EDITORIAL

On a subject no one wants to read about (about which no one wants to read?)

May R. Berenbaum, *Editor-in-Chief, PNAS*

Long before I developed an interest in entomology, I had an interest in etymology—the study of word origins and evolution. An interest in word coinage is certainly compatible with a career in science. As a matter of course, scientists develop new concepts, discover new materials, invent new instruments, or describe new species, and have an ongoing need for new words to communicate with others about them. As scientific subjects have increasingly permeated popular culture, scientific words have crept into common parlance. According to the *Oxford English Dictionary*, the popular science magazine *Scientific American* has been the source for the first English usage of almost 6,000 words, including many that were at first strictly technical scientific terms. On this list are such familiar words as "bleach," "clone," "computing," "radio," and even "editor" (1).

Although research advances can lead to word coinage, they can also lead to its devaluation. Psychologist (and eugenics advocate) Henry Goddard (2), for example, felt the need to identify "feebleminded" individuals who, when evaluated with an intelligence test developed by Alfred Binet and Theodore Simon to estimate mental age, tested in the range of 8–12 years of age and therefore could not be classified as imbeciles (with a mental age of 3–7) or idiots (with a mental age of 2 or younger). As Goddard (2) explained, "One of the most helpful things that we can do, would be to distinctly mark out the limits of this class and help the general public to understand that they are a special group and require special treatment,[*sic*]—in institutions when possible." The word he proposed was "the noun from the Greek word meaning foolish, 'moronia,' and these children might be called 'morons.'" This word was originally coined strictly as a medical diagnosis; needless to say, the word no longer has a scientific definition or in fact utility in any scientific context at all. Ironically, many components of the Binet test that made it subject to bias were based on word use or language facility.

May R. Berenbaum.

So, scientific terms evolve in concert with evidence-based knowledge. There are, though, other components of scientific communication—grammar, including syntax and diction, among them—that evolve at a much slower pace. Conforming to established standards of written language has long been a requirement of scientific communication and, at least in scientific journals, editors have had a responsibility to oversee compliance. The scientific journal as it exists today originated in the 19th century (3), as did the set phrase "editor-in-chief." At that time, copyediting became a publication bulwark—the process by which confusing inconsistencies, time-wasting repetition, misleading inaccuracies, significant omissions, and any other departures from standard usage that may distract a reader are, if not eliminated, then at least minimized. Today, dozens of websites refer to various versions of the "4 Cs" (or occasionally the "5 Cs") of copyediting, which include some combinations or permutations of clarity, coherency, consistency, conciseness, comprehensibility, and correctness, but all converge on the "Cardinal C"—communication. Copyediting done well in theory benefits the author, the publisher, and the reader.

"Communication" remains in the vocabulary of scientific publishing—for example, as a category of manuscript ("Rapid Communications") and as an element of a journal name (*Nature Communications*)—not as a vestigial remnant but as a vital part of the enterprise. The goal of communicating effectively is also why grammar, with its arcane, baffling, or even irritating "rules," continues to matter. With the rise of digital publishing, attendant demands for economy and immediacy have diminished the role of copyeditor. The demands are particularly acute in journalism. As *The New York Times* editorial board member Lawrence Downs (4) lamented, "...in that world of the perpetual present tense—post it now, fix it later, update constantly—old-time, persnickety editing may be a luxury.... It will be an artisanal product,

www.pnas.org/cgi/doi/10.1073/pnas.1920932117

like monastery honey and wooden yachts." Scientific publishing is catching up to journalism in this regard.

The seemingly exponential increase in the number of scientific journals of late has intensified competition and incentivized efforts to economize on production costs and speed up publication times for manuscripts. Eliminating one step in the production process long considered integral—in-house copyediting (or, in some cases, any copyediting at all)—has been embraced by some journals as a way to gain a competitive edge. There's sad irony in the fact that, as the scientific community increases demands for transparency, accuracy, and reproducibility, the words that help to achieve those commendable goals are now receiving so much less scrutiny. Errors in syntax, word choice, spelling, and even punctuation, for example, can make descriptions of methods sufficiently mystifying that reproducing a study with a high degree of fidelity can become an impossible task. A few subject–verb disagreements in a manuscript with stellar scientific content should not lead automatically to rejection, but text that is so compromised by awkward writing benefits neither authors nor readers and can blunt the impact of the scientific content.

