



1-31-1997

Comparative Power of Three Author-Attribution Techniques for Differentiating Authors

John B. Archer
Redcon, Inc.

John L. Hilton
Brigham Young University

G. Bruce Schaalje
Brigham Young University

Follow this and additional works at: <https://scholarsarchive.byu.edu/jbms>

BYU ScholarsArchive Citation

Archer, John B.; Hilton, John L.; and Schaalje, G. Bruce (1997) "Comparative Power of Three Author-Attribution Techniques for Differentiating Authors," *Journal of Book of Mormon Studies*: Vol. 6 : No. 1 , Article 5.

Available at: <https://scholarsarchive.byu.edu/jbms/vol6/iss1/5>

This Feature Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in *Journal of Book of Mormon Studies* by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.



NEAL A. MAXWELL INSTITUTE
FOR RELIGIOUS SCHOLARSHIP

BRIGHAM YOUNG UNIVERSITY • PROVO, UTAH

Title Comparative Power of Three Author-Attribution Techniques for Differentiating Authors

Author(s) G. Bruce Schaalje, John L. Hilton, and John B. Archer

Reference *Journal of Book of Mormon Studies* 6/1 (1997): 47-63.

ISSN 1065-9366 (print), 2168-3158 (online)

Abstract Over the last twenty years, various objective author-attribution techniques have been applied to the English Book of Mormon in order to shed light on the question of multiple authorship of Book of Mormon texts. Two methods, one based on rates of use of noncontextual words and one based on word-pattern ratios, measure patterns consistent with multiple authorship in the Book of Mormon. Another method, based on vocabulary-richness measures, suggests that only one author is involved. These apparently contradictory results are reconciled by showing that for texts of known authorship, the method based on vocabulary-richness measures is not as powerful in discerning differences among authors as are the other methods, especially for works translated into English by a single translator.

Comparative Power of Three Author- Attribution Techniques for Differentiating Authors

G. Bruce Schaalje, John L. Hilton, and John B. Archer

Abstract: Over the last twenty years, various objective author-attribution techniques have been applied to the English Book of Mormon in order to shed light on the question of multiple authorship of Book of Mormon texts. Two methods, one based on rates of use of noncontextual words and one based on word-pattern ratios, measure patterns consistent with multiple authorship in the Book of Mormon. Another method, based on vocabulary-richness measures, suggests that only one author is involved. These apparently contradictory results are reconciled by showing that for texts of known authorship, the method based on vocabulary-richness measures is not as powerful in discerning differences among authors as are the other methods, especially for works translated into English by a single translator.

Two dollar-bill changers are available in the building where we work. One is of an older style, but it is our favorite. It recognizes that a dollar bill is not bogus even when the bill is old and washed out. The modern changer is more conservative. The dollar bill has to be crisp and bold to convince this machine that it is not counterfeit. Both machines are valid dollar-bill changers in the sense that they give change when they are absolutely sure that a real dollar bill has been fed into them. Neither machine has been replaced, so we can assume that neither machine makes errors in the sense of getting fooled by counterfeit bills. But it would be a

mistake to conclude that the piece of paper in your hand is a counterfeit dollar bill just because the conservative machine in the main lobby will not accept it. If you were trying to detect counterfeit bills, the old north-wing machine would be much more useful. When it does not accept a bill, you can be fairly sure that something about the bill is really strange. You can think of the old north-wing machine as being more powerful in discerning the difference between real and counterfeit money.

What has this story to do with authorship analysis? Several objective author-attribution techniques are in current use, all oriented around the idea of assigning numerical measures to various aspects of authors' styles in an attempt to answer questions about texts of unknown or disputed authorship.¹ These techniques, which have proliferated and gained popularity since the advent of accessible high-speed computers, are like bill changers. If it is suspected, for example, that a literary text traditionally ascribed to Shakespeare was not in fact written by Shakespeare, both the controversial text and others known to have been authored by Shakespeare can be examined using an objective author-attribution technique. If the technique reveals a large statistical difference between the controversial text and the known Shakespearean texts, such strong evidence implies that Shakespeare did not write the controversial text. But if only a small difference is found, we cannot make any conclusion unless we know how powerful the attribution technique is in discriminating among authors. The test we used may be like the bill changer in the main lobby—too conservative to pick out the real difference.

