

Understanding What Software-Based Editing Can – and Can't – Do for You

Some things that authors should consider



I have, on several occasions, spent time going through the report produced when a software editing tool was used to assess one of my client's manuscripts. Such reports bring into sharp focus some of the limitations of these tools and the importance for a writer to understand the tool's recommendations in the proper context.

The examples in this tutorial are from actual reports. They are not meant as a criticism of any software company or the product that produced the report. They are intended only to be illustrative of the caution that's needed in blindly applying the recommendations of any software editing tool without consideration of the context.

Such tools can be helpful at a certain level, and the summation at the end of the tutorial acknowledges this value. Authors should, however, be aware of the limitations of software and make informed decisions about their own manuscripts.

Happy writing!



The Self-Editing Conundrum

For many, editing is a frustrating or even painful process that leaves you wondering if it's even worth it to continue



- Writers are storytellers
- With some exceptions, we're not expert grammarians
- We're not experts in the knowledge or application of style guides
- We may or may not be expert spellers, and constantly going back and forth to the dictionary is highly disruptive
- We know the story too well, so it's easy to overlook things

So what's an author to do?

a) If you want to do as much self-editing as you can

- Use tutorials to fill your knowledge gaps on grammar and style (there are many available, including on this website)
- Use that dictionary – this is the copy editing phase of your writing journey, after all
- Consider a software editing tool for suggestions
- Get as far as you can on your own and then engage a professional editor to help you polish the work

b) Engage a professional editor early in the process to benefit from their knowledge of grammar and style and to be your partner in polishing the manuscript



The Pitfalls of Software-Based Editing

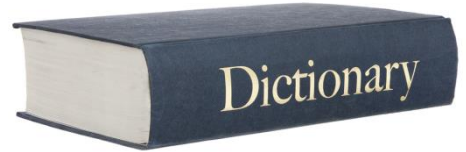
- This is somewhat over-simplified to help non-technical people easily understand
- Software can only apply the rules that are programmed into it
- Editing software works by comparing your manuscript to its known set of rules for the elements of grammar, style, spelling, etc.
- Some of these rules will be hard-coded into the program
- If the tool uses some modern, AI-like techniques, some of its internal rules may be dynamically changed by “learning” from the manuscripts it assesses
- Rules can only take you so far. Nuance and context are equally important – and often, more so.
- Software is not yet a flawless replacement for the human mind.

- References to “CMoS” in this guide are for the *Chicago Manual of Style, 17th Edition*
- References to M-W are to Merriam-Webster’s Dictionary, both print and online



The Pitfalls: Spelling

- Software references its internal dictionary(ies)
- Know what these are and know if multiple languages are available
- Understand the limitations of what's provided



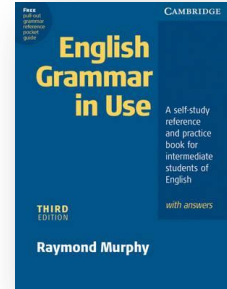
Where software may give false reports

- **Non-English words**
Unless the software can use multiple languages simultaneously, all non-English words will be flagged as spelling errors.
- **Colloquialisms or dialect**
Many tools flag colloquialisms (such as “oughta” or “gonna”) as spelling errors. These are perfectly acceptable in fiction (per CMOs) and often occur in dialogue.
- **Proper names**
Common proper names won't be flagged, but things even slightly out of the ordinary will be identified as spelling errors.
- **Place names and other geographical names**
Software may accept “Atlantic Ocean,” but I've recently seen an example in which “Tyrrhenian Sea” was flagged as a spelling error, despite being a properly spelled geographical feature.
- **Words that are actually in the dictionary**
A recent report I saw flagged “indignance” as a spelling error, despite the fact that it appears in M-W.
- **Past tenses**
A recent report I saw flagged “knelt” as the British spelling and not US spelling, despite the fact that “knelt” is M-W's first preference for the past tense of kneel.
- **Odds and Ends**
A recent report I saw flagged “l'd” as a spelling error when it occurred in the phrase “l'd have been lost if you hadn't shown me the path.”



