

A computational approach to the poetry of Katherine Philips

Wright, Gillian; Taylor, Kathleen

DOI:

[10.1080/09699082.2016.1179399](https://doi.org/10.1080/09699082.2016.1179399)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Wright, G & Taylor, K 2016, 'A computational approach to the poetry of Katherine Philips', *Women's Writing*.
<https://doi.org/10.1080/09699082.2016.1179399>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

This is an Accepted Manuscript of an article published by Taylor & Francis in *Women's Writing* on 28th June 2016, available online:
<http://www.tandfonline.com/10.1080/09699082.2016.1179399>

Checked 21/7/2016

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Kathleen Taylor and Gillian Wright

A computational approach to the poetry of Katherine Philips

Katherine Philips's poetry survives in several early manuscript and printed sources. While these witnesses have been extensively studied by literary scholars, much about their textual origins and relationships remains unclear. In this article, we apply computational methods to analyse such aspects of these witnesses as similarity, compilation, and organization. We also consider how use of such methods can shed light on long-standing literary questions about Philips's poems.

Keywords: compilation, computational analysis, digital humanities, poems, statistics, textual relationships

Introduction

The purpose of this article is to investigate the relationships between four major manuscript and two print witnesses to the poetry of Katherine Philips. These witnesses—the Tutin, Dering, Clarke, and Rosania manuscripts, and the printed *Poems* of 1664 and 1667—were produced between the mid 1650s and the late 1660s, and collectively comprise the principal early collections of Philips's verse.¹ While all these collections have been extensively analysed using traditional literary methods, much remains unclear about how each was produced and how they relate to one another.² Our aim is to investigate whether, and if so how, a computational approach can contribute to understanding these issues and can prompt new ideas and hypotheses about Philips's work.

We combine simple quantitative methods and literary-historical information to address the following questions. Can such methods offer insights into the textual relationships between early witnesses to Philips's poetry? Can they shed light on cases where manuscript witnesses include physical evidence suggestive of missing poems? Can they identify patterns in the ordering of poems across witnesses? If so, what can these patterns tell us about relationships between witnesses?

Our quantitative methods extend recent innovative work in the digital humanities to Philips scholarship, and may also prove applicable to other poets with similarly complex histories of textual transmission.³ Applying such methods to the Philips corpus enables a complex dataset to be analysed more systematically, and expeditiously, than by conventional means, and with less risk of researcher bias. It also allows for much easier visualization and comparison of complex materials, making patterns in the data clearer (e.g. Figure 3(a) below). Furthermore, quantifying relationships between witnesses (e.g. using correlation and curve-fitting, see below) can strengthen traditional literary-historical interpretations and generate new questions and ideas. We believe that our results are not only consistent with the findings of traditional Philips scholarship but are also a fruitful source of new ideas about her work.

Materials and methods

Selection of sources and poems

Our six sources—the four manuscripts and two printed editions—were chosen because of their close relationship with Philips herself and because each represents the first known witness to a significant number of her poems.⁴ We excluded any manuscripts deriving from printed texts and also later printed collections with no independent authority. Another

early witness, the Cardiff manuscript, was excluded from most aspects of our analysis because of its short extent (it includes just 14 poems by Philips, three of which are of doubtful authenticity), which renders comparisons with longer collections problematic.⁵ We did, however, include Cardiff in our analysis of similarities as a means of testing the validity of the method.

We consider only the content and ordering of poems within each witness, ignoring issues of textual variation within poems. We identify poems which tend to be placed together in different witnesses. We also develop measures of similarity between witnesses, allowing us to draw inferences about their structure and development and to compare these with known data about the history of Philips's poetry production.⁶

Witness coding

For each witness, its poems' titles were transcribed into a table of contents. Incomplete poems were counted as whole poems. Poems in the hand of a later scribe were treated separately.⁷ The numbers of poems were, respectively: Tutin, 57; Dering, 76; Clarke, 73; Rosania, 91; 1664, 75; and 1667, 116. All were then merged into an overall list of 121 of Philips's poems.⁸ Titles were listed in the forms provided by 1667 (the most extensive of the six witnesses); some were shortened for ease of reference. Spellings were modernized, and the list was alphabetically ordered.⁹

Each poem was then given a unique identifying number, POEMCODE, based on its position in the list. (For the complete list, with POEMCODES and poem numbers in *CWI*, see the Appendix.) Each table of contents was recoded into numeric (vector) format by replacing titles with their equivalent POEMCODES.¹⁰ This enabled us to treat the data quantitatively, considerably expediting comparisons between witnesses and rendering visualizations much easier. The lack of content markers also allowed inferences to be drawn without reference to—and hence bias by—pre-existing knowledge about the poems involved.¹¹

Each data point in a witness vector represents a poem in that witness, and its index gives its location within the witness. Thus 1667₍₃₎, the third poem in that edition, has POEMCODE 17 in our overall list; its modernized title is “Arion on a Dolphin”. The Tutin manuscript is a special case, in that its poems are arranged in two separate sequences, proceeding from each end of the manuscript and overlapping centrally. We therefore created two vectors, Tutin1 and Tutin2, which contained one sequence (of philosophical poems) followed by the other (more personal works), and *vice versa*.

Missing pages or stubs were coded as zero, but only if they obviously interrupted transcription (see below).

Data analysis

Initial coding was carried out using Microsoft Excel. Data analysis, curve-fitting, and visualization were done using Excel, the high-level programming environment Octave, and the data-analysis programme SPSS.¹²

Which poems are in which witnesses?

Similarities between witnesses

To assess the similarity of two manuscripts is not straightforward, even if one simplifies—as we have done—by considering merely the poems' identity rather than their textual content, thereby reducing each witness to a list of numbers. Statistical methods do not offer any single “similarity” measure; instead one must choose whichever seems most appropriate to the materials and context. We therefore compared witnesses using Sorenson's quotient of similarity, QS.¹³ This measure is widely used in ecology to compare biodiversity (number of species) between two environments. Extending by analogy from species to poems

and from environments to witnesses, QS computes how many poems two witnesses share, irrespective of their ordering.

$$QS = \frac{(2 * (\text{number of poems in both witnesses}))}{((\text{number of poems in witness 1}) + (\text{number of poems in witness 2}))}$$

Table 1 gives the QS values for each pair of witnesses, in descending order of similarity from 1 (identical poems) to 0 (no poems in common).

Table 1: similarity values between witnesses

QS	Witnesses				Poems in Common
	(Name, Number of Poems)				
0.99	1664	75	Clarke	73	73
0.97	1664	75	Dering	76	73
0.97	Dering	76	Clarke	73	72
0.90	Dering	76	Rosania	91	75
0.89	1664	75	Rosania	91	74
0.88	Rosania	91	Clarke	73	72
0.87	1667	116	Rosania	91	90
0.83	Tutin	57	Clarke	73	54
0.82	1664	75	Tutin	57	54
0.80	Dering	76	Tutin	57	53
0.79	1667	116	1664	75	75
0.78	1667	116	Dering	76	75
0.77	1667	116	Clarke	73	73
0.72	Rosania	91	Tutin	57	53
0.65	1667	116	Tutin	57	56

The QS indicates that, in terms of their poetic content, 1664 is quite like Rosania, but is most similar to the Dering and Clarke manuscripts (which are themselves very similar). 1667 is quite similar to 1664, Clarke, and Dering, but most similar to Rosania. Tutin bears some resemblance to Clarke, Dering, and 1664, but less to 1667. The two printed editions are somewhat alike, but much less similar than, say, Dering and 1664.

Our similarity data fit with what is known about the periods during which the witnesses were produced: Tutin assembled in the 1650s, Dering and Clarke compiled in 1662-3, and Rosania completed between Philips's death in 1664 and the publication of the 1667 edition. As a further check, we compared results for another early collection of Philips's poetry, the Cardiff manuscript. Given Cardiff's shorter length and the doubtful authenticity of some of its contents, we predicted that comparing it with other witnesses should produce much lower QS values. This was the case. QS values ranged from 0.25 (comparing Cardiff with 1664) to 0.17 (Cardiff/1667), considerably lower than the lowest value of 0.65 in Table 1. This confirms that QS is successfully capturing similarity of content between witnesses.

Unique, frequent, and missing poems

The QS measure quantifies the extent to which witnesses share poems. As implied by the high values in Table 1, many poems are shared between five or more witnesses.¹⁴ A few of Philips's poems, however, are each found only in one or two of our six witnesses.¹⁵ Furthermore, some poems appear to be missing from some witnesses, a possibility supported by physical evidence of missing material in two of our witnesses, Dering and Tutin.¹⁶

The Dering manuscript, otherwise transcribed on consecutive leaves, has two blank pages between Poems 8 and 107. The Tutin manuscript now contains 28 stubs, each representing an excised leaf.¹⁷ However, further examination indicates that not all of these stubs are plausible loci for once-extant poems. Some stubs, for instance, occur mid-poem but do not interrupt the transcription of the text: thus, the stub midway through Poem 7, "A Friend", does not interrupt the text of the poem or the catchwording of the leaves.¹⁸ The most likely explanation in such instances is that the leaf was excised by Philips herself, probably due to a transcription error.¹⁹ A few further stubs can also be readily explained: thus, for instance, the missing leaf between Poems 115 and 78 evidently once contained the final stanza of Poem 115, which runs to three stanzas in all other witnesses (including a further manuscript in Philips's hand) but has only two in Tutin.²⁰ When all such stubs are excluded, ten possible sites for missing poems remain.

