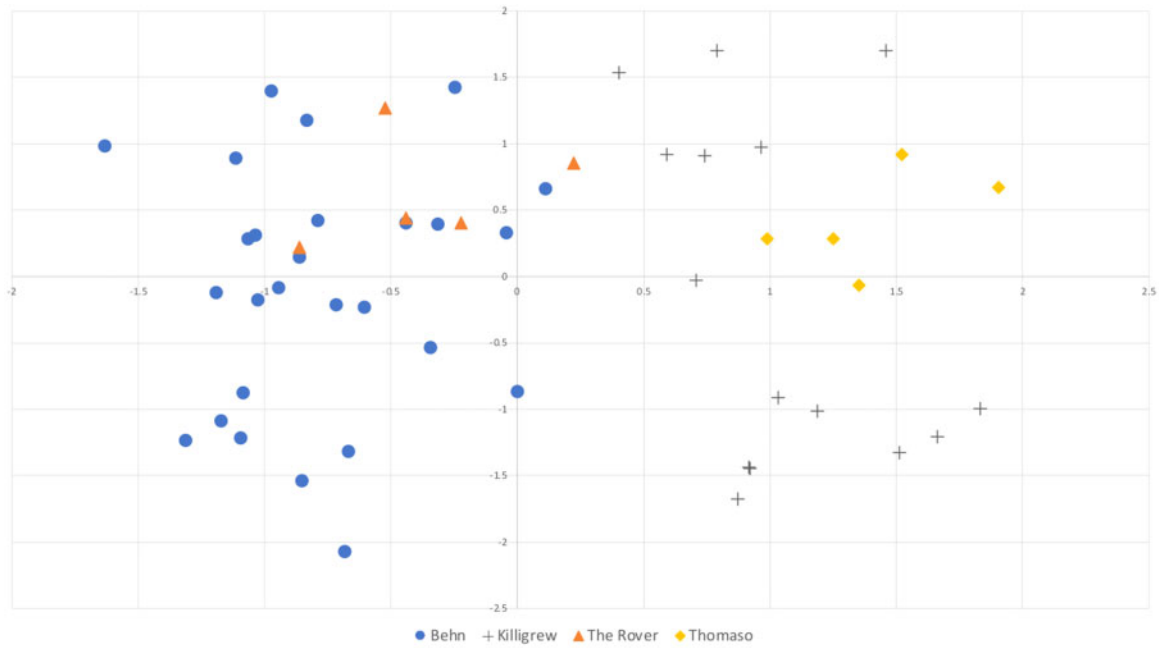
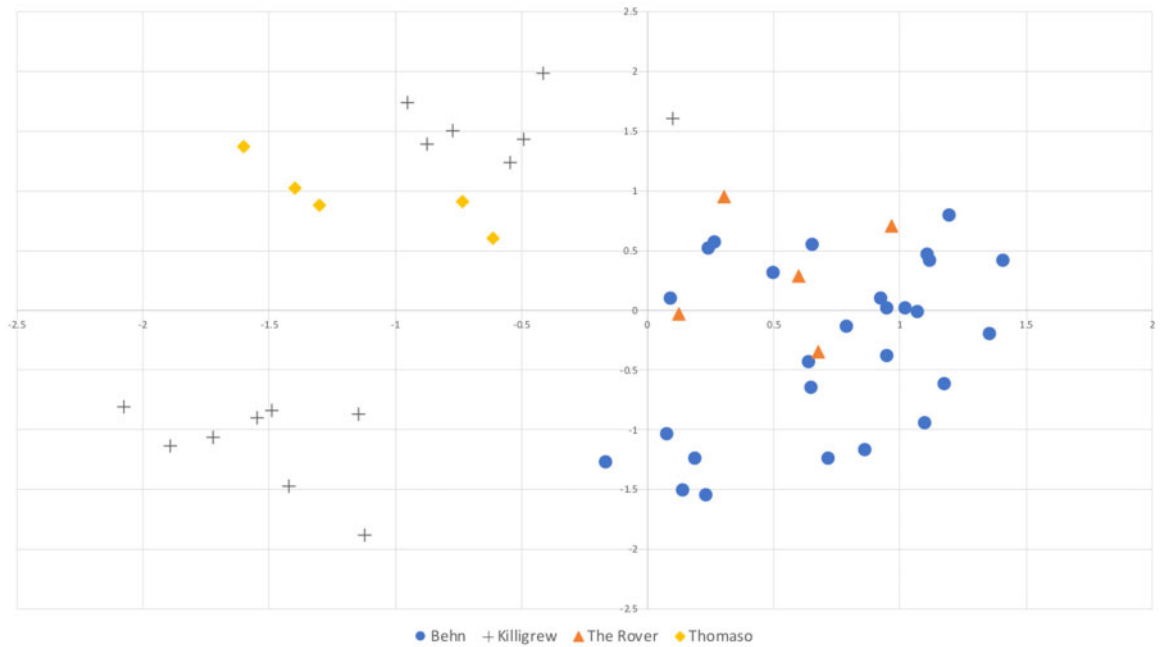**Fig. 4** PCA of a sample of plays by Behn and Killigrew using 100 MFW. *The Rover* and *Thomaso* added

Fig. 5 PCA of a sample of plays by Behn and Killigrew including *The Rover* and *Thomaso* using 500 most frequent 2-grams

Table 1 Results for Delta analysis, comparing *The Rover* with plays by Behn, Killigrew, and D'Urfey

The Rover (1)	Delta *z*-scores
Behn_CityHeiress	−1.16
Behn_TownFopp	−1.02
Behn_SirPatientFancy	−0.88
Killigrew_Thomaso	−0.40
Killigrew_ParsonsWedding	−0.09
D'Urfey_MadamFickle	0.36
D'Urfey_FoolTurnedCritic	0.60
D'Urfey_LoveForMoney	0.63
Killigrew_Bellamira	1.95

No pronouns, 60% culled, 500MFW.

PCA, with the tragedy *Bellamira* scoring much lower on the *y*-axis than Killigrew's comedies.

5.3 Delta analysis

Delta analysis is the third descriptive statistical perspective we apply to precursory authorial style in Behn's *The Rover*. The analysis uses Hoover's (2019) workbooks to undertake traditional full text comparative analysis of plays by Behn, Killigrew and D'Urfey, producing a ranked-list of likeness with the test text. In all iterations, Behn's plays (*The City Heiress*, *The Town Fopp*, and *Sir Patient Fancy* are used in this test) are ranked first to third, regardless of the configuration of the test (e.g. Table 1). Interestingly, tests with 60% and 80% culling, as well as those with no culling, rank Killigrew's comedies *Thomaso* and *Parson's Weddings* fourth and fifth, respectively: the test identifies them as stylistically more like *The Rover* than D'Urfey's comedies. These results appear to concur with Hobby's assessment of Behn's use of her source text, in which Killigrew's work is diffused within dialogue of Behn's devising: *Thomaso* runs below the surface but, at the level of high-frequency vocabulary and even function words (as captured in the tests using 100 or 200 MFW, for example). Thus, it is detectable at a quantitative level when the texts are analysed in full, rather than requiring sub-section granularity.

