

An exercise in non-ideal authorship attribution: the mysterious Maria Ward

David L. Hoover and Shervin Hess

Department of English, New York University, UK

Abstract

The dangers of computational approaches to authorship attribution in the absence of an adequate set of training texts for the claimant authors are well known. This study aims to show, however, that significant progress can be made even where conditions are quite problematic. We investigate a difficult authorship question involving three texts, ostensibly by three authors, each of whom wrote nothing else. Only one of the texts can be unquestionably ascribed to a known author, and this author has been suggested as the true author of one of the two remaining texts. We investigate these three texts, along with similar texts by other authors, using cluster analysis, Delta analysis, *t*-testing, and PCA. We also create simulations of our authorship problem using sets of three texts of known authorship by one, two, and three authors. We test these sets using correct and incorrect assumptions of authorial difference, and then compare the results with analyses of our three texts based on the same range of assumptions. By combining information from all of these tests, we achieve what we believe is a persuasive, if not conclusive, solution to a significant and long-standing question concerning the authorship of Maria Warda's violently anti-Mormon *Female Life Among the Mormons*. At the same time, we demonstrate methods for making progress in cases where conditions are less than ideal.

Correspondence:

David L. Hoover,
Department of English,
New York University,
19 University Place,
5th Floor, New York,
NY 10003, USA.

E-mail:

david.hoover@nyu.edu

1. Introduction

We investigate here an authorship question involving *Female Life Among the Mormons* (Ward, 1855; *Female Life*, below).¹ This rabidly anti-Mormon first-person autobiographical account, purportedly written by the wife of a Mormon Elder, sold huge numbers of copies and remained in print for at least thirty-five years. Though published anonymously, the narrative reveals its author as Maria Ward. Yet the book has long been tentatively identified as a pseudonymous work of fiction by Mrs B. G. Ferris. The identity of the author of *Female Life*

strikes us as an intrinsically interesting question that has ramifications for religious studies, history, politics, and literary studies, especially when we consider that Mrs B. G. Ferris was the wife of a former Secretary of Utah Territory.

Because traditional authorship investigation has been unable to supply a firm attribution, a computational investigation seems appropriate. Yet the conditions for such an investigation are far from ideal: no other writings by Maria Ward or Mrs B. G. Ferris are known, though there is one other book that Maria Ward claims to have edited. In such unpromising cases, one valuable

cautionary voice makes the following unequivocal recommendation: 'If you are planning a study and cannot get the correct electronic texts, or you realize that control texts do not exist, do not do the study' (Rudman, 1997, p. 354). This is excellent advice for an investigator looking for a problem upon which to test authorship methods. Unfortunately the authorship of *Female Life*, like many other interesting and compelling authorship attribution questions in which traditional methods of determining authorship have been unsuccessful, lacks such adequate texts and other information. What if we are not willing just to give up? We present here a series of approaches that we believe can shed a great deal of light on this difficult problem. These approaches should also prove useful in other attribution problems in which the question is compelling but less-than-ideal conditions preclude a computational investigation based on best practices.

2. Historical Context

The Second Great Awakening suffused mid-nineteenth-century America with a host of new faiths, a new brand of politics, and a surge of religious bigotry aimed at non-Protestants. Catholics and Jews bore the brunt of this intolerance, but as the nation grappled with the issue of slavery, another moral dilemma drew the attention of both the nation and Congress: polygamy. The practice of polygamy, or plural marriage, cast a shadow over the Mormon Church, still in its infancy, and accounts of it horrified the American people, who cried out for an end to what they saw as enslavement of white women (Gordon, 2002, pp. 47ff).

The Mormon Church had been attacked almost from its beginnings, but when it first publicly approved plural marriage in 1852, the same year *Uncle Tom's Cabin* was published, this combination of events stimulated the writing of anti-Mormon novels, which would number more than fifty by the first decade of the twentieth century (Arrington and Haupt, 1968, p. 244). The scandalous scenarios painted by the anti-polygamy

novelists became wildly popular. Storylines often followed a beautiful young heroine's experience of being kidnapped, enslaved, or otherwise mesmerized by a depraved Mormon husband with an insatiable appetite for collecting wives. Popular anti-Mormon literature includes a surprising number of texts by well known authors. Chapters about Mormons appear in Mark Twain's *Roughing It* (1872), though, characteristically, his barbs are in the service of humor. Robert Louis Stevenson includes a chapter on the Mormons, 'Story of the Destroying Angel,' in *The Dynamiter* (1885) (written with his wife Fanny). The antagonists in Arthur Conan Doyle's first Sherlock Holmes novel, *A Study in Scarlet* (1887) are a band of Mormon men who kidnap gentile women. Zane Grey's *Rider's of the Purple Sage* (1912) and *The Rainbow Trail* (1915), both portray numerous atrocities committed by Mormons, and Jack London treats the infamous Mountain Meadows massacre in his *The Star Rover* (1915).

As public outcry against polygamy heightened, driven in large part by popular literature, members of Congress grew uneasy over the power that the Mormon Church wielded in the Utah Territory. The recently formed Republican party used anti-polygamy as one of its campaigning platforms, referring to polygamy and slavery as the 'Twin Relics of Barbarism' (Gordon, 2002, p. 55). Fears that the remote territory would become a rogue, theocratic state prompted Congress to enact legislation aimed at tightening control over the Mormon government. As early as 1856 Congress had begun voting on measures that sought to punish and prevent the practice of polygamy in the Territories of the USA and other places (H. R. 433, 1856). By 1862, the Morrill Anti-Bigamy Act was signed into law, banning plural marriage and limiting Church ownership of US territory to \$50,000. The law was ineffective, however, largely because of the remote situation of Utah Territory and a lack of funds for enforcement (Gordon, 2002, pp. 81–3). The next round of anti-polygamy legislation was enacted twenty-five years later in 1887: under the Edmunds–Tucker Act, civil marriage licenses were required,

wives were allowed to testify against their husbands, and polygamy could be punished by a fine of up to \$800 (Woodworth-Ney, 2008, pp. 156–7).

The causes of anti-Mormon feeling were complex, but both Mormon and literary historians have argued that four books published in 1855 and 1856 served as the nucleus of anti-Mormon propaganda and encouraged the litany of anti-polygamy and anti-Mormon literature published throughout the late nineteenth and early twentieth centuries: Alfreda Eva Bell's *Boadicea, the Mormon Wife* (Bell, 1855), Orvilla S. Belisle's *Mormonism Unveiled* (Belisle, 1855), Metta Victor's *Mormon Wives* (Victor, 1856), and *Female Life Among the Mormons* (Ward, 1855; Arrington and Haupt, 1968, pp. 244–5; Gordon, 2002, pp. 29–32).

The last of these books stands apart for several reasons to be explored here. Worldcat categorizes the other three books as fiction, and Bell, Belisle, and Victor already enjoyed some reputation in literary circles. *Female Life*, in contrast, is an anonymous, first-person narrative, framed as an autobiographical account. When it was first published, a *New York Times* book review both panned the book and labeled it fiction: 'An anonymous work like this, making grave accusations against a body, and against an individual, is not entitled too much consideration at any rate. For our part, we cannot bring ourselves to believe its details are true' (Review, 1855, p. 3).

Despite unfavorable reviews, the first edition of *Female Life* sold at least 40,000 copies and was later reprinted in Danish, French, Hungarian, Swedish, and German (Worldcat, 2001–2008). It also went through several subsequent editions with varying titles and remained in print at least until 1890. Both the name Maria Ward and *Female Life* have long perplexed scholars. Though the work is now widely regarded as fiction, questions surrounding its authorship linger. Because it was one of the most widely-read works of anti-Mormon propaganda, determining its authorship may lead to a more nuanced understanding of the politics at play during the formative years of the relationship between the USA and the Mormon Church.

3. Overview of *Female Life*

Female Life chronicles a period in the life of a woman from Onondaga County, New York who marries a Mormon man, referred to only as 'Mr Ward.' She is quickly 'hypnotized' by the man and agrees to wed. Over the course of their migration west, she experiences many of the turning points in Mormon history, from the persecution and murder of Joseph Smith to the colonization of Utah. The narrator marries and converts willingly, but is horrified by the polygamous practices she witnesses once in Utah Territory. She does not live in a polygamous household, but she illustrates the harsh conditions of the women who do. Throughout the text, she highlights the brutal treatment of those women who protest against the Church:

One poor woman, who had told an emigrant, in the hearing of a Mormon elder, that polygamy was a system of abominations, and who repeated a few of her troubles and sufferings, was taken one night when she stepped out for water, gagged, carried a mile into the woods, stripped nude, tied to a tree, and scourged till the blood ran from her wounds to the ground (Ward, 1855, p. 429)

The narrative also documents rules that plural wives were expected to follow and the punishments inflicted for disobedience. If two wives quarrel, for example, 'the one who commences the quarrel, to receive the punishment, which varies in degree from three lashes to twenty-five' (Ward, 1855, p. 314). *Female Life* concludes with the author's escape from Salt Lake City, but does not explain where she fled to, or what became of her.

4. Maria Ward as Author

The opening chapter of *Female Life* reveals that the author hails from near Ithaca, New York, but a genealogical search for Maria Ward in all of New York during her life yields no results. A search of the genealogical records at Brigham Young University likewise fails to produce a Maria Ward or any

Mormon elders with the surname Ward in Utah during the relevant time period. Other prominent characters appearing in the novel were also absent from the BYU genealogical records.

Maria Ward is also credited as having edited *The Husband in Utah* (Ward, 1857), later reprinted as *Male Life Among the Mormons* (Ward, 1863), presumably to emphasize the parallel with the very popular *Female Life*. The later edition is apparently printed from the same plates: it is identical down to the hyphenation and the signatures, and the running heading remains 'The Husband in Utah.' In her preface to *Male Life*, as we will refer to this book, Maria Ward claims to have gathered and published the letters of her husband's late nephew Austin, and promises that she has not 'changed or omitted a sentence' (Ward, 1857, p. vi). The title page of both editions also explicitly refers to Maria Ward as 'the author of "Female Life Among the Mormons."'

A Library of Congress copyright registry search yields no documentation for Maria Ward, Austin Ward, *Female Life* or *Male Life*. The library database Worldcat holds records for twenty-nine editions of *Female Life*, eleven of which are categorized as fiction. All editions list Maria Ward as the author, but eight list an additional author: Mrs B. G. Ferris. Among the twenty-three Worldcat records for Austin Ward's book, including editions in Danish and Swedish, a few categorize it as fiction, but most as 'description and travel'; one record suggests that Maria and Austin Ward are pseudonyms.