Can our computational approach not only identify similarities between witnesses but use them to suggest plausible candidates for the missing poems? To answer this question we must first look at relationships between the witnesses themselves.

"Missing links"

The witnesses considered in this article were not the only copies of Philips's poetry extant in the 1650s and 1660s. Surviving documents include the printed miscellany *Poems, by several persons* (Dublin, 1663), which includes three of Philips's poems.²¹ In addition, it can be inferred that further witnesses, now lost, must once have existed in order to source the new materials in the later manuscripts and printed texts. By assessing which poems are extant in which witnesses, it may be possible to shed some light on the minimum number of (presumably manuscript) witnesses that must have existed to account for the contents of the printed editions.

The editor of 1667 claimed to have made this posthumous collection of Philips's poetry as complete as possible.²² Implicit in his/her account of the volume's construction is the existence of multiple early witnesses from which 1667 had been compiled. Since 1667 is known to have drawn heavily on 1664, we therefore asked: what is the most parsimonious number of witnesses from which the poems in 1664 and 1667 could have been assembled? Our conclusion is that at least two, and perhaps up to five, additional manuscript witnesses were required (designated Witnesses A-E below; see also **Figure 1**). Our reasoning is set out below.

Compilation of 1664

Of the 57 poems in Tutin, 54 are also in Clarke.²³ However, Clarke also contains 19 further poems.²⁴ At least one of these was once included in Tutin, which is known to have lost leaves at some point in its history: a manuscript copy of poem 11, which once formed part of the Tutin manuscript, now survives at the University of Kentucky.²⁵ However, it is highly unlikely that Tutin ever included all 19 of the poems found in Clarke, several of which date from 1660 or later, given the probable date of Tutin (which Philips seems not to have worked on after 1658) and the order and location of the stubs (discussed below). The most parsimonious assumption is that these 19 poems derive from a single manuscript (which we designate Witness A).

All of Clarke's 73 poems are included in 1664 (the QS score of 0.99 indicates near-identity); so again, the simplest inference is that Clarke is a substantive source for this edition.²⁶ Clarke is a more plausible source than the very similar Dering manuscript because the latter includes a poem (no. 109) which is not in 1664. While it is conceivable that poem 109 could have been deliberately left out of the printed edition, this seems unlikely, as there is no obvious reason for its omission—and poems once included in Philips witnesses tend to be retained. (That Clarke is a more likely source than Dering for 1664 is supported by the ordering of poems in these three witnesses; see below.)

1664 contains two poems additional to Clarke's 73: nos 41 and 112. The former is in Dering, and both are in Rosania. However, Rosania is known to have been completed after the 1664 edition, so it is unlikely to be amongst the latter's sources.²⁷ What, then, was the editor's source for Poem 112—and for Poem 41—if as we propose Clarke was the principal source for 1664, rather than Dering? We (parsimoniously) designate the unknown source(s) of these poems Witness B. Witness B is likely to be quite an early production, since it is a potential source both for 1664 and for the copies of Poem 41 in Dering and in *Poems, by several persons*.

We infer, therefore, that at least one and possibly two additional witnesses to Philips's poetry were available to the compiler of 1664.

Compilation of 1667

The 1667 edition drew heavily on 1664, using all 75 of its poems. Of those poems in 1667 which are not in 1664, 15 full-text poems are also in Rosania. These include two poems preserved as title-only in Dering. There remain 23 poems in 1667 which do not appear in the four manuscripts discussed here. A further source, Witness C, is necessary to supply these texts.

Moreover, there are three Tutin poems whose status with respect to 1667 is puzzling. One is poem 22 ("Epitaph on Mr. John Lloyd"), which is in Tutin but not in 1667. The other two are nos 57 and 87, incomplete in Tutin and found otherwise only in 1667. It is conceivable that the editor may have used Tutin but deliberately omitted Poem 22, perhaps not believing it to be by Philips. It is an epitaph written as if by Lloyd's wife, so doubt as to its authorship is understandable. However, it is equally possible that the decision to omit all three poems was made by the compiler of Clarke (on grounds of incompleteness for nos 57 and 87 and doubtful authorship for no. 22), and was by this means transmitted to later editors of Philips's work; thus, the editor of 1667 need not have had access to Tutin.²⁸ (The Tutin-1667 QS score of 0.65 is the lowest in Table 1, noticeably lower than for Clarke-1664 or Clarke-1667.) However, he or she must have obtained the full versions of Poems 57 and 87 from somewhere. We therefore infer the existence of a further intermediate document, Witness D, used by the 1667 editor.

The fact that 1667 contains many poems which are also found in Rosania (but not earlier witnesses) suggests, *prima facie*, that Rosania could have been a source for the printed edition. There is, however, one anomaly: Poem 109 ("To the Lady Mary Butler"), which is

extant in Rosania, Dering, and *Poems, by several persons*, is not included in 1667. How can this omission be explained? One possibility is to postulate an additional intermediate document, Witness E, containing some of Rosania's poems but not this one, which was used to help compile the second edition. As with 1664, there is no obvious reason for Poem 109 to have been deliberately omitted from 1667.

It is possible that Witnesses C, D, and E could be one collection of Philips's poetry, containing the later material unique to 1667, the poems first collected in Rosania (though perhaps not no. 109), and the Tutin poems nos 57 and 87 (but not no. 22), and probably other material as well.

Assembling 1667, therefore, would have required at least one and possibly up to three additional witnesses, as well as the one or two required for 1664. Thus, we conclude that there are, or were, at least two further manuscript collections of Philips's work circulating in the mid-1660s, and perhaps up to five or more, as well as the four discussed here (plus Cardiff and *Poems, by several persons*).

What can we learn from poem ordering?

Having described the content of the witnesses, we next considered the ways in which poems were ordered within them. For example, are certain poems consistently placed together, leading to "clustering" of poem sequences? Where there is variation between witnesses, can we identify "cut and pasted" clusters which are identical or similar between witnesses? Or is there much more evidence of reorganization (which would suggest a repurposing of the witness) and/or addition (which might suggest additional sources)?

For example, plotting poems' positions for 1664 and 1667 (see [Figure 2](#)) shows their considerable overlap, interspersed with poems new in 1667. This suggests that the editor of 1667 worked primarily by interpolating additional poems within the pre-existing edition's structure, as well as adding extra material at the end. Only minor changes were made to the order of poems. In 1664, by contrast, our findings suggest that the editor was not just collecting material, but was purposefully re-ordering and re-presenting it.

Correlations

To quantify the relationship between the orders of poems in pairs of witnesses such as the two editions, we calculated the correlation coefficient r between the positions of poems extant in both witnesses, for all possible witness pairs. A correlation coefficient of $r = 0$ would indicate no obvious relationship between the order of poems in one witness and that in the other. A perfect positive correlation ($r = 1$) would represent exactly the same order in both witnesses; a perfect negative correlation ($r = -1$) would represent exactly the opposite order. Typically values of r between 0 and 0.3 are considered weak, those between 0.3 and 0.5 medium, and higher values are considered strong.²⁹ So, for example, the correlation between 1664 and 1667 is extremely strong ($r = 0.96$).

Our default assumption is that our four manuscript witnesses were independently created each for the purposes of its owner and/or scribe. If so, one might expect the correlations between their poems' ordering to be low. If they served as sources for 1664 and 1667, the correlations between them and the editions should be higher. In general, a higher correlation between two witnesses is consistent with the possibility that one may have served as a source for the other.

What we found is that the correlations divide naturally into four groups: weak (r between -0.27 and 0.27), positive (r between 0.37 and 0.39), strongly positive (r greater than 0.75), and negative (r between -0.40 and -0.70). Figure 1 shows the correlations between witnesses; it also shows our postulated links and additional witnesses.

The weak group includes the correlations between Clarke and Rosania, Clarke and Dering, and Dering and Rosania. Although these three manuscripts contain many of the same poems (as shown by their high QS values), their arrangement differs considerably. This suggests that they were assembled by three independent agents organising the manuscripts for their own purposes, albeit perhaps drawing on common sources. Dering and Rosania show weak correlations with 1664 and 1667, whereas the editions' correlations with Clarke are stronger (see below). This strengthens our previous inference that Clarke is a more likely source than either Dering or Rosania for 1664.³⁰

For Tutin, which as noted earlier can be read from either end, Tutin1 and Tutin2 have distinctly different relationships with other witnesses (see Figure 1). Tutin1 is strongly negatively correlated with both editions ($r = -0.68/-0.42$ for 1664/1667 respectively), whereas for Tutin2 the coefficients are positive ($r = 0.39$ for both editions). We interpret this as implying that 1664 was more likely to have used Tutin2 than Tutin1 as a basis for its ordering, if it used Tutin directly at all. Both versions of Tutin have weak correlations with Rosania, suggesting that this manuscript did not draw heavily on Tutin with respect to the arrangement of its poems. Tutin1 is strongly correlated with Clarke ($r = 0.76$), whereas Dering correlates very highly with Tutin2 ($r = 0.95$). The converse relationships—between Tutin2 and Clarke, and Tutin1 and Dering—are insignificant. This suggests that the compilers of Clarke and Dering, if they had access to the Tutin manuscript, probably approached it from opposite ends when drawing up their own manuscripts.

