

On Early Style: A Stylochronometric Critique of Late Style in Literature

Jonathan Reeve (jonathan.reeve@columbia.edu)

1 The Problem

The idea has persisted for centuries: that artists, whether writers, painters, or musicians, significantly change their artistic styles in their final years, or in the final stages of their careers. If this sounds vague, it is. ‘Late style,’ as it is known, or *Spätstil* in the German critical tradition, has been used to explain the last works of, among others, Shakespeare, Beethoven, Goethe, Proust and Picasso. The late style of these artists have been the primary concern of numerous recent monographs (see, for example, Wheeldon, 2009; Solomon, 2004; Neumann, n.d.; Smit, 1988; Lipking, 1984; Van Dam, 2016). Anecdotal evidence certainly abounds. But is this a phenomenon that can be tested with a large collection of works?

One especially influential book in this field is Edward Said’s *On Late Style: Music and Literature Against the Grain*, a work which expands Theodor Adorno’s theories of late style in Beethoven to literature, covering Marcel Proust, Jean Genet, Thomas Mann, and C.P. Cavafy, among others (Said, 2006). Said’s book is itself a work of late style, and was posthumously compiled from his speeches, notes, and course materials: in the ’90s, Said had taught a course at Columbia University called Late Style, which included required readings in Sophocles’s *Oedipus at Colonus*, Euripedes’s *Bacchae*, Shakespeare’s *Winter’s Tale*, Mann’s *Doctor Faustus*, Beckett’s *Endgame*, poems by Gerard Manley Hopkins, Proust’s *Time Regained*, poems by Cavafy, and two works by Genet (Said, 1989). In the experiment that follows, I’ll test the works of some of these writers for the statistical significance of their late periods.

What Said and Adorno mean, precisely, by ‘late’ or ‘lateness’ isn’t always clear, however. Michael Wood’s introduction to his volume seems to summarize Said’s conception, while acknowledging the slipperiness of the term:

It’s worth pausing over the delicately shifting meanings of the word late, ranging from missed appointments through the cycles of nature to vanished life. Most frequently perhaps late just means “too late,” later than we should be, not on time. But late evenings, late blossoms, and late autumns are perfectly punctual—there isn’t another clock or calendar they are supposed to match. Dead persons have certainly got themselves beyond time, but then what difficult temporal longing lurks in our calling them “late”? Lateness doesn’t name a single relation to time, but it always brings time in its wake. It is a way of remembering time, whether it is missed or met or gone. (5)

In introducing his argument, Said posits ‘three great human episodes common to all cultures and traditions,’ the third of which he intends to study (19). The third episode is characterized by the contrariness that is implied in the subtitle, *Music and Literature Against the Grain*. This opposition, or incompatibility, may be directed towards to one’s own time, one’s prior works, or the contemporary status quo. It ‘involves a nonharmonious, nonserene tension, and above all, a sort of deliberately unproductive productiveness going *against*’ (21). This is also Adorno’s view, whom Said quotes here: ‘the maturity of the late works,’ according to Adorno, “do not resemble the kind one finds in fruit. They are ... not round, but furrowed, even ravaged. Devoid of sweetness, bitter and spiny, they do not surrender themselves to mere delectation’ (26).

If we are generous, we might read ‘the late works’ here as ‘these particular late works,’ rather than ‘all late works,’ or even ‘most.’ Yet Said implicitly extends his argument well beyond just the works he discusses in their volume, as do most critics writing about late style. If this is a fair characterization, it is a very bold and far-reaching claim, and one which is testable.

Despite this, there are few sustained critiques of late style in English literary studies¹. Two works that approach a critique include Gordan McMullan’s recent monograph *Shakespeare and the Idea of Late Writing*, and the volume of essays he edits, *Late Style and Its Discontents* (McMullan, 2007; McMullan and Smiles, 2016). McMullan traces the history of the idea, which he sees as an extension of man-of-genius schools of criticism, especially those surrounding Shakespeare and Beethoven (17). He also notes that *man* of genius is seemingly an important designation for earlier critics of the late style tradition, since they rarely, if ever, laud women artists as possessing this characteristic (ibid.). (And in fact, Said’s syllabus features no works by women writers.)

The genealogy of late style as a critical concept, interesting though it is, is not my concern here, however. Instead, I propose an experiment for testing late style itself, at least as far as may be measured with text analysis. This is an ideal job for a quantitative approach, one that could analyze thousands of novels at a time.

2 Operationalizing ‘Late’

Translating a literary critical claim into an algorithm is always a ‘lossy’ operation. There are so many definitions of both ‘late’ and ‘style’ that it would be impractical to design experiments that test them all. But we could choose a few of the most compelling definitions of each, and design experiments around them, using well-tested methods. We should begin by choosing definitions for ‘late.’ This is not easy, since Said’s argument, for one, depends on the term’s unresolved ambiguities, many of which are contradictory. We might be able to mitigate

this problem somewhat, however, by operationalizing several different meanings of the word. First, let's take Said's notion of 'three great human episodes' literally.

2.1 Late as Final Third

Most who write about 'late style' of an artist describe it as an outgrowth of a late period, which itself contrasts with an early period, and sometimes a middle period. Only in a rare cases are artists' chronologies segmented much further, as in Picasso's 'blue' and 'rose' periods. It follows, then, that one way to operationalize lateness is to begin by dividing the works of a writer into two or three categories.

One way to accomplish this is to treat the vector of publication dates² as a one-dimensional clustering problem, and to leverage Jenks Natural Breaks Optimization to find the largest gaps in vector, optimizing the Goodness of Variance Fit by incrementing the number of classes until we arrive at three. We might then be able to look for what Adorno terms the 'caesura,' or definitive break that separates an artist's earlier from late styles.

2.2 Lateness as Proximity to Death

Another way of interpreting lateness is to understand it as a writer's growing awareness of his or her mortality. Writers themselves are likely the best judges of their own health, and so barring a sudden, violent death, they are likely aware of their failing health for some time—perhaps even years. This is McMullan's primary interpretation of lateness, whose *Shakespeare and the Idea of Late Writing* is subtitled *Authorship in the Proximity of Death*. There are certainly writers who have struggled with debilitating illness. In 'Cough Prints and Other Intimacies: Considerations in Editing Lawrence's Later Verse,' for instance, Christopher Pollnitz remarks on blood spattered across D.H. Lawrence's later manuscripts, brought on by Lawrence's tuberculosis (Pollnitz, 1995). And some read Virginia Woolf's last novel, *Between the Acts* in light of her impending suicide (Kenney, 1975).

Since Project Gutenberg maintains author birth and death data, (even though they don't maintain publication dates) this made this one of the easier analyses to perform. To operationalize this, I assigned 'early,' 'middle,' and 'late' categories, as before, only instead of basing them on publication years or their gaps, I assigned them according to when they were published with respect to the year of that writer's death. If it was published within five years, I assigned it to the 'late' category; if within ten, I assigned it to 'middle,' and 'early' to the rest. These are fairly liberal choices, I freely admit, but the object was to determine whether, should there be no statistically discernable late style, there might yet be a very late style.

2.3 Others

Other methods for operationalizing lateness would include considering ‘late style’ to be an expression of Altersstil, or the style of old age. Operationalizing Altersstil might mean calculating the age of the author at each publication, and designating it Altersstil after the age of, say, 65. But would this be affected by changes in life expectancies over time? Or would this definition be different for women, who typically have longer life expectancies? Due to the complexities of these questions, and the number of other ways of considering ‘style,’ I leave these other definitions for future studies.

