DOCUMENT RESUME

ED 280 060 CS 210 374

AUTHOR Smye, Randy
TITLE Style and Usage Software: Mentor, not Judge.
PUB DATE 19 Mar 87
NOTE 15p.; Paper presented at the Annual Meeting of the Conference on College Composition and Communication (38th, Atlanta, GA, March 19-21, 1987).
PUB TYPE Viewpoints (120) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Computer Assisted Instruction; *Computer Software; Computer Software Reviews; Feedback; Higher Education; *Revision (Written Composition); *Word Processing; Writing Evaluation; Writing Improvement; *Writing Instruction; *Writing Processes

ABSTRACT

Computer software style and usage checkers can encourage students' recursive revision strategies. For example, HOMER is based on the revision pedagogy presented in Richard Lanham's "Revising Prose," while Grammatik II focuses on readability, passive voice, and possibly misused words or phrases. Writer's Workbench "Style" (a UNIX program) provides information on sentence type, word usage, and percentages for sentence beginnings. However, most programs share the same weaknesses: a linear and judgmental composition philosophy and an inappropriate revision model. A combination of Writer's Workbench and Emacs (a Microsoft "word" type of environment that has allowed for creation of a word processor through flexible use of the VT220 keyboard) results in a technology that encourages style and usage revision at any stage of the student's composing process. For example, a student can view simultaneously the text being edited and a Writer's Workbench text analysis. This program can also be used as an instructional tool by writing a script with blank spaces for inserting variables taken from the program's statistical text analysis. The resulting feedback, specific to the assignment's rhetorical mode of an assignment and combined with the computer software technology, encourages both deep structure and surface feature revisions. (SRT)

Style and Usage Software: Mentor not Judge

by Randy Smye, Sheridan College
for the C.C.C. Annual Convention
in Atlanta, March 19, 1987

One of my most interesting recent journeys was both intellectual and geographic: a 1983 trip to visit Charles Smith at CSU in Fort Collins. My brain at the time was filled with the excitement of a recent prewriting project. The related field trials were showing me an exciting new world of computer assisted student writing. I was not looking forward to seeing WWB in action. Grammar and correctness were the furthest from my mind. But that trip to the mountains surprised me. In a congenial way Charles proceeded to anatomize how he had taken apart the bits of **Writer's Workbench** and made it his own. With a shock of recognition, on the trip back to Denver, I realized that in fact Charles had opened the door for a revision paradigm that was as heuristic, contextual and interactive as some of the latest exemplary prewriting software on the market.

That's really the point I would like to make with you today. In contrast to recent prewriting programs, computer assisted revision often seems to be more summative or evaluative, focusing more on the surface correctness of a final product. Unfortunately many instructors have allowed the "traditional" paradigm to direct our use of technology. The limits of speed and memory often rule that style and usage commentary has to lurk like a red-pencilling terror at the end of students' writing processes. However, I'd like to show you how the determined instructor can use style and usage checkers (for instance **HOMER**, **GRAMMATIK** and **Writer's Workbench**) for

writing exercises that encourage a recursive pattern of student re-vision. Let's look at several instructional design ideas that can create "process centered" revision software which appears heuristic, contextual and interactive.

First, what is a style and usage checker? One simple, "elegant" little program is Michael Cohen's **HOMER** which runs on the Apple IIe. Typically students compose an assignment and then choose to have **HOMER** read their texts to find four different word types simultaneously: prepositions, *be* verbs, words containing *tion* or *sion*, and vague words. The program keeps track of how often each word type appears in a text and also counts the total number of words and sentences. Students can choose several options for displaying the resultant information and statistics. The strength of **HOMER** is its integration with a specific revision pedagogy: Richard Lanham's "paramedic model" of stylistics as described in his *Revising Prose* (1979).