Taken together, the results of these tests suggest there is a lexical similarity between *The Rover* and Behn's other plays, as well as evidence of a more diffused stylistic presence of Killigrew (especially

Thomaso) within *The Rover*, primarily in the first half of the play. The retention of Killigrew's language appears to manifest most strongly at the level of function words, and is most prominent and traceable in the RD tests.

6 Case 2: *The Counterfeit Bridegroom*

The Counterfeit Bridegroom (CB) is a reworking of Thomas Middleton's *No Wit*. It was first attributed to Behn in 1832 in John Genest's *Some Account of the English Stage, from the Restoration in 1660 to 1830*, who asserts that 'it does not appear who altered Middleton's play – but it is so much improved, that it seems probable that Mrs. Behn was the person who made the alteration – 2 or 3 new scenes are added – and the Widow's marriage [...] is much better managed than in the original play' (Genest, 1832, p. 213). The attribution, therefore, is based on an impressionistic and rather romantic assessment of internal evidence. As an attribution case, the investigation would theoretically follow the verification model, in which the text is analysed to identify any characteristics that show a similarity with its hypothetical author. However, as Genest's comments suggest, the text bears a strong resemblance to its source work, and this complicates the analysis of its authorial signal.

Challinor's (forthcoming) close reading of *CB* and *No Wit* identify the relationship between the two plays on an act-by-act basis. Differing degrees of modification to the source text are identified in Acts 1, 3, 4, and 5, while Act 2 is considered to closely follow Middleton's original dialogue (see Table 2; adapted from Challinor, forthcoming).

In an adapted play, such as *CB*, the investigation of style needs to try to account for the possible 'interference' of the precursory author. In addition to texts by Thomas Middleton and Aphra Behn, therefore, plays by two of Behn's contemporaries, Thomas D'Urfey and Edward Ravenscroft, are also included for comparison to assist with the verification of Behn's involvement as the Restoration adaptor. In the analysis of *CB*, the precursory authorial signal is identifiable in a greater range of tests than in the previous case study of Behn's *The*

Table 2 Scene-by-scene comparison of CB with No Wit

Scene	Description
1.1	Follows the shape of Middleton and conveys much of the same information; the language is often similar and many phrases are exactly the same.
1.2	While the scene serves a similar dramatic function to the parallel scene in Middleton, most of the dialogue is particular to CB.
2.1	Follows Middleton closely, often verbatim.
2.2	Follows Middleton very closely, often lifting chunks of dialogue verbatim (including the Dutch and cod-Dutch).
3.1	Broadly follows Middleton.
3.2	Much of the dialogue is original to CB, though the scene covers the same plot points as Middleton.
4.1	The masque itself follows Middleton quite closely; the encounter between Mrs Hadland and Noble is the creation of CB author.
5.1	Original to CB, though a few phrases are recycled.
5.2	Entirely original to CB.
5.3	Much of the dialogue is original to CB, though some phrases are taken directly from Middleton.

Adapted from Challinor (forthcoming).

Rover—arguably because the manner of adaptation is less sophisticated. Our quantitative results for all three descriptive statistical tests identify similarities between Middleton's No Wit and CB in specific parts of the play, especially in Act 2. As well as attesting to the stylistic trace of precursory authorship, the results also allow the analysis to focus more precisely on the stylistic characteristics of the other sections of the play, in order to assess the veracity of the attribution to Aphra Behn.

6.1 Rolling Delta

The Rolling Delta (RD) tests were conducted following the same principles and methods as those used in the first case study. However, as well as using overlapping segments, the play was split into acts to allow the qualitative claims of the literary editors to be evaluated.[2] Our discussion focusses primarily on the latter, partly because they are more transparent textual units.

The Middleton 'maximal likeness' scores, when compared to those of the other authors, show a pronounced similarity with CB for the first four acts of the play (Fig. 6). This trend maps onto the

critical analysis reproduced in Table 2. Notably, while the verbatim replication of dialogue might be expected to show up in a non-culled analysis of 968 MFW, the same trend is captured with a more discriminatory word list (Fig. 7), excepting Act 4 which now shows a similar profile to Act 5 in having a greater likeness with Behn than Middleton.

In the play-by-play charting of maximal likeness, the source text No Wit shows a strong and unique similarity with the test text CB in the first three acts (Fig. 8). Again, in Acts 4 and 5, the similarity reduces. Middleton's other plays included in the test (Chaste Maid and Widow) follow the same trend, albeit with higher Delta scores; as in the first case study, this can be interpreted as a potential trace of the precursory authorial signal in the adapted test text distinct from the source text.

In both of our case studies, therefore, the source text and other works representative of the precursory author show a stylistic similarity with the test text, indicating that the resonance of a source in its adaptation is identifiable using quantitative stylistic methods developed for synchronic cases of collaboration. Further corroboration is provided in Fig. 9, which presents the results for overlapping segment analysis. The test uses a different selection of texts for the four authors, but the trends—importantly—are the same: No Wit is most like CB in Acts 1–3, particularly in Act 2.

