2-23-1990

Can Authors Alter Their Wordprints? Faulkner's Narrators in As I Lay Dying

Tim Hiatt

John Hilton

CAN AUTHORS ALTER THEIR WORDPRINTS?
FAULKNER'S NARRATORS IN *As I Lay Dying*

Tim Hiatt and John Hilton
Brigham Young University

William Faulkner's *As I Lay Dying* is told in a unique way: fifteen narrators speak from their own points of view. The language of each narrator contributes greatly to each's uniqueness. Just how well-characterized are the narrators?

To answer the question, we measured some of the narrators in *As I Lay Dying* using current wordprinting techniques. Wordprinting or stylometry is the study of words which appear in patterns specific to authors. These patterns are thought to be somewhat like an identifying fingerprint---that they can help identify authors of anonymous texts. These wordprinting measurements will help us determine if Faulkner can alter his own literary "fingerprint."

A wordprint, as measured in current wordprinting methods, usually consists of studying the author's use of non-contextual words. To explain how we obtained our measurement of Faulkner's wordprint, we will first describe what non-contextual words are and how they are used to determine a "wordprint." We will then describe the wordprinting process we undertook for *As I Lay Dying*, as well as for *Pylon* and *Light in August*, two other works by Faulkner. Finally, we will show how Faulkner's manipulation of non-contextual words and other word use patterns helps characterize the narrators so well, and how he seems to have set himself apart from all other measured authors.

Contextual and non-contextual words

Authors try to simulate different voices for different characters when they write in the first person. They achieve a definite flavor in their characters by differing the topics their characters think or speak about, their attitudes about their surroundings, and the events they narrate. Authors can include dialects to help differentiate between characters; they can raise the social level of the spoken words; they can riddle the character's thoughts with clichés and sayings that other roles wouldn't use. These are all external methods of changing the authorial voice between characters and they are all commonly employed.

All these methods are contextual: they depend on the context the author is writing under. But wordprinting measures things that do not strongly depend on context: words such as *and*, *the*, *of*, *for*, *it*, and so forth. The idea behind wordprinting is that usage of these non-contextual words varies little with changing context, that the occurance rates of these words among all works by one author will be approximately the same.

For example, in *Moby Dick*, the word "whale" appears many times. The same word appears less in *Billy Budd* (using a "use rate" or percentage measurement---which lets us compare works of differing length), even though both were written by Melville. It is of course incorrect to conclude that the two works were written by different authors on the basis of a contextual or context-related word or words.

If the subconscious usage patterns for non-contextual words are specific to each author (which idea is hypothesized in current wordprinting), then detection of these patterns will reveal the author's

subconscious hand in all his works. These statistical patterns can, however, be influenced somewhat by externally imposing forms didactic, expository, narrative, or other literary forms. Therefore, patterns measured between authors are best taken from free flow writing from texts of the same literary form. All else being equal, the statistical patterns are somewhat like an identifying fingerprint: it is found for most authors to be specific to the author when the author freely writes in similar genres.

Wordprints used for determining authorship

Documents that have no reputed author, or disputed authors, have been assigned authors by detecting the non-contextual word patterns and matching them with known patterns of candidate authors. Such studies are the main thrust of applied wordprinting analysis. One of the originators of literary stylometry or wordprinting, A. Q. Morton, briefly outlines the history of such analysis. Greek originally was the language of wordprinting, and English language studies started later with the Shakespeare/Bacon/Marlowe authorship controversy.

The late nineteenth century saw the use of several numerical methods in trying to solve the Shakespeare authorship question. One of the first numerical methods was measuring the average number of characters in each word---the word length. It was found that Shakespeare's average word length was four, while Bacon's was three and Marlowe's was four (Morton 184). This result is interesting, but of course inconclusive. More specific tests were required---tests that would show overarching patterns peculiar to individual writers.

Later, tests using sentence length, punctuation style, and verb tense were tried, again with interesting but inconclusive results. It was Morton who eventually refined the idea of analyzing the parts of speech to include only the non-contextual word patterns.

A recent example of a more modern wordprinting method was done at BYU by Larsen, Rencher, and Layton, then of the BYU Statistics Department. They used very advanced statistical techniques to demonstrate numerical differences in word use rates among the purported writers in the Book of Mormon. They concluded that the texts they examined were written by several distinct authors and that none of the modern candidates tested for Book of Mormon authorship wrote any of that text (Larsen 234). However, D. James Croft, a statistician from the University of Utah, pointed out several areas in which he thought the Larsen group's study needed improvement before definitive conclusions could be reached (see Croft 20-21).

