

A Technology-based Solution to Reduce Time Spent Identifying and Commenting Writing Errors in Research Papers

Charles L. McDonald, Jr., Ph.D.
Texas A & M University - Texarkana
Charles.McDonald@tamut.edu

Theresa McDonald, Ph.D.
Texarkana College
tmcdonald@texarkanacollege.edu

Abstract

The increasing demand for Web-based courses and larger classes is forcing professors to seek technology-based solutions to address increased workloads (Mupinga & Maughan, p. 17). Four years ago, two professors, with a history of software development, felt that their productivity was being suppressed by spending too many hours commenting grammar errors in electronically submitted research papers. These professors felt that a more efficient methodology to provide consistency and detail in commenting writing concerns was needed to better serve their students (Bonnell, p. 290).

Introduction

The *Emerging Technologies in MIS* graduate course at Texas A&M University-Texarkana was used for this study. It is a Web-based course in which proposals, submissions, feedback, and evaluations are communicated via electronic media. Each student was required to develop six 2,000-word research papers and each research paper required a preapproved proposal. Proposals were received, reviewed, and returned via e-mail, usually within 24 hours. Successful proposals addressed an appropriate topic, contained a short abstract, included a statement of scope, and cited at least four credible references. All cited references must be available online via electronic databases. Additional references could be added to enhance content during the paper's development. As the course addressed emerging technologies, references could not be more than six months old without the professor's approval.

Student's research papers were submitted via the www.turnitin.com site. Student accounts were established on Turnitin's site to include three folders for each paper. Submissions to the draft folder provided students with originality reports concerning plagiarism without adding their paper to Turnitin's database (Key Questions Students Ask About Turnitin). Students were allowed to resubmit papers into the draft folder as needed. The second folder was provided for a student's first submission to the professor and, if the paper was returned for revisions, the third folder was available for final submissions. The second and third folders allowed only one submission each.

As the semester progressed, professors would retrieve submitted papers from the www.turnitin.com site. Non-plagiarized submissions were reviewed and either accepted for credit or, in most cases, returned via e-mail to the student for revisions with comments that identified grammar, style, and content concerns. On most first submissions, a significant portion of the professor's reviewing effort was spent commenting grammar problems. Returned papers could be revised and resubmitted for a final professor's review. Final submissions received either an accepted or rejected grade.

Problem

The increasing demand for Web-based courses and larger classes is forcing professors to seek technology-based solutions to address increased workloads (Mupinga & Maughan, p. 17). Four years ago, two professors, with a history of software development, felt that their productivity was being suppressed by spending too many hours commenting grammar errors in electronically submitted research papers. Each professor had responsibility for evaluating about 450 proposals and research papers from graduate students enrolled in the Web-based *Emerging Technologies in MIS* course at Texas A&M University-Texarkana. Given that an average-quality 2,000-word submission required about 30 undisturbed minutes for evaluation and that a poorly written paper demanded at least an hour, it was difficult to provide each submission with consistent quality analysis and feedback when evaluating multiple electronically submitted research papers. These professors felt that a more efficient methodology to provide consistency and detail in commenting writing concerns was needed to better serve their students (Bonnell, p. 290).

Design of the software tool - capabilities and interface

To identify writing concerns that would need to be addressed by a software solution, repetitive errors were documented from previously evaluated papers. A non-inclusive list of our initial findings is depicted in Table 1.

Table 1

Clichés, colloquial expressions, and conversational writing
Sentences longer than 30 words
Misuse of conjunctive adverbs
Starting a sentence with inappropriate words
Use of first or second person
Repeated words
Use of words not listed in the dictionary
Ending a sentence with a preposition
Use of weak word choices
Use of a slash to imply word choices
Use of one-sentence paragraphs
Use of etcetera in formal writing
Ending sentence with a preposition
Use of contractions
Paragraphs with more than 10 sentences
Excessive use of specific words
Other grammatical errors

As papers were submitted in Microsoft's Word format, it seemed appropriate to address the solution using Microsoft's Visual Basic for Applications (VBA) to develop a set of macro programs to include a menu for the user interface (Marino, p. 137). As the application was intended to assist in editing research papers, it was named *Edit Assist*.

