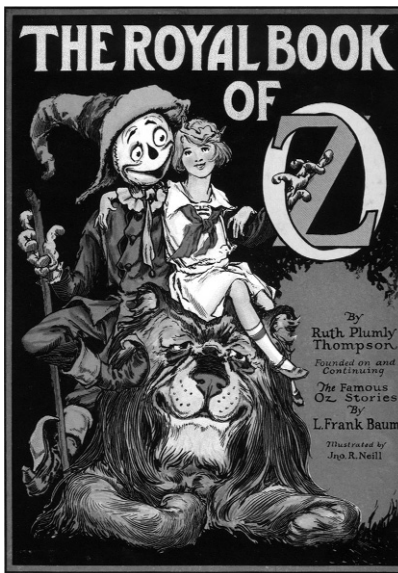
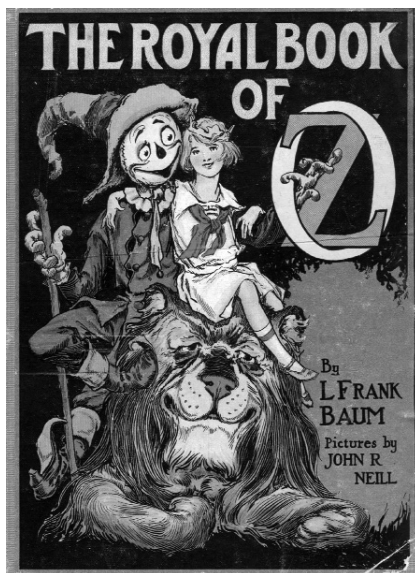

Does only “The Oz” himself know who wrote
The Royal Book of Oz?

Who Wrote the 15th Book of Oz? An Application of Multivariate Analysis to Authorship Attribution

José Nilo G. Binongo



Who is the author of *The Royal Book of Oz*?

The Royal Historians of Oz

Lyman Frank Baum (1856–1919), according to Martin Gardner, was “America’s greatest writer of children’s fantasy.” “His *Wonderful Wizard of Oz* has long been the nation’s best known, best loved native fairy tale” (Gardner and Nye 1957, p. 19). Although initially rejected by publishers in Chicago,

it became a children’s best seller by Christmas of 1900. Such an instant success led biographer Russell MacFall to write a chapter about 1900, calling it Baum’s *annus mirabilis* (Baum and MacFall 1961).

After writing the first Oz book, Baum could not separate himself from the fantasy land he created. In less than two decades, he produced a series con-

sisting of 14 books. He consequently earned the title, “The Royal Historian of Oz.” Nye describes the scenario in the early 1900s:

The Wizard was apparently written with no intention of supplying a sequel; it is a complete unit, with nothing in it to anticipate a successor, much less thirteen of them. . . . [Baum did try] to end the series in 1910 with *The Emerald City of Oz*, but he was driven back to Oz by the demands of his readers...

Finally, promising that “as long as you care to read them I shall try to write them,” he resigned himself to at least one Oz story each year. (Gardner and Nye, p. 5)

Even in frail health, Baum tried hard to keep that promise. Carpenter and Shirley (1992, p. 117) write the details:

In 1918, Frank agreed to have his gallbladder removed. By this time he had written two extra Oz books, *The Magic of Oz* and *Glinta of Oz . . .* They were to be published if he became too ill to write a new book each year. .

. . . [After his surgery, he] tried to adjust to his new way of life by writing a little each day. Propped up on pillows, he finished *The Tin Woodman of Oz* for a 1918 publication date.

Unfortunately, Baum's condition took a turn for the worse. On May 5, 1919, he suffered a stroke. "Baum's generous heart, unlike the fine velvet heart of the Tin Woodman, was not replaceable" (Gardner and Nye, p. 40). He passed away the next day. His last words were: "Now we can cross the Shifting Sands" (Baum and MacFall, p. 275).

Interpreting a conversation between Ozma and Dorothy in Baum's last book, Riley (1997) opines that "if Baum had been able to write more stories, there would have been more plots involving Ozma settling disputes and extending her benevolent rule to all the undiscovered corners of Oz" (pp. 222–223). With Baum's death, the publishers Reilly & Lee (formerly Reilly & Britton) had to find someone to continue writing tales about Ozma's reign. The motive appeared to be commercial: nineteen years after the publication of the first book, the series was still very popular and profitable. Reilly & Lee negotiated a compromise with the widow Maud Baum, promising "a fixed royalty for her and Baum's heirs on every Oz book, no matter who was the author" (Baum and MacFall, p. 277).

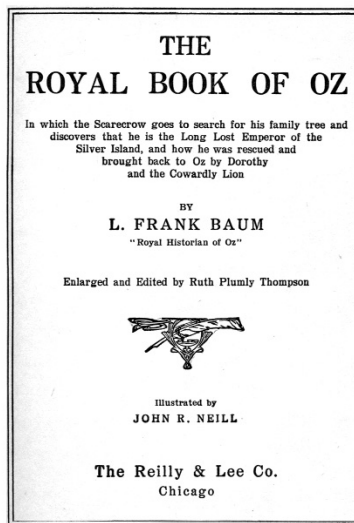
Ruth Plumly Thompson (1891–1976), an established children's writer, took on the job as Baum's successor. Although 35 years younger than Baum, Thompson started her career at 23, when she wrote a page for children every week in the Philadelphia *Public Ledger*. Reilly & Lee were not disappointed: Thompson, like her predecessor, was a prolific writer. Every Christmas season from 1921 to 1939, readers enjoyed a new Oz adventure by Thompson. According to Riley (1997, p. 232), for many American children of the time, Oz and Santa Claus were twin images of Christmas. The classic MGM movie with Judy Garland was released in 1939. By that year, Thompson had published the 33rd Oz book.