The Larsen group's study, as well as Croft's suggestions, were taken into account by a group of scientists at Berkeley who were also investigating the wordprinting idea. Led by John L. Hilton and Kenneth D. Jenkins, the Berkeley group developed an independent computerized method employing the best research available at that point, along with Morton's ideas on non-contextual word patterns and his series of word pattern tests.

It is important to note that wordprinting is an advancing science: the Larsen study illustrates that point and tells us to proceed with caution, especially when dialects and regionalisms are involved. To produce an accurate and believable result, meticulous measurements and attention to detail must underlie the best wordprinting procedures available.

John Hilton moved to Provo recently, and brought the Berkeley group's advanced wordprinting methods with him. Thus began our wordprinting studies on William Faulkner.

Wordprinting Faulkner's works

Our purpose in examining *As I Lay Dying* was not to determine authorship; Faulkner obviously wrote the story unassisted. Instead, we intended to use statistical methods to see if Faulkner was able to create distinct narrative voices---distinct to the point of challenging one hypothesis of current wordprinting.

Hilton and others have previously performed wordprinting measurements with many different authors, four of which are included in this study to provide typical perspective. Two of our examples, Samuel Clemens and Robert Heinlein, were trying, in their representative texts, to create different narrators. As we will point out later, unlike Faulkner, Clemens's narrators are grammatically and statistically equal, and so are Heinlein's.

In addition to *As I Lay Dying*, we wordprinted two other Faulkner works: *Pylon* and *Light in August*, to examine Faulkner's inner consistency. Time restrictions kept us from evaluating Faulkner's essays or personal writings. Before we discuss our findings, we will briefly outline the mechanics of wordprinting.

The process begins with the text of the story in computer-readable form. We divide the story by chapters and group the chapters by narrator. We eliminate all identifiable quotations from other characters from each group of chapters, so that what remains represents the thoughts and words of only that narrator. Each of these files are then converted by the computer to a series of codes reflecting the parts of speech each words represents: "n" for noun, "v" for verb, etc., leaving intact the non-contextual words for analysis in their original state. Computer programs process this coded file representing the grammatical patterns of each of the sentences in the story to find the raw word occurrence data for each non-contextual word. The results of this processing are combined with more passes through the coded file to identify the different word patterns that Morton outlined, and to gather more data for other tests. Once all this data is generated for each of the single author files, the test results for each are compared with results from the others (Hilton 3-18).

Results of computations

Of the fifteen narrators in *As I Lay Dying*, only three had enough text to yield statistically significant samples: Darl, Tull, and Vardaman. Also, for comparison, we include data from *Pylon* and *Light in August*. They are both third-person narratives, while *As I Lay Dying* is in the first-person: some of the differences found in the study will exist because we are comparing different literary styles.

We used three different measurements to examine the Faulkner texts. First, the occurrences of each non-contextual word were counted and totaled. We then calculated the ratios of Morton's word patterns. Finally, we analyzed the rate at which unique words are introduced into the text. We will report the results beginning with the non-contextual word occurrences, then continue with the Morton word pattern ratios, then conclude with a discussion of the new word introduction rate and related vocabulary.

Non-contextual word usage

A chart showing the non-contextual words used (per 5000 words of text) appears in figure 1. Decimals appear because the larger texts were broken into several 5000 word segments with the average value recorded. Included in the chart are measurements from *Pylon* and *Light in August* for comparison.

	Pylon	L.I.A.	Darl	Tull	Vardaman
A	108	105	89	111	58
AN	12	8.8	12	0	3
AND	262	211	183	204	230
ANY	4.7	3.2	3	1	0
ALL	10	16.6	12.3	19	17
BUT	18.2	25.6	15	25	34
BE	8	14	10	23	17
BY	22	12.6	9.6	8	2
I	.75	1.8	51	69	149
IN	75	83.2	74	61	85
IT	53	72	92	172	132
NO	7.2	13.8	8	13	1
NOT	35	65	49	71	115
OF	136	90	90	66	37
THAT	35	78	44	55	14
THE	455	353	375	281	332
TO	79	105	82	100	109
WITH	48	45	40	33	14
"TO BE"s	49	121	110	164	223
HAVE	9.5	13.8	10.3	14	13
UP	13.2	12	30.6	44	44
UPON	7	11.4	9.6	0	1
WITHOUT	12.5	8.8	7.6	3	2
YOU	1	0	28	36	25
AS	42	38	40	12	11
VERBs	808	1011	970	1183	1145
ADJs	1029	892	821	685	685
length of sentence (in words)	27.5	17.4	13.6	13.6	10.2

Fig. 1: Average number of occurences of non-contextual words in 5000 word segments of texts.

