Authorship of Pauline Epistles Revisited

Jacques Savoy

Computer Science Department, University of Neuchatel,
Rue Emile Argand 11, 2000 Neuchâtel, Switzerland

Jacques.Savoy@unine.ch

Abstract

The name Paul appears in thirteen epistles, but is he the real author? According to different biblical scholars, the number of letters really attributed to Paul varies from four to thirteen, with a majority agreeing on seven. This paper proposes to revisit this authorship attribution problem by considering two effective methods (Burrows' Delta, Labbé's intertextual distance). Based on these results, a hierarchical clustering is then applied showing that four clusters can be derived, namely {*Colossians*-*Ephesians*}, {1 and 2 *Thessalonians*}, {*Titus*, 1 and 2 *Timothy*}, and {*Romans*, *Galatians*, 1 and 2 *Corinthians*}. Moreover, a verification method based on the impostors' strategy indicates clearly that the group {*Colossians*-*Ephesians*} is written by the same author who seems not to be Paul. The same conclusion can be found for the cluster {*Titus*, 1 and 2 *Timothy*}. The *Letter to Philemon* stays as a singleton, without any close stylistic relationship with the other epistles. Finally, a group of four letters {*Romans*, *Galatians*, 1 and 2 *Corinthians*} is certainly written by the same author (Paul), but the verification protocol also indicates that 2 *Corinthians* is related to 1 *Thessalonians,* rendering a clear and simple interpretation difficult.

Keywords

Authorship attribution; Authorship verification, Text categorization; Machine learning.

Introduction

The *Pauline Epistles* or *Letters of Paul* correspond to fourteen letters attributed to Paul the apostle (Decaux, 2003) and written between around AD 47 (estimated year for the *Epistle to the Galatians*, or in short *Galatians*) to around AD 68 (2 *Timothy*). Belonging to the New Testament, these letters are the oldest Christian writings and represent a classical authorship attribution problem (Love, 2002), (Juola, 2006). According to the Orthodox tradition, these fourteen letters were written by Paul, but the Catholic canon attributed only thirteen to Paul. The last one (*Hebrews*), written anonymously, corresponds to Paul's doctrine but it is generally admitted that the true author is not Paul. With the different Protestantism churches, the number of letters authored by Paul varies from four to thirteen.

Viewed as sacred text, people believe that the authorship of these letters cannot be assessed because they correspond to words of God. Biblical scholars have however debated the authorship of these letters from very early on, as, for example, with Marcion of Sinope (85-160) or St Jerome (347-420) (Love, 2002). Although it is impossible to solve this question with certainty,

this study will provide a new view on this problem by applying two recent and effective authorship attribution methods. To achieve a better view of the results of those automatic attribution methods, a hierarchical clustering is then applied. Finally, this question is also analyzed with a verification protocol determining whether or not a given author (Paul in this study) did in fact write a given book (Koppel *et al*., 2009), (Kocher & Savoy, 2017a), (Koppel & Seidman, 2018).

The rest of this paper is organized as follows. The next section presents the state of the art while Section 3 describes the corpora used in our experiments. Section 4 exposes the results obtained by two attribution approaches used in conjunction with a clustering method. A set of verification tests are discussed in Section 5 and the last section presents the main findings of this study.

State of the Art

Like other text categorization tasks (Sebastiani, 2002), (Manning & Schütze, 2000), an effective authorship attribution model (Juola, 2006), (Zheng *et al*., 2006), (Stamatatos, 2009) must represent each text according to a set of stylistic features reflecting the author's style. To achieve this, a first family of methods suggests defining an invariant stylistic measure (Holmes, 1998) reflecting the particular style of a writer and varying from one author to another. For example, de Morgan (1851) suggests using the word length as a stylistic indicator for determining the authorship of the *Epistle to the Hebrews*. In a similar way, Morton, (1978) analyzed the authorship of the Pauline letters based on sentence length distribution. None of these measures has proven very satisfactory (Holmes, 1998), (Love, 2002), (Baayen, 2008).

As a second paradigm, multivariate analysis can be applied to project each document representative into a reduced space under the assumption that texts written by the same author will appear close together. Some of the main approaches applicable here are principal component analysis (PCA) (Binongo & Smith, 1999), (Craig & Kinney, 2009), hierarchical clustering (Labbé & Labbé, 2001), (Labbé, 2007), (Tuzzi & Cortelazzo, 2018), or discriminant analysis (Jockers & Witten, 2010). As stylistic features, these approaches tend to employ the top 50 to 500 most frequent word-types (MFT), as well as some POS information.

Third, different distance-based measures have been suggested. As well-known strategies, one can mention Burrows' Delta (Burrows, 2002), (Evert *et al*., 2017) using the top *m* most frequent word-tokens (with m = 40 to 1,000), the Kullback-Leibler divergence (Zhao & Zobel, 2007) using a predefined set of 363 English words, or Labbé's method (Labbé, 2007) based on the entire vocabulary and opting for a variant of the Tanimoto distance, an approach found effective for AA (Kocher & Savoy, 2017b).

As an example, the Tanimoto distance between Text A and Text B (denoted D(A, B)) is computed according to Equation 1, where rtf_{iA} denotes the relative frequency of the ith term (for i = 1, 2, …, m). The value m represents the number of selected terms or features.

$$\text{D}(A,B) = \sum_{i=1}^{m} \left| \text{rtf}_{iA} - \text{rtf}_{iB} \right| \Big/ \sum_{i=1}^{m} \text{Max}\left(rtf_{iA},\, rtf_{iB}\right) \quad (1)$$

With the Burrows' Delta model, the relative term frequency rtf_{iA} of each term t_i in the Text A is computed, as well as the mean ($mean_i$), and standard deviation (s_i) of that term over all texts

belonging to the corpus. From these values, a standardised frequency is computer as shown in Equation 2.

$$\text{Z score}(t_{iA}) = (rfr_{iA} - mean_i) / s_i \quad (2)$$

Then, the distance between Text A and B (denoted $\Delta(A,B)$) is computed according to Equation 3. When, for all terms, the Z score values are very similar, the distance between the two texts is small, indicating that both texts might have been authored by the same writer.

$$\Delta(A,B) = {}^{1}/_{m} \cdot \sum_{i=1}^{m} \left| \text{Z score}(t_{iA}) - \text{Z score}(t_{iB}) \right| \quad (3)$$

As a fourth family of methods, various machine learning approaches have been suggested (Stamatatos, 2009), (Jockers & Witten, 2010) as, for example, decision trees, back-propagation neural networks, *k*-NN, random forests, and support vector machines (SVM), the latter being a popular approach in various CLEF campaigns (Stamatatos *et al*., 2015). Zheng *et al*. (2006) found that SVM and neural networks tended to produce similar performance levels that are significantly better than those achieved by decision trees. The *k*-NN approach tended to produce better effectiveness than both the naïve Bayes or decision tree (Zhao & Zobel, 2007). Jockers & Witten (2010) showed that the Delta scheme could surpass performance levels achieved by the SVM method.