Within the MS/DOS domain we find **Grammatik II**, a wanderer from the old CP/M operating system. **Grammatik**, typical of several business writing style checkers, focuses on readability, passive voice, and possibly misused words or phrases. You'll find a useful collection of these types of programs reviewed in the May, 1986 **PC Magazine**. A common characteristic of all these checkers is the familiar "write-then-check" model. The single tasking, limited memory of the PC forces linearity. Writers must compose with their word processor, save, quit, load the analysis software, run the program on a file, and then wait patiently while disks whirl to produce an output file. Since these programs use substantial dictionaries, and since

typical college papers are several pages long, the typical PC we find in the colleges will often groan for 5 or 10 minutes to produce its output. The revision process works in a world as separate from the composing process as your judgmental red marks are from the sweaty, immediate world of your students' composing processes. The technology reflects an inappropriate composing paradigm: linear and judgmental.

My own particular favourite is the UNIX program *style*, one of the most widely available, but often overlooked, revision aids available to college level composition students. *Style*, at the core of **Writer's Workbench**, is available as a general utility hidden in the documentation of many recent implentations of the UNIX/Xenix operating systems. In a sense, if a university or college has purchased the operating system, *style* is available at no additional cost. Not only does *style* comment on readability, numbers of prepositions, and "to be" verbs, but *style* also attempts a more sophisticated analysis. The writer can find information on sentence type (for instance, simple, complex, compound, and compound-complex), word usage (that is the percentage of various parts of speech), and as well, percentages for sentence beginnings.

Style relies on another powerful UNIX tool, *parts* which assigns parts of speech to each word in a given text file. *Parts* uses a small dictionary of function words and irregular verb forms and word endings to classify words. The program then classifies any remaining words by looking for relations between words, using rules of English structure. Primarily *style* then prints a summary table of statistics produced by *parts*. But secondarily *style* can also be flagged to print out sentences with certain

characteristics; for instance, all sentences with a passive voice, nominalization, or a readability higher than a specified number. Here is a typical *style* table.

```
readability grades:
(Kincaid) 18.5  (auto) 19.6  (Coleman-Liau) 14.6  (Flesch) 17.0 (19.2)

sentence info:
no. sent 6 no. wds 191
av sent leng 31.8 av word leng 5.32
no. questions 0 no. imperatives 0
no. content wds 101  52.9%  av leng 7.49
short sent (<27) 33% (2) long sent (>42)  17% (1)
longest sent 47 wds at sent 6; shortest sent 17 wds at sent 5

sentence types:
simple  33% (2) complex  50% (3)
compound  0% (0) compound-complex  17% (1)

word usage:
verb types as % of total verbs
tobe  62% (8) aux  31% (4) inf  0% (0)
passives as % of non-inf verbs  38% (5)
types as % of total
prep 16.8% (32) conj 4.7% (9) adv 5.2% (10)
noun 27.7% (53) adj 14.1% (27) pron 3.1% (6)
nominalizations  7 % (14)

sentence beginnings:
subject opener: noun (0) pron (0) pos (0) adj (3) art (1) tot  67%
prep  33% (2) adv  0% (0)
verb  0% (0)  sub_conj  0% (0) conj  0% (0)
expletives  0% (0)
```

A wealth of information! But what to do with it?

Style apparently shares all the linear features of its MS/DOS poorer relations. The writer must drop out of composing and wait patiently for an output file. Of course with this program another barrier appears. Until recently few English teachers had a UNIX facility available to run *style* for their 'writing' students. UNIX software and documentation was expensive and so were the powerful computers required.

I've tried to surface two issues with style and usage checkers: the technology forces us to accept an inappropriate revision model and that the program output results lack a supporting instructional context. Let's face these two issues. First, the technology. The relentless increase in computer speed and memory and the proportional drop in price can now affect our composition curriculum. The new family of products with 68020 and 80386 chips brings to our classrooms the possibility of a multi-tasking and recursive writing environment. No longer do we have to suspend the composing process and wait several minutes for revision feedback. New software can integrate prewriting, writing and post-writing heuristics.

