e-ISSN 2724-3923

magazén
Vol. 1 – Num. 2 – December 2020

A Brief History of the Theory and Practice of Computational Literary Criticism (1963-2020)

Chris Beausang
Maynooth University, Ireland

Abstract This paper will construct a history of computational literary criticism (CLS) which has engaged statistical methods by providing an historical account of the journal articles as well as other publications which have advanced the field to the most significant extent since 1963. This paper divides the history of CLS into three distinct epochs, within each of which the methods and theories CLS scholars utilise undergo significant qualitative transformation. The decisive factor in each of these epochs is CLS' relationship to traditional literary criticism. Partly as a result of this, CLS scholarship initially cleaves to organic theories of literary style and adopts a highly polemicised opposition to then-regnant post-structuralist theories of authorship.

Keywords Computational literary criticism. Literary criticism. Digital humanities. Delta. Machine learning.

Summary 1 Introduction. – 2 Embryonic CLS (1963-1979). – 3 PCA & Proto-Delta (1980-1990). – 4 Delta, Results and Prospects (2000-2020). – 5 Conclusion.

Edizioni Ca'Foscari

Peer review

Submitted	2020-10-09
Accepted	2020-11-02
Published	2020-12-22

Open access

Citation Beausang, C. (2020). "A Brief History of the Theory and Practice of Computational Literary Criticism (1963-2020)". *magazén*, 1(2), 181-202.

DOI 10.30687/mag/2724-3923/2020/02/002

1 Introduction

The heterogeneity of computational literary studies (CLS) can render it difficult to map. Its being integral to the digital humanities has the consequence that it would not be uncommon for an article or a chapter which marks a significant advance within CLS to appear in the same publication as articles written on topics as varied as 3D modelling or database ontology. This paper will nevertheless make an effort to marshal a significant proportion of the field's research output into an historical narrative which is capable of encompassing developments underway in the field since the early sixties. It is in the eighties and nineties that we begin to see previously regnant methods consistently outperformed by multivariate approaches in which ~100 of the most frequent words (MFWs) in a text are quantified. In the oughts and tens CLS scholars extended these methods further, analysing thousands of words and treating texts more or less in their entirety. In accounting for these three phases, this chapter will emphasise particular works of scholarship which have been instrumental in transforming one epoch into the next. Any chronology which accounts for the discipline's history will be a generalising one and will require the omission or simplification of particular phenomena. Some articles anticipate transformations within the discipline which are later to take place and, as we will also see, CLS scholars are sometimes prone to continuing to use methods which have been shown to be inadequate. The periodisation here proposed allows us to introduce both superstructural and infrastructural causes in considering the history of the discipline. The tendency to focus on isolated formal features in the discipline's early days is symptomatic of its inclination to reverse the death of the author at the hands of figures such as Roland Barthes and Michel Foucault, re-emphasising the individual agency and style of the author. In its early history therefore, CLS manifests an inclination towards more romantic theories of authorship, a tendency which results in the development of methods which assume that all texts written by different authors are differentiable on the basis of parameters which are wholly arbitrary. As we will see, it is not until the success of John Burrows' Delta method that this notion begins to be challenged.

2 Embryonic CLS (1963-1979)

As Jack Grieve notes, there is a long history of mathematics being brought to bear on the study of attributing authorship, reaching back to the nineteenth century (Grieve 2007, 251). However, in his history of the field, David Holmes identifies the first instance of modern stylometry in Frederick Mosteller and David Wallace's attempts to iden-

tify authorship in the twelve pseudonymously written essays and articles in *The Federalist Papers* written by Alexander Hamilton, James Madison and John Jay in the late eighteenth century. Mosteller and Wallace attribute authorship on the basis of similar rates at which function words are deployed in the text, such as prepositions, conjunctions and articles (Holmes 1998, 112). Fred Damerau seems to be the first to account for the use of function words from a theoretical perspective, citing W.J. Paisley's theory of "minor encoding habits". According to Paisley, in turn drawing from theories developed in the field of art history, indices of personal style can be found in minor, but highly common, features of a work. They should not vary significantly between works produced by the same author but should vary significantly between works produced by different authors. In satisfying these criteria, Damerau identifies function words as being most suitable (Damerau 1975, 271-2). Damerau's approach is perfectly logical and given that it is wholly appropriate to reduce most authorship attribution problems to what Burrows refers to as a "closed game", where there is a restricted set of texts and candidate authors, function words seem to be capable of providing promising results. However, this culminates in the assumption that authorship is in and of itself a guarantor of a distinctive or individual style which suffuses the work in its entirety. It is therefore assumed that each text produced by a single author is statistically homogenous and that any given quantity of features identified in a text written by one author will be statistically distinct from the same feature in a text written by another author, under the assumption that this can be confirmed through the use of Mann-Whitney, chi-square, Student's *t* and Fisher tests. As this paper proceeds, we will see that the influence of this assumption is detrimental. Barron Brainerd's work, in its capacity to identify and willingness to test the resilient assumption of intra-authorial heterogeneity, represents an exception and Brainerd is therefore among the first to identify many of the drawbacks associated with the use of the chi-square method when applied to literary texts (Brainerd 1975, 161; 1979, 5-12).

Though there are significant numbers of papers in *Computers and the Humanities*' (*C&H*) and *Literary and Linguistic Computing*'s (*LLC*) early history which abide by sound statistical and methodological practice, such as Paule Sainte-Marie, Pierre Robillard and Paul Bratley's application of principal component analysis (PCA) to 44 MFWs in 30 plays written by Molière (Sainte-Marie, Robillard, Bratley 1973, 136) and Brainerd's application of cluster analysis in order to differentiate novels from romances (Brainerd 1973, 267), many CLS articles until the nineties can be characterised by the arbitrariness of their methodological approaches. Sampling, variable selection and statistical measurements are often adopted and applied without explicit reasoning or reference to previously undertaken studies with-

in which the efficacy of these methods have been validated. Citations are also less common in early articles than they later become, and this has the effect that the precise rationale for any given procedure being carried out is more often assumed than explained. Robert Cluett's analyses of part-of-speech (POS) entities in Restoration-era prose and John Foley's analyses of stress patterns in Beowulf represent another tendency rife at this early stage in CLS history, which takes a heuristic approach to drawing conclusions rather than using proven mathematical techniques (Cluett 1971, 264-8; Foley 1978, 78). These defects can probably be accounted for by bearing in mind the nascency of the field. As M.W.A. Smith notes, at the time of writing in 1987 there was no extant corpus of studies undertaken which had successfully inculcated an understanding of statistical best practice when analysing literary texts and CLS scholars could not benefit from a corpus of articles on which to base their approaches in the same way a would-be CLS scholar could today (Smith 1987, 146). Other constraints which exert a significant influence on the early scholarship include the available infrastructure. Computational memory limits, which would have been a factor in experimental design, go some way also in explaining the methodological focus we see on quantifying the frequencies of a very small number of function words. Computing was also expensive and, prior to the sharing of digital texts via the internet, each researcher would need to build their own corpus (Sainte-Marie, Robillard, Bratley 1973, 131-2; Sula, Hill 2019, 191).

