


Cluster Analysis in Tracing Textual Dependencies – a Case of Psalm 6 in 16th-century English Devotional Manuals

Jerzy Wójcik <jwojcik_at_kul_dot_pl>, The John Paul II Catholic University of Lublin  <https://orcid.org/0000-0001-5283-9017>

Abstract

This article uses cluster analysis in order to track textual affinities and identify the sources of different versions of historical texts on the basis text of Psalm 6 found in the 16th-century English manuals of devotion. The article offers a brief overview of the manuals of prayer examined, describes methods of cluster analysis used within the present work, and shows how cluster analysis can enrich and guide traditional philological knowledge.

1. Introduction

The purpose of this contribution is to trace textual affinities between psalm versions contained in several manuals of devotion^[1] when they first started to be printed in English in the 1530s. The rapidly growing popularity of this publication type in the era when new English translations of the Psalms were beginning to emerge additionally invites the question concerning the sources of the psalm versions selected by the publishers of the manuals, while the changing religious views of the three successive Tudor monarchs makes the investigation of the issue within their reigns especially interesting. As an average manual contained about 50 psalms, while new manuals were brought out with increasing frequency and so were new translations of the Bible, it seems that a thorough investigation of the issue would benefit from methods available within digital humanities, which are designed to process algorithmically large amounts of textual data.^[2] The technique I will use in this contribution is that of cluster analysis, which is a relatively novel approach to tracking textual affinities and identifying the sources of different versions of historical texts.^[3] To the best of my knowledge, cluster analysis has not been used in the analysis of textual dependencies between different versions of English biblical texts.^[4] To assess the applicability of this method I applied it to the prose translations of Psalm 6 as found in twenty publications printed between 1530 and 1557. The choice of Psalm 6 as the subject of the present analysis was dictated by the fact that it is one of the so-called Seven Penitential Psalms, which form an invariable part of many manuals of devotion, guaranteeing the availability of its text for comparison. The approach used within the present work proved to independently produce results which agree with philological knowledge and historical facts, and to be able to detect textual affinities that have so far remained unnoticed.

1

2. The analysed texts

The tradition of religious manuals of devotion for laypeople can be traced back to the mid-13th century ([de Hamel 1998]; [Erler 1999 (2008)]; [Duffy 1992 (2002)]; [Duffy 2011]; [Kennedy 2014]), when books which were intended to guide their daily devotions and provide help in participating in the services of the Church, first started to appear.^[5] Initially, these manuals – Books of Hours – contained only Latin material but around the second half of the 14th century exclusively English manuals started to emerge, where both the scriptural material and prayers were offered in translation ([Butterworth 1953]; [Hargreaves 1956]; [Harris-Matthews 1980]; [Kennedy 2014]; [Sutherland 2015]; [Sutherland 2017]; [Charzyńska-Wójcik and Wójcik forthcoming]).^[6] As observed by Eamon Duffy, psalms always

2

constituted an important part of these prayer books, reflecting their special place in Christian devotional life and spirituality [Duffy 1992 (2002), 156]. It needs to be stressed, however, that manuals of prayer contained only a selection of most popular psalms. Among these were the so-called Seven Penitential Psalms, the gradual Psalms^[7] and a slightly varying selection of other psalms. An average Book of Hours contained about 50 different psalms (cf. [de Hamel 1998, 138]; [Morey 2000, 182]). As noted by numerous scholars ([de Hamel 1998]; [Eler 1999 (2008)]; [Duffy 1992 (2002)]; [Kennedy 2014]), Books of Hours were the most popular book of the Middle Ages.

3

The appearance of print in the second half of the 15th century, which made possible the mass-production of books, turned printed Books of Hours into the best-seller of the late medieval and early modern book trade. “Books of Hours were among the first books to be efficiently mass-produced” [Duffy 1992 (2002), 211]. Charles Butterworth notes that up to 1523, printed primers were exclusively in Latin, but soon started to show first non-scriptural vernacular material, and as of 1529 also the psalms in English [Butterworth 1953, 5]. This is by no means surprising, since new English translations of Biblical texts associated with such figures as William Tyndale (the author of a new translation of the New Testament printed in Worms in 1526) or George Joye^[8] (the author of the first printed English Psalter which appeared in 1530 in Antwerp) marked just a beginning of a long succession of English biblical translations, which appeared throughout the 16th century (and afterwards), shaping the history of English Reformation and culminating in the King James Bible of 1611. The sheer number of the 16th-century manuals of devotion and new Psalter translations makes it interesting to try and analyse which Psalter versions were used as sources of psalms appearing in printed manuals of devotion.

The analysis offered here concentrates on Psalm 6 appearing in thirteen manuals printed between 1530 and 1557, i.e. during the reign of Henry VIII (1509-1547), who severed ties with Rome, Edward VI (1547-1553), who transformed the Church of England into an openly Protestant body, and Mary I (1553-1558), who re-introduced Rome-based Catholicism, at least at the level of prescribed practices. The publications selected for analysis include seven manuals printed during Henry VIII's rule, i.e. *Ortulus anime* from 1530^[9] (STC 13828.4), Marshall's primer from 1534 (STC 15986), Godfray's primer from 1535 (STC 15988a), Rouen primer from 1536 (STC 15993), Redman's primer from 1537 (STC 15997)^[10], *Manual of prayers* from 1539 (STC 16009), and Henry VIII's primer from 1545 (STC 16034). Three manuals come from the time of Edward VI's rule, i.e. the *Book of Common Prayer* from 1549^[11] (STC 16270a), Primer from 1552 (STC 16057), and *Book of Common Prayer* from 1552 (STC 16288). Finally, there are three primers printed during Mary I's reign, i.e. Caly's Primer from 1555 (STC 16062), Wayland's Primers from 1555 (STC 16063), and 1557 (STC 16080).

4

The text of Psalm 6 contained in these thirteen manuals was compared with the text of this psalm appearing in seven prose^[12] translations of the Psalter which were in circulation at the time when the manuals were published either as new translations of the Psalter or as part of complete translations of the Bible. These seven prose translations of the Psalms used in the examination were: George Joye's English Psalter translated from the Latin text of Martin Bucer first published in 1530 (STC 2370), George Joye's English Psalter translated from the Latin text of Huldrych Zwingli first published in 1534 (STC 2372), Psalms from Coverdale's first complete Bible issued in 1535 (STC 2063), Psalms from Coverdale's second complete Bible (the Great Bible) first issued in 1539 (STC 2068), Psalms from Richard Taverner's Bible issued in 1539 (STC 2067), 1539 edition of Coverdale's Psalter translated from the Latin of Johannes Campensis (first printed in 1535) (STC 2372.6), and Coverdale's Psalter translated from the Vulgate in 1540 (STC 2368). The list in (1) below provides all the analysed sources of Psalm 6 arranged in chronological order.