If words seem a natural way to generate a text surrogate, other studies have suggested using the letter occurrence frequencies (Kjell, 1994) or the distribution of short sequences of letters (character *n*-grams) (Juola, 2006) (Stamatatos *et al*., 2015). As demonstrated by Kešelj *et al*. (2003), such a representation can produce high performance levels. Kocher & Savoy (2018) tend to confirm this findings with *n* values between 5 to 7. Mikros & Perifanos (2013) suggest considering a combining word-based and letter *n*-gram representations, for example, based on a mixed of *m* most frequent letter 2-grams and 3-grams together with the *m* most frequent words. When adopting such a strategy, the final decision is more difficult to explain to the user (e.g., what is the stylistic meaning of a frequent use of "ui"?) and the best value for *n* must be determined empirically. In addition, the fingerprint of an author can also be identified by the POS tags distribution or short sequences of such tags (Kocher & Savoy, 2018). Such text representations do not usually produce the best performance levels but can be used as useful complementary information (Zheng *et al*., 2006), (Juola, 2006).

Finally, to solve the verification question (determining whether or not a given author did in fact write a given text), some modifications of these strategies must be done. In this context, the training sample contains texts written by a single author who might also be the writer of the query document. To achieve this, the disputed text (denoted Q and assumed to be written by A) can be processed as a whole or as a sequence of *c* chunks (e.g., each composed of 500 word tokens). The result obtained by these *c* subparts of Q determine the final answer (Koppel *et al*., 2007). As a variant, a set of other possible writers called impostors (with a text sample for each of them) can be included. A set of binary classifiers is trained to learn models for A vs. not-A, B vs. not-B, etc. The *c* chunks of the doubtful text are then classified according to the learned models, and, if a preponderance of chunks is classified as A, then we conclude that A is the real author (Koppel & Winter, 2014). Kocher & Savoy (2018) proposed to compute the distance between Q and the impostors using the *m* most frequent word-tokens based only on Q. After a few iterations, if the intertextual distance with A is small compared to the others, the system suggests that the real author is A. Finally, Koppel & Seidman, (2018) suggest to iterate by

selecting only a fraction of the entire *m* features. The final decision depends on the similarity between Q and texts written by A or by the impostors.

Evaluation Corpus

As apostle Paul is a member of the Hellenistic community, the language of his letters is Ancient Greek prose. In this language, our corpus contains in total twenty-one epistles, when adding the letters attributed to James, Peter, John, and Jude. This collection has been downloaded from the website *Bibelwissenschaft.de*. As a second version, an English version (the *King James* edition) was also extracted from the same website. Different information about these letters are reported in Table 1 (given in the order specified in the New Testament). In the last two columns, the size, in words, is provided for both the Ancient Greek and English versions[1].

As indicated in Table 1, the letters contain less than 10,000 words (with two exceptions for the English version). With a mean length clearly below 5,000 words (Greek: 2,480.8; English: 3,332.3), a reliable authorship attribution is rather difficult to achieve (Eder, 2015), (Savoy, 2018). The texts are however of high quality regarding the spelling and additional elements (e.g. verse numbers, running titles) have been removed.

#	Title	Author	Size (Greek)	Size (English)
1	Romans	Paul	8,233	10,952
2	1 Corinthians	Paul	8,048	11,020
3	2 Corinthians	Paul	5,245	7,059
4	Galatians	Paul	2,617	3,596
5	Ephesians	Paul	2,741	3,466
6	Philippians	Paul	1,884	2,532
7	Colossians	Paul	1,801	2,305
8	1 Thessalonians	Paul	1,696	2,125
9	2 Thessalonians	Paul	938	1,170
10	1 Timothy	Paul?	1,866	2,660
11	2 Timothy	Paul?	1,441	1,947
12	Titus	Paul?	779	1,066
13	Philemon	Paul	388	505
14	Hebrews	Paul?	5,739	7,891
15	James	James	2,024	2,707
16	1 Peter	Peter	1,914	2,883
17	2 Peter	Peter	1,213	1,767
18	1 John	John	2,476	2,915
19	2 John	John	283	352
20	3 John	John	256	346
21	Jude	Jude	515	714

Table 1: List of the twenty-one epistles of the New Testament

Concerning the authorship, in the first thirteen letters, one can read the name of the author as Paul in the first or second verse (e.g. “Paul, a servant of Jesus Christ, called to be an apostle”). Sometimes a second name could appear (e.g., “Paul, a prisoner of Jesus Christ, and Timothy

[1] The website *Bibelwissenschaft.de* gives only few information about the translation process of the Pauline epistles.

our brother, unto Philemon our dearly beloved"). From this list, only *Hebrews* and *1 John* appear anonymous.

In the writings attributed to Paul, the *Hebrews* and the pastoral epistles[2] (1 and 2 *Timothy*, and *Titus* (addressed to an individual rather than a community)) are the most disputed letters (indicated with a "?" in Table 1). Ignoring these four letters, one can assume that the remaining ten epistles can be attributed to Paul. This solution corresponds to our first hypothesis denoted *Ten Letters* (*Romans*, 1 and 2 *Corinthians*, *Galatians*, *Ephesians*, *Philippians*, *Colossians*, 1 and 2 *Thessalonians*, and *Philemon*).