This simple but subtle point was not initially understood by Holmes, who computed various measures of "vocabulary richness" for segments of text drawn from the Book of Mormon, the Doctrine and Covenants, the book of Abraham, Isaiah, and personal writings of Joseph Smith.² These measures reflect aspects of a writer's working vocabulary, such as its size and the writer's habits for drawing upon it. Using statistical methods of investi-

¹ David I. Holmes, "Authorship Attribution," *Computers and the Humanities* 28 (1994): 87-106.

² David I. Holmes, "A Stylometric Analysis of Mormon Scripture and Related Texts," *Journal of the Royal Statistical Society Series A* 155 (1992): 91-120.

gating differences among entities for which several numerical measures are available, Holmes showed that based on vocabulary-richness measures, the texts seemed to fall into three distinct groups: (1) Isaiah texts, (2) segments of Joseph Smith's personal writings, and (3) all the rest. Because texts ascribed to different Book of Mormon authors did not segregate on a prophet-by-prophet basis nor differ very much from Doctrine and Covenants or book of Abraham texts, Holmes concluded that they were all written by the same author. He proposed that they were all the work of Joseph Smith and that they differed in vocabulary-richness from Joseph Smith's personal writings only because Smith was apparently able to write in a distinct "prophetic voice" when he desired.³ Holmes did not recognize that his conclusions would only be reasonable if his vocabulary-based author-attribution technique could be shown to be very powerful in distinguishing among authors.

Holmes was not aware that his findings about the similarity of working vocabularies used by different Book of Mormon prophets was not original. Hilton reported that "new word introduction rates" in Book of Mormon writings ascribed to different prophets were very similar.⁴ Holmes was also not aware that in a separate study, Hilton had used certain noncontextual word-pattern ratios as an author-attribution technique and had thereby shown that Book of Mormon texts attributable to Nephi and Alma differed significantly.⁵ However, Holmes was aware that Larsen, Rencher, and Layton had applied yet another objective author-attribution technique to Book of Mormon writings and had also shown that writings of different Book of Mormon prophets differed significantly in their rates of use of common noncontextual words.⁶ Holmes argued that his technique must be preferable to that of

³ David I. Holmes, "Vocabulary Richness and the Prophetic Voice," *Literary & Linguistic Computing* 6 (1991): 259-68.

⁴ John L. Hilton, "Some Book of Mormon 'Word Print' Measurements Using 'Wrap-around' Block Counting" (Provo, Utah: FARMS, 1988).

⁵ John L. Hilton, "On Verifying Wordprint Studies: Book of Mormon Authorship," *BYU Studies* 30/3 (1990): 89-108; also available as a FARMS reprint.

⁶ Wayne A. Larsen, Alvin C. Rencher, and Tim Layton, "Who Wrote the Book of Mormon? An Analysis of Wordprints," *BYU Studies* 20/3 (1980): 225-51.

Larsen et al. because his method used all textual words in its calculations, but he provided no support, empirical or theoretical, to validate this statement.⁷ It is interesting, therefore, that in a recent paper Holmes reversed his position and praised the use of non-contextual word frequencies when he found that authorship attribution based on vocabulary richness was not able to segregate *Federalist Papers* texts attributed to Hamilton, Madison, and Jay as clearly as the method based on rates of use of common non-contextual words.⁸

It seems entirely possible that texts of different authorship but translated by a single translator, as the English Book of Mormon texts are claimed to be, could exhibit the vocabulary richness of the translator, but still have unique rates of use of noncontextual words and word patterns common to the original authors. If so, the findings of Holmes do not give any weight to the position that Joseph Smith was the sole author of the Book of Mormon.

The purpose of this study is to use texts of known authorship to investigate the relative power of each of the three author-attribution techniques mentioned above. Both original nontranslated works and translated works are used in this study. This information will be helpful in correctly interpreting results of studies for which differences are *not* detected.

Author-Attribution Techniques

Many objective author-attribution techniques are in current use; however, because of their connection to work on the Book of Mormon, we concentrate on three techniques—methods based on measures of vocabulary richness, on the rates of use of common noncontextual words, and on noncontextual word-pattern ratios. The various measures will be referred to generically as “stylometric measures.” Most of these measures are corrected for the length of the text, but to further guarantee that text length did not influence the outcome, we used texts of 5,000 words each in the current study.

⁷ Holmes, “Stylometric Analysis,” 98.