The early polemics which we find in the first issues of *C&H* are illustrative as regards the 'theory wars', a consistent feature of CLS discourse. It is Louis Tonko Milic who initiates this dialogue, both in *A Quantitative Approach to the Style of Jonathan Swift* (1967) and in two articles which argue for the significance and contributions computing may potentially make to the study of literature. Milic's arguments are based on the capacity of computing to alert the critic or analyst to patterns and trends which are not detectable via traditional, qualitative approaches. This is particularly important as, from Milic's perspective, words which are traditionally deployed in the interrogation or analysis of style in literary criticism are vague or impressionistic. Milic partly attributes this to the blurring of the boundary between literary criticism and social theory (Milic 1967, 27-8, 38, 54). In solving this problem, Milic wished to facilitate a synthesis between computation and the creative intuition which has historically predominated within literary criticism rather than automating the latter out of existence (1966, 5). Milic begins from the notion that syntax may provide a deep and unifying structure or promising a starting point for quantitative approaches (1967, 32, 79) and proceeds by dividing words into twenty-four different grammar-types, looking at how the means of these word-types increase or decrease in Swift's writings over time. Milic then carries out close readings

e-ISSN 2724-3923

of these grammar types in their context within the works (1967, 32, 79, 174, 205, 272). It is Emmanuel Mesthene who presents the first sceptical response, arguing that for all the precision and accuracy which computational tools have the potential to introduce, they also bring bias to literary-critical research as computing cannot serve as a neutrally clarifying agent (Mesthene 1969, 2). Bruce A. Beatie cites C.P. Snow's essay "The Two Cultures" (1959), in locating literary studies within a school of thought totally opposed to that of statistics (Beatie 1979, 186-7). Susan Wittig objects to CLS on the basis of a more overt commitment to post-structuralism, which envisions the text as an ineffable system of exchange which resists all forms of hierarchical categorisation (Anderson 1983, 68). This is utterly contrary to the ways in which natural language processing (NLP) and linguistic analysis require us to regard text (Wittig 1977, 211-2). Despite being written more than half a century ago, these three critics broadly anticipate the two opposed positions we now confront in considering CLS' relationship with the broader literary-critical milieu, even to the present day. Milic, on the one hand, emphasises the capacity of computation to allow the critic to exceed their individual point of view and potentially gain access to an hypothesised deep structure, while CLS' detractors mount an overall objection to CLS in principle, refraining from engaging with statistical methods themselves or a history of their application on the basis that empiricism is an inveterately instrumentalised and insufficiently reflexive form of knowledge production. As this chapter continues, we will see that these two positions and the tensions residing within them are crucial to any account of CLS' history.

3 PCA & Proto-Delta (1980-1990)

In the eighties we see John Burrows publish analyses that anticipate the Delta method he would later develop. Burrows begins by focusing on the changing rates at which modal auxiliaries are used in six novels written by Jane Austen (Burrows 1986, 9). Though Burrows argues his approach allows for the treatment of texts in their entirety, against literary criticism's historical tendency to focus on highly specific features of a work, in his focus on modal auxiliaries and how they relate to sentence length, Burrows remains constrained within the framework he aims to supersede (20-3). In his second article in *C&H*, Burrows attempts to quantitatively differentiate three different narrative categories which he identifies as being at work in Austen's novels; dialogue, 'pure narrative' – here meaning the voice of the narrator alone – and 'character narrative', here meaning the voice of the narrator mediated by the thoughts or feelings of a particular character, elsewhere referred to within literary criticism as 'free in-

direct discourse'. Burrows first correlates the frequencies of a list of function words which appear in each of these three categories, then applies a statistical transformation to these correlation coefficients. The aim of this method, PCA, is to reduce the dimensionality of a dataset consisting of a large number of variables. This is achieved by combining these variables into new variables called 'principal components'. Each principal component encompasses a specific amount of the variation within the original data, to the extent that a two-dimensional visualisation is generally sufficient to provide an insight into the data's underlying structure (Binongo, Smith 1999; Joliffe 2004, 1). Burrows applies this method in a series of distinct permutations, firstly separating the three different narrative types by gender then by character, describing each time the clustering patterns which can be observed in relation to the literary-critical discourse surrounding Austen (Burrows 1987, 64-9). In his third article, Burrows applies his method to fifteen other nineteenth-century novelists. As before, Burrows is invested in identifying a unique and individual style for each author and though his graph has no temporal component, he argues that each author's *oeuvre* clusters chronologically and that Austen, George Eliot and Elizabeth Gaskell's relative distance from the other authors justifies reading their styles as individual, erecting a movement away from neo-classical prose styles which otherwise predominated in the late eighteenth and early nineteenth centuries (Burrows 1989, 318; Holmes 1998, 113).

Given their capacity to cluster texts on the basis of authorship, genre and era, function words remain central within CLS and we see a number of studies emerge which continue to demonstrate the efficacy of the method (Burrows 1992, 91-103; Craig 1991, 183-5; 1999, 222-40; Tse, Tweedie, Frischer 1998, 141-6). We also see further interrogations of PCA in and of itself in Binongo and Smith's investigations into its mathematical principles (1999, 445-64). Penelope Gurney's and Lyman W. Gurney's application of PCA to MFWs significantly outperforms attempts to attribute authorship on the basis of vocabulary richness, a statistic which is calculated by dividing the number of unique word types by the number of words in the text overall (Gurney P., Gurney L. 1998, 119-30). This result is replicated by Fiona Tweedie and R. Harald Baayen, who note that even measurements for vocabulary richness which are independent of text length are unsuccessful in discriminating texts on the basis of their authorship (Tweedie, Baayen 1998, 323-50). Attempts to identify a length-independent means of quantifying a text's lexical richness, for the logical reason that a shorter text will have far more unique word-types than a longer one, are a consistent fixture of CLS discourse, as we see in Philippe Thoiron's diversity or entropy-based method (1986) or John Baker's attempts to quantify the pace at which new vocabulary enters a writer's work (1988, 38-9). The centrality of vocabulary

richness to CLS may be attributed to theories of intra-authorial heterogeneity, but also to the measure's relative simplicity and comprehensibility. This is probably also the case for the persistence of measurements based on sentence, word and syllable lengths, which are also plagued by similar issues relating to reproducibility (Aoyama, Constable 1999). Gurney and Gurney recommend incorporating more MFWs into future analyses, computing space allowing (1998, 119-30).