3 Operationalizing ‘Style’

3.1 Stylometry

The most obvious way to operationalize differences in style, is to employ the techniques of stylometry, or the statistical measurement of writing style. Although there have been a number of different approaches to stylometry over the years, one of the most successful, and one used in forensic linguistics today, is based on the most frequent words of a text. This analysis follows from the observation that frequent words tend to be function words, (*a, the, and*, etc), and infrequent words tend to carry content (*blue, sunset, Kent*). Separating style from content, by dividing words according to their ranked frequencies, has been very successful at authorship attribution, and to a lesser extent to the attribution of gender or nationality (Love, 2002; Rybicki, 2015). In fact, the proportions of single frequent words, like *an*, may even be used to tell with surprising accuracy whether a writer is likely to be British or American. Notable papers employing stylometry include John Burrows’s seminal 1987 *Computation into Criticism*, which deals with Jane Austen’s novels, and Hugh Craig and Athur Kinney’s 2009 *Shakespeare, Computers, and the Mystery of Authorship*, which gives detailed attributions for segments of the bard’s collaboration with Christopher Marlowe (Burrows, 1987; Craig and Kinney, 2009).

Stylometry has also been used on a few occasions to study diachronic changes in a single-author corpus. David Hoover, in ‘Corpus Stylistics, Stylometry, and the Styles of Henry James,’ uses stylometric methods to show differences between James’s much-discussed early and late periods (Hoover, 2007). Carmen Klaussner and Carl Vogel use this method to predict chronology in James and Mark Twain, and arrive at a root mean squared error of ± 7.2 years (Klaussner and Vogel, 2015). Richard Forsyth, in a recent study of what he terms ‘stylochronometry,’ differentiates between ‘younger Yeats’ and ‘older Yeats,’ devising along the way a measurement he calls a ‘youthful Yeatsian index.’ (Forsyth, 1999) Finally, Dirk van Hulle and Mike Kestemont use stylometry to periodize Samuel Beckett’s works, and find an instance of late style in his last work, *Worstword Ho*. ‘[T]he unique position of *Worstword Ho* in the stylometry analyses,’ they conclude, ‘indicates that, when Beckett was almost eighty

years old, he still managed to write a work that was stylistically so innovative that it constitutes yet another turning point' (van Hulle and Kestemont, 2016, p.196). These are all remarkable case-studies of single authors, but do these trends hold true at a bigger scale?

The methods of these, and other stylometric studies, are to compare the writing styles of texts by comparing the frequencies of the first 100-800 words. Sometimes this is done using Burrows's Delta method, or variations thereof, but often these frequencies are compared with a multivariate technique like Cluster Analysis, or Principal Component Analysis (PCA) (Burrows, 2002; Smith and Aldridge, 2011). To choose the appropriate technique, I ran a grid search over roughly 9,000 possible parameter combinations, including vectorization techniques (adjusting document frequency with inverse document frequency using TF-IDF, for instance), numbers of most frequent words, dimensionality reduction techniques, and clustering techniques. The best-performing methods and parameters were document frequency vectors of the top 800 most frequent words, reduced with PCA to five principal components, and clustered in these five dimensions with a Bayesian Gaussian mixture model with a maximum number clusters of three. The advantage of using the Bayesian model here is that it can infer the numbers of clusters, thereby accomodating writers whose styles have one cluster (no distinct styles), two clusters (early and late styles), or three clusters (early, middle, and late styles).

Since we're working in 5-dimensional space, we can quantify the distinctiveness of a document as we would a vector, by taking the L_5 norm, also denoted $\|x\|_5$, and defined as

$$\|x\|_5 = \left(\sum_i \|x_i\|^5 \right)^{\frac{1}{5}}$$

where x_i are the PCA-reduced word frequencies of each text. A high L_5 could be read as a proxy for the stylistic distinctiveness of the text, and a low L_5 could be read as a proxy for how stylistically typical the text is with respect to the writer's corpus.

3.2 Style as Content: Word Vectors

Paradoxically, style may also be construed as content. We might read Said's description of late style as 'bitter' and 'spiny' as describing the *content* of the lateness, instead of its *manner*, i.e. *what* is said, rather than *how* it is said—not a complaining manner of speaking, for instance, but a complaint itself. Taken as such, stylometry would not be able to detect this style-as-content. To test this, we would instead need a method that takes content into account.

For this, I chose to use word embeddings, first developed at Google, by Tomas Mikolov, et al (Mikolov et al., 2013). Since word embeddings encode the collocations of words with

their contexts in millions of documents, they are often used as useful proxies for meaning, encoded as n-dimensional vectors. As vectors, I can reuse the rest of my clustering pipeline. I chose to use Stanford University's Global Vectors for Word Representation, 300-dimensional vectors pre-trained on millions of web documents (Pennington et al., 2014). Of course, there will be some anachrony at work here, as contemporary web pages do not exhibit the same semantics as nineteenth-century novels, but, after all, we want to compare meanings as they are construed by readers today, rather than insist on historical accuracy.

Since I use raw document frequencies as preliminary vectors for both of these corpora, I had to apply caution with respect to document length. Since it's unfair to compare the frequency of a word, especially an infrequent word, in two documents of very different lengths (a given word simply has more of a chance to appear in a longer text), I first split each text in two. This allows me to see whether the style of a text is self-similar, thereby acting as a sanity check for the stylometric parameters. From there, I draw random samples from each text segment, each with the length of the shortest segment. I then perform 10 iterations of the entire analysis—random sampling, dimensionality reduction, and clustering—to ensure that these samples are fairly representative, and any resulting metric represents the average of all of these trials.

3.3 Other Operationalizations

There are many other possible ways to test style, of course. A third way is to assume that style is reflected not in the words used, but in their syntactic relations, that the composition of a sentence reveals the most about style. It would be impractical to consider all possible operationalizations of literary style here, but a few more warrant mentioning. One is style in the wake of success: how might a growing readership change a writer's style? (I suspect that James Joyce's nine-year attempt at publishing *Dubliners* would have been much longer if *Finnegan's Wake* were instead his first book.) Is style-after-success a more statistically significant shift than that which one might find in old age, or faced with impending death? Similarly, would other major life events affect a writer's literary style, like marriage, parenthood, or graduation from an MFA program?

4 Corpora

The texts I chose to analyze belong to two corpora: one curated corpus, composed of writers either directly discussed by Said, or otherwise well-known for their periodicity; and another, much larger, automatically generated corpus. The curated corpus includes the writers Marcel Proust, Jean Genet, and Thomas Mann, who are either directly discussed in *On Late Style*, or appear on Said's course syllabus³. (I used the original-language versions of all of these texts,

and adjusted my tokenizer for these languages.) To these, I added writers either well known for their periodicity (James and Dickens), or their relative stylistic uniformity (George Meredith, Willa Cather). I also included a few popular but now lesser-studied writers, Mary Augusta Ward (who published as Mrs. Humphry Ward), and Arnold Bennett, the villain, one might say, of Virginia Woolf's famous essay 'Mr. Bennett and Mrs. Brown.' I manually annotated each text with publication dates, and cleaned up the OCR using a semi-automated method.