Current primitive examples point the way. RAM resident programs like Borland's **Lightning** and Living Videotext's **Ready** already allow a degree of integration for spelling and outlining for MS/DOS word processing. A Macintosh loaded with Apple **Switcher**, 3.0 **Microsoft Word**, and communications software allows the writer to do the same, but adds the possibility for easy spooling of an analysis task to a remote UNIX cpu running **Writer's Workbench**. However, the most exciting signposts for the future are the VIRTUE Workstations under development at Carnegie Mellon's ANDREW Project and the INTERMEDIA Project at Brown University. Here developers are working with Sun workstations which typically involve large bit-mapped screens as well as the speed and multi-tasking of UNIX running on 68020 cpus. Finally, the Writers Workstation has arrived.

Let me illustrate the possibilities for a more current revision paradigm from my own experience at York University where I've been spending this

year as Director of the new Computer Assisted Writing Centre. After evaluating what text editors were available under UNIX, the standard ex/vi editor and the Rand 'E' editor, we began to adapt Richard Stallman's GNU **Emacs** (4th edition, Version 17) available from his Free Software Foundation. **Emacs** was originally developed at MIT as a programmer's environment for working on UNIX multi-tasking, multi-user computer projects. This has turned out to be an exciting choice for us, because **Emacs** (a Microsoft **Word** type of environment) has allowed York faculty to create their own word processor through an impressively flexible use of the VT220 keyboard. For instance, a student while composing in **Emacs** can press F19 and watch simultaneous **Writer's Workbench** analysis appear in a window. Of course this eliminates lines at the printer or those mind numbing waits for an output file while your terminal sits frozen, incapable of doing anything else. But for me the major benefit is that finally we have a technology that encourages style and usage revision at any stage of the student's composing process.

Here is a rough idea what *style* looks like running within **Emacs**. The top half of the screen is a text available for editing while the output is sent into a buffer in the bottom half of the screen. The output buffer could be copied into the text or saved as a separate file. The writer could make revisions in the source text, push F19 again and watch the effect of the revisions on the screen below.

```
The York Centre will to a degree differ from some already
existing models of university computer writing centres typically
serving composition courses or various word processing needs. York's
ambition is to have the Centre provide a special resource for three
areas of need: writing intensive courses, individualized instruction
by the Writing Workshop and Essay Tutoring Centre, and the more
traditional need for undergraduate word processing and text analysis.
Writing intensive courses at York are regular academic courses
designed to improve student writing in a wide range of disciplines.
More generally, they are courses that employ what has come to be
-----Emacs: test            (Text Fill)----Top-------------------------
/usr/local/wwb/bin/style -mm -li test
readability grades:
     (Kincaid) 16.1  (auto) 16.5  (Coleman-Liau) 14.7  (Flesch) 17.0 (24.6)
sentence info:
     no. sent 8 no. wds 201
     av sent leng 25.1 av word leng 5.38
     no. questions 0 no. imperatives 0
     no. content wds 122  60.7%  av leng 6.98
     short sent (<20) 13% (1) long sent (>35)  13% (1)
     longest sent 41 wds at sent 2; shortest sent 10 wds at sent 8
sentence types:
--**-Emacs: *Shell Command Output*    (Text Fill)----Top--------------------
```

Despite the marvels of **Emacs** and the power of a computer like the Sun 3/160s we use, the problem I raised earlier still exists. How do we use this software? To summarize the problems with most style and usage output, you could characterize the results as the same type of information overload that students must deal with when they receive back their innocent papers blotched with the zealous instructor's supposedly helpful notes. A wealth of information! But what to do with it?

For some of you, who may have gone through the gentle massage of Roger Garrison or Donald Murray, the answer may be easy. Build a series of student writing conferences that focus on a limited series of relevant revision priorities. Avoid the disincentive dump of extensive commentary.

From this perspective the limits of **HOMER** are a virtue: commentary on a very specific, manageable revision domain. I believe it's essential to place style and usage commentary early on in a composing process of multiple drafts; and that we must adapt the commentary to focus on one or two salient points relevant to where students are in any stage of a composition curriculum.