Discussions concerning comments on previously reviewed papers led to identifying specific needs, which concluded that the software tool would need to perform in three modes. In the automatic mode, it would need to scan documents locating common grammar and writing errors and insert helpful and descriptive comments. In the semi-automatic mode, it would need to insert predesigned comments for highlighted text by clicking one of about twenty menu choices. In the manual mode, it would need to insert blank comments for the evaluator's remarks. A utility button was needed to delete a group of comments associated with highlighted text. The menu buttons text and associated comment text are depicted in Table 2.

Table 2

Button Text	Comment Text
Blank Comment	
Parallel Construction	Parallel construction error - The members of a series must be all nouns, all infinitives, all prepositional phrases, all gerunds, or all clauses.
Choice of Words	Questionable, inappropriate, weak choice, or order of word(s) - needs revision
Grammar Error(s)	Text contains grammar error(s) - possibly missing word(s)
Avoid Clichés	Avoid clichés, colloquial expressions, and conversational writing
Sentence Structure	Sentence subject/verb tense or structure problem - needs revision
Nebulous Content	Nebulous, rambling, or confusing content - needs clarification of meaning, focus on topic, or deletion
Personal Scenario	Avoid reflecting scenarios or personal opinions unsupported by references
Delete and Revise	Delete text and revise sentence if necessary
Redundant Content	Redundant content - addressed previously in document
Split Infinitive	Avoid split infinitives - revise as needed
Punctuation Error	This text contains one or more error(s) in punctuation (e.g., a missing comma)
Incomplete Sentence	This is an incomplete sentence - revise as needed
Acronym not Defined	Acronym not properly defined (e.g., 'Computer Based Training (CBT)')
Wordy Writing	Avoid wordiness or disorganized focus of topic (e.g., replace 'in spite of the fact that' with 'even though')
Choppy Sentence(s)	In formal writing, choppy sentences (rapid switching of topics using short sentences) are to be avoided. Consider adding conjunctions, clauses, subjects, and/or verbs to make your writing more intricate.
Misuse of 'the'	Avoid using 'the' to imply 'all' (e.g., 'The pets are an important part of childhood development.')
Use Lower Case	Do not place words in proper case without cause
Content Problem	Inaccurate, incomplete, or misleading content
Reference Format	References contain incorrect formatting - see Reference Examples page on course's website or visit the APA Style Form Guide site at

Button Text	Comment Text
	www.apastyle.org
Review Stopped	The professor's review stopped here. This does not imply that all errors above or below this point were commented. Consider contacting the writing lab for assistance in developing sentences.
Delete Comments	Deletes comments associated with highlighted text
Report Summary	Inserts first comment at top of document. See Figure 1.
Scan Document	Scans the open document applying about 200 grammar rules to insert comments for each highlighted discovery

During alpha testing of the software, it was realized that a summary report to include statistics was needed. Figure 1 depicts a summary report generated during alpha testing. Although passive voice is noted in the sample summary report and is a very reliable rule, it was deactivated for use in this course.

Figure 1

Comment [1]:]: Summary Report:
 172 concerns are commented
 Left justification was not used throughout document
 Multiple fonts were utilized in document
 The word "that" was used 37 times.
 Note that a grammar or sentence structure problem is usually created by simply deleting the excessively used word without revising the sentence.

The use of passive voice was commented. Although the use of passive voice is not a grammar error, it is a stylistic issue that can weaken the clarity of writing. Although most reviewers prefer the active voice, the use of passive voice may be appropriate when developing technical or scientific papers.

Document Statistics:
 Words - 2,331
 Characters - 11073
 Paragraphs - 29
 Sentences - 119
 Sentences per Paragraph - 5.1
 Words per Sentence - 16.9
 Characters per Word - 5.1
 Flesch Reading Ease - 44.7
 Flesch-Kincaid Grade Level - 11

Writers should not assume that all errors are commented.

The Flesch Reading Ease and Flesch-Kincaid Grade Level measurements are included in each summary report. In 1948, Dr. Rudolph Flesch, an Austrian, devised what is considered to be the most accurate readability formula, the Flesch Reading Ease Formula. The range of measurement is 0 to 100 with higher numbers assuring the text is easier to read. As an example, the average 5th grader should easily understand text written with a

Readability Ease (RE) of 90 to 100; whereas, a score of 0 to 30 would be more appropriate for college graduates (The Flesch Reading Ease Readability Formula). Later in the year, Dr. Flesch co-authored the Flesch Grade Level Readability Formula with Dr. John P. Kincaid. In 1976, it was modified by the US Navy and renamed the Flesch-Kincaid Grade Level. Most educators have since adopted it as a standard. The results can be interpreted as grade levels; although, in Microsoft Word, 12 is the maximum grade level reported (The Flesch Grade Level Readability Formula).