The second corpus I automatically generated using [Corpus-DB](#), a SQL database wrapper and REST API I created for Project Gutenberg and other texts, with enhanced metadata aggregated from DBPedia and library APIs. I queried this database for all plain text files categorized with the Library of Congress Classification 'PR' (British Literature) that are written by writers with more than eight total works available in English. I chose to narrow these to this LCC so as to minimize international stylochronometric difference, and I chose nine as a reasonable number for dividing into groups of three—had I included writers who wrote only two novels, the clustering algorithm might unfairly ascribe early style to one, and late style to the other. The generated corpus originally contained works by 141 writers, but only 51 of those writers had publication data publicly available on DBPedia. The final generated corpus features roughly 900 works from these 51 writers.

5 Metrics

I define the 'lateness' of a work as how different the work's style is, statistically speaking, from the rest of the author's corpus. Formally, this is the absolute value of the average deviation from the mean, in five-dimensional PCA vector space, of the centroids of the single-author corpus's samples. The vectors of the samples themselves are either PCA-reduced document frequency (DF) vectors, or style-as-content word embedding vectors. Since PCA collapses many of the distances here, this is a proportional value, but still one which is valuable in comparison to others in the same vector space.

The measure may be adapted to compute the 'lateness' of a writer, by finding the centroid, in 5-D, of the Bayesian-clustered category containing the works with the latest publication dates (automatically detected 'late works'), and subtracting the Euclidean distance to the centroids of the early works. A positive lateness score would indicate a measure of 'late style,' a score of zero would indicate no statistically significant stylistic variation at all, and a negative score would indicate 'early style,' or statistically significant differences between early works and the mean⁴.

A writer's lateness may be further adapted to compute the periodicity of a writer, by comparing the periods pre-clustered according to their publication dates with the periods

inferred by stylometric clustering. Thus, if one-dimensional clustering of an author's publication chronology, using Jenks Natural Breaks Optimization, returns a categorization scheme for three clusters, those three clusters are compared with the clusters inferred by the Bayesian Gaussian mixture model using an Adjusted Rand Index, which computes the mutual information of these clusters, regardless of their labels or numbers (Rand, 1971).

6 Results: Curated Corpus

Fig. 1 shows the mean latenesses and mean periodicities of the writers in the curated corpus. Four writers have noticeable lateness scores: Mann, Genet, Cather, and to a lesser extent, James. Since Said devotes chapters to both Mann and Genet in *On Late Style*, and more than one volume has been devoted to the late style of James (see, for example, Smit, 1988), this shouldn't be very surprising. In fact, this seems to confirm Said's intuitions, at least with respect to these particular exemplars. The bigger picture isn't as convincing, however: Dickens, Scott, Meredith, and Woolf show very miniscule late styles; Conrad and Bennett show negative latenesses (their early styles are slightly more distinct), and both Proust and Ward show highly negative late styles.

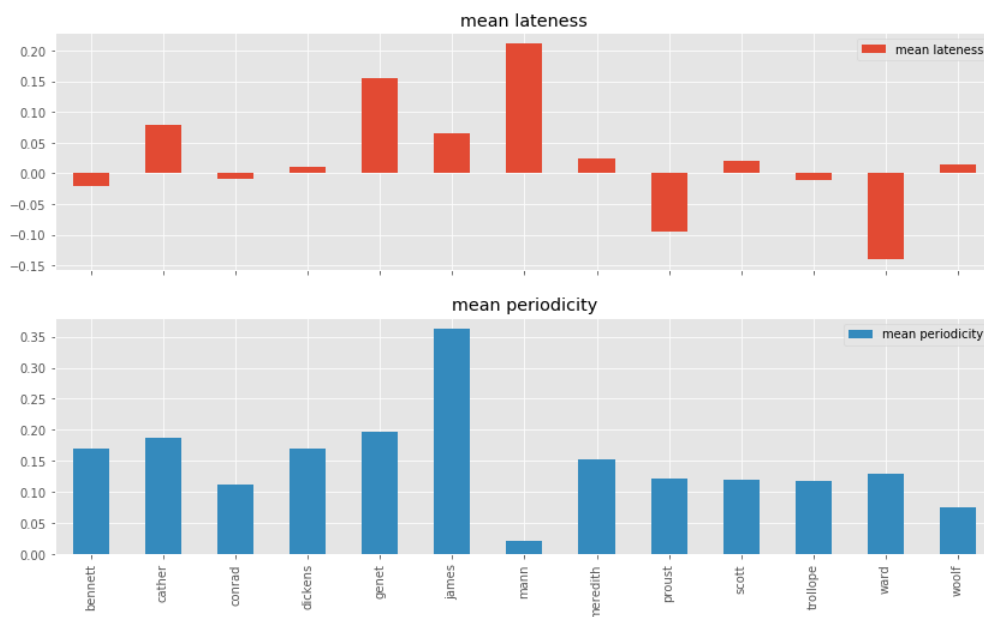


Fig. 1: Average Latenesses and Periodicities

The average lateness score for this corpus is 0.03, a very weak case for late style, and this is a corpus whose writers were selected largely because they are known for lateness.

Curiously, the lateness scores don't align with the periodicities. There, it seems that most writers exhibit roughly the same periodicities, excepting Mann, who scores much lower, and James, who scores much higher. A few of these deserve a closer look.

6.1 Case Study: Henry James

Fig. 2 shows the first two components of the PCA-reduced vector space of Henry James's novels, with each novel represented as two samples. The sizes of the points represent their relative chronology: smaller points for earlier works, and larger points for later works. The colors of the points represent how they are clustered by the Bayesian Gaussian mixture model, in five dimensions.

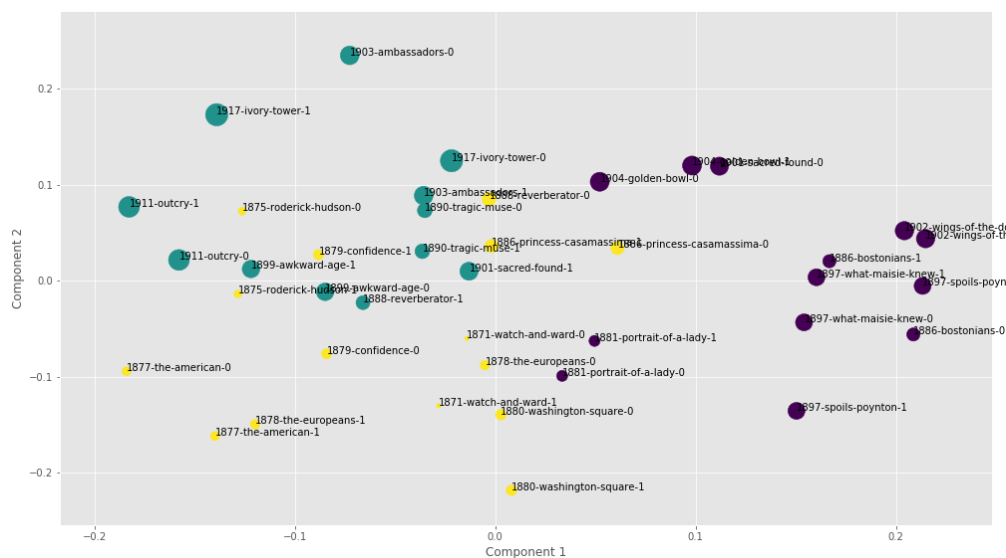


Fig. 2: James PCA

This is one of the most uniform and successful clustering experiments in this corpus: the first, yellow cluster contains only novels published in the '70s and '80s. The second, purple cluster contains mostly novels from the '90s and early 1900s. The third, mostly the 1900s and 1910s.