With the case of CB, of course, there is the secondary question of the identity of its adaptor. Comparison with The Rover case study is revealing. In that test (see, e.g. Fig. 3), Behn stood out in the RD analysis, separate from her contemporary D'Urfey as well as sharing points of transition with Killigrew. In CB results, there is no strong evidence to suggest there is a greater likeness between the anonymous text and Behn's comedies than between CB and the plays of D'Urfey, Ravenscroft, or others by Middleton. Admittedly, the measure of maximal likeness gives Behn the lowest scores of her contemporaries but—as Burrows (2010) acknowledges—it is unwise to take a moderate marker of difference as strong evidence of authorial involvement. Even in Fig. 9, which could suggest Behn's involvement in Act 5 due to the low score for a segment from The City Heiress, the finding is
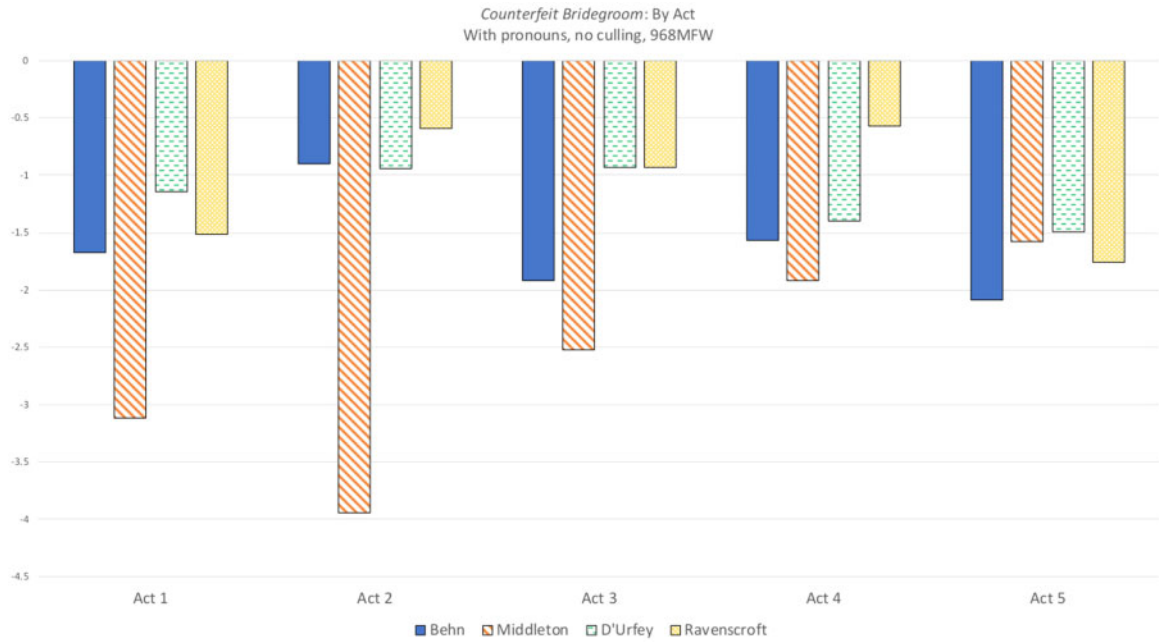
Fig. 6 RD comparing *CB* with Behn, D'Urfey, Middleton, and Ravenscroft, using lowest score for each author per act. With pronouns, no culling, 968 MFW
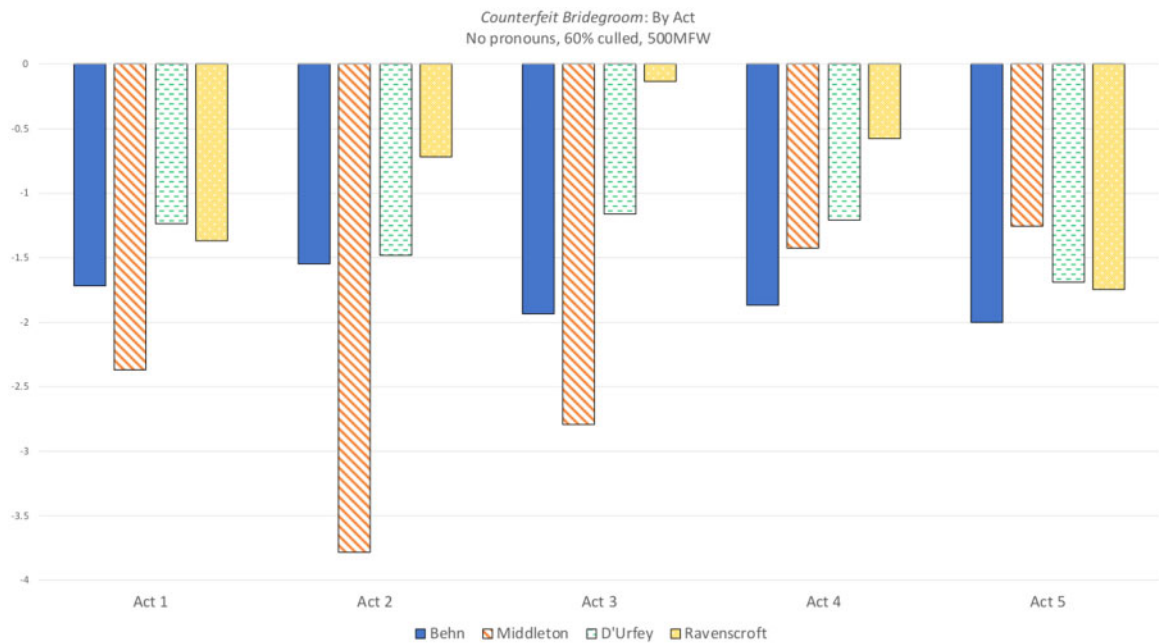

Fig. 7 RD comparing *CB* with Behn, D'Urfey, Middleton, and Ravenscroft, using lowest score for each author per act. No pronouns, 60% culled, 500MFW
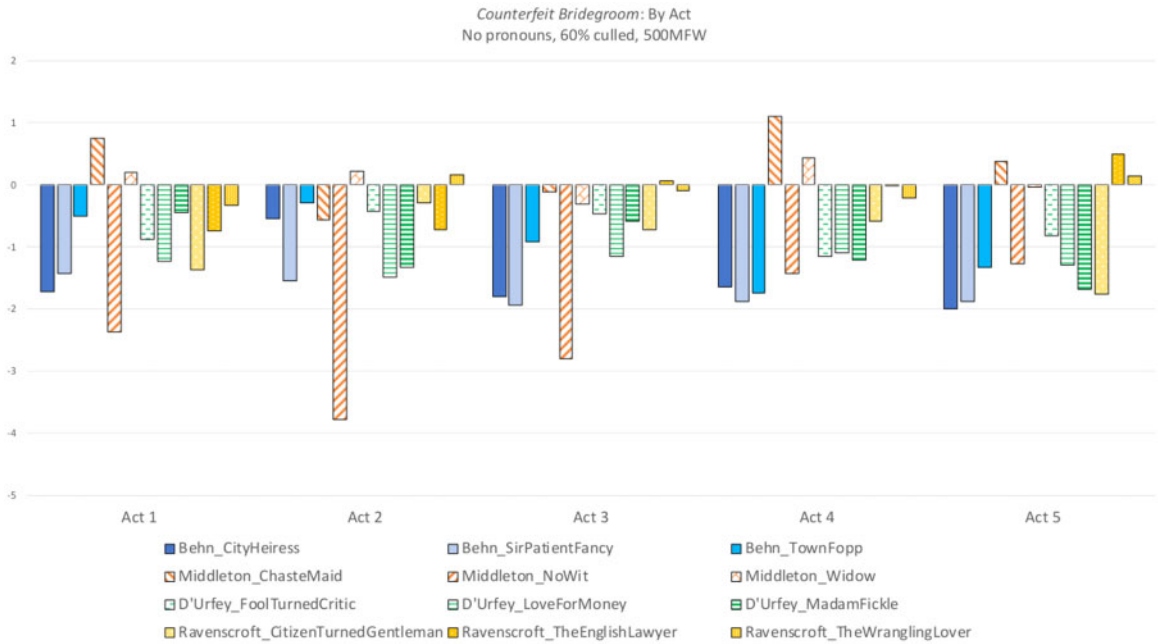
Fig. 8 RD analysis of *CB* by act, comparing Behn, D'Urfey, Middleton, and Ravenscroft by play. No pronouns, 60% culled, 500 MFW
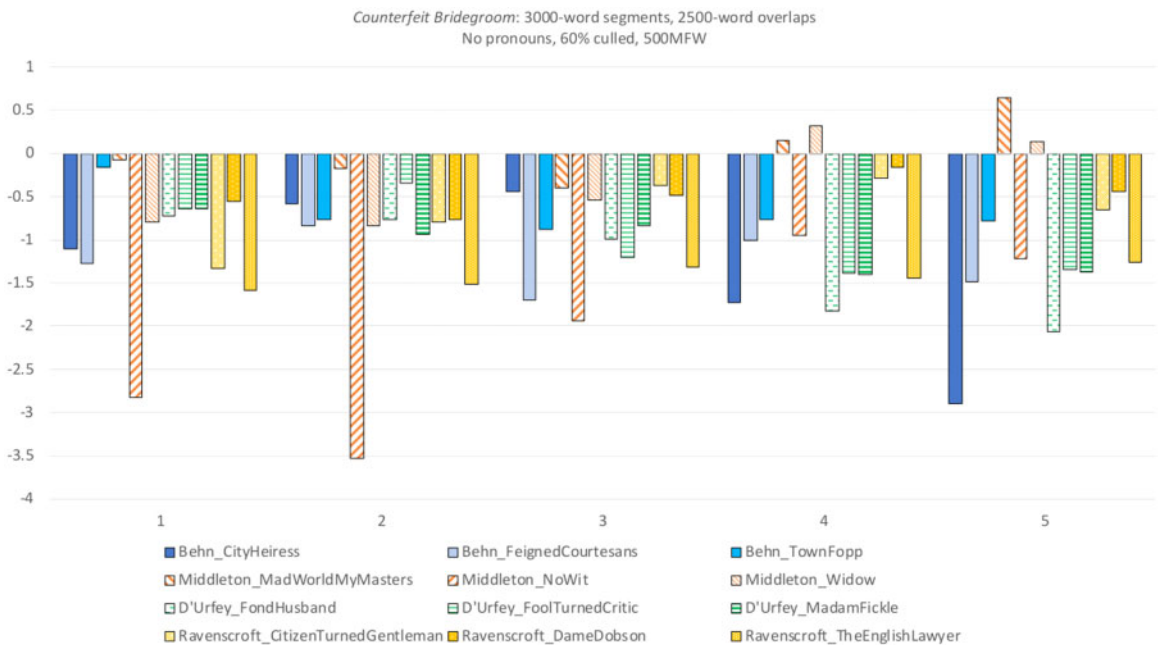
Fig. 9 RD analysis of *CB* by overlapping 3,000-word segments, comparing Behn, D'Urfey, Middleton, and Ravenscroft by play. No pronouns, 60% culled, 500 MFW