Clarke is strongly correlated with both editions. However, the correlation coefficients for the relationships between Clarke and the editions are negative: $r = -0.69$ for 1664 and -0.67 for 1667. This suggests that the order is to some extent reversed between Clarke and 1664; of this more shortly. Since Clarke and Tutin1 also have a strong positive correlation, while Tutin1 is negatively correlated with 1664 and 1667, the data fit our suggestion that Tutin1 was incorporated into Clarke, which was then used in the compilation of 1664.

Visual alignment

Our QS measure assessed the similarity of the poetic content between pairs of witnesses, but we are also interested in how similarly poems are ordered within witnesses. To assess this “order homology”, we considered pairs of witnesses for which we already suspected a possible relationship based on their strong correlation coefficients, i.e. Tutin1 and Clarke, Tutin2 and Dering, and Clarke and 1664. We used visual alignment to facilitate comparisons of poem order, as shown for all six witnesses in [Figure 3\(a\)](#).

Tutin and Clarke: Figure 3(a) clearly shows that both the poetic content and the order of the poems are very similar between Tutin1 (first column) and Clarke (fifth column). There are exceptions: Tutin1's initial two poems (nos 8 and 42) are near the end of Clarke and poem pairs 60 and 99, and 96 and 97, are reversed. Clarke contains 12 extra poems: no. 100, inserted midway through the manuscript, and a sequence of 11 extra poems included towards the end of the manuscript. The obvious explanation—that these latter poems were added later (either to Clarke itself or to its immediate source Witness B)—is borne out by their contents: they comprise Philips's elegy on her stepdaughter (dated May 1660) and the post-Restoration royal poems, as well as “The Enquiry”, which on this evidence is also likely to derive from the early 1660s.³¹

Tutin and Dering: comparing Tutin2 (Figure 3(a), third column) and Dering (sixth column) likewise shows a similar pattern: some sequences of poems match exactly, others are slightly rearranged. Most of the poems found in Dering but not Tutin are concentrated in two areas within the former: a sequence of 10 poems transcribed towards the end of Dering (Poems 67-51-113-17-120-48-45-28-74-46) and a further sequence of five added at its very end (Poems 114-68-65-109-41). The sequence of 10 includes all the poems found in Clarke's concluding 11, in a similar order; we speculate that these poems may likewise have derived

from Witness B. The remaining poem, no. 114, is the first in Dering's concluding sequence of five. The other four members of Dering's concluding sequence are also likely or certain to derive from late 1662 or early 1663, and so are probably late additions to the manuscript.³²

Filling in the gaps

We next used visual alignment to address the question of what might have filled the gaps in the Tutin and Dering manuscripts. (Figure 3(a) shows the gaps as zeros with thick borders.) For Tutin1, the gaps between Poems 8 and 42, 69 and 36, 37 and 26, 76 and 110, and 96 and 57 have no corresponding poem in Clarke.³³ We infer that these gaps may not have been stubs (in the sense of a poem or poems deliberately removed from the witness), but pages removed for some other reason, such as scribal error.

For the gap in Tutin between Poems 10 and 32, Clarke has Poem 2; for the gap between Poems 82 and 61, Clarke has Poem 81. For the gap between Poems 99 and 117 in Tutin, Clarke has Poems 111 and 72, suggesting that the gap may originally have contained two leaves.³⁴ Between Poems 103 and 115, Clarke has Poem 118. Between Poems 18 and 89, Clarke has Poems 4 and 11, again implying a gap of multiple leaves.³⁵ Figure 3(a) (second and fourth columns) shows our inferred poem orders for Tutin1 and Tutin2 with the gaps filled.

The gap in Dering falls between Poems 8 ("A Prayer") and 107 ("To the honoured Lady E. C."). Here, Clarke has Poem 42 ("On Controversies in Religion"). These three poems are also found together in Tutin2, and they are close neighbours in 1664 and 1667. We infer that Poem 42, 74 lines in length, could have been intended to fill the two-page gap in Dering.

In sum, our quantitative approach has identified plausible candidates for the gaps in Tutin and Dering, and can also discriminate between genuinely excised poems and "false" gaps.

Introducing clusters—Clarke and 1664

A negative correlation, hinting at reverse ordering, necessitates a different approach to the question of order homology between Clarke and 1664. We therefore used visual alignment to look for common clusters: groups of poems found together in Clarke and also in 1664—albeit perhaps in a different location. We used Clarke as the basis for this analysis because of our working hypothesis that Clarke may have been a source for 1664. It therefore made sense to identify clusters in Clarke initially, and then look at how these may have been transmitted to the 1664 edition.

To identify clusters we used the following algorithm. Taking the first poem from Clarke, we found its location in 1664. We then looked at closest neighbour poems in both witnesses. If these matched, they were considered part of a cluster in Clarke. This was repeated for the next closest neighbours until there was a mismatch. However, we were allowing for the possibility of poem rearrangement within clusters, so we did not restrict the matching to strict identity (that is, the clusters did not have to be in the same order for both witnesses, but they did have to contain the same poems). If there was a mismatch, therefore, we looked at further neighbour poems until it was clear that there was no order homology between the witnesses. A new cluster was then begun. In this way we worked through all the poems in Clarke, assigning every poem in the witness to a cluster (some clusters will contain only one poem).

For example, Clarke begins with Poems 27-19-71-69. The same group is found almost at the end of 1664, rearranged as 71-69-27-19. In Clarke, the group is followed by Poems 36, 34, 30, 52, whereas in 1664, the group is followed by Poems 112 and 41. These two groups have no poem in common so are clearly part of different clusters. This gives us an initial

cluster in Clarke of four poems (labelled C1; clusters were numbered sequentially according to their order within Clarke). We note that this method is most effective for identifying longer clusters, since with pairs and singletons it can sometimes be unclear as to whether they are part of one larger cluster, especially when the analysis is expanded to include other witnesses. We have therefore focused primarily, although not exclusively, on the longest clusters in our analysis: those with six or more poems (all other clusters contained four or fewer poems).

As shown in Figure 3(a), we identified 21 groups of poems making up Clarke's total of 73. However, on closer examination it became apparent that one triplet and a group of seven, which in Clarke are separated by Poem 52, in 1664 are grouped together as one cluster of 10 poems. We therefore considered these two groups a single cluster (C2) in both witnesses: namely a cluster of 11 poems in Clarke, from which Poem 52 is relocated elsewhere in 1664. (One additional cluster, C21, comprises poems extant in 1664 but not in Clarke.)

The longest clusters are therefore as follows: C5 (containing 13 poems), C2 (11), C4 and C17 (nine poems each), and C12 (a sextet).

To quantify the relationship between clusters in Clarke and 1664, we took the starting position of each cluster within each witness and plotted the data for 1664 against Clarke. We used curve-fitting techniques in SPSS to determine whether the resulting pattern was best described by a straight line—expected if the clusters were in a similar order in both witnesses—or some other curve, i.e. inverse, quadratic, cubic, power, exponential. The best fit (highest correlation coefficient) was found for a quadratic curve ($r = 0.62$, vs. linear $r = 0.14$). This means that, on the whole, clusters found near the start of the Clarke manuscript are near the end of 1664, while clusters in the middle of Clarke are centrally located in 1664. Clusters located towards the end of Clarke are more scattered through 1664. We interpret the data as suggesting that the editor of 1664, if using Clarke as a source:

- a) thought of its poems as clustering into groups—clusters which were to a considerable but variable extent maintained between the two witnesses;
- b) rearranged most of these clusters' locations within the overall structure, while keeping the clusters intact;
- c) rearranged the order of poems within individual clusters, while keeping the poems in the cluster together;
- d) did more relocation of clusters at the beginning and the end than in the central part of the edition, tending to move early Clarke clusters to later in the edition, while moving later Clarke clusters more diversely.

We infer that the editor of 1664 was not merely compiling, but purposefully organizing, the material available to him or her.

Cluster transmission

We next considered whether any of the five longest poem clusters found in Clarke and 1664 (C2, C4, C5, C12, and C17) could also be found in other witnesses. If so, this would suggest either that (a) the witnesses were drawing on a common source or (b) that the poems in the cluster have some obvious common factor, such as theme or addressee, which makes it natural to group them together; or both.

We used Octave to match the five clusters across witnesses, allowing for rearrangements within each cluster. As before, this initial analysis was carried out without attention to the content of the poems. Our findings indicate that clusters frequently but not invariably survive from witness to witness, and that this survival is probably due to a combination of shared sources and common factors.

Reading for source: C5, C2, C4, C12

These clusters are all intact, albeit with minor rearrangements, in Tutin, Dering, and 1667, but are extensively rearranged in Rosania. To cite a representative example, the 13-poem cluster C2 is complete, albeit rearranged, in 1664 and is slightly re-ordered in Dering, where it occurs earlier in the witness. All C2's poems are also found in 1667, but two (nos 106 and 117) have been split off from the rest of the cluster. All but one of the poems—no. 100—are clustered in Tutin2, and with an order very similar to Clarke (if our interpretation of how Tutin's gaps were originally filled is correct). In Rosania, however, all 13 poems are extant but scattered through the text.

Similarly, C5, C4, and C12 originated in Tutin and were preserved by the creators of Clarke, Dering, 1664, and 1667, presumably because each was thought to cohere into an ordered group. By contrast, the creator of Rosania, in which all three clusters are dispersed, either did not share this view or had other editorial priorities.