Some German biblical scholars in the 19th century (Morton, 1978) have proposed another view. For them (e.g., (Baur, 1845)), only four letters (*Romans*, *Galatians*, 1 and 2 *Corinthians*) were written or dictated by Paul (Love, 2002). This second hypothesis is called *Four Letters*. Thus, the number of genuine letters varies from four to thirteen (or fourteen), with seven being the choice of numerous scholars (the *Seven Letters* hypothesis corresponds to *Romans*, 1 and 2 *Corinthians*, *Galatians*, *Philippians*, 1 *Thessalonians*, and *Philemon*) (Wall, 2002), (Aune, 2010). To visualize what scholars think about the Pauline authorship, one can see McGrath's chart[3].

This authorship question has been debated for centuries, grounded on both internal evidence (comments within the letters, vocabulary or rhetoric analysis) and external evidence (references with other texts together with the historical context) (Love, 2002). The controversy persists with opposite views, as for example:

> "It seems to me that much of the evidence regarding authorship of the Pastorals is sufficiently ambiguous that the issue cannot be decided." (Porter, 1995, p. 121).

As pre-processing, it is useful to remove the first verses corresponding to general greetings including an identification of the author and recipients, as, for example, the first two verses of the *Letter to the Colossians* "1: Paul, an apostle of Jesus Christ by the will of God, and Timotheus our brother, 2: To the saints and faithful brethren in Christ which are at Colosse: Grace be unto you, and peace, from God our Father and the Lord Jesus Christ.…". Some examples of translational variations into English are reported in Table A.1 in the Appendix.

As for other languages, the most frequent tokens in Ancient Greek correspond to determiners (e.g., ὁ, οἱ, ἡ, τὸ (the)), prepositions (πρὸς (to)), conjunctions (καὶ (and, the most frequent word in our corpus)), pronouns (ἐγὼ (I), μέ (me), σύ (you), αὐτοῦ (it, him)) or modal verb forms (ἐστι, ἐστιν (is)). As Ancient Greek has three genders (masculine, feminine, neutral), three numbers (singular, dual, and plural), and four grammatical cases (nominative, accusative, genitive, and dative), the morphology can be viewed as more complex compared to English. Thus, there is no single translation of the definite determiner *the* but twenty-three possible words (e.g., ὁ (masc. sing. nom.), οἱ (masc. plur. nom.), ἡ (femi. sing. nom.), τὸ (neut. sing. nom.), etc.). Finally, the word order is relatively free compared to English.

---

[2] *Letter to Philemon* is sometimes considered as the last pastoral epistle.

[3] Available at http://coolingtwilight.com/pauline-authorship-survey-chart/.

Authorship Attribution and Clustering Experiments

To visualize the distances between all twenty-one epistles, Labbé's intertextual distance (Labbé & Labbé, 2001), (Labbé, 2007) was applied based on all tokens having an occurrence frequency larger than two. Removing word having an absolute frequency smaller or equal to two reduces the vocabulary size of around 50%. This practice can be viewed as a feature selection procedure. Moreover, those infrequent terms tend to be marginal in describing an author's style.

More precisely, the distance between Text A and Text B is computed according to Equation 4 where n_A indicates the length of Text A (in number of tokens), and tf_{iA} denotes the absolute frequency of the ith term (for $i = 1, 2, \ldots, m$). The value m represents the vocabulary length. It is rare that both texts have the same length, so let us assume that Text B is the longer. To reduce the longer text to the size of the smaller, each of the term frequencies (in our case tf_{iB}) is multiplied by the ratio of the two text lengths, as indicated in the second part of Equation 4.

$$D(A,B) = \sum_{i=1}^{m} \left| tf_{iA} - \widehat{tf_{iB}} \right| / (2 \cdot n_A) \qquad \text{with } \widehat{tf_{iB}} = tf_{iB} \cdot n_A / n_B \tag{4}$$

This intertextual distance returns a value between 0 and 1 depending on the lexical overlap between two texts. When two texts are identical, the distance is 0. The largest distance of 1 would appear when the two books have nothing in common (e.g., one is in Ancient Greek and the other in English). Between these two limits, the distance value depends on the number of terms appearing in both novels, and their occurrence frequencies.

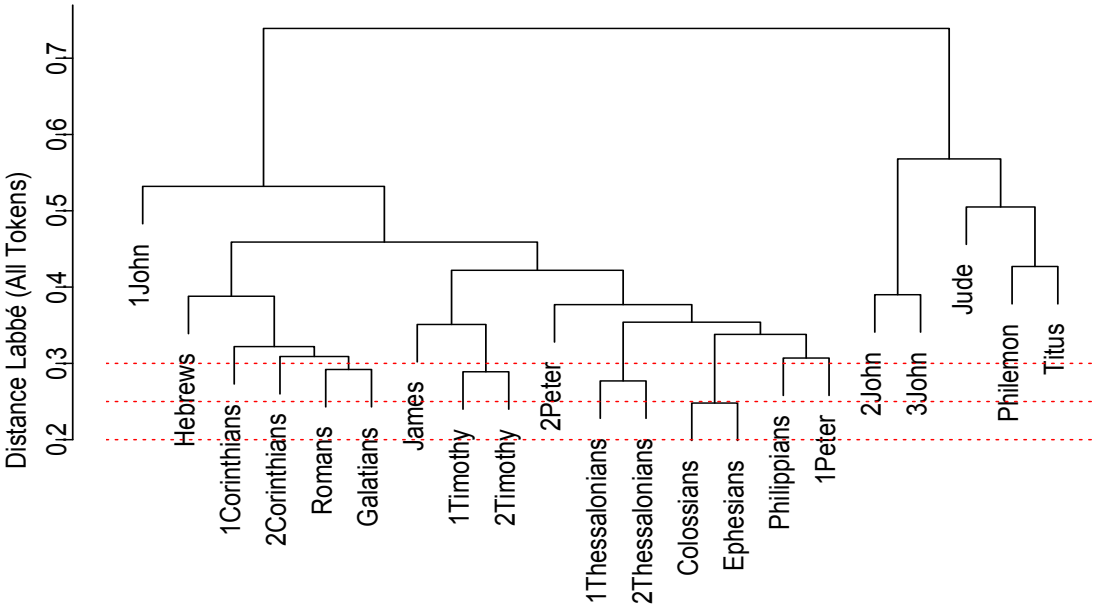

Figure 1: The twenty-one epistles regrouped using all tokens
(Complete link, Labbé distance, Greek version)

Instead of reporting the resulting symmetric matrix (21 x 21 values), a hierarchical clustering was applied. Figure 1 shows the outcome achieved with the complete link method (Greek version). Adopting the complete link to merge two clusters implies that all members of the first group must be similar to all members of the second cluster.