Use of the software

Papers were reviewed by downloading files to a folder and clicking a filename to open a document in Microsoft Word. The *Edit Assist* menu was available anytime Word was open. In Word 2003, an easy to use resizable, dropdown menu is available, but in Word 2007, only a fixed position *Ribbon* is available. The scanner runs in either version; although, the usage of a ribbon requires more mouse navigation to reach menu buttons. By clicking the Scan Document menu button, about 200 grammar and writing rules are activated to scan and comment common writing concerns in the open document. Once the scan completes, in about 30 seconds, the professor could start reviewing the paper utilizing menu options to add predesigned comments or blank comments to sections of highlighted text. When the reviewer either completed the paper's review or stopped the paper's review because of excessive errors, a menu button was available to insert the summary report as the first comment. After saving the reviewed document, it was returned to the student for resubmission or recorded as an accepted or rejected grade.

Data Analysis

The purpose of this study was to determine if an automated tool could be developed that, in use, would reduce the professor's mundane editing efforts and provide students quality feedback concerning writing concerns in research papers (Collis, De Boer & Slotman, p. 307). During the spring 2007 semester, 896 research papers were evaluated using this tool. Of this collection, 762 were retained for this study. Papers were excluded from this study if they revealed evidence of plagiarism or represented a single submission. Single submissions included papers that were either accepted for credit or not resubmitted for evaluation. To provide evidence of feedback, only papers that were submitted and resubmitted were used in this study. Grades assigned to papers were outside of the scope of this study.

To collect meaningful data for this study, a macro was created using VBA that would scan all comments in each paper to detect occurrences of specific words and document the topic of each comment. The result of this effort provided data grouped by columns for first and second submissions. Table 3 depicts data from six paper submissions. The top two rows identify the submission number (first or second) and the total number of comments on each submission. The next section of rows represents the number of comments that were automatically inserted by the scanning software. The lower section of rows depicts the number of comments that were generated via menu selections. The bottom row reveals the number of manually entered comments for each paper.

Table 3

Submission	1	2	1	2	1	2	1	2	1	2	1	2
<i>Error Type</i> Overused words	1	0	1	0	1	1	1	1	1	0	0	0
Long sentences	4	0	0	0	4	1	6	0	1	0	0	0
Use of clichés	2	1	4	0	1	0	0	1	1	0	1	0
Use of slashes	0	0	0	0	0	0	0	0	0	0	0	0
Long paragraphs	0	0	0	0	0	0	0	0	0	0	0	0
Avoid starting a sentence....	6	1	0	0	0	0	2	0	0	0	0	0
One-sentence paragraphs	0	0	0	0	0	0	0	0	0	0	0	0
First or second person	1	0	0	0	0	0	3	0	0	0	0	0
Overuse of "he or she"	0	0	0	0	0	0	0	0	0	0	0	0
Comma missing "which"	1	0	15	0	6	0	1	0	0	0	0	0
avoid use of contractions	1	0	0	0	1	0	0	0	0	0	0	0
Misuse of conjunctive adverb	2	0	2	0	0	0	1	0	0	0	0	0
Ending with a preposition	0	0	2	0	1	0	0	0	0	0	0	0
Internet	4	0	0	0	0	0	1	0	0	0	0	0
Weak choice of words	0	0	0	0	0	0	0	0	0	0	0	0
Revise sentence	0	0	0	0	0	0	0	0	0	0	0	0
Avoid ending a sentence	1	0	2	0	1	0	0	0	1	0	0	0
Misc. grammar errors	10	3	7	2	15	3	18	2	17	2	8	2
Software generated - Totals	33	5	33	2	30	5	33	4	21	2	9	2
<i>Menu Buttons</i> Parallel Construction	0	0	0	0	0	0	0	0	0	0	0	0
Choice of Words	0	0	3	0	0	1	0	0	2	0	0	0
Grammar Error(s)	1	0	0	0	1	1	0	0	0	0	0	0
Avoid Clichés	1	0	3	0	1	0	8	0	2	0	0	0
Sentence Structure	3	0	7	0	4	1	0	0	0	0	0	0
Nebulous Content	1	0	1	0	3	0	0	0	1	0	0	0
Personal Scenario	0	0	0	0	0	0	0	0	0	0	0	0
Delete and Revise	1	0	1	0	3	0	2	0	1	0	1	0
Redundant Content	0	0	1	0	0	0	0	0	0	0	0	0
Split Infinitive	0	0	0	0	1	0	0	0	0	0	0	0
Punctuation Error	6	0	1	0	3	0	5	0	1	0	0	0
Incomplete Sentence	1	0	0	0	0	1	0	0	0	0	0	0
Acronym not Defined	0	0	0	0	0	0	0	0	1	0	0	0
Wordy Writing	4	0	0	0	0	0	1	0	2	0	1	0
Choppy Sentence(s)	0	0	0	0	0	0	0	0	0	0	0	0
Misuse of "the"	0	0	0	0	0	0	0	0	0	0	0	0
Use Lower Case	0	0	0	0	0	0	0	0	0	0	0	0
Content Problem	0	0	0	0	0	0	0	0	0	0	0	0
Reference Format	0	0	0	0	0	0	0	0	0	0	0	0
Auto-menu - Totals	18	0	17	0	16	4	16	0	10	0	2	0
Typed comments - Totals	15	1	17	1	18	4	5	1	2	0	7	0
All comments - Totals	66	6	67	3	64	13	54	5	33	2	18	2