5

(1) The list of analysed sources of Psalm 6

6

01 *Ortulus anime* from 1530 (STC 13828.4)

02 George Joye's English Psalter translated from the Latin text of Martin Bucer; first published in 1530 (STC 2370)

03 George Joye's English Psalter translated from the Latin text of Huldrych Zwingli; first published in 1534 (STC 2372)

04 Marshall's primer from 1534 (STC 15986)

- 05 Godfray's primer from 1535 (STC 15988a)
- 06 Psalms from Coverdale's first complete Bible issued in 1535 (STC 2063)
- 07 Rouen primer from 1536 (STC 15993)
- 08 Redman's primer from 1537 (STC 15997)
- 09 *Manual of prayers* from 1539 (STC 16009)
- 10 Psalms from Coverdale's second complete Bible, known as the Great Bible; first issued in 1539 (STC 2068)
- 11 Psalms from Richard Taverner's Bible issued in 1539 (STC 2067)
- 12 1539 edition of Coverdale's Psalter translated from the Latin of Johannes Campensis; first printed in 1535 (STC 2372.6)
- 13 Coverdale's Psalter translated from the Vulgate; issued in 1540 (STC 2368)
- 14 Henry VIII's primer from 1545 (STC 16034)
- 15 *Book of Common Prayer* from 1549 (STC 16270a)
- 16 Primer from 1552 (STC 16057)
- 17 *Book of Common Prayer* from 1552 (STC16288)
- 18 Caly's primer from 1555 (STC 16062)
- 19 Wayland's primer from 1555 (STC 16063)
- 20 Wayland's primer from 1557 (STC 16080)

3. Methodology

The technique which will be used for detecting textual dependencies between the twenty analysed texts of Psalm 6 is that of cluster analysis.^[13] Cluster analysis or simply clustering is a process of partitioning a set of data objects (or observations) into subsets (i.e. clusters) so that objects within a cluster have high similarity, but are very dissimilar to objects in other clusters. Dissimilarities and similarities are assessed based on the attribute values describing the objects and often involve distance measures [Han et al. 2012, 443]. As observed by Hermann Moisl, “[t]he fundamental intuition underlying cluster analysis is that data distributions contain clusters when the data objects can be partitioned into groups on the basis of their relative similarity such that the objects in any group are more similar to one another than they are to objects in other groups, given some definition of similarity” [Moisl 2015, 155].

7

The literature typically divides clustering methods (algorithms) into two major categories, non-hierarchical (partitioning) and hierarchical, depending on the kind of output a given method generates [Moisl 2015, 156].^[14] Han et al. give the following formalisation of a non-hierarchical (partitioning) method: “given a set of n objects, a partitioning method constructs k partitions of the data, where each partition represents a cluster and $k \leq n$. That is, it divides the data into k groups such that each group must contain at least one object” [Han et al. 2012, 448]. An important feature of partitioning methods is that the number of partitions k is a parameter that has to be pre-specified by the user. This method of clustering is non-hierarchical as it performs a one-level partitioning on data sets.

8

In contrast, hierarchical clustering methods perform a grouping of data objects into a hierarchy of clusters. Two major types of hierarchical clustering are distinguished in the literature, i.e. agglomerative hierarchical clustering and divisive hierarchical clustering [Han et al. 2012, 459]. An agglomerative hierarchical clustering method uses a bottom-up strategy and starts by considering each object a cluster of its own and iteratively merges clusters into larger clusters, until all the objects are in a single cluster. A divisive hierarchical clustering method employs a top-down strategy. It starts by placing all objects in one cluster and then divides this initial cluster into several smaller subclusters, and recursively partitions those clusters into smaller ones. The partitioning process continues until each cluster at the lowest level either contains only one object, or the objects within a cluster are sufficiently similar to each other. The result of employing hierarchical clustering algorithms is a tree-based hierarchical representation of the objects, known as a *dendrogram* [Kassambara 2017, 67]. Agglomerative hierarchical clustering is more commonly used than divisive clustering since it is more tractable computationally and it is often the only method available in clustering software packages [Moisl

9

Cluster analysis has been successfully applied in different areas of linguistics as well as other fields ranging from document classification through data mining or speech processing to quantitative stylometry or author attribution [Moisl 2015, 279]. For example, hierarchical clustering and accompanying dendrograms demonstrated their usefulness in detecting relationships between Old English poems [Drout et al. 2011] or, when applied to biblical texts, in examining linguistic variation within Biblical Hebrew [van der Schans et al. 2020]. It should be observed, however, that hierarchical cluster analysis has certain inherent limitations which should be addressed before we embark on the analysis of Psalm 6 data here.

The first limitation concerns the size of the analysed texts. The shortest text analysed here contains 154 words (texts 14, 16), the longest has 264 words (text 12), with an average word count of the twenty text analysed being 174,6 words.

[15] While it is true that applying standard clustering techniques to group short text data creates issues especially with text representation for the purpose of measuring distance [Majid et al. 2022], it has to be stressed that the data analysed here are not a random collection of any short texts. It is known in advance that all the analysed texts of Psalm 6 are quite similar to one another as they ultimately come from the same biblical source text, albeit via different routes. Some of these texts represent translations into English of (different) Latin renditions directly from the Hebrew original, while others are translations of the Latin Vulgate, itself being a translation from Hebrew via a Greek intermediary. By using clustering techniques, we want to find out whether these relatively small textual differences will allow us to indentify the sources of psalm texts as they appear in the 16th-century manuals of devotion. An additional advantage of using clustering on the texts of psalms is that it has a potential of overcoming problems of using an inherently subjective and imprecise language used in the literature to describe textual dependencies and mutual relationship of these psalm texts in terms of “revisions”, “deep revisions”, or calling them “practically new translations”. As observed by Charzyńska-Wójcik and Wójcik [Charzyńska-Wójcik and Wójcik 2022, 213], this inevitably introduces confusion and does not contribute to propelling our knowledge of psalm translations and their revisions.

The second limitation of hierarchical cluster analysis is that it leaves to the user to decide how many clusters the data contain [Moisl 2015, 215], while the structure of the dendrogram representing the division of the data depends on the different cluster joining criteria and can typically generate different trees for the same data.^[16] As noted by Moisl [Moisl 2015, 216], the traditional solution to these limitations of the method is that an expert in the domain from which the data is taken should select the analysis which seems most reasonable in terms of what is known about the research area.

Before applying hierarchical clustering algorithms to the texts of Psalm 6, it is also important to note certain complexities that are inherent to analysing any English text produced before spelling standardisation.^[17] As noted above, the application of a clustering algorithm requires the use of a certain measure of similarity between the classified objects (i.e. different texts of Psalm 6). In the analysis presented in this paper I will employ the cosine similarity^[18] to measure the level of similarity between the compared texts of Psalm 6. Cosine similarity measures the cosine of an angle between vectors, which represent compared texts [Han et al. 2012, 77–78]. A cosine value of 0 means that the two vectors (each representing a text) are at 90 degrees to each other and have no match (the texts are completely different, i.e. they do not share a single item). The closer the cosine value to 1, the smaller the angle and the greater the match (similarity) between vectors (texts). The cosine similarity of 1 means that the compared texts are identical. What transpires from the above is that the compared texts have to be represented as vectors. This is done within a bag of words model [Welbers et al. 2017, 246], where words are assumed to appear independently and the order of words is not taken into account. Each word corresponds to a dimension in the resulting data space and each document then becomes a vector consisting of non-negative values on each dimension [Huang 2008, 50]. This results in the so-called document-term matrix (DTM), in which rows are documents (texts), columns are terms (words), and cells indicate how often each word occurs in each text. The advantage of this representation is that it allows the data to be analysed with vector and matrix algebra, effectively moving from text to numbers [Welbers et al. 2017, 252].