Moving from bottom to top, the distance between clusters increases. On the bottom part, four homogenous clusters can be found with a distance limit of 0.3. First, with the smallest distance (0.248) between all pairs, the *Colossians* and *Ephesians* depict a very similar style indicating that they might have been written by the same author. The second group with a similar style is formed by 1 and 2 *Thessalonians* (distance 0.277). The third cluster is composed by 1 and 2 *Timothy* (distance 0.289). Finally, and close to a distance of 0.3, one can see the cluster corresponding to the *Four Letters* hypothesis (namely *Romans*, 1 and 2 *Corinthians*, and *Galatians*) with a small distance (0.292) between *Romans* and *Galatians* (or 0.293 between *Romans* and 2 *Corinthians*, and 0.308 between *Romans* and 1 *Corinthians*).

With a higher distance (or a smaller similarity), a cluster can be generated with *Philippians* and *1 Peter* (0.308). As the distance increases, the certainty that the same author wrote both texts decreases. Considering clusters with a higher value than 0.3 is problematic. Defining such a threshold is always partially subjective. However, with Labbé's intertextual distance, some calibration experiments have been performed (Labbé & Labbé, 2001) indicating that values smaller than 0.2 indicate that the two texts (same text genre, similar topics) have been written by the same author. Between 0.2 and 0.3 the two texts might have been written by the same author, usually those texts have different text genres, or a large temporal gap between them. These limits are given for texts having more than 10,000 words (which is not the same in this study). Thus, we suggest to relax this limit up to 0.3, a value also used in (Labbé, 2007).

Overall, this picture could be interpreted as evidence in favor of Baur's (1845) thesis specifying that Paul wrote himself only four letters (*Romans*, 1 and 2 *Colossians*, and *Galatians*). Moreover, Figure 1 does not provide support for the *Seven Letters* nor the *Ten Letters* hypothesis. Paul does not seem to be the real author of seven letters (the previous four, plus *Philippians*, 1 *Thessalonians*, and *Philemon*, the last three appearing in different clusters).

Table 2 Ranked list of Labbé's distances between two letters (All tokens, Greek version)

Rank	Dist.	Letter	Letter
1	0.2484	*Colossians*	*Ephesians*
2	0.2771	*1 Thessalonians*	*2 Thessalonians*
3	0.2890	*1 Timothy*	*2 Timothy*
4	0.2915	*Romans*	*Galatians*
5	0.2931	*Romans*	*2 Corinthians*
6	0.2999	*Philippians*	*Colossians*
7	0.3075	*Philippians*	*1 Peter*
8	0.3079	*Romans*	*1 Corinthians*
9	0.3091	*2 Corinthians*	*Galatians*
10	0.3127	*Philippians*	*2 Timothy*
…	…	...	...

In addition, Table 2 reports the ten smallest values achieved between letter pairs when using Labbé's distance function. As one can see, the four clusters are represented by seven pairs over ten (the remaining three are related to *Philippians*). The same information obtained with the English version is depicted in the Appendix. Comparing the two ranked lists, the strong relationship between 1 and 2 *Timothy* (3rd rank in Table 2) does not appear in the top ten in the English version.

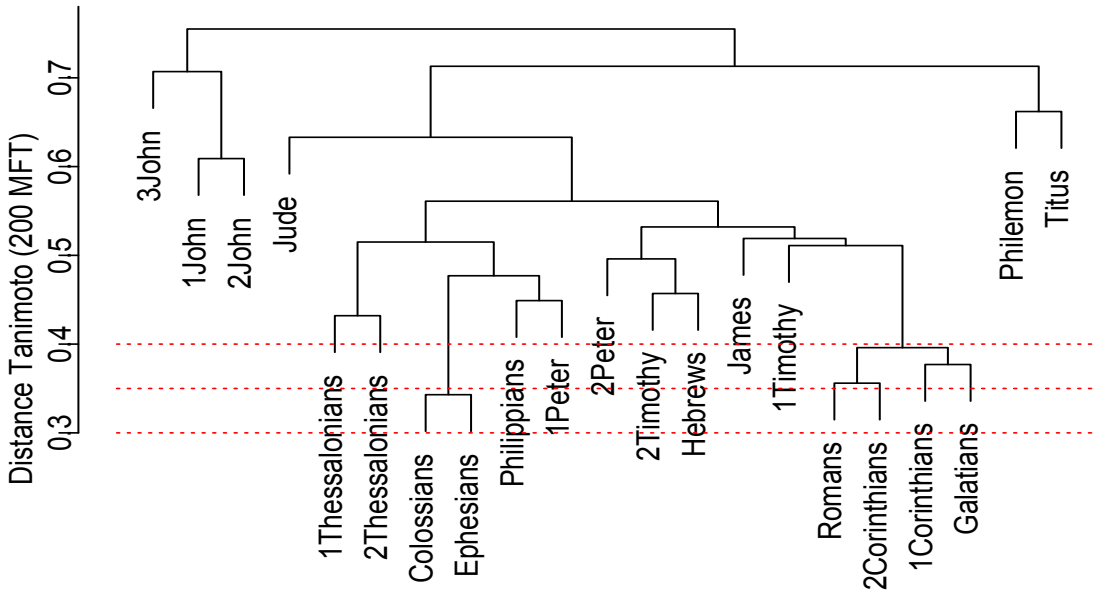

Figure 2: The twenty-one epistles regrouped using the 200 MFT (complete link, Tanimoto distance, Greek version)

To confirm these findings, Figure 2 depicts the stylistic relationships between the twenty-one letters that can be obtained using the 200 most frequent tokens (MFT) (Greek version). Such a text representation puts emphasis on the frequent style markers of each text. This limit of 200 seems subjective. A recent study (Savoy, 2015) shows however that considering between 200 to 500 most frequent terms tends to produce the highest performance levels. Moreover, in the current case, some of the epistles are rather short (and some have less than 200 distinct words, e.g., *Philemon*, *2 John*). Finally, the distance between two texts is computed according to the Tanimoto distance (see Equation 1), a function found effective for authorship attribution (Kocher & Savoy, 2017b).