5. Mr and Mrs B. G. Ferris

Mrs B. G. (Cornelia) Ferris, a native of Ithaca, New York, accompanied her husband B. G. (Benjamin) Ferris to Utah from August 1852 to April of 1853 during his tenure as Secretary of Utah Territory. Mr Ferris had completed only six months of his four-year term when they returned to Ithaca, where each of them published a book on Mormonism and Utah. Mr Ferris's *Utah and the Mormons* (B. G. Ferris, 1854) presents a history of Mormonism and Utah, and Mrs Ferris's *The Mormons at Home* (C. Ferris, 1856) is a compilation

of her letters written during their journey to and from Utah and their time there.

Like many Federal appointees to Utah Territory, B. G. Ferris found himself at odds with the Mormon leadership. He was strongly opposed to the practices and leadership of the Church, as is reflected in his book, which chronicles the history of the Mormon Church in a less than favorable light. Soon after his return, members of Congress were consulting with Mr Ferris about Utah and the Mormon government there (House of Representatives, 1854). Ferris later makes his perspective on Mormonism and the Church-run Utah government violently clear in a letter to the *New York Times*: 'The unscrupulous set of villains, congregated in the valley of the Great Salt Lake, would put Pandemonium to shame; and the time has come when their crimes can no longer go unwhipped of justice without national disgrace' (B. G. Ferris, 1857, p. 1).

Mrs Ferris' book did not stir as much controversy as her husband's, but her letters bear an underlying anti-Mormon sentiment throughout, and are sometimes quite harsh. When she meets two of the three wives of a U. S. Marshal, for example, she comments:

What a strange spectacle! Here was an elderly woman, apparently of fair intelligence, and correct notions of propriety, in whom the feelings and instincts of womanhood may be supposed to have become fixed and permanent habits of thought, yielding all that is valuable to a ridiculous system of imposture—in other words, becoming a concubine. (C. Ferris, 1856, p. 138)

A little later, after labeling the Mormons 'worshippers of Beelzebub,' she says, 'Rig up a system that will fasten itself upon the superstitious credulity and animal appetites of the stupidly ignorant, and you are ready for a flourishing business' (p. 142). These sentiments certainly accord well with *Female Life*.

We do not know who first theorized that Mrs Ferris wrote *Female Life*, but one contributing factor may have been that Maria Ward and Mrs Ferris were both from the same part of upstate New York. In any case, as early as 1873, a respected bibliography lists both Ward and Ferris as its

authors (Sabin, 1873). Leonard Arrington, considered one of the twentieth-century's most important Mormon historians, attempted to confirm this theory in the 1950s. After nearly a decade of research, he concluded 'that Mrs Ferris' book, *The Mormons at Home*... provided the basis for *Male Life Among the Mormons*, a sequel to *Female Life*, but that some other Eastern woman, not yet discovered, was the real author of both the 'Maria Ward' books' (Arrington and Haupt, 1968, p. 253). It is tempting to suggest a connection between the attribution of *Female Life* to Mrs Ferris and the fact that the Mormon part of her own book was first published under the heading 'Life Among the Mormons' in *Putnam's Monthly Magazine* (C. Ferris, 1855). The same running heading appears in the book form of *The Mormons at Home* and in the 1872 edition of *Female Life*.

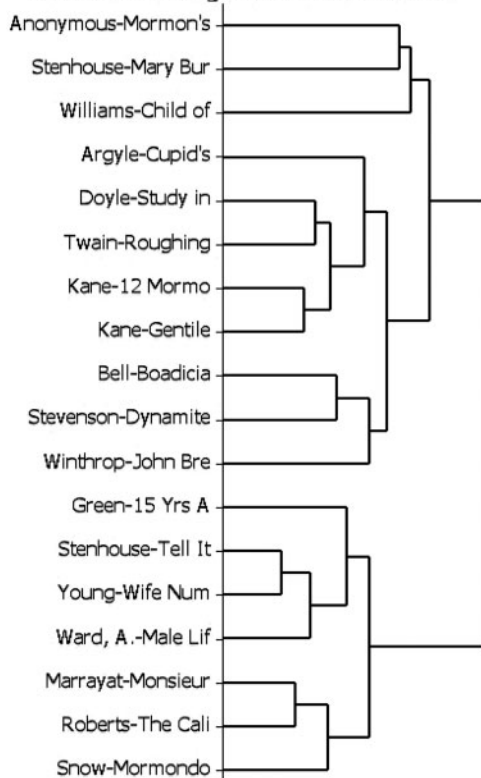
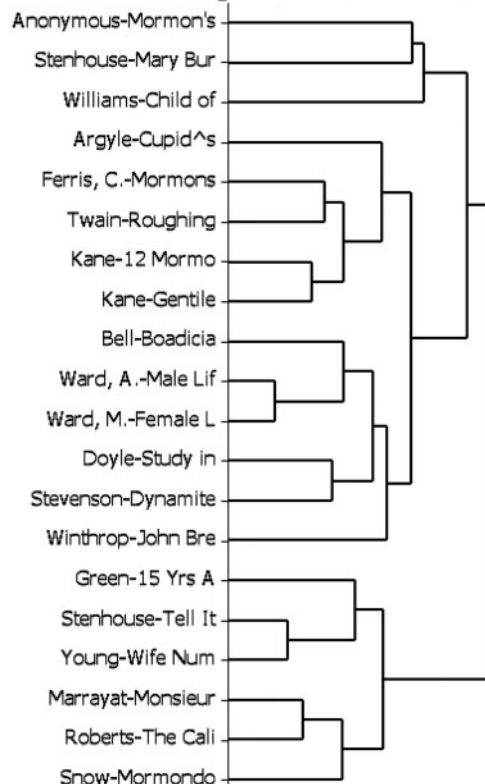
6. A Computational Analysis

A computational investigation of the authorship of *Female Life* presents serious challenges. So far as we can determine from external information, the putative author, the pseudonymous 'Maria Ward,' wrote only this book. Likewise, the suggested author, Mrs Ferris, wrote only *The Mormons at Home*. These conditions do not meet the requirements that are normally set for a valid non-traditional authorship attribution (see, for example, Rudman, 1997), and the situation is made even more difficult by the problems of genre and gender. *Female Life* and *The Mormons at Home* are both putatively non-fictional first-person narratives by women, but the fact that Maria Ward is a pseudonym leaves open the possibility that her book is actually fiction. This possibility is made nearly certain by the presence of some extraordinary errors, including having Joseph Smith shot dead from horseback after helping a Mormon woman break out of jail (see Etzenhouser, 1894, pp. 156–9, for a pro-Mormon account of some other errors). There is also no direct evidence that the author is a woman. In addition, only about the middle third of Mrs Ferris's book actually deals with the Mormons; the beginning and end narrate the trip to and from Utah.

Adding to the complexity are the two related books by Austin Ward and Benjamin Ferris. The most intriguing for comparison with *Female Life* is Austin Ward's *Male Life*, which, as noted above, Maria Ward claims to have edited. Naturally, if 'Maria Ward' is a pseudonym, 'Austin Ward' must also be a pseudonym, leaving open the possibility that Maria Ward actually also wrote *Male Life*. The fictional status of *Male Life* is asserted unequivocally by Sir Richard Burton, who spent three weeks in Salt Lake City in 1860; he calls *Male Life* 'pure invention' (Burton, 1862, p. 207). Second, we have Benjamin Ferris's *Utah and the Mormons* (B. G. Ferris, 1854) for comparison; though it is third-person nonfiction, it is of interest because it slightly predates *Female Life*, *Male Life*, and *The Mormons at Home* and was written by Mrs Ferris's husband. (The fact that Mr Ferris's book tells the true story of Joseph Smith's death casts further doubt on Mrs Ferris's authorship of *Female Life*.) Finally, there is no dearth of contemporary first-person and third-person accounts of Mormon life, some fiction, some non-fiction, and some of the same doubtful status as *Female Life* and *Male Life*.

6.1 Mormon-related first-person texts

Given the complexity of the situation and the lack of definite training texts by Maria Ward and Mrs Ferris, we begin with a group of Mormon-related first-person fiction and nonfiction texts (minimum 7,000 words). Any similarity between Mrs Ferris and Maria Ward will be more significant in the presence of many other candidate authors. In Fig. 1, based on the 800 MFW, only two authors, Kane and Stenhouse, are represented by pairs of texts, and Stenhouse's two texts radically fail to cluster (all analyses from the 500 to the 900 MFW are similar). (All cluster analyses were performed in Minitab (2005), with Ward linkage, squared Euclidean distance, and standardized variables.) However, 'StenhouseMa' is an excerpt from Stenhouse's *Tell it All* that purports to be Mary Burton's story in her own words, so that this cannot be considered a definite error (Mary Burton's story is also only about one tenth as long as the rest of *Tell it All*). Adding Mrs Ferris and Maria Ward (see Fig. 2) produces a dendrogram that provides no support for

800 MFW 18 Large 1st Person Mormon**Fig. 1** Eighteen first-person Mormon-related texts, without M. Ward or C. Ferris**800 MFW 20 Large 1st Person Mormon****Fig. 2** Twenty first-person Mormon-related texts, with M. Ward and C. Ferris

Mrs Ferris's authorship of *Female Life*: the two authors are in widely separated clusters. *Female Life* and *Male Life* are the first two books to cluster, which shows that they are the most similar books in this set, and therefore suggests that Maria and Austin Ward may be the same person.