⁸ David I. Holmes and D. I. Forsyth, “The Federalist Revisited: New Directions in Authorship Attribution,” *Literary and Linguistic Computing* 10 (1995): 111–27.

Holmes suggested five measures of vocabulary richness (VR) for use in studying disputed authorship questions.⁹ The first two measures, which he termed *hapax legomena* (R) and *hapax dislegomena* (V_2/V), are counts of once-used and twice-used words, respectively, standardized by the length of the text. Two of the other three measures are related to specific probability models for vocabulary usage, but will neither be used nor discussed further here because Holmes shows that all three are somewhat redundant and concludes that "for characterizing the differences between the textual samples, therefore, only variables R and V_2/V need to be computed."¹⁰

Larsen et al. based their work on the frequency of occurrence of thirty-eight common noncontextual words (NCW) such as *and* and *the* (see Larsen et al. for a list of the thirty-eight words).¹¹ In this paper we compute the frequency of occurrence of the following twenty common words, in alphabetical order: *a, all, an, and, any, as, but, by, in, it, no, not, of, that, the, to, up, upon, with, without*.

Hilton calculated sixty-five noncontextual word-pattern ratios (WPR) (originally suggested by Morton).¹² Examples of such ratios include the number of times *a* appears as the first word of a sentence divided by the number of sentences; the number of times *and* is followed by an adjective divided by the number of times *and* is used; and the number of times *any* is used divided by the number of times *any* and *all* are used. All sixty-five word-pattern ratios were calculated for all texts in this study.¹³

Holmes, Hilton, and Larsen et al. each used a different statistical method in connection with their stylometric measures to discern authorship differences among texts. For ease of comparison and to eliminate differences ascribable to statistical methods, we used a single statistical method, discriminant analysis,¹⁴ to

⁹ Holmes, "Stylometric Analysis," 92-5.

¹⁰ Ibid., 116.

¹¹ Larsen et al., "Who Wrote the Book of Mormon?" 247.

¹² Hilton, "On Verifying Wordprint Studies," 96. A. Q. Morton, *Literary Detection: How to Prove Authorship and Fraud in Literature and Documents* (New York: Scribner's Sons, 1978); also personal communication.

¹³ Hilton, "On Verifying Wordprint Studies," 104.

¹⁴ Alvin C. Rencher, *Methods of Multivariate Analysis* (New York: Wiley, 1995), 296-349.

quantify the degree of separation of the texts due to authors for all three techniques. Under this method a mathematical rule for assigning texts to authors is developed based on the stylometric measures. The rule is then applied to each of the texts, and an indicator of the degree of separation of the texts according to author is the percentage of texts correctly classified. Two variants of this method were used: (1) the *restitution* approach by which the texts used to develop the rule were also classified by the rule and (2) the *cross-validation* approach by which each text in turn is classified using a rule developed with that text left out. Either variant is useful for purposes of comparing the author-attribution techniques, but the cross-validation approach has the additional benefit that it gives a better idea of how successful we might expect to be in assigning a text of unknown authorship to the correct author using the technique.

Because the sets of measures for two of the techniques (NCW and WPR) were large, they were subjected to principal components analysis¹⁵ in order to reduce the dimensionality. This method uses the correlation structure of a large set of measures to generate a small set (usually two or three) of composite stylometric measures, called principal components, which contain most of the information carried by the large set. The development of the principal components is valid in that it is carried out blind to the actual authorship of the texts.

SAS software was used to carry out the discriminant analysis and principal components analysis computations.¹⁶ A BASIC program was used to compute the stylometric measures.

Texts

The original nontranslated 5,000-word texts of known authorship ("control texts") chosen for this study (table 1) included a number of literary genres and covered a fairly large time span. Their use was also based, in part, on availability. No claim is made that these texts represent an optimal set of texts for which to evaluate the power of author-attribution techniques. However, they

¹⁵ Ibid., 415–44.

¹⁶ SAS Institute Incorporated, *SAS/STAT User's Guide, Version 6, Fourth Edition* (Cary, N.C.: SAS, 1990).

were chosen before the application of any of these techniques to them and so can be considered unbiased with regard to displaying differences in power among the techniques.