The Pitfalls: Grammar (1 of 2)

- Software has grammar rules built in
- Grammar is sometimes dependent on context
- Grammar rules have changed since the 19th century



Where software may give false reports

▪ Dialogue

Not every character speaks grammatically perfect English. Your dialogue needs to be true to the unique voice of each character, whether it's poor grammar, regional dialect, or written-out pronunciations of a character's accent. Software will often flag these things as bad grammar. When it does, you should ignore it.

Software will often flag turns of phrase used in dialogue as poor grammar when that's what a character would say in real life, even if it's not how you would write it if it were part of the narrative portion of the story. Again, feel free to ignore it.

▪ Ending a sentence with a preposition

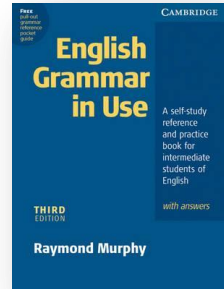
The prohibition on this has been debunked. It was a 19th century rule concocted by a grammarian who believed that English grammar should follow the principles of Latin grammar in order to be "high minded." Ignore any suggestion by software that ending a sentence with a preposition is prohibited – just be aware of flow and use the construct that will be the most natural for the reader.

▪ Split infinitives

The proscription on use of split infinitives is archaic advice from the 19th and early 20th centuries. Grammarians now fully support appropriate usage of split infinitives. CMOs Sections 5.108 and 5.171 not only allow the use of split infinitives but even encourage the practice, especially when the modifier impacts the reader's understanding of the verb usage. Ignore any suggestion by software that this is prohibited – just be aware of the meaning you're trying to convey and use the construct that's most suitable.



The Pitfalls: Grammar (2 of 2)



Where software may give false reports

■ Context

Software can often fail to understand the context in which a certain phrase or construct is used (dialogue is just a special case of this). Here are some examples from recent reports I've seen where the software erroneously flagged a grammar error because it failed to grasp the context:

The sentence: “. . . the Germans are giving orders for an all-out push to the beach.”

The context is of American officers listening in to German radio communications. The software claimed “The word "for" is not used after "order" in this context. Try just using 'orders an'.” This misses the context entirely. One might write “the captain orders an all-out push” – but that's not the context of this usage.

A hyphenated word (straight-on) used as an adverb. The recommendation from the software followed the rules for adjectives and completely missed that the usage was as an adverb.

The sentence: “Are you nuts?” The software reported “The pronoun 'you' must be used with a non-third-person form of a verb: 'nut'.” It somehow decided that “nuts” was being used as a verb when, in fact, it was an adjective.

The sentence: “It's kind of a nuisance.” The software reported “Don't include 'a' after a classification term. Use simply 'kind of'.” Wrong advice. To say “It's kind of nuisance” would be completely incorrect.

The phrase: “as much like the one I had as I could make it” when discussing a child's toy. The software reported “Comparison is written 'as much 'as'.” The software completely missed that the “as much . . . as” comparison is actually in the sentence and that the phrase “like the one I had” is just a qualifying descriptor within the comparison.

Do not blindly accept the grammar assessment and recommendations of a software tool without considering the context.



The Pitfalls: Punctuation

- In my experience, software editing tools seem to be a bit better in this area
- But there are still areas where they get things wrong



Where software may give false reports

■ Ellipses

If you follow CMoS and use spaced periods (. . .) for ellipses, many software tools will falsely flag these as periods that don't require a space before the period. In a recent report I saw, the tool reported 477 instances of erroneous spaces before periods – they were all CMoS-style ellipses.

■ En dashes

Some tools will try to enforce the rule that there should be no space on either side of an en dash. This rule is true for em dashes but is more flexible for en dashes. If you use en dashes with a space on each side, either for rhythm or for visually setting off an explanatory phrase in a complex sentence (both of which I recommend to clients), ignore these false flags.