countered by the similar trajectory of D'Urfey's plays. There is no clear, unanimous patterning of Behn's comedies versus the other texts included in the test. On the basis of RD, Genest's attribution of her (adaptive) authorship cannot be confirmed.

Of course, the RD analysis explicitly addresses the lexical features relevant to Middleton's authorial style: that was the primary purpose in exploring the nature of precursory authorial style. Removing Middleton from the test—bearing in mind what we now know about the likeness between *CB* and the source text *No Wit*—offers a more focussed perspective on the Restoration elements of the play. Fig. 10 shows the results by act.

The results suggest that, in some acts, there is a greater similarity between Behn's plays and *CB* than between the anonymous play and those of her two contemporaries, D'Urfey and Ravenscroft. Acts 1 and 3 show the greatest likeness with segments from Behn's *The City Heiress* and *Sir Patient Fancy*, with a substantial gap to the lowest scoring segments of the other plays in the test. Acts 2, 4, and 5 show greater parity between Behn and D'Urfey, with Ravenscroft remaining a distant third place

throughout. Although Behn's play segments are among the lowest scoring in the test, this is not to the same degree identified in the case study of *The Rover*. The evidence thus suggests Behn's involvement is plausible, but the results are not definitive.

6.2 PCA

Unlike the case study of *The Rover*, the PCA picks up on similarities between parts of *CB* and Middleton, particularly in relation to Act 2. The disputed text and the source text were dropped into a PCA comparing plays by Behn and Middleton using 100 MFW (Fig. 11). *No Wit* groups comfortably with Middleton, and Act 2 of *CB* also positions closely with this authorial grouping. The segment representing Act 1 of *CB* intermingles with a subset of Behn segments that score lower on Component 1 than the rest of the Behn group, while the remaining *CB* segments (representing Acts 3, 4, and 5) score moderately lower on both components than Behn's plays. The PCA results correspond with the RD results in that Middleton's style dominates Act 2, and that this likeness is identifiable on the basis of the MFWs (Fig. 11). The influence of
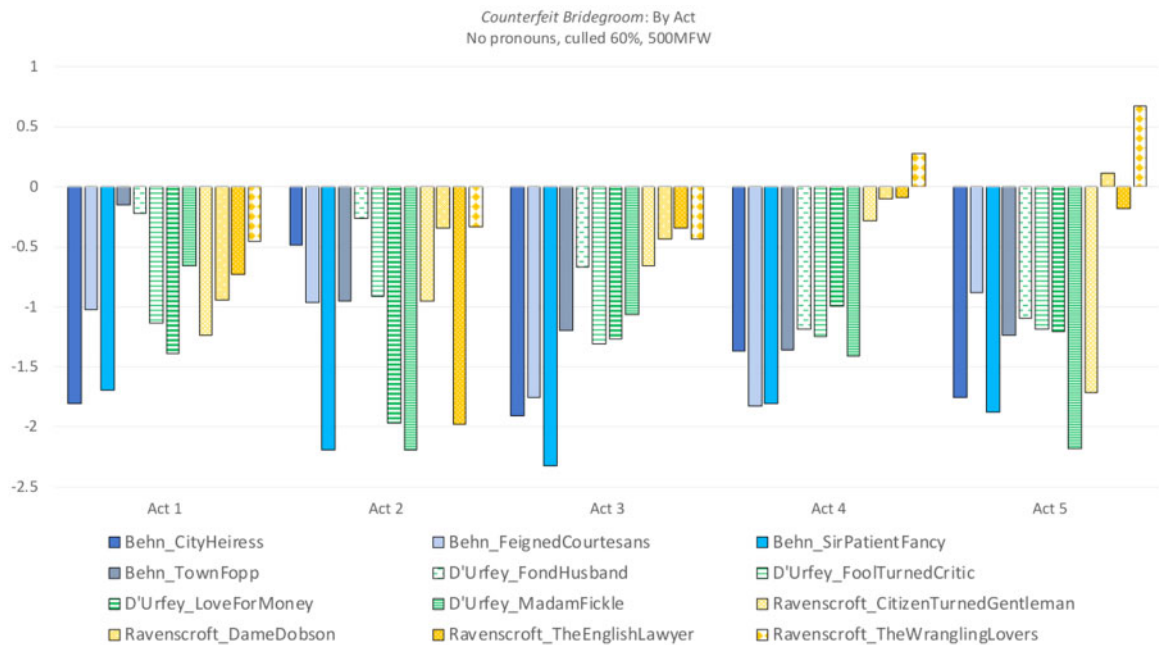
Fig. 10 RD analysis of *CB* by act, comparing Behn, D'Urfey, and Ravenscroft by play. No pronouns, 60% culled, 500 MFW
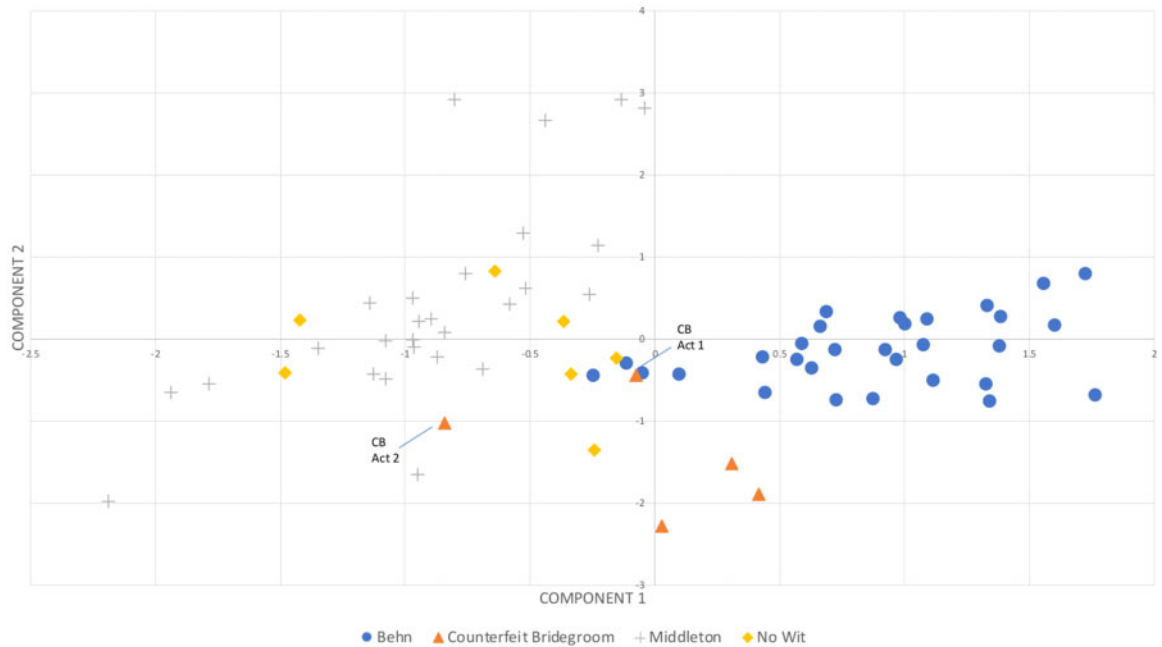
Fig. 11 PCA comparing Behn, Middleton, and CB by act, using 100 MFW