Table 1. Control Texts	
Author	Texts
Samuel Clemens	2 selections from <i>The Complete Short Stories of Mark Twain</i> , 1 from "Extracts from Adam's Diary" and 1 from "Eve's Diary"; 1 selection from "Early Days" in <i>Mark Twain's Autobiography</i> ; 1 selection from <i>Does the Race of Man Love a Lord?</i>
Oliver Cowdery	4 selections of religious discourse and biographical essay in the <i>Messenger and Advocate</i> , entitled "Letters to W. W. Phelps"
Robert Heinlein	2 selections from <i>The Number of the Beast</i> , 1 representing the character Hilda and the other representing the character Deety; 4 selections from <i>Revolt in 2100</i>
Samuel Johnson	2 selections from <i>The Rambler</i> ; one selection from <i>The Idler</i> ; 2 selections from <i>A Journey to the Western Islands of Scotland</i> ; 1 selection from <i>The Fountains: A Fairy Tale</i>
Joseph Smith	2 selections of letters to his wife and friends from <i>The Personal Writings of Joseph Smith</i> ; 1 selection from "Joseph Smith—History" in the Pearl of Great Price
Harry Steinbauer	2 selections from "The Novella," a commentary in <i>Twelve German Novella</i> ; 1 selection from <i>Heine and Cecile Furtado: A Reconsideration</i>

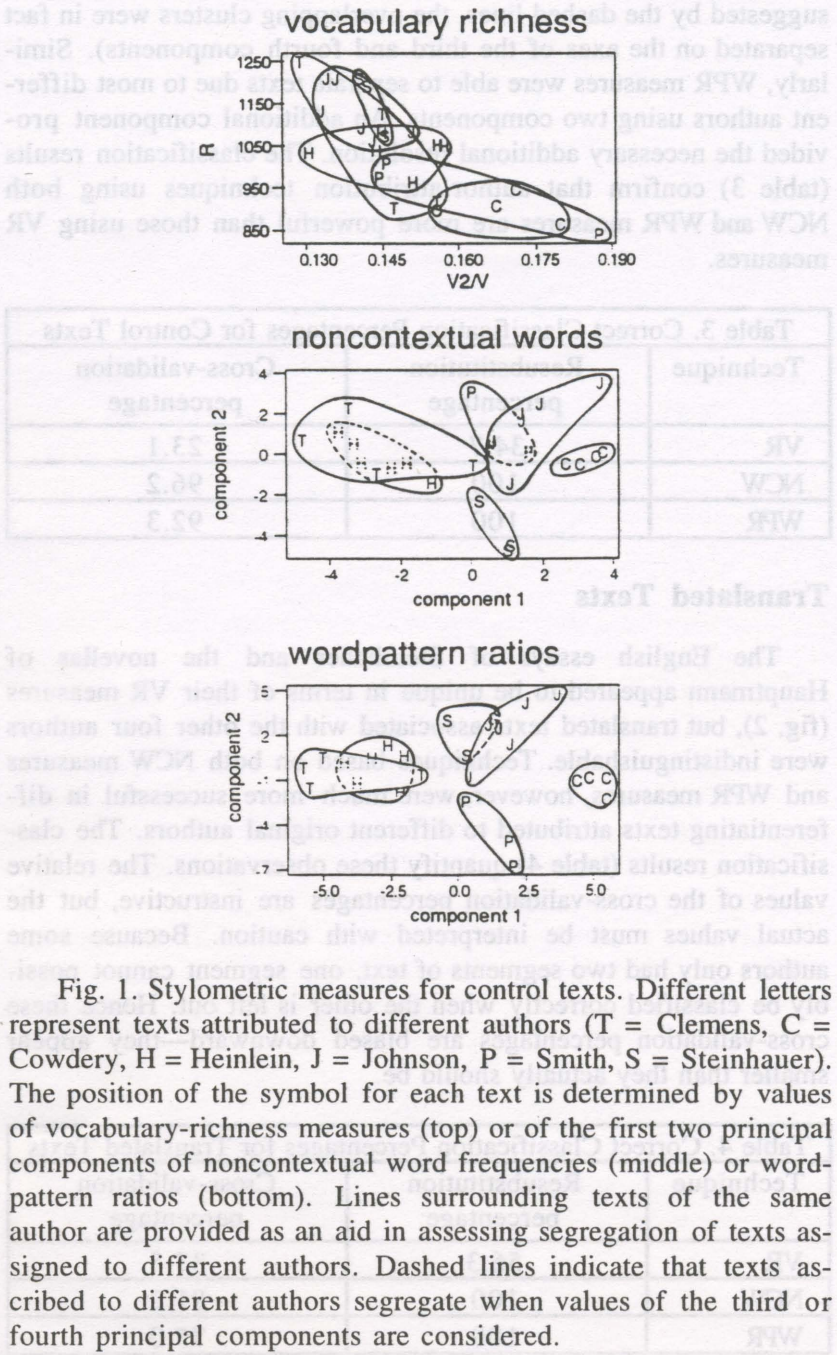
The translated texts used in this study (table 2) are all from a set of German novellas translated by Steinhauer.¹⁷ This set of translated works is of particular interest because the texts were written in German by different authors but are of the same genre and were translated by a single translator to English. In addition, original untranslated essays written in English by Steinhauer himself are available in the same book. Those novellas for which at least two 5,000-word texts could be extracted were used in this study.