6.2 First-person and third-person Mormon and non-Mormon texts

We turn, then, to an analysis of thirty-five texts (minimum 7,000 words), including Benjamin Ferris's *Utah and the Mormons*, with two texts by each author wherever possible, though that means comparing texts about Mormons with texts in which they play no part. This allows us to train the method on texts similar to our texts of interest,

select the kind of analysis that is most effective on the texts of known authorship, and then insert Ward and Ferris into the set, testing with the same settings. The known texts by other authors compensate, to some extent, for a lack of training texts by our primary authors.²

The fact that ten of the thirteen pairs of texts by known authors cluster together in Fig. 3, based on the 400 MFW words, in spite of the fact that most of them contain one Mormon and one non-Mormon text, shows that this analysis is capturing known authorship very effectively. (The 400–600 MFW produce the same number of correct clusters for the known authors.) Although Doyle's two texts do not cluster correctly, they are near neighbors, and the other two failures may not really be errors. As before, Mary Burton's story does not

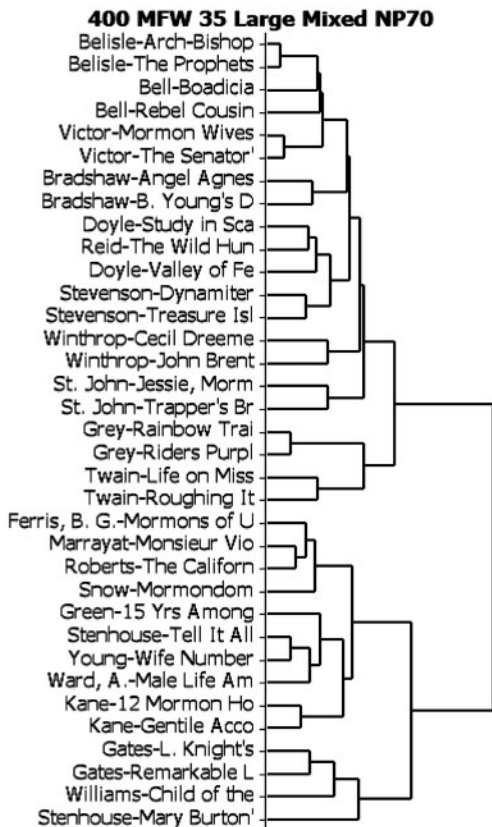


Fig. 3 Thirty-five mixed texts, without M. Ward or C. Ferris

cluster with Stenhouse's section in her own voice. Bell's texts also fail to cluster, but her non-Mormon text, *Rebel Cousins*, is supposedly the autobiography of Bertha Stephens, 'prepared' by Bell, so that this, too, is not definitely an error. Further complications in authorship among these texts exist (for example, *The Dynamiter* was co-written by Stevenson and his wife); some of these will be discussed briefly below. This analysis also shows that, even though nearly all of the 400 words on which Fig. 3 is based are content words, the different subject matter in the Mormon and Non-Mormon texts does not prevent texts by the same author from clustering. Given the large number of texts, their varied subject matter and point of view, and the fact that some are fiction and some non-fiction, these are strong results.

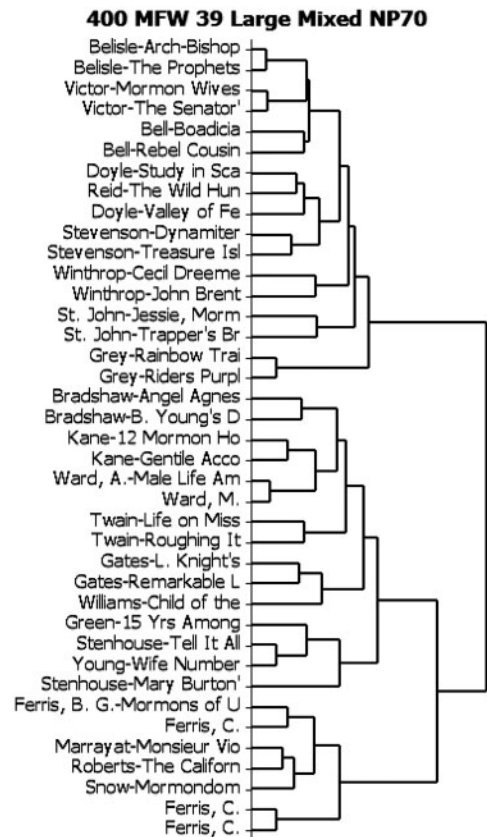


Fig. 4 Thirty-nine mixed texts, with M. Ward or C. Ferris

Adding Maria Ward and Mrs Ferris produces the dendrogram shown in Fig. 4. For this analysis, we have divided Mrs Ferris's *The Mormons at Home* into three sections, so as to isolate the Mormon-related second part. Even with four new texts added, the same ten pairs of texts of known authorship cluster together here as in Fig. 3, and the two Bell texts now also cluster together. Here, as in Fig. 2, Maria and Austin Ward form a very tight cluster that is widely separated from Mrs Ferris. The clustering of the Mormon section of Mrs Ferris's book with her husband's before these join her non-Mormon sections suggests the possibility of some collaboration between Mr and Mrs Ferris. Although analyses based on other numbers of words vary somewhat, most of the known pairs cluster in most analyses and the Wards never cluster with Mrs Ferris.

6.3 A smaller narrative-only Mormon set

Turning in a new direction, we compare only the Mormon section of *The Mormons at Home* with other Mormon texts, limit ourselves to narrative by deleting all dialogue, and divide the texts into sections of about 6,500 words. All analyses based on the 300–800 MFW are very similar to that shown in Fig. 5, based on the 700 MFW: Williams's texts form a separate main cluster, and the Ward texts form a single cluster in which there is some intermingling of sections from the two texts; all other pairs or groups of sections by a single author cluster correctly. In only one analysis, based on the 900 MFW, does the cluster containing Mr and Mrs Ferris's text join in a single larger cluster with the one containing Austin and Maria Wards' texts. Even there, the two clusters are not

very similar, and the analysis based on the 990 MFW returns to the pattern shown in Fig. 5. The upshot of all these tests is that they do not suggest that Mrs Ferris wrote *A Female Life*, but they again suggest that Maria and Austin Ward are the same person. What is especially telling is the mingling of sections by Maria and Austin Ward, which shows that their texts are very similar. Such mingling is relatively unusual even in cases of two texts by a single author.

6.4 Delta analysis

Although Delta analysis is problematic here because of the complexity of the case, a series of analyses with different configurations of primary and secondary authors points in the same direction. Mrs Ferris is never strongly identified as the most likely author for *A Female Life*. She is sometimes identified as the most likely author of *Male Life*, but only when Maria Ward is not included among the primary authors. In fact, when either of the Wards is included among the primary authors, he or she is regularly and strongly identified as the author of the other Ward text in the secondary set.

We report here only one of our tests, a Delta-Lz (0.7) analysis with seventeen primary and seventeen secondary texts—fourteen by authors in the primary set. *Male Life* and *The Mormons at Home* are in the primary and *Female Life* in the secondary set. We analyze the 200, 300, ..., 1,000, 1,500, 2,000, ..., 4,000 MFW, no pronouns, culled at 70% (see Burrows, 2002, Hoover, 2004a,b for discussion). In each of the twelve analyses up to the 2,500 MFW, there are only two or three errors, a total of thirty-three in all, involving just four different texts. There are two errors (600 and 700 MFW) for *Angel Agnes*, three errors (200, 300 and 400 MFW) for *Cecil Dream*, seven errors for *The Blunders of a Bashful Man* (200, 300, 400, 1,000–2,500 MFW), nine errors for the beginning of B. G. Ferris's *Utah and the Mormons* (500–2,500 MFW), and twelve errors for St. John's *The Trapper's Bride* (200–2,500 MFW). Leaving aside the five relatively isolated errors for *Angel Agnes* and *Cecil Dream*, the rest of the errors are intriguing. *The Blunders of a Bashful Man*, a dime novel, was actually published under the

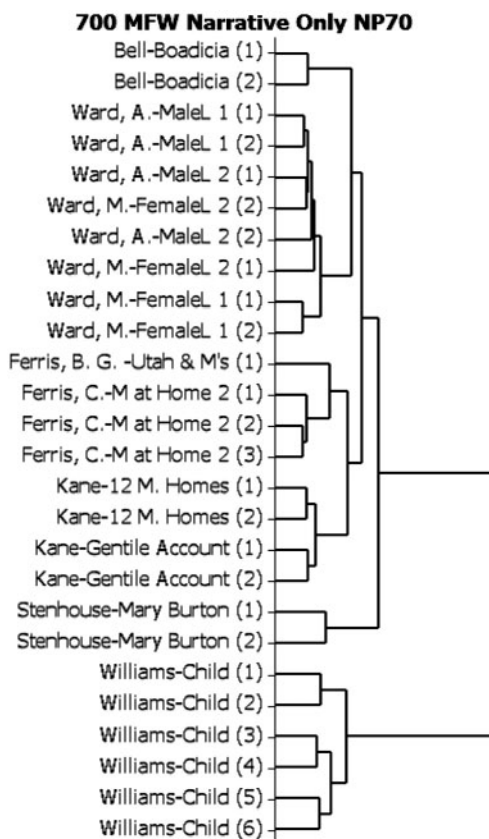


Fig. 5 Nine Mormon-related texts, narrative only

name Walter Gray. Although this is thought to be a pseudonym of Metta Victor, the authorship of such ephemeral novels is very confused and uncertain; her other novel in this analysis, *The Senator's Son*, is universally attributed to her. Furthermore, the attribution of the beginning of B. G. Ferris's book (which relates the journey to Utah and discusses its geography) to Mrs Ferris accords with the cluster analysis in Fig. 4, and may be the result of collaboration. The twelve errors for St. John's *The Trapper's Bride* are even more interesting: this novel is incorrectly attributed in all of our Delta analyses. The attribution to St. John of *Jessie, The Mormon's Daughter* (1861) (his novel in the primary set), however, may rest on a confusion. *Jessie* was anonymously published except for the notation, 'By the author of *The Blue Dwarf*,' and, while St. John did publish a book by this title in 1870, there was an earlier (1861) novel with this title by Lady Esther Hope (also a pseudonym). With the Delta-Lz analysis even more reliable than the results initially suggest, the attribution of *Female Life* to Austin Ward in all analyses from the 200 to the 4,000 MFW, with Delta-Z scores of three or higher in the most accurate analyses, strongly confirms the implication of the cluster analyses. It seems very likely that Austin and Maria Ward are two pseudonyms for the same person, who is not likely to be Mrs Ferris.

6.5 *t*-tests

The results presented above already seem to us strong evidence that Mrs Ferris is not the author of *Female Life* and that Austin and Maria Ward are the same person. Given the non-ideal nature of the problem, however, it seems worthwhile to examine some further approaches using *t*-testing. At first glance, *t*-testing seems problematic for this case. In a classic and often cited article, Burrows (1992) uses *t*-tests to distinguish Austen from James, but he begins his hypothetical authorship problem with the knowledge that the anonymous text is by either James or Austen. This allows him to use *t*-tests on the combined vocabularies of groups of James and Austen texts to identify a set of words that strongly differentiate the two authors—a set that can then be used to test which

author is likeliest to have written the anonymous text. In the present case, we have only three texts, ostensibly by three authors, and only the attribution of *The Mormons at Home* to Mrs Ferris is beyond question.

Nevertheless, we believe that some additional progress in this and other similar difficult cases is possible by performing three different sets of *t*-tests. We select two of our three texts and assume that they are by different authors, use *t*-tests to identify words that strongly differentiate them, and then test the third text against this set of words. We follow the same procedure for the other two possible pairs of texts and compare the results. For this procedure to make sense, however, we need to know what to expect if each set of initial assumptions is correct and what to expect if it is incorrect. We begin, therefore, with sets of known texts for each of the three possibilities, where the three texts are by one, two, or three authors.