Beginning with an obvious difference that stems from third and first-person narration, we see that there is an average of less than one "you" in the third-person narratives, while 28, 36, and 25 "you"s appear in the Darl, Tull, and Vardaman narratives. Coupled with this is use of "I": under 2 average in the third-person novels and 51, 69, and 149 for Darl, Tull, and Vardaman.

The difference between first and third-person narration also appears in the numbers of verbs and adjectives used. Tull and Vardaman use less verbs than Darl, Pylon, and Light in August; but the third-person novels use more adjectives.

Faulkner's characterizations emerge in the use of "a" and "an." We see that the ratio of "an"s to "a"s for the novels and Darl is about one "an" for every ten "a"s (10%). But Tull (0%) and Vardaman (6%) have significantly less "an"s. Faulkner could be viewing the word "an" as a word to be used by only the educated or articulate. Darl uses "an" at a rate similar to Faulkner's third-person narratives, which narratives use a more ornate or full style. Tull and Vardaman seem to use a more colloquial grammar, which explains the lack of "an." In fact, no "an"s appear in Tull's narrative, and all the occurences of "an" in Vardaman's text are in a single paragraph, which also includes the only use of "upon" for Tull or Vardaman.

In As I Lay Dying, just before the paragraph containing the "an"s, Vardaman drives away the doctor's team and buggy, and Darl comes to see what happened. Vardaman then describes a

dismembered, disassociated Darl in animal terms. The three "an"s appear in his description---along with the atypical poetic-sounding language in Vardaman's perception. Part of the paragraph says:

> It is as though the dark were resolving him out of his integrity, into an unrelated scattering of components . . . an illusion of a co-ordinated whole . . . within which, detached and secret and familiar, an *is* different from my *is*. I see him dissolve . . . and float upon the dark in fading solution; all one yet neither; all either yet none. (1556)

For this paragraph only, the vocabulary widened and the imagery grew complex. The style sounds more like the imaginative Darl and his introspective observations than the young boy Vardaman. Faulkner's use of "an" here alerts us to the change in register: Vardaman's speech briefly becomes eloquent, as indicated in part by his exceptional use of "an."

If sentence length is a function of language complexity, then Vardaman is the simplest text at 10.2 words per sentence, and *Pylon* the most elaborate at 27.5 words. Sentence length is also influenced by literary form, so it is fair to say that the difference is in part caused by the difference between narrative points of view.

Due also to Faulkner's technique of writing Tull and Vardaman in a lower register than Darl or the two novels is the frequent use of "to be" verbs in Tull and Vardaman. Apparently Faulkner's experience told him that the less articulate will use forms of the verb "be" more often. For this reason, Tull and Vardaman use "to be" verbs almost twice as much as Darl, *Pylon*, and *Light in August*.

From the evidence we present, it appears Faulkner viewed many of these normally defined non-contextual words as contextual in the sense that they help contribute to the peculiar quality of his characterizations. Many words (*an, any, by, of, with, upon, without*, and *as*) appear to be restricted by Faulkner for use by complex narrators, while others (*it, not*, and *up*) are used more by the simple language narrators. Still other words (*all, but, in, that, the, to*, and *have*) seem to be uniform through all the texts---truly non-contextual, or at least not manipulated by Faulkner.

Morton's word patterns

The non-contextual word occurence data we analyzed helps one understand how Faulkner manipulated language to meet his needs. However, more subtle methods of orchestration appear when one observes the patterns in which these non-contextual words occur.

A. Q. Morton's word patterns provide a more comprehensive measurement of what would normally be thought of as "subconscious" usage patterns. These word patterns measure where non-contextual words appear in the sentence in relation to the number of sentences. They also measure the placement of modifying or related words in relation to non-contextual words, and the rate at which new words are introduced adjacent to the non-contextual key words. Figure 2 displays the Morton word patterns, with a guide for interpretation.