The Royal Book of Oz

The 15th book in the series is now believed by many to be Thompson's first work. Moore (1974, p. 89) reports:

Notes and a fragmentary draft of . . . *The Royal Book of Oz* were presumably turned over by his publisher to a successor, Ruth Plumly Thompson, but no one seems to know exactly how much of this book was really Baum's work. . . .

The complication started in 1921, the year the 15th book was published. Baum's name was on the cover, and Thompson was acknowledged only as having "enlarged and edited" the work.



Title page of Oz sequel written by Thompson

In her letter after the title page, Mrs. Baum explained that the book was based on "some unfinished notes" her husband had left. Three decades later, Oz chronicler Jack Snow (1954) wrote that the 15th book was, in reality, Thompson's own work. The current predominant view sides with Snow. Hearn (2000, p. lxxxv), for example, asserts that

Thompson did *not* base the story on any notes Baum left behind as stated in the introduction. *The Royal Book of Oz* was entirely her own work. Thompson did not slavishly imitate what Baum had done, but instead built on his creation.

The story about Baum's unfinished notes seems to have been fabricated by the publishers in an effort to ease the transition and, presumably, to guarantee sales. With the Del Rey edition in 1985, Thompson was identified as the author of the book. Following her name on the title page was the qualification, "Founded on and continuing the famous Oz stories by L. Frank Baum."

In 2001, a Dover edition of the book was released. As noted by the publisher, this is "an unabridged republication of the [1921] work." As such, this edition credits Baum as the author on the front and back cover. However, the explanation on the back cover is not very clear:

THE ROYAL BOOK OF OZ L. Frank Baum

Before his death, L. Frank Baum, the "Royal Historian of Oz," left behind notes for more delightful stories about the adventures of Dorothy and her friends in Oz. This charming sequel, based on those notes and written by Ruth Plumly Thompson, fully captures the excitement, imaginative spirit, and playfulness of Baum's own work.

Although Baum's name follows the title of the book, Thompson is acknowledged as having written the work.

Can statistics help shed light on the authorship of the 15th book? That is, is it possible to show statistically that the writing style in the 15th book is, indeed, more like Thompson's than Baum's? Critics like Riley (1997) claim that "[Thompson's] writing style was quite different from Baum's" (p. 233). Can we actually show using statistical tools that each author used a consistently distinct style when narrating an Oz story?

A Non-Traditional Method of Attributing Authorship

Albeit described by humanities scholars as "non-traditional," the use of quantitative techniques to resolve the authorship of a disputed text has been around for more than a century. Binongo and Smith (1996) traced one

of the first uses of the word “stylometry” in the work of Wincenty Lutoslawski (1897). Calling his method “stylometric,” Lutoslawski attempted a computational procedure to establish a relative chronology of Plato’s dialogues. Reflecting on his approach, he wrote:

this . . . new science of style . . . will enable us to decide questions of authenticity and chronology of literary works. . . . This future science of stylometry may improve our methods beyond the limits of imagination. (p. 193)

Lutoslawski’s words proved prophetic. Progress in stylometry has been made possible through technological advancement. Even the early work of Mosteller and Wallace on the authorship of “The Federalist Papers” relied on technology that was available in the 1960s. With the proliferation of electronic texts on the Internet and in libraries all over the globe, some researchers have considered looking at literary style from a quantitative viewpoint.

In the present investigation, we differentiate between literary styles using a graphical approach. In particular, we will demonstrate graphically that the works of Baum and Thompson are diametrically different, even as they write about the same fairy land. The same tool is then used to decide which of the two Historians of Oz is more likely to have written the 15th book. While it is true that stylometric methods do not rely on subjective impressions, quantification does not always produce results that can be trusted. Thus, prior to the attribution proper, it is important to check the reliability of the proposed method using works known to have been written by either author.

Preparing the Data Set

All of Baum’s 14 books and the 15th book can be downloaded from the Internet. Project Gutenberg’s website has machine-readable versions of texts that are no longer under copyright. As for Thompson, only seven of her books, as of this writing, are in the public domain. Thus, copies of the books were purchased and scanned using charac-

- ★ Although deemed aesthetically uninteresting by many literary scholars, writers cannot avoid using function words to construct even the most basic sentences. Function words constitute the skeleton around which the body of any text is built.
- ★ They belong to a closed class of words: this class does not admit new vocabulary as language evolves. Neither do the words in this class easily become archaic; instead, they remain part of the language for several generations.
- ★ With little semantic meaning, they are least dependent on context.
- ★ With the exception of auxiliary verbs and pronouns, a number of them are not inflected and thus appear in one form.
- ★ In short, they are typically irreplaceable, are much more frequent, more reliable, and more stable than content words.
- ★ They are of interest to researchers particularly because they are not easily affected by a writer’s conscious use of the language. Their usage may therefore reveal idiosyncratic patterns in a writer’s style.

Figure 1. Why function words?