the source text in the other acts—which have a more diffuse configuration of Middleton's play and Restoration innovations—does not, however, appear to permeate through to this lexical stratum.

This has implications for the possible likeness between the less derivative acts in *CB* and Behn's dramatic style. Other tests, comparing Middleton and other Restoration playwrights and using a range of lexical measures (100 and 500 MFW, as well as 2- and 3-grams), offer no clear evidence that Acts 1, 3, 4, and 5 are more like Behn's plays than those of her contemporaries. Figures 12 and 13 show representative results from these comparison tests, comparing Middleton with D'Urfey and Ravenscroft, respectively, using 100 MFW. While Act 2 of *CB* uniformly groups with Middleton's plays, the remaining acts align more closely (to different degrees) with the candidate author, all scoring above zero on Component 1. Act 5, which showed the greatest similarity with Behn's plays in the Middleton-included RD tests, does not show a similarly distinctive alignment in the PCA analyses. The close placement of Acts 1, 3, 4, and 5 of *CB* in

relation to each candidate author in these tests precludes a positive attribution specifically to Behn; rather, it is more likely that such stylistic affinities are determined by temporal similarities in dramatic style. Thus, PCA leaves us at something of an impasse with regard to the likelihood of Behn's involvement in the adaptation of *CB*, although it is able to identify the similarities with the source text, mirroring the results of the RD tests.

6.3 Delta and hierarchical cluster analysis

The full-text Delta results show a more mixed picture than the Behn/Killigrew distribution observed in the first case study. The top-ranked authors vary on an act-by-act basis, and there is minimal consistency between tests—with the exception that Act 2 of *CB* is persistently most like Middleton's source text *No Wit*. This is of interest, given that the tests show a good rate of accuracy (typically 80–100%) when ranking secondary test texts with known authorship (i.e. a Middleton test text is most like another
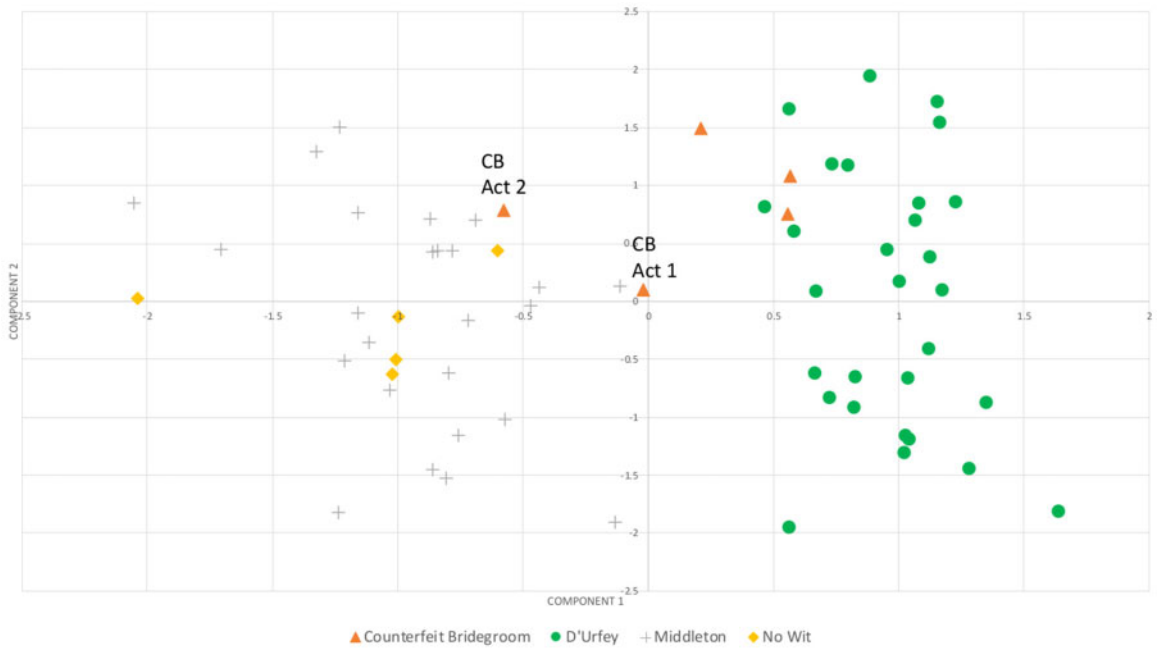
Fig. 12 PCA comparing D'Urfey, Middleton, and CB by act, using 100 MFW
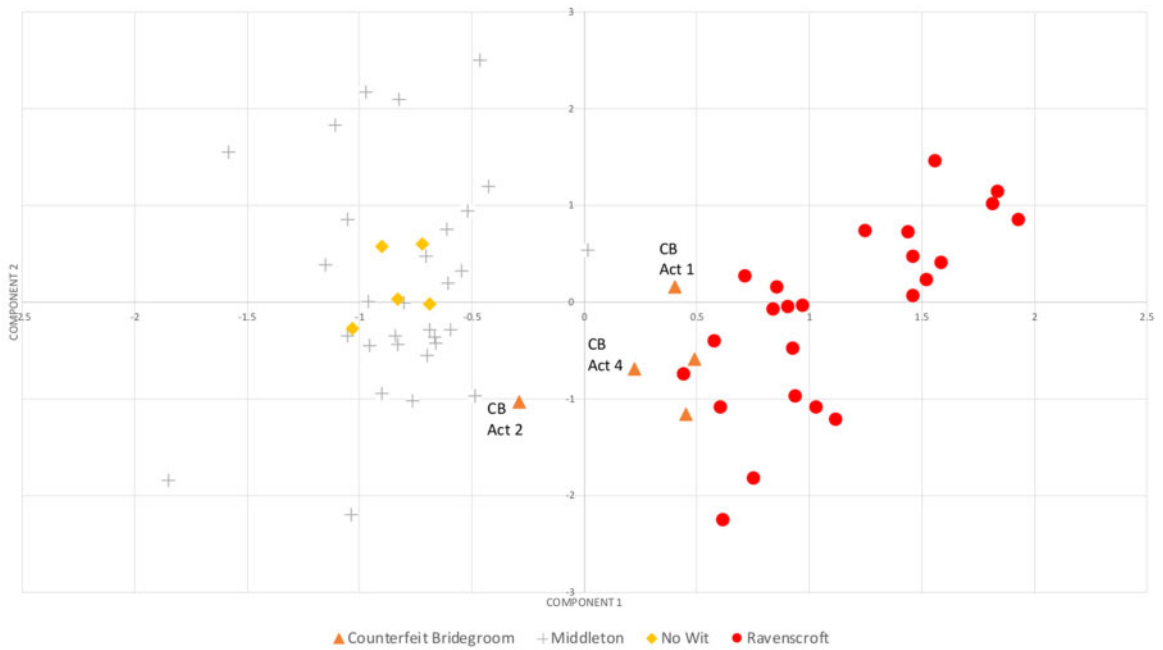