Reading for source: C17

The nine-poem C17 differs from the above-cited clusters in two key respects: it is not found in Tutin, and it is preserved (albeit rearranged) in the Rosania manuscript. In all other witnesses, the cluster is very similar (and invariably begins with Poem 51). The most divergent witness is Dering, which has all nine poems, but with one (Poem 114) moved elsewhere and another (Poem 28) interpolated within the cluster. The fact that C17 is so similar in all witnesses—even Rosania, in which materials are so often rearranged—suggests an especially strong unifying principle for its poems; and indeed, C17 will be familiar to most readers of Philips's poetry, since it comprises the post-Restoration royal poems.

Common factors

Close reading of the Philips clusters identifies several common factors linking their constituent poems. Of these, the two most notable are date of composition and theme or subject (as in C17). Clusters may be linked by one or more factors; details are given in [Figure 3\(b\)](#).

Below, we analyse three of the longer clusters—C12, C5, and C4—to assess their principles of coherence. We show how cluster coherence and transmission can help to account for the construction of two Philips witnesses, 1664 and Dering (and also for clusters lacking any obvious date- or theme-based coherence). We also show how attending to coherence or anomalies within clusters may shed light on two long-standing critical debates about Philips's writing: namely, which of her poems caused controversy with her early reader "J. J.", and how her poetry relates to Francis Finch's prose treatise *Friendship*.

Cluster coherence

C12 represents a comparatively straightforward example of single-factor cluster coherence. Its six poems—like approximately three-quarters of Philips's poems—are all addressed to or otherwise concerned with named individuals: 26 ("Friendship's Mystery, To my dearest Lucasia"), 18 ("Content, to my dearest Lucasia"), 4 ("A Dialogue of Absence 'twixt Lucasia and Orinda"), 11 ("A Sea voyage from Tenby to Bristol, begun September 5 1652. sent from Bristol to Lucasia"), 89 ("To my dearest Sister, Mrs. C. P. on her marriage"), and 76 ("To Mr. Henry Lawes"). Of the two most frequently named people in C12, Lucasia is a presence in the first four poems, but not the other two, while Henry Lawes, the addressee of the final poem, set to music both "Friendship's Mystery" and "A Dialogue of Absence" but had no known involvement with the other three. Neither Lucasia nor Lawes, therefore, unifies all six poems, which also span a variety of forms and genres. With both addressee and genre discounted, the factor most likely to link C12 is date of composition: all of the datable poems in this cluster derive from the period 1652-5.³⁶ A link by date would explain the otherwise odd inclusion of "To my dearest Sister", which does not obviously

resemble C12's other poems. It would also explain the presence of poems on both Lucasia and Lawes, given that Philips's relationships with both seem to have flourished in these years.

C5's common factors include both date and theme. Like C12, this cluster consists almost exclusively of poems on or addressed to named individuals, including several on Rosania and Lucasia.³⁷ All of its datable poems can be linked with the period between February 1650 (Poem 60) and September 1652 (Poem 101).³⁸ Furthermore, all of the Rosania and Lucasia poems included in this cluster derive from a specific period within Philips's relationships with both women: namely, the latter stages of her friendship with Rosania and the start of that with Lucasia. Poems deriving from earlier in her friendship with Rosania or later in her friendship with Lucasia are not included; nor are poems such as 77 and 9 which are datable to 1651 but stand outside the matrix of Rosania/Lucasia poems.³⁹ It seems likely, therefore, that subject as well as date was involved in Philips's decision to group these poems in the (holograph) Tutin manuscript, a grouping reproduced in most later witnesses. The exception—the Rosania manuscript, in which these poems are dispersed—may be attributed to tact on the part of its compiler. Scattering these poems would have helped to obscure the apparent decline in Philips's relationship with the manuscript's addressee.

This thematic/chronological reading of C5 is consolidated by the evidence of a further cluster, C4. Of its nine poems, all but one concern named friends of Philips's, including Rosania; the exception (Poem 24, "Friendship") is readily explicable, given its obvious overlap of subject with other C4 poems such as "To Mrs. M. A. at parting" which also theorize friendship. While not all of the poems in this cluster can be dated, most of those that can derive from the period 1649-51.⁴⁰ The temporal overlap between C4 and C5 further confirms the role of theme as well as date in contributing to cluster coherence, as the Rosania poems in C4 derive from the (for Philips) happier period before Mary Aubrey's marriage to William Montagu, while those of C5 mainly chart the later and more problematic phase in the two women's relationship. The one poem that seems to resist this date- and theme-based distinction between C4 and C5 may in fact prove the rule. The C4 poem "Rosania shadowed whilst Mrs. Mary Aubrey" belongs topically and temporally with the pre-marital poems of C4, but may have become associated with the post-marital poems of C5 because of the implicit allusion to Aubrey's marriage in its title.

Clusters and witness construction: 1664

Scholars have observed how the royal poems (our C17) are used in the 1664 edition (and later in 1667) to provide a strategic opening statement of Philips's royalism.⁴¹ This royalist effect in the printed editions is enhanced by their inclusion of Poems 118 ("Upon the double Murder of King Charles I") and 52 ("On the third of September, 1651", a meditation on the Battle of Worcester) immediately before and after C17. These poems never sit alongside C17 in earlier manuscript collections, where Poem 118 invariably circulates as a singleton (C9) and Poem 52 forms part of the philosophically focused C2. Their redeployment as top and tail for C17 in 1664 (and later 1667) testifies to the 1664 compiler's particular concern to shape the beginning of this collection.

The 1664 edition terminates with three clusters: C2, C1, and C21. The inclusion of C21, which is not extant in Clarke, at the very end of 1664 can be attributed to practicality: its contents, Poems 112 and 41, are among the most recent poems included in 1664 and may have reached the printer late in the preparation of the volume. Of the remaining two clusters, C2 is the most religious of the philosophically focused clusters (it includes Poems 66 and 1, "Submission" and "2 Corinthians 5.19"), while C1 (comprising Poems 71, "The World"; 69, "The Soul"; 27, "Happiness"; and 19, "Death") is the only cluster to comprise exclusively religious material. That the compiler of 1664 actively wanted to emphasize religion at the conclusion of the volume is further demonstrated by the rearrangement of poems within both C2 and C1 in this witness. C2, which does not conclude with Poems 66 and 1 in earlier

manuscripts, does so here, while C1 is revised from an order shared by all previous witnesses (from 27-19-71-69 to 71-69-27-19) to culminate not with the this-worldly preoccupations of “The World” and “The Soul” but with the spiritual consolation and quietism of “Happiness” and “Death”. The 1664 compiler, evidently keen to portray Philips as a good royalist at the start of the volume, was equally careful to present her as a good Christian at the end.

Clusters and witness construction: Dering

As noted above, non-Tutin material is incorporated into Dering towards the latter’s end. A sequence of ten new poems comprises Poem 67, “The Enquiry”, as well as all but one of C17 (the royal poems), and Poem 28, “In memory of F. P.”, while the five poems at the very end of Dering comprise the remaining poem from C17 (no. 114), the two title-only poems (68 and 65), and Poems 109 and 41. The poems intervening between these two sequences comprise C1 and C2, as well as the thematically diverse C19. Had Poem 42 been added to the manuscript, as we suggest was Dering’s intention, they would also have included the similarly diverse C20.

Comparison with cluster arrangement in other early witnesses, notably Tutin and Clarke, helps to explain not only the construction and organization of Dering but also the origins of C19 and C20. C19 comprises the last two complete poems in the personal sequence in Tutin—59 (“Parting with Lucasia, a Song”) and 14 (“Against Pleasure, Set by Dr. Coleman”)—as well as the first in the philosophical sequence, Poem 8 (“A Prayer”).⁴² C20, comparably, includes the second poem in Tutin’s philosophical sequence, no. 42 (“On Controversies in Religion”) and the final item in its personal sequence, 107 (“To the honoured Lady E. C.”), extant as title-only. The miscellaneity of these two clusters can thus be explained, albeit incompletely, by the two-sequence structure of Tutin, which is bridged in both C19 and C20. Notwithstanding their somewhat ad hoc origins, these clusters—which also occur at the end of Clarke—were subsequently transmitted to later collections of Philips’s work; both are reproduced in 1664 and 1667.

C19 and C20 aside, the concluding sections of Dering can be accounted for through a combination of thematic, chronological, bibliographical, and biographical factors. Thematic concerns are signalled by the (near-) preservation of C17, as well as of C1 and C2. That the latter two clusters are transcribed so close to the end of the manuscript may indicate that Dering, like the 1664 compiler, originally intended to conclude his collection on a philosophical-cum-religious note. Chronological factors probably explain the inclusion in or near C17 of “The Enquiry” and “In memory of F. P.”. The date of F. P.’s death—24 May 1660—associates Philips’s elegy on her step-daughter with the eight earlier C17 poems, which commemorate events between May 1660 and February 1662; as noted above, the inclusion of “The Enquiry” immediately before C17 in both Dering and Clarke suggests that it, too, is likely to date from this period. Bibliography and biography combine with chronology to account for the otherwise odd detachment of a single C17 poem (no. 114) from the rest of its cluster. The event celebrated in Poem 114, Catherine of Braganza’s arrival in Portsmouth in May 1662, postdates the latest event commemorated in the earlier royal poems (the death of Elizabeth of Bohemia in February 1662) by a few—but, for Philips, crucial—months. Philips is known to have learnt of Catherine’s arrival almost immediately before her departure for Ireland in June 1662.⁴³ Poem 114 thus seems likely to have been written during or after Philips’s journey to Ireland—a key transitional point in her life—and may have been textually separate from the other royal poems in Dering’s source.⁴⁴ Its inclusion later in the manuscript, close to poems attributable to Philips’s year in Ireland (definitely 68 and 109 and probably also 65 and 41), further indicates that towards the end of his work on the manuscript, Dering was probably adding poems (or parts of poems) as he received them. As in the 1664 volume, this late addition of newly composed material partially obscures the planned shape of the collection.