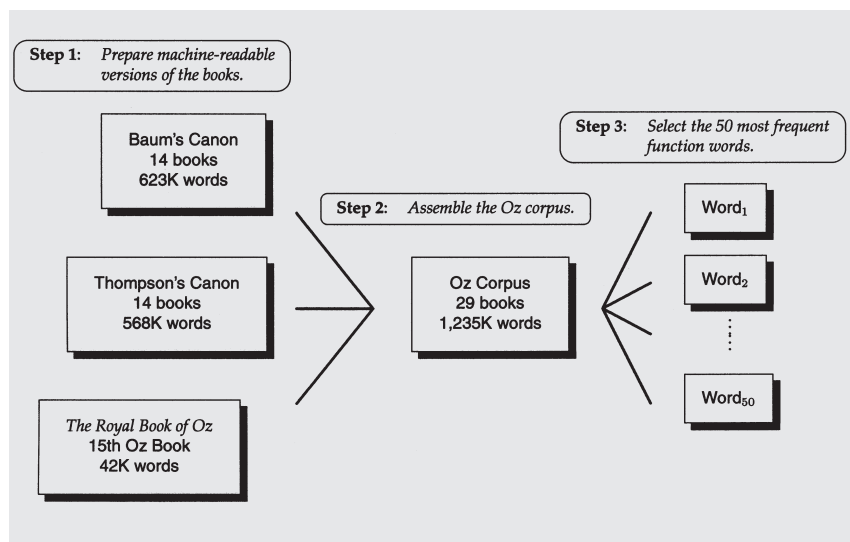


Figure 2. Finding 50 variables to describe each text.

ter recognition software. Proofreading both the downloaded and scanned texts was at once the most time-consuming and the most enjoyable part of the study, as it provided the opportunity to reread the timeless tales. The texts were edited in that whenever an author (particularly Baum) altered the spelling of a character’s words to suggest a peculiar accent, the misspellings were corrected. When a particular sentence contained many misspellings, the sentence was removed. Additionally, portions in the narrative that were not written in prose (e.g., the Patchwork Girl’s idiolect) were excluded.

Many of the books are out of print. Even after hopping from one library to another, four of Thompson’s books—*The Yellow Knight of Oz*, *The Purple Prince of Oz*, *Ojo in Oz*, and *Speedy in Oz*—were still missing. The recent

reprinting of all of Thompson’s books by the Books of Wonder should allow the inclusion of the missing works in a future study.

A computer program was then written to count the number of occurrences of each of the words used. Of these words, function words of the highest frequency were selected. The reasons for choosing them (and not content words) are listed in Figure 1. Among the parts of speech, function words are made up of pronouns, auxiliary verbs, prepositions, conjunctions, determiners, and degree adverbs. These parts of speech have a more grammatical than lexical function. Nonetheless, while the preposition “of” has little semantic content, the preposition “below” signifies position and can alter meaning. To prevent the latter type of words from being considered, only the 50 most fre-

the (6.7%)	with (0.7%)	up (0.3%)	into (0.2%)	just (0.2%)
and (3.7%)	but (0.7%)	no (0.3%)	now (0.2%)	very (0.2%)
to (2.6%)	for (0.7%)	out (0.3%)	down (0.2%)	where (0.2%)
a/an (2.3%)	at (0.6%)	what (0.3%)	over (0.2%)	before (0.2%)
of (2.1%)	this/these (0.5%)	then (0.3%)	back (0.2%)	upon (0.1%)
in (1.3%)	so (0.5%)	if (0.3%)	or (0.2%)	about (0.1%)
that/those (1.0%)	all (0.5%)	there (0.3%)	well (0.2%)	after (0.1%)
it (1.0%)	on (0.5%)	by (0.3%)	which (0.2%)	more (0.1%)
not (0.9%)	from (0.4%)	who (0.3%)	how (0.2%)	why (0.1%)
as (0.7%)	one/ones (0.3%)	when (0.2%)	here (0.2%)	some (0.1%)

Figure 3. Fifty stylistic variables.

quent function words qualified. In calculating each word's average rate of occurrence, the books were assigned equal weight, irrespective of their lengths. In a corpus populated by over a million words, those on top of the frequency spectrum were the least context-dependent words.

Figure 2 shows a schematic diagram of how the 50 function words were obtained. The qualifying words are displayed in Figure 3. (The numbers in parentheses are the average rate of occurrence.) Auxiliary verbs are noticeably absent from the list: they are a bit difficult to handle in that one root can take more than one inflectional suffix. Personal pronouns are also missing: they undergo even more inflections, depending on case, number, person, and gender. Another reason for their exclusion is that the sex of the characters can confound the outcome.

The English language is replete with aural ambiguities. The words "for" and "but," for instance, are both prepositions and conjunctions. Likewise, "that" functions both as a subordinating conjunction and as a demonstrative pronoun. The word "to" is used both as an infinitive and as a preposition. In this study, the different meanings of such homographs are not separated.

Finally, high frequency words are generally monosyllabic and unstressed and are thus prone to contraction. Contractions which involved one of the 50 function words were expanded. For example, "'twasn't" was changed into "it was not." Admittedly, some contractions could be separated in more than one way: "he'd," for instance, depending on the context, could mean, "he would" or "he had." Because counts for auxiliary verbs were not included in

the analysis, it did not matter which expansion was chosen. The word "cannot," albeit not technically a contraction, was also broken into two words. Words like "one's" or "somebody's" were left untouched: a crude computer program, incapable of examining the context, could not tell the difference between a legitimate contraction and a possessive adjective.

With the automated expansion of contractions pertinent to the study, the electronic versions of the Oz books have more words than the printed versions. To illustrate, *The Wizard of Oz* gained some 200 words, an increase of 0.5% in its original length.

