Anglistentag 2017 Regensburg

Proceedings

edited by

Anne-Julia Zwierlein, Jochen Petzold,
Katharina Boehm and Martin Decker

MB 040. 015-39

Universität Tübingen
Brechtbau - Bibliothek

WVT Wissenschaftlicher Verlag Trier

Anglistentag 2017 Regensburg
Proceedings
ed. by Anne-Julia Zwierlein, Jochen Petzold,
Katharina Boehm and Martin Decker
Trier: WVT Wissenschaftlicher Verlag Trier, 2018
(Proceedings of the Conference of the German
Association for the Study of English; Vol. 39)
ISBN 978-3-86821-767-4

Umschlaggestaltung: Brigitta Disseldorf

© WVT Wissenschaftlicher Verlag Trier, 2018
ISBN 978-3-86821-767-4

Alle Rechte vorbehalten
Nachdruck oder Vervielfältigung nur mit
ausdrücklicher Genehmigung des Verlags

Gedruckt auf alterungsbeständigem
und säurefreiem Papier
Printed in Germany

WVT Wissenschaftlicher Verlag Trier
Bergstraße 27, 54295 Trier
Postfach 4005, 54230 Trier
Tel.: (0651) 41503 / 9943344, Fax: 41504
Internet: http://www.wvttrier.de
E-Mail: wvt@wvttrier.de

Proceedings of the Conference of the German Association for the Study of English

Volume XXXIX

Contents

Anne-Julia Zwierlein, Jochen Petzold, Katharina Boehm
and Martin Decker (Regensburg)
Preface ix

Section I: Digital Humanities: The Role of the Digital in English Philology

Sabine Bartsch (Darmstadt) and Ilka Mindt (Paderborn)
Digital Humanities: The Role of the Digital in English Philology 3

Helen Baker and Tony McEnery (Lancaster)
Using Corpora to Explore Shadows from the Past:
Early Modern Male Prostitution, an Evasive Marginal Identity 9

Nicola Glaubitz (Frankfurt)
Zooming in, Zooming out: The Debate on Close and Distant Reading
and the Case for Critical Digital Humanities 21

Matthias Bauer (Tübingen) and Angelika Zirker (Tübingen and Berlin)
Shakespeare and Stylometrics:
Character Style Paradox and Unique Parallels 31

Annika Elstermann (Heidelberg)
The Power of Digital Publishing:
Rethinking Knowledge Distribution in English Philology 39

Manfred Markus (Innsbruck)
Digital Humanities: A New Departure in English Dialectology
(Based on *EDD Online*) 51

Christian Mair (Freiburg)
When All Englishes Are Everywhere: Media Globalisation
and Its Implications for Digital Corpora and World English Studies 65

Ulrike Schneider and Matthias Eitelmann (Mainz)
Turning Digital in English Linguistics: Challenges and Opportunities 73

Josef Schmied and Matthias Hofmann (Chemnitz)
New Digital Methodologies for Old Grammar Problems:
Corpus Analyses and Eye-Tracking to Discover
Non-Native English Article Usage Preferences 85

Peter Uhrig (Erlangen-Nuremberg)
NewsScape and the Distributed Little Red Hen Lab:
A Digital Infrastructure for the Large-Scale Analysis of TV Broadcasts 99

Section II: The Value of Economic Criticism Reconsidered: Approaching Literature and Culture through the Lens of Economics

Ellen Grünkemeier (Hanover), Nora Pleßke (Magdeburg), and Joanna Rostek (Giessen)
The Value of Economic Criticism Reconsidered:
Approaching Literature and Culture through the Lens of Economics 117

Natalie Roxburgh (Siegen)
Rethinking 'Moral Economy' through Recent Economic Criticism 127

Benjamin Kohlmann (Freiburg)
Cognitive Sympathy and the Laissez-Faire Economy:
Adam Smith and Harriet Martineau 135

Barbara Straumann (Zurich)
Admired and Reviled: Figures of Finance Capitalism
in Martin Scorsese's *The Wolf of Wall Street* (2013)
and Anthony Trollope's *The Way We Live Now* (1875) 145

Dorothee Birke (Aarhus)
Economies of Space: The Value of Home in Contemporary British Theatre 159

Melissa Kennedy (Passau and Vienna)
Imaginary Economics: Alternative Ways of Thinking Literary Studies
and the Economy 169