Clusters and the critical debate: “J. J.” and Francis Finch

C5 includes two poems—72 (“To Antenor, on a Paper of mine which J. J. threatens to publish to prejudice him”) and 116 (“To the truly competent judge of Honour, Lucasia, upon a scandalous Libel made by J. J.”)—that testify to an early controversy in the reception of Philips’s poetry. The “Paper” or “Libel” mentioned in these titles is often assumed in Philips scholarship to be her poem “Upon the double Murder of King Charles I” (no. 118).⁴⁵ However, no version of C5 includes Poem 118, and the only direct evidence for the association is provided by the (linked) headings of all three poems in the Rosania manuscript, which explicitly identify “Upon the double Murder” with Philips’s “Paper”.⁴⁶ This evidence is far from conclusive, as it is possible that the Rosania compiler was either mistaken or deliberately attempting to mislead readers.⁴⁷ Internal evidence in “To Antenor” and “To the truly competent judge of Honour, Lucasia” does little to explain to why Philips’s “Paper” had proved so controversial; nonetheless, the focus of the latter—honour—does not seem entirely consistent with the identification of the controversial poem as “Upon the double Murder”, which although clearly a political embarrassment to James Philips is less clearly an offence against his (or his wife’s) honour. If, however, the offending “Paper” was not “Upon the double Murder”, might it be among the other C5 poems? If this were the case, the most likely candidates to have provoked J. J. are such passionate friendship poems as “To the truly noble Mrs. Anne Owen” or “To my excellent Lucasia, on our friendship”, which, on this interpretation, J. J. might have threatened to make the subject of a sexual scandal. This might in turn explain why Philips addressed one of her poems of self-vindication (“To the truly competent judge of Honour”) to Lucasia, and why in this poem she represents the issue at stake as one of honour, and portrays herself as misunderstood and slandered by J. J.⁴⁸ Had J. J. interpreted “Upon the double Murder” as expressing royalist sympathies, this would not have been a misunderstanding.

Another puzzle generated by cluster analysis is the inclusion of “Mr. Francis Finch, the excellent Palaemon”, but not “To the noble Palaemon, on his incomparable Discourse of Friendship”, in C4. Like (though less overtly than) “To the noble Palaemon”, “Mr. Francis Finch” appears to allude to Finch’s treatise *Friendship*; it claims, for instance, that Finch had “rescu’d gasping friendship when / The bell toll’d for her funeral with men”.⁴⁹ However, although no date is given on the title page, *Friendship* carries internal dates of 30 October 1653 and 30 March 1654.⁵⁰ It thus postdates by 3-4 years the other datable poems within C4. A final discrepancy is that, while “Mr. Francis Finch” sits alongside Rosania/Mary Aubrey poems in C4, Finch’s treatise *Friendship* is closely associated with Philips’s friendship with Lucasia/Anne Owen: its prefatory letter is addressed “To the truly honourable Mrs A.O.” and the main text is headed “D. Noble Lucasia-Orinda”.⁵¹ For “Mr. Francis Finch” to keep company in so many witnesses with the Rosania rather than the Lucasia poems is counterintuitive.

If the inclusion of “Mr. Francis Finch” within C4 is not to be dismissed as an intractable anomaly, one possible explanation should be considered. This is that Philips’s poem does not, in fact, allude to the printed *Friendship* of 1654 but to an earlier, presumably manuscript, version of the same text. If this earlier version of *Friendship* were, for instance, to have been circulated in the late 1640s, Philips’s response to it could have been written around the same time as the other C4 poems, in 1649-1650. One might further speculate that Finch later revised the manuscript *Friendship* into the printed version of 1654, *inter alia* adding the references to Lucasia/Anne Owen, and that it was in response to this later version that “To the noble Palaemon” was written. This hypothesis, though currently unprovable, would account for several otherwise puzzling aspects not only of Philips’s Palaemon poems but also of *Friendship* itself.⁵² If Philips’s poems were each responding to a different moment in the textual history of *Friendship*, this would explain why, for instance, the two were

transcribed so far apart in the Tutin manuscript, which is largely ordered by chronology in the personal sequence.⁵³ If the original version of *Friendship* were not addressed to Anne Owen—or to both Owen and Philips, as “Lucasia-Orinda” implies—then the inclusion of “Mr Francis Finch” alongside the Rosania poems in C4 would no longer seem problematic. It might also explain why the main text of *Friendship*, despite the “Lucasia-Orinda” heading, so often seems to assume a single addressee.⁵⁴ An early dating of the manuscript *Friendship* would further explain why, despite Finch’s evidently royalist sympathies, the treatise never refers to the execution of Charles I. Indeed, the clear implication of several of Finch’s allusions in *Friendship* is that Charles was still alive at the time of writing.⁵⁵

Conclusion

The application of computational methods to literary research is a relatively new but fast-developing area.⁵⁶ For such methods to prove their worth, they must be shown to be consistent with existing knowledge, productive of new ideas, and capable of going beyond, or even challenging, current scholarship in the field. In this article we have described and tested quantitative methods specifically designed to elucidate the early textual transmission of Katherine Philips’s verse. We have used these methods to analyse and clarify the relationships between the extant early witnesses to Philips’s poetry and to draw inferences about missing witnesses and texts. Through our identification of poetic clusters, we have also shed light on the processes through which several of these witnesses were compiled, and have generated new readings of the J. J./Antenor controversy and the Finch poems. We hope that in doing so we have not only modelled new and potentially transferable methods of literary analysis but have also demonstrated their value to Philips research.

NOTES

¹ National Library of Wales MS 775B (Tutin); University of Texas at Austin, Pre-1700 MS 151 (Dering); Worcester College, Oxford, MS 6.13 (Clarke); National Library of Wales MS 776B (Rosania); *Poems. By the Incomparable, Mrs. K. P.* (1664); *Poems By the most deservedly Admired Mrs. Katherine Philips* (1667). Tutin is holograph; Dering, Clarke, and Rosania are scribal manuscripts deriving from Philips’s circle.

² Studies include Peter Beal’s *Catalogue of English Literary Manuscripts* <www.celm-ms.org.uk> (CELM); Elizabeth Hageman, “Making a Good Impression: Early Texts of Poems and Letters by Katherine Philips, the ‘Matchless Orinda,’” *South Central Review* 11.2 (1994): 39-65 and “Traacherous Accidents and the Abominable Printing of Katherine Philips’s 1664 *Poems*,” *New Ways of Looking at Old Texts*, vol. III, ed. W. Speed Hill (Tempe, AZ: Arizona Center for Medieval and Renaissance Studies in conjunction with Renaissance English Text Society, 2004) 85-95; and Gillian Wright, *Producing Women’s Poetry, 1600-1730: Text and Paratext, Manuscript and Print* (Cambridge: Cambridge UP, 2013) 97-145. See also *The Collected Works of Katherine Philips, The Matchless Orinda*, ed. Patrick Thomas, G. Greer, and R. Little, 3 vols (Stump Cross: Stump Cross Books, 1990-3), Vol. 1, hereafter cited as *CW1*.

³ Recent applications of quantitative methods to textual issues include Ruth Ahnert and Sebastian E. Ahnert, “Protestant Letter Networks in the Reign of Mary I: A Quantitative Approach,” *English Literary History* 82.1 (2015): 1-33; Michelle Levy and Mark Perry, “Distantly Reading the Romantic Canon: Quantifying Gender in Current Anthologies,” *Women’s Writing* 22.2 (2015): 132-55; Jonathan Hope and Michael Witmore, “Quantification and the language of later Shakespeare,” *Société Française Shakespeare Actes du Congrès* 31

(2014): 123-49; and Alan B. Farmer and Zachary Lesser, "What Is Print Popularity? A Map of the Elizabethan Book Trade," *The Elizabethan Top Ten: Defining Print Popularity in Early Modern England*, ed. Andy Kesson and Emma Smith (Farnham: Ashgate, 2013) 19-54.

⁴ *CELM*; Wright, *Producing Women's Poetry*.

⁵ Cardiff Central Library, MS 2.1073. The doubtful poems in the Cardiff manuscript (*CWI* nos 126-8) were included in our overall list of poems, but were used only in our QS analysis.

⁶ On the relationships between these early Philips sources, see Hageman, "Tracherous Accidents" and Wright.

⁷ For example, poem 82 is a later addition to the Rosania manuscript. See below.

⁸ Any ambiguities due, for instance, to differing titles in different witnesses, were resolved by checking first lines.

⁹ Thus, the poem known in many Philips witnesses as 'La Grandeur d'Esprit' and in 1667 as 'A Resvery' is here referred to as 'A Reverie'.

¹⁰ All coding was done by one author and then checked by the other.

¹¹ Since witnesses have varying numbers of poems, the witness vectors were padded with zeroes to make them all the same length; this is purely to facilitate analysis. Our method resembles Hope and Witmore (note 3, above) in using quantitative methods as a starting-point for more traditional literary analysis.

¹² For Octave (an open-source equivalent of MATLAB), see <https://www.gnu.org/software/octave/> (accessed 8 Mar. 2016).