As illustrated in Figure 4, each book was partitioned into blocks of 5,000 words. Dividing books into smaller parts gave us a chance to examine within-book variability. The initial data set was a matrix of 223 text blocks and 50 words. Because the variability of the words differed greatly, standardization

was performed. This transformation reduced the variance of each variable to unity.

The Method of Dimensional Reduction

The text blocks are each described by 50 variables. Located in hyperspace, it is thus not possible to visualize their positions. However, using the multivariate statistical technique of principal component analysis (PCA), we can take a glimpse of their relative positions in 50-dimensional space.

Essentially, what happens within our application of the procedure is the reduction of 50 dimensions into 2. The best 2-dimensional approximation is found by rotating the original 50 axes to new axes so that the latter represent directions of decreasing variability. The scalar projections of the text blocks on the rotated axes, which are the coordinates of the text blocks with respect to the new orthonormal basis, are referred to as their PC scores. Interestingly, the new axes also define the plane for which the sum of the squared perpendicular distances of the text blocks from the plane is minimized. Indeed, PCA satisfies a few optimal algebraic and geometric properties; the interested reader is referred to Jolliffe's reference (1986). The particular application of

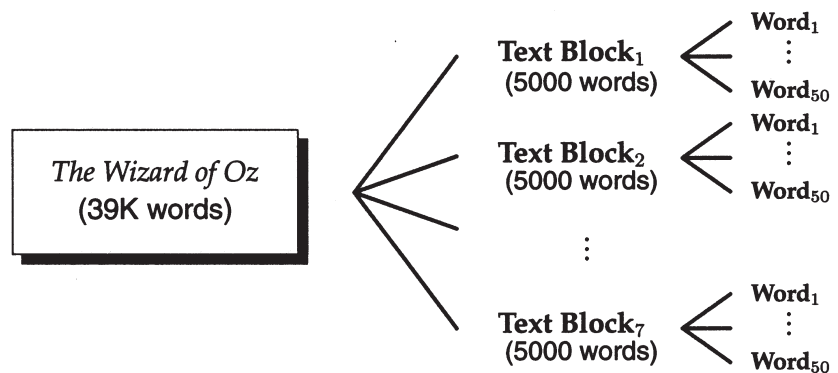


Figure 4. Obtaining the matrix of n observations (text blocks) and 50 variables (word counts).

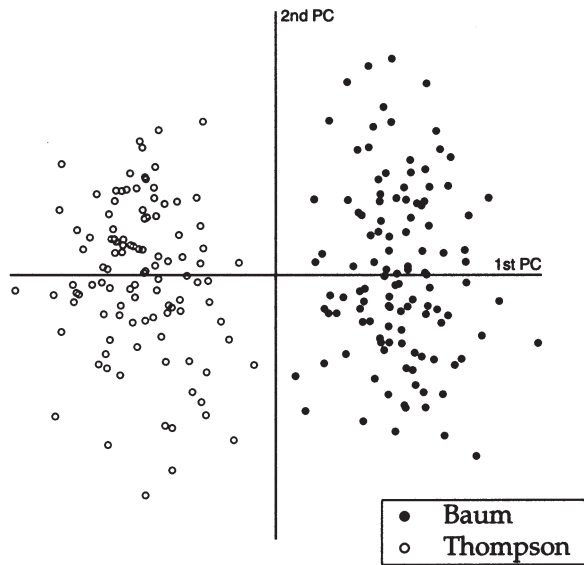


Figure 5. Baum vs. Thompson.

the technique to stylistic analysis is detailed in Binongo and Smith (1999a).

Because the PCs account for rapidly decreasing portions of the total variance, most of the remaining PCs may be replaced (admittedly with some loss of information) by their means, which have been corrected to zero. In our particular application, the 50 PC scores for the j th text block are approximated

by $(y_{1j}, y_{2j}, 0, 0, \dots, 0)$, where y_{ij} represents the i th PC score of the j th text block, $j = 1, 2, \dots, n$. Plotting the points (y_1, y_2) in a scatterplot, we obtain the best two-dimensional approximation of the original conceptual picture in 50 dimensions. Following Pearson (1901), we refer to this two-dimensional representation as the best-fitting plane.

PCA, therefore, provides a parsimonious representation rather than a classification of the data. However, by focusing on those factors which account for the largest sources of variation, a map of the text blocks can be drawn in which natural clusters are often revealed. The interpoint distances in the lower-dimensional representation indicate the degree of affinity in hyperspace.

Baum vs. Thompson

Figure 5 shows the projections of Baum's and Thompson's text blocks on the best-fitting plane. The 1st PC accounts for 20% of the total variation; the 2nd PC 7%. Albeit explaining only a fifth of the variability in the data set, the 1st PC clearly separates the two authors. The 2nd PC does not appear to contain additional authorship information; we could thus have projected the text blocks onto what Jackson (1991, p. 6) calls "the orthogonal regression line." Nonetheless, we plot the 2nd PC to produce a readable graph, bearing in mind that our primary interest lies in the 1st PC.