Findings

During this course, there were 16,020 comments inserted across 742 research papers. The Scan Document button automatically inserted 5,823, or 36%, of grammar related comments. Use of the highlight text and select a menu button method (auto-menu method) generated 4,742, or 30%, of the comments. The remainder, 5,455 comments, or 34%, were entered via the highlight text, select Blank Comment, and manually type a comment that usually related to content or the paper's organization.

First submissions contained 5,249 software-generated comments; whereas, second submissions contained only 574 comments. The auto-menu comments represented 3,909 comments on first submissions, which reduced to 833 comments on second submissions. Typed comments totaled 4,478 comments on the first submissions, but fell to 977 comments on second submissions.

Papers averaged 36 comments on first submissions and 6.3 comments on second submissions. The average number of software-generated comments on first submissions was 14, which reduced to 1.5 on second submissions. The average number of auto-menu generated comments on first submissions was 10, which dropped to 2.2 on second submissions. The average number of typed comments on first submissions was 12, which condensed to 2.6 on second submissions.

Although the number of deletions of incorrect auto-generated comments was not documented, some papers did not require any adjustments; others required three or four comments to be removed. Most of these issues were associated with the rule to determine whether a comma was appropriate to precede the word "which". After the semester, this rule was deleted as the decision appears to be too subjective for rule-based decisions. Occasionally, there were comments that required attention, such as an author named "Dr. You" being commented as a "Use of second person". The authors felt that, overall, the auto-generated comments were correct more than 95% of the time.

By using *Edit Assist*, 66% of comments were either auto-generated or inserted via a mouse click after highlighting text. Although, the remaining 34% of comments required manual entry, *Edit Assist* eased the operation by eliminating layers of menu navigation to insert a blank comment. Both authors feel that their workload concerning the editing of research papers was reduced by about 60%. This software will not evaluate content, but it does eliminate a majority of the mundane editing efforts associated with the review of research papers.

References

Bonnel, Wanda "Improving Feedback to Students in Online Courses" *Nursing Education Perspectives*, Vol. 29, No. 5, October 2008, p. 290-293.

Collis, B., W. De Boer & K. Slotman "Feedback for Web-based Assignments" *Journal of Computer Assisted Learning*, Vol. 17, 2001, p. 307.

“Key Questions Students Ask About Turnitin” 3 Oct. 2009
http://www.turnitin.com/resources/documentation/turnitin/sales/7_Things_Educators_Should_Know.pdf

Marino, Dale J. “Physiologically Based Pharmacokinetic Modeling Using Microsoft Excel and Visual Basic for Applications” *Toxicology Mechanisms and Methods*, Vol. 15, No. 2, Mar/Apr 2005, p. 137-154.

Mupinga, Daison M. and George R. Maughan “Web-based Instruction and Community College Faculty Workload” *College Teaching*, Vol. 56, No. 1, Winter 2008, p. 17-21.

“The Flesch Grade Level Readability Formula” *Readability Formulas*. 3 Oct. 2009
<http://www.readabilityformulas.com/flesch-grade-level-readability-formula.php>

“The Flesch Reading Ease Readability Formula” *Readability Formulas*. 3 Oct. 2009
<http://www.readabilityformulas.com/flesch-reading-ease-readability-formula.php>