¹³ Our QS measure is also called the Sorensen-Dice coefficient; see, for example, Lee R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology* 26 (1945): 297-302. Alternative measures of similarity/difference include Euclidean distance, Hamming distance, Mahalanobis distance, and Jaccard index. We chose Sorensen-Dice because the analogy between species and poems is intuitive and the measure simple to interpret.

¹⁴ Frequently shared poems include 21 found in five witnesses and 52 found in all six witnesses. Five of the 52 are also in Cardiff (Poems 20, 29, 50, 60, 99). See the Appendix for titles and Figure 3 for witness contents.

¹⁵ Dering, Clarke, Rosania, and 1664 have no unique poems. Poem 22 is found only in Tutin; for another witness, beyond those surveyed here, see *CWI* 317. Poem 87 (title-only in Tutin) is provided in full in 1667, which also supplies five stanzas for Poem 57 (two stanzas in Tutin). Poem 95 ("To my Lord Duke of Ormond") also survives in an autograph manuscript, National Library of Wales, NLW MS 21702 E, fols 158r-9v; see *CELM*, PSK 437. Furthermore, 23 other poems are attested only in 1667, and 14 poems are only in Rosania and 1667 (see Figure 3 for details). Almost all of the poems considered here are in 1667; the exceptions are poems 22, 109 (only in Rosania and Dering), and the three doubtful poems in Cardiff (nos 35, 54, 102).

¹⁶ The highest QS values are for Clarke and Dering, and Dering and 1664; yet when comparing these very similar witnesses we note the following differences. Firstly, although all of the poems from Clarke are in 1664 (and 1667), Dering lacks Poem 42 and Rosania Poem 82; the latter was added to the witness by a later scribe. Dering includes all or part of four poems which Clarke lacks: nos 68 and 65 (both title-only), 109 and 41; these are also included in full in Rosania. Poems 65 and 68 are not in 1664.

¹⁷ Hageman, "Tracherous Accidents" 90.

¹⁸ The stub occurs between pp. 160 and 161 of Tutin.

¹⁹ Other stubs that fall into this category include those between pp. 64 and 65, and between pp. 84 and 85.

²⁰ See also Hageman, "Tracherous Accidents" 90.

²¹ Poems 41 (“Ode upon Retirement”), 68 (“The Irish Greyhound”), and 109 (“To the Lady Mary Butler”).

²² *Poems* (1667), sig. a1v.

²³ The three missing poems are 22, 57, and 87 (all discussed elsewhere in this section).

²⁴ Clarke also includes a full text of one poem (no. 107), which is included as title-only in Tutin.

²⁵ University of Kentucky, W. Hugh Peal Collection, Accession No. 8379.

²⁶ Hageman reaches a similar conclusion on literary grounds (“Treacherous Accidents” 88). Compared with traditional methods, however, our QS measure is more systematic, less apt to be influenced by prior expertise, and extends across more witnesses.

²⁷ The Rosania manuscript is generally thought to have been compiled after Philips’s death (to which its preface alludes, 5-7); however, it is also possible that it was begun earlier and completed posthumously.

²⁸ This is to assume the “parsimonious” model described above. It is likely, however, that at least one further manuscript intervened between Tutin and Clarke. See Hageman, “Treacherous Accidents” 93.

²⁹ The influential statistician Jacob Cohen suggested that $r = 0.5$ signalled a large, and $r = 0.3$ a medium effect (J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (New York: Routledge Academic, 1988)). N.B. We obtained correlation coefficients for poem vectors with and without gaps. The results were not materially different. Coefficients cited in the text are for vectors with gaps coded as zeros.

³⁰ Of course, the fact that two witnesses have similar poem ordering does not prove that one was a direct source for the other; for example, both could have relied on a third witness. However, our approach precludes multiplying potential witnesses unnecessarily.

³¹ As suggested by Hageman, “Treacherous Accidents” 91.

³² They include, for instance, Poem 109, “To the Lady Mary Butler”, which dates from October 1662 (when Butler was married), and “The Irish Greyhound”, which was presumably written during Philips’s year in Ireland.

³³ For the last of these gaps, similar conclusions to ours are reached by Hageman (“Treacherous Accidents” 92).

³⁴ Hageman identifies only a single page stub at this point, but notes as problematic the apparent omission from Tutin of “To the Queen of inconstancy” (Poem 111; “Treacherous Accidents” 91-2). Poems 111 and 72 together total 54 lines—too long for a single leaf if Philips was consistently writing on rectos only.

³⁵ Hageman (“Treacherous Accidents” 91) identifies three page stubs at this point. The two poems total 84 lines.

³⁶ “A Sea voyage” is ascribed to September 1652; Philips’s sister-in-law Cicely was married in October 1653, and “Friendship’s Mystery” and “To Mr. Henry Lawes” were both published in Lawes’s *Second Book of Ayres and Dialogues* in 1655. See *CWI* 337, 270, 269.

³⁷ Poem 33 (“Injuria amicitiae”) is addressed only to a “Lovely Apostate”—presumably unnamed for reasons of tact.

³⁸ Both poems are dated by Philips in the Tutin manuscript.

³⁹ See *CWI* 338-9. Poems 77 and 9 (which together comprise C7) may also have been perceived by Philips as having a Welsh focus.

⁴⁰ *CWI* 359-60, 357, 356. The exception, Poem 40 (“Mr. Francis Finch, the excellent Palaemon”) is discussed below.

⁴¹ Hageman, “Making a Good Impression” 43; Wright 128-9.

⁴² This is to exclude poem 22, ignored in all subsequent recensions of Tutin.

⁴³ Katherine Philips, *Letters from Orinda to Poliarchus* (1705) 45-50.

⁴⁴ The near-contemporary Clarke groups Poem 114 with the other royal poems of C17. That Clarke is more up-to-date in terms of order while Dering is more up-to-date in terms of contents suggests a divergence in the textual transmission of Philips's poetry shortly after her arrival in Ireland.

⁴⁵ *CWI* 346.

⁴⁶ Rosania MS 265-7, 235-7.

⁴⁷ The latter possibility may also be supported by the sheer detail with which the titles of these three poems insist on the association—detail unequalled elsewhere in the Rosania manuscript.

⁴⁸ *CWI* 114, line 4; 116, line 54.

⁴⁹ *CWI* 144-5 (lines 49-50).

⁵⁰ Francis Finch, *Friendship* 36 and sig. A3r.

⁵¹ Finch sig. A2r and 1.

⁵² It would also, admittedly, generate new anomalies: most obviously, Finch's claim that *Friendship* was "first penn'd by your [i.e. A. O.'s] Commands" (sig. A2r).

⁵³ "Mr. Francis Finch" and "On the noble Palaemon" are transcribed, respectively, on Tutin 9-13 and 95-7.

⁵⁴ See, for example, Finch's references to a singular "Mistresse" (1) and "Soul" (35-6).

⁵⁵ E.g. *Friendship* 7 and 28.

⁵⁶ Indicative studies include John Burrows, "Andrew Marvell and the 'painter satires': a computational approach to their authorship," *Modern Language Review* 100.2 (2005): 281-97; Hugh Craig, "Stylistic Analysis and Authorship Studies," *A Companion to Digital Humanities*, ed. S. Schreibman, R. Siemens, and J. Unsworth (Oxford: Blackwell, 2004) 271-88; and Michaela Mahlberg, *Corpus Stylistics and Dickens's Fiction* (New York: Routledge, 2013).

Kathleen Taylor trained in neuroscience at the universities of Oxford and Stirling and is currently a research visitor at the Department of Physiology, Anatomy, and Genetics at Oxford. She has published three books on psychology and neuroscience with Oxford University Press. Her fourth book, *The Fragile Brain*, is about the neuroscience of dementia, and is currently in press. *Address*: neurotaylor.com [email: neurotaylor@btinternet.com]

Gillian Wright is senior lecturer in English Literature at the University of Birmingham. Her publications include *Producing Women's Poetry, 1600-1730: Text and Paratext, Manuscript and Print* (Cambridge: Cambridge UP, 2013). She is currently editing Aphra Behn's poetry for Cambridge University Press. *Address*: Department of English Literature, University of Birmingham, Birmingham B12 2TT, United Kingdom. [email: g.wright@bham.ac.uk]

PoemCode	CWI Number	Poem Title
1	71	2 Corinthians 5.19. God was in Christ
2	61	A Country Life
3	80	A Dialogue between Lucasia and Rosania
4	19	A Dialogue of Absence 'twixt Lucasia and Orinda
5	97	A Dialogue of friendship multiplied
6	84	A Farewell to Rosania
7	64	A Friend
8	48	A Prayer
9	22	A retired friendship, to Ardelia
10	60	A Reverie
11	16	A Sea voyage from Tenby to Bristol
12	100	A Triton to Lucasia going to Sea
13	96	Against Love
14	47	Against Pleasure, Set by Dr. Coleman
15	108	An Answer to another persuading a Lady to Marriage
16	82	An Epitaph on my honoured Mother Mrs. Philips
17	3	Arion on a Dolphin
18	18	Content
19	75	Death
20	55	Engraven on Mr. John Collier's Tombstone
21	88	Epitaph on her Son H. P.
22	123	Epitaph on Mr. John Lloyd
23	110	Epitaph on my truly honoured Publius Scipio
24	57	Friendship
25	29	Friendship in Emblem, or the Seal
26	17	Friendship's Mystery
27	74	Happiness
28	30	In memory of F. P.
29	51	In memory of Mr. Cartwright.
30	67	In memory of Mrs E. H.
31	31	In memory of that excellent person Mrs. Mary Lloyd
32	63	In memory of the most justly honoured, Mrs. Owen of Orielton
33	38	Injuria Amicitiae
34	66	Invitation to the Country
35	127	Juliana and Amaranta
36	65	L' Accord du Bien
37	27	Lucasia
38	109	Lucasia and Orinda parting with Pastora and Phillis
39	83	Lucasia, Rosania, and Orinda parting at a Fountain
40	52	Mr. Francis Finch, the excellent Palaemon
41	77	Ode upon Retirement