The outliers here, with the highest lateness scores, are one sample from *The Ambassadors* (1903), both samples of *The Sacred Fount* (1901), and earlier works. These are not exactly James's latest works—but then again, James does not have a very high lateness score, comparatively. His periodicity, however, is remarkable, and that is reflected in this analysis.

6.2 Case Study: Marcel Proust

Marcel Proust is one of the more interesting cases here, since Fig. 3 shows that the most distinctive of Proust's works are not his early or late works, but those in the middle. Proust's *Le Temps retrouvé*, known in English as *Finding Time Again* and *The Past Recaptured*, appears on Said's syllabus with the possibly self-translated hybrid title *Time Recaptured*. In his introduction to *On Late Style*, Michael Wood notes that Said's papers for his Late Style course contains passages from Proust, following a note about "conversion of time into space" (Said,

2006, pp.5–7). Proust’s novel, the final installment of his magnum opus *À la recherche du temps perdu*, deals explicitly with time and aging. As Adam Watt describes one such passage: “after long illness and absence from society life, the Narrator returns to one last matinée at which he meets many figures from his distant past. Time has changed them, aged and distorted their faces, their gait” (Watt, 2011, p.17). In particular, the narrator sees M. d’Argencourt, whom he regards “as a puppet, a trembling puppet with a beard of white wool” (quoted in Watt, 2011, p.17). This is a picture of old age in which the body is no longer in control of itself (Watt, 2011, p.17). Since this is Proust’s last novel, we are tempted to read this as a projection of the sick and dying Proust himself. But the publication chronology of *À la recherche* complicates this view.

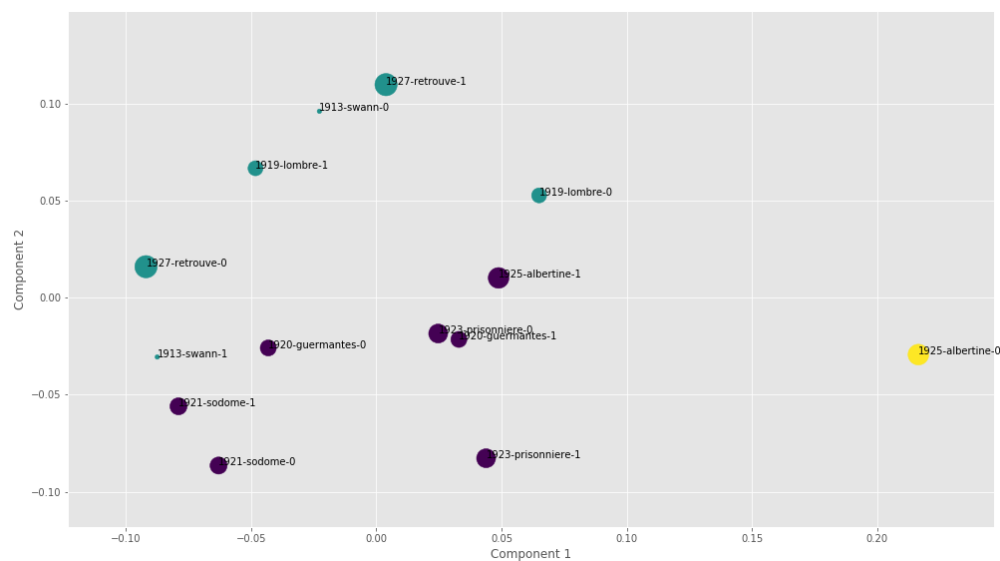


Fig. 3: PCA of Proust

Proust oversaw the publication of the first four volumes in the series, *Du côté du chez Swann*, *À l’ombre des jeunes filles en fleurs*, *Le Côté de Guermantes*, and *Sodome et Gomorrhe*, but the last three, *La Prisonnière*, *Albertine disparue* (also called *La Fugitive*), and *Le Temps retrouvé*, were all published posthumously from his manuscripts. Although Proust died in 1922, the last three works were published in 1923, 1925, and 1927, respectively. But they weren’t written in this order. Marion Schmid writes that Proust had written large parts of *Le Temps retrouvé* between 1910 and 1911, before the publication of the first volume (Schmid, 2001, p.64). In 1913 he imagined *À la recherche* to have three parts: *Du Côté du chez Swann*, *Le Côté de Guermantes*, and the third *Le Temps retrouvé* (Schmid, 2001, p.66). Proust expanded the novel in 1914, but expanded it from the middle, adding the ‘Albertine cycle’ of *La Prisonnière* and *Albertine disparue*, while the beginning and the end remained the same

(Schmid, 2001, p.67). This seems to explain how these works are statistical outliers here. Said was right about Proust's late style, after all, but not in the way he imagined.

6.3 Case Study: Thomas Mann

Thomas Mann was an appropriate subject for this study since he is so frequently discussed in Said's *On Late Style*, and since *Doktor Faustus* appears on Said's syllabus. Mann had a long writing career, publishing the short story collection *Der kleine Herr Friedemann* in 1898 at age 23, and the novella *Die Betrogene* in 1953, at age 78, two years before his death in 1955. *Doktor Faustus* was published fairly late, as well, in 1947. Said's interest in *Doktor Faustus* owes in part to Adorno's contributions, and to the fact that it directly treats, as well as exhibits, late style, such as in Adrian Leverkühn's discussion of late Beethoven, a passage which Said identifies as "pure Adorno":

Beethoven's art had overgrown itself, risen out of the habitual regions of tradition, even before the startled gaze of human eyes, into spheres of the entirely and utterly and nothing—but personal—an ego painfully isolated in the absolute, isolated too from sense by the loss of his hearing; lonely prince of a realm of spirits, from whom now only a chilling breath issued to terrify his most willing contemporaries, standing as they did aghast at these communications of which only at moments, only by exception, they could understand anything at all. (Said, 2006, pp.23–24)

Said cites this passage as a further definition of late style, as given here by Mann and Adorno, and also as an illustration of Adorno's preoccupation with "the figure of the aging, deaf and isolated composer," one whose late works "often communicate an impression of being unfinished" (Said, 2006, pp.24–25). But is *Doktor Faustus* itself an *instance* of late style? It was one of Mann's last works, created when he was 72 years old, only eight years before his death. Yet critics typically regard it as a stylistically and aesthetically consummate work of Mann's, not only not "unfinished," or in Adorno's terms, "ravaged," "furrowed," and "bitter," but Mann's greatest achievement. As J.P. Stern puts it, "the impression ... we are left with is one of immense narrative ease behind the complex syntactic devices ... Of this idea and mode of life *Doctor Faustus* is the final and greatest embodiment" (Stern, 1989, p.738).