36	A(fws)/#
37	AN(fws)/#
38	AND(fws)/#
39	IN(fws)/#
40	IT(fws)/#
41	IT(lws)/#
42	OF(fws)/#
43	OF(2nd lws)/#
44	THE(fws)/#
45	THE(2nd lws)/#
46	WITH(2nd lws)/#
47	A(2nd lws)/A
48	A(fb adj)/A
49	A(fb x AND)/A
50	A(fb x OF)/A
51	A x A/A
52	A x x A/A
53	AND(fb adj)/AND
54	AND(fb THE)/AND
55	AND(fb x OF)/AND
56	AND x AND/AND
57	AND x x AND/AND
58	AS x AS/AS
59	AS x x AS/AS
60	BE(fb A)/BE
61	BE(pb TO)/BE
62	BUT(fb A)/BUT
63	BY(fb THE)/BY
64	I(fb AM)/I
65	I(fb HAVE)/I
66	I x I/I
67	I x x I/I
68	IN(fb A)/IN
69	IN(fb THE)/IN
70	OF(fb A)/OF
71	OF(fb THE)/OF
72	OF(fb x AND)/OF
73	THE(pb AND)/THE
74	THE(pb OF)/THE
75	THE(pb IN)/THE
76	THE(pb TO)/THE
77	THE(fb x AND)/THE
78	THE(fb x THE)/THE
79	THE(fb x x THE)/THE
80	TO(fb BE)/TO
81	TO(fb THE)/TO
82	TO x TO/TO
83	TO x x TO/TO
84	YOU x YOU/YOU
85	YOU x x YOU/YOU
86	TO(between)VERBs/TO
88	AN/AN+A
89	ANY/ANY+ALL
90	NO/NO+NOT
91	UP/UP+UPON
92	WITH/WITHOUT+WITH
94	A(r)/A(r+1)use only
95	AND(r)/AND(r+1)only
96	IN(r)/IN(r+1)only
97	IT(r)/IT(r+1)only
98	I(r)/I(r+1)use only
99	OF(r)/OF(r+1)only
100	THAT(r)/THAT(r+1)
101	THE(r)/THAT(r+1)only
102	TO(r)/TO(r+1)only

Guide to Figure 2.

#	the number of end of sentence markers
fws	first word in sentence
lws	last word in sentence
2nd lws	second to last word in sentence
fb	followed by
pb	preceded by
x	any word
x x	any two words
r	word to the right is unique within the current block of 1000 words
r+1	the words to the right and left that are unique within the current block of 1000 words

Note: Each test here is numbered from 36 to 102. Tests 87 and 93 were omitted because they did not apply to these texts.

Fig. 2: List of Morton's word pattern ratios.

Each of these tests yields a ratio (or percentage). For instance, test 36 is A(fws)/#. This formula stands for the percentage of times that "a" is the first word in a sentence divided by the total number of sentences. Test 54 reads AND(fb THE)/AND which calculates the number of times "and" is followed by "the" as a percentage of the total number of "and"s. In other words, the construction "AND followed by THE" will occur this percent of all uses of "and" by this author.

To make the comparison between two texts using the 64 Morton word pattern ratios, each word pattern ratio measured from the first text are compared to the same patterns in the second text. When there is enough difference between the two ratios to be statistcally significant, we identify it as a null-hypothesis rejection, or simply a "rejection."

Before looking at the number of rejections found in our texts (Darl, Tull, Vardaman, *Pylon*, and *Light in August*), let us give some examples of how works by other authors compare using the same system of Morton word pattern tests for null-hypothesis rejections.

For our study, a number of non-disputed control authors provided a standard from which to judge the results of the Faulkner comparisons. Among these works and their authors were "Everyone Loves a Lord" (essay), "Early Days" (autobiography), "Diary of Adam," and "Diary of Eve" (fictional narratives) by Samuel Clemens; *The Number of the Beast* (novel) by Robert Heinlein, "Rambler," "Idler," (essays) and "Journey to the Western Islands of Scotland" (travelogue) by Samuel Johnson, and various essays by translator Harry Steinhauer.