Perhaps most impressive in this result is that, as one may recall, frequency of occurrence was the only

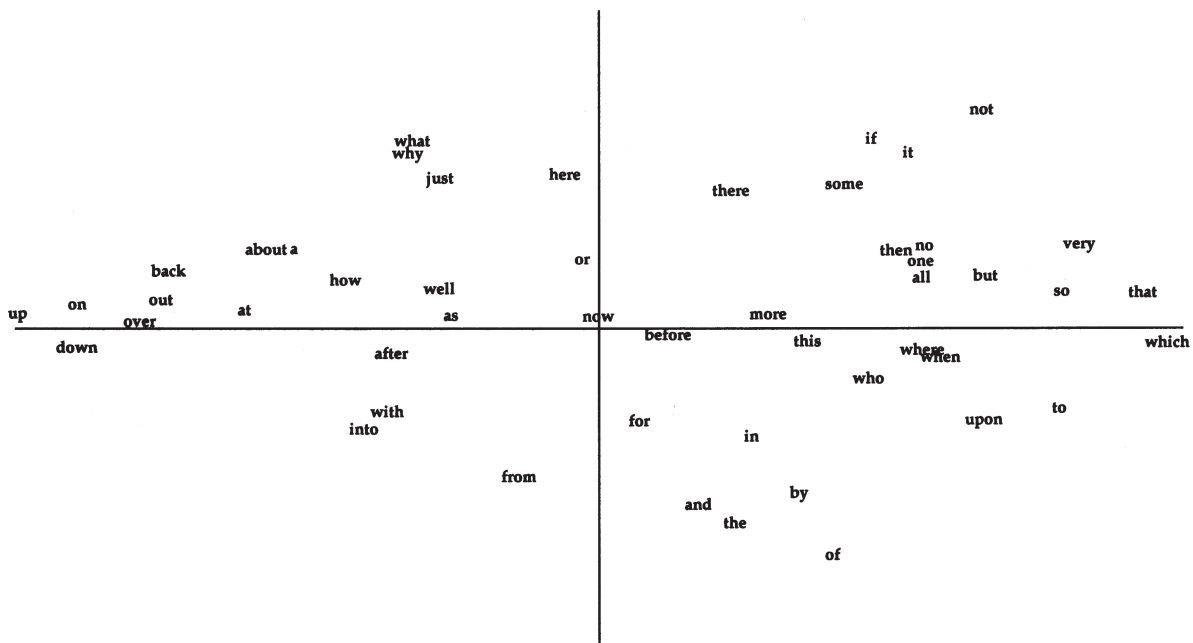


Figure 6. Component loadings.

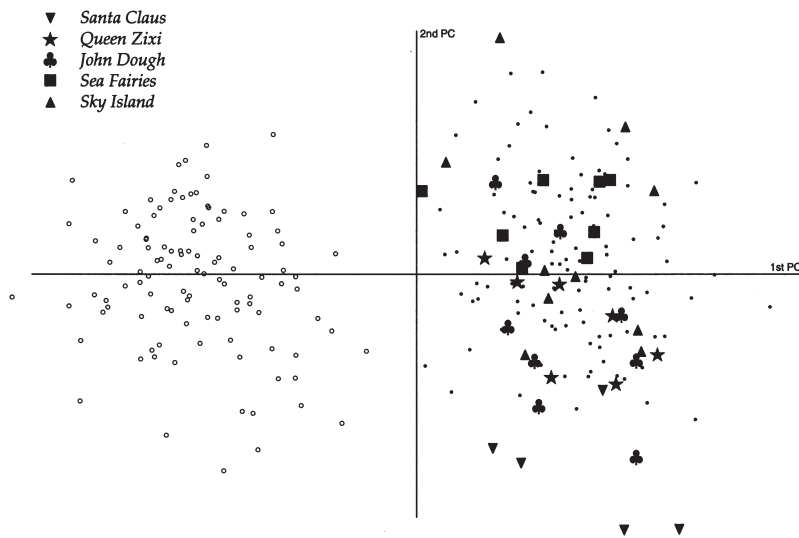


Figure 7. Baum's non-canonical works.

criterion used in the selection of variables. Even in the absence of prior information regarding how well the words distinguish between the two authors, the separation in Figure 5 is quite sharp.

What actually differentiates the styles of the two authors? The loadings on the 1st PC, which are correlations between the PC and the function words, suggest Thompson's tendency to use words indicating position—"up," "down," "on," "over," "out," and "back"—more frequently than Baum. An examination of the raw data reveals that, for these words, Thompson's average rates of usage are about twice as Baum's. Baum, on the other hand, prefers "which" and "that." Moreover, he has a greater propensity for negative words: "but," "not," and "no." Figure 6 shows the rest of the authors' contrasting uses of the function words considered in this study.

External Validation

In addition to writing the Oz series, Baum published ten full-length fantasies, four volumes of short stories, and approximately twenty miscellaneous fairy tales (Riley 1997, p. 7). Copies of some of these non-Oz works were obtained. These will now be treated as anonymous texts. In the process, we shall find out if the graphical technique as described above suc-

ceeds in ascribing the authorship of these books to their author.

Figure 7 shows the positions of the text blocks taken from Baum's *The Life and Adventures of Santa Claus* (1902), *Queen Zixi of Ix* (1905), *John Dough and the Cherub* (1906), *The Sea Fairies* (1911), and *Sky Island* (1912) on the best-fitting plane. On the background are the Oz books. With no exception, all of the text blocks from the five non-canonical works align with Baum's Oz books, not Thompson's.

The books in Figure 7, though technically not part of the Oz canon, are in one way or another related to the world of Oz. In *The Sea Fairies*, for example, readers encounter Trot and Cap'n Bill, whose travel to Oz is documented in *The Scarecrow of Oz*, the 9th canonical book. Button-Bright and Polychrome, who first appeared in *The Road to Oz*, the 5th Oz book, meet Trot and Cap'n Bill in *The Sky Island*.

The next two books, *The Master Key* (1901) and *The Enchanted Island of Yew* (1903), have little or no connection to Oz. Still, as Figure 8 reveals, they have much more in common with Baum's works than with Thompson's.