42	44	On Controversies in Religion
43	68	On Rosania's Apostacy, and Lucasia's friendship
44	89	On the death of my Lord Rich
45	8	On the death of the illustrious Duke of Gloucester
46	10	On the death of the Queen of Bohemia
47	105	On the death of the truly honourable Sir Walter Lloyd Knight
48	4	On the fair weather just at the Coronation
49	94	On the first of January 1657
50	56	On the little Regina Collier
51	2	On the numerous Access of the English
52	11	On the third of September, 1651
53	86	On the Welsh Language
54	128	On Argalus his vindication to Rosania
55	106	Orinda to Lucasia
56	93	Orinda to Lucasia parting October 1661
57	101	Orinda upon little Hector Philips
58	112	Parting with a friend
59	46	Parting with Lucasia, a Song
60	41	Philoclea's parting
61	34	Rosania shadowed whilst Mrs. Mary Aubrey
62	98	Rosania to Lucasia on her Letters
63	37	Rosania's private Marriage
64	81	Song to the Tune of Adieu Phillis
65	79	Song, to the tune of Sommes nous pas trop heureux
66	70	Submission
67	58	The Enquiry
68	78	The Irish Greyhound
69	73	The Soul
70	90	The Virgin
71	72	The World
72	33	To Antenor, on a Paper of mine
73	107	To Celimena
74	9	To her Royal Highness the Duchess of York
75	116	To his Grace Gilbert Lord Archbishop of Canterbury
76	15	To Mr Henry Lawes
77	21	To Mr. Henry Vaughan
78	24	To Mr. J. B. the noble Cratander
79	111	To Mr. Samuel Cooper
80	62	To Mrs. Wogan
81	53	To Mrs. M. A. at parting
82	49	To Mrs. M. A. upon absence
83	50	To Mrs. Mary Aubrey

84	23	To Mrs. Mary Carne
85	99	To my Antenor, March 16 1661/2
86	54	To my dearest Antenor, on his parting
87	92	To my dearest friend Mrs A. Owen upon her greatest loss
88	113	To my dearest friend upon her shunning Grandeur
89	20	To my dearest Sister, Mrs. C. P. on her marriage
90	36	To my excellent Lucasia, on our friendship
91	85	To my Lady Anne Boyle
92	69	To my Lady Elizabeth Boyle
93	95	To my Lady M. Cavendish
94	115	To my Lord and Lady Dungannon
95	103	To my Lord Duke of Ormond
96	43	To my Lucasia
97	59	To my Lucasia, in defence of declared friendship
98	114	To Pastora
99	40	To Philaster
100	39	To Regina Collier
101	42	To Rosania, now Mrs. Montagu
102	126	To Sir Amorous La Foole
103	14	To Sir Edward Dering
104	104	To the Countess of Roscommon
105	87	To the Countess of Thanet
106	25	To the Excellent Mrs. Anne Owen, upon her receiving the name of Lucasia
107	45	To the honoured Lady E. C.
108	102	To the Lady E. Boyle
109	125	To the Lady Mary Butler
110	12	To the noble Palaemon
111	35	To the Queen of inconstancy, Regina Collier
112	76	To the Queen's Majesty, on her late sickness and Recovery
113	6	To the Queen-Mother's Majesty
114	5	To the Queen's Majesty on her arrival at Portsmouth
115	13	To the Right honourable Alice Countess of Carbury
116	32	To the truly competent judge of Honour, Lucasia
117	26	To the truly noble Mrs. Anne Owen, On my first Approaches
118	1	Upon the double Murder of King Charles I
119	91	Upon the graving of her name upon a Tree in Barn-Elms Walks
120	7	Upon the Princess Royal
121	28	Wiston Vault

Figure 1: Postulated links between witnesses

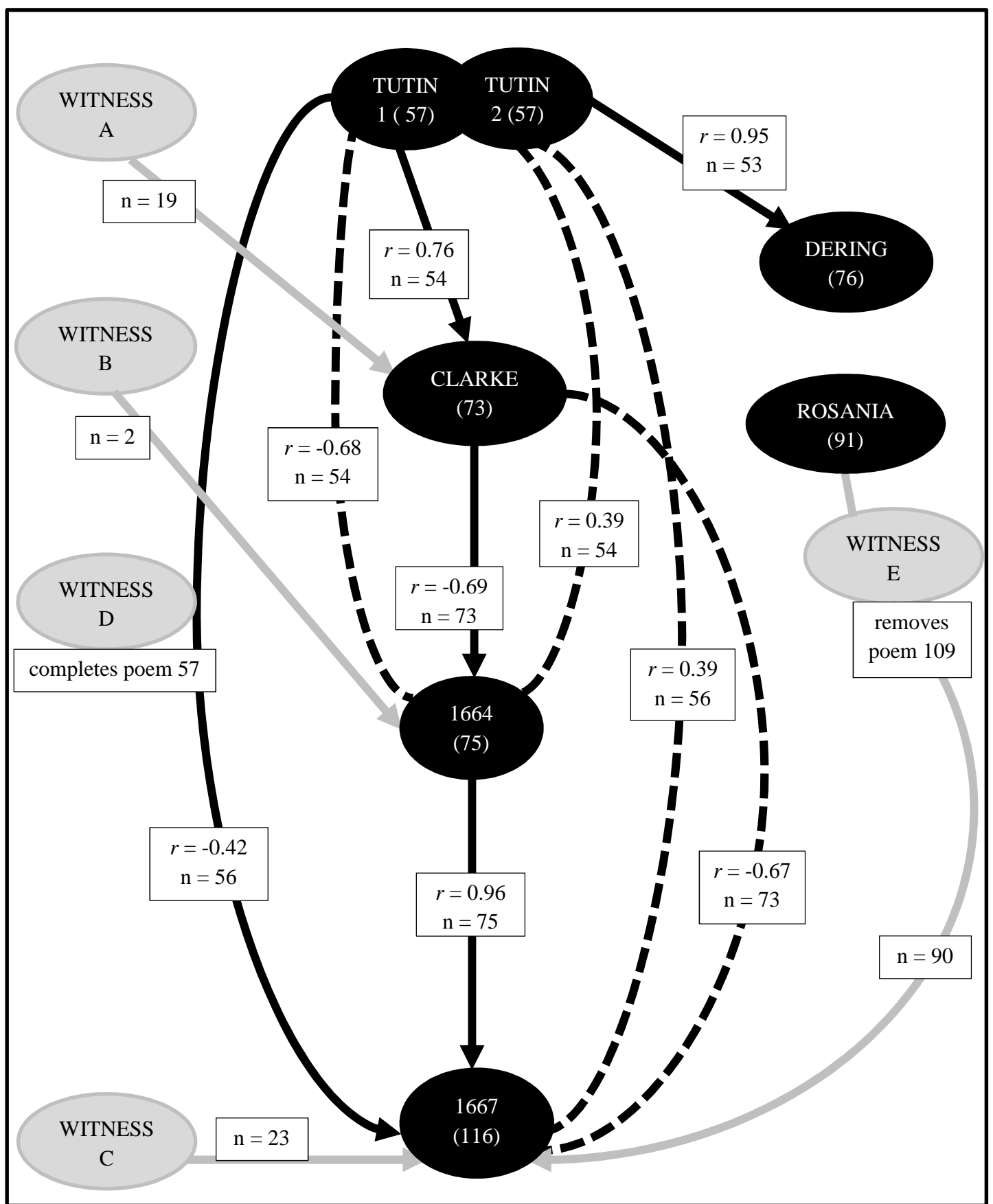


Figure 1 shows potential connections between the six Philips witnesses (black ovals) assessed in this article. Numbers of poems per witness are shown in brackets. Solid lines with arrows indicate proposed contributions from sources to later works. Dashed lines indicate shared poems and significant correlations between witnesses. Each connection is labelled with the correlation coefficient (r , see below) and with the number of shared poems (n). Weak correlations are not shown. Black ovals and lines indicate actual witnesses; grey ovals and lines indicate potential missing witnesses.