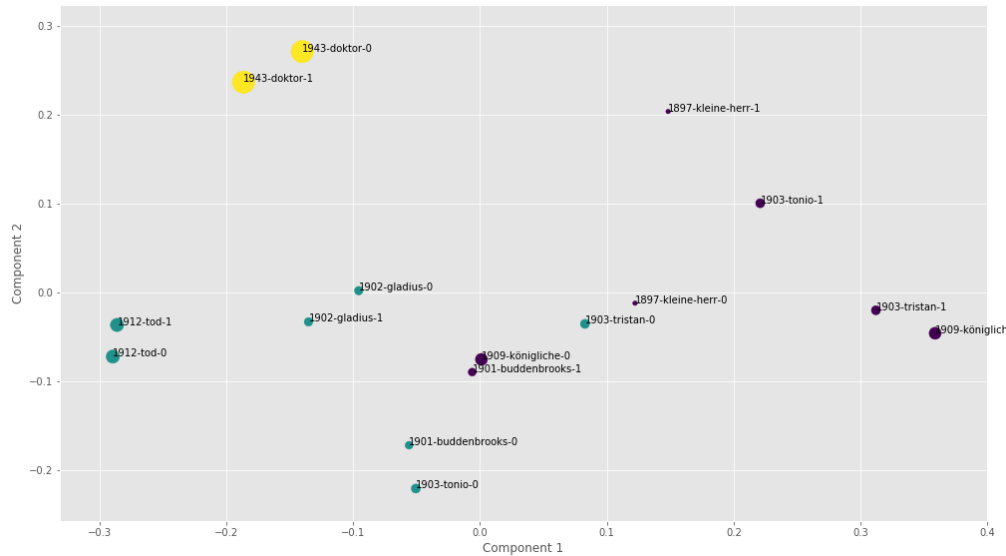


Fig. 4: PCA of Thomas Mann Novels

Fig. 4 shows the results of the PCA of Mann’s novels. Unfortunately, there are gaps in this corpus, due to the unavailability of electronic texts after the 1920s, which explains the absence of orange or green-colored points here. The most typical Mann novels occur closest to the mean, at (0,0), and the most atypical novels are the outliers. On the whole, the PCA shows principal components that align fairly well with chronology: the earliest works are in the southeast, and the chronology progresses fairly evenly toward the northwest.

Doktor Faustus, represented in its three parts as the three large red circles in the northwest corner of the chart, is fairly atypical. The beginning of the novel—the 0th part—is the most atypical of the three, while the two remaining parts are more typically Mannean. The most atypical work, however, the one with the highest L_5 norm, is not *Doktor Faustus* at all, but Mann’s very early work, *Der kleine Herr Friedemann*. As a collection of short stories, this probably explains why the style is so divergent from his other works. *Königliche Hoheit* (“Royal Highness”), another relatively early work of Mann’s, has the second highest L_5 norm. As a romantic comedy and “a story of princely initiation told in the easy style of an ironic fairy tale for grown-ups” this is atypical for Mann, and generically experimental (Stern, 1989, p.737).

Interestingly, although it does not have a high L_5 norm, the most proximate novel in the galaxy of *Doktor Faustus* is *Der Tod in Venedig* (“Death in Venice”). Said recognizes this novel as an instance of late style that is stylistically, although not chronologically, late. The titular death of the novel is a suggestion of mortality that Said correlates with late style. As he explains, “within Mann’s novella, Aschenbach’s half-aware and yet inevitable voyage to Venice induces in the reader the sense that because of various premonitions and past

associations (e.g., Wagner’s own death there in 1883) and its own peculiar character, Venice is a place where one finds a special finality” (Said, 2006, p.164). Said argues that it is “paradoxical” that the work is an early one in the author’s chronology, due to its “autumnal and even at time elegiac qualities.” Of Benjamin Britten’s operatic adaptation, too, he questions whether it “can be regarded as being in more than a chronological sense a *last work*” (Said, 2006, p.163). In this sense, we might view the proximity of *Der Tod in Venedig* with *Doktor Faustus* as an indication that the work is, as Said suggests, late.

6.4 Case Study: Virginia Woolf

The novels of Virginia Woolf present another good opportunity for stylometric analysis, since they resist periodicity. Recall that, in Fig. 1 above, Woolf has the lowest periodicity scores of all the writers in this smaller, curated corpus. This is also an effect noticeable in Fig. 5. The ‘latest,’ or most statistically distinct, novel in Woolf’s corpus is in fact her 1931 experimental novel, *The Waves*. This should come as no surprise to those that have read it: its form, which in earlier drafts Woolf termed a ‘play-poem,’ stretches the genre of the novel toward poetry and/or drama. In fact, as the colors of this figure indicate, the clustering algorithm has categorized *The Waves* as its own cluster, indicating it considers this book to be the only one that exhibits Woolf’s late style. Woolf published two more novels, however: *The Years* (1937) and *Between the Acts* (1941). *The Years* clusters together with Woolf’s first novel, *The Voyage Out* (1915), and *Between the Acts* clusters together with her third, *Jacob’s Room* (1922).

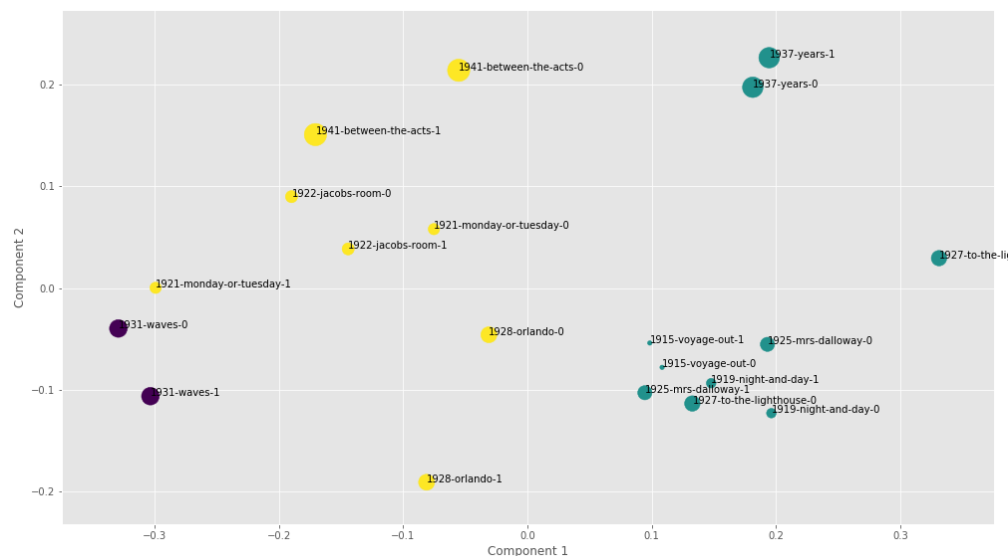


Fig. 5: PCA of Woolf Novels

Interestingly, although most of the novels here cohere—that is, both the first and second samples appear close to each other—one novel, *To the Lighthouse*, appears with its two samples in very different areas of this vector space, and thus at highly divergent lateness values. The first part of the novel appears in a cluster with *The Voyage Out* (1915), *Night and Day* (1919), and *Mrs Dalloway* (1925), and carries a relatively low lateness score (L_5 norm) of 0.19. In contrast, the second part is the second most distinct sample of all, after *The Waves*, with a lateness score of 0.37. It is clear that the stylometric analysis here detects the middle section of *To the Lighthouse*, an experimental section with a very distinct voice, which Woolf originally wrote and published as a separate work.

Here again, of course, we can make more of a case for a cohesive early, rather than late style, since all Woolf's 1910s novel cluster together, along with *Mrs Dalloway* and the first part of *To the Lighthouse*. The late works, in contrast, appear all over this chart. It appears that, early in Woolf's career, her style was fairly regular, but like James Joyce and other modernists, she became bolder in her experimentation later in her life.