Author	Texts
Harry Steinhauer	3 English selections as listed in table 1
Christoph Wieland	2 selections from <i>Love and Friendship Tested</i>
Heinrich von Kleist	3 selections from <i>Michael Kohlhaas</i>
Ernst Hoffmann	2 selections from <i>Mademoiselle de Scudery</i>
Theodore Fontane	2 selections from <i>Stine</i>
Gerhart Hauptmann	3 selections from <i>The Heretic of Soana</i>

Control Texts

With few exceptions, VR measures were unable to distinguish texts attributed to different authors (fig. 1). Even texts written in such different genres and time periods as those attributed to Samuel Johnson and Robert Heinlein were not differentiable using VR measures. Note that Mark Twain's writings span almost the whole range of R values as he attempts to make his writings represent different people (Adam and Eve). In contrast, NCW measures were able to differentiate texts attributed to most authors by using just the first two principal components. Using two additional components, almost perfect separation of authors is achieved (as

¹⁷ Harry Steinhauer, trans. and ed., *Twelve German Novellas* (Berkeley: University of California, 1977).



suggested by the dashed lines, the overlapping clusters were in fact separated on the axes of the third and fourth components). Similarly, WPR measures were able to separate texts due to most different authors using two components. An additional component provided the necessary additional resolution. The classification results (table 3) confirm that author-attribution techniques using both NCW and WPR measures are more powerful than those using VR measures.

Technique	Resubstitution percentage	Cross-validation percentage
VR	34.7	23.1
NCW	100	96.2
WPR	100	92.3

Translated Texts

The English essays of Steinhauer and the novellas of Hauptmann appeared to be unique in terms of their VR measures (fig. 2), but translated texts associated with the other four authors were indistinguishable. Techniques based on both NCW measures and WPR measures, however, were much more successful in differentiating texts attributed to different original authors. The classification results (table 4) quantify these observations. The relative values of the cross-validation percentages are instructive, but the actual values must be interpreted with caution. Because some authors only had two segments of text, one segment cannot possibly be classified correctly when the other is left out. Hence these cross-validation percentages are biased downward—they appear smaller than they actually should be.

Technique	Resubstitution percentage	Cross-validation percentage
VR	56.3	37.5
NCW	100	81.2
WPR	100	75.0

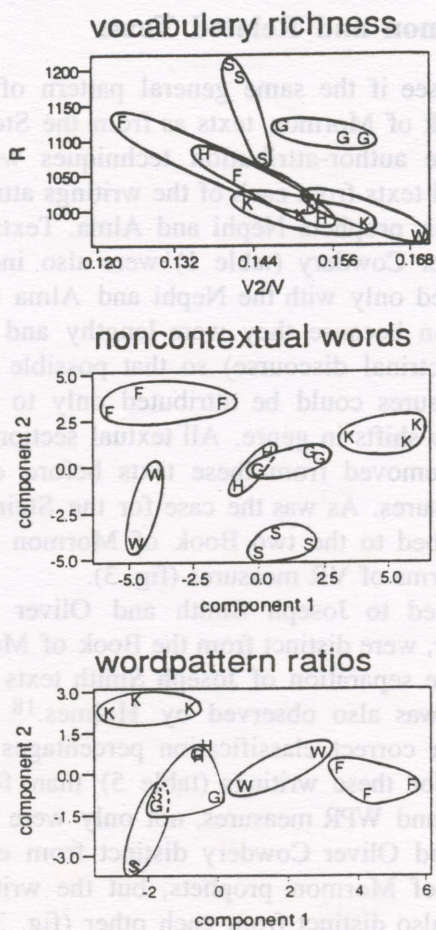


Fig. 2. Stylometric measures for translations. Different letters represent texts due to different authors (S = Steinhauer, W = Wieland, K = von Kleist, H = Hoffmann, F = Fontane, G = Hauptmann). The position of the symbol for each text is determined by values of vocabulary-richness measures (top) or of the first two principal components of noncontextual word frequencies (middle) or word-pattern ratios (bottom). Lines surrounding texts of the same author are provided as an aid in assessing segregation of texts due to different authors. Dashed lines indicate that texts due to different authors segregate when values of the third or fourth principal components are considered.