We use the rigorous Mann-Whitney statistical calculation to measure the numbers of null-hypothesis rejections at 95% probability as we tested each pair of texts with all of Morton's 64 word pattern tests. We found that as each control authors' texts were paired with another text by the same author (called a within author test) an average of 2.9 rejections were measured from Morton's word patterns. In other words, 2.9 of Morton's 64 word pattern tests on average were very different when the same author writes in different texts.

In comparing the control authors' works where the tested pairs of texts were written by two different authors (called between author tests) we find between each an average of 7.4 of Morton's ratios were measured as significantly different. These results verify a basic assumption of current wordprint studies: they indicate that normally authors vary between each other in quantifiable amounts much more than they differ from themselves in their non-contextual patterns. It is interesting to note that the average number of rejections for each of the control authors do not vary much: the averaged within author comparisons are as low as 2 rejections (Heinlein) and as high as 3.5 (Johnson); between author averaged rejections range from 7 (Steinhauer) to 7.9 (Johnson).

The relative uniformity of the within author comparisons (2.9 average rejections with extremes of individual texts measuring from 0 to 6) is especially significant because in two cases, the authors tried to simulate different narrators. The Heinlein text was written so that each chapter was narrated by Hilda and Deety. Both characters showed non-contextual word patterns that vary little between the two parts (suggesting a unity of style on the part of the author despite the differing characters). The other example of a control author trying to simulate different narrative voices is Clemens in his diaries of Adam and Eve. They average only 2.8 rejections when the word pattern ratios are compared---almost the average of all within author works by all the control authors. Both Heinlein and Clemens are respected authors in their fields, but both are unable to create distinct characters on the subconscious non-contextual word level.

The number of rejections measured for Faulkner's works, however, were in clear contrast with the control authors' results. Figure 3 displays the results.

texts:	compared within author, same text	compared with control texts	compared with Faulkner texts
control texts	2.9	7.4	9.8
Darl	2.5	10.1	6.1
Tull	*	9.3	6.5+
Vardaman	*	13.8	9.1+
Pylon	1.5	10.5	5.7
Light in Aug	2.8	8.4	5.9

Fig. 3: Average number of null-hypothesis rejections when comparing various texts with the same texts, with the control texts, and with Faulkner's texts, using Morton's word pattern ratios.
* = text size too small to compare with itself
+ = a single measurement compared with Faulkner measurements

Looking at the first column, we see that Faulkner's usage patterns for non-contextual words are very stable within the same character or text. Realizing that high numbers of rejections show larger differences in usage patterns, these averages indicate that Faulkner is surprisingly consistent inside each character or text he writes.

The second column shows the average number of between author rejections for each when compared with the 12 control authors' texts. Faulkner's higher average rejections suggest that his texts are also surprisingly different from the control texts than the control texts are from each other.

The third column is where we find the most unexpected contrast. When each control text is compared with the other control texts by different authors, the average rejections total 7.4. But comparing each Faulkner text with other Faulkner characters and texts results in rejections ranging from 5.7 to 9.1. If Faulkner had measured according to previous observation, we would have expected an average rejection of around 2.9 for his within author testing. It would seem that Faulkner is manipulating his "subconscious" word usage patterns in an atypical and consistent way.

Sifting through the computer printouts, we found certain tests consistently rejecting between one text or character and all or most of the others. Examining these particular tests closely should yield an understanding of how Faulkner was able to do what no other author that we know of could do.

Test 40 yields the percentage of sentences that begin with "it" (the test reads IT(fws)/#, see figure 1 above). For Tull, 11% of his sentences begin with "it," while for the three blocks of Darl and for Vardaman their percentages are 4, .9, 4.4, and 5.5% respectively. The increased use of "it" must correspond to the way Faulkner viewed the character. Faulkner wrote Tull as a fairly uneducated observer speaking in his own socio-economic dialect. Because Tull mostly narrates what he sees going on around him, and because he is not nearly as articulate as Darl, Tull usually employs "it" as a pronoun taking the place of the noun mentioned in the previous sentence. Where Darl would use a different, descriptive noun to refer to a previous noun, Faulkner wrote Tull to be less intelligent-sounding by using just "it."

Test 88 is related to the previous discussion about the use of "a" and "an." The test gives the percentage of times when "an" is used of all occurences of "an" or "a." The three sections of Darl use "an" 8.6, 12.5, and 15.6%, while Vardaman and Tull use "an" 4.6 and 0%. Again, Faulkner's

apparent conscious manipulation through interpreting an educational or social dialect explains the difference.