Concurrent with the development of reliable multivariate statistical techniques in CLS, we also see previously regnant methods challenged for their failures to operate reliably. Thomas Merriam, for example, demonstrates the unreliability of 'proportionate pairs', a method used by A.Q. Morton, which assumes that particular pairs of words which exist in a fixed ratio to one another between texts are suggestive of shared authorship. Merriam demonstrates that more than random variation can often be observed in works produced by the same author (Merriam 1989, 252-3) while Michael Hilton and Holmes demonstrate the inadequacy of another method developed by Morton, wherein the incidence of two formal features are plotted on a line graph. The two lines are then superimposed on one another and it is determined that any instances in which these lines deviate from one another are indicative of the intervention of a second author. Hilton and Holmes propose a more statistically rigorous variant of this approach, which incorporates the weighting of particular features, but concludes that even with these improvements, they fail to reliably attribute authorship (Hilton, Holmes 1993, 73-80; Holmes 1998, 114). Smith also publishes a number of articles which challenge the use of chi-square tests, on the basis that they are prone to delivering Type II errors (Smith 1985, 3-10) as well as Morton's correspondence analyses, based on obtaining corresponding values of particular words in particular positions and collocation analyses, which quantify occurrences of a prescribed word either followed or preceded by a second prescribed word (Holmes 1998, 202; Smith 1987, 145-6). Smith goes on to criticise CLS scholars for using methods which are insufficiently rigorous and proposes instead analysing the rates at which the first word in every speech appears per 1,000 words in the works of six Elizabethan-era playwrights. Smith demonstrates his method's capacity to correctly identify John Webster as the most likely candidate of the six to have authored *The Duchess of Malfi* (1614) and Ben Jonson as the most likely to have authored *The Alchemist* (1610). On the basis of the seeming capacity of this method to function, Smith proposes George Wilkins as being the most likely to have authored *Pericles* (1619) (Smith 1988, 34-7). In the late eighties and early nineties, we see studies which continue to draw from discredited approaches such as the chi-square tests (McColly 1987, 174), the visual inspection of visualisations (Anderson, McMaster 1989, 343-5; Irizarry 1993, 88; Philippides 1988, 4), but these in-

creasingly represent the exception. Even in instances in which PCA is not deployed, in favour of more generic visualisation of distances, analyses employ increasing numbers of variables (Greenwood 1992, 44-7; 1993, 216-9; Irizarry 1991, 176-8). While approaches such as these neglect to aggregate the results of these dendrograms or line graphs as one would within the context of bootstrapping, they still represent the movement of CLS towards holistic analyses of text and a heterogenous number of quantitative methods.

Criticism of CLS in this era continues to maintain the inadequacy of scientific methods operationalised within literary criticism. Both Roseanne Potter and W. van Peer argue that literary studies weigh evidence in a way which is qualitatively different to statistics, which by necessity requires overlooking the process-like nature of literary expression (Potter 1988, 94; van Peer 1989, 303). The difficulty in providing an account of these debates is that neither side, whether they happen to be invested in maintaining a strong post-structuralist current within literary criticism or CLS scholars who wish to render literary studies more empirical, are interested in clarifying or examining what the other side is doing. Even though the milieu at this time would seem to be ripe for the contribution of a scholar versed in both the historical of statistical methods and continental philosophy, such a synthesis unfortunately never materialises. Rather, the strawman which roughly equates one to reactionary politics and the other to an incoherent admixture of feminism and relativism, remains rife. We only need to consider Fortier's arguments that post-structuralist approaches to literature have moved beyond 'sense and reason' (Fortier 1991, 193) or Milic's that postmodernism, as manifested within the strain Milic regards as responsible for the death of the author, is nothing more than a mixture of 'victimisation theory', and 'Marxism' (Milic 1991, 394) to identify how much more heat than light has been generated in CLS scholars' engagements with literary theory.

4 Delta, Results and Prospects (2000-2020)

Burrows first presents the Delta method in 2001 in an attempt to move CLS beyond the quantification of authorship from within the context of the closed game, wherein only two or three authors may be presented as probable candidates within an analysis. The Delta method's capacity to incorporate large numbers of authors, Burrows contends, will allow for the development of CLS analyses which do not close off potential avenues of interpretation before the analysis has begun. Burrows' first use of the Delta method begins by identifying 30 MFWs, disambiguating some of his chosen MFWs on the basis of their grammatical function and expressing each MFW's frequency as a percentage of the number of words in the text overall. The dis-

tribution of each word is then normalised, such that each frequency is expressed in terms of the number of standard deviations it resides from the mean. The 'Delta score' is the mean of the differences between each word's normalised frequency. Through the use of this method, Burrows demonstrates that works by John Milton are less dissimilar to one another than they are to the works of twenty-four other seventeenth-century English poets. Burrows tests Delta with 150, 120, 100, 80, 60 and finally 40 MFWs, observing a decrease in attributional accuracy with each decline in quantified MFWs (Burrows 2002, 272-82). In an article published in Blackwell's *Companion to Digital Humanities* (2004) which analyses forty seventeenth- and eighteenth-century poems, Burrows divides his 150 chosen MFWs into three groups based on subjective readings of their function and applies Delta to each of them separately, trying to identify which of the three cohorts could be considered to be more denotative of authorship as compared with genre (Burrows 2004).

Before the use of the Delta method was taken up to a significant extent, Hoover published a number of articles which involved the application of distance measurements to word frequencies, albeit without normalising or relativising them. Hoover compares how rates of successful attribution are changed by altering the number of MFWs, sample size, methods of computing distance, or removing dialogue, pronouns or texts with a first-person narrator from the analysis. Hoover's analyses replicate Burrows' most significant overall finding, that the quantification of more MFWs increases the rate at which a text is successfully attributed and the most frequent bigrams such as 'it is', 'to the' and 'of the' may be even more effective in this regard (Hoover 2001, 421-38; 2002, 157-76; 2003, 261-82). As G. Bruce Schaalje et al. demonstrate, Delta does not quite allow CLS to definitely break from the problem of the closed game. By virtue of the way in which Delta operates, it in fact tends towards the generation of false positives if it is applied as a means of attributing authorship (Schaalje et al. 2011, 71-88). Scholars such as Patrick Juola have suggested a means by which Delta's tendency to do so can be reduced, by introducing a distractor corpus of true negatives, thereby raising the bar of similarity required if a text is to be identified as the most similar to any other (Juola 2015, i100-13). Even if Juola's proposed adjustment is successful, the central problematic remains in place and is in fact implicit in Burrows' initial terms of reference. Delta is thus best conceived as a means of analysing style in relational terms, rather than as a means of settling instances of contentious authorship. Yet it is a peculiarity of the early discourse that Delta's capacity to consider style in this manner is not considered to any significant extent. We see this in the context of two studies undertaken by Hoover. Hoover is firstly reticent to incorporate additional function words into an analysis, on the basis that this will lead to the quantification

of formal features which are within the conscious control of the author (Hoover 2004b). In a second study, Hoover attempts to improve attributional success by removing textual features such as contractions or personal pronouns from the analysis and then applies Delta to one or two texts divided into a number of different parts in order to see if Delta will cluster them with one another. Hoover's methods therefore again attempt to return to smaller-scale qualitative readings which emphasise the decisive impact of specific formal features (Hoover 2004c). The ongoing influence of Paisley's theory of minor encoding habits is the best means of accounting for why it is that the results of Delta analyses are so consistently passed over, despite the efforts of scholars such as David Mannion and Peter Dixon, who dispute Hoover's and others' focus on unconscious formal features in favour of understanding some other features as being consciously deployed (Mannion, Dixon 2004).