Fig. 13 PCA comparing Ravenscroft, Middleton, and CB by act, using 100 MFW

Table 3 Marker words out of the 100 MFW in author/text & forty-four contemporary plays corpus (independent *t*-test; *p* < 0.001)

	CB	Middleton (comedies)	Behn (comedies)
More frequency	for, how, my, **now**, **sir**	**sir**, it, **now**, in, never, there, why, that, one, them	<u>how</u>, <u>sir</u>, this, thee, why, which, so, who, oh
Less frequent	<u>any</u>, thou	and, to, madam, with, do, of, who, be, very, an	in, <u>any</u>, them, never, the, your

Items in bold are marker words that are shared by *CB* and Middleton, and items underlined are marker words shared by *CB* and Behn.

Middleton play, a Behn test text is most like other Behn plays, and so on).

A curated assessment of lexical markers—that is, a wordlist including items known to be statistically significant in distinguishing a specific author's style—does not provide clear-cut confirmation of Behn's authorship of *CB*, either, although the precursory authorship remains more consistently identifiable. The curated wordlist was compiled following *t*-test results comparing Behn's plays and Middleton's plays with a Restoration reference corpus (comprising forty-four comedies in total, including those of D'Urfey and Ravenscroft; Table 3). The *t*-tests, following Craig (2000), identify lexical items that occur at a statistically significant frequency in one authorial corpus compared to a reference corpus, based on the 100 MFW in the combined corpora.

Using the thirty-two authorial marker words in a Delta cluster analysis provides a curated picture that profiles all texts using lexical features most characteristic of Thomas Middleton and Aphra Behn. Cluster analyses 'force' texts to group together on a scale of diminishing similarity: the closer together two texts are on the visualization, the greater their similarity. The analyses were conducted using Stylo for R.

Fig. 14 shows the results of a cluster analysis comparing Middleton (black), Behn (red), Ravenscroft (orange), and D'Urfey (blue). Act 2 of *CB* (green) groups with Middleton as expected given the inclusion of Middleton marker-words in the list and the consistent association of this act with its precursory author. Acts 3–5 group on their own branch, part of an off-shoot with Act 1 of Behn's *Sir Patient Fancy*, *The Feigned Courtesans*, and Ravenscroft's *The English Lawyer*. The rest of Behn's plays cluster together on a separate branch.

Given that it is the segments from Behn's plays that move to a different branch, rather than the act from *CB* clustering with Behn's works—as with Middleton—the results suggest a weaker kind of stylistic similarity than that identified for the source text. Overall, the positioning of *CB* on a separate branch to the three main authorial groupings suggests that, using this curated word list, there is no strong stylistic similarity between *CB* as a whole and these authors.

Interestingly, the test foregrounds Ravenscroft's *The English Lawyer* as stylistically fragmented, different to his other works. This may reflect the play's complex adaptive history: Ravenscroft adapted *The English Lawyer* from Robert Codrington's translation (1662) of George Ruggle's Latin play *Ignoramus* (published 1630), which was itself based on Giambattista della Porta's *commedia erudita* play *La trappolaria* (see Langbaine, 1691, p. 420). In effect, it has transitioned through four authors writing in three different languages over the course of almost a century (*Ignoramous* was written in 1615), and this diffuse heritage may be captured here in the cluster analyses.

Removing Ravenscroft's plays entirely provides a different picture—and one more resemblant of the act-by-act RD analyses. In Fig. 15, the authorial grouping is consistent across the different play acts. For *CB*, Act 2 clusters with Middleton, further attesting to the precursory signal in this part of the disputed text. The rest of the play, however, clusters on its own sub-branch, which is most proximate to Behn's *The City Heiress* and *Town Fopp*. However, the different clustering position of the disputed play urges caution. While Middleton's precursory role is consistently identified, the variable distribution of *CB* across tests in which Behn is included suggests that the cluster analysis in Fig. 15 is most likely
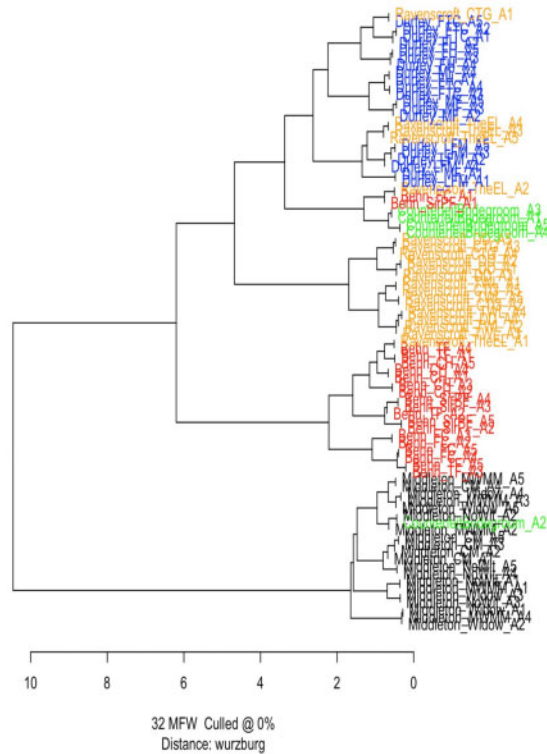
Fig. 14 Hierarchical Delta clusters analysis of plays by Behn (red), D'Urfey (blue)惠, Middleton (black), and Ravenscroft (orange), plus *CB* (green), using thirty-two marker words for Behn and Middleton

responding to the greatest similarity out of the included candidates. The choice of authors and the application of different statistical tests produce different and sometimes contradictory results.

In sum, this could suggest that Behn was not involved in *CB* at all, with the identified likenesses linked to temporal similarities in style, for example. Alternatively, the variable results may indicate that she was one of two or more collaborators involved in preparing the play for the stage, but that her contribution is occluded through the combination of authorial styles—the same way that the traces of Middleton, as precursory author, are diluted in the latter half of the play. A narrative could be woven around the collaborative activities of the Restoration playwrights; yet the inconsistency is also a strong indicator that other aspects of style may be at play. It may be that none of the Restoration authors

included in the tests here were involved in *CB*, and that the play's (main) adaptor remains untested and therefore unidentified. Perhaps the only certainty, to paraphrase Burrows, is the present uncertainty.