Book of Mormon and Related Texts

In order to see if the same general pattern of results is obtained from Book of Mormon texts as from the Steinhauer translations, the three author-attribution techniques were applied to three 5,000-word texts from each of the writings attributable to the Book of Mormon prophets Nephi and Alma. Texts from Joseph Smith and Oliver Cowdery (table 1) were also included in this study. We worked only with the Nephi and Alma texts from the Book of Mormon because they were lengthy and written in the same genre (doctrinal discourse) so that possible differences in stylometric measures could be attributed only to author differences and not to shifts in genre. All textual sections of historical narrative were removed from these texts before computing the stylometric measures. As was the case for the Steinhauer translations, texts ascribed to the two Book of Mormon prophets were not distinct in terms of VR measures (fig. 3).

Texts ascribed to Joseph Smith and Oliver Cowdery personally, however, were distinct from the Book of Mormon texts in VR measures; the separation of Joseph Smith texts from Book of Mormon texts was also observed by Holmes.¹⁸ Consequently, somewhat higher correct classification percentages based on VR were observed for these writings (table 5) than for the control texts. For NCW and WPR measures, not only were the writings of Joseph Smith and Oliver Cowdery distinct from each other and from the Book of Mormon prophets, but the writings of Nephi and Alma were also distinct from each other (fig. 3). The correct classification percentages for NCW and WPR measures were much higher than for VR (table 5). We conclude, therefore, that no stylometric evidence disproves Joseph Smith's claim that he was the translator of works written by multiple foreign-language authors.

¹⁸ Holmes, "Stylometric Analysis," 109, 116.