Hoover is the first analyst who aims to further optimise Delta by making quantitative adjustments to Burrows' original method. Hoover does so by treating positive and negative *z*-transformed relative frequencies differently, either by focusing on higher values, or squaring and summing positive and negative means in a number of different permutations. None of these approaches are successful in outperforming Delta outright (Hoover 2004a, 477-95) but proposed modifications are still widely applied and compared with one another, as in Holmes and Daniel W. Crofts (Holmes, Crofts 2010, 179-97). Daumantas Stanikūnas, Justina Madravickaitė and Tomas Krilavičius apply a further modification known as Eder's Delta, which applies weights to frequencies in order to moderate the influence of infrequent word-types (Stanikūnas, Madravickaitė, Krilavičius 2015, 1-7). Shlomo Argamon also attempts to improve Delta on mathematical grounds. Argamon (2008) points out that Burrows normalises word distributions by mean and standard deviation, an approach which would only make sense if the word frequencies were distributed normally, but applies a Manhattan distance, which assumes a Laplace distribution. Stefan Evert et al. (2017), in a subsequent publication which systematically assesses Delta's performance against that of its subsequent improvements, confirm that, based on results obtained from both English and German reference corpora, word frequency distributions are better represented by a normal than by a Laplace distribution. Given this instance of statistical error, Argamon proposes three improvements. The first is Linear Delta, which retains Manhattan distance but normalises the relative frequencies according to median and spread. The second is Quadratic Delta, which retains Burrows' method of normalising, but applies the more mathematically sound Euclidean distance to the word frequencies and finally, on the basis of Delta's doubtful assumption that word frequencies are independent, introduces a third adjustment, Rotated

Delta, which performs a whitening transformation on the word frequencies in order to render them independent from one another (Argamon 2008). Despite their greater degree of mathematical legitimacy, however, Argamon's approaches do not outperform classic Delta (Evert et al. 2017, 8; Jannidis et al. 2015). Peter W.H. Smith and W. Aldridge argue that, on the basis of the assumptions which Euclidean distance makes and the fact that its accuracy decreases as dimensionality – i.e. the number of MFWs we apply the distance measurement to – increases (Smith, Aldridge 2011), there may be an upper limit beyond which we should not quantify words when conducting a Delta analysis. Smith and Aldridge propose 200-300 MFWs as this upper limit, though, as Fotis Jannidis et al. argue, this figure is probably quite low and it may be a product of the fact that Smith and Alridge's study was based on an analysis of a corpus of poetic texts (Jannidis et al. 2015). Jacques Savoy's study, which applies Kullback-Leibler divergence, Burrows' Classic Delta and chi-square in a bid to identify the optimal number of differentiators, argues for between 300 and 500 terms (Savoy 2013). In demonstrating that cosine distance outperforms classic Delta, Evert et al. also note distinct behaviours at higher MFW frequency ranks; classic Delta peaks at ~1000-1500 MFWs and thereafter maintains more erratic behaviour, whereas cosine distance plateaus (2017, 14). Jan Rybicki and Maciej Eder not only quantify up to 3000 MFWs but also test particular strata, attempting to identify if Delta's success may be specific to a particular frequency rank. On the basis of the results obtained, Rybicki and Eder recommend quantifying the first 3000 MFWs (Rybicki, Eder 2011). Alexis Antonia, Hugh Craig and Jack Elliott investigate whether larger *n*-grams as opposed to individual words are more likely to correctly attribute authorship and find that the efficacy of the parameter varies from corpus to corpus (Antonia, Craig, Elliott 2014). Antonia, Craig and Elliott's conclusion that the optimal parameters and measures vary between corpora seems to be confirmed by studies such as Enrico Tuccinardi's, who demonstrates that character grams are more suitable in shorter documents (Tuccinardi 2016) and Lisa Pearl, Kristine Lu and Anousheh Haghighi's analysis of idiolect in epistolary literature, which allows for the weighting of some features as being more important than others (Pearl, Lu, Haghighi 2017). These findings culminate in the developing tendency towards the application of a diversity of methods applied to a similarly diverse set of parameters, for example discriminative words, word lengths, character-based frequency analysis, word-length, POS tags, measures for vocabulary richness, to which vector space representation PCA, hierarchical clustering, SVM, random forests, *k*-nearest neighbours, Delta or rolling Delta, the application of Delta to sequential windows of text, may be applied (Gladwin, Lavin, Look 2017; Hou, Jiang 2016; Saccenti, Tenori 2015; Sayoud 2012). Rybicki and Eder (2011) attempt

to generalise Delta's functionality by applying it to other languages, attaining high levels of success in French, German, Hungarian and Italian corpora but poorer results for Latin and Polish. Richard S. Forsyth and Phoenix W.Y. Lam as well as Rybicki and Magda Heydel apply Delta to translated texts, in an attempt to identify whether the stylistic signal of the author or translator predominates. Both find that the signal of the original author is more powerful, but the presence of different translators can be identified by comparing two different translations of the same author's works (Forsyth, Lam 2014; Rybicki, Heydel 2013, 708-17). Through the use of bootstrap consensus trees and network analysis, which involve the representation of texts and the relationships between them as discrete entities (Eder 2017), Changsoo Lee (2017) demonstrates that the further two languages are apart linguistically, the more likely it is that the translator's writing style will exert itself in comparison to that of the author.

The basic positions we confront in engaging the debates concerning the supposed incompatibility of CLS within literary criticism will by this stage of this paper be familiar and the argument that CLS is both overly generalising and insufficiently reflexive as a form of scholarly inquiry remains the predominant point of attack (Gooding 2013). However, we have not yet considered CLS scholars who have made a virtue of this charge to a certain extent, as in the literary criticism of Franco Moretti. It should be noted that Moretti's major works such as *Atlas of the European Novel* (1998), *Distant Reading* (2005), *Graphs, Maps, Trees* (2007) and *The Bourgeois* (2013) are not substantively computational or statistical in their approaches, but rather use maps, spreadsheets and diagrams in order to illustrate what are often quite traditional literary-critical hypotheses, holding out the possibility that literary criticism might aspire to the ambition and scope of quantitative sociology (Moretti 2007, 4-30; 2012, 67). Moretti's most notorious argument, that the development of industrial capitalism in nineteenth-century Europe (2012, 16-8; 2013, 14-21) paves the way for the emergence of modernist literature is not in and of itself a controversial one; this axiom more or less undergirds a significant amount of literary criticism conducted from a Marxian perspective. Moretti's reception has more to do with what is perceived as his method's apologia for the literary-critical school referred to as the world literary system as it has been developed by Pascale Casanova (Cleary 2006). Criticism of this school has been trenchant from postcolonial scholars such as Emily Apter and Christopher Prendergast on the basis of its tendency towards national chauvinism, imperialist logic and uncritical handling of the relationship between modernisation and the canonisation of literature (Prendergast 2004; Apter 2013, 42-58). However, the publication of Moretti's writings, and responses to them, in pre-eminent venues such as *n+1* and *New Left Review* (Allison et al. 2012; Moretti 2020) has the consequence that

these criticisms have a tendency to assume the shape of criticisms of CLS in general, despite the lack of actual quantification in Moretti's work. That Moretti's far less provocatively post-political analyses of POS tags and word frequencies in the context of the Stanford Lit Lab (Algee-Hewitt, Heuser, Moretti 2015; Allison et al. 2013, 2011) have not been critiqued to the same extent attests to the fact that it is Moretti's more traditional literary-critical work which can be criticised on the basis of its Eurocentricity