One important consequence of representing texts as vectors in the context of documents produced before standardisation is the need for normalising spelling before any meaningful analysis can be performed. As mentioned

above, the early Modern English texts analysed in this paper were produced before the standardisation of orthography. This means that, for example, the Modern English *anguish* can be found spelled in the following three ways *anguisshe*, *anguyshe*, or *anguysshe*, while *avoid* shows as many as six different spellings *auoid*, *auoide*, *auoyd*, *auoyde*, *avoide*, *avoyde*. It is quite clear that this sort of variation has to be disregarded if meaningful results of similarity measurements are to be obtained. Consequently, in the process of preparing the texts for analysis different spellings of the same word or morpheme were normalised by adopting one consistent spelling across all analysed texts. Moreover, all punctuation was removed and all words spelled with a capital letter were turned to lower case, as these likewise show lack of any consistency in the early Modern period. Normalisation was performed with the use of software called VARD. VARD – from VARiant Detector is a tool ([Baron and Rayson 2008] and [Baron and Rayson 2009]) designed specially to assist research on historical data featuring spelling variation, particularly eMnE texts. Removing punctuation and lowercasing was performed in RStudio with the *quanteda* package [Benoit et al. 2018]. By way of illustration, (2) below presents an original (2a)^[19] and normalised text (2b) of the first lines of Psalm 6 from the first of the texts analysed here, i.e. *Ortulus anime* from 1530 (STC 13828.4).

(2) *Ortulus anime* from 1530 (STC 13828.4)

15

a. original

AH Lorde / rebuke me not in thy wrathe: nether chasten me in thyne anger

But deale fauourably with me (O lorde) for full sore broken am I: heale me (lorde) for my bones are all to shaken.

b. normalised

oh lord rebuke me not in thy wrath neither chasten me in thy anger

but deal favourably with me oh lord for full sore broken am I heal me lord for my bones are all to-shaken.

The grouping of the texts of Psalm 6 into dendrograms was performed using RStudio. The *quanteda* R package was used for creating a document-term matrix and computing cosine similarities between psalms for each pair of psalms, resulting in a 20 x 20 = 400 scores, half of which are redundant, as the similarity between every two texts is the same, whether text A is compared to Text B or the other way around. The R base function *hclust()* was then employed to create a hierarchical tree on the basis of the cosine similarity matrix generated with the *quanteda* package. Since the *hclust()* function uses a distance (dissimilarity) metric between texts, the similarity scores matrix obtained in the previous step was converted into a distance matrix.^[20] Dendrograms representing hierarchical trees were generated with the help of the *factoextra* package [Kassambra and Mundt 2020]. The results of these operations are presented in the next section.

16

4. Results

Table 1 below shows a sample of a document-term matrix and represents a stage in the analysis at which compared texts are turned into vectors. The whole matrix contains 281 columns corresponding to 281 distinct words used in the compared texts of Psalm 6.

17

text	oh	lord	rebuke	me	not	in	thy	wrath	neither	chasten	anger	but	deal
01 <i>Ortulus anime</i> from 1530 (STC 13828.4)	2	8	1	6	1	5	3	1	1	1	2	2	1
02 George Joye's English Psalter from 1530 translated from the Latin of	2	8	1	6	1	5	3	1	1	1	2	2	1

Martin Bucer (STC 2370)													
03 George Joye's English Psalter from 1534 translated from the Latin of Huldrych Zwingli (STC 2372)	1	8	1	7	2	5	3	1	1	0	1	2	0
04 Marshall's primer from 1534 (STC 15986)	2	8	1	6	1	5	3	1	1	1	2	2	1
05 Godfray's primer from 1535 (STC 15988a)	2	8	1	6	1	5	3	1	1	1	2	2	1
06 Psalms from Coverdale's first complete Bible from 1535 (STC 2063)	7	8	1	6	2	5	3	0	0	1	1	1	0
07 Rouen primer from 1536 (STC 15993)	0	8	1	6	1	5	3	0	1	0	1	1	0
08 Redman's primer from 1537 (STC 15997)	0	8	1	6	1	5	3	0	1	1	1	1	0
09 <i>Manual of prayers</i> from 1539 (STC 16009)	0	8	1	6	1	5	3	0	1	1	1	1	0
10 Psalms from Coverdale's Great Bible from 1539 (STC 2068)	5	8	1	7	1	4	3	0	1	1	0	1	0
11 Psalms from Richard Taverner's Bible from 1539 (STC 2067)	2	8	1	6	2	5	3	0	0	1	1	1	0
12 1539	4	6	0	9	2	2	2	1	1	1	0	2	0

edition of Coverdale's Psalter translated from the Latin of Johannes Campensis; first printed in 1535 (STC 2372.6)													
13 Coverdale's Psalter translated from the Vulgate from 1540 (STC 2368)	3	8	1	6	1	5	3	1	1	1	0	1	0
14 Henry VIII's primer from 1545 (STC 16034)	1	8	1	6	1	5	3	0	0	0	1	1	0
15 <i>Book of Common Prayer</i> from 1549 (STC 16270a)	5	8	1	7	1	4	3	0	1	1	0	1	0
16 Primer from 1552 (STC 16057)	2	8	1	6	1	5	3	0	0	0	1	1	0
17 <i>Book of Common Prayer</i> from 1552 (STC16288)	5	8	1	7	1	4	3	0	1	1	0	1	0
18 Caly's primer from 1555 (STC 16062)	0	8	1	6	1	5	3	0	1	0	1	1	0
19 Wayland's primer from 1555 (STC 16063)	0	8	1	6	1	5	3	0	1	1	1	1	0
20 Wayland's primer from 1557 (STC 16080)	0	8	1	6	1	5	3	0	1	0	1	1	0

Table 1.

In the next step, cosine similarity scores for each pair of the analysed twenty texts were computed. The heat map found in the Appendix provides the resulting similarity scores. Finally, the dendrogram representing relations between the analysed texts was obtained, as shown in Figure 1 below.