6.5.1 *Three texts by the same author*

In this scenario, any assumption of a difference in authorship will obviously be false. Our first test on three texts by a single author involves sections of 13,000–22,000 words of Robert Louis Stevenson's *Kidnapped*, *Treasure Island*, and *The Master of Ballantrae*, roughly the size of the Mormon part of *The Mormons at Home*. (We avoid *The Dynamiter* from our Mormon set because it was jointly written by Fanny Stevenson.) We begin by assuming that *Kidnapped* and *Treasure Island* are by different authors and use *t*-tests to identify the words that are most effective in differentiating the two texts. Dividing the two texts into ten sections of about 1,350 words each and *t*-testing the 3,000 MFW yields forty-six words with $P < 0.05$. We add *The Master of Ballantrae* as an unknown, divide the three novels into 4,000-word sections, and perform a principal components analysis (PCA) based on these forty-six words (see Fig. 6). *The Master of Ballantrae*, though grouped between the test texts, is not obviously distinguished from *Kidnapped*.³ The very small number of discriminating words and the relatively small scale of the graph result partly from the shortness of the texts, but they

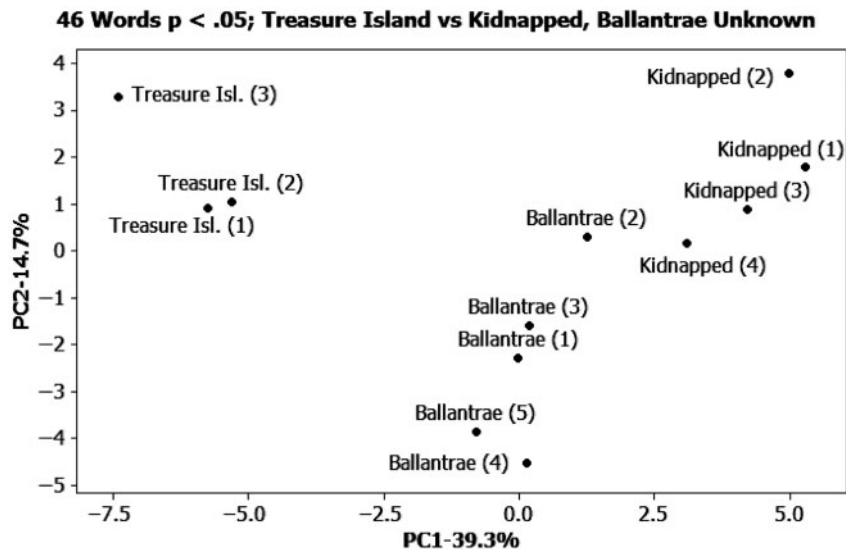


Fig. 6 Stevenson: *Treasure Island* versus *Kidnapped*, with *Ballantrae* unknown

also reflect the fact that the test is being asked to distinguish two texts by the same author, texts that are apparently quite similar.

The second test of three texts by the same author involves David Graham Phillips's *The Conflict*, *The Grain of Dust*, and *The Price She Paid*, in this case using the whole novels (about 100,000 words each) and assuming that *The Conflict* and *The Grain of Dust* are by different authors. Testing the 3,000 most frequent words for discriminators yields 233 words with $P < 0.05$ (thirteen sections of 7,000 words). A PCA of the three novels (10,000-word sections) groups *Price* between the test texts (see Fig. 7). These longer texts produce a larger number of discriminating words and the separation is clearer between the two original novels and the unknown. The range is also somewhat wider, and the novels are separate on both the first and the second component (barely on the latter).

The final test of three texts by a single author involves Glasgow's *The Deliverance*, *Life and Gabriella*, and *Virginia*, using the complete novels (about 125,000–150,000 words) and assuming that *Deliverance* and *Gabriella* are by different authors. Testing the 3,000 most frequent words for

discriminators yields 662 words with $P < 0.05$ (eighteen sections of 7,500 words). A PCA of the three novels (10,000-word sections), shown in Fig. 8, generally separates *Virginia* and *Gabriella* on component two, but, as in Fig. 6, there is considerable intermingling. Adding *The Wheel of Life* to the same analysis as an additional unknown crowds *Virginia* and *Gabriella* closer together, and produces a tight, distinct *Wheel* cluster. There are a great many discriminating words and the scale is quite large on the first component. This analysis suggests that *Deliverance* is quite different from the other two novels and that those two are quite similar to each other. This is confirmed by redoing the analysis assuming *Gabriella* and *Virginia* are different. This analysis yields only 327 discriminator words and a PCA produces a graph very similar to that in Fig. 7.

These results suggest that, if Mrs Ferris wrote *Female Life*, *Male Life*, and *The Mormons at Home*, it may be difficult to tell. With long texts, there is enough variation available so that there can be a large number of discriminators and a large range on the first principal component even when all three texts are by the same author. Furthermore, the results can differ a great deal even if there is

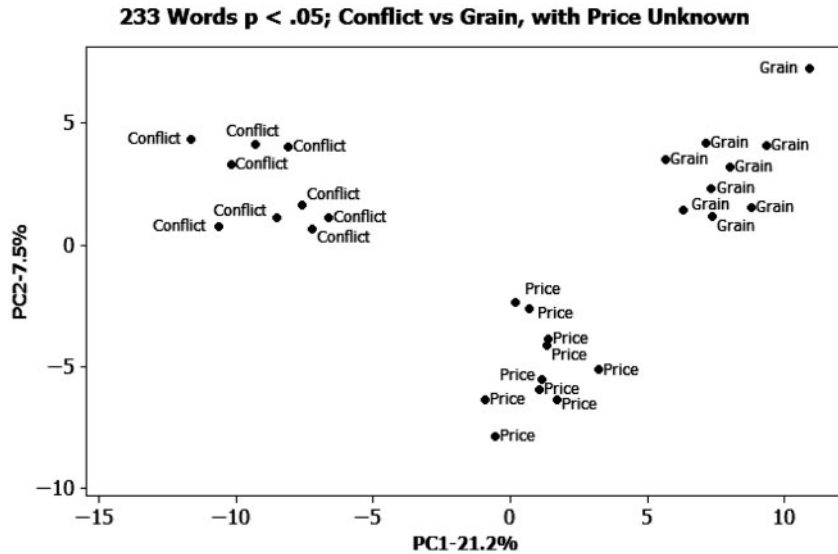


Fig. 7 Phillips: *Conflict* versus *Grain*, with *Price* unknown

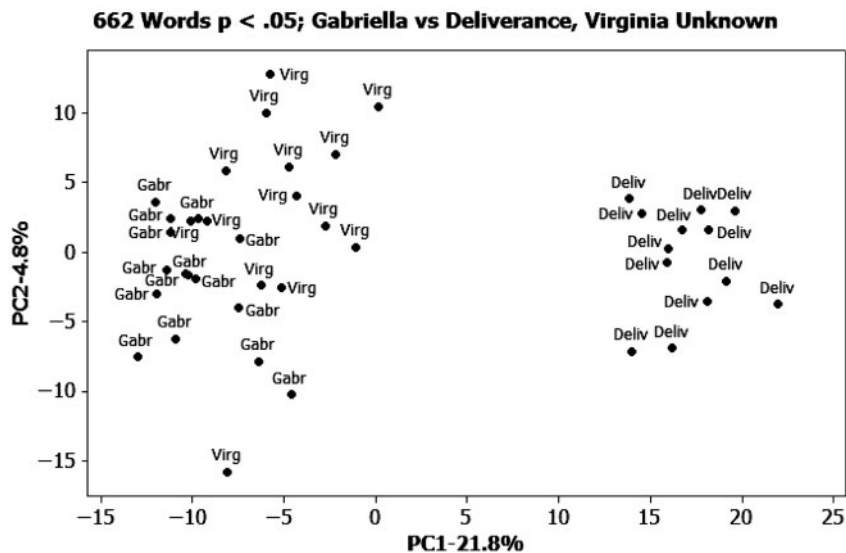


Fig. 8 Glasgow: *Gabriella* versus *Deliverance*, with *Virginia* unknown

only one author involved, depending on how similar the two texts are that are assumed to be by the same author. So far as we know, no one has ever suggested that Mrs Ferris also wrote *Male Life*, but, given the problem of pseudonyms, we must allow for this possibility.

6.5.2 Three texts by two authors

With three texts by two authors, one assumption of difference will be correct and the other incorrect, and the results of our analysis should be recognizably different for the two assumptions (there are actually two different correct assumptions).

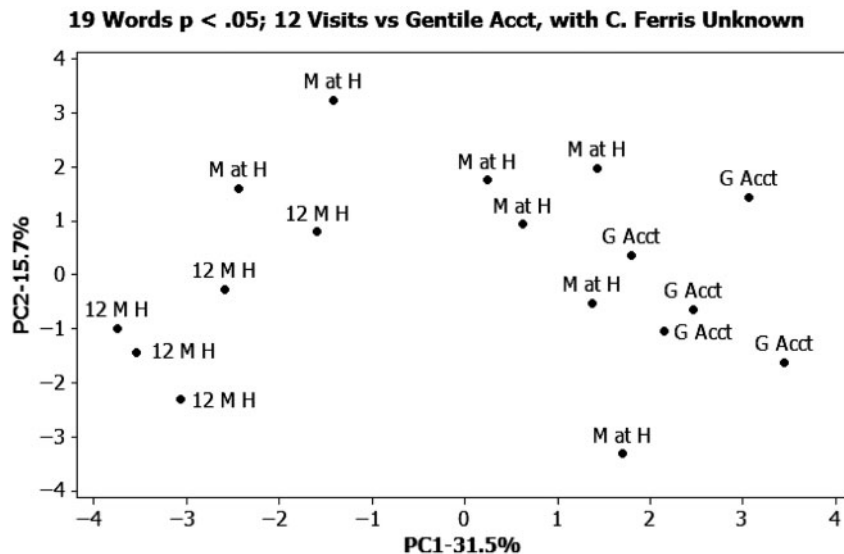


Fig. 9 Kane and C. Ferris: *Twelve versus Gentile*, with *Mormons at Home* unknown

The first test of three texts by two authors involves two texts by Kane, *Twelve Mormon Homes* and *A Gentile Account of Life in Utah's Dixie*, and the Mormon part of Mrs Ferris's *The Mormons at Home*, first with the two Kane texts assumed to be by different authors. Testing the 1,500 MFW yields only nineteen words with $P < 0.05$ (eight sections of 1,600 words). A PCA of these texts (2,600-word sections) groups *The Mormons at Home* between the two test texts but not separately from them (see Fig. 9). Note the small number of discriminating words and the small range on component one. Clearly the shortness of these texts affects these characteristics (the Kane texts are only about 13,000 words each), but the main issue seems to be that the t -tests discover a few words that distinguish two very similar texts by Kane, but these are not sufficient to distinguish Ferris from Kane. Knowing that the tests are based on a false assumption, there is no reason to expect that they should. Using the same texts, but assuming that Kane's *Twelve Mormon Homes* and Ferris's *The Mormons at Home* are by different authors, shows quite a different pattern. Testing the 1,500 MFW yields fifty-five words with $P < 0.05$ (nine sections of 1,500 words). A PCA of these texts (2,600-word sections)

separates Kane's *Gentile Account* distinctly from *The Mormons at Home*, but not from Kane's other text, *Twelve Mormon Homes* (see Fig. 10). Note the larger number of discriminating words and the larger range on component one, showing how much more difference there is between the two authors than between the two Kane texts. The alternate correct assumption, testing Kane's *Gentile Account* against *The Mormons at Home* the same way, yields forty-nine discriminating words; a PCA produces a pattern very similar to that in Fig. 10. (We will not test the second correct assumptions below.)