Critics who continue to maintain CLS scholars' dependence on reductive or categorical reasoning at this time begin to advocate for more exploratory or interpretative approaches (Escobar 2016, 85; Sinclair 2003) and we might consider Steven Ramsay, Joanna Drucker, Bethany Nowviskie and Jerome McGann symptomatic of this tendency, given their proposals that humanities computing reconfigure itself as a synthesis of theory, statistics and aesthetics. In seeking to locate a common ground between the works of these critics, we might identify their joint rejection of ground truth. The bureaucratic overtones of any reductive striving towards 'accuracy' is eschewed, in favour of a focus on a generative or procedural critical project which may emerge from the transformation of texts, according to the notion that deformance, re-mediation, translation and misprision form crucial parts of the critical enterprise (Ramsay 2011, x; Drucker 2014; Rockwell 2003). The difficulty in considering the work of these critics within the context of CLS is that, even though they may provide novel and engaging philosophical insights, they do not engage to a significant extent with the actuality of statistical approaches and it is as a result impossible, on the basis of their writings, to arrive at practical steps towards the implementation of a provisional or exploratory CLS. There is also a tendency at work in such criticisms to overlook the changing nature of CLS over time.

While, as we have seen, some early CLS scholars may well have had a propensity to overstate the significance of their results, by the 2000s we can see that the promises to reconstruct literary criticism on a foundation more hospitable to scientific rigour in order to exorcise the spectre of post-structuralism have given way to comparisons with endeavours such as sociology, economics or state planning, all of which have long histories of applying statistics in critical and reflective ways. Burrows, for example, asserts that, as it would be an impossibility for a demographer to identify 'pure' instances of the social phenomenon they aim to quantify, whether class, race or gender, the use of spectra or 'fuzzy logic' becomes essential. Burrows' more pragmatic twining of empirical and intuitive analysis undergirded by a growing body of scholarship go a long way towards rejecting the caricature of CLS which sceptics identify as operating within the discipline (Burrows 2018, 725). How much the field of CLS

e-ISSN 2724-3923

can be said to have advanced in this regard can be seen in the work of Taylor Arnold and Lauren Tilton (2019, 4-14) which, in its simultaneous consideration of both Barthes and the functionality of machine learning, takes strides in combining the actual mechanics of computing and theoretical criticism as well as the changing nature of other anti-CLS articles which now circulate. While the criticisms Nan Z. Da presents in *Critical Inquiry* are partly inhibited by their aim to de-legitimise the quantification of literature in general, Da's article still represents a paradigm shift, in that it argues that computational or statistical methods are widely misunderstood or not implemented properly within. Implicit within Da's analysis (2019) then, is the notion that the field could be improved on these bases. Katharine Bode, in a response to Da's article, also notes this distinction, as well as the greater degree of care which needs to be taken in critiquing CLS on the basis of its scientism, given the pivot from objective to greater amounts of subjectivity and uncertainty which are made possible via the modelling of machine learning outputs (Bode 2019). These methods, when they are first operationalised in the nineties, function more or less as black boxes. CLS scholars do not expend significant amount of time examining the actual functionality of the algorithms themselves; the emphasis is more often placed on the algorithm's capacity to identify an optimal number of classes having been given them at the outset. In many ways this is to be expected at an early stage in CLS' history, given that, when it is applied to research questions such as Shakespeare's authorship, there is a relatively constrained set of probable candidates. In this sense, machine learning methods are used in more or less the same way as PCA is, as a means of dimension reduction, rather than grappling with the capacity of the method in and of itself. We might compare this with Ted Underwood's 2014 project, "Understanding Genre in a Collection of a Million Volumes", which aimed to classify page-level data into one of three categories, either prose, poetry or drama. In the course of this project, Underwood demonstrates how the two paradigms of knowledge production held to be in opposition for almost the entirety of CLS' history, the statistical and literary aspects, may be synthesised. Underwood notes that, as literary critics do not understand genre empirically, but rather socially, it makes no sense to enforce a rigid either/or classification, but rather an approach based on a spectrum. Approaches arising from the field of machine learning, with its capacity to score goodness of fit as a figure between zero and one, zero representing total uncertainty and one representing absolute certainty, is uniquely suited. A further safeguard against empirical reductionism is erected by cross-validating the obtained results with human judgement, specifically a group of five readers who were recruited in order to classify literary data page by page through the use of a GUI purpose-built for the project. Through the labour ex-

erted by these readers, who labelled all pages in 414 books, training data for the project was obtained, which was instrumental in the algorithm attaining an agreement rate of 94.5% in identifying prose as opposed to poetry, fiction as opposed to nonfiction and body text as opposed to paratext. The statistical model which was constructed on the basis of this training data was found to be less accurate than human judgment by a margin of just 0.9%. In this way, Underwood's utilisation of machine learning points to the capacity of CLS to utilise ambiguity and shades of difference within an empirical approach (Underwood 2014, 8-12).

5 Conclusion

In providing a history of the development of CLS, this paper has demonstrated that, from an early stage in CLS' history, the frequencies of an undifferentiated selection of high-frequency word types were highly effective in clustering texts together on the basis of their authorship. However, CLS scholars aimed to challenge the predominance of post-structuralist theories of authorship and, as a result, CLS was from its inception subject to robust criticism from a cohort of literary critics who were more invested in theoretical readings and who charged CLS critics as operating within a politically reactionary and reductive form of knowledge production. In response, CLS cleaved from an early stage in its history to organic theories of authorship and a focus on unconsciously deployed formal features within the work. The original discovery regarding the efficacy of highly frequent word types is consequently elided for a significant period of time in favour of focuses on the individual contributions of particular words or word types insofar as these can be re-integrated within a traditional or qualitative literary-critical reading. This remains the case even after Burrows develops the Delta method on which subsequent CLS scholars develop improvements; these analyses are noteworthy for their focus on particular words and apparent reluctance to move into higher and higher frequency strata. Yet again this did not change until scholars such as Maciej Eder and Jan Rybicki enact a sequence of benchmark analyses which make the superiority of quantifying thousands of MFWs irrefutable, as well as the development of highly effective unsupervised machine learning techniques optimised for large datasets with thousands of parameters, within which manual intervention would become impractical or inefficient. The development of machine learning represents a significant riposte to the most well-worn arguments against CLS and will no doubt have a significant role to play in the development of the field in the future.

Bibliography

Algee-Hewitt, M.; Heuser, R.; Moretti, F. (2015). "On Paragraphs. Scale, Themes, and Narrative Form". *Stanford Lit Lab*, Stanford. https://litlab.stanford.edu/LiteraryLabPamphlet10.pdf.