At my home institution, Sheridan College (Brampton, Ontario), we have developed several implementations of 'style' to give feedback specific to the appropriate rhetorical mode of an assignment. In several sections of Sheridan's first year Writing Lab course, for instance, students use *style* on four different occasions while word processing multiple drafts of four different writing assignments. Within each assignment the students use *style* for a different purpose. For instance, after the description assignment first draft, *style* feedback focuses on descriptive language and and sentence combining for sentence variety. The feedback limits itself to points students have previously covered in our **Sheridan Prewriter** 'Description' dialogue.

Style is a marvelous tool for an instructional designer. You simply write a script with holes available for inserting variables read from the *style* statistical table. What follows is the feedback for a description assignment. The notes include standard AT&T **Writer's Workbench** commentary mixed with my own text. My programmer and I spent about three hours producing it: not a major task.

****
/usr/randy/bench RANDY SMYE Thu Jun 26 10:55:49 1986
****

Here is some help in discovering ideas for the revision of your DESCRIPTION essay.

STYLE

The readability of your text ranges between grade 14.6 and grade 19.6. Good, clear messages generally fall between grades 7-11.

You've used a total of 191 words, with an average sentence length of 31.8 words per sentence. Professional writers aim for between 17-21 words per sentence.

You have 8 forms of the verb "to be". Descriptive writing should try to use concrete action words in place of the weaker "be" verbs. You might consider that in your revision.

You have 2 simple sentences and 3 complex sentences. For good sentence variety the difference between these figures should be less than 20%. While considering sentence variety, note that your sentences start 67% of the time with the subject of the sentence. That percentage will drop as you revise by using more prepositions, adverbs, verbs, or various joining words and phrases for sentence openers.

Adjectives supply descriptive detail for your narrative. Your sample contains 27 adjectives. Try increasing the number of adjectives that describe colour, sound, or size.

ABSTRACTION

Texts differ in the extent to which they refer to concrete objects and abstract ideas. Concrete objects, places, or things can be seen, heard, felt, smelled, or tasted. Abstract ideas, on the other hand, cannot be experienced by our senses. From the results of psychological research, we know that concrete texts are easier to read, easier to use, and easier to remember.

In your sample, 5.2 percent of the words are abstract words, which is a high score.

Texts with more than about 2% abstract words are abstract. A sample of Plato's Symposium has 4.41% abstract words. It begins, "For we have a custom, and according to our custom any service to another under the idea that he will be improved..." On the other hand, a text with no abstract words contains this passage: "The explosion of a gaseous mixture such as hydrogen and oxygen,..."

One way to improve such text would be to add concrete examples to explain the abstract ideas.

List of your abstract words

5 cost
1 professional
1 permission
1 original
1 development

1 amount

NOMINALIZATIONS

A nominalization is a noun created from a verb, such as "transformation" or "admittance". Research suggests that people will remember your description more effectively when you cut down the fuzziness of your nominalizations. Where possible turn your nominalizations back into their original direct verbs. For example "discussion" could become "discuss":
His discussion of the term paper was . . . .
He discussed the term paper . . . .
If you have used any nominalizations you will find them highlighted in their sentences before the final table of statistics. The table summarizes important statistics for a descriptive essay.

The *development* of increasingly efficient devices for the mechanical *reproduction* of printed material has meant that the amount of unauthorized *reproduction* of copyrighted materials has increased to a point where publishers and others are suffering serious economic effects.
sentence length: 38 COMPLEX : begins with article

Continued widespread unauthorized *reproduction* can only have the effect in the long run of decreasing the *publication* of specialized materials and increasing the cost of those that are published.
sentence length: 29 COMPOUND-COMPLEX :passive:: begins with adjective

In many cases the savings gained by *reproduction* rather than purchase of original material are considerably less than commonly supposed.
sentence length: 20 COMPLEX : begins with preposition