With a distance limit of 0.4, Figure 2 confirms the presence of the two main clusters, namely the *Colossians-Ephesians*, and the *Four Letters* (*Romans*, 1 and 2 *Corinthians*, *Galatians*). This last group appears more distinctively than in Figure 1. The smallest distance (0.343) can be found between *Colossians* and *Ephesians*, and the second smallest (0.356) occurs between *Romans* and 2 *Corinthians*. Again, this figure supports Baur's hypothesis but not the *Seven* or *Ten Letters* hypothesis.

Moreover, the cluster 1 and 2 *Thessalonians* is generated with a higher distance (it was the second smallest distance in Figure 1), while the cluster with 1 and 2 *Timothy* does not appear. Again, *Philemon* appears far away from the other six letters attributed to Paul according to the *Seven Letters* hypothesis.

Table 3 Ranked lists produced by the Delta model (200 MFT, Greek version)

Rank	Distance	Letter	Distance	Letter	Distance	Letter	Distance	Letter
		Romans		*1 Corinthians*		*1 Thessalonians*		*Philemon*
1	0.744	*2 Corinthians*	0.715	*Romans*	0.808	*2 Thessaloni-*	1.137	*2 Timothy*
2	0.754	*Galatians*	0.735	*2 Corinthi-*	0.847	*2 Corinthians*	1.173	*Philippians*
3	0.775	*Ephesians*	0.811	*Galatians*	0.885	*1 Peter*	1.258	*Ephesians*

Based on the same feature set (200 MFT), Burrows' Delta (Burrows, 2002) has then been employed to compute the distance between letter pairs. Table 3 reports the three closest letters for four selected texts. The first two examples (*Romans*, *1 Corinthians*) show the strong similarity between the *Four Letters*. If we accept the hypothesis that Paul is the true author of the *Seven Letters*, the last two examples in Table 3 do not clearly corroborate this assumption. Both present a stylistic similarity with letters not written by Paul (e.g., *1 Peter*) or with letters usually not attributed to Paul (*2 Thesalonians*, *2 Timothy*). For the last example, the Delta distances are clearly higher and close to twice the values reported in the first two examples, indicating a low certainty that those texts are authored by the same person. The same outcome achieved with the English version is available in the Appendix (see Table A.3).

Using the Delta distance, Figure 3 reports the resulting dendrogram (complete link) when each letter is represented by the 200 most frequent tokens (MFT). This figure confirms the presence of the four main clusters found in Figure 1.

The clustering and the corresponding ranked lists achieved by Labbé's or Delta models tend to favor the *Four Letters* hypothesis. Moreover, we can, with some certainty, assume that the same author wrote 1 and 2 *Thessalonians*, and the same writer can be detected behind both *Colossians* and *Ephesians*. These three clusters could have been written by a unique author, or by two or three distinct persons. In addition, Figure 3 does not corroborate the *Seven Letters* hypothesis proposing that Paul is the author of the *Four Letters* plus 1 *Thessalonians*, *Philippians* and *Philemon*. Our study indicates that 1 *Thessalonians* is related to the style of 2 *Thessalonians*, while *Philippians* is loosely connected to clusters *Colossians-Ephesians* and 1 and 2 *Thessalonians* corresponding to letters usually not attributed to Paul. For *Philemon*, the answer is less clear, mainly because this text is rather short (388 words), rendering a reliable assignment difficult.

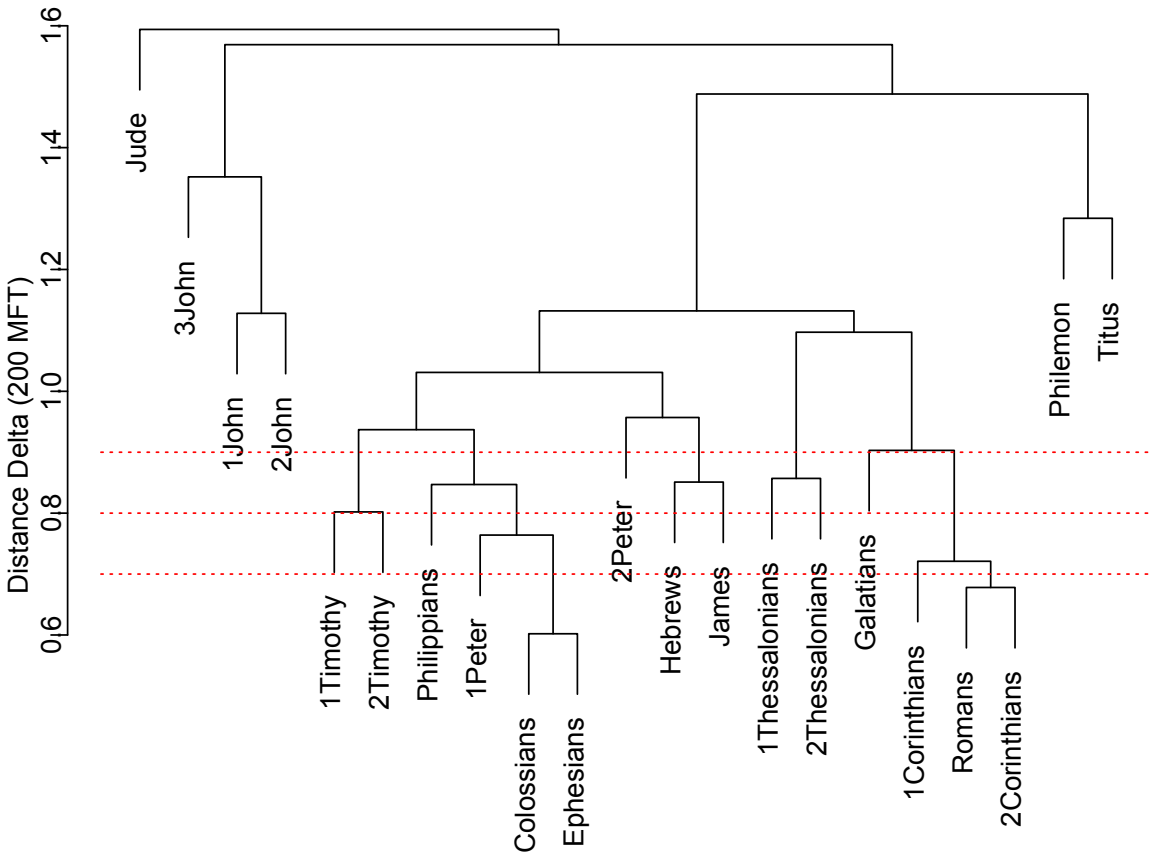

Figure 3: The twenty-one epistles represented by the 200 MFT (complete link, Delta distance, Greek version)

