CHECKING THE CHECKERS: COMPUTER EDITING PACKAGES

MARIE DOWLER   CATHERINE McEVOY
RYERSON POLYTECHNICAL   LOYALIST COLLEGE
INSTITUTE

ABSTRACT

Dowler and McEvoy compare three commercial software packages for checking English grammar and style: *RightWriter, Grammatik* and *Correct Grammar.* They have tested these programs on real and simulated examples of student writing, and have used them in the classroom and remedial centre with both ESL and native speakers. They compare these packages for reliability and ease of use, and speculate on the future for them in an academic setting.

Computer programs like *RightWriter, Grammatik* and *Correct Grammar* are best sellers, but do they have a place in the classroom or the remedial writing centre? It is not easy to correct fundamental grammatical and stylistic errors when there are three or four in every sentence. It is depressing for the student and repetitive and monotonous work for the teacher, whose training would be better employed in the human work of argument analysis and word choice. The large number of students now entering Canadian colleges and universities with English as their second language (ESL students) has compounded the problem. Thus a computer program that could reliably point out and explain grammatical mistakes, just as a spelling checker can now catch 99% of errors, would be a blessing indeed. Such grammar-checking programs might be valid teaching devices in the classroom: as they all explain their suggested changes and ask the user to verify them, they require the user to think about language.

This paper is limited to a discussion of three readily available DOS-based programs which will run on most personal computers found in an academic environment: *RightWriter* (Que Software), *Grammatik* (Reference Software International), and *Correct Grammar* (WordStar International). All cost under $120, need 640K or less of RAM, less than 5MB of hard disk space, and DOS 3.0. A colour monitor is recommended, though not essential. All provide interactive on-screen correcting, a printed copy of the edited file, and tutorial help on grammar topics.[1]

DEVELOPMENT OF STYLE AND GRAMMAR CHECKERS

The first style and grammar checkers which would run on personal computers with their limited memory were pattern matchers. *Grammatik* is the most sophisticated. For example, a phrase like "at this point in time" is considered wordy and redundant; the computer program treats it as a pattern, flags it every time it occurs in the text it is checking, and suggests a replacement. Over ten years of development, *Grammatik* has been elaborated to match complex patterns, including specified positions for any pattern in its sentences. *Grammatik* also allows its users great flexibility. They can select among the patterns the program offers; they can add their own patterns; they can even, with some programming, develop their own grammatical rules for the program to follow.

While pattern matching may help to reduce wordy phrases, split infinitives and other deviations of style, it is a relatively simple tool. *Grammatik*, for example, has been programmed to flag gender specific terms, and thus advises that the word "woman" had better be avoided and "person" used instead, even when the text it is correcting is about pregnancy and childbirth. Pattern matching is too crude to tackle the complex problems of correct grammar. All the programs, for example, associate "every" with a singular noun. Thus they flag as incorrect the phrase "every 17 minutes."

Long before pattern matching reached its present stage of development, other programs were using a different method based on the work of linguists. The best known of these is *Writer's Workbench*,

running under the UNIX operating system. *Writer's Workbench* parses text. It can classify words correctly 80-95% of the time. But a full version of its analysis of any text is daunting in length and complexity, even for a user with grammatical knowledge. Its component programs have to be broken down and selected for novice users.[2]

This parsing method of grammar checking was not available for personal computers until memory became cheaper and more available. Parsing involves context, and context takes enormous amounts of memory. Researchers at MIT found that eight sentences in their samples could be parsed in over 300 ways each (Wallraft 69). The partial solution to this problem was a set of algorithms based on the work of Francis and Kucera at Brown University. Francis and Kucera compiled the "Brown Corpus of Present-Day Edited American English," a collection of over 1,000,000 words from 500 samples of continuous written discourse, analysed to show the frequency of occurrence of each word.[3] The Brown Corpus was analysed into a "grammatically tagged" version. From his work on this analysis, Kucera designed the algorithms, the computing principles, used by *Correct Grammar*. The power of these algorithms comes from their attempt to deal with ambiguity by calculating the likelihood of any particular word having a particular grammatical function.