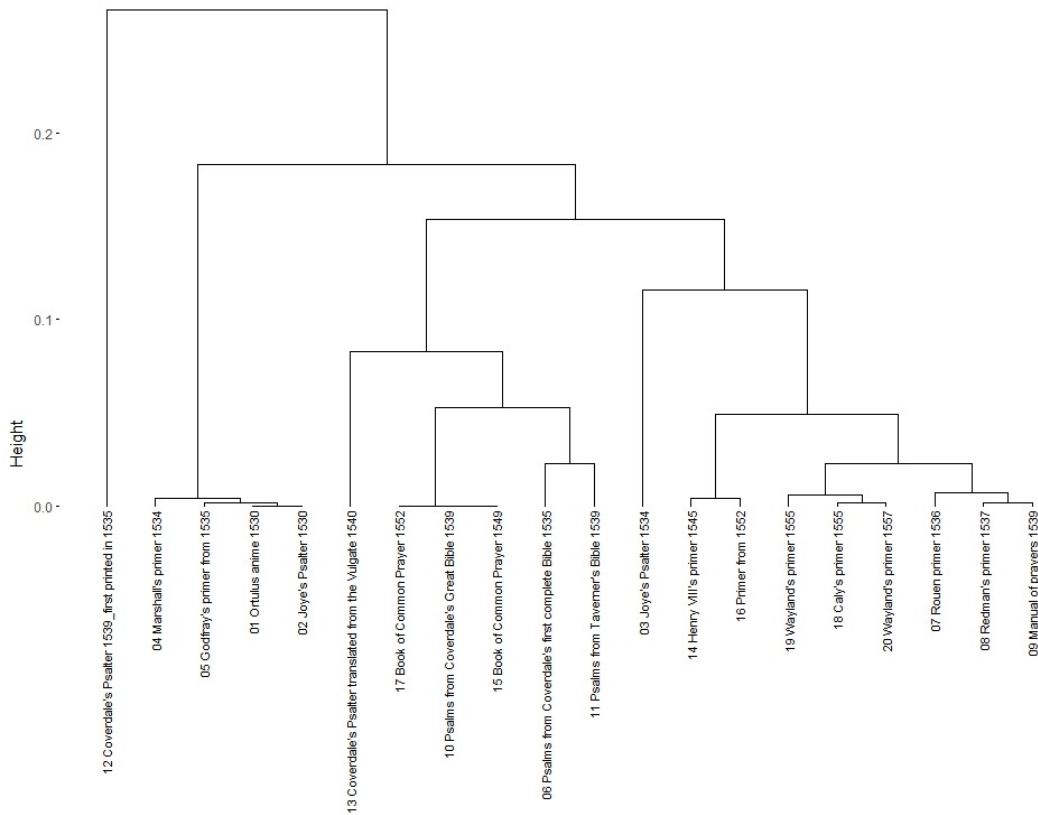


Figure 1. Cluster Dendrogram

The agglomeration method which was used as an argument of the *hclust()* function was that of complete linkage.^[21] In a hierarchical dendrogram like the one above, the distance (difference) between texts (or clustered groups of texts) is represented by the length of the vertical line connecting the grouped objects, i.e. by the height at which a connection in a dendrogram is made. A grouping of objects at height 0, for example, means that a given set contains identical objects whose dissimilarity is 0 (this is the case of texts number 17, 10, and 15, where the vertical length of a line connecting the objects in a group is 0). The bigger the distance (dissimilarity) between the grouped objects, the higher the level at which the vertical line connecting the groupings is found. In the diagram above, for example, text 12 is represented as most unlike all the texts. For one thing, text 12 does not cluster with any other texts and, what is more, it connects to the rest of the texts at the highest level.

The inspection of the data represented in the hierarchical dendrogram above reveals the existence of clear groupings within the texts of Psalm 6 found in the manuals of prayer, their dependence on the existing English translations of the Psalms, as well as their mutual affinities. Quite clearly, there are three groups of texts visible, forming clusters in the dendrogram. The first cluster contains texts 01, 02, 04, and 05. The earliest among the texts in this group is Joye's Psalter from 1530 (02), which appeared in January of 1530, a few months before *Ortilus anime* (01) [Butterworth 1953, 23]. As mentioned earlier, Joye's Psalter from 1530 was the first English printed Psalter, where psalms were translated from the Latin text of Martin Bucer. Later in the same year, Joye published a revised edition of his now lost Primer under the title *Ortilus anime* (01) [Juhász 2014, 21]. Two other texts which cluster with 01 and 02 are Marshall's primer from 1534 (04) (the first book printed in England which contained the English text of Psalms [Butterworth 1953, 50] and Godfray's primer from 1535 (05). The dendrogram clearly identifies the text of Psalm 6 from manuals 01, 04, and 05 as being derived directly from Joye's 1530 translation of the Psalter.

The second group visible in the dendrogram consists of texts 13, 17, 10, 15, 06, and 11. Among these, there are two editions of the *Book of Common Prayer* from 1549 and 1552, three versions of the Bible (10, 06, and 11), and Coverdale's Psalter from 1540 (number 13). Within this group, two editions of the *Book of Common Prayer* from 1549 and 1552 contain the text of Psalm 6, which is identical to that found in the Psalms from Coverdale's Great Bible from 1539 (10). This correctly represents a well-known historical fact: all scriptural material printed in the *Book of Common Prayer* was taken from the Coverdale's 1539 Bible ([Daniell 2003]; [Charzyńska-Wójcik 2021]).^[22] What is more, Psalms from Coverdale's first complete Bible from 1535 (06) are clearly the source of Psalms in Taverner's Bible from 1539 (11), as the grouping in the dendrogram shows their close affinity. This accurately captures what we know about these texts (cf. [Daniell 2003, 193]; [Charzyńska-Wójcik and Wójcik 2022]) since Taverner's 1539 Bible is a revision of Matthews Bible of 1537, which in turn relied on Coverdale's first Bible from 1535 for the Psalms. In contrast, Psalm 6 in Coverdale's 1535 Bible (06) shows more differences with respect to his 1539 Bible (11), also a fact reported in the literature of the topic. Jacobs [Jacobs 2013], Norton [Norton 2000], and Ferguson [Ferguson 2011] call Coverdale's 1539 Bible a (slight) revision of his 1535 Bible. The final text in this group is number 13, i.e. Coverdale's Psalter translated from the Vulgate in 1540, which is least like the other texts in the group, indicating the relations obtaining between the different versions of Coverdale's translations of the Psalms from 1535, 1539, and 1540. As is well-known, Coverdale's 1535 Bible was in Coverdale's own words: "faithfully and truly translated out of Douche [i.e. German] and Latyn [i.e. the Vulgate]". On the other hand, Coverdale's 1540 Psalter was translated "out of the common texte in Latyne", i.e. the Vulgate. The higher level at which the text of Coverdale's 1540 Psalm 6 attaches to the rest of this cluster reflects the differences of the source texts used by Coverdale for the two translations.

The third group of texts identifiable in the dendrogram contains nine texts within which three further subgroupings can be discerned. The first subgroup contains the texts from three primers: (07) Rouen primer from 1536, (08) Redman's primer from 1537, and (09) *Manual of prayers* from 1539. These are quite similar to the next three primers forming a second subgroup, i.e. texts (18) Caly's Primer from 1555, (19) Wayland's Primer from 1555, and (20) Wayland's Primer from 1557, all of which seem to continue the same textual line as the first subgroup. Next comes the third subgroup, comprising (14) Henry VIII's primer from 1545 and (16) a Primer from 1552. Clearly the Primer from 1552 is based on Henry VIII's 1545 primer and both show a relatively high level of reliance on Primers from 1536 and 1537 (07, 08) and a *Manual of prayers* from 1539 (09). The final text within the third cluster is (03) Joye's Psalter from 1534, which forms a branch of its own within this cluster and only attaches to the remaining groupings at a relatively high level. What is interesting in this context is that the text of Joye's Psalter from 1534 (03) is the only translation of the whole Psalter which is found within this cluster comprising the nine texts (in chronological order 03, 07, 08, 09, 14, 16, 18, 19, and 20) and as such is a natural candidate for the source text behind all the text versions found in this cluster. At the same time, the high level at which it attaches in the dendrogram is quite different from what we observed in the case of other clusters, when the source translations (Joye's 1530 translation, or Psalms from Coverdale's Great Bible) showed a high degree of affinity with the texts of psalms from the manuals. This suggests a relation of a different nature between Joye's Psalter from 1534 to the text of Psalm 6 within this cluster. In this context it is interesting to note that Butterworth [Butterworth 1953, 134] claims that the Rouen Primer from 1536 (chronologically the earliest primer in this cluster) introduced a new translation of the Psalms and "cut loose" [Butterworth 1953, 134] from the earlier tradition and based its version on the accompanying Latin Vulgate text in the margin. The dendrogram read in the light of Butterworth's observation indicates that the eight manuals grouped in this cluster all represent the same translation based on the Vulgate. At the same time, the presence of the text of Joye's 1534 translation naturally implies some relationship, albeit not a direct one, of this translation to the text of the Vulgate. This implication, however, stands in stark contrast to what is known about the source of Joye's 1534 translation of the Psalms, which is based on Huldrych Zwingli's new Latin translation ([Butterworth and Chester 1962, 129]; [Juhász 2014, 23]). This contradiction needs to be accounted for as it seems that the clustering algorithm incorrectly grouped the text of Joye's 1534 translation with the texts based on the Vulgate. It turns out, however, that the observed contradiction is only apparent, as transpires from a comparison of the two Latin sources, i.e. the Vulgate and Zwingli's Latin texts of Psalm 6. While these two Latin texts are clearly different they do show some striking similarities as illustrated by the first and last verses of Psalm 6 from these two Latin texts presented in (3).

a. The Vulgate

Domine ne in furore tuo arguas me: neque in ira tua corripas me.