Verification Experiments

The results of the previous experiments do not present clear attributions. In particular, behind the *Seven Letters* hypothesis, one can find one to four possible authors (as depicted in Figures 1 to 3). To resolve this question, recent verification methods have been applied under the assumption that the *Ten Letters* hypothesis is true. In this procedure, the last seven letters (*James*, 1 and 2 *Peter*, 1, 2, 3 *John*, and *Jude*) will form the set of impostors. Following the method suggested by Koppel & Seidman (2018), each letter pair of the *Ten Letters* will be compared with the seven impostors. In this case, the impostors correspond to similar texts, written in the same language (Ancient Greek), genre (theological letters), and approximate same period (from around 40 to around 80 AD).

Following Koppel & Seidman (2018), each text is represented by 3, 4, 5 and overlapping 6-grams (four different feature types). For example, from the phrase "we give thanks", the following 4-grams are generated: {"we_g", "e_gi", "_giv", "give", …, "anks", } where "_" indicates a space. To reflect the stylistic markers, the 200, 400, or 800 most frequent *n*-grams have been used to form the feature set.

As described in Koppel & Seidman (2018), not only one but 100 iterations are performed, each randomly considering 50% of the entire feature set (e.g., one iteration is based on 50% of the

400 most frequent *n*-grams). To compute the distance between letter pairs, the Manhattan and Tanimoto (see Eq. 1) functions have been used. Thus in total 2,400 verifications have been performed (100 iterations x 4 feature types x 3 feature sizes x 2 distance measures).

From these author verification results, Figure 4 has been generated in which solid lines connect epistles pairs found to be written by the same author by more than 2,000 verification iterations. The dotted lines link letter pairs written by the same author according to 1,500 to 2,000 verification iterations.

Figure 4 indicates that the cluster *Colossians-Ephesians* (*Col-Eph*) appears distinctly from the rest and have been written by the same author. This writer seems to be distinct from the others as no link was found between these two texts and any others.

The letter to *Philemon* is alone, without any stylistic relationship with the others. The rather short length of this letter (388 words, see Table 1) can explain the difficulty in establishing a reliable assignment to this letter. Thus, it is rather difficult to specify if its author is distinct from the others or not. Having only 191 distinct words in this letter to *Philemon*, a reliable authorship attribution is rather difficult to achieve (Eder, 2015), (Savoy, 2018). As a third group, one can see 1 and 2 *Timothy* and *Titus* (the three pastoral letters) (*1Ti – 2Ti – Titus*).

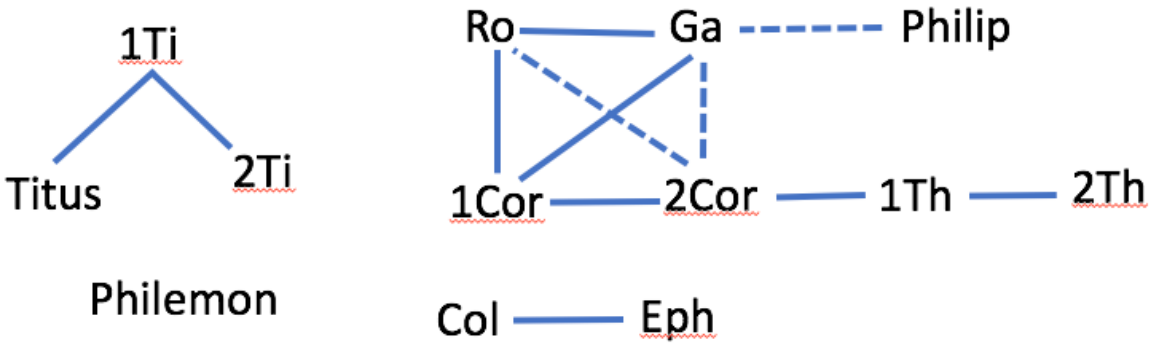

Figure 4: Graph achieved with 2,400 verifications performed over the thirteen epistles attributed to Paul

From these three groups, the verification results do not reveal any clear stylistic proximity with the seven other letters. Thus, it seems that the pastoral letters have been authored by a district person.

In Figure 4, a kernel is formed by the *Four Letters* confirming the same author behind this set. This result corroborates Baur's hypothesis. Figure 4 however exhibits a link from 2 *Corinthians* (*2Cor*) to 1 *Thessalonians* (*1Th*), and a weak link from *Galatians* (*Ga*) to *Philippians*. These two links contradict the *Four Letters* hypothesis.

Can Figure 4 justify the *Seven Letters* hypothesis (the previous four, plus *Philippians*, 1 *Thessalonians*, and *Philemon*)? First, the *Letter to Philemon* is too short to expect a clear stylistic relationship with the others. Thus, it cannot be excluded that its author is the same as the writer behind the previous four letters. Second, these four epistles have a link with both *Philippians* and 1 *Thessalonians*. This finding supports the *Seven Letters* hypothesis.

However, two concerns can be put forward. First, both *Philippians* and *1 Thessalonians* do not have any strong stylistic relationship between them, or with the other members in the kernel.

Second, the strong relationship between 1 and 2 *Thessalonians* cannot be explained by the *Seven Letters* hypothesis.

Finally, the *Ten Letters* hypothesis (*Romans*, 1 and 2 *Corinthians*, *Galatians*, *Ephesians*, *Philippians*, *Colossians*, 1 and 2 *Thessalonians*, and *Philemon*) is partially supported by Figure 4. In this case, all depicted links can be explained. *Philemon* stays isolated as explained previously. In addition, one can argue that the same author could be behind both the cluster *Colossians-Ephesians* and the larger cluster formed by the seven texts. The three pastoral letters remain separately.