When it was first introduced, *Correct Grammar* was better at finding grammatical errors than either *RightWriter*, another pattern matcher, or *Grammatik*. But it mostly ignored deviations of style. Now all three programs combine the two methods. *Correct Grammar* in its latest edition (Version 4, 1991) seems to have added little in the way of parsing accuracy, but has included some style-checking elements and has increased the flexibility it gives users to customize it to their needs. On the other hand, *RightWriter* and *Grammatik*, by adding parsing to their arsenals, have greatly improved their ability to find grammatical errors while remaining strong in pattern matching. However none of the programs is yet reliable enough for a weak student writer to use unassisted.[4]

RELIABILITY

To assess the accuracy of these programs we used two tests. Our first test was based on sentences taken from the multiple-choice section of the College Diagnostic Writing Test, developed as a screening tool in 1982 by the Ontario Community College system. These sentences were augmented with others chosen from experience and frustration with certain problem areas, in particular, apostrophes, clichés, and lexical redundancies. Sixty grammar or style problems were identified. When we scored the computer programs, we found that in many cases the error was partially identified: for example, an apostrophe error might appear in a list of potentially misspelled words. In such cases we awarded the program a generous 0.5 for error detection.

To indicate the success rates in problem areas more clearly, we have arranged the errors in six categories; only the sixth category, style and word choice, contains stylistic elements such as passive voice. The other five categories refer to grammatical errors: "the students thinks," "their is," "lets try," "less opportunities." The following table indicates the success rate of the three programmes: of 60 errors, *Right Writer* identified 23; *Grammatik* 25.5 and *Correct Grammar* 22.

errors		*Right Writer* v. 5	*Grammatik* v. 5	*Correct Grammar* v. 4
subj./vb. agreement	(4)	0	2	2
other grammatical	(15)	2.5	6.5	9
apostrophes	(7)	3.5	4	2
other punctuation	(8)	2	1	1
hyphens/typing	(4)	1.5	2.5	2
style/word choice	(22)	13.5	9.5	6
TOTAL	(60)	23	25.5	22

Figure 1. Errors Detected in Test Sentences

The second test used three sets of student papers. Two of these sets were examination papers from a first-year composition course, assessed D or F, one set written by native English speakers and the other by ESL students. We felt these papers were the toughest test we could set the checkers. The third set contained fewer mistakes. It consisted of extracts from student papers written for a core professional course, selected by an instructor concerned about the standard of writing. Rather than categorizing errors for this set in the same way, we used the classification system suggested by Maxine Hairston (1981). Hairston polled approximately 101 professionals from all areas (except English teaching), showing them 65 faulty sentences and asking them which errors, in her words, "bothered them a lot," "bothered them somewhat," "bothered them a little," and "hardly bothered them at all" (Hairston, 795). From their replies she compiled 5 classes of errors: those that were unacceptable (here marked X); very serious (1); serious (2); moderately serious (3); minor or unimportant (4). Figure 2 indicates the number of errors in each class identified by the three packages.[5]

Category	Human	RWriter	Grammatik	CGrammar
X	73	9.5	12	17
1	106	18	42	47
2	138	10.5	18	12.5
3	57	13.5	8	7
4	34	19	17.5	10.5

Fig. 2. Errors Identified using Hairston's Categories

Although the quantitative figures produced by the two tests are not impressive, at least some of the more blatant errors are correctly diagnosed. Every program is correctly diagnosing more errors now than its previous version was a year ago. *RightWriter* and *Grammatik* have more than doubled their scores. The programs can catch subject/verb agreement error, as long as it is not complicated by intervening words or reversed sentence order. This is helpful for ESL students, but less useful for native English speakers. Incorrect verb forms are generally accurately flagged. Both *Grammatik* and *Correct Grammar* offered corrections for

"I haven't ate" and "my ears were froze," though only *Correct Grammar* caught "If he had only went." The programs excelled in identifying the passive voice. They also picked up clichés and wordy expressions. Unfortunately the advice both programs give is sometimes inappropriate, sometimes just wrong. Sometimes the analysis given is correct, but following the advice for correction leads to further errors. All programs flagged some correct expressions as incorrect. In fact, as grammar and style checkers have become more reliable in identifying errors, the frequency of non-errors flagged has also increased.