Section III: Women in Men's Shoes: 'Sheroes' in Literatures and Cultures of the British Isles, the Commonwealth and Beyond

Caroline Lusin and Christine Schwanecke (Mannheim)
Women in Men's Shoes: 'Sheroes' in Literatures and Cultures
of the British Isles, the Commonwealth and Beyond 183

Wolfgang Funk (Mainz)
The Philosopher's Wife: Amy Levy's Rewriting of Xanthippe 189

Wieland Schwanebeck (Dresden)
"You're not the usual kind of Private Eye": The Deconstruction
of the Whodunit in P.D. James's Cordelia Gray Mysteries 201

Johannes Fehrle (Mannheim)
"If I get an outfit can I be cowboy, too":
Female Cowboys in the Revisionist Canadian Western 211

Stefanie Schäfer (Erlangen-Nuremberg)
Marooning James Bond: Cleopatra Jones, Blaxploitation
and the Imperial Mystique 221

Philip Jacobi (Passau)
Daddy Issues: Wet T-Shirt Feminism and the Contemporary
Video Game Heroine 233

Section IV: The Reformation in the English-Speaking World

Ralf Haekel (Göttingen and Giessen), Lukas Lammers (Berlin) and Kirsten Sandrock (Göttingen)
The Reformation in the English-Speaking World 245

Anne Enderwitz (Berlin)
Reformation and Beyond: Shakespeare's *Henriad* 251

Katrin Röder (Potsdam)
The Reformation and the Ottomans: On Religious Anxiety, Imperial Envy
and Moral Diversity in Late Elizabethan and Early Jacobean Drama 259

Felix C.H. Sprang (Siegen)
The Reformation of Modern Poetry: Reform and Form 271

Katrin Berndt (Halle-Wittenberg)
Protestantism and Regional Identity in Alice Munro's Short Fiction 281

Section V: Adaptations, Creations and Transformations: Teaching Literature Today

Gabriele Blell (Hanover) and Maria Eisenmann (Würzburg)
Adaptations, Creations and Transformations: Teaching Literature Today 295

Christiane Lütge (Munich)
Digital, Transcultural and Global?
Reconsidering the Role of Literature in the EFL Classroom 299

Susanne Heinz (Kiel)
Researching Multimodal Reader Response(s) in the EFL Classroom 311

Claudia Deetjen (Würzburg)
Teaching William Shakespeare's *The Tempest* through Graphic Novels 325

Christian Ludwig (Karlsruhe)
Joe Sacco's Comics Journalism in the English Classroom: *Kushinagar* 337

Kerstin Theinert (Weingarten)
Training Reading Fluency:
Multilingual Readers' Theatre and Forms of Literary Learning 351

Notes on Contributors 363

So, Richard Jean (2017): "'All Models Are Wrong,'" *PMLA* 132.3, 668-673

Venus, Jochen (2009): "Mikro/Makro: Zur Wissens- und Technikgeschichte einer Eigentümlichen Unterscheidung," in: Köster, Ingo; Schubert, Kai (eds.): *Medien in Raum und Zeit: Maßverhältnisse des Medialen*. Bielefeld: transcript, 47-61

Wilkens, Matthew (2012): "Canons, Close Reading and the Evolution of Method," in: Gold, Matthew K. (ed.): *Debates in the Digital Humanities*. Minneapolis: U of Minnesota P, 249-258

MATTHIAS BAUER (TÜBINGEN)
AND ANGELIKA ZIRKER (TÜBINGEN AND BERLIN)

Shakespeare and Stylometrics: Character Style Paradox and Unique Parallels

Stylometric analysis goes back to the age before the computer was introduced in the humanities. It has gained new acumen with the ability to work with big data, i.e. huge quantities of text. In what follows, we will look at stylometric analysis in the context of Shakespeare. After a brief overview as to how stylometrics has evolved over the years, we will attempt a critical evaluation of these approaches and focus on two: Principal Component Analysis, conducted extensively by Hugh Craig and his co-authors, and n-grams, a method used prominently by Gary Taylor. Shakespeare and his contemporaries are, in our view, apt candidates to address these matters as quite a few of the works published under his name are products of co-authorship – according to the *New Oxford Shakespeare: Authorship Companion* (Taylor and Egan 2017) up to 17.[1] This makes him one of the focal points of authorship attribution studies that use stylometric methods.