Erubescant, et conturbentur (vehementer) omnes inimici mei: convertantur et erubescant valde velociter.

b. Zwingli (1532)

Domine ne quæso in ira tua arguas me, et in furore tuo ne corripas me.

Erubescant ac turbabuntur uehementer omnes inimici mei, mutabuntur, et erubescant subito

Viewed from this perspective, the presence of (03) Joye's Psalter from 1534 in the cluster is no longer a surprise, as it correctly captures the similarity of Zwingli's Latin to the Vulgate for Psalm 6.^[23]

Finally, as mentioned above, Psalm 6 from the 1539 edition of Coverdale's Psalter translated from the Latin of Johannes Campensis forms a branch of its own and attaches to the rest of the dendrogram at the highest level, clearly indicating the greatest dissimilarity to the remaining texts. This is by no means surprising, as the literature refers to the text of this translation as a paraphrase [Ferguson 2011, 154] so it does not resemble any of the remaining Psalter translations nor does it constitute a source of any of the examined manuals.

5. Conclusion

The study of Psalm 6 contained in the examined publications has shown that cluster analysis can be successfully applied to interpreting textual affinities. First of all, it turned out that the hierarchical clustering algorithm employed here resulted in the groupings which are consistent with historical facts and with the relationships between the analysed texts presented in the specialist literature on this subject. This demonstrates the accuracy of the method relying on strict mathematical criteria for the analysis of textual data. Secondly, the method's inherently objective nature has the capacity to overcome preconceptions in approaching textual analyses and pinpointing unexpected similarities, as demonstrated by the text of Psalm 6 in Joye's 1534 Psalter. As we have seen above, the grouping of Joye's 1534 Psalter with the texts based on the Vulgate has revealed a hitherto unknown affinity of Zwingli's Latin to the Vulgate.^[24] It needs to be emphasised, however, that the results obtained here express textual affinities exclusively with respect to Psalm 6 with no claim for generalisations. To determine textual relations with respect to the other psalms printed in the manuals further research along these lines is needed. Thirdly, the applied method offers reliable results in the form of dendrograms which present relations between texts in a straightforward way providing a wealth of information. Finally, cluster analysis, although performed here on just twenty texts, can be applied to a potentially unlimited number of objects, taking advantage of the developing body of textual data available in the digital format. What transpires from the above is that cluster analysis can assist traditional philological examinations by objectively processing large amounts of data and, in effect, drawing scholarly attention to textual affinities which have so far remained unnoticed.

Acknowledgments

The author would like to thank Prof. Magdalena Charzyńska-Wójcik and Dr. Kinga Lis for their valuable comments on earlier drafts of this paper.

Appendix

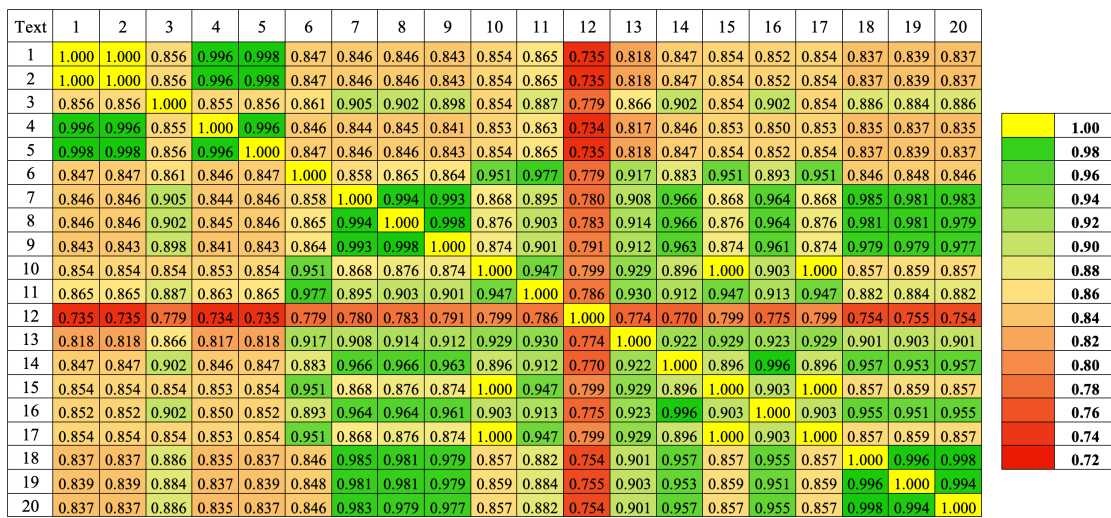


Figure 2. Heat map with cosine similarity scores between twenty texts compared

Sources (chronologically)

Ortulus anime the garden of the soule ... Alternate title: Hortulus animae. English. 1530.; Garden of the soule. Bibliographic name/number: STC (2nd ed.) / 13828.4. Anonymous. [288] p. Antwerp: by me Francis Foxe [i.e. M. de Keyser], 1530.

The Psalter of Dauid in Englishe purely a[n]d faithfully tra[n]slated aftir the texte of Feline: euey Psalme hauynge his argument before, declarynge brefly thentente [and] substance of the wholl Psalme. Alternate title: Bible. O.T. Psalms. English. Joye. Bibliographic name/number: STC (2nd ed.) / 2370. Anonymous. 34, 34-235, [4] leaves. Antwerp: In the year of oure lorde 1530. the. 16. daye of lanuary by me Francis foxe [i.e. Martin de Keyser], 1530.

Dauids Psalter, diligently and faithfully tra[n]slated by George loye, with breif arguments before euey Psalme, declaring the effecte therof. Alternate title: Bible. O.T. Psalms. English. Joye. Bibliographic name/number: STC (2nd ed.) / 2372. Anonymous. 221, [3] leaves. Antwerp: [Maryne Emperowr], 1534.

A prymer in Englyshe with certeyn prayers [et] godly meditations, very necessary for all people that vnderstonde not the Latyne tongue. Cum priuilegio regali. Alternate title: Book of hours (Salisbury); *Ortulus anime.* Bibliographic name/number: STC (2nd ed.) / 15986. Anonymous; Catholic Church. [288] p. London: In Fletestrete by Johan Byddell. Dwellyng next to Flete Brydye at the signe of our Lady of pytye. for Wylllyam Marshall, 1534.

A primer in Englysshe with dyuers prayers & godly meditations. The contentes. ... Cum priuilegio regali. Alternate title: Book of hours.; *Ortulus anime.* Bibliographic name/number: STC (2nd ed.) / 15988a. Anonymous; Church of England. [270] p. London: By Thomas Godfray, 1535.

Biblia the Bible, that is, the holy Scripture of the Olde and New Testament, faithfully and truly translated out of Douche and Latyn in to Englishe. Alternate title: Bible. English. Coverdale. Bibliographic name/number: Darlow & Moule (Rev. 1968), 18; STC (2nd ed.) / 2063. Anonymous. [8], xc, cxx, liij, cij; lxxxi, [1], cxiiij, [1] leaves :. Cologne: Printed by E. Cervicornus and J. Soter?, 1535.

[This prymer in Englyshe and in Laten is newly tra[n]slatyed after the Laten texte.] Alternate title: Book of hours (Salisbury). Bibliographic name/number: STC (2nd ed.) / 15993. Anonymous; Church of England. folios 9-181 [1] p. :. Rouen: [by N. le Roux?], 1536.

[This prymer in Englyshe and in Laten ...] Alternate title: Liturgies. Hours. Salisbury Bibliographic name/number: STC (2nd ed.) / 15997. Anonymous; Church of England. [264] p. :. London: printed by R. Redman, 1537.