Finally, text blocks from Baum's short story compilations—*American Fairy Tales* (1901), *The Magical Monarch of Mo* (1903), and *Animal Fairy Tales* (1969)—and from Thomp-

son's *The Wizard Way-up and Other Wonders* are displayed in Figure 9. There is a slight overlap near the mean of the 1st PC. This aberration may be due to the difference in genre. Moreover, a short story is by itself a complete, stand-alone literary work. Because many of them have less than 5000 words, stories in each collection were concatenated to form one contiguous unit. Consequently, the highlighted text blocks in Figure 9 may consist of two (or more) different stories.

The Attribution Stage

Because the disputed work, *The Royal Book of Oz*, is part of the canon, checking how the proposed method performs on non-canonical works is not really necessary. Nonetheless, the nearly perfect concordance between author and text in the preceding section increases our confidence in our approach. Convinced that the method works, we are ready to tackle the problem of the authorship of *The Royal Book of Oz*.

Figure 10 confirms the present-day consensus on the authorship of the 15th book. From a statistical standpoint, this book is more likely to have been written in Thompson's hand.

The last book by Baum, *Glinda of Oz*, was published a year after his death. Gardner (Gardner and Nye, p. 40) reports that during the last year and a half of his life, Baum wrote a rough draft of this work. One of his sons edited the published version. Figure 10 unambiguously shows that the 14th book behaves very much like the compositions of the first Oz Historian.

A nagging question that often arises in an observational study such as the present one is whether the observed separation between the two authors' literary productions can, in fact, be attributed to the difference in authorial style. Both Baum and Thompson wrote in the same genre and about the same setting. Although new personalities often appeared, the key players—Dorothy, Ozma, Glinda, and the Wizard—remained unchanged. What about the date of composition? Because the 14th and

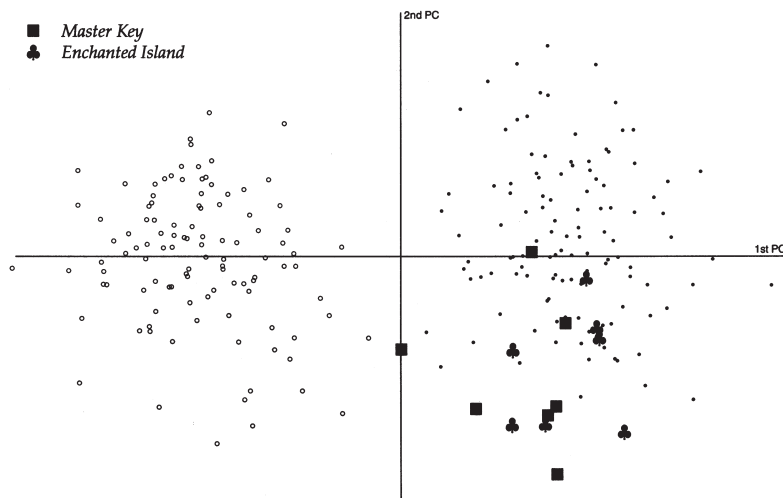


Figure 8. Baum's non-canonical works.

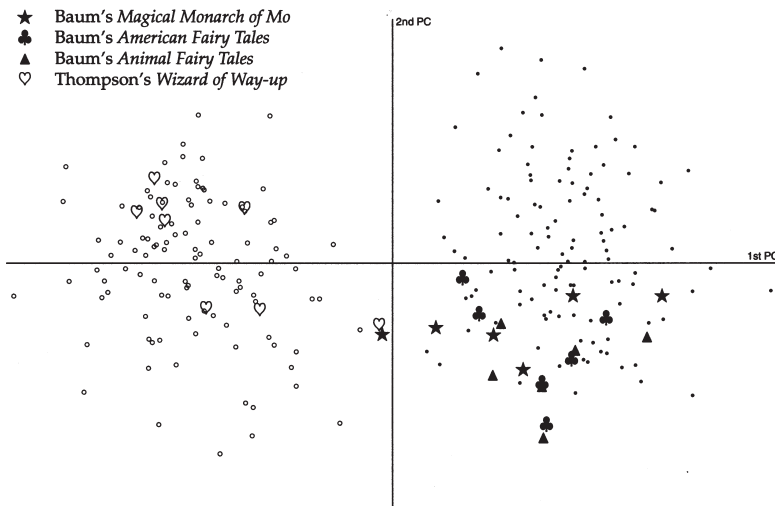


Figure 9. Baum's and Thompson's short stories.

15th books are almost contemporaneous (relative to the other books in the series) and yet are found to have distinct styles, there is reason to believe that the stylistic gulf depicted in Figure 10 is due to the authorial factor.

What if a Third Author Exists?

Just before the centennial anniversary of the publication of *The Wizard of Oz*, popular mathematics writer, Mar-

tin Gardner, surprised readers with his own Oz adventure, which he titled, *Visitors from Oz* (1998). In his words:

All my life I have hesitated about writing another Oz book. . . . My respect for Baum's "canon" was . . . so great that I felt it was a kind of sacrilege to write an Oz book that was not about the "real" Oz.

As the centennial anniversary . . . approaches, I have finally

persuaded myself that it is impossible to damage the Royal Historian's reputation by perpetrating a new Oz book. (p. xv)

Gardner's knowledge of the world of Oz and its creator cannot be contested. In 1957, he published the first biographical essay on Baum. He also wrote the introduction to five paperback reprints of Baum's books. In the preface to his own Oz book, he wrote: "I have done my best to come close to Baum's simple style, and to be faithful to the whimsical land and characters he created" (p. xv). Even with such unwavering fidelity, Gardner, as the 1st PC in Figure 11A reveals, actually writes more like Thompson than the creator of Oz.