Figure 2: Comparing the first and second editions of Philips's poetry, in 1664 and 1667

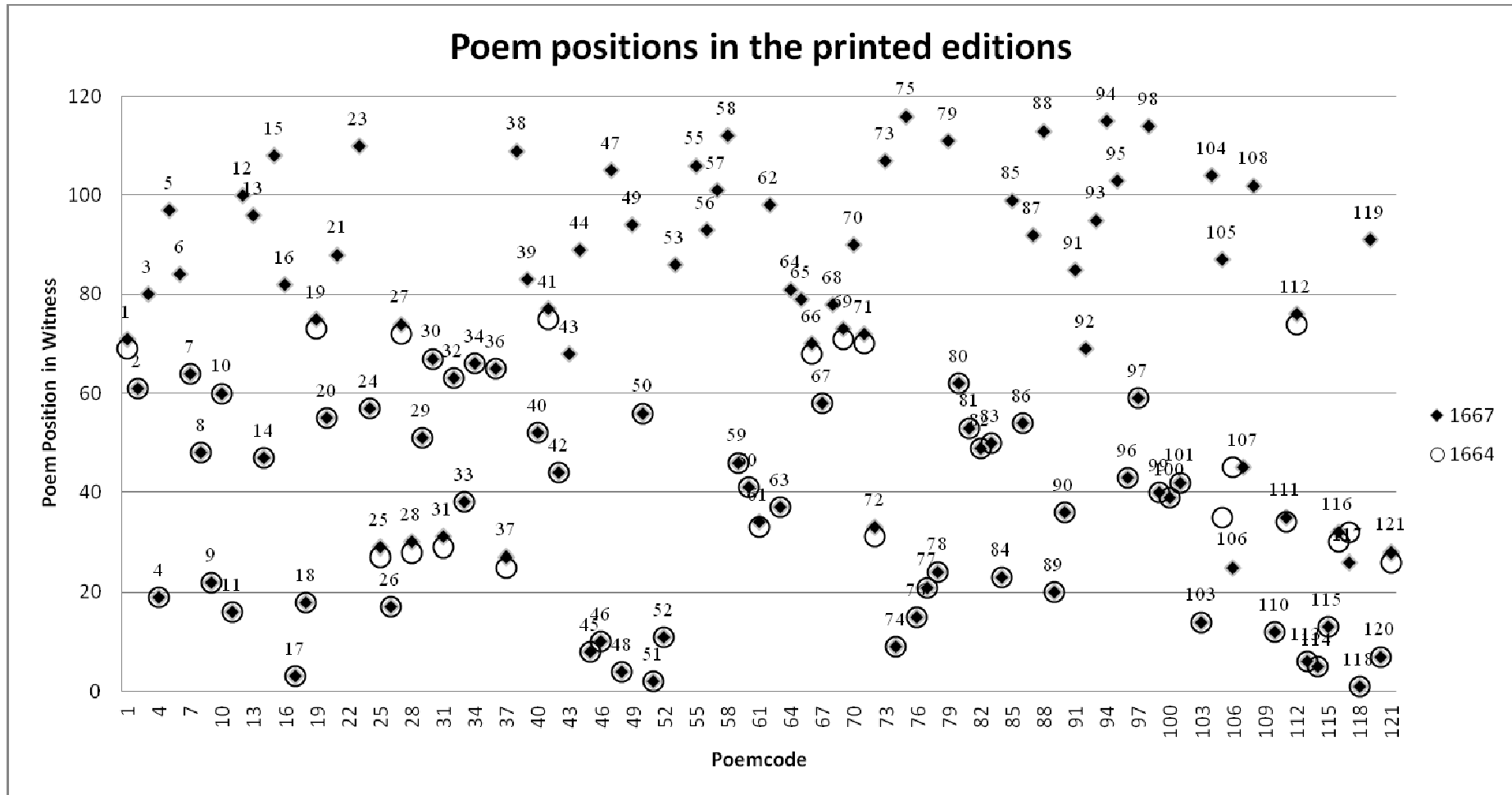


Figure 2 shows the poetic content of the 1664 and 1667 editions (horizontal axis), listed by POEMCODE (i.e. alphabetical order by title). The vertical axis shows the poems' positions within the two witnesses (e.g. Poem 118 (see lower right) is at position 1 for both). Data shown as circles are for the 1664 edition, black diamonds represent 1667 poems. The data are labelled with their POEMCODE. Areas of overlap (circles touching or containing diamonds) represent poems for which the two editions' ordering is closely aligned. The Figure shows that there is considerable overlap for shared poems. These regions of overlap are interspersed with poems new to the second edition.

Figure 3(a). Orders of poems in six Philips witnesses, with clusters colour-coded

Tutin1	Tutin1_FILL	Tutin2	Tutin2_FILL	Clarke	Dering	Dering_FILL	Rosania	1664	1667
8	8	86	86	27	86	86	111	118	118
0	42	20	20	19	20	20	11	51	51
42	27	50	50	71	50	50	116	17	17
27	19	29	29	69	29	29	101	48	48
19	71	40	40	36	83	83	64	114	114
71	69	83	83	34	40	40	107	113	113
69	36	82	82	30	82	82	59	120	120
0	34	0	81	52	81	81	10	45	45
36	30	61	61	7	61	61	14	74	74
34	52	63	63	66	63	63	34	46	46
30	7	101	101	1	101	101	71	52	52
52	66	33	33	80	100	100	69	110	110
7	1	60	60	10	99	99	67	115	115
66	80	99	99	2	33	33	1	103	103
1	10	0	111	32	60	60	86	76	76
80	2	117	72	24	111	111	72	11	11
10	32	106	117	86	72	72	118	26	26
0	24	116	106	20	117	117	2	18	18
32	86	90	116	50	106	106	31	4	4
24	20	84	90	29	90	90	61	89	89
86	50	77	84	40	116	116	84	77	77
20	29	9	77	83	84	84	37	9	9
50	40	103	9	82	77	77	110	84	84
29	83	0	103	81	9	9	3	78	78
40	82	115	118	61	118	118	92	37	106
83	81	78	115	63	103	103	112	121	117
82	61	37	78	101	115	115	103	25	37
0	63	0	37	33	78	78	68	28	121
61	101	26	26	100	37	37	51	31	25
63	33	18	18	99	26	26	17	116	28
101	60	0	4	60	4	4	45	72	31
33	99	0	11	111	18	18	120	117	116
60	111	89	89	72	11	11	113	61	72
99	72	76	76	117	89	89	48	111	61
0	117	0	110	106	76	76	74	106	111
117	106	110	121	116	110	110	46	90	90
106	116	121	25	90	121	121	114	63	63
116	90	25	97	84	25	25	28	33	33
90	84	97	96	77	96	96	109	100	100
84	77	96	57	9	97	97	41	99	99
77	9	0	87	103	67	67	65	60	60
9	103	57	31	118	51	51	83	101	101
103	118	87	59	115	113	113	81	96	96
0	115	31	14	78	17	17	9	42	42
115	78	59	22	37	120	120	33	107	107
78	37	14	107	26	48	48	90	59	59
37	26	22	8	18	45	45	26	14	14

0	18	107	42	4	28	28	18	8	8
26	4	8	27	11	74	74	4	82	82
18	11	0	19	89	46	46	121	83	83
0	89	42	71	76	27	27	97	29	29
0	76	27	69	110	19	19	7	40	40
89	110	19	36	121	71	71	16	81	81
76	121	71	34	25	69	69	39	86	86
0	25	69	30	96	36	36	91	20	20
110	97	0	52	97	34	34	6	50	50
121	96	36	7	67	30	30	53	24	24
25	57	34	66	51	52	52	105	67	67
97	87	30	1	113	7	7	21	97	97
96	31	52	80	17	66	66	44	10	10
0	59	7	10	120	1	1	75	2	2
57	14	66	2	45	80	80	70	80	80
87	22	1	32	48	10	10	119	32	32
31	107	80	24	74	2	2	29	7	7
59		10		46	32	32	40	36	36
14		0		114	24	24	20	34	34
22		32		28	31	31	50	30	30
107		24		31	59	59	100	66	43
				59	14	14	99	1	92
				14	8	8	60	71	66
				8	0	42	78	69	1
				42	0	107	63	27	71
				107	107	114	117	19	69
					114	68	106	112	27
					68	65	25	41	19
					65	109	52		112
					109	41	96		41
					41		24		68
							115		65
							76		3
							89		64
							77		16
							80		39
							32		6
							30		91
							42		53
							8		105
							36		21
							66		44
							27		70
							19		119

Continued see below

1667 also has Poems 87, 56, 49, 93, 13, 5, 62, 85, 12, 57, 108, 95, 104, 47, 55, 73, 15, 38, 23, 79, 58, 88, 98, 94, 75.

Figure 3(b). Clusters identified from Clarke, with contents and possible linking factors

CLUSTER NUMBER	LENGTH	POSSIBLE LINKING FACTOR(s)	POEMS
C1	4	Religious/ philosophical	27, 19, 71, 69
C2	11	Philosophical	36, 34, 30, 52, 7, 66, 1, 80, 10, 2, 32
C4	9	Date/friendship	24, 86, 20, 50, 29, 40, 83, 82, 81
C5	13	Date/friendship	61, 63, 101, 33, 100, 99, 60, 111, 72, 117, 106, 116, 90
C6	1	N/A	84
C7	2	Date/Welsh?	77, 9
C8	1	N/A	103
C9	1	N/A	118
C10	1	N/A	115
C11	2	Bibliographical	78, 37
C12	6	Date	26, 18, 4, 11, 89, 76
C13	1	N/A	110
C14	2	Friendship/memory	121, 25
C15	1	N/A	96
C16	2	Friendship/ philosophical	97, 67
C17	9	Royal	51, 113, 17, 120, 45, 48, 74, 46, 114
C18	2	Elegy	28, 31
C19	3	Bibliographical	59, 14, 8
C20	2	Bibliographical	42, 107
C21	2	Date	112, 41

The columns in Figure 3(a) display the poetic content of our six witnesses, including both versions of Tutin, in the order in which they occur in each witness. Sequences of poems identified as clusters (see text) are colour-coded (see Figure 3(b) for details). For Tutin and Dering, the original layouts are shown, with gaps coded as zeros and outlined in black. Also shown are the layouts constructed from our analysis of the gaps (“_FILL” columns). Figure 3(b) gives the length and our suggested linking factors for each cluster, with constituent poems, colour-coded to match Figure 3(a).