This prymer in Englyshe and in Latyn is newly correctyd thys presente yere of our Lorde M.CCCC.XXXVIII. Bibliographic name/number: STC (2nd ed.) / 16008. Anonymous.[272] p. : b ill. s.l.: R. Redman, 1538.

The manual of prayers or the prymer in Englysh & Laten set out at length, whose contentes the reader by y[e] prologe next after the kale[n]der, shal sone perceauce, and there in shall se brefly the order of the whole boke. / Set forth by lhon by Goddes grace, at the Kynges callyng, Byshoppe of Rochester at the comaun demente [sic] of the ryghte honorable lorde Thomas Crumwell, lorde priuie seale, vicegerent to the Kynges hyghnes. Alternate title: Book of hours. Bibliographic name/number: STC (2nd ed.) / 16009. Anonymous; Church of England. [356+] p. :. London: by me John Wayland in saynt Du[n]stones parysh at the signe of the blewe Garland next to the Temple bare, 1539.

The Byble in Englyshe that is to saye the content of all the holy scrypture, both of ye olde and newe testament, truly translated after the veryte of the Hebrue and Greke textes, by ye dylygent studye of dyuerse excellent learned men, expert in the forsayde tonges. Alternate title: Bible. English. Great Bible. Bibliographic name/number: Darlow & Moule (Rev. 1968), 46; STC (2nd ed.) / 2068. Anonymous. [6], lxxxiiij; cxxiiij, [1], cxxvj, cxxxix-cxxxiiij, lxj [i.e. lxxx], ciiij, [1] leaves :. Paris: Prynted by [Francis Regnault, and in London by] Rychard Grafton [and] Edward Whitchurch. Cum priuilegio ad imprimendum solum, 1539.

The most sacred Bible, whiche is the Holy Scripture conteyning the Old and New Testament / translated into English, and newly recognised with great diligence after most faythful exemplars, by Rychard Taverner. Alternate title: Bible. English. Taverner. 1539. Bibliographic name/number: STC (2nd ed.) / 2067. Anonymous. [32], CCXXX [i.e. 460], LXXXXI [i.e. 182], [2], LXXV [i.e. 150], [2], CI [i.e. 190], [5] p. London: Prynted at London in Fletestrete at the sygne of the Sonne by John Byddell, for Thomas Barthlet, 1539.

A paraphrasis vpon all the Psalmes of Daud, made by Iohannes Campensis, reader of the Hebrue lecture in the vniuersite of Louane, and translated out of Latine into Englysshe. Alternate title: Bible. O.T. Psalms. English. Campen. Bibliographic name/number: STC (2nd ed.) / 2372.6. Anonymous. [320] p. London: Prynted in the house of Thomas Gybson, 1539.

The Psalter or boke of Psalmes both in Latyn and Englyshe. wyth a kalender, & a table the more eassyer and lyghtlyer to fynde the psalmes contayned therin. Alternate title: Bible. O.T. Psalms. Latin. Vulgate.; Bible. O.T. Psalms. English. Coverdale. Bibliographic name/number: STC (2nd ed.) / 2368. Anonymous. [8], cxxviii leaves :. London: Ricardus grafton excudebat. Cum priuilegio ad imprimendum solum, 1540.

The primer, set forth by the Kynges maiestie and his clergie, to be taught lerned, [and] read: and none other to be used throughout all his dominions. Alternate title: Book of hours. Bibliographic name/number: STC (2nd ed.) / 16034. Anonymous; Church of England. [308] p. London: VVithin the precinct of the late dissolved house of the gray Friers, by Richard Grafton printer to the Princes grace, 1545.

The booke of the common prayer and administracion of the sacramentes, and other rites and ceremonies of the Church: after the vse of the Church of England. Alternate title: Liturgies. Book of common prayer. Bibliographic name/number: STC (2nd ed.) / 16270a. Anonymous; Church of England. [10], clvii, [1] leaves :. London: in officina Edouardi Whitchurche [and Nicholas Hill] Cum priuilegio ad imprimendum solum, 1549.

The primer, and catechisme, sette furthe by the kynges highnes and his clergie, to be taught, learned, and redde, of all his louing subiectes al other set apart corrected accordyng to the statute, made in the thirde and iiiij. yere, of our souereigne Lordes the kynges maiestie reigne. Bibliographic name/number: STC (2nd ed.) / 16057. Anonymous; Catholic Church. [324] p. London: by Richard Grafton, printer to the Kynges Maiestie, 1552.

The booke of common prayer and adminystracion of the sacramentes, and other rytes and ceremonies in the Church of Englande. Alternate title: Book of common prayer. 1552; Psalter, or psalmes of Daud. Bibliographic name/number: STC (2nd ed.) / 16288. Anonymous; Church of England. [439] p. London: by in officina Edouardi whitchurche [sic], 1552.

[The primer in English and Latin, after Salisburie vse, set out at length with manie praiers and goodly pictures, newly imprinted this present yeare, 1555] Alternate title: Book of hours. Salisbury. Bibliographic name/number: STC (2nd ed.) / 16062. Anonymous; Catholic Church. [372+] p. :. London: In æibus Roberti Caly, 1555.

[The primer in Englyshe (after the vse of Sarum)] Alternate title: Book of hours. Bibliographic name/number: STC (2nd ed.) / 16063. Anonymous; Catholic Church. [371+] p. :. London: J. Wailande, 1555.

The prymer in Englyshe and Latine after Salisbury vse: set out at length wyth many prayers and goodlye pyctures. Alternate title: Book of hours (Salisbury). Bibliographic name/number: STC (2nd ed.) / 16080. Anonymous; Catholic Church. [416] p. :. London: By the assygnes of Ihon Wayland, forbyddyng all other to prynt thys or any other prymer, 1557.

Notes

[1] Manual of devotion is used here as an umbrella term for the variously titled productions (*Ortulus Animae*, *Primer*, *Manual of prayers*, etc.) whose common denominator is that, among many other texts, they offer psalms in English.

[2] The current contribution concentrates only on Psalm 6 in the twenty analysed texts, which is not a particularly large corpus. The method, however, can be applied to other psalms both in the same manuals as well as in other historical texts, subject to the availability of these sources in the digital format.

[3] For a recent example of using clustering in the analysis of authorship attribution of historical texts, see [Tikhonov and Müller 2022].

[4] See [Charzyńska-Wójcik and Wójcik 2022] for an analysis of Psalm 129, using cosine distance text similarity measurements.

[5] This was an off-shot of the Lateran Council of 1215, whose decrees demanded that the laity be offered more guidance in their spiritual needs [Scott-Stokes 2006, 4]. A primer or Book of Hours offered the lay in a simplified form a devotional frame both easy to navigate, in contrast to liturgical books for the religious, and adaptable to the daily routines of their users.

[6] Some researchers use the term *Primer* in contrast to *Book of Hours* to differentiate between vernacular prayer books and Latin manuals. Others use the term interchangeably, a practice I am going to follow here.

[7] The seven Penitential Psalms are psalms 6, 31, 37, 50, 101, 129, and 142 in the Vulgate numbering, and psalms 6, 32, 38, 51, 102, 130, and 143 in the Hebrew numbering, while the gradual Psalms are the fifteen psalms 119-133 (in Hebrew 120-134).

[8] In fact, Joye's first publication was the *Primer*, which was also the first English Primer to be printed. Unfortunately, no copies of this publication have survived. It probably contained the text of more than thirty Psalms [Butterworth and Chester 1962, 52]. For details of Joye's role in English Reformation, see [Juhász 2014]. Joye's Psalter translations are discussed in detail in [Wójcik 2016], [Wójcik 2019] and [Charzyńska-Wójcik and Wójcik forthcoming].