Our second test of three texts by two authors involves Phillips's *The Conflict* and *The Grain of Dust*, and Maria Ward's *Female Life*, first with Phillips's texts assumed different. Testing the 3,000 MFW for discriminators yields 233 words with $P < 0.05$ (thirteen sections of 7,000 words). A PCA of these novels (10,000-word sections) groups the sections of *Female Life* in a separate cluster, distinct from the two Phillips novels on both the first and second principal component (see Fig. 11). Using the same three texts, we next assume that *The Conflict* and *Female Life* are by different authors. Testing the 3,000 MFW yields 613 words with $P < 0.05$ (fifteen sections of 6,700 words). A PCA of the novels

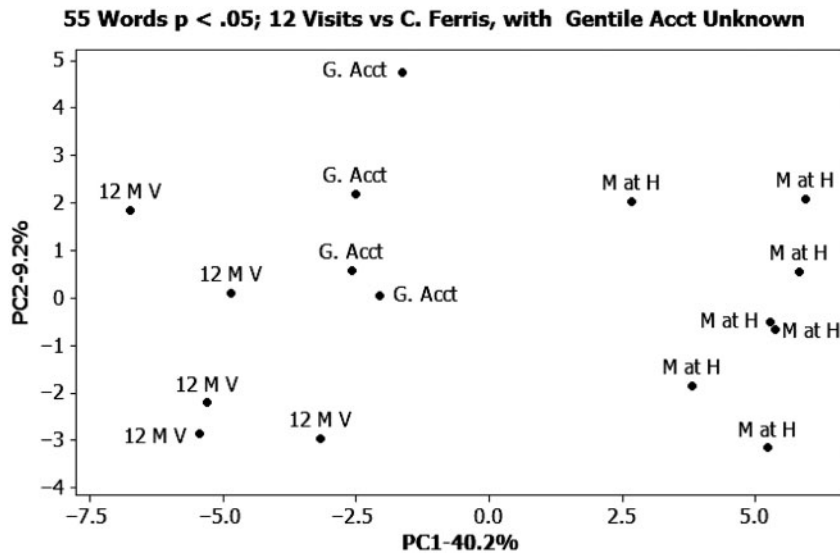


Fig. 10 Kane and C. Ferris: *Twelve versus Mormons at Home*, with *Gentile* unknown

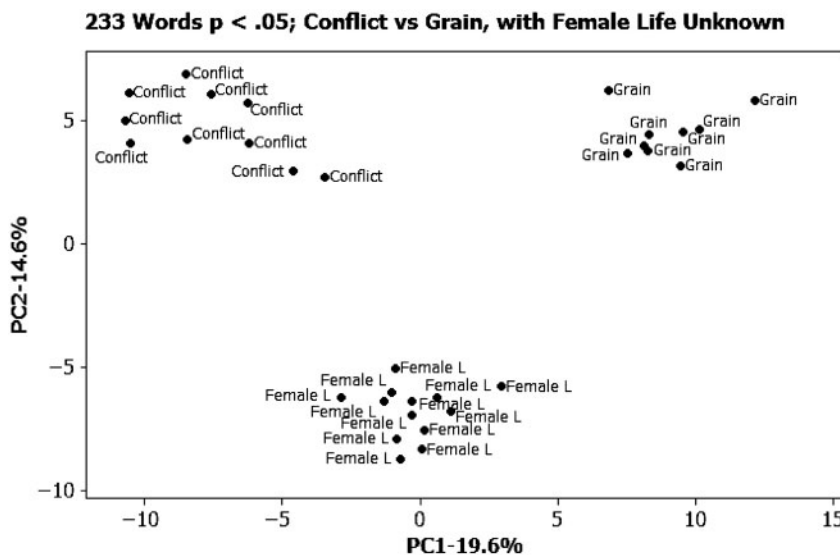


Fig. 11 Phillips and M. Ward: *Conflict versus Grain*, with *Female Life* unknown

(10,000-word sections) widely separates *Female Life* from the two Phillips novels on PC1, but separates the two Phillips novels only on PC2 (see Fig. 12). Note that the correct assumption here produces more than twice as many discriminating words

and a much larger scale on the first component. In spite of the fact that the two Phillips novels were easy to differentiate, the words that distinguish *Female Life* and *Conflict* also do an excellent job of distinguishing *Female Life* and *Grain*.

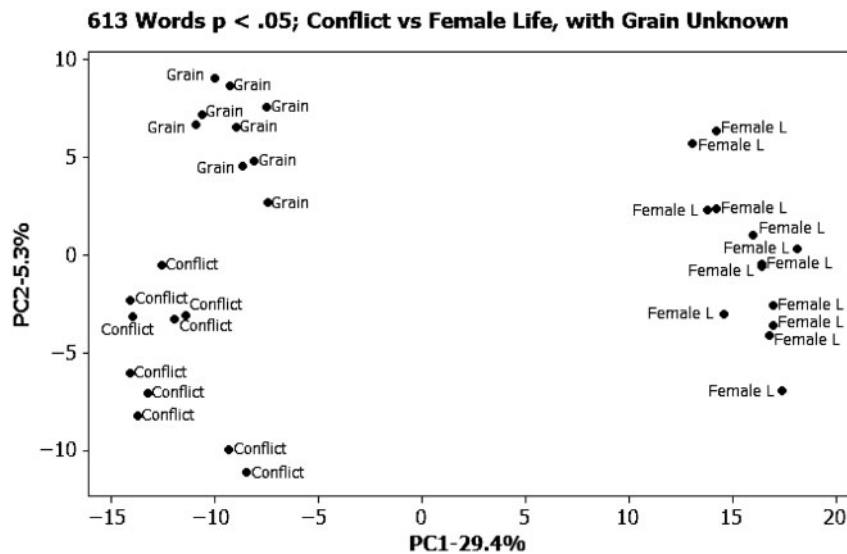


Fig. 12 Phillips and M. Ward: *Conflict* versus *Female Life*, with *Grain* unknown

Our final test of three texts by two authors involves Grey's *The Rainbow Trail* and *Riders of the Purple Sage* and Phillips's *The Conflict*, first with Grey's texts assumed different. Testing the 3,000 MFW for discriminators yields 275 words with $P < 0.05$ (thirteen sections of 7,500 words). A PCA of these novels (10,000-word sections) groups them in three widely separate clusters, with *The Conflict* distinct on both component one and two, as shown in Fig. 13. Note the relatively small number of discriminators and the relatively small range on component one, given the length of the texts. In contrast, testing the 3,000 MFW of *Conflict* and *Riders* for discriminators yields 677 words (more than twice as many as when the two Phillips texts were assumed different) with $P < 0.05$ (fifteen sections each of 6,400 words). Again, a PCA of these novels (10,000-word sections) widely separates Grey's novels from *Conflict*, but the two Grey novels intermingle thoroughly (see Fig. 14), and the range on component one is far larger than in Fig. 13. These results reflect Grey's rather formulaic style and perhaps also the fact that *Rainbow* is a sequel to *Riders*.

The results for all the scenarios in which two of the three texts are by one author suggest that it may

be possible to use information about the size of the list of discriminating words and the range on component one to support an argument for the attribution of the texts. That is, if the texts are tested in two different ways, texts by different authors tend to produce larger numbers of discriminators and a larger range on component one than texts by the same author. A closer look at Figs 9–14 also shows that, when an incorrect assumption of difference is made, there is a much larger proportion of variance in the second component. That is, when two texts by the same author are assumed to be by different authors, testing a third author's text naturally increases the amount of variance captured by component two, which is actually measuring the true authorship difference. When texts by different authors are used for the t -tests and a text by one of them is tested with PCA, a much larger proportion of the total variance is captured in the first, the authorship component, leaving the second component to capture the smaller intra-authorial differences.

6.5.3 Three texts by three authors

In this scenario, an assumption of authorial difference between any two texts will be correct. Our first