1. Shakespeare and Stylometrics: State of the Art

When it comes to the stylometric analysis of Shakespeare (and others), three approaches are dominant: frequency, i.e. how often are particular words used, combination, i.e. their syntagmatic distribution, and proximity. These approaches have not fundamentally changed with the introduction of the computer.[2] From early on, *The Two Noble Kinsmen* (*TNK*) has been a candidate for stylometric analysis. According to the *New Oxford Shakespeare: Authorship Companion* (2017, 590), Harold Littledale (1876), for instance, looked for verbal parallels in *The Two Noble Kinsmen* and other plays by Shakespeare and Fletcher. Other approaches were suggested by H. Dugdale Sykes (1916), who focused on metrical evidence,[3] and Alfred Hart, who selected 1,000 "rarer words" (1934, 274) from the play and compared them to the indisputably attributed plays of Shakespeare and Fletcher. Sykes argues that Massinger also had a hand in the play and comes to three conclusions (1916, 137): (1) traces of Massinger's language can be found in the non-Fletcherian parts of *TNK*, (2) "reminiscences of Shakespeare are characteristic of Massinger who has 'continual touches showing that some passage of Shakespeare was running in his head,'" (*ibid.*) and (3) the language of the female characters "could not conceivably have been put by Shakespeare into the mouths of virtuous women, but is typical of Massinger's heroines" (*ibid.*). When we look, however, at his juxtaposition of passages from *TNK* with excerpts from a variety of Massinger's plays, it becomes difficult to see parallels.[4]

---

[1] See, e.g., Alberge (2016).

[2] For a more comprehensive account, see the *New Oxford Shakespeare: Authorship Companion* (2017); and Love (2002).

[3] See also Chambers (1930).

Hart, unsurprisingly, one might think, finds the "rare" words of five Shakespeare plays also in *TNK* but leaves out factors like stylistic imitation, character idiolect etc. and is limited by his own memory and subjective judgement (see Vickers 2017). Thus it is easy to see in those older approaches what David Holmes maintained in 1994: "[a]ll authorship studies begin with a choice of criteria believed to characterize authors" (qtd. in Vickers 2016, 10), i.e. certain preconceived ideas of authors and their stylistic habits.

What is perhaps more surprising is that this tendency continues to this very day, even though the methods have become much more sophisticated. At least there is a belief that frequency, combination and proximity give us the key to author identification. The first is used, for example, by Burrows and Craig (2017, 194), who mix methods of frequency count to achieve better results; i.e. Delta, "which calculates a [...] degree of difference [...] between a disputed text and a series of authors by combining individual differences in word counts", Iota, which uses words "that appear in the first author's works and never in the second's" (*ibid.*), and Zeta, based on frequencies of mid-ranked words which are preferred by an author. Combination is used, for example, by Egan who strives to "quantify writers' preferences for putting particular words in particular orders" (2016, 232); and proximity is the choice of Jackson (e.g. 2014) and Vickers (e.g. 2016), who employs anti-plagiarism software to discover strings of words uniquely mentioned in a disputed text and in one other play (Vickers 2017, 102).[5]

Warning voices such as Rudman point out that it is important to keep in mind how style changes over time (2016, 319); this makes it even more difficult to identify stable features of style based on either of those three methodological principles. Another caveat is editorial: in how far do printed texts, such as Shakespeare's Folio, actually reflect what the playwrights wrote? This is especially relevant when variant word forms such as *ye* and *you* are used for purposes of author identification.

2. Critical Evaluation

In the following, we would like to select two examples of current work in stylometrics in order to reflect in slightly greater detail on the uses of stylistic computing and to contribute to the discussion of its pros and cons. Both examples combine a particular methodological approach, frequency of words and strings of words, with a particular goal. Both cases aim at identifying authors – the traditional and overarching aim of stylometrics – while at the same time we have chosen them because they go beyond that aim and take other objectives into view, in particular character style and the stylistic relationship between texts. Put simply, the question has been asked and debated whether character and/or intertextual features such as genre trump author and vice versa. In other words, what happens if dramatic authors imitate the styles and voices of people they have invented and the styles of texts belonging to the genre, theme etc. they have chosen?

---

[4] An example referred to by Sykes is the First Queen in *TNK*: "... what to do quickly / Is not done rashly; your first thought is more / Than others' laboured meditance" and Francisco (*The Duke of Milan* IV.i.): "They ... without a blush / Would swear that I, by nature, had more knowledge / Than others could acquire by any labour" (qtd. in Sykes 1916, 140). Though the idea expressed is similar, it is not characteristic enough to become evidence of identifiable influence.