Any true estimate of the cost of *reproduction* must include not only the cost of the paper used but also the cost of rental or *amortization* and *maintenance* of the *equipment* and the costs of clerical or professional time involved.
sentence length: 40 SIMPLE : begins with adjective

These points are brought to your *attention* with a view to ensuring that the *reproduction* of copyrighted material be strictly limited to that available under the law , and that in other cases *reproduction* be avoided unless the written *permission* of the owner of the copyright is obtained.
sentence length: 47 COMPLEX :passive:: begins with adjective

readability grades:
19.6 - 14.6

sentence info:
no. sent 6 no. wds 191
av sent leng 31.8
no. questions 0 no. imperatives 0
no. content wds 101 52.9%

sentence types:
simple 33% (2) complex 50% (3)
compound 0% (0) compound-complex 17% (1)

word usage:
verb types as % of total verbs
tobe 62% (8) aux 31% (4) inf 0% (0)

passives as % of non-inf verbs 38% (5)
types as % of total
prep 16.8% (32) conj 4.7% (9) adv 5.2% (10)
noun 27.7% (53) adj 14.1% (27) pron 3.1% (6)
nominalizations 7 % (14)

sentence beginnings:
subject opener: noun (0) pron (0) pos (0) adj (3) art (1) tot 67%
prep 33% (2) adv 0% (0)
verb 0% (0) sub_conj 0% (0) conj 0% (0)
expletives 0% (0)

I'm still quite smug about having created this. But in all fairness I must credit Charles Smith and Kate Kiefer at CSU for their initial work in the early 80's on their own **Writer's Workbench** scripts. As composition specialists become more familiar with the concepts of software design, perhaps we may begin to see a brisk trade in these types of revision scripts, in the same way as many of us are developing computer aided prewriting heuristics.

Standing at the brink of a new domain of computer assisted revision, do we see any pedagogical effects to warrant the time and money we're forever seeking? Perhaps yes. Sheridan's research shows that students, with the encouragement of this revision software, seem to pursue in their subsequent drafts deep structure revisions as well as more obvious surface feature changes. Let me present one small experiment I tried.

For our freshman 'basic writers', our descriptive writing goals include descriptive language but also sentence variety. After a first draft, students encounter a few sentence combining excerpts from O'Hare and Memerling's **The Writer's Work** (1980). We show uncombined kernel sentences in

comparison to the professional writer's original. In both cases we include 'style' output for the texts involved.

Here is a specific example.

Kernels

No two classes of object could be more different. A meteor is a speck of matter. The speck is usually smaller than a grain of sand. The speck of matter burns itself up by friction. It tears through the outer layers of Earth's atmosphere. A comet may be millions of times larger than the entire Earth. A comet may dominate the night sky for weeks on end. A comet may look like a searchlight. The comet is really great. The searchlight shines across the stars. Something is not surprising. Such an object always caused alarm. The object was portentous. It appeared in the heavens. Calpurnia said to Caesar something. "Beggars die. There are not comets seen. The heavens themselves blaze forth the death of princes."
(O'Hare and Memerling)

Readability -- Kincaid: 4.0; Coleman-Liau: 6.7; Flesch: 7.1, 78.9
Sentences -- Average length: 6.9; Simple: 89%; Complex: 11%; Compound: 0%; Compound-Complex: 0%; Subject openers: 94%

Original

No two classes of object could be more different. A meteor is a speck of matter, usually smaller than a grain of sand, which burns itself up by friction as it tears through the outer layers of Earth's atmosphere. But a comet may be billions of times larger than the entire Earth, and may dominate the night sky for weeks on end. A really great comet may look like a searchlight shining across the stars, and it is not surprising that such a portentous object always caused alarm when it appeared in the heavens. As Calpurnia said to Caesar: "When beggars die, there are no comets seen; the heavens themselves blaze forth the death of princes."
(Arthur C. Clarke)

Readability -- Kincaid: 10.0; Coleman-Liau: 9.0; Flesch: 8.5, 64.6
Sentences -- Average length: 23.2; Simple: 20%; Complex: 40%; Compound: 0%; Compound-Complex: 40%; Subject openers: 60%

Of course at this stage the students also have *style* feedback on their own description first draft. In class I will remark briefly on the use of "linking" words: the term "linking" usually is adequate enough for me to avoid such deadly terms as adverbial and relative clauses. The students then work in small groups looking for similarities in their own work to either the kernels or the professional's original.