This *Ten Letters* hypothesis cannot however be accepted without difficulties. Those are related with missing stylistic links between the cluster *Colossians-Ephesians* and the seven other texts. Moreover, as for the *Seven Letters* hypothesis, Figure 4 does not exhibit links between *Philippians* and the rest, nor between 1 and 2 *Thessalonians* and the rest.

Conclusion

The authorship of the Pauline epistles has been the subject of various studies (Morton, 1978), (Love, 2002). In this paper, two computer-based authorship methods (Burrows' Delta (Burrows, 2002), and intertextual distance (Labbé, 2007)) have been applied. Based on their outcome, a hierarchical clustering method (complete link) was applied to identify groups of letters depicting similar styles. Even if this study is unable to reveal the true author of all epistles, we were able to clearly identify three groups. The first homogeneous stylistic cluster regroups *Romans*, 1 and 2 *Corinthians*, and *Galatians*. This set corresponds to the four letters attributed to Paul by Baur (1845) and Morton (1978). In the second group, one can find two letters (*Colossians* and *Ephesians*) probably authored by the same person. The third cluster corresponds to 1 and 2 *Thessalonians*, with an indication that both letters might have been written by the same author. These three clusters could have been written by a unique author, or by two or three distinct persons.

These results do not corroborate the hypothesis that Paul is the author of seven letters (*Romans*, 1 and 2 *Corinthians*, *Galatians, Philippians,* 1 *Thessalonians,* and *Philemon*) (Wall, 2002). The strong similarity between 1 and 2 *Thessalonians* represents a major objection to this hypothesis.

With the *Ten Letters* hypothesis (with the addition of 2 *Thessalonians*, *Colossians*, and *Ephesians*), the main concern is explaining how the same author can adopt distinct styles corresponding to the three clusters ({*Romans*, 1 and 2 *Corinthians*, *Galatians*}, {1 and 2 *Thessalonians*}, {*Colossians, Ephesians*}), together with the difficulty of assigning *Philippians* and *Philemon* to one of the previous three clusters.

These attribution results must be taken with caution because ten epistles over fourteen have less than 5,000 words (see Table 1). According to previous studies (Eder, 2015), such a small text length renders an authorship attribution less trustworthy. As an extreme case, the short length of *Philemon* (388 words) implies that a reliable assignment is rather difficult.

The result of our verification experiment mainly confirms these findings. The four letters seem to be written by the same author, as well as 1 and 2 *Thessalonians*. However, the verification results indicate a clear link between 2 *Corinthians* and 1 *Thessalonians* but the latter

does not have a clear stylistic relationship with *Romans*, 1 *Corinthians*, and *Galatians*. Similarly, the *Letter to the Philippians* has some relationship with the *Galatians*, but not with the other three. Thus, if the *Four Letters* hypothesis is true, it is difficult to explain the stylistic relationship of a single letter of this group with either *Philippians* or 1 *Thessalonians*.

In summary, this study cannot clearly confirm one of the underlying hypotheses (*Four, Seven, or Ten Letters*). Future research might capture more precisely stylistic aspects using more complex schemes as, for example, based on a combination of text representations (Mikros & Perifanos, 2013), adopt more successful feature selection strategies (Savoy, 2015) as well as more effective distance functions or classifiers. In addition, the resulting decision must be relatively simple to interpret and with some degree of belief or probability estimate assigned to the proposed attribution (Savoy, 2016).

Finally, the text itself contains ambiguous authorship. For example, in *Romans*, one can read in the beginning 1:1 "Paul, a servant of Jesus Christ, called to be an apostle, separated unto the gospel of God,". Later, in the same letter, a second author name appears 22:16 "I Tertius, who wrote this epistle, salute you in the Lord." This passage is usually analyzed by scholars as evidence of the presence of an amanuensis in the writing process and could reflect the culture of that time (Aland, 1961).

References

Aland, K. (1961). The problem of anonymity and pseudonymity in Christian literature of the first two centuries. *Journal of Theological Studies*, 1(1), 39-49.

Aune, D.E. (2010). *The Blackwell Companion of the New Testament*. Oxford: Wiley-Blackwell.

Baayen, H.R. (2008). *Analysis Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

Baur, F. C. (1845). *Paulus, der Apostel Jesu Christi, sein Leben und Wirken, seine Briefe und seine Lehre*. Stuttgart: Becher & Müller.

Binongo, J.N.G., & Smith, M.W. (1999). The application of principal component analysis to stylometry. *Literary and Linguistic Computing,* 14(4), 445-465.

Burrows, J.F. (2002). Delta: A measure of stylistic difference and a guide to likely authorship. *Literary and Linguistic Computing*, 17, 267-287.

Craig, H., & Kinney, A.F. (Eds). (2009). *Shakespeare, Computers, and the Mystery of Authorship*. Cambridge: Cambridge University Press.

de Morgan, A. (1851). Letter to Rev. Heald 18/08/1851," in *Memoirs of Augustus de Morgan by his wife Sophia Elizabeth de Morgan with Selections from his Letters*, (S. Elizabeth and D. Morgan, eds.), London: Longman's Green and Co., 1851/1882.

Decaux, A. (2003). *L'avorton de Dieu, une vie de saint Paul*. Paris: Perrin.

Eder, M. (2015). Does size matter? Authorship attribution, small samples, big problem. *Digital Scholarship in the Humanities*, 30(2), 167-182.

Evert, S., Proisl, T., Jannidis, F., Reger, I., Pielström, S., Schöch, C., & Vitt, T. (2017). Understanding and explaining Delta measures for authorship attribution. Digital Scholarship in the Humanities, 32, (2), ii4–ii16.

Holmes, D.I. (1998). The evolution of stylometry in humanities scholarship. *Literary and Linguistic Computing*, 13(3), 111-117.

Jockers, M.L., & Witten, D.M. (2010). A comparative study of machine learning methods for authorship attribution. *Literary and Linguistic Computing*, 25, 215-223.

Juola, P. (2006). Authorship attribution. *Foundations and Trends in Information Retrieval*, 1, 1-145.

Kešelj, V., Peng, F., Cercone, N., & Thomas, C. (2003). N-gram-based author profiles for authorship attribution. *Proceedings of the Conference Pacific Association for Computational Linguistics*, pp. 255-264, ACL.