[9] *Ortulus Anime* is a revised 1530 edition of Joye's lost *Primer* from 1529.

[10] The original 1536 edition is in the possession of the Bibliothèque Nationale in Paris. It does not have STC and is not available through EEBO and for this reason I had to rely on a 1537 edition in this study.

[11] The *Book of Common Prayer* represents one of the important products of the English Reformation and was the first prayer book to include the complete forms of service for daily and Sunday worship in English.

[12] The analysed manuals of devotion contained exclusively prose renditions of the psalms. Consequently, it is only natural to look at the new prose translations of the Psalter appearing throughout the 16th century in an attempt to identify the sources of textual material found in these manuals. Observe, however, that the clustering method employed here could just as well be used for analysing textual affinities between prose and metrical (poetic) psalms, since the similarity/distance measure which is a fundamental entity of all clustering algorithms would be able to capture the difference between poetic and prose renditions of the Psalms.

[13] For a detailed discussion of different clustering methods, see [Han et al. 2012, Ch.10 and Ch.11] or [Moisl 2015, 156ff]. Comprehensive accounts of cluster analysis can be found in numerous handbooks on the subject, e.g. [Duda et al. 2001] or [Everitt et al. 2001].

[14] This is by no means the only possible categorization of clustering methods. Various clustering methods may possess features of several overlapping categories. Consequently, the relevant literature provides different classifications of clustering algorithms [Han et al. 2012, 448]. See Han et al. [Han et al. 2012, 448ff] for a discussion.

[15] See Majid et al. [Majid et al. 2022] for a recent discussion and survey of problems associated with short text clustering, as well as ways of overcoming the problems of short text sparseness, dimensionality, and lack of information.

[16] See the discussion in Moisl [Moisl 2015, 216ff], where various cluster-joining criteria (single linkage, complete linkage, centroid linkage, average linkage, or Ward's method) are discussed and compared.

[17] It is generally assumed that English spelling standardisation had not been completed before the end of the 17th century ([Scragg 1974]; [Salmon 1999]; [Görlach 2001]; [Nevalainen 2012]). Clearly, the period covered by this study predates even the beginning of spelling standardisation.

[18] Gomaa and Fahmy [Gomaa and Fahmy 2013] or Wang and Dong [Wang and Dong 2020] offer an overview and comparison of different text similarity measurements. The cosine measure is the most common measures used for computing similarity between documents [Steinbach et al. 2000, 5].

[19] Abbreviations present in the original edition have been silently expanded.

[20] The relationship between cosine similarity and cosine distance is captured by the following formula: Cosine Distance = 1 - Cosine Similarity. That means that a cosine similarity of 1 (identical texts) corresponds to a cosine distance of 0; and conversely, a cosine similarity of 0 (completely different texts) corresponds to a cosine distance of 1.

[21] For a discussion of possible cluster agglomeration methods (also known as linkage methods) which are used at subsequent steps in a tree-building sequence see, for example, Moisl [Moisl 2015, 208ff].

[22] It is interesting to note that the wording of Coverdale's Psalms from the 1539 Bible was retained in all editions of the Book of Common Prayer up to the 1960s while other scriptural material was replaced already in 1662 by the text of the King James Bible of 1611 [Daniell 2003, 488].

[23] For more on Zwingli's Latin translation of the Psalms from Hebrew and his reverence and reliance on the Septuagint, see Potter [Potter 1979, 47](1979, 47). For his competence in Greek, Hebrew and Latin, see Gordon [Gordon 2015, 160] and Pietkiewicz [Pietkiewicz 2020, 333, 338] (also on the authority of [Jones 1983] and [Newman 1925]).

[24] As observed by an anonymous reviewer, the similarity between the English translations based on the Vulgate and Zwingli's Latin, which resulted in Joye's 1534 Psalter being associated with renditions from the Vulgate in the dendrogram, may simply follow from the fact that both are vernacular translations of two Latin versions of the same text. Observe, however, that the same can be stated about other translations from Latin analysed here. Recall, for example, that Joye's 1530 Psalter (02) is a translation from Martin Bucer's Latin text, while Coverdale's 1539 Great Bible (10) is based on the Vulgate and German. Nevertheless, these texts formed separate clusters together with the manuals based on them. In this light, the fact that Joye's 1534 Psalter is grouped with translations based on the Vulgate does reveal an unexpected fact about this translation as regards the text of Psalm 6.

Works Cited

- Baron and Rayson 2008** Baron, A. and P. Rayson. (2008) "VARD 2: A Tool for Dealing with Spelling Variation in Historical Corpora", *Proceedings of the Postgraduate Conference in Corpus Linguistics*, Aston University, Birmingham, UK, 22 May 2008.
- Baron and Rayson 2009** Baron, A. and P. Rayson. (2009) "Automatic Standardization of Texts Containing Spelling Variation: How Much Training Data Do You Need?". In M. Mahlberg, V. González-Díaz, and C. Smith (eds), *Proceedings of the Corpus Linguistics Conference 2009*.
- Benoit et al. 2018** Benoit K, K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, and A. Matsuo. (2018) "quanteda: An R Package for the Quantitative Analysis of Textual Data", *Journal of Open Source Software* 3.30, pp. 774. DOI: 10.21105/joss.00774.
- Butterworth 1953** Butterworth, C. C. (1953) *The English Primers (1529-1545)*. Philadelphia: University of Pennsylvania Press.
- Butterworth and Chester 1962** Butterworth, C. C. and A. G. Chester. (1962) *George Joye (1495?-1553). A Chapter in the History of the English Bible and the English Reformation*. Philadelphia: University of Pennsylvania Press.
- Charzyńska-Wójcik 2014** Charzyńska-Wójcik, M. (2014) "The Secrets of a Sixteenth-Century Psalter: In Praise of Circumstances", *Language and Literary Studies of Warsaw* 4, pp. 135-158.
- Charzyńska-Wójcik 2021** Charzyńska-Wójcik, M. (2021) "Familiarity and Favour: Towards Assessing Psalm Translations", *Linguistica Silesiana* 42, pp. 43-77. DOI: 10.24425/linsi.2021.137231.
- Charzyńska-Wójcik and Wójcik 2022** Charzyńska-Wójcik, M. and J. Wójcik. (2022) "Similarity Measurements in Tracing Textual Affinities: A Study of Psalm 129 in 16th-century Devotional Manuals", *Token: A Journal of English Linguistics* 14, pp. 191-220.
- Charzyńska-Wójcik and Wójcik forthcoming** Charzyńska-Wójcik, M. and J. Wójcik. (In prep.) *Measure for Measure: Comparing Psalm Versions in Medieval English Books of Hours*.
- Daniell 2003** Daniell, D. (2003) *The Bible in English. Its History and Influence*. New Haven, CT and London: Yale University Press.
- Drout et al. 2011** Drout, M. D. C., M. J. Kahn, M. D. LeBlanc, and Ch. Nelson. (2011) "Of Dendrogrammatology: Lexomic Methods for Analyzing Relationships among Old English Poems", *The Journal of English and Germanic Philology* 110(3), pp. 301-336. DOI: <https://doi.org/10.5406/jenglgermphil.110.3.0301>.
- Duda et al. 2001** Duda, R., P. Hart, and D. Stork. (2001) *Pattern Classification, 2nd ed.*, New York: John Wiley and Sons, Inc.
- Duffy 1992 (2002)** Duffy, E. (1992 [2022]) *The Stripping of the Altars: Traditional Religion in England, 1400–1580*. New Haven and London: Yale University Press.
- Duffy 2011** Duffy, E. (2011) *Marking the Hours. English People and their Prayers 1240-1570*. New Haven and London: Yale University Press.