EFFECTIVE USE

The use of style and grammar checkers met with some success in a pilot project conducted at Loyalist College in 1990-91. *Correct Grammar* and *Grammatik* were incorporated into a Writing with Computers course. The students, all native speakers of English, were required to use one of the checkers with one of their assignments. The programs were introduced as tools to improve correctness, with the caution that they were not omniscient: if the students disagreed with the checking program and did not want to change a particular sentence, that was fine. No other demands were made.

In open-ended evaluation questionnaires, all students reported the checkers as helpful. When they were asked whether they would recommend making the programs available to the following year's students, the answer was a unanimous yes. Stronger writing students found the packages less useful, while several weak writing students began to use them on all assignments. All students found the revision suggestions and "help screens" useful.[6]

These suggestions and "help" screens are imperative if the checkers are to be used in a teaching environment. For example, all the programs use the term "passive voice." But grammatically weak students do not understand the term, and so they do not know how to make alterations. They need simple tutorials and on-line help if they are to explore grammatical concepts and think about their writing.

Also, if computer style and grammar checkers are to be useful in teaching, these programs must be made easy to customize. The advice the programs give can be misunderstood or lead to fresh errors. For example, reviewing the sentence, "Wanting to win a place on the Olympic team, an athlete prepares himself not only by exercising and he follows a rigid diet," *Right Writer* suggests "Split into two sentences." We know where to make the split, but a weak student is likely to place the period after "team," unless a warning is given. *Grammatik*'s response to the same sentence is "A long sentence may be difficult for your reader to read. Rewrite it so that it contains only one thought." This is of little use to teachers who have just spent days on the principles of subordination. They may wish the screen to read, "This is a long sentence where you are most likely to make mistakes. Check it carefully." Currently they can define the length of a "long sentence," but they cannot change the advice that appears. Only *Grammatik* and *Correct Grammar* allow some of this kind of customizing. All examples of phrases that *Grammatik* will flag can be inspected. The user can order the program to ignore some or even all of them, can alter the advice that appears with them, and can add new phrases to be flagged. Unfortunately *Grammatik* will not allow the writing teacher to change the advice screens for common grammatical rules so that they reflect personal teaching style and use the terms the students are used to. This flexibility has to be one of the keys to the programs' increased usefulness in a teaching/learning environment.

Another key is the level of interaction with the user. Errors are presently missed or miscorrected, because the computer, using its algorithms, assumes a certain structure for a sentence, but does not ask the user to ratify the assumption. Here is a sample sentence from an ESL student:

But in order to help reducing the percentage of assault of women and stop the number of violent images and pictures in movies and magazines government should take a step forward.

All that *Right Writer* and *Grammatik* could do with this was to note that it was a long sentence, that "in order to" was "wordy" and should be replaced by "to." But *Correct Grammar* suggested that "magazines

government" should have an apostrophe. Thus it seemed to have broken the sentence into the following parts:

> But / in order to help reducing the percentage of assault of women and stop / the number of violent pictures in movies and magazines government should take a step forward.

If the user could be shown this sentence division and allowed to correct it, some of the problems of multiple parsing possibilities could be avoided. At the least, the user would be forced to study the sentence.

Correct Grammar was the first of the three programs to take a step towards identification of sentence structure. In some, but not all of its flaggings of subject/verb agreement errors, it highlights the subject and verb, as it identifies them, in different colours. This is a great help. However, reviewing a sample from our first test, "We look forward to a better tomorrow where we wont make mountains out of molehills...," *Correct Grammar* gave the dubious advice: "Consider wonts instead of wont." We know that "wont" is not part of the intended subject of "make"; so do most of our students. If *Correct Grammar* had identified subject and verb in this example, the students might have recognized not only the program's mistake but also their own. The latest update of *Grammatik* (5 1991) offers an on-line analysis of each sentence into its parts of speech, though it does not identify subjects and objects. *RightWriter* (Version 5, 1992) has an even more detailed "parse tree" accessible at all times. These are a great help to the writing teacher, who can see where and why the program is making mistakes, but they are not yet in a form that the weak student writer can use.

To summarize, grammar and style checkers are improving with each new version. As the programs' parsing abilities increase, as they provide better tutorial help, and as they provide for customizing, they will become useful tools in the classroom, remedial writing centre, or study. None is reliable enough to stand alone yet. Ten years ago, many weak student writers' attitude was that they didn't need to know about apostrophes (or spelling, or pronouns, or ...) because their secretaries would make corrections for them. Today, many students believe they still don't need to know because a computer will do it for them. If it

does nothing else, use of grammar and style checking programs in the classroom now will dispel that myth.