7 Conclusion

Our discussion of asynchronous collaboration has sought to address the question of whether the authorial signal of a precursory author is identifiable in a given text, alongside that of the adaptor, or adaptors, of that text, and how significant this stylistic signal may be for investigations of authorship attribution. Across two case studies of dramatic texts (*The Rover* and *CB*), our results determined that different methods (RD, PCA, Delta, and Delta Hierarchical Cluster Analysis) can identify
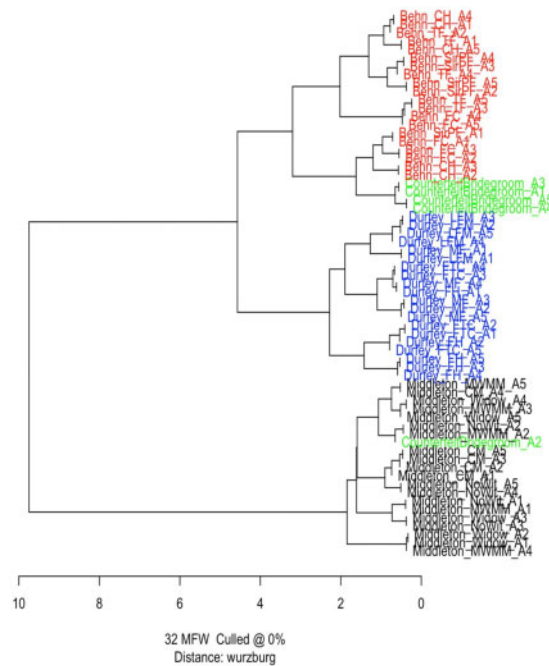
Fig. 15 Hierarchical Delta clusters analysis of plays by Behn (red), D'Urfey (blue)瓩, and Middleton (black), plus *CB* (green), using marker words for Behn and Middleton

precursory authorship in the two test texts considered here. When applied to *The Rover*, RD, and Delta analyses locate clear signs of Killigrew's *Thomaso*, as well as traces of other Killigrew texts that register a level of similarity with *The Rover*, regardless of temporal distance. As Hobby suggests, this indicates the subtle diffusion of Killigrew's style throughout Behn's dialogue, a style that is captured by quantitative measures because it consists of high-frequency words, including function words. The PCA, cluster analyses with curated word lists and RD also consistently foreground Middleton's style in Act 2 of *CB* regardless of the tests' parameters. These tests showed repeatedly that stylistic features in Act 2 correlate with those of Middleton, and RD demonstrated that traces of his style are also identifiable in Acts 1 and 3. The results achieved through quantitative measures, therefore, support the critical interpretation of the plays offered by Challinor (see Table 2). It is likely that such strong Act 2 similarities are down to the retention of Middleton's content and function words, while similarities in Acts 1

and 3 may represent different degrees of precursory and adaptive authorship that include both word-types in the MFWs.

Contrary to the identification of Middleton's stylistic traits in *CB* and Behn's clear stylistic presence in *The Rover*, the evidence for her authorship of *CB* is not clear-cut. Moreover, isolating any single Restoration candidate as potential adaptor from the authors included in the contemporary drama corpus was not possible. When RD, excluding Middleton's texts, was applied in order to localize the analysis to Restoration candidates, Behn's similarity to *CB* was greater than that of the other authors included in the test. However, additional PCA, Delta, and Cluster analyses failed to identify her as a likely author with any consistency or certainty. It may be that *CB* actually features the styles of multiple collaborating Restoration authors, or an author(s) not included in our comparison corpora, making a positive attribution more difficult. What is certain is that a range of methods should be employed in this kind of attribution study, not only to

detect the different stylistic layers that are revealed depending on the test and its parameters, but also to properly explore the reasons that such patterns might exist.

In effect, the results highlight the importance of accounting for the possible contributions of precursory authors in early modern drama, although perhaps not quite in the ways anticipated at the start of this investigation. In the Restoration period, unacknowledged borrowing from earlier plays was extensive, and source texts can be tracked to varying levels in Behn's *The Rover* and the anonymous play *CB*. What was unanticipated was the potential impact that adaptation also had on the texts in the comparison corpus, with stylistic properties cutting across authorial and temporal levels. The example of Ravenscroft's *English Lawyer* demonstrates the often convoluted textual histories of adapted plays, and the continuing challenges of disentangling factors of authorship from chronology and translation.

Moreover, the different ways in which our source texts interact with their adapted versions raises questions about how we interpret the style of palimpsestic adaptations in general: for example, do they tend to register more at the level of topic or high-frequency lexis, or a mixture of both? Or, do they have their own distinctive profile, as some studies have shown to be true of translations (Rybicki and Heydel, 2013; Lee, 2018; Lynch and Vogel, 2018)? Where authorship attribution has focussed on the narrowing or elimination of potential confounding variables, such as genre and chronology, in the search for the authorial 'linguistic fingerprint', our results have shown that the question of time cannot be removed from the analysis of early modern drama without the loss of important lexical data relevant to other dimensions of style, including authorship. The challenge for computational stylometry is to find the best methods for uncovering stylistic variation in these complex and layered texts, whether they are straight adaptations of earlier plays, translations, reworkings of other genres, or all of the above, and how a theoretical framework dedicated to an authorial 'fingerprint' can satisfactorily accommodate the permutations of topic, genre and time into our understanding of literary style and computational stylistic methods.

References

Auerbach, D. (2018). 'A Cannon's Burst Discharged against a Ruinated Wall': a critique of quantitative methods in Shakespearean authorial attribution. *Authorship*, **7**(2):1–16. https://doi.org/10.21825/aj.v7i2.9737.

Barber, R. (2019). Marlowe and overreaching: a misuse of stylometry. *Digital Scholarship in the Humanities*, **34**(1): 1–12.

Baron, A. (2017). *VARD 2 (version 2.5.4)*. Lancaster: UCRELS, University of Lancaster. http://ucrel.lancs.ac.uk/vard/about/ (accessed 2 December 2019).

Binongo, J. N. G. (2003). Who wrote the 15th book of Oz? An application of multivariate analysis to authorship attribution. *Chance*, **16**(2): 9–17.

Brown, P. (2017). *Early Modern Theatre People and Their Social Networks*. Unpublished Ph.D. thesis, De Monfort University. uk.bl.ethos.732356.

Burrows, J. (2007). All the way through: testing for authorship in different frequency strata. *Literary and Linguistic Computing*, **22**(1): 27–47.

Burrows, J. (2010) Never say always again: reflections on the numbers game, the wisbey lecture for 2006, King's College, London. In McCarty, W. (ed.), *Text and Genre in Reconstruction*. London: OpenBook Publishers, pp. 23–35.

Challinor, J. (forthcoming). Headnote to The Counterfeit Bridegroom. In Bowditch, C., Evans, M., Hobby, E. and Wright, G. (eds), *The Cambridge Edition of the Works of*

Aphra Behn Vol II. Cambridge: Cambridge University Press.

Craig, H. (2000). Is the author really dead? An empirical study of authorship in English Renaissance Drama. *Empirical Studies of the Arts*, **18**(2): 119–34.