[5] Vickers's (2017) defence against Egan (2016) is weak: Egan criticised that the unique strings are sometimes quite common in non-dramatic printed texts, while Vickers claims that language use in drama is not influenced by other forms of discourse.

2.1 Word Frequency (PCA) and Character Style

As regards characters, one may ask whether they speak in such an individual manner that it is impossible to detect their common author? Is the way in which Rosalind (in *As You Like It*) speaks, for example, very different from the king's speech in *Henry V*, written roughly at the same time? And, if this is the case, is the difference bigger than the difference of each of them from a character of the same age, rank, sex and genre in Jonson or Fletcher? These questions are fascinating because they provide a particular challenge that we might call the *character style paradox*: any author identification by means of style is based on a notion that is expressed by a classical topos, perhaps best known in its French version by Conte de Buffon, "[l]e style est l'homme même" (see Müller 1981, 40-51).[6] If human beings are individuals, their use of language must (quintessentially) show this.[7] But it is paradoxical that the more this notion is proved by literary authors in the creation of their characters, i.e. by making their characters speak in distinctive and unique ways, the more they disprove it as regards themselves, i.e. the less characteristic does their own style become.

We think that this paradox has not yet sufficiently been taken into account in computer-based stylometric analyses. In the field of Shakespeare studies, the most prominent examples of this approach are to be found in the work of Burrows and Craig; in particular, Craig (2008) and Burrows and Craig (2012) are concerned with "Shakespeare Characters and Common Words" and "Authors and Characters" respectively. In both papers, the authors establish their findings by working with lists of the 50 (2008) and 100 (2012) most frequent words in the corpus of plays to be analysed. These words are all function words (articles, pronouns, etc.),[8] a feature that in the eyes of many scholars makes word-count a more reliable method for the identification of (authorial) style because "they take less of their colour from their context [...] than lexical words" (Craig 2008, 283), they are "topic-independent" (Juola 2008, 265) and thus allow for greater individuality among writers (whereas tragedy or the character of a king might require typical lexical words, authorial variety may be seen better in the distribution of function words).[9] The assumption is that function words give us a clue to an author's stylistic DNA or at least to his or her formed habits. We will see some problems of this when looking at Craig's and Burrows' findings as regards characters.

---

[6] See also Love on Erasmus, *Patristic Scholarship*: "Style then is very much *l'homme-même*" (2002, 21).

[7] See Juola who speaks of the "authorial fingerprint" (2008, 239) that can be detected in people's writings. Juola also refers to the discovery of a "human stylome" (*ibid.*) by van Halteren et al. (2005).

[8] It is not absolutely clear to us whether Burrows and Craig (2012) are using the 100 most frequent function words or the 100 most frequent words (e.g. "let", which appears among the most frequent 50 words in Craig 2008, see 285, Fig. 24, does not appear on the list of the most frequent 100 words in Burrows and Craig 2012, see 293 n8).

[9] This is why it is surprising that Vickers maintains that a "computation of function words [...] may tell you something about the characters but cannot reliably indicate authorship" (Vickers 2009, 42). See also Vickers (2016).

In both papers, the method of Principle Component Analysis is used, which serves to find "a line of differentiation through the counts which accounts for the greatest amount of variation in it" (Craig 2008, 283). The idea is that the greatest variation (in our case in the use of function words) will tell us most about the stylistic differences. In the data space resulting from the frequencies of words used by the different characters, the first principle component "is the axis on which the data has the most variance" (Juola 2008, 259) and the second principle component "captures the next greatest variance, and so forth" (*ibid.*). One of the graphs of Craig's findings (Fig. 23, Craig 2008, 284), shows that the greatest difference is between Warwick (in the first part of *Henry VI*) and Pandarus (in *Troilus and Cressida*) according to the first principle component, and between Menenius Agrippa (in *Coriolanus*) and Romeo and Juliet according to the second. This only becomes meaningful if we look at the distribution of words provided by Craig (Fig. 24, *ibid.*, 285), where we find the most variance between "and" and "our" on the one side and "is", "not" and "I" on the other, and the second greatest variance between "his", "him" and "he" versus "thy", "thee" and "thou".

