Copy Editors Can Play a Role in the Detection and Elimination of "Tortured Phrases"

Jaime A Teixeira da Silva

Copy editors possess a very specific skill set, including linguistic fluidity. In scholarly communication, the work by copy editors translates into text in papers that is scientifically accurate,[1] which can increase the chance of acceptance following peer review.[2] Five professional features of copy editors that contribute to the accuracy of scientific communication are "clarity, coherency, consistency, conciseness, comprehensibility, and correctness."[3(p4)] These elements improve a paper's style and readability after peer review and prior to proof development and publication.[4,5] In academic publishing, copy editors are also required to verify the accuracy of references.[6]

When the accuracy of jargon to describe scientific findings is absent, the text's scientific message becomes ambiguous or misleading.[7] Copy editors, either in-house or externally contracted, contribute to the accuracy of a journal's content, usually at the last step of each manuscript's processing. In an attempt to trim costs and maximize profits, however, some journals or publishers might cut copy editors from the quality control workflow.[3] Yet, this decision comes at a risk, namely that erroneous or ambiguous text introduced by authors into their papers may be undetected by peer reviewers and editors, and thus end up in published papers. Although—to the author's knowledge—there are no economic analyses that have assessed the costs associated with "cleaning up" erroneous literature at the postpublication stage, relative to the employment costs that were saved by removing the copy editor, the current volume of retractions, especially those related to poor quality control,[8] suggests that scientific and linguistic errors could have been avoided, to some extent, had a copy editor existed (i.e., had journals not cut proverbial financial "corners").

Jaime A Teixeira da Silva (https://orcid.org/0000-0003-3299-2772; jaimetex@yahoo.com) is an independent researcher, Kagawa-ken, Japan.

https://doi.org/10.36591/SE-4703-22

The role of copyediting in scientific translations,[9] although an important extension of this discussion, is not explicitly considered in this paper.

"Tortured Phrases" Distort the Language and Scientific Prose of a Paper

The inaccurate description of science and scientific terms, whether these be the background information in a paper's introduction or discussion, or more technical aspects in the methods or results, can dilute the impact of peer-reviewed scientific literature, and thus its accuracy.[10] "Tortured phrases" is a relatively new term to describe a linguistic phenomenon in which established scientific terms and jargon have been replaced by unconventional or inaccurate ones, usually as a result of the use of artificial intelligence (AI), such as machine-generated translations.[11] This implies that there are both a human element, as well as an AI component, that leads to the creation of a tortured phrase. That process, as well as the level of AI-human dependency, might depend on several factors, such as the level of English proficiency, the authors' level of scientific experience in both research and publishing, or the reliance on AI, such as an online thesaurus, to generate text.

The existence of tortured phrases is not limited to peer-reviewed literature, for which copyediting is generally expected, but may also be found in preprints, where copyediting rarely, if ever, exists.[12] The presence of tortured phrases has also been associated with cases of plagiarism and other ethical infractions[11] and is one strategy to avoid the detection of textual similarities and/or plagiarism, by using these odd or unconventional terms.[13] Tortured phrases can thus serve as primers to detect potentially problematic papers.[14]

A hypothetical example follows. A novice researcher in the environmental sciences is not aware that the correct scientific term is "heavy metals." They use an online translator (a form of AI) to translate text from their native language, with the output "substantial metals." Thinking that this is the veritable scientific translation, this novice researcher incorporates this tortured phrase into their paper, which is not noticed by the coauthors, peer reviewers, and editors.

CONTINUED

However, a copy editor with experience in the environmental sciences would surely be able to detect this linguistic red flag and correct it prior to the paper's publication.

In one documented set of examples in stress-related biochemical and physiological studies, and as a subset of the environmental sciences, the term "acid" was replaced, via the use of an online translator or thesaurus, by the term "corrosive," such that established jargon (in this case hormones) like "abscisic acid," "jasmonic acid," and "salicylic acid," became incorrectly represented by "abscisic corrosive," "jasmonic corrosive," and "salicylic corrosive."[15] Similarly, Parkinson's disease might be erroneously represented by the tortured phrases "Parkinson's malady," "Parkinson's ailment," "Parkinson's infection," and "Parkinson's sickness."[16]

Another possibility is where a run of text through plagiarism detection software reveals a high level of textual similarity. An author, in their desire to reduce that level of textual similarity, opts for a radical procedure, including the conversion of established terms into alternative terms, which can be achieved with an online thesaurus. This, too, results in a textual output that deviates from established jargon. In this case, an attempt to avoid plagiarism can lead to the production of tortured phrases.[11]

Can AI Replace Human Copyediting Endeavor?

Ultimately, the distortion of accurate science communication, via the introduction of tortured phrases, leads to the emergence of science miscommunication, or inaccurate science communication. In light of the compromised quality of scientific writing that may arise when copyediting is insufficient, when inexperienced copy editors are employed,[17] or when quality is sacrificed at the expense of pressures to publish greater volumes,[18] a desire and need by academia and the publishing industry may arise to be able to detect and transform tortured phrases into standardized terms and jargon by using AI.