7 Results

7.1 Generated Corpus

The above findings may be interesting, but the corpus is consciously biased, since I chose to test the works of writers whom critics have already selected as exhibiting late style. A much larger corpus should mitigate this somewhat. Fig. 6 shows average lateness scores for all the writers in the generated corpus. Of special note here are the authorial appellations *Anonymous* and *Various*, which of course include the works of more than one person. I've included these just out of curiosity, but am surprised to see that they have almost the lowest lateness scores.

At a glance, it is apparent that there are more negative scores than positive ones, and that the negative scores show up to a 50% stronger signal. The mean lateness score for the corpus is -0.019, meaning that, among these 51 authors, early style is more statistically significant than late.

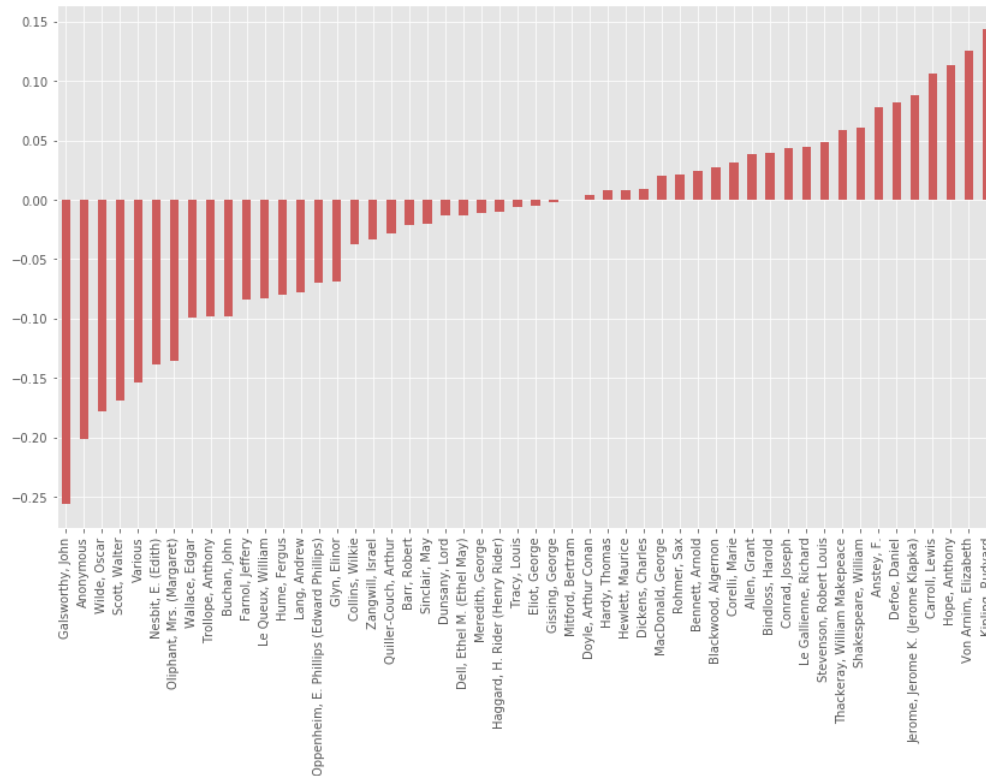


Fig. 6: Generated Corpus, Lateness Scores

Again, the periodicities do not show the same trends. Fig. 7 shows that most writers show strongly-defined periods (or that the periodicities of most writers may be computationally predicted). We might say, reading this graph, that John Galsworthy has the most regular style of these writers, and Lewis Carroll the most periodic. Perhaps Galsworthy's regularity might be owed to his long-running series, *The Forsyte Saga*. Carroll's periodicity is surprising, since he mostly writes childrens' fantasies. But maybe it is precisely the fantastic signal—that of imagination—that bleeds into style, and thus periodicity?

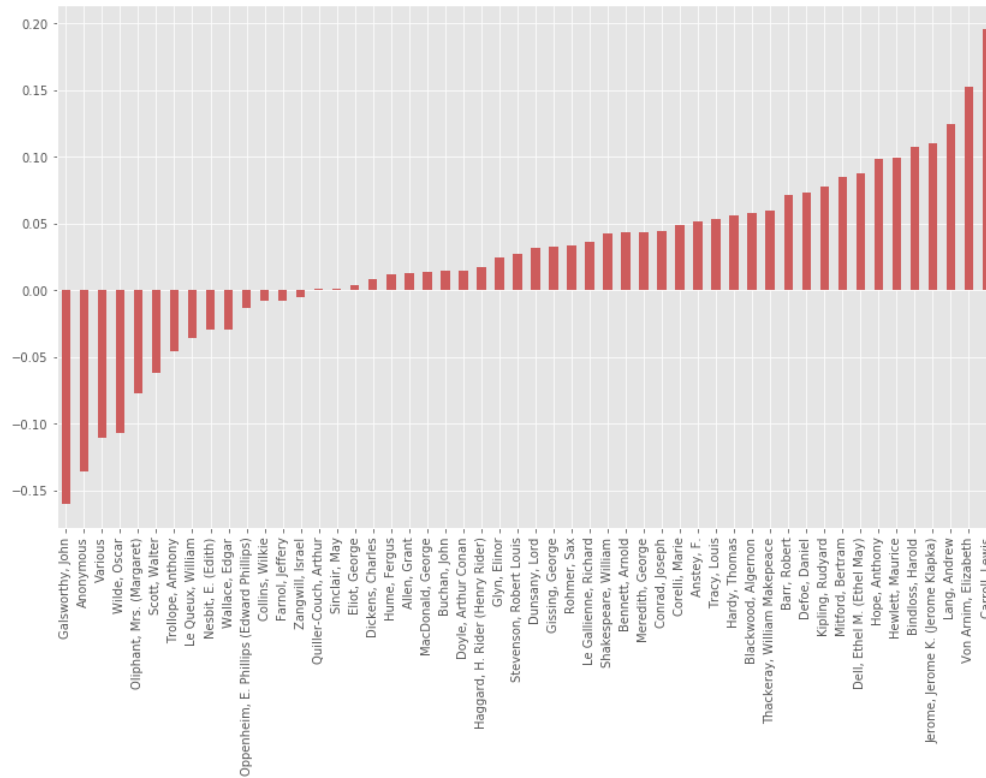


Fig. 7: Generated Corpus, Periodicities

This picture changes fairly dramatically if we shift our definition of lateness to mean proximity to death. In my opinionated operationalization of this effect, most writers exhibit late style. The abundance of writers showing lateness scores of exactly zero means that these are writers who have no texts (in my corpus, at least), published within five or ten years before their deaths.

There are also striking changes from other interpretations of style. Most noticeably, Shakespeare moves from showing only a minor lateness score to showing the highest, and almost double that of the rest. In a more dramatic shift, Oscar Wilde moves from having the fourth-lowest lateness score to the second-highest, by this metric.