Test 75 shows the percentage that "in the" occurs of all uses of "the." Vardaman consistently uses "in the" more (16.3%) than the other characters (6, 7.5, 5.7, and 7.4%). Vardaman's phrasing is influenced by his character: being a young boy, he observes and comments on things using language familiar to a young boy. He comments on his mother "in the box" and people and horses "in the river," and "in the water." Vardaman uses the prepositional "in" instead of mentioning where the object is and letting the reader remember. Faulkner apparently believed that a youngster's mind would work that way.

Related to test 75 is test 69, which shows the percentage that "in the" occurs of all uses of "in." Briefly, Vardaman uses "in the" 61% of all his "in"s, while the others use "in the" for 35, 30, 34.7, and 34% of the "in"s.

These tests give an idea of how Faulkner was able to manipulate non-contextual word usage. He did so by using those words contextually (as a part of the subject matter) or by varying the use manually by conforming to a predetermined character description that Faulkner himself would have had in mind.

He was able to change his word usage patterns enough for his characters to seem, statistically, to be written by a different author. This finding seems to put the simple measurements for wordprinting into a questionable light. However, these tests let us see how it is that Faulkner changed his style. Obviously, more complicated measurements will be required to correctly wordprint Faulkner than the other authors. For wordprinting to always be valid, further testing and refinement of the current method of computing this data are necessary.

Tests 94 to 102 refer to the rates at which new words are introduced into the text adjacent to the key non-contextual words. That data is discussed in the next section in more detail.

New word introduction rate

The new word introduction rate provides a measurement of working vocabulary. This rate represents the number of unique words introduced per 100 words of text averaged over each 500 consecutive words of text.

To provide perspective on the new word introduction rate, figure 4 illustrates differing vocabularies at work.
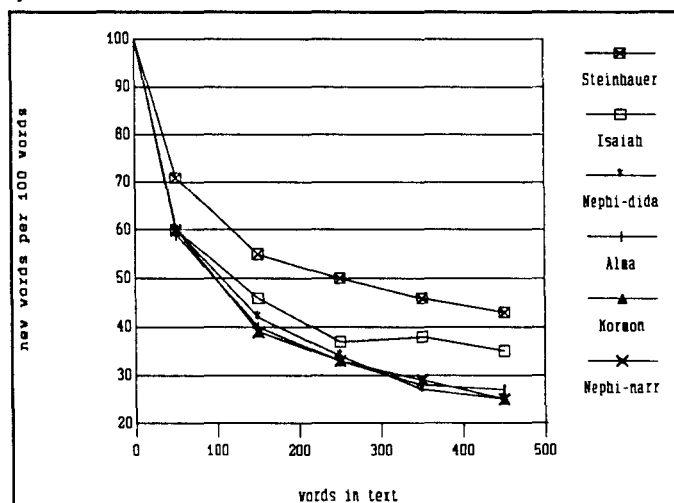

Fig. 4: Contrasting new word introduction rates among the control authors.

The most noticable feature of Figure 4 is that all purported Book of Mormon writers have almost identical low new word use rates. None stand out as being articulate among the three writers here, and these figures agree with what would be expected from Joseph Smith's early working vocabulary. He simply did not have, early in life, an extensive vocabulary. Had Joseph Smith been able to use a larger vocabulary, differences in articulateness might have appeared between the writers in his

translation. We notice also that Steinhauer has a very high rate of introducing unique words. That would mean he could, if he wished, restrict his vocabulary to simulate a less articulate writer.

Now we examine figure 5, a similar graph with only Faulkner's works plotted on it. Since these works were written by Faulkner, we can assume that his personal vocabulary (or at least his new word introduction rate) is quite high, because *Pylon* closely parallels the translator Steinhauer's rate seen in figure 4, and *Pylon* is not necessarily Faulkner at his most articulate.

The thoughtful, eloquent Darl is next, followed closely by *Light in August*. The level of vocabulary present in Darl is very comparable to the narrators of the two novels. This fact brings to mind an interesting speculation: did Faulkner write a first-person narrated short story from Darl's point of view, and then decide to embellish it with other points of view? The narrative sophistication is certainly there, stylistically, to carry the burden of narration.