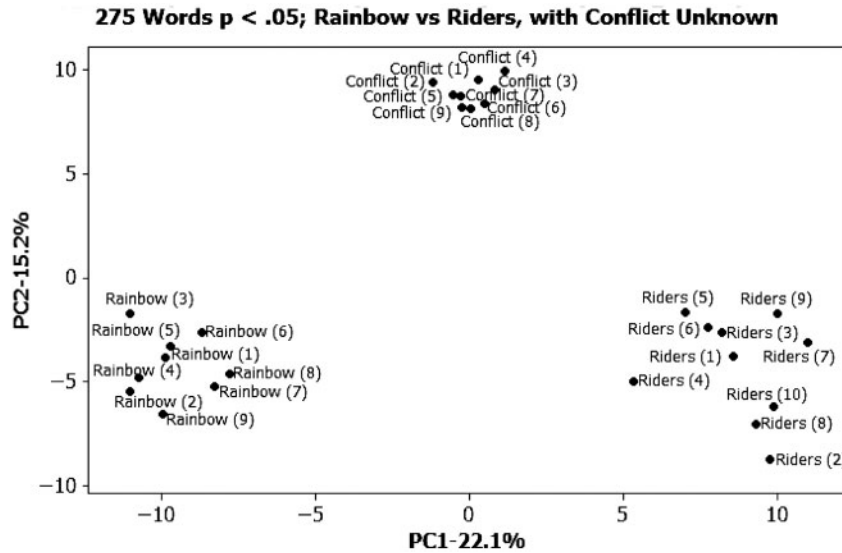


Fig. 13 Grey and Phillips: *Rainbow* versus *Riders*, with *Conflict* unknown

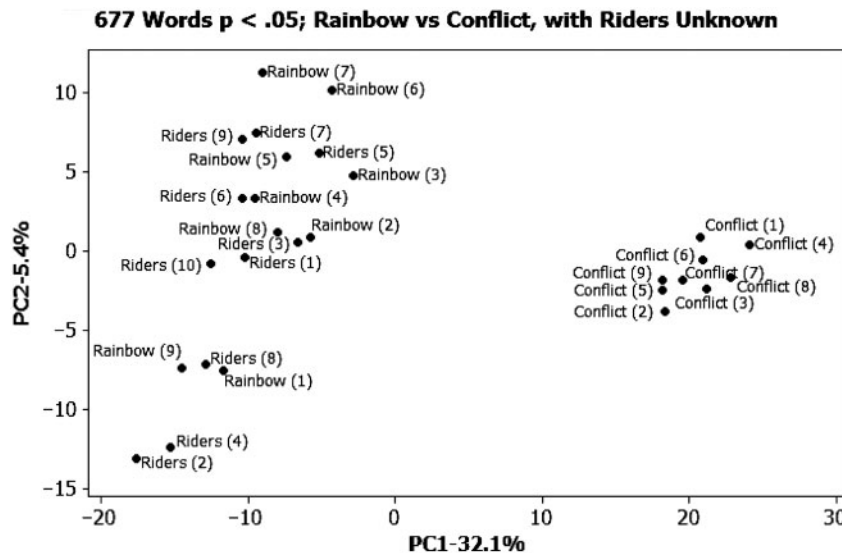


Fig. 14 Grey and Phillips: *Rainbow* versus *Conflict*, with *Riders* unknown

test involves Maria Ward's *Female Life*, Stenhouse's *Tell it All*, and Belisle's *The Prophets*, first with Ward and Stenhouse assumed different. Testing the 3,000 MFW of *Female Life* and *Tell it All* for discriminators yields 681 words with $P < 0.05$ (seventeen sections of 7,800 words). A PCA of these texts

(10,000-word sections) shows *The Prophets* in a distinct cluster between the other two on the first component and very distinct on the second, as shown in Fig. 15. Note again the large number of discriminators and the large range on both components. Testing these three texts in the other two possible

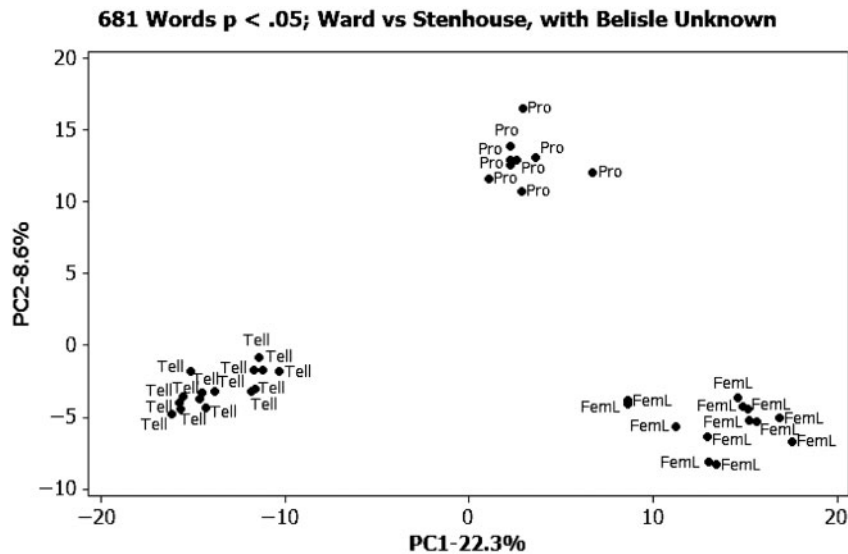


Fig. 15 Ward, Stenhouse, Belisle: *Female Life* versus *Tell It All*, with *Prophets* unknown

configurations gives extremely similar results, with about 600 discriminating words, a large range on both components, and a clear separation of the three texts. Two other sets of three texts by three authors produce the same predictable pattern and need not be discussed further, but a final set involving Mrs Ferris deserves a brief comment. Testing the 3,000 MFW of *The Mormons at Home* against Victor's *Mormon Wives* for discriminators yields 448 words with $P < 0.05$ (thirteen sections of 4,200 words). The pattern shown in Fig. 16 is similar to that in Fig. 15, except that the two sections of C. Ferris that closely correspond to the Mormon part are pulled toward Young, and, to a lesser extent, toward Victor.

6.5.4 Maria Ward, Austin Ward, Mrs Ferris, and the authorship of *Female Life*

We are now ready for tests on *Female Life*, *Male Life*, and *The Mormons at Home*, and we begin by assuming Austin Ward and Maria Ward are different. Testing the 3,000 MFW of *Female Life* and *Male Life* for discriminators yields 224 words with $P < 0.05$ (sixteen sections of 4,300 words). A PCA of these novels (roughly 9,500-word sections)

separates *The Mormons at Home* from *Female Life* on the first component, but from *Male Life* only on the second component, with the Mormon sections very close to *Male Life* (see Fig. 17). By itself, this graph suggests that Ferris might have written or influenced *Male Life*, but it does not suggest any connection between her and *Female Life*.

The other two assumptions of difference studied in the same way produce the very similar results seen in Figs 18 and 19. (Both analyses test the 3,000 MFW in thirteen sections of 4,200 words; each PCA uses sections of about 9,500 words.) In both cases, Ferris is quite separate from the Wards, whose texts intermingle, and the Mormon sections of *The Mormons at Home* are pulled toward the Wards, especially toward *Male Life*.

As we have seen in several analyses above, the Mormon third of Ferris's book naturally shows more affinity to other Mormon-related texts than the travel sections at the beginning and end; therefore, for a final set of tests, we limit ourselves to the Mormon part of *The Mormons at Home* (about 19,000 words). Second, because *The Mormons at Home* is almost wholly narration, while the two Ward novels have a great deal of dialogue,

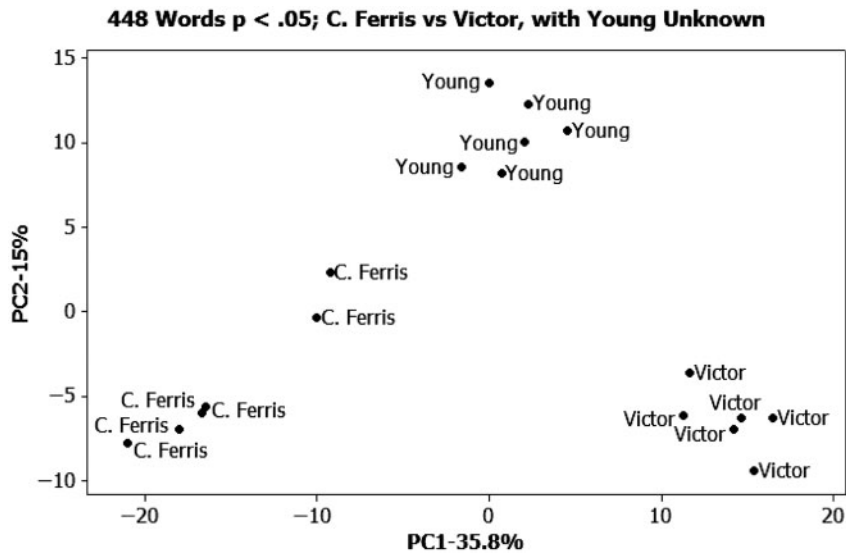


Fig. 16 C. Ferris, Victor, Young: *Mormons Home* versus *Mormon Wives*, with *Wife 19* unknown

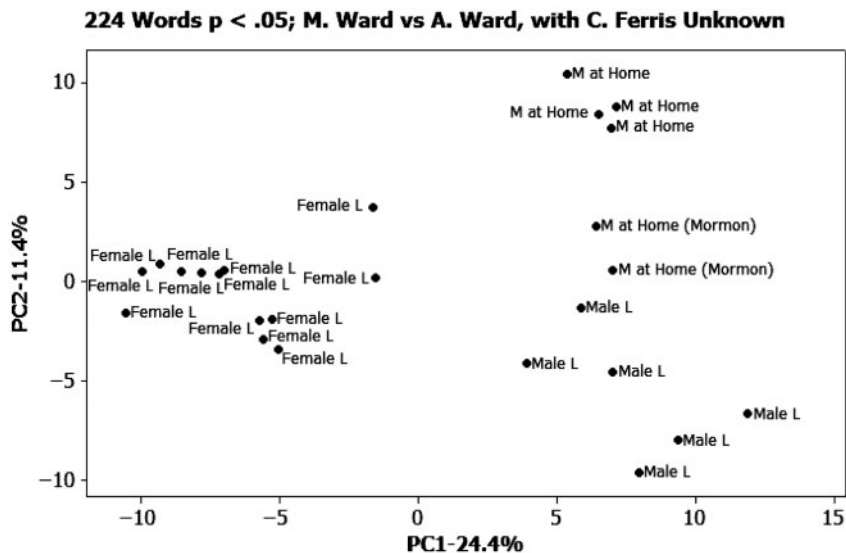


Fig. 17 M. Ward, A. Ward, C. Ferris: *Female Life* versus *Male Life*, with *Mormons Home* unknown

we restrict the analysis to narrative, removing the few hundred words of dialogue from *The Mormons at Home*, and then extracting about 27,000 words of pure narration from *Female Life* and *Male Life* (concentrating on long narrative passages).