We may well ask what is to be inferred from the data. Craig puts it, for example, like that: "Pandarus (to the right in illustration 23) is a character formed of negation, querulously undercutting and anxiously re-directing" (*ibid.*). In other words, he finds the frequent use of "not" fits in with his interpretation of the character which he obviously derives from the coherent text since the words on the right-hand side of the graph do not tell you anything about undercutting and redirecting. As regards Romeo and Juliet, "their abundance of thou, thee and thy is a measure of the focus of their spoken parts on each other" (*ibid.*, 287) – again a result that confirms the obvious. More informative is perhaps the statement that some characters, such as Lear and Prospero, use "thou" quite frequently, against the trend of language development in the time of Shakespeare's career (*ibid.*). In any case, what we do get from this kind of statistical evidence is an impression of characters speaking differently from others as regards the use of function words, and differently from average uses of those words. An interpretation of those uses is difficult without a close reading of other features of their style and of what those characters are actually saying.

Furthermore, the graphs make us realise the character style paradox: the less the use of function words "means" in the sense of telling us something about a character's convictions, intentions etc., the more does it become imitative of the common difference between speakers, the inherited or acquired setup of a person's style. Craig's interpretations (e.g. of Pandarus) thus, strictly speaking, work against his own method of grounding differences between characters on the 'unconscious' use of function words, as if those characters were real-life persons. What they are, of course, is statements by their author, only apparently uttered by different real-life persons. It thus must remain unclear what the distribution of words indicated by the PCA plot actually tells us: the amount of success in making dramatic characters speak as differently as people do in real life, or the amount of success in characterising dramatic characters by means of their style. This is by no means the same.

This is a problem which, we think, also obfuscates Burrows' and Craig's attempt "to show the literary fundamentals of the relations between character styles and authorial styles in one particular field" (2012, 293), as we see in their PCA plot of Shakespeare's versus Fletcher's characters (Fig. 1, *ibid.*, 295). While it is clearly meant to show (and apparently does show, even though we are not given any figures that might tell us whether the differences are statistically valid) that the use of function words distinguishes all Shakespeare characters from all Fletcher characters, it invalidates the interpretative conclusions drawn from the use of function words, unless we want to claim that the word choice tells us something about Shakespeare's and Fletcher's character (and not just characters). This is of course absurd. Thus, when Burrows and Craig interpret the data and say that "a handful of Fletcher characters are more sententious and less concerned with the immediate and individual than any Shakespeare characters" (*ibid.*, 298-299), they are doing the same thing as with Pandarus and others, i.e. regard the function word use as being in agreement with what we know about the characters from their attitudes revealed by the content of their speeches. If this agreement exists (and it is not shown in detail how this can be worked out), what we learn is that Shakespeare created different kinds of characters than Fletcher, not necessarily that we can use this method to identify authors. This becomes even more evident when we look at the other comparisons worked out by Burrows and Craig, which show, for example, that Jonson's and Chapman's characters cannot be separated that easily, with 18.8 per cent of characters misclassified (Table 2, *ibid.*, 307) and many more very close to each other in their use of function words. If we look at their Fig. 7 (*ibid.*, 305), we see that especially the comedy characters are virtually indistinguishable by author. Against Burrows' and Craig's own conclusion, we may say that character and genre combined may easily trump author. This is perhaps no coincidence since comedy characters (especially of the "humour" kind adopted by Chapman and Jonson) tend to be more similar to each other. Accordingly, we think that the best use of the method lies in identifying and describing the variety of styles authors are able to create for their characters. Authors may thus be distinguished by their different scopes of style rather than by the function words themselves.

2.2 Taylor and n-grams/Collocations

Taylor takes the findings of Craig's PCA on the 1602 Additions to Kyd's *The Spanish Tragedy* as a starting point for his own analysis of these additions (see Craig 2009). His criticism of Craig is based on two major aspects: for one, he sees a problem in the false positives that Craig's approach created, i.e. wrong attributions of non-Shakespearean passages (Taylor 2017, 247)[10]; and, secondly, he criticises the implicit assumption (not only by Craig) that the 1602 Additions were single-authored (see *ibid.*, 248). His own approach hence does not focus on PCA, but on unique parallels in what he calls "micro-attribution of small blocks of texts" (*ibid.*, 249), i.e. he analyses n-grams and collocations in the first 173 words of the first addition. He incorporates "variant grammatical forms, whenever there were no matches for grammatically identical word strings" as well as "[d]isjunct trigrams and quadgrams" (*ibid.*, 250) whenever "no exact strings" could be found; he also includes "variant spellings" (*ibid.*, 255). The corpus used for his approach is the Literature Online database (LION) as well as Early English Books Online – Text Creation Partnership (EEBO-TCP) in a later run of the experiment).