Allison, S.; Gemma, M.; Heuser, R.; Moretti, F.; Tevel, A.; Yamboliev, I. (2013). "Style at the Scale of the Sentence". *Stanford Lit Lab*, Stanford. https://litlab.stanford.edu/LiteraryLabPamphlet5.pdf.

Allison, S.; Heuser, R.; Jockers, M.; Moretti, F.; Witmore, M. (2011). "Quantitative Formalism. An Experiment". *Stanford Lit Lab*, Stanford. https://litlab.stanford.edu/LiteraryLabPamphlet1.pdf.

Allison, S.; Heuser, R.; Jockers, M.; Moretti, F.; Witmore, M. (2012). "Quantitative Formalism". *n+1*, (13). https://nplusonemag.com/issue-13/essays/quantitative-formalism-an-experiment/.

Anderson, C.W.; McMaster, G.E. (1989). "Quantification of Rewriting by the Brothers Grimm". *Computers and the Humanities*, 23(4/5), 341-6. https://doi.org/10.1007/bf02176639.

Anderson, P. (1983). *In the Tracks of Historical Materialism. The Wellek Library Lectures*. London: Verso.

Antonia, A.; Craig, H.; Elliott, J. (2014). "Language Chunking, Data Sparseness, and the Value of a Long Marker List". *Literary and Linguistic Computing*, 29(2), 147-63. https://doi.org/10.1093/llc/fqt028.

Aoyama, H.; Constable, J. (1999). "Word Length Frequency and Distribution in English: Part I. Prose". *Literary and Linguistic Computing*, 14(3), 339-58. https://doi.org/10.1093/llc/14.3.339.

Apter, E. (2013). *Against World Literature. On the Politics of Untranslatability*. London: Verso.

Argamon, S. (2008). "Interpreting Burrows's Delta". *Literary and Linguistic Computing*, 23(2), 131-47. https://doi.org/10.1093/llc/fqn003.

Arnold, T.; Tilton, L. (2019). "Distant Viewing". *Digital Scholarship in the Humanities*, 34, Suppl. 1, i3-i16. https://doi.org/10.1093/digitalsh/fqz013.

Baayen, H.; Van Halteren, H.; Tweedie, F. (1996). "Outside the Cave of Shadows". *Literary and Linguistic Computing*, 11(3), 121-31. https://doi.org/10.1093/llc/11.3.121.

Baker, J.C. (1988). "Pace". *Literary and Linguistic Computing*, 3(1), 36-9. https://doi.org/10.1093/llc/3.1.36.

Beatie, B.A. (1979). "Measurement and the Study of Literature". *Computers and the Humanities*, 13(3), 185-94. https://doi.org/10.1007/bf02395096.

Binongo, J.N.G.; Smith, M.W.A. (1999). "The Application of Principal Component Analysis to Stylometry". *Literary and Linguistic Computing*, 14(4), 445-65. https://doi.org/10.1093/llc/14.4.445.

Bode, K. (2019). "Computational Literary Studies: A Critical Inquiry Online Forum". https://critinq.wordpress.com/2019/03/31/computational-literary-studies-a-critical-inquiry-online-forum/.

Brainerd, B. (1973). "On the Distinction Between a Novel and a Romance". *Computers and the Humanities*, 7(5), 259-70. https://doi.org/10.1007/bf02395426.

Brainerd, B. (1975). "Statistical Analysis of Lexical Data Using Chi-squared and Related Distributions". *Computers and the Humanities*, 9(4), 161-78. https://doi.org/10.1007/bf02402331.

Brainerd, B. (1979). "Pronouns and Genre in Shakespeare's Drama". *Computers and the Humanities*, 13(1), 3-16. https://doi.org/10.1007/bf02744988.
Burrows, J. (2002). "'Delta'". *Literary and Linguistic Computing*, 17(3), 267-87.
Burrows, J. (2018). "Rho-grams and Rho-sets". *Digital Scholarship in the Humanities*, 33(4), 724-47. https://doi.org/10.1093/llc/17.3.267.
Burrows, J.F. (1986). "Modal Verbs and Moral Principles". *Literary and Linguistic Computing*, 1(1), 9-23. https://doi.org/10.1093/llc/1.1.9.
Burrows, J.F. (1987). "Word-Patterns and Story-Shapes". *Literary and Linguistic Computing*, 2(2), 61-70. https://doi.org/10.1093/llc/2.2.61.
Burrows, J.F. (1989). "'An Ocean Where Each Kind...'". *Computers and the Humanities*, 23(4/5), 309-21. https://doi.org/10.1007/BF02176636.
Burrows, J.F. (1992). "Not Unless You Ask Nicely". *Literary and Linguistic Computing*, 7(2), 91-109. https://doi.org/10.1093/llc/7.2.91.
Burrows, J.F. (2004). "Textual Analysis". Schreibman, S.; Siemens, R.; Unsworth, J. (eds), *A Companion to Digital Humanities*. Oxford: Blackwell. https://doi.org/10.1002/9780470999875.ch23.
Cleary, J. (2006). "The World Literary System". *Field Day Review*, 2, 196-219.
Cluett, R. (1971). "Style, Precept, Personality". *Computers and the Humanities*, 5(5), 257-77. https://doi.org/10.1007/bf02402207.
Craig, D.H. (1991). "Plural Pronouns in Roman Plays by Shakespeare and Jonson". *Literary and Linguistic Computing*, 6(3), 180-6. https://doi.org/10.1093/llc/6.3.180.
Craig, H. (1999). "Contrast and Change in the Idiolects of Ben Jonson". *Computers and the Humanities*, 33(3), 221-40. https://doi.org/10.1023/A:1002032032618.
Da, N.Z. (2019). "The Computational Case against Computational Literary Studies". *Critical Inquiry*, 45(3), 601-39. https://doi.org/10.1086/702594.
Damerau, F.J. (1975). "The Use of Function Word Frequencies as Indicators of Style". *Computers and the Humanities*, 9(6), 271-80. https://doi.org/10.1007/bf02396290.
Dixon, P.; Mannion, D. (1993). "Goldsmith's Periodical Essays: A Statistical Analysis of Eleven Doubtful Cases". *Literary and Linguistic Computing*, 8(1), 1-19. https://doi.org/10.1093/llc/8.1.1.
Drucker, J. (2014). *Graphesis. Visual Forms of Knowledge Production*. Cambridge: Harvard University Press.
Eder, M. (2017). "Visualization in Stylometry: Cluster Analysis Using Networks". *Digital Scholarship in the Humanities*, 32(1), 50-64. https://doi.org/10.1093/llc/fqv061.
Escobar, M.V. (2016). "The Essay/ontology Workflow, Challenges in Combining Formal and Interpretive Methods". *Literary and Linguistic Computing*, 31(1), 84-94. https://doi.org/10.1093/llc/fqu071.
Evert, S.; Proisl, T.; Jannidis, F.; Reger, I.; Pielström, S.; Schöch, C.; Vitt, T. (2017). "Understanding and Explaining Delta Measures for Authorship Attribution". *Digital Scholarship in the Humanities*, 32(2), ii4-ii16. https://doi.org/10.1093/llc/fqx023.
Foley, J.M. (1978). "A Computer Analysis of Metrical Patterns in Beowulf". *Computers and the Humanities*, 12(1/2), 71-80. https://doi.org/10.1007/bf02392918.
Forsyth, R.S. (1999). "Stylochronometry with Substrings, or: A Poet Young and Old". *Literary and Linguistic Computing*, 14(4), 467-77. https://doi.org/10.1093/llc/14.4.467.