Years ago, PNAS extensively copyedited manuscripts to improve readability, but the process was costly and time-consuming, and, as PNAS staff members tell me, seemed to make authors more upset than appreciative. To avoid changing the intended meaning of the authors or giving them extra work to do to verify or object to extensive copyedits, PNAS now restricts its copyediting to correcting grammatical and punctuation errors, marking awkward or confusing phrases, and maintaining consistent journal style. Costs are managed with the use of standard style manuals rather than creating a highly customized style for the journal. We state in our Instructions for Authors: Language-Editing Services: "Prior to submission, authors who believe their manuscripts would benefit from professional editing are encouraged to use a language-editing service (see list here). PNAS does not take responsibility for or endorse these services, and their use has no bearing on acceptance of a manuscript for publication."

Irrespective of the potential future of scientific journals, which are, as more than a few editorial writers have pointed out lately, "17th century institutions" (e.g., ref. 5), it seems likely that the written word will remain the primary means of disseminating scientific knowledge for a long time. Accordingly, it is in the best interest of the scientific community to work toward improving the quality not only of published data but also of the language used to introduce, report, and interpret those data.

If anything, unambiguous, logical, concise, and, yes, grammatically correct writing has become more, not less, important given recent changes in the scientific enterprise. For one thing, citation analysis and digital search depend on accuracy. Among the most important tasks of copyeditors today is to ensure that manuscripts are both human-readable and machine-readable so that they are properly linked to external databases and searchable. Simple spelling, typographical, or punctuation errors can condemn a paper to bibliographic obscurity. Wates and Campbell (6) carried out an analysis of the copyediting function by comparing the text in the author's version and the publisher's version of 189 manuscripts selected at random from 23 journals published by Blackwell. The greatest proportion of changes made as a result of the copyediting and proof-correction process, 42.7%, involved the accuracy of references. In addition to affecting the potential impact of publications as reflected by citations, a significant concern among authors, such errors may bias literature reviews and metaanalyses via omission of critical data. As an example of how long citation errors can persist, due to some confusion around its official publication date in 1964, the classic paper on plant–insect coevolution by Ehrlich and Raven (7) was often mistakenly cited as "Ehrlich and Raven (1965)" soon after publication; this incorrect citation has continued to appear in bibliographies for at least five decades (e.g., Grunzweig et al., ref. 8).

Lingua Franca

In addition, the scientific research enterprise has become global. Scientists who are native English speakers frequently struggle with every aspect of writing scientific papers in their own language, despite having been exposed to the rudiments of English as a written language throughout their formative years. As English has become a lingua franca for scientific communication, the proportion of papers written in English by nonnative speakers without such lifelong familiarity with the language is increasing; the frequency of errors arising from less than complete fluency is likely to increase as well. Between 2005 and 2010, for example, submissions to journals that used the ScholarOne manuscript submission system from the United States increased 177% while the submissions from China increased 484%, with the United States' share dropping 3.3% and China's share increasing 5.5% (9). This trend creates challenges for both native English readers and nonnative readers; understanding written English can depend on how closely the text conforms to accepted usage for both groups. The use of a lingua franca for science can create biases and inefficiencies that under some circumstances might become life-threatening (e.g., ref. 10), but it has become a necessity for global information exchange. That said, varying levels of familiarity with the lingua franca can create inequities and underrepresentation in areas of scholarship where data are urgently needed. Understanding global phenomena requires evaluation of evidence from all affected regions, and incorrect English usage can lead to omissions that can obscure or distort patterns. That papers from critical regions published in languages other than English are frequently missed entirely is understandable, but relevant papers from regions of the world where English is not the primary language that otherwise meet quality standards (such as peer review) may be discounted or omitted entirely due to problems with comprehensibility or concerns about quality due to "bad grammar," as is the case with news articles (11).

The Rise of Team Science

Another recent trend creates challenges—the rise of team science. A National Research Council report on team science (12) noted that the proportion of all science and engineering papers written by more than one author reached 90% by 2013. Although 95% were written by 10 or fewer authors, there are papers in the literature listing hundreds or even thousands of authors. While increasing the number of scientists working together offers many advantages—including access to broader expertise and more sophisticated instrumentation—it also presents challenges associated with communication and coordination. There is little precedent in the history of English-language literature for cooperatively producing manuscripts with so many authors, particularly when the authors come with different discipline-specific research philosophies and specialized vocabularies. Obtaining external reviewers for these manuscripts can be difficult given that the most qualified reviewers may have technical conflicts of interest by virtue of institutional associations or collaborations with authors. Clearly, there are ways to harmonize and integrate text provided by large numbers of authors within a single manuscript—the high-energy physics community offers a sterling example of "communitarian" culture, including team authorship—but other research communities may need their own unique solutions.