---

[10] "Shakespeare has an exceptionally large, varied, and stable canon, which increases the sheer quantity and variety of data in his canon, in ways that may disproportionately weight results in his favour" (Taylor 2017, 247).

The following table shows his findings on the basis of this analysis (*ibid.*, 251):[11]

to night frolicke and mery	Heywood, *Woman Killed with Kindness* ('merry, pleasant, / And frolick it to night')
frolicke and mery	Heywood, *Edward IV, Part One* ('a mery mate, / So frolycke, and')
had no custome to	Heywood, *Rape of Lucrece* ('hath beene no custom … to')
to stay out … late	Heywood, *Lucrece* ('to stay out late')
he may be in his	Lodowick Carlell, *Passionate Lover, Part 2*
be in his chamber	Marlowe, *Faustus* (both versions)
some go see	Heywood, *If You Know not me, Part One* ('some one goe see')
it besides he	William Heminge, *Fatal Contract*
is so generally	Anonymous, *Two Wise Men and all the rest fools*
generally beloved	Jonson, *Magnetic Lady*
did grace him	Middleton and Dekker, *Bloody Banquet* 4.2 (Middleton scene)
waiting on his cup	Heywood, *Four Prentices* ('wait upon his cup').
assure he	Anonymous, *Look about you*
I wonder how this fellow	Middleton, *Phoenix*
got his clothes	Fletcher, *Women Pleas'd*
Pedro … the Duke of Castiles	Thomas Kyd, *Spanish Tragedy*
do ye hear me	Henry Porter, *Two Angry Women of Abingdon*
blush not man	Thomas Lodge, *Wounds of Civil War*

His intermediate result from this analysis consists in the fact that there was no "single unique parallel with Shakespeare" (*ibid.*, 252), "the language of this passage is less like the language of Shakespeare, at any stage of his career and in any genre, than it is like the language of at least ten other playwrights" (*ibid.*). He also finds that "a single unique parallel, in a passage of this length, is insignificant, a mere linguistic coincidence" (*ibid.*). The two playwrights that have more than one unique parallel are Middleton and Heywood, and Taylor continues his research on the basis of these two in a follow-up experiment. The second experiment used EEBO-TCP, a larger corpus, and Taylor eliminated "parallels that (a) involve a string of five successive words, or (b) occur more than ten times in the EEBO-TCP search, or (c) belong to authors not alive or not possibly writing for the theatre in 1594-1602" (*ibid.*, 254), which left him with eight parallels only. Most of them are from Heywood, which leads him to the conclusion that Heywood is the "leading candidate" (*ibid.*) as author for the first addition.

As Shakespeare has recently most often been claimed to be the author of the additions, Taylor ran Shakespeare against Heywood (again in EEBO-TCP), which again leaves Heywood as the most likely author. With regard to the remainder of the first addition as well as the Painter's scene, he drew on research conducted by John V. Nance (2017) to eventually conclude that the 1602 Additions are, altogether, most likely the collaborative work of Heywood and Shakespeare. So far so good.

---

[11] Taylor's commentary on the parallels has been left out here for the sake of space.

To evaluate Taylor's method, it makes sense to focus on two items: "do ye hear me" – which, according to Taylor, has a unique parallel in Porter – and "waiting on his cup" with its "unique parallel" in Heywood. It is surprising that "do ye hear me" is supposed to have only one unique parallel in the whole corpus of early modern drama. A simple concordance search of Shakespeare's works yields the following results[12]:

> *Much Ado About Nothing* 5.1: Do you hear me
> *The Tempest* 2.1: Do you not hear me speak?
> *3 Henry VI* 5.5: Didst thou not hear me

The substitution of "ye" with "you" and "thou" is, according to Taylor, legitimate, as he includes spelling variants, and as "ye" is a notorious candidate for printers' substitutions.[13] In the light of Taylor's own criteria – the acceptance of disjunct trigrams and quadgrams, incorporation of variant grammatical forms and spelling variants –, it is rather surprising that he did not come up with these results. And if we move on to the next example, "waiting on his cup", in the *Spanish Tragedy* 1602 Additions, we come to the conclusion that he is measuring the text with different standards as he, in this case, accepts an even more different quadgram – "wait upon his cup" in Heywood – as a unique parallel.