In any case, Tull is less sophisticated, and Vardaman is the least prone to use unique words. The different plotted lines show, then, yet another way in which Faulkner was able to create realistic characters: he restricted the vocabulary to conform to the character he was creating.
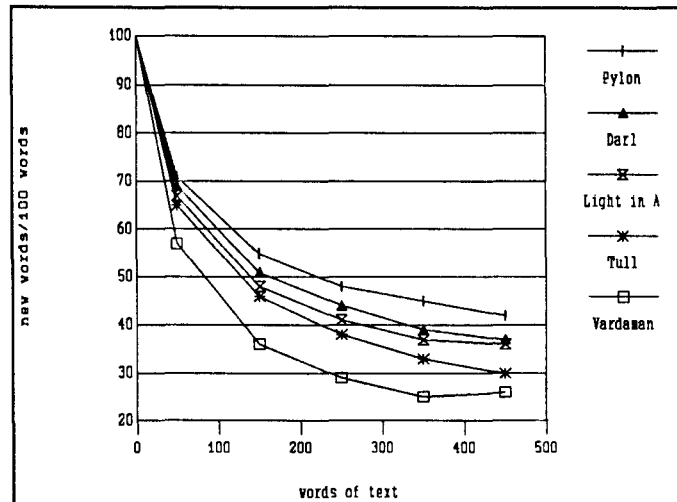

Fig. 5: New word use rates for Faulkner's works.

Conclusion

Looking back on this study, there are a couple things we would like to do differently. First, it would be good to verify the computations made by the Berkeley group's computer programs. We have made several checks and not found anything wrong, but a full verification would be better.

Also, we have not yet studied some of Faulkner's essays. It would be interesting to see if his personal essays are stable within themselves and if his wordprint matches any of his narrators or characters. It would be especially interesting if we found a match with Darl, or another of his fictional characters.

But the results we do have are astounding. Faulkner's ability to vary "non-contextual" word usage patterns seems to contradict a simple thesis of wordprinting. Yet, as we can see through our analysis of the Morton word pattern tests that were rejected, Faulkner was able to change his patterns using authorial techniques. Faulkner seems to have taken into account each narrator's socio-economic position, age, and education, as well as other intangibles like attitude and experience when creating these characters, at least as far as the within author measurements are concerned.

Therefore, although Faulkner broke the mold in creating these different narrators, his methods were discernable using this wordprinting technique. For the thesis of wordprinting ("a writer can not change his subconscious writing habits") to be rigorously investigated, however, it will require improvements in the current measuring technique which may reveal the more subtle patterns of usage by an author as sophisticated as William Faulkner. If more revealing or complex tests are devised,

we may find that Faulkner is more uniform than he now appears. It is a matter for further research.

Also of further interest is rounding up more texts in computer readable format written by sophisticated authors for further control author testing. A good candidate is *Ulysses* by James Joyce.

Up to now, wordprinting studies have been conducted with authors and texts with the purpose of assigning authorship. As far as we know, wordprinting has not heretofore been used as a literary tool. Thus, more works from the literary canon should be analyzed to see if the word patterns vary as much with other skilled authors as they do with Faulkner.

Basing this study on "non-contextual" words does show that these words are relative: Faulkner was able to manipulate them---thus making them contextual. Such contextual use would indicate that our list of what are normally non-contextual words need not always be used as such. Reconsideration of the simpler measurements of wordprinting techniques seems to be in order if we wish to measure authors such as Faulkner who apparently can manipulate "non-contextual" words in a contextual way.

This study has been a circuitous journey to find that Faulkner is an unusually complex and very skilled writer. Further studies may, or may not, bring to light more authors of similar complexity, but as of this writing he stands unique in his ability to characterize.

Works Cited

Croft, D. James. "Book of Mormon 'Wordprints' Reexamined." *Sunstone* March-April 1981: 15-21.

Faulkner, William. *As I Lay Dying*. Vol. 2 of *The Norton Anthology of American Literature*. Ed. Nina Baym et al. 3rd ed. 2 vols. New York: Norton, 1989.

Hilton, John L., and Kenneth D. Jenkins. "On Maximizing Author Identification by Measuring 5000 Word Texts." Provo: available through F.A.R.M.S., 1987.

Larsen, Wayne A., Alvin C. Rencher, and Tim Layton. "Who Wrote the Book of Mormon? An Analysis of Wordprints." *BYU Studies* 20.3 (Spring 1980): 225-251.

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