■ Commas with words such as “either,” “anyway,” “nonetheless,” etc.

When CMoS changed the rules such that a comma is no longer required before “too” or “also” unless it's essential for clarity in a complex sentence, they apply the new rule to other words that appear at the end of a sentence. You can – and should – ignore such false flags if they don't make sense in your sentences.

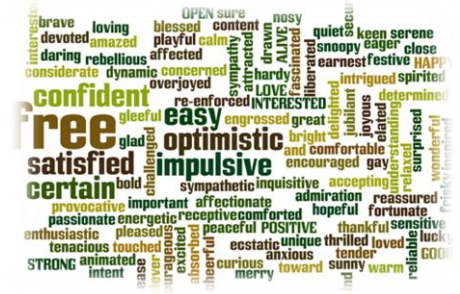
■ Commas before the conjunction in compound sentences

In general, such flags would be correct. However, there are exceptions. If the two independent clauses are short and closely related, it is perfectly permissible to omit the comma before the conjunction. CMoS gives plenty of examples in which the absence of the comma is actually preferred. Feel free to ignore recommendations to add the comma anyway. One such report I saw went on to describe the exceptions but then suggested you might want to add the comma anyway because you can never be wrong doing that. I completely disagree with that assessment because there are many situations in which the addition of the comma between short clauses spoils the pacing and rhythm that the author is trying to create.



The Pitfalls: Word Usage

- Software is very good at statistical analyses
- This is an area where a software tool can help you identify repetition, pet words, overused words, padding words, and superfluous words (like “that”)



But you should still apply an appropriate dose of caution

- **Is the repeated word part of a character’s voice?**
If so, leave them alone – just be sure it isn’t so overdone that readers will get bored.
- **What’s the frequency of the repetition?**
Does the word appear 140 times in a 90,000-word manuscript? That’s less than 2/10 of 1% of the words in the manuscript, so it’s not particularly excessive (unless all the instances occur within a 2000-word section of text, for example). On the other hand, does the word appear 5,000 times in a 75,000-word manuscript? That’s egregious and needs attention.
- **Is the repetition appropriate for the story?**
One report I saw complained that the word “father” occurred 167 times in the manuscript and some of the instances should be eliminated or changed. Well, in this particular novel, the protagonist’s father was a major character – and the narration was first person. So the only way the protagonist could address his father or discuss him in third person with other characters was to use the word “father.” In this context, the repetition is actually quite essential to the story, so the software’s complaint should be ignored.
- **Just how nit-picky is the software being?**
Does it seem like there’s some sort of internal requirement to find at least some specific number of words used repetitively? You can tell this if it starts complaining that a word occurs less than 50 times in a typical-length novel. Always do the math (divide the number of repetitions by the total number of words in your manuscript). What percentage is the repetition of the total word-count in your manuscript? If it’s fractions of a percent (0.00x), you can safely ignore it. If it starts creeping above 5%, then it’s probably worth investigating.



Pros and Cons of Software Tools

So how should I think about the editorial suggestions from a software tool?

Where software tools may be effective

- Statistical analysis of word usage
- Typos and spelling errors – up to a point
- Most punctuation
- Pointing you in the right direction for some things to consider

Where software tools may not be reliable

- When your manuscript follows a specific style guide, such as CMOS or your publisher's style guide
- When context matters – particularly in matters of grammar and forms of usage (words or phrases)
- Spelling – particularly if your manuscript includes anything outside of formal, standard English
- Punctuation that is essential to the rhythm and pacing of the narrative

Where software tools may not be your friend

- When their reporting is so overwhelming you don't even know how to think about it or where to start
- When most complaints are inappropriate if context is properly understood
- When you've already invested in a collaboration with a professional editor

If you and your editor have already worked to polish your novel, a software editing tool may, on occasion, provide an insight or two that might otherwise turn up in a final proofread. However, **the tool should not be given priority over the combined judgment of the author and the human editor.**