When we assume that Maria Ward and Austin Ward are different, testing the 2,000 MFW of *Female Life* and *Male Life* for discriminators yields fifty-two words with $P < 0.05$ (twelve sections of 2,050 words). A PCA of the novels (4,800-word sections) places *The Mormons at Home* in a cluster

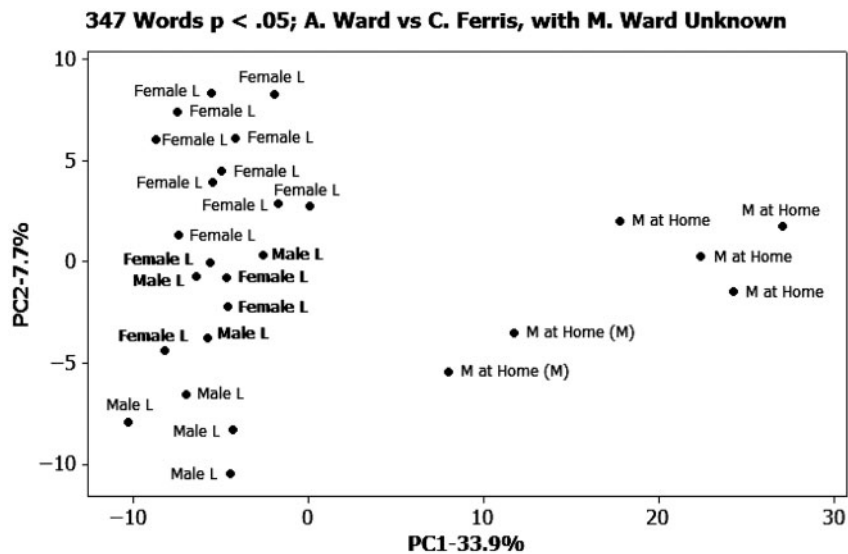


Fig. 18 M. Ward, A. Ward, C. Ferris: *Male Life* versus *Mormons Home*, with *Female Life* unknown

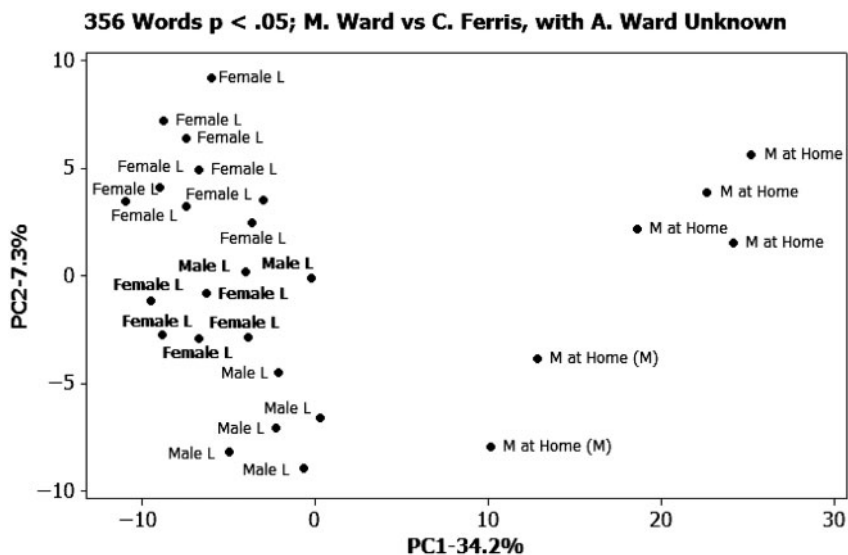


Fig. 19 M. Ward, A. Ward, C. Ferris: *Female Life* versus *Mormons Home*, with *Male Life* unknown

that is distinct from the Ward texts on both components (see Fig. 20). For comparison, if we replace *The Mormons at Home* with Young's *Wife Number 19* and perform a PCA (4,800-word sections) with the same fifty-two words, sections of *Wife Number*

19 overlap slightly with the other two texts on component one and are much less distinct on component two, showing that Mrs Ferris is less like the Wards than Young is, at least with respect to these fifty-two words. The large proportion of variance in

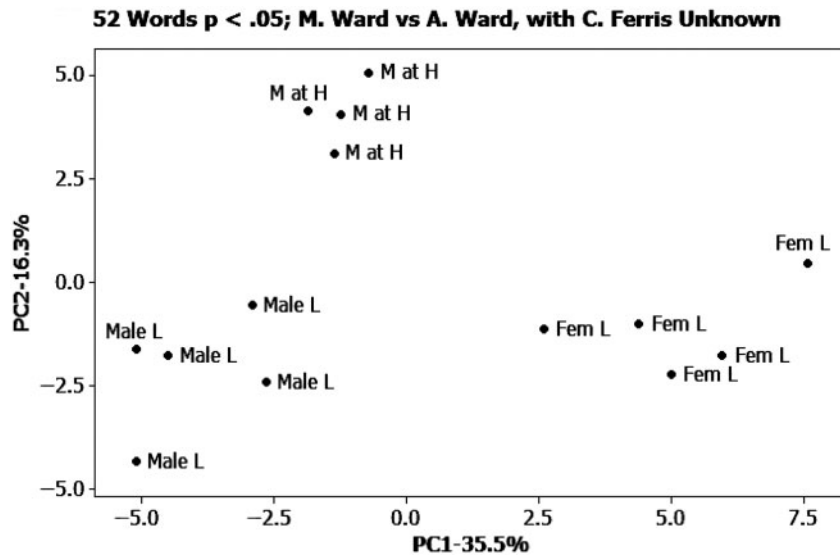


Fig. 20 M. Ward, A. Ward, C. Ferris: Ward versus Ward, with Ferris unknown

the second component in Fig. 20 and the fact that there are only fifty-two discriminator words, even with these relatively long texts, again suggest that the Wards are one person.

Finally, assuming Austin Ward and Mrs Ferris are different and testing the 2,000 MFW of *Male Life* and *The Mormons at Home* for discriminators yields seventy-two words with $P < 0.05$ (eight sections of 2,400 words). For Maria Ward and Mrs Ferris the same kind of analysis yields ninety-nine words $P < 0.05$. For both scenarios, a PCA (4,800-word sections) widely separates Ferris from the Wards, but the sections of *Female Life* and *Male Life* intermingle (see Figs 21 and 22). The larger number of discriminators and the much smaller proportion of variance in the second component compared with Fig. 20 again suggests the Wards are one person and that Mrs Ferris is not likely to be that person. In spite of our removal of two sources of possible non-authorial difference (limiting ourselves to narrative and to the Mormon-related part of Ferris's book, which was pulled toward other Mormon-related texts above), Ferris remains distinct from both Wards, making this analysis seem especially persuasive.

7. Conclusion

We believe that the results above provide very strong evidence that Mrs Ferris did not write *Female Life*, but that whoever did so also wrote *Male Life*. This conclusion is based on the entire series of tests. First, cluster analysis of first-person Mormon-related texts groups Maria and Austin Ward over a wide range of analyses, but does not group Mrs Ferris and Maria Ward. This remains true when we add more texts, both first-person and third-person texts, and both Mormon-related and not. This larger set includes many pairs of texts by the same author, which allows us to select the range of analyses that are most accurate on the known texts for testing our Ferris and Ward texts. In effect, we use pairs of texts by known authors as proxies for the additional texts we would like to have had by our target authors. We also get confirmatory results when test only the narrative (divided into sections) of a smaller group of first-person texts, including our target authors. This analysis consistently groups the sections of texts by the same author, again groups the two Wards together, but does not group Mrs. Ferris and Maria Ward.

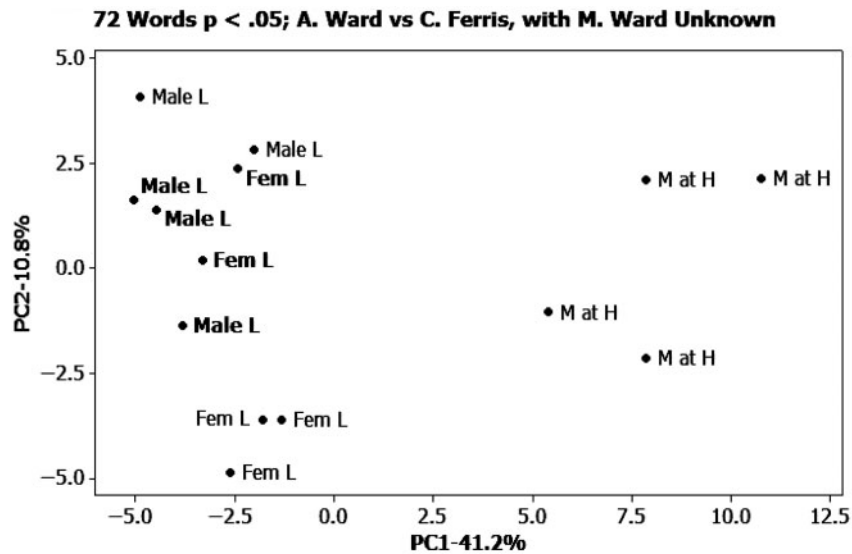


Fig. 21 M. Ward, A. Ward, C. Ferris: A. Ward versus Ferris, with M. Ward unknown

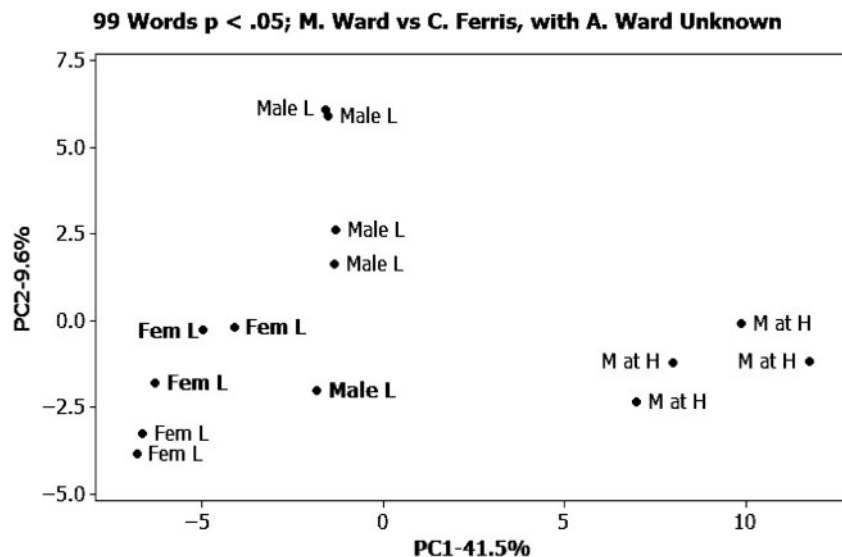


Fig. 22 M. Ward, A. Ward, C. Ferris: A. Ward versus Ferris, with A. Ward unknown

The fact that the sections of the two Ward books intermingle seems to strongly support the proposition that the two Wards are the same person. Delta analysis, in which we test with various configurations of our target authors in the primary and secondary sets, strongly points in the same direction.

Finally, a series of *t*-tests on three texts by one, two, and three known authors with both correct and incorrect assumptions of difference shows that, not surprisingly, *t*-tests based on correct assumptions of authorship difference produce much larger numbers of discriminating words than those based on

incorrect assumptions. Performing PCA based on the various sets of discriminating words shows further that, for any given group of texts, word sets based on incorrect assumptions produce graphs with narrower ranges of values on the first principal component and with much larger proportions of variance in the second component. With three texts, contrasting correct and incorrect assumptions is obviously only possible when there are two authors involved: with one author, all assumptions of difference are incorrect, with three, all are correct. In this non-ideal authorship problem, our three texts strongly mirror the results on known texts involving two authors rather than one or three. However, rather than confirming the long-standing tentative attribution of *Female Life* to Mrs Ferris, our analysis strongly attributes *Female Life* and *Male Life* to a single author who is not Mrs Ferris. However bigoted *The Mormons at Home* shows Mrs Ferris to be, we can confidently, if not conclusively, acquit her of responsibility for *Female Life*. We believe that this kind of approach should be of use in other non-ideal authorship attribution problems.