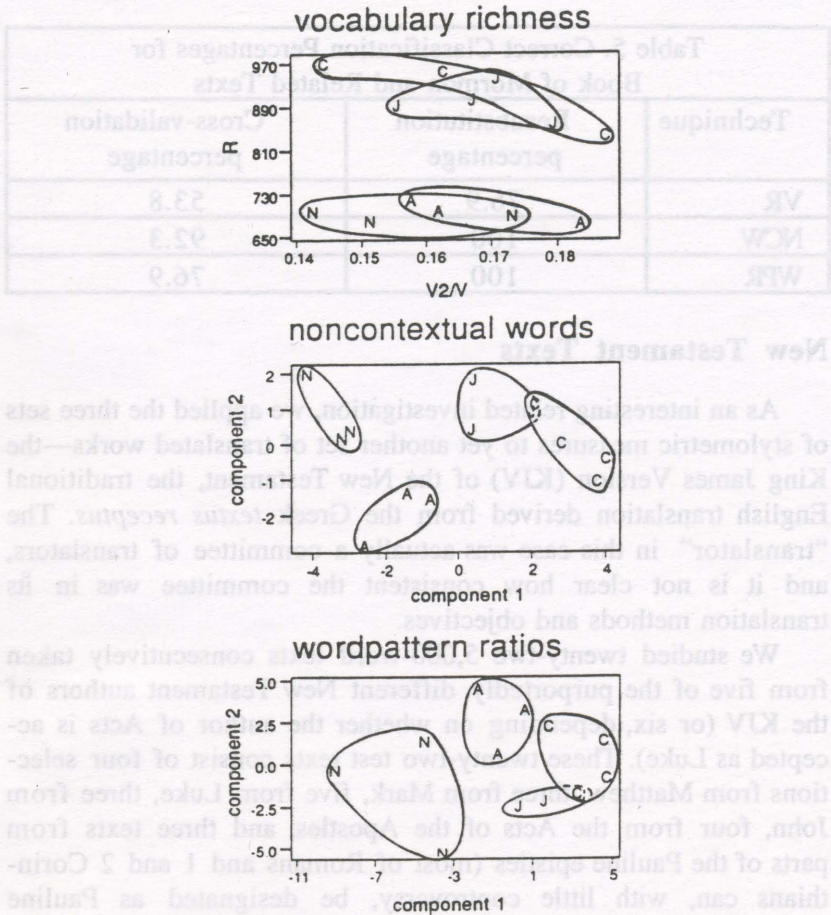


Fig. 3. Stylometric measures for Book of Mormon and related texts. Different letters represent texts attributed to different prophets or authors (N = Nephi, A = Alma, J = Joseph Smith, C = Oliver Cowdery). The position of the symbol for each text is determined by values of vocabulary-richness measures (top) or of the first two principal components of noncontextual word frequencies (middle) or word-pattern ratios (bottom). Lines surrounding texts of the same author are provided as an aid in assessing segregation of texts ascribed to different authors. Dashed lines indicate that texts attributed to different authors segregate when values of the third or fourth principal components are considered.

Table 5. Correct Classification Percentages for
Book of Mormon and Related Texts

Technique	Resubstitution percentage	Cross-validation percentage
VR	76.9	53.8
NCW	100	92.3
WPR	100	76.9

New Testament Texts

As an interesting related investigation, we applied the three sets of stylometric measures to yet another set of translated works—the King James Version (KJV) of the New Testament, the traditional English translation derived from the Greek *textus receptus*. The “translator” in this case was actually a committee of translators, and it is not clear how consistent the committee was in its translation methods and objectives.

We studied twenty-two 5,000-word texts consecutively taken from five of the purportedly different New Testament authors of the KJV (or six, depending on whether the author of Acts is accepted as Luke). These twenty-two test texts consist of four selections from Matthew, three from Mark, five from Luke, three from John, four from the Acts of the Apostles, and three texts from parts of the Pauline epistles (most of Romans and 1 and 2 Corinthians can, with little controversy, be designated as Pauline according to previous stylometric measurements of the Greek).¹⁹

Other than the texts from the Gospel of John, which had very low vocabulary richness, few differences attributable to authors could be discerned using VR measures (fig. 4). Using NCW measures, especially WPR measures, enough clustering frequently permits segregation of the texts according to authors. Except for the shaded area covering the five texts from the Gospel of Luke, the segregation of the translated English wordings for these New Testament authors approaches that of our different English writing control authors or Steinhauer’s English translations of his

¹⁹ Morton, *Literary Detection*, 182–3.