The Future of Persnickety

English is far from the ideal language for writing about science. Among other things, it has rules that are regularly broken (e.g., irregular verbs, which do not conform to standard inflection patterns), meanings that seem to change arbitrarily depending on context (e.g., phrasal verbs, which change meaning depending on the presence or absence of other words, as does "run," depending on whether it's followed by "up," "over," "down," or "across"), idiosyncratic idioms, and an infuriatingly inconsistent spelling system. It also lacks some of the useful features of other languages. That English lacks a unique third-person nongendered singular pronoun is at the heart of current conversations about pronoun use in the context of changing cultural attitudes toward gender identity and inclusivity. The inadequacies of English in providing flexibility to accommodate certain aspects of changing culture are layered on top of universal changes in style preferences that have evolved in response to cultural selective pressures. Passive voice, for example, was heavily favored for scientific writing for much of the 20th century, but today there is a clear preference for active voice for reasons of "clarity and conciseness" (13). Passive voice, though, still has a few passionate advocates.

Compared to the availability of tools to assist authors with other elements of manuscripts, there's vanishingly little available to assist authors with developing or refining their technical writing skills. Beyond the phenomenal digital tools available to scholars for creating those other elements—figures and statistical analyses, for example—there is no shortage of workshops or even formal coursework to help refine those skills. Such is not the case for students attempting to learn discipline-specific writing skills, despite the fact that science writing is considered a key competency, at least in biology education (14).

At present, digital tools for composing and correcting text are remarkable, but they're not yet up to the task of editing complex scientific text. Perhaps the best way forward is for academic institutions to work together, in concert with scientific societies and journal publishers, to develop mechanisms for providing that guidance in the form of webinars or workshops, if introducing formal courses in writing into already crowded curricula isn't feasible. Monks today not only continue to make honey, they've added fudge, cheese, and lavender oil to their repertoire; surely the scientific community can devise new and effective ways to keep scientific communication concise, coherent, and clear for the foreseeable future.

Acknowledgments

I thank Adam Dolezal for insightful comments on the text (particularly those relating to monastery honey) and Richard Leskosky for sharing with me his linguistics insights and for recalling the time that he searched the University of Illinois Library digital catalogue to find the diary of 19th century experimental physicist Michael Faraday and instead discovered *Faraday's Dairy*.

1 R. D. Fields, How *Scientific American* helps shape the English language (2018). https://blogs.scientificamerican.com/observations/how-scientific-american-helps-shape-the-english-language/. Accessed 10 December 2019.
2 H. H. Goddard, Four hundred feeble-minded children classified by the Binet method. *J. General Psychol.* **17**, 386–397 (1910).
3 J. R. Topham, The scientific, the literary and the popular: Commerce and the reimagining of the scientific journal in Britain, 1813–1825. *Notes Rec. R. Soc. Lond.* **70**, 305–324 (2016).
4 L. Downes, In a changing world of news, an elegy for copy editors. *NY Times*, 16 June 2008. https://www.nytimes.com/2008/06/16/opinion/16mon4.html. Accessed 10 December 2019.
5 P. Walter, D. Mullins, From symbiont to parasite: The evolution of for-profit science publishing. *Mol. Biol. Cell* **30**, 2537–2542 (2019).
6 E. Wates, R. Campbell, Author's version vs. publisher's version: An analysis of the copy-editing function. *Learn. Publ.* **20**, 121–129 (2007).
7 P. R. Ehrlich, P. H. Raven, Butterflies and plants: A study in coevolution. *Evolution* **18**, 586–608 (1964).
8 L. Grunzweig, D. J. Spiering, A. Labatore, R. J. Warren II, Non-native plant invader renders suitable habitat unsuitable. *Arthropod-Plant Interactions* **9**, 577–583 (2015).
9 K. Hyland, Academic publishing and the myth of linguistic injustice. *J. Second Lang. Writ.* **31**, 58–69 (2016).
10 D. Cyranoski, Bird flu data languish in Chinese journals. *Nature* **430**, 955 (2004).
11 A. Appelman, A. M. Schmierbach, Make nomistake? Exploring cognitive and perceptual effects of grammatical errors in news articles. *Journal. Mass Commun. Q.* **95**, 930–947 (2018).
12 National Research Council, *Enhancing the Effectiveness of Team Science* (The National Academies Press, Washington, DC, 2015).
13 P. A. Yeong, The passive voice in scientific writing. The current norm in science journals. *J. Sci. Communication* **13**, 1–16 (2014).
14 K. Grzyb, W. Snyder, K. G. Field, Learning to write like a scientist: A writing-intensive course for microbiology/health science students. *J. Microbiol. Biol. Educ.* **19**, 19.1.10 (2018).