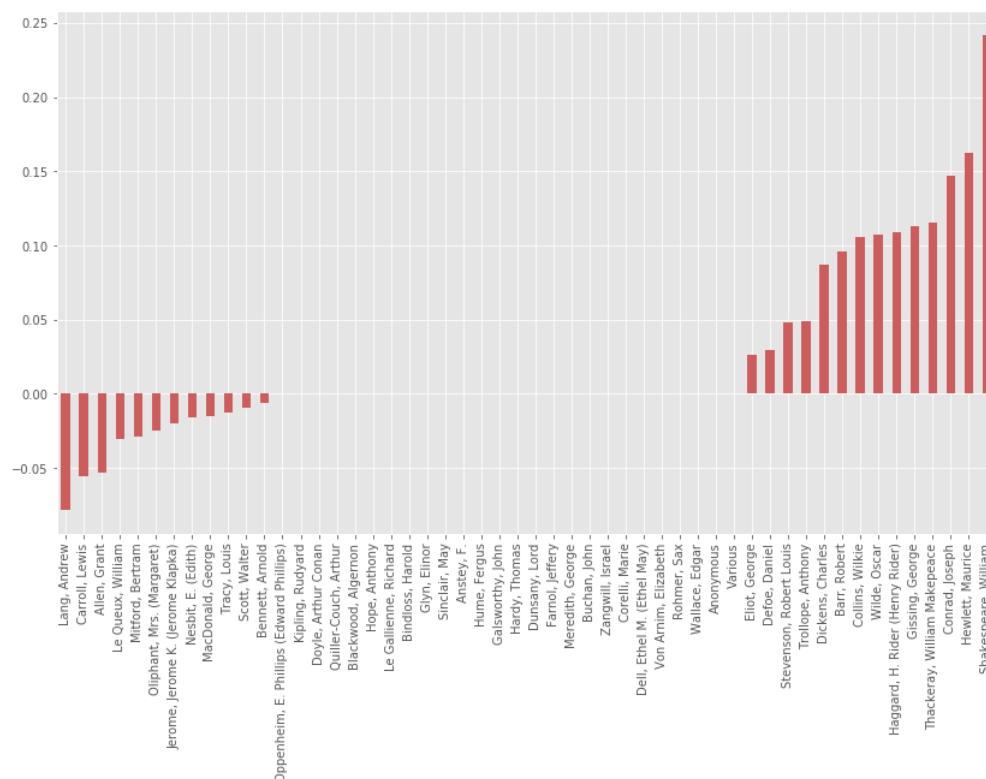


Fig. 8: Generated Corpus, Lateness as Time to Death

However, this provides only a very weak case for late style. The mean lateness score of all 51 writers is only 0.022, and if we remove Shakespeare it falls to 0.018.

7.2 Style as Content

If we consider style as content, as in Fig. 9, the picture is not very different from that of style as style. A few writers change position, from showing distinctive early styles to showing distinctive late styles: Oscar Wilde changes polarity like this, for instance. Here again, the mean lateness value is negative, -0.029, meaning that early style is more distinctive than late according to this measure, and that it is even more distinctive than when we operationalize style using stylometry.

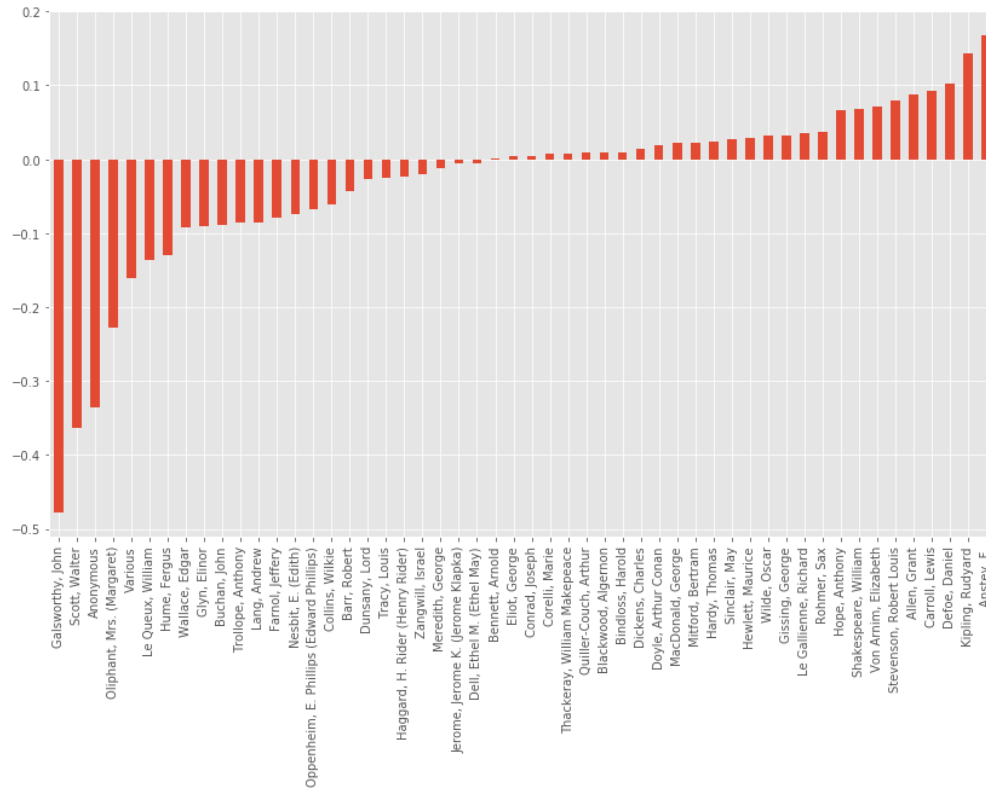


Fig. 9: Generated Corpus, Word Embeddings, Lateness

Periodicities for these writers, as measured by style-as-content, shift somewhat, as shown in Fig. 10. According to this metric, most writers now show negative periodicities, meaning that their works are not easily clustered into distinct periods on the basis of their words' meanings alone.

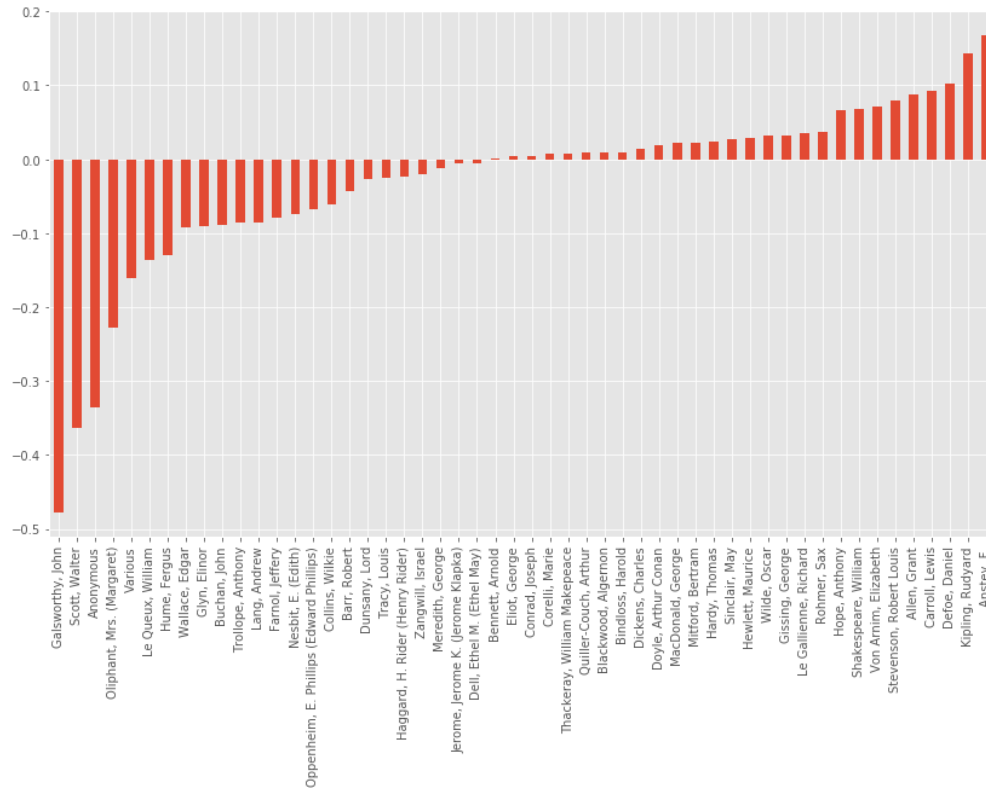


Fig. 10: Generated Corpus, Word Embeddings, Periodicities

8 Conclusions

Here is a summary of results:

1. In a test of 900 books by 51 British writers, early style is more statistically significant than late style.
2. Yet in a test of hand-picked writers known for their late style, late style is more significant.
3. Most writers show two or three statistically discernable stylistic periods, when style is interpreted as the way in which one writes, i.e. with which function words.
4. When we interpret style as content, however, these periods are much less pronounced, and most writers do not have statistically discernable styles (negative periodicities).
5. When we interpret style as content, late style is less pronounced, and early style is more pronounced. Here again, early style is more statistically significant.

It is clear from this analysis that ‘late style,’ as a critical concept, should be treated with suspicion, at best, and that ‘early style’ should be recognized as a better candidate for critical attention.

8.1 Further Work

Beyond simply testing and disputing the claims of Said and Adorno, the techniques developed here may be applicable to other problems in single-author corpora. Further analyses might include investigations into the effects of marriage on a writer, or parenthood. Furthermore, comparative stylochronometry might also prove fruitful: a preliminary comparison of Jane Austen and George Eliot novels show that while the two begin by writing in very different styles, their later works are stylometrically very similar. Another analysis might also involve a comparison of literary biographies with the works of that author. More work in this area is certainly necessary.

References

- Burrows, J.** (2002) 'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing*. 17 (3), 267–287.
- Burrows, J. F.** (1987) *Computation into criticism: A study of Jane Austen's novels and an experiment in method*. Clarendon Press.
- Craig, H. & Kinney, A. F.** (2009) *Shakespeare, Computers, and the Mystery of Authorship*. Cambridge University Press.
- Forsyth, R. S.** (1999) Stylochronometry with substrings, or: A poet young and old. *Literary and Linguistic Computing*. 14 (4), 467–478.
- Hoover, D. L.** (2007) Corpus Stylistics, Stylometry, and the Styles of Henry James. *Style*. 41 (2), 174–203.
- van Hulle, D. & Kestemont, M.** (2016) Periodizing Samuel Beckett's Works: A Stylochronometric Approach. *Style*. 50 (2), 172–202.
- Kenney, S. M.** (1975) Two endings: Virginia Woolf's suicide and "between the acts.". *University of Toronto Quarterly*. 44 (4), 265–289.
- Klaussner, C. & Vogel, C.** (2015) 'Stylochronometry: Timeline prediction in stylometric analysis', in *International Conference on Innovative Techniques and Applications of Artificial Intelligence*. 2015 Springer. pp. 91–106.
- Lipking, L.** (1984) *The life of the poet: Beginning and ending poetic careers*. Beginning and ending poetic careers. University of Chicago Press.
- Love, H.** (2002) *Attributing Authorship: An Introduction*. Cambridge University Press.
- McMullan, G.** (2007) *Shakespeare and the Idea of Late Writing: Authorship in the Proximity of Death*. Cambridge University Press.
- McMullan, G. & Smiles, S.** (2016) *Late style and its discontents: Essays in art, literature, and music*. Oxford University Press.

- Mikolov, T. et al.** (2013) 'Distributed representations of words and phrases and their compositionality', in C. J. C. Burges et al. (eds.) *Advances in neural information processing systems* 26. Curran Associates, Inc. pp. 3111–3119. <http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>.
- Neumann, E.** (n.d.) 'Art and time', in *Art and the creative unconscious*.
- Pennington, J. et al.** (2014) 'GloVe: Global Vectors for Word Representation', in *Empirical Methods in Natural Language Processing (EMNLP)*. 2014 pp. 1532–1543. <http://www.aclweb.org/anthology/D14-1162>.
- Pollnitz, C.** (1995) Cough-prints and other intimacies: Considerations in editing lawrence's later verse'. *Editing DH Lawrence*.
- Rand, W. M.** (1971) Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*. 66 (336), 846–850. <https://www.tandfonline.com/doi/abs/10.1080/01621459.1971.10482356>.
- Rybicki, J.** (2015) Vive la différence: Tracing the (Authorial) Gender Signal by Multivariate Analysis of Word Frequencies. *Digital Scholarship in the Humanities*. fqv023. <http://dsh.oxfordjournals.org/content/early/2015/07/07/llc.fqv023> (Accessed 12 July 2016).
- Said, E. W.** (1989) *Edward Said Papers*.
- Said, E. W.** (2006) *On late style : Music and literature against the grain*. New York: Pantheon Books.
- Schmid, M.** (2001) 'The birth and development of A la recherche du temps perdu', in Richard Bales (ed.) *The Cambridge Companion to Proust*. Cambridge: Cambridge University Press. pp. 58–73. <http://universitypublishingonline.org/ref/id/companions/CBO9780511999727A008> (Accessed 5 May 2016).
- Smit, D.** (1988) *The language of a master: Theories of style and the late writing of henry james*. Southern Illinois University Press.
- Smith, P. W. H. & Aldridge, W.** (2011) Improving Authorship Attribution: Optimizing Burrows' Delta Method*. *Journal of Quantitative Linguistics*. 18 (1), 63–88. <http://dx.doi.org/10.1080/09296174.2011.533591> (Accessed 19 May 2017).
- Solomon, M.** (2004) *Late beethoven: Music, thought, imagination*. University of California Press.
- Stern, J. P.** (1989) 'Thomas Mann', in *European writers*. New York: Scribner.
- Van Dam, F.** (2016) *Anthony trollope's late style: Victorian liberalism and literary form*. Edinburgh critical studies in victorian culture. Edinburgh University Press.
- Watt, A. A. (. A.** (2011) *The Cambridge Introduction to Marcel Proust*. Cambridge: Cambridge University Press.

Wheeldon, M. (2009) *Debussy's late style*. Musical meaning and interpretation. Indiana University Press.

1. After I'd completed the first draft of this paper, I learned of two other works-in-progress that test the idea of late style with computational text analysis: an unpublished work by Andrew Piper, and an article by Simone Rebora and Massimo Salgaro: "Is 'Late Style' Measurable? A Stylometric Analysis of Johann Wolfgang Goethe's, Robert Musil's, and Franz Kafka's Late Works," which will soon appear in *Elephant & Castle* 18.↵
2. It would be far better, of course, to use the dates these works were written, rather than published, but that data is much harder to find in a machine-readable form. I decided that, since the majority of books are published in their writers' lifetimes, and chronologically according to when they're written, the gains in accuracy would not be worth the time it would take to find this data.↵
3. Since many of these texts are not yet out-of-copyright, obtaining them in some cases has meant special-ordering their delivery from France, cutting the binding with a 'book guillotine,' and scanning them with a sheet-feeder.↵
4. The source code for all of these computations may be found at <https://github.com/JonathanReeve/late-style-PCA>↵