Unlike *The Royal Book of Oz*, whose text blocks scatter around what is visibly Thompson's region of the plane in Figure 10, text blocks from Gardner's *Visitors from Oz* appear to cluster in a smaller region in Figure 11. This merits a follow-up investigation. Zooming in on the region does not help. To derive more information from each text block, we increase the size of each text block from 5,000 to 10,000 words. Figure 11B shows that doubling the text block size allows the 2nd PC to capture additional authorship information.

When Figures 7–10 are redone using 10,000 words per text block for each individual book tested previously, no separation in the 2nd PC similar to that observed in Figure 11B occurs. If anything, the increase in block size only sharpens the discrimination between Baum and Thompson, completely eliminating the overlap in Figure 9.

The Fingerprint Metaphor

The acquisition of an adequate number of sample texts from each candidate author is a sine qua non of every successful attribution. In our present study, there was plenty of opportunity to study each author's stylistic variation: Baum and Thompson produced hundreds of thousands of words in the same genre. Why do we need thousands of words? Unlike fingerprints, an author's habits of style are

not always immediately obvious, are not “immune to variation from infancy to the grave” (Slater 1988, p. 54), and can be expressed only in relation to another author’s. The authorial signals which stylometrists perceive are often faint, and a clearer view may be achieved with additional compatible data. Smith (1992, p. 194) issues a word of caution: “Reputable practitioners would hesitate to claim that stylometry is as precise as fingerprinting.” Michaelson, Morton, & Hamilton-Smith’s (1979) notion of stylometry as “fingerprinting the mind” is thus a considerable exaggeration. The metaphors of “signature” or “voice” are perhaps more appropriate because they permit a degree of change while retaining a writer’s individuality. Furthermore, even though stylometry cannot offer criteria which are as uniquely defined as fingerprints are, it is not necessarily invalidated as a technique for attributing authorship.

The Advantage of the Graphical over the Inferential

Using the multivariate statistical technique of principal component

analysis, we are able to reduce 50 dimensions into 2 dimensions. As the first two PCs account for the two largest proportions of the total variation, the construction of the best-fitting plane ensures that the most important information is retained. The clusters that are formed are natural; the visual display is faithful to the data. The interpoint distances of the text blocks in the lower-dimensional representation provide us with a picture of their relative affinities in the original hyperspace.

The largest source of variation in the data sets appears to be the authorial factor. Baum’s books are differentiated from Thompson’s in that the former’s books are on one side of the mean of the 1st PC and the latter’s on the other side. With this unerring consistency, we have confidence in our identification of Thompson as the author of the 15th book.

Nonetheless, it must be mentioned that stylometric investigations are typically observational studies, not controlled experiments. These investigations are, in fact, subject to an incalculable degree of uncertainty due to unknown and unknowable factors which can affect the production of texts. Faced with this limitation, stylometrists can be confident in their

ascription only after they have acquired a collection of comparable writing samples sufficient to assess the variability of authors’ styles. The use of the most frequent function words may reduce the confounding influence of genre and date of composition (which is also related to writing maturity) on the outcome; still, their extraneous effect cannot be dismissed entirely. There is always a need to conduct internal and, whenever possible, external validation tests. It is conceivable that the writing styles of a pair of authors cannot be distinguished using the proposed method.

PCA used as a graphical tool, however, has an advantage over methods which report numerical probabilities: the validity of PCA as a technique for identifying authors does not depend on unverifiable statistical assumptions. Indeed, its use has gained popularity among stylometrists, who often employ a multivariate approach in their investigations. Interestingly, the method may do more than distinguish between two authors. When used to analyze the works of one author, it may also isolate difference in literary genre. In a previous study (Binongo and Smith 1999b), Oscar Wilde’s plays were distinguished from his essays on the best-fitting plane.

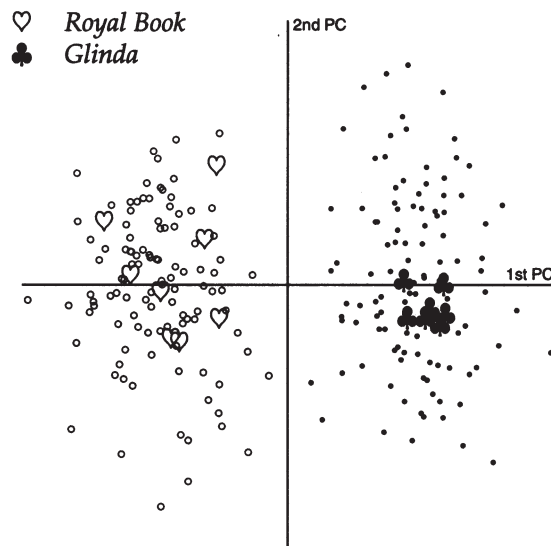


Figure 10. *The Royal Book of Oz* (1921).

Conclusion

The statistical analysis in this article reveals that the writing style in the 15th Book of Oz is more compatible with Thompson’s than with Baum’s. Whether Baum left unfinished notes which Thompson later used as a springboard for writing the book cannot be verified by statistics. However, the present investigation provides objective, independent evidence supporting the literary consensus that the book was written in Thompson’s pen (in addition to Hearn 2000, mentioned earlier, see Gardner and Nye 1957, p. 40; Carpenter and Shirley 1992, p. 124; Riley 1997, p. 232).