Forsyth, R.S.; Holmes, D.I.; Tse, E.K. (1999). "Cicero, Sigonio, and Burrows: Investigating the Authenticity of the Consolatio". *Literary and Linguistic Computing*, 14(3), 375-400. https://doi.org/10.1093/llc/14.3.375.

Forsyth, R.S.; Lam, P.W.Y. (2014). "Found in Translation: To What Extent is Authorial Discriminability Preserved by Translators?". *Literary and Linguistic Computing*, 29(2), 199-217. https://doi.org/10.1093/llc/fqt018.

Fortier, P.A. (1991). "Theory, Methods and Applications: Some Examples in French Literature". *Literary and Linguistic Computing*, 6(3), 192-6. https://doi.org/10.1093/llc/6.3.192.

Gladwin, A.A.G.; Lavin, M.J.; 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.

Gooding, P. (2013). "Mass Digitization and the Garbage Dump: The Conflicting Needs of Quantitative and Qualitative Methods". *Literary and Linguistic Computing*, 28(3), 425-31. https://doi.org/10.1093/llc/fqs054.

Greenwood, H.H. (1992). "St Paul Revisited–a Computational Result". *Literary and Linguistic Computing*, 7(1), 43-47. https://doi.org/10.1093/llc/7.1.43.

Greenwood, H.H. (1993). "St Paul Revisited–Word Clusters in Multidimensional Space". *Literary and Linguistic Computing*, 8(4), 211-19. https://doi.org/10.1093/llc/8.4.211.

Grieve, J. (2007). "Quantitative Authorship Attribution: An Evaluation of Techniques". *Literary and Linguistic Computing*, 22(3), 251-70. https://doi.org/10.1093/llc/fqm020.

Gurney, P.J.; Gurney, L.W. (1998). "Authorship Attribution of the Scriptores Historiae Augustae". *Literary and Linguistic Computing*, 13(3), 119-31. https://doi.org/10.1093/llc/13.3.119.

Hilton, M.L.; Holmes, D.I. (1993). "An Assessment of Cumulative Sum Charts for Authorship Attribution". *Literary and Linguistic Computing*, 8(2), 73-80. https://doi.org/10.1093/llc/8.2.73.

Holmes, D.I. (1998). "The Evolution of Stylometry in Humanities Scholarship". *Literary and Linguistic Computing*, 13(3), 111-17. https://doi.org/10.1093/llc/13.3.111.

Holmes, D.I.; Crofts, D.W. (2010). "The Diary of a Public Man: A Case Study in Traditional and Non-traditional Authorship Attribution". *Literary and Linguistic Computing*, 25(2), 179-97. https://doi.org/10.1093/llc/fqq005.

Hoover, D.L. (2001). "Statistical Stylistics and Authorship Attribution: An Empirical Investigation". *Literary and Linguistic Computing*, 16(4), 421-44. https://doi.org/10.1093/llc/16.4.421.

Hoover, D.L. (2002). "Frequent Word Sequences and Statistical Stylistics". *Literary and Linguistic Computing*, 17(2), 157-80. https://doi.org/10.1093/llc/17.2.157.

Hoover, D.L. (2003). "Frequent Collocations and Authorial Style". *Literary and Linguistic Computing*, 18(3), 261-86. https://doi.org/10.1093/llc/18.3.261.

Hoover, D.L. (2004a). "Delta Prime?". *Literary and Linguistic Computing*, 19(4), 477-95. https://doi.org/10.1093/llc/19.4.477.

Hoover, D.L. (2004b). "Multivariate Analysis and the Study of Style Variation". *Literary and Linguistic Computing*, 18(3), 341-60. https://doi.org/10.1093/llc/18.4.341.

Hoover, D.L. (2004c). "Testing Burrows's Delta". *Literary and Linguistic Computing*, 19(4), 453-75. https://doi.org/10.1093/llc/19.4.453.
Hou, R.; Jiang, M. (2016). "Analysis on Chinese quantitative stylistic features based on text mining". *Digital Scholarship in the Humanities*, 31(2), 357-67. https://doi.org/10.1093/llc/fqu067.
Irizarry, E. (1991). "One Writer, Two Authors: Resolving the Polemic of Latin America's First Published Novel". *Literary and Linguistic Computing*, 6(3), 175-9. https://doi.org/10.1093/llc/6.3.175.
Irizarry, E. (1993). "The Two Authors of Columbus' Diary". *Computers and the Humanities*, 27(2), 85-92. https://doi.org/10.1007/bf01830301.
Jannidis, F.; Pielström, S; Schöch, C.; Vitt, T. (2015). "Improving Burrows' Delta – An Empirical Evaluation of Text Distance Measures". *Digital Humanities Conference 2015*. https://www.researchgate.net/publication/280086768_Improving_Burrows%27_Delta_-_An_empirical_evaluation_of_text_distance_measures.
Joliffe, I.T. (2004). *Principal Components Analysis*. New York: Springer.
Juola, P. (2015). "The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions". *Digital Scholarship in the Humanities*, 30, Suppl. 1, i100-i13. https://doi.org/10.1093/llc/fqv040.
Lee, C. (2017). "Do Language Combinations Affect Translators' Stylistic Visibility in Translated Texts?" *Digital Scholarship in the Humanities*, 33(3), 592-603. https://doi.org/10.1093/llc/fqx056.
Mannion, D.; Dixon, P. (2004). "Sentence-length and Authorship Attribution: the Case of Oliver Goldsmith". *Literary and Linguistic Computing*, 19(4), 497-508. https://doi.org/10.1093/llc/19.4.497.
McColly, W.B. (1987). "Style and Structure in the Middle English Poem 'Cleanness'". *Computers and the Humanities*, 21(3), 169-176. https://doi.org/10.1007/bf02252793.
Merriam, T. (1989). "An Experiment with the Federalist Papers". *Computers and the Humanities*, 23(3), 251-4. https://doi.org/10.1007/bf00056147.
Mesthene, E.G. (1969). "Technology and Humanistic Values". *Computers and the Humanities*, 4(1), 1-10. https://doi.org/10.1007/bf02393443.
Milic, L. (1991). "Progress in Stylistics: Theory, Statistics, Computers". *Computers and the Humanities*, 25(6), 393-400. https://doi.org/10.1007/bf00141189.
Milic, L.T. (1966). "The Next Step". *Computers and the Humanities*, 1(1), 3-6. https://doi.org/10.1007/bf00188010.
Milic, L.T. (1967). *A Quantitative Approach to the Style of Jonathan Swift*. Berlin: Mouton & Co.
Moretti, F. (1998). *Atlas of the European Novel 1800-1900*. London: Verso.
Moretti, F. (2007). *Graphs, Maps, Trees. Abstract Models for a Literary History*. London: Verso.
Moretti, F. (2012). *Distant Reading*. London: Verso.
Moretti, F. (2013). *The Bourgeois. Between History and Literature*. London: Verso.
Moretti, F. (2020). "The Roads to Rome". *New Left Review*, 2(124), 125-36.
Pearl, L.; Lu, K.; Haghighi, A. (2017). "The Character in the Letter: Epistolary attribution in Samuel Richardson's *Clarissa*". *Digital Scholarship in the Humanities*, 32(2), 355-76. https://doi.org/10.1093/llc/fqw007.
Philippides, D.M.L. (1988). "Literary Detection in the Erotokritos and The Sacrifice of Abraham". *Literary and Linguistic Computing*, 3(1), 1-11. https://doi.org/10.1093/llc/3.1.1.