During the course of our investigation, and because of it, we uncovered one additional strong piece of traditional evidence that Mrs Ferris is not the author of *Female Life*. Some early analyses with texts divided into sections showed that some sections of Fanny Stenhouse's *Tell it All* fail to group with the rest. Besides alerting us to the presence of Mary Burton's story, purportedly in her own words, this discovery led us to look at Stenhouse's book more closely. *Tell it All* is a well-known first-person nonfictional account actually written by an ex-Mormon. It is what *Female Life* only claims to be. Following a preface by Harriet Beecher Stowe, Stenhouse's own preface comments on two previous books claiming to be by ex-Mormons, though without naming them. The first is probably Nelson Winch Green's *Fifteen Years Among the Mormons* (1860). Stenhouse says the second was 'first published nearly twenty years ago,' and describes it as follows:

This book was first published by a New York firm, and being supposed by the innocent public to be genuine, it had an extraordinary circulation—forty or fifty thousand being

sold. The publishers, however, failed, and the stereotype plates passed into other hands. Subsequently the work having come under the notice of a subscription firm at Hartford, they negotiated for the use of the plates. One word of the heading of each page was cut out, a new title was selected, some old illustrations and a few new ones were added, and an ancient steel-plate portrait, which had once done duty in some book of poetry or illustrated volume of fashionable beauties of years ago, was vamped up, and the supposed signature of the fictitious author was engraved beneath it. This book, now re-christened, and *apparently* a new volume, was launched upon the market. It is at the present moment advertised in many local newspapers, and the confiding public cheerfully buy it under the impression that it is the genuine production of a Mormon woman. Such is the history of some of the so-called autobiographies which have appeared. (Stenhouse, 1875, p. xii)

This description perfectly matches *Female Life*, which was published in New York in 1855, 'nearly twenty years' before Stenhouse's book (the original copyright is 1874). It was republished in 1872 and again in 1873 as *The Mormon Wife* by the Hartford Publishing Company (Ward, 1872). Although Maria Ward is still not named explicitly as the author of this edition, her signature does appear, as Stenhouse describes it, under a 'vamped up' picture of her, and the running head (normally the book's title) appears as 'Life Among the Mormons,' with the word 'Female' removed from the original edition's 'Female Life Among the Mormons.' The true author of *Female Life*, and almost certainly *Male Life* as well, may never be known, but the careful detail Stenhouse gives about the later edition of *Female Life* leads us to believe that she herself *did* know who wrote it. We wish that Mrs Stenhouse had followed her own injunction to 'tell it all,' rather than leaving us with the following tantalizing, but still anonymous, attribution of *Female Life Among the Mormons*:

It was professedly written by the wife of a Mormon elder; but it was really the

production of an old lady in New Jersey, who had never even been out to Utah, and who drew entirely upon her own imagination for all that she could not adapt from other sensational writers on Mormonism. (p. xii)

References

- Arrington, L. and Haupt, J.** (1968). Intolerable Zion: the image of Mormonism in nineteenth century American literature. *Western Humanities Review*, **22**: 243–60.
- Bell, A.** (1855). *Boadicea: The Mormon Wife*. Baltimore: A.R. Orton. Available online at <http://www.lettrs.indiana.edu/web/w/wright2/about.html> (accessed 11 January 2008).
- Belisle, O.** (1855). *The Prophets; Or, Mormonism Unveiled*. Philadelphia: Wm. White Smith. Available online at <http://books.google.com/books?id=gOLER7yPy9UC> (accessed 23 February 2008).
- Burrows, J. F.** (1992). Not unless you ask nicely: the interpretative nexus between analysis and information. *LLC*, **7**(2): 91–109.
- Burrows, J. F.** (2002). ‘Delta’: a measure of stylistic difference and a guide to likely authorship. *LLC*, **17**(3): 267–87.
- Burton, R. F.** (1862). *The City of the Saints: And Across the Rocky Mountains to California*. New York: Harper & Brothers. Available online at <http://quod.lib.umich.edu/m/moagrp/> (accessed 17 January 2008).
- Etzenhouser, R.** (1894). *From Palmyra to Independence*. Independence, MO: Ensign Pub. Available online at <http://www.archive.org/details/frompalmyranewyo00etzerich> (Accessed 28 January 2008). Also available at <http://www.sidneyrigdon.com/1894Etz2.htm#pg151a>.
- Ferris, B. G.** (1854). *Utah and the Mormons*. New York: Harper & Brothers. Available online at <http://www.archive.org/details/utahmormonshisto00ferrrich> (accessed 25 January 2008).
- Ferris, B. G.** (1857). Affairs in Utah; condition of affairs among the Mormons—interesting letter from Hon. B. G. Ferris. *New York Daily Times*, May 23: 1. Available online at ProQuest Historical Newspapers, Bobst Library, NYU; cited in Furniss, Norman F., *The Mormon Conflict: 1850–59*. Cambridge: Yale University Press, 1960: 39 n 43.
- Ferris, C.** (1855). Life among the Mormons. *Putnam’s Monthly Magazine*. 6.32–6.36. Available online at http://lcweb2.loc.gov/ammem/ndlpcoop/moahtml/title/putn_vols.html (accessed 31 March 2009).
- Ferris, C.** (1856). *The Mormons at Home*. New York: Dix & Edwards. Available online at http://www.archive.org/stream/mormonsathomewit00ferrrich/mormonsathomewit00ferrrich_djvu.txt (accessed 8 January 2008).
- Gordon, S. B.** (2002). *The Mormon Question: Polygamy and Constitutional Conflict in Nineteenth-Century America*. Chapel Hill, NC: University Of North Carolina Press. Available online at <http://ezproxy.library.nyu.edu:6305/lib/nyulibrary/edf.action?p00=&docID=10022612> (accessed 27 March 2009).
- Hoover, D. L.** (2004a). Testing Burrows’s Delta. *LLC*, **19**(4): 453–75.
- Hoover, D. L.** (2004b). Delta prime? *LLC*, **19**(4): 477–95.
- Hoover, D. L.** (2008). 1856 Mormon tale. Expert Commentator. *The History Detectives*, Season 6, Episode 1. LION Productions, PBS, June 30, 2008.
- House of Representatives** (1854). The History Center in Tompkins County (New York). Letter from House of Representatives to Benjamin Ferris.
- H. R. 433** (1856). *Journal of the House of Representatives of the United States*, Friday, June 27, 1856: 1120.
- Minitab** (2005). Release 14. State College, PA: Minitab, Inc.
- Review** (1855). Review of *Female Life Among the Mormons*, *New York Times*, July 14, 1855: 3. Available online at ProQuest Historical Newspapers, Bobst Library, NYU.
- Rudman, J.** (1997). The state of authorship attribution studies: some problems and solutions. *Computers and the Humanities*, **31**: 351–65.
- Sabin, J.** (1873). *Dictionary of Books Relating to America. From its Discovery to the Present Time*. Vol. VI. New York: J. Sabin & Sons. Available online at http://ia311219.us.archive.org/1/items/dictionaryofbook06sabi/dictionaryofbook06sabi_bw.pdf (accessed 2 April 2009).
- Stenhouse, F.** (1875). *Tell it All: The Story of a Life’s Experience in Mormonism*. Hartford, Conn.: A. D. Worthington. Available online at <http://quod.lib.umich.edu/cgi/t/text/text-idx?c=moa;idno=ANT6396> (accessed 8 January 2008).
- Victor, M.** (1856). *Mormon Wives: A Narrative of Facts Stranger Than Fiction*. New York: Derby & Jackson.

Available online at <http://www.lettrs.indiana.edu/web/w/wright2/> (accessed 11 January 2008).

Ward, A. (1857). In Maria, Ward (ed.), *The Husband in Utah*. New York: Derby & Jackson. Available online at <http://books.google.com/books?id=FucQAAAAIAAJ&oe=UTF-8> (accessed 8 January 2008).

Ward, A. (1863). In Maria, Ward (ed.), *Male Life Among the Mormons*. Philadelphia: John E. Potter and Company, 1863. Available online at <http://www.archive.org/details/malelifeamongmor00wardrich> (accessed 8 January 2008).

Ward, M. (1855). *Female Life Among the Mormons; A Narrative of Many Years' Personal Experience. By the Wife of a Mormon Elder, Recently from Utah*. New York: J. C. Derby. Available online at <http://www.archive.org/details/femalelifeamongm00wardrich> (accessed 8 January 2008).

Ward, M. (1872). *The Mormon Wife; A Life Story of the Sacrifices, Sorrows and Sufferings of Woman. A Narrative of Many Years' Personal Experience*. Hartford, Conn.: Hartford Pub. Co. Available online at <http://pds.lib.harvard.edu/pds/view/2801633?n=4&imagesize=1200&jp2Res=.25> (accessed 28 January 2008).

Worldcat. (2001–2008). Dublin, Ohio: OCLC Online Computer Library Center, Inc. Available online at <http://www.worldcat.org/>.

Woodworth-Ney, L. (2008). *Women in the American West*. Santa Barbara, Calif.: ABC-CLIO <http://books.google.com/books?id=4eu0WhwUJVYC> (accessed 27 March 2009).

Notes

- 1 An earlier stage of this research was presented on an episode of the PBS series *The History Detectives* (Hoover, 2008).
- 2 The order of presentation does not follow that of the investigation itself. We began with a small set of texts, ran preliminary tests, gradually discovered and added more texts, and eventually removed some of the shortest texts. The results presented here thus grew out of dozens of earlier tests that will not be discussed; all of the earlier results, however, were generally compatible with those described here.
- 3 We divide the shortest text as evenly as possible into at least eight and typically twelve or more sections to give a reasonable number of sections of at least 1,200 words on which the t-tests can operate. We then divide the other text into sections of the same size. If one of the two texts is significantly longer, we drop the final section (which contains any remainder) and as many other sections as necessary (at random) to arrive at an equal number of sections for each text. For longer texts, we use longer sections and test larger numbers of words. The sizes and numbers of sections into which the texts are divided for PCA are arbitrary to some extent; we have aimed at a readable graph with enough sections to allow for some scatter. In most of the PCA analyses presented here, the first two principal components account for less than half the total variance; such results must be interpreted with caution.