Even though there is a level of irony in suggesting that AI is able to correct errors introduced by AI (e.g., the tortured phrases), especially if the introduction of those errors into a scientific paper was overseen by human authors, provisional evidence shows that one form of AI, ChatGPT, has the ability to reverse tortured phrases,[19] thereby allowing accurate scientific jargon to be reintroduced into papers. The use of this large language model (LLM) as a potential solution to the introduction of tortured phrases into scientific papers, namely to remove them and replace them with proper scientific jargon, would serve not only authors who may not have high English proficiency, or may have limited scientific experience in research and publishing, but also journal editors who have to, for whatever reason, remove a quality control step that involves copy editors. In other words, this opens up the opportunity for AI to serve as a copy editor, a role that has traditionally only been reserved for humans in scientific publishing, or to support experienced copy editors, without replacing the human element. In such cases, journals and publishers have the ethical responsibility of ensuring that their use of AI is properly acknowledged in a published paper, just as authors are held to this requirement.[20]

Conclusion

This commentary advocates for the need for copy editors in any journal due to their linguistic and technical skills. Where peer reviewers or editors might fail in quality control, especially of finer-scale details in the text, such as distorted technical language or jargon, in the form of tortured phrases, experienced and well-versed copy editors would be able to detect such irregular terminology, and either eliminate it prior to the paper being published, or alert editors of ethical infractions associated with their use, such as those related to plagiarism. There is a need to ensure that specialized copy editors form part of the quality control chain, so that the integrity of standard scientific terms and jargon is guaranteed prior to the publication of papers. Provisional evidence shows that tortured phrases can be detected and corrected by AI (e.g., ChatGPT), suggesting that this LLM could serve in a copyediting role in the future, although this requires extensive testing of different LLMs. Given that the employment of human copy editors would be threatened, the financial and technical feasibility of replacing them with AI/LLM-based copyediting, and not merely lending support, needs to be assessed. That debate has already begun.[21]

References and Links

1. Wates E, Campbell R. Author's version vs. publisher's version: an analysis of the copy-editing function. Learned Publ. 2007;20:121–129. https://doi.org/10.1087/174148507X185090.
2. Roth R. Understanding the importance of copyediting in peer-reviewed manuscripts. Sci Ed. 2019;42:51–54.
3. Berenbaum MR. On a subject no one wants to read about (about which no one wants to read?). Proc Nat Acad Sci USA. 2020;117:4–6. https://doi.org/10.1073/pnas.1920932117.
4. Goodman D, Dowson S, Yaremchuk J. Open access and accuracy: author-archived manuscripts vs. published articles. Learned Publ. 2007;20:203–215. https://doi.org/10.1087/095315107X204012.
5. Pillière L. Imposing a norm: The invisible marks of copy-editors. In: Pillière L, Andrieu W, Kerfelec V, Lewis D, editors. Standardising English: Norms and Margins in the History of the English Language. Cambridge, UK: Cambridge University Press; 2018. p. 251–276. https://doi.org/10.1017/9781108120470.
6. Ngatuvai M, Autrey C, McKenny M, Elkbuli A. Significance and implications of accurate and proper citations in clinical research studies. Ann Med Surg. 2021;72:102841. https://doi.org/10.1016/j.amsu.2021.102841.
7. Kueffer C, Larson BMH. Responsible use of language in scientific writing and science communication. BioScience. 2014;64:719–724. https://doi.org/10.1093/biosci/biu084.
8. Oransky I. Retractions are increasing, but not enough. Nature. 2022;608:9. https://doi.org/10.1038/d41586-022-02071-6.

CONTINUED

9. Solum K. The tacit influence of the copy-editor in literary translation. Perspectives. 2018;26:543–559. https://doi.org/10.1080/0907676X.2018.1458888.
10. Teixeira da Silva JA. Tortured phrases dilute the specificity of medical jargon. J Health Soc Sci. 2022;7:137–140. https://doi.org/10.19204/2022/TRTR2.
11. Cabanac G, Labbé C, Magazinov A. Tortured phrases: a dubious writing style emerging in science. Evidence of critical issues affecting established journals. arXiv. 2021. https://doi.org/10.48550/arXiv.2107.06751.
12. Teixeira da Silva JA. "Tortured phrases" in preprints. Curr Med Res Opinion 2023;39:785–787. https://doi.org/10.1080/03007995.2023.2201098.
13. Teixeira da Silva JA. "CRISPR" mutations: inaccurate linguistic variations and misrepresentation of the CRISPR acronym. CRISPR J. 2024;7:3–4. https://doi.org/10.1089/crispr.2024.0005.
14. Teixeira da Silva JA. "Tortured phrases" in COVID-19 literature: can they serve as epistemic markers to assess the integrity of medical information? Phil Med. 2023;4:1–12. https://doi.org/10.5195/pom.2023.164.
15. Teixeira da Silva JA. "Corrosive" acids, inaccurate forms of salicylic, jasmonic, gibberellic and abscisic acids, in the plant literature. [in press]. Vegetos. https://doi.org/10.1007/s42535-023-00772-5.
16. Teixeira da Silva JA. The nomencluratural misrepresentation of Parkinson's disease. Neurol Sci. 2023;44:2179–2180. https://doi.org/10.1007/s10072-023-06672-5.
17. McKinley J, Rose H. Standards of English in academic writing: the authors respond. J Second Language Writing. 2019;44:114–116. https://doi.org/10.1016/j.jslw.2019.04.004.
18. Mosser GF, Wohl H, Fine GA. Alone in publand: leaving academics to themselves. Amer Sociol. 2016;47:238–252. https://doi.org/10.1007/s12108-015-9281-x.
19. Teixeira da Silva JA, Tsigaris P. ChatGPT's ability to reverse "tortured phrases" into standardized English and scientific jargon: relevance to nurse educators and researchers. Nurse Edu. 2024;49:E161. https://doi.org/10.1097/NNE.0000000000001290.
20. Teixeira da Silva JA, Daly T, Türp JC, Sabel BA, Kendall G. Undeclared use of third party service providers in academic publishing is unethical: an epistemic reflection and scoping review. [in press]. Naunyn-Schmiedeberg's Arch Pharmacol. 2024. https://doi.org/10.1007/s00210-024-03177-6.
21. Sebastian F, Barron R. Artificial intelligence: what the future holds for multilingual authors and editing professionals. Sci Ed. 2024;47:38–42. https://doi.org/10.36591/SE-D-4702-01.