Potter, R.G. (1988). "Literary Criticism and Literary Computing: The Difficulties of a Synthesis". *Computers and the Humanities*, 22(2), 91-7. https://doi.org/10.1007/bf00057648.

Prendergast, C. (2004). "The World Republic of Letters". Prendergast, C. (ed.), *Debating World Literature*. London: Verso, 1-25.

Ramsay, S. (2011). *Reading Machines*. *Towards an Algorithmic Criticism*. Champaign: University of Illinois Press.

Rockwell, G. (2003). "What is Text Analysis, Really?". *Literary and Linguistic Computing*, 18(2), 209-19. https://doi.org/10.1093/llc/18.2.209.

Rybicki, J.; Eder, M. (2011). "Deeper Delta Across Genres and Languages: Do We Really Need the Most Frequent Words?" *Literary and Linguistic Computing*, 26(3), 315-21. https://doi.org/10.1093/llc/fqr031.

Rybicki, J.; 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.

Saccenti, E.; Tenori, L. (2015). "Multivariate Modeling of the Collaboration Between Luigi Illica and Giuseppe Giacosa for the Librettos of Three Operas by Giacomo Puccini". *Digital Scholarship in the Humanities*, 30(3), 405-22. https://doi.org/10.1093/llc/fqu006.

Sainte-Marie, P.; Robillard, P.; Bratley, P. (1973). "An Application of Principal Component Analysis to the Works of Molière". *Computers and the Humanities*, 7(3), 131-7. https://doi.org/10.1007/bf02403851.

Savoy, J. (2013). "Comparative Evaluation of Term Selection Functions for Authorship Attribution". *Digital Scholarship in the Humanities*, 30(2), 246-61. https://doi.org/10.1093/llc/fqt047.

Sayoud, H. (2012). "Author Discrimination Between the Holy Quran and Prophet's Statements". *Literary and Linguistic Computing*, 27(4), 427-44. https://doi.org/10.1093/llc/fqs014.

Schaalje, G.B.; Fields, P.J.; Roper, M.; Snow, G.L. (2011). "Extended Nearest Shrunken Centroid Classification: A New Method for Open-Set Authorship Attribution of Texts of Varying Sizes". *Literary and Linguistic Computing*, 26(1), 71-88. https://doi.org/10.1093/llc/fqq029.

Sinclair, S. (2003). "Computer-Assisted Reading: Reconceiving Text Analysis". *Literary and Linguistic Computing*, 18(2), 175-84. https://doi.org/10.1093/llc/18.2.175.

Smith, M.W.A. (1985). "An Investigation of Morton's Method to Distinguish Elizabethan Playwrights". *Computers and the Humanities*, 19(1), 3-21. https://doi.org/10.1007/bf02259614.

Smith, M.W.A. (1986). "A Critical Review of Word-links as a Method for Investigating Shakespearean Chronology and Authorship". *Literary and Linguistic Computing*, 1(4), 202-6. https://doi.org/10.1093/llc/1.4.202.

Smith, M.W.A. (1987). "Hapax Legomena in Prescribed Positions: An Investigation of Recent Proposals to Resolve Problems of Authorship". *Literary and Linguistic Computing*, 2(3), 145-52. https://doi.org/10.1093/llc/2.3.145.

Smith, M.W.A. (1988). "The Authorship of Acts I and II of *Pericles*: A New Approach Using First Words of Speeches". *Computers and the Humanities*, 22(1), 23-41. https://doi.org/10.1007/bf00056347.

Smith, P.W.H.; Aldridge, W. (2011). "Improving Authorship Attribution: Optimizing Burrows' Delta Method*". *Journal of Quantitative Linguistics*, 18(1), 63-88. https://doi.org/10.1080/09296174.2011.533591.

Stanikūnas, D.; Mandravickaitė, J.; Krilavičius, T. (2015). "Comparison of Distance and Similarity Measures for Stylometric Analysis of Lithuanian Texts". *International Conference for Young Researchers in Informatics, Mathematics and Engineering*. http://ceur-ws.org/Vol-1852/p01.pdf.

Sula, C.A.; Hill, H.V. (2019). "The Early History of Digital Humanities: an Analysis of Computers and the Humanities (1966-2004) and Literary and Linguistic Computing (1986-2004)". *Digital Scholarship in the Humanities*, 34, Suppl. 1, i190-i206. https://doi.org/10.1093/llc/fqz072.

Thoiron, P. (1986). "Diversity Index and Entropy as Measures of Lexical Richness". *Computers and the Humanities*, 20(3), 197-202. https://doi.org/10.1007/bf02404461.

Tse, E.K.; Tweedie, F.J.; Frischer, B.D. (1998). "Unravelling the Purple Thread: Function Word Variability and the *Scriptores Historiae Augustae*". *Literary and Linguistic Computing*, 13(3), 141-9. https://doi.org/10.1093/llc/13.3.141.

Tuccinardi, E. (2016). "An Application of a Profile-Based Method for Authorship Verification: Investigating the Authenticity of Pliny the Younger's Letter to Trajan Concerning the Christians". *Digital Scholarship in the Humanities*, 32(2), 435-47. https://doi.org/10.1093/llc/fqw001.

Tweedie, F.J.; Baayen, R.H. (1998). "How Variable May a Constant be? Measures of Lexical Richness in Perspective". *Computers and the Humanities*, 32(5), 323-52. https://doi.org/10.1023/A:1001749303137.

Underwood, T. (2014). "Understanding Genre in a Collection of a Million Volumes. Interim Report". https://doi.org/10.6084/m9.figshare.1281251.

van Peer, W. (1989). "Quantitative Studies of Literature. A Critique and an Outlook". *Computers and the Humanities*, 23(4/5), 301-7. https://doi.org/10.1007/bf02176635.

Wittig, S. (1977). "The Computer and the Concept of Text". *Computers and the Humanities*, 11(4), 211-15. https://doi.org/10.1007/bf02396857.