Craig, H. (2019). *Intelligent Archive* (Rosella IA 3.0). Newcastle: Centre for Literary and Linguistic Computing. https://www.newcastle.edu.au/research-and-innovation/centre/education-arts/cllc/intelligent-archive (accessed 2 May 2019).

Craig, H. and Burrows, J. (2012). A collaboration about a collaboration: the authorship of King Henry VI, Part Three. In Deegan, M. and McCarty, W. (eds), *Collaborative Research in the Digital Humanities*. Farnham: Ashgate, pp. 27–65.

Da, Nan Z. (2019). The computational case against computational literary studies. *Critical Inquiry*, **45**(3): 601–39.

Eder, M. (2016). Rolling stylometry. *Digital Scholarship in the Humanities*, **31**(3): 457–69.

Eder, M., Rybicki, J., and Kestemont, M. (2016). Stylometry with R: a package for computational text analysis. *The R Journal*, **8**(1): 107–21.

Erne, L. (2013). *Shakespeare as Literary Dramatist*. 2nd edn. Cambridge: Cambridge University Press.

Freebury-Jones, D. (2017) Kyd and Shakespeare: authorship versus influence. *Authorship*, **6**(1): 1–24. http://dx.doi.org/10.21825/aj.v6i1.4833

García-Reidy, A. (2019). Deconstructing the authorship of Siempre Ayuda La Verdad: a play by Lope de Vega? *Neophilologus* **103**(4): 493–510. https://doi.org/10.1007/s11061-019-09607-8.

Genest, J. (1832). *Some Account of the English Stage, from the Restoration in 1660 to 1830*. Vol. **I**. Bath: H. E. Carrington.

Gladwin, A. A. G., Lavin, M. J., and Look, D. M. (2017). Stylometry and collaborative authorship: Eddy, lovecraft, and 'The Loved Dead'. *Digital Scholarship in the Humanities*, **32**(1): 123–40. https://doi.org/10.1093/llc/fqv026.

Harbage, A. (1940). Elizabethan: restoration palimpsest. *The Modern Language Review*, **35**(3): 287.

Hobby, E. (forthcoming). Headnote to Aphra Behn's The Rover. In Bowditch, C., Evans, M., Hobby, E. and Wright, G. (eds), *The Cambridge Edition of the Words of Aphra Behn Vol II*. Cambridge: Cambridge University Press.

Hoover, D.L. (2019). *The Delta Spreadsheets*. https://wp.nyu.edu/exceltextanalysis/deltaspreadsheets/ (accessed 18 September 2019).

Ilsemann, H. (2016). The two Oldcastles of London. *Digital Scholarship in the Humanities*, **32**(4): 788–96.

Ilsemann, H. (2018). Christopher Marlowe: hype and hoax. *Digital Scholarship in the Humanities*, **33**(4): 788–820.

Jockers, M. L. (2013). *Macroanalysis: Digital Methods and Literary History*. Urbana, IL: University of Illinois Press.

Kestemont, M. (2014). Function words in authorship attribution. From black magic to theory? In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*. Gothenburg: Association for Computational Linguistics, pp. 59–66.

Kewes, P. (1998). Authorship and appropriation: writing for the stage in England, 1660-1710. *Oxford English Monographs*. Oxford/New York: Clarendon Press/Oxford University Press.

Langbaine, G. (1691). *An Account of the English Dramatick Poets*. Oxford: George West and Henry Clements.

Lee, C. (2018). Do language combinations affect translators's stylistic visibility in translated texts? *Digital Scholarship in the Humanities*, **33**(3): 592–603.

Love, H. (2002). *Attributing Authorship: An Introduction*. New York: Cambridge University Press.

Lynch, G. and Vogel C. (2018). The translator's visibility: detecting translatorial fingerprints in contemporaneous parallel translations. *Computer Speech & Language*, **52**: 79–104.

Orgel, S. (1992). What is a text? In Kastan, D. S. and Stallybrass, P. (eds), *Staging the Renaissance: Reinterpretations of Elizabethan and Jacobean Drama*. New York: Routledge, pp. 83–7.

Potter, L. (2008). Involuntary and voluntary poetic collaboration: The Passionate Pilgrim and Love's Martyr. In Drábek, P., Kolinská, K., and Nicholls, M. (eds), *Shakespeare and His Collaborators over the Centuries*. Newcastle Upon Tyne: Cambridge Scholars Publishing, pp. 5–20.

Rybicki, J. (2016). Vive La Différence: tracing the (authorial) gender signal by multivariate analysis of word frequencies. *Digital Scholarship in the Humanities*, **31**(4): 746–61.

Rybicki, J. and Heydel, M. (2013). The stylistics and stylometry of collaborative translation: Woolf's Night and Day in Polish. *Literary and Linguistic Computing*, **28**(4): 708–17. https://doi.org/10.1093/llc/fqt027.

Rybicki, J., Hoover, D., and Kestemont M. (2014). Collaborative authorship: Conrad, Ford and Rolling Delta. *Literary and Linguistic Computing*, **29**(3): 422–31.

Schöberlein, S. (2016). Poe or not Poe? A stylometric analysis of Edgar Allan Poe's disputed writings. *Digital Scholarship in the Humanities*, **32**(3): 643–59.

Stallybrass, P. (1992). Shakespeare, the individual, and the text. In Grossberg, L., Nelson, P., and Treicher, P. (eds), *Cultural Studies*. New York: Routledge, pp. 593–612.

Stamatatos, E. (2009). A survey of modern authorship attribution methods. *Journal of the American Society for Information Science and Technology*, **60**(3): 538–56.

Stamatatos, E. (2013). On the robustness of authorship attribution based on character n-gram features. *Journal of Law and Policy*, **21**(2): 421–39.

Todd, J. M. (2017). *Aphra Behn: A Secret Life*. London: Fentum Press.

Underwood, T. (2019). *Distant Horizons: Digital Evidence and Literary Change*. Chicago, IL: The University of Chicago Press.

van Halteren, H., Baayen, H., Tweedie, F., Haverkort, M., and Neijt, A. (2005). New machine learning methods demonstrate the existence of a human stylome. *Journal of Quantitative Linguistics*, **12**(1): 65–77.

Vickers, B. (2018). Introduction. *Authorship*, **7**(2): 1–10. https://doi.org/10.21825/aj.v7i2.9734.

Weidman, S. G., and O'Sullivan, J. (2018). The limits of distinctive words: re-evaluating literature's gender marker debate. *Digital Scholarship in the Humanities*, **33**(2): 374–90.

Wright, D. (2017). Using word n-grams to identify authors and idiolects: a corpus approach to a forensic linguistic problem. *International Journal of Corpus Linguistics*, **22**(2): 212–41.

Zitner, S. (1984). *The Knight of the Burning Pestle*, by Beaumont, F. *The Revels Plays*. Manchester: Manchester University Press.

Notes

1 The inclusion/exclusion of pronouns has a similar, moderate impact.

2 The segments and overlap sizes are different to those used in *The Rover* analysis: 3,000-word segments with 2,500-word overlaps. This is due to the greater quantity and length of texts under investigation, which required alternative segmentation criteria in order to be able to include all segments of the texts in the preconfigured Delta spreadsheet (Hoover, 2019).