The goals for a subsequent second draft include work on concrete language but also work on reducing "kernels" and increasing the "relationships between ideas and word pictures". I've had the sense that my students' marked relish for this task is somehow related to the *style* software's ability to act as a concrete monitoring or feedback mechanism. I suspect that somehow the cognitive task of revision is becoming clearer for them, more goal oriented. Whatever the source, the results are impressive. Here is a before and after snapshot typical of what I'm finding.

Pat's First Draft
I parked the car and began to hunt around the car lot. The flags that surrounded the lot were clapping in the warm, gentle breeze. I began to admire the new cars that were on display but I wasn't alone for long. A largely built salesman with small beady eyes approached me. He right away tried to sell me a new car but I quickly disappointed him. I told him what I was looking for and he took me to a few cars which did not interest me at all. I was about to turn away when he tried one last time. He took me toward a bright red vehicle. I think this was it. The body was in excellent shape and there weren't very many scratches or dents. It was a 1979 Mustang with a 6-cylinder engine and I was surprised at how clean it was. The interior was black vinyl, very sleek-looking and it had its original AM radio. I took it for a test drive and was quite impressed. This was the car for me.

Readability -- Kincaid: 3.3; Coleman-Liau: 4.5; Flesch: 5.5, 94.8
Sentences -- Average length: 13.1; Simple: 36%; Complex: 14%; Compound: 21%; Compound-Complex: 29%; Subject openers: 93%

Pat's Second Draft
As soon as I parked the car I began to hunt around the car lot. While the flags clapped in the warm, gentle breeze, I admired the new cars that were on display. Suddenly, a largely built salesman with small beady eyes touched my shoulder. Quickly he tried to sell me a new car but I let him know right away that I was not interested. After explaining to him what I did want, he took me toward a few likely suspects. Again, there was nothing which appealed to me. Then, I noticed a bright red vehicle parked in the rear of the lot. Holding back my excitement, I enthusiastically approached the car. It was a 1979 Mustang with a 6-cylinder engine and I was surprised at how clean it was. The exterior was in excellent shape with hardly any scratches or dents. The sleek-looking black vinyl interior impressed me greatly, and it was equipped with its original AM radio. After taking it for a test drive I knew this was the car for me.

Readability -- Kincaid: 5.4; Coleman-Liau: 6.9; Flesch: 6.7, 82.9
Sentences -- Average length: 14.8; Simple: 42%; Complex: 42%; Compound: 8%; Compound-Complex: 8%; Subject openers: 33%

Note the increase in readability, average sentence length, and sentence type distribution, and too the decrease in subject openers.

In conclusion then I would urge you to re-examine style and usage software from two perspectives. First, the new technologies are finally bringing within our financial reach the possibility of a revision paradigm that is more interactive and recursive, more immediately available at any stage of the composing process. And second, we've got to mould the technology to match our curriculum goals. I remember Ray Rodrigues once remarking how prewriting software was in its infancy sharing the infancy of invention heuristics. Revision software is in a similar infancy shared by recent research which is only now beginning to probe the mysteries of the reviser's mental processes.

References

Lanham, Richard A. ***Revising Prose***. New York: Charles Scribner's Sons, 1979.

Memering, Dean and O'Hare, Frank. ***The Writer's Work:*** *Guide to Effective Composition* Englewood Cliffs: Prentice-Hall, 1980.

Randy Smye
Project Manager
Computer Assisted Writing Projects
Research and Development
Sheridan College of Applied Arts and Technology
Trafalgar Road
Oakville, Ontario
L6H 2L1
Canada
(416) 845-9430 ext. 277