NOTES

1. All will run on a PC XT, though an AT is recommended for *Grammatik; Correct Writer* requires only 512K RAM and DOS 2.0 or higher. Other style and grammar checkers are available, if the user has the access to the necessary hardware. *Writer's Workbench* (AT & T) requires the UNIX operating system, and a new contender, *PowerEdit* (Artificial Linguistics), recommends a 386, 2MB RAM and 10MB of hard disk space. It also costs $295.00

    Outside the scope of this paper are three computer programs developed within and specifically for college writing courses: *HBJ Writer* (originally known as WANDAH, developed at UCLA, published by Harcourt Brace Jovanovitch); William Wresch's *Writer's Helper* (Conduit, University of Iowa) and Elaine and John Thiesmeyer's *Editor* (Modern Language Association). These programs have some of the functions of grammar and style checkers directed to the general professional user, and they also contain components directed to composition classes, such as suggestions for invention, for paragraph structure, and argument analysis.

2. As well as working under the UNIX not DOS operating system, it is also very expensive. AT & T have marketed their Collegiate Edition as a complete writing laboratory, with a central computer serving individual terminals, and costing over $25,000. Randy Smye at Sheridan College in Toronto has successfully adapted its subsets for basic writing courses.

3. See H. Kucera and W. Nelson Francis, (1967), *Computational Analysis of Present-Day American English (Providence, Rhode Island: Brown UP)*; W. Nelson Francis, (c. 1982), *Frequency Analysis of English Usage: Lexicon and Grammar* (Boston: Houghton Mifflin); Geoffrey Leech, (1991), "The State of the Art in Corpus Linguistics" (in Karin Aijmer and Bengt Altenberg, (1991), *English Corpus Linguistics* (London: Longman).

4.   In fairness to the programmers, the three programs are directed to the professional, commercial world, not to schools. The programmers could assume that the users would have some writing competence, that on the whole they would make occasional, discrete grammatical mistakes, and would have learned some bad stylistic habits, particularly overuse of the passive voice and probably of too many words altogether. For such writers, all three programs can offer accurate and helpful suggestions.

5.   This test indicates our bias in pedagogy towards work-place practice. Those favouring academic criteria should treat Hairston's first two categories as one, including nonstandard verb forms, subject/verb agreement, fragments, run-on (though non comma-spliced)sentences, adverb/adjective confusion, faulty grammatical parallelism. However, they might disagree over category 2, which treats tense shifts, dangling modifiers, and pronoun case confusion as equivalents to omitting a comma after "however," or in a series, and all of these as worse than a comma splice, which is found in category 3.

6.   *Correct Grammar* was seen as giving more useful suggestions. The students were using version 3 (1990) and version 4 of *Grammatik*. Version 5 of *Grammatik* (1992) has enhanced "help" messages and screens.

REFERENCES

Hairston, M. (1981). "Not All Errors Are Created Equal: Non-academic Readers in the Professions Respond to Lapses in Usage." *College English*, 43: 8, 794-806.

Leech, G. (1991). "The State of the Art in Corpus Linguistics." In K. Aijmer and B. Altenberg, *English Corpus Linguistics*. London: Longman. 8-29.

Wallraft, B. (Jan. 1988). "The Literate Computer." *Atlantic Monthly*, 261, 64-71.

Computer Software

Correct Grammar 4. (1991). WordStar International, Novato, California.

Grammatik 5. (1991). Reference Software International, San Francisco.

RightWriter 5. (1992). Que Software (Prentice Hall Computer Publishing), Carmel, Indiana.

PowerEdit. (1991). Artificial Linguistics, Dallas, Texas.

Writer's Workbench. (1985). Collegiate Edition Software, AT & T Information Systems.

* * * * * * * * * *

Marie Dowler is a member of the English Department and Director of the Writing Centre at Ryerson Polytechnical Institute, Toronto; Catherine McEvoy teaches Communications and Writing with Computers for the Human Studies Division at Loyalist College, Belleville, Ontario. She is currently seconded to develop a computer-assisted learning centre, CAL CENTRAL.