Besides these methodological questions that we may ask, the approach begs another question as well, and this leads us back to Craig. In the context of his own research on the 1602 Additions, Craig notes:

> The Additions, then, have some intimate connections with the drama of their day. Some of these may well be allusions to passages and phrases the audience could be expected to know well. Others are equally direct but so adapted to the communicative purposes of the moment that they seem (as best one can tell) better explained by the same mind reverting to idiosyncratic habits in the process of composition. This sort of evidence […] requires the exercise of judgment, and so is not altogether objective, but it is clear that there are some strong links between the vocabulary and phrasing of the Additions and canonical Shakespeare in particular. (2009, 170-171)

What he seems to leave out completely – and this brings us back to our criticism of character style analysis by PCA – is the fact that an author may wittingly or unwittingly imitate and allude to another author, in order to make his audience recognise the intertextual link. The individual aesthetic quality of a text here comes into play again, and we may wonder what exactly it actually means when strings and collocations are identical. Just as both the individual and the genre-specific style of an invented character can work against determining an author's stylistic identity by means of word frequency, the widely practised culture of imitation works against identifying authors by identical sequences of words – only think of the educational practise of writing down in commonplace books not only "vocabula usus quotidiana", but also, and especially, "vocabula rara, exquisita […] idiomata et formulas loquendi" (Vives 1537, 6). Stylometrics, we think, should leave behind its fixation on author identification and lead to a number of other exciting questions about what authors are doing with words.

---

[12] See <https://www.opensourceshakespeare.org/concordance/> [last accessed 13 January 2018].

[13] See Horton (1994, 322): "Many common function words have several variant spellings". See also Love (2002, 105).

References

Alberge, Dalya (2016): "Christopher Marlowe Credited as One of Shakespeare's Co-Writers," *The Guardian*, 23 October 2016. <https://www.theguardian.com/culture/2016/oct/23/christopher-marlowe-credited-as-one-of-shakespeares-co-writers> [last accessed 17 January 2018]

Burrows, John (2017): "The Joker in the Pack? Marlowe, Kyd, and the Co-authorship of *Henry VI*, Part 3," in: Taylor, Gary; Egan, Gabriel (eds.): *The New Oxford Shakespeare: Authorship Companion*. Oxford: Oxford UP, 194-217

-----; Craig, Hugh (2012): "Authors and Characters," *English Studies* 93.3, 292-309

Chambers, Edmund K. (1930): *William Shakespeare: A Study of Facts and Problems*, 2 vols. Oxford: Clarendon. Vol. 1

Craig, Hugh (2008): "'Speak, that I may see thee': Shakespeare Characters and Common Words," *Shakespeare Survey* 61, 281-288

----- (2009): "The 1602 Additions to *The Spanish Tragedy*," in: Craig, Hugh; Kinney, Arthur (eds.): *Shakespeare, Computers, and the Mystery of Authorship*. Cambridge: Cambridge UP, 162-180

----- (2010): "Style, Statistics, and New Models of Authorship," *Early Modern Literary Studies* 15.1. <https://extra.shu.ac.uk/emls/15-1/craistyl.htm> [last accessed 13 January 2018]

----- et al. (2016): "A Novel Clustering Methodology Based on Modularity Optimisation for Detecting Authorship Affinities in Shakespearean Era Plays," *PLoS ONE* 11.8, 1-27

Egan, Gabriel et al. (2016): "Attributing the Authorship of the Henry VI Plays by Word Adjacency," *Shakespeare Quarterly* 67.2, 232-256

van Halteren, Hans et al. (2005): "New Machine Learning Methods Demonstrate the Existence of a Human Stylome," *Journal of Quantitative Linguistics* 12.1, 65-77

Hart, Alfred (1934): "Shakespeare and the Vocabulary of 'The Two Noble Kinsmen,'" *Review of English Studies: A Quarterly Journal of English Literature and The English Language* 10.39, 274-287

Horton, Thomas B. (1994): "Distinguishing Shakespeare from Fletcher through Function Words," *Shakespeare Studies* 22, 314-335

Jackson, MacDonald P. (2014): *Determining the Shakespeare Canon: 'Arden of Faversham' and 'A Lover's Complaint.'* Oxford: Oxford UP

Juola, Patrick (2008): "Authorship Attribution," *Foundations and Trends in Information Retrieval* 1.3, 233-334