Kocher, M., & Savoy J. (2017a). A simple and efficient algorithm for authorship verification. *Journal of the American Society for Information Science and Technology*, 68(1), 259-269.

Kocher, M., & Savoy, J. (2017b). Distance measures in author profiling. *Information Processing & Management*, 53(5), 1103-1119.

Kocher, M., & Savoy, J. (2018). Evaluation of text representation schemes and distance measures for authorship linking. *Digital Scholarship in the Humanities,* to appear, https://doi.org/10.1093/llc/fqy013.

Kjell, B. (1994). Authorship Determination using letter pair frequencies features with neural networks classifiers. *Literary and Linguistic Computing*, 9(2), 119–124.

Koppel, M., Schler, J., & Bonchek-Dokow, E. (2007). Measuring differentiability: Unmasking pseudonymous authors. *Journal of Machine Learning research*, 8(6), 1261-1276.

Koppel, M., Schler, J., & Argamon, S. (2009). Computational methods in authorship attribution. *Journal of the American Society for Information Science & Technology*, 60(1), 9-26.

Koppel, M., & Winter, Y. (2014). Determining If two documents are by the same author. *Journal of American Society for Information Science & Technology,* 65(1), 178-187.

Koppel, M., & Seidman, S. (2018). Detecting pseudoepigraphic texts using novel similarity measures. *Digital Scholarship in the Humanities,* 33(1), 72-81.

Labbé, C., & Labbé, D. (2001). Inter-textual distance and authorship attribution. Corneille and Molière. *Journal of Quantitative Linguistics*, 8(3), 213-231.

Labbé, D. (2007). Experiments on authorship attribution by intertextual distance in English. *Journal of Quantitative Linguistics*, 14, 33-80.

Love, H. (2002). *Attributing Authorship: An Introduction*. Cambridge: Cambridge University Press.

Manning, C.D., & Schütze, H. (2000). *Foundations of Statistical Natural Language Processing*. Cambridge: The MIT Press.

Mikros, G. K., & Perifanos, K. (2013). Authorship attribution in Greek tweets using multilevel author's n-gram profiles. In E. Hovy, V. Markman, C. H. Martell & D. Uthus (Eds.), Proceedings AAAI Spring Symposium "Analyzing Microtext", (pp. 17-23). Palo Alto (CA): The AAAI Press.

Morton, A.Q. (1978). *Literary Detection. How to Prove Authorship and Fraud in Literature and Documents*. New York: Charles Scribner's Sons.

Porter, S.E. (1995). Pauline authorship and the pastoral epistles: Implications for canon. *Bulletin for Biblical Research* 5, 105-123.

Savoy, J. (2015). Comparative evaluation of term selection functions for authorship attribution. *Digital Scholarship in the Humanities*, 30(2), 246-261.

Savoy, J. (2016). Estimating the probability of an authorship attribution. *Journal of the American Society for Information Science and Technology*, 67(6), 1462-1472.

Savoy, J. (2018). Is Starnone really the author behind Ferrante? *Digital Scholarship in the Humanities*, 33(4), 902-918.
Sebastiani, F. (2002). Machine learning in automatic text categorization. *ACM Computing Survey*, 14, 1-27.
Stamatatos, E. (2009). A survey of modern authorship attribution methods. *Journal of the American Society for Information Science & Technology*, 60, 433-214.
Stamatatos, E., Potthast, M., Rangel, F., Rosso, P., & Stein, B. (2015). Overview of the PAN/CLEF 2015 Evaluation Lab. In Josiane Mothe *et al*., editors, *Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings 6th International Conference of the CLEF Initiative (CLEF 15)*, 518-538, Berlin: Springer.
Tuzzi, A., & Cortelazzo, M. (2018). What is Elena Ferrante? A comparative analysis of a secretive bestselling Italian writer. *Digital Scholarship in the Humanities*, 33(3), 685-702.
Wall, R. (2002). *New Interpreter's Bible Vol. X*. Nashville: Abingdon Press.
Zhao, Y., & Zobel, J. (2007). Entropy-based authorship search in large document collection. In *Proceedings ECIR* (pp. 381-392). Berlin: Springer, LNCS #4425.
Zheng, R., Li, J., Chen, H., & Huang, Z. (2006). A framework for authorship identification of online messages: writing-style features and classification techniques. *Journal of the American Society for Information Science & Technology*, 57(12), 378-393.

Appendix

Table A.1. Translation variations

Version	Philemon 1:1
King James	Paul, a prisoner of Jesus Christ, and Timothy our brother, unto Philemon our dearly beloved,
New American Standard Bible	Paul, a prisoner of Christ Jesus, and Timothy our brother, To Philemon our beloved brother and fellow worker,
New Living Translation	This letter is from Paul, a prisoner for preaching the Good News about Christ Jesus, and from our brother Timothy. I am writing to Philemon, our beloved co-worker,
Young's Literal Translation	Paul, a prisoner of Christ Jesus, and Timotheus the brother, to Philemon our beloved and fellow-worker,

Table A.2. Ranked list of Labbé's distances between two letters (All tokens, English version)

Rank	Distance	Letter	Letter
1	0.2191	Colossians	Ephesians
2	0.2241	Romans	1 Corinthians
3	0.2527	Romans	Galatians
4	0.2569	Romans	2 Corinthians
5	0.2580	1 Thessalonians	2 Thessalonians
6	0.2591	1 Corinthians	2 Corinthians
7	0.2593	Colossians	1 Peter
8	0.2594	Philippians	Colossians
9	0.2615	Romans	Hebrews
10	0.2651	2 Corinthians	Galatians

…	…	...	...

Table A.3. Ranked lists produced by the Delta model (200 MFT, English version)

Rank	Distance	Letter	Distance	Letter	Distance	Letter	Distance	Letter
	Romans		*1 Corinthians*		*1 Thessalonians*		*Philemon*	
1	0.688	*Galatians*	0.673	*Romans*	0.786	*2 Thessalonian*	1.272	*Philippians*
2	0.691	*2 Corinthians*	0.681	*2 Corinthians*	0.912	*1 Peter*	1.401	*2 Corinthian*
3	0.742	*1 Corinthians*	0.742	*James*	0.922	*2 Corinthians*	1.450	*Galatians*