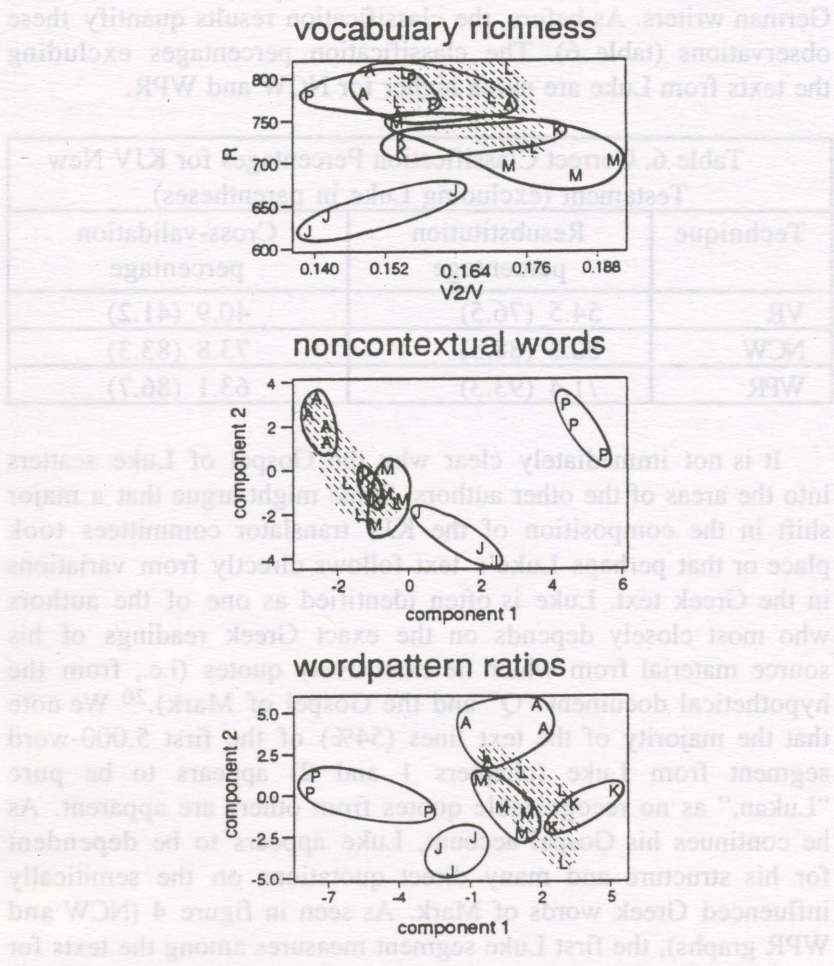


Fig. 4. Stylometric measures for the KJV New Testament. Different letters represent texts due to different authors (M = Matthew, K = Mark, L = Luke, J = John, A = Acts of the Apostles, P = Pauline Epistles). The position of the symbol for each text is determined by values of vocabulary-richness measures (top) or of the first two principal components of noncontextual word frequencies (middle) or word-pattern ratios (bottom). Lines (or shading in the case of Luke) surrounding texts of the same author are provided as an aid in assessing segregation of texts due to different authors.

German writers. As before, the classification results quantify these observations (table 6). The classification percentages excluding the texts from Luke are much higher for NCW and WPR.

Technique	Resubstitution percentage	Cross-validation percentage
VR	54.5 (76.5)	40.9 (41.2)
NCW	80.3 (88.3)	73.8 (83.3)
WPR	71.4 (93.3)	63.1 (86.7)

It is not immediately clear why the Gospel of Luke scatters into the areas of the other authors. Some might argue that a major shift in the composition of the KJV translator committees took place or that perhaps Luke's text follows directly from variations in the Greek text. Luke is often identified as one of the authors who most closely depends on the exact Greek readings of his source material from which he extensively quotes (i.e., from the hypothetical document "Q" and the Gospel of Mark).²⁰ We note that the majority of the text lines (54%) of the first 5,000-word segment from Luke (chapters 1 and 2) appears to be pure "Lukan," as no recognizable quotes from others are apparent. As he continues his Gospel account, Luke appears to be dependent for his structure and many direct quotations on the semantically influenced Greek words of Mark. As seen in figure 4 (NCW and WPR graphs), the first Luke segment measures among the texts for Acts, which are traditionally thought to be pure Lukan. Especially in the NCW graph, it appears that the four other Luke Gospel texts are scattered around the Mark and Matthew cluster. It has been observed that in the Greek text, Matthew quotes even more extensively from Mark than did Luke while he cleaned up Mark's colloquial Greek. Therefore, the overlapping of the Matthew and Mark clusters for NCW measurements in figure 4 (but not for WPR) might in part be explained in differing abilities of the two procedures to sense this kind of change in the Greek as reflected in the English translations. Nevertheless, regardless of possible

²⁰ Roger R. Keller, personal communication.

explanations for the scatter of the sections of Luke, the English words of the KJV from the other five tested New Testament authors show a clear and nonambiguous author clustering. Only two explanations are apparent for this clustering: (1) a consistent major shifting by the KJV translators occurred precisely with each of the New Testament books, or (2) a measurable underlying unique pattern for each of these authors existed in the Greek text itself and was translated into the KJV English. The first explanation seems unlikely both in a historic context and because the NCW and WPR measures of the first chapters of Luke lie within the area of the Acts.

Conclusions

From our studies of texts of known authorship, it is clear that vocabulary-richness measures do not generally have good power for differentiating texts according to authors. Thus in author-attribution studies, a lack of difference between texts for vocabulary-richness measures does not imply no difference in authorship of the texts and certainly does not imply that differences detected using other stylistometric measures should be negated.

On the other hand, both noncontextual word frequencies and word-pattern ratios seem to have relatively good differentiating power. Author-attribution methods based on these measures would seem to be the first choice. Vocabulary-richness measures may still be very informative and useful, but their application to detect differences and especially similarities among texts of questionable authorship has severe limitations.

In light of our results for translated works and texts from the Book of Mormon, the fact that writings attributed to different Book of Mormon prophets have similar vocabulary richness but distinct frequencies of noncontextual words and word-pattern ratios is completely consistent with Joseph Smith's educational level and his account of the translation process. This conclusion is strengthened by the fact that translated writings attributed to different New Testament authors also show similar vocabulary richness but display distinct frequencies of noncontextual words and word-pattern ratios.