Littledale, Harold (1876): "Bibliography," in: *The Two Noble Kinsmen. Reprint of the 1634 Quarto*. London: New Shakespeare Society

Love, Harold (2002): *Attributing Authorship: An Introduction*. Cambridge: Cambridge UP

Nance, John V. (2017): "Shakespeare and the Painter's Part," in: Taylor, Gary; Egan, Gabriel (eds.): *The New Oxford Shakespeare: Authorship Companion*. Oxford: Oxford UP, 261-277

Rudman, Joseph (2016): "Non-Traditional Authorship Attribution Studies of William Shakespeare's Canon: Some Caveats," *Journal of Early Modern Studies* 5, 307-328

Sykes, H. Dugdale (1916): "'The Authorship of 'The Two Noble Kinsmen,'" *The Modern Language Review* 11.2, 136-148

Taylor, Gary (2016): "Collaboration 2016," in: Callaghan, Dympna; Gossett, Suzanne (eds.): *Shakespeare in Our Time: A Shakespeare Association of America Collection*. London: Bloomsbury, 141-149

----- (2017): "Did Shakespeare Write 'The Spanish Tragedy Additions'?" in: Taylor, Gary; Egan, Gabriel (eds.): *The New Oxford Shakespeare: Authorship Companion*. Oxford: Oxford UP, 246-260

-----; Egan, Gabriel (eds.) (2017): *The New Oxford Shakespeare: Authorship Companion*. Oxford: Oxford UP

Vickers, Brian (2009): "The Marriage of Philology and Informatics," *British Academy Review* 14, 41-44

----- (2016): "The Misuse of Function Words in Shakespeare Authorship Studies." <https://www.etrap.eu/wp-content/uploads/2016/12/2016-11-30-vickers-misuse.pdf> [last accessed 13 January 2018]

----- (2017): "Shakespeare and the 1602 Additions to 'The Spanish Tragedy': A Method Vindicated," *Shakespeare* 13.1, 101-106

Vives, Juan Luis (1537) : *De Ratione Studii Puerilis Epistolae Duae*. Basile. <https://books.google.de/books?hl=de&id=vTU8AAAAcAAJ&q> [last accessed 13 January 2018]

ANNIKA ELSTERMANN (HEIDELBERG)

The Power of Digital Publishing: Rethinking Knowledge Distribution in English Philology

1. The "Digital" in Academic Research

As electronic media and devices increasingly shape our culture and our society, it comes as no surprise that they have also become a significant factor in those disciplines that examine how culture and society work. The digital has found its way into the humanities, and English philology is no exception: literary scholars are analysing computer games and hyperpoetry, while linguists are investigating tweets and other born digital corpora. While more and more research is being conducted on electronic source material, researchers also increasingly utilise digital methods and approaches for their analyses of digital as well as traditional texts – corpus linguistics and its equivalent in literary studies, distant reading, being examples of quantitative methods enabled by the calculating capacities of computers. As our field moves ever closer towards the digital, we need to consider the impact of new technologies not only on our subject matter and methodology, but also on how we publish the results of our research.

The fundamental idea behind every publication process is of course the same: making the results of a research project public, distributing knowledge. Beyond that, however, digital publishing comes with its own implications, advantages and potential issues, which are distinct from those of print publishing. Because they offer so many new opportunities, electronic publications – for example digital editions of journals – have quickly become the norm in many disciplines; now, we must take a closer look at those new opportunities, many of which remain virtually untapped, and determine how they can best serve academic communication of research in English philology.

A transfer of digital sources and approaches to traditional print publications might lead to some form of unnecessary reduction. Digital publications, however, no longer need to adhere to the constraints of print, and so-called 'born digital' research output can potentially benefit from alternative forms of knowledge organisation, hyperlinks, interactive visualisation, connection to corpora and research databases, and countless other options.

In addition to changing the form or medium of research presentation, the possibilities of electronic publication have opened up new avenues for the dissemination of research output. It can be argued that the subscription-based model of academic journals, for example, is a remnant of a waning print age, and is no longer adequate in a time when knowledge can instantly be distributed worldwide at little to no extra cost (see Suber 2005, 231). Additionally, an ethical argument can, of course, be made with regards to who can access that knowledge and whether affluence should be a factor.

This chapter will explore some of the questions and implications of (open) access. The goal is not to present a one-size-fits-all solution or definitive answers, but rather to