Literature, indeed, is not without a dimension that is open to statistical scrutiny.

Acknowledgments

I am grateful to Collegiate Schools for the Brinkley Grant, which allowed me

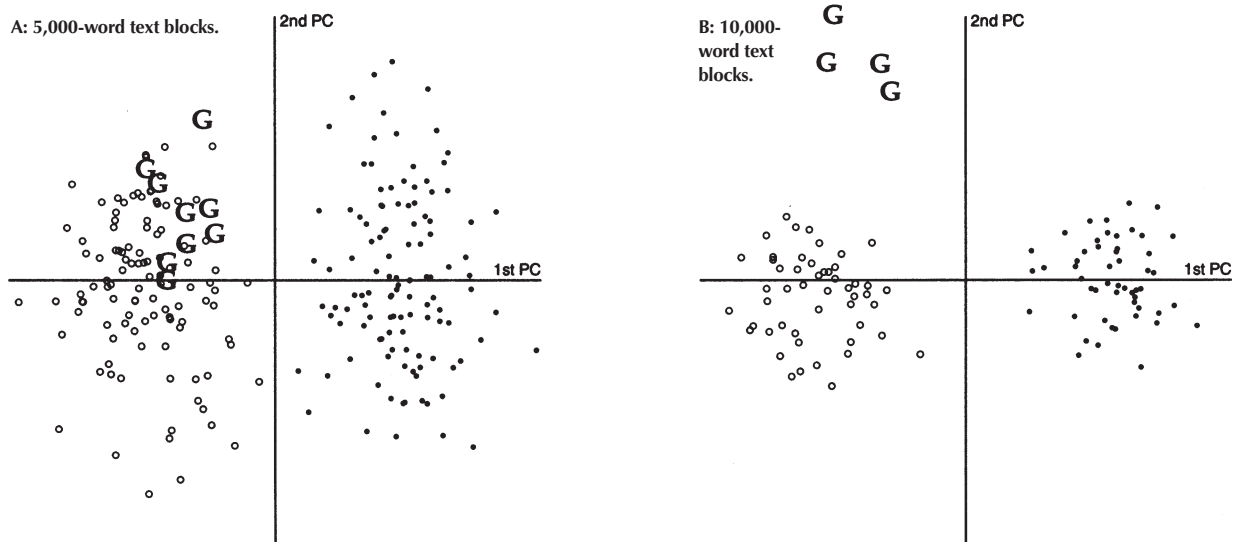



Figure 11. Gardner's Visitors from Oz (1998).

to complete this study and to present the findings at the 2002 Conference of the Classification Society of North America in Madison, WI. Thanks are also due to Hampton Ford for going over the manuscript. 

References

- Baum, F.J. and MacFall, R.P. (1961), *To Please a Child: A Biography of L. Frank Baum, Royal Historian of Oz*. Chicago: Reilly & Lee.
- Baum, L.F. [Thompson, R.P.] (2001), *The Royal Book of Oz*. New York: Dover Publications.
- Binongo, J.N.G. and Smith, M.W.A. (1996), "Stylometry," *Notes and Queries* 241: 448–52.
- Binongo, J.N.G. and Smith, M.W.A. (1999a), "The Application of Principal Component Analysis to Stylometry," *Literary and Linguistic Computing* 14: 445–65.
- Binongo, J.N.G. and Smith, M.W.A. (1999b), "A Bridge Between Statistics and Literature: The Graphs of Oscar Wilde's Literary Genres," *Journal of Applied Statistics* 26: 781–87.
- Carpenter, A.S. and Shirley, J. (1992), *L. Frank Baum: Royal Historian of Oz*. Minneapolis: Lerner Publications Company.
- Gardner, M. (1998), *Visitors from Oz*. New York: St. Martin's Press.
- Gardner, M. and Nye, R.B. (1957), *The Wizard of Oz & Who He Was*. East Lansing: Michigan State University Press.
- Hearn, M.P. (2000), *The Annotated Wizard of Oz*. New York: W.W. Norton & Company.
- Jackson, J.E. (1991), *A User's Guide to Principal Components*. New York: John Wiley & Sons.
- Jolliffe, I.T. (1986), *Principal Component Analysis*. New York: Springer-Verlag.
- Lutoslawski, W. (1897), *The Origin and Growth of Plato's Logic*. London: Longmans.
- Michaelson, S., Morton, A.Q., and Hamilton-Smith, N. (1979), "Fingerprinting the Mind," *Endeavour* 3: 171–175.
- Moore, R. (1974), *Wonderful Wizard, Marvelous Land*. Bowling Green: Bowling Green University Popular Press.
- Mosteller, F. and Wallace, D.L. (1984), *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. 2nd ed. of *Inference and Disputed Authorship: The Federalist* (1964). New York: Springer-Verlag.
- Pearson, K. (1901) "On lines and planes of closest fit to system of points in space," *Philosophy Magazine* 6: 559–72.
- Riley, M.O. (1997), *Oz and Beyond: The Fantasy World of L. Frank Baum*. Lawrence: University Press of Kansas.
- Slater, E. (1988), *The Problem of The Reign of King Edward III: A Statistical Approach*. Cambridge: Cambridge University Press.
- Smith, M.W.A. (1992), "Bridging the Cultural Divide." *Scriblerian* 24: 192–5.
- Snow, J. (1954), *Who's Who in Oz*. Chicago: Reilly & Lee.