- Erler 1999 (2008)** Erler, M.C. (1999 [2008]) "Devotional Literature." In L. Hellinga and J. B. Trap (eds), *The Cambridge History of the Book in Britain. Vol. 3, 1400-1557*. Cambridge: Cambridge University Press, pp. 495-525.
- Everitt et al. 2001** Everitt, B., S. Landau, and M. Leese. (2001) *Cluster Analysis. 5th ed.* New York: John Wiley and Sons, Inc.
- Ferguson 2011** Ferguson, J. H. (2011) "Miles Coverdale and the Claims of a Paraphrase." In L. P. Austern, K. Boyd, D. McBride and L. Orvis (eds), *Psalms in the Early Modern World*. Ashgate: Farnham, pp. 137-154.
- Gomaa and Fahmy 2013** Gomaa, W. and A. Fahmy. (2013) "A Survey of Text Similarity Approaches", *International Journal of Computer Applications* 68(13), pp. 13-18.
- Gordon 2015** Gordon, B. (2015) "Huldrych Zwingli", *The Expository Times* 126(4), pp. 157-168.
- Görlach 2001** Görlach, M. (2001) *Eighteenth-century English*. Heidelberg: Universitätsverlag C. Winter.
- Han et al. 2012** Han, J., M. Kamber, and J. Pei. (2012) *Data Mining: Concepts and Techniques. 3rd ed.* Waltham, MA: Morgan Kaufmann Publishers.
- Hargreaves 1956** Hargreaves, H. (1956) "The Middle English Primers and the Wycliffite Bible", *The Modern Language Review* 51(2), pp. 215-217.
- Harris-Matthews 1980** Harris-Matthews, J. M. (1980) "Lay Devotions in Late Medieval English Manuscripts." PhD. thesis, University of Cambridge. <https://www.repository.cam.ac.uk/handle/1810/250830>.
- Huang 2008** Huang, A. (2008) "Similarity Measures for Text Document Clustering", *New Zealand Computer Science Research Student Conference (NZCSRSC08)*, pp. 49-56.
- Jacobs 2013** Jacobs, A. (2013) *The Book of Common Prayer. A Biography*. Princeton, NJ: Princeton University Press.
- Jones 1983** Jones, G. L. (1983) *The Discovery of Hebrew in Tudor England: A Third Language*. Manchester and New Hampshire: Manchester University Press.
- Juhász 2014** Juhász, G. M. (2014) *Translating Resurrection. The Debate between William Tyndale and George Joye in Its Historical and Theological Context*. Leiden and Boston: Brill.
- Kassambara 2017** Kassambara A. (2017) "Practical Guide to Cluster Analysis in R." STHDA .
- Kassambara and Mundt 2020** Kassambara A. and F. Mundt. (2020) "factoextra: Extract and Visualize the Results of Multivariate Data Analyses." R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>.
- Kennedy 2014** Kennedy, K. E. (2014) "Reintroducing the English Books of Hours, or 'English Primers'", *Speculum*, 89(3), pp. 693-723.
- Majid et al. 2022** Majid, A., S. Tiun, N. Omar, and N. Sani. (2022) "Short Text Clustering Algorithms, Application and Challenges: A Survey", *Applied Sciences*, 13(1), 342. DOI: <https://doi.org/10.3390/app13010342>.
- Moisl 2015** Moisl, H. (2015) *Cluster Analysis for Corpus Linguistics*. Mouton de Gruyter, Berlin.
- Morey 2000** Morey, J. H. (2000) *Book and Verse. A Guide to Middle English Biblical Literature*. University of Illinois Press, Urbana and Chicago.
- Nevalainen 2012** Nevalainen, T. (2012) "Variable Focusing in English Spelling between 1400 and 1600." In S. Baddeley and A. Voeste (eds), *Orthographies in Early Modern Europe*. Berlin and New York: Mouton de Gruyter, pp. 127-166.
- Newman 1925** Newman, L. I. (1925) *Jewish Influence on Christian Reform Movements*. New York: Columbia University Press.
- Norton 2000** Norton, D. (2000) *The History of the English Bible as Literature*. Cambridge: Cambridge University Press.
- Pietkiewicz 2020** Pietkiewicz, R. (2020). "Ulryk Zwingli jako biblista", *Rocznik Teologiczny*, 62(2), pp. 331-415.
- Potter 1979** Potter, G. R. (1979) "Zwingli and the Book of Psalms", *The Sixteenth Century Journal* 10(2) pp. 42-50.
- R Core Team 2020** R Core Team. (2020) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Salmon 1999** Salmon, V. (1999) "Orthography and Punctuation". In R. Lass (ed), *The Cambridge History of the English Language*, Vol. 3, 1476–1776. Cambridge: Cambridge University Press, pp. 13-55.

- Scott-Stokes 2006** Scott-Stokes, Ch. (2006) *Women's Books of Hours in Medieval England*. Cambridge: Boydell and Brewer.
- Scragg 1974** Scragg, D. G. (1974) *A History of English Spelling*. Manchester: Manchester University Press.
- Steinbach et al. 2000** Steinbach, M., G. Karypis, and K. Vipin. (2000) "A Comparison of Document Clustering Techniques". In *Proceedings of the International KDD Workshop on Text Mining 1999*.
- Sutherland 2015** Sutherland, A. (2015) *English Psalms in the Middle Ages 1300-1450*. Oxford: Oxford University Press.
- Sutherland 2017** Sutherland, A. (2017) "The Wycliffite Psalms." In E. Solopova (ed), *The Wycliffite Bible: Origin, History and Interpretation*. Leiden and Boston: Brill, pp. 183-201.
- Tikhonov and Müller 2022** Tikhonov, A. and K. Müller. (2022) "Scribe versus Authorship Attribution and Clustering in Historic Czech Manuscripts: A Case Study with Visual and Linguistic Features", *Digital Scholarship in the Humanities*, 37(1), pp. 254-263. DOI: <https://doi.org/10.1093/lc/fqaa004>.
- Wang and Dong 2020** Wang, J. and Dong, Y. (2020) "Measurement of Text Similarity: A Survey", *Information* 11(9), p. 421. DOI: <https://doi.org/10.3390/info11090421>.
- Welbers et al. 2017** Welbers, K., W. Van Atteveldt, and K. Benoit. (2017) "Text Analysis in R", *Communication Methods and Measures* 11(4), pp. 245-265.
- Wójcik 2016** Wójcik, J. (2016) "The Orthographical Representation of {-NESS}, {-SHIP}, {-DOM} and {-HOOD} in the First Printed Editions of English Psalms", *Roczniki Humanistyczne* 64(5), pp. 63-78.
- Wójcik 2019** Wójcik, J. (2019) "The First English Printed Psalters – George Joye's Translations and their Editions", *Roczniki Humanistyczne* 67(5), pp. 143-154. DOI: <http://dx.doi.org/10.18290/rh.2019.67.5-8>.
- de Hamel 1998** de Hamel, C. (1998) "Books of Hours 'Imaging' the Word." In J. L. Sharpe III and K. van Kampen (eds), *The Bible as Book: The Manuscript Tradition*. London: British Library and Newcastle, DE: Oak Knoll Press, pp. 137-143.
- van der Schans et al. 2020** van der Schans, Y., D. Ruhe, W. van Peursen, and S. Bhulai (2020) "Clustering Biblical Texts Using Recurrent Neural Networks". DOI: <https://doi.org/10.5281/zenodo.4003509>.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.