Published in final edited form as:

Acad Med. 2012 January ; 87(1): . doi:10.1097/ACM.0b013e31823aad7.

Publication Ethics and the Emerging Scientific Workforce: Understanding 'Plagiarism' in a Global Context

Carrie Cameron, PhD,

associate director of the Cancer Prevention Research Training Program, Division of Cancer Prevention and Population Sciences and instructor, Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

Hui Zhao, PhD, and

senior data analyst, Department of Breast Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

Michelle K. McHugh, PhD

postdoctoral fellow, Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

Abstract

Scientific publication has long been dominated by the English language and is rapidly moving towards near complete hegemony of English, while the majority of the world's publishing scientists are not native English speakers. This imbalance has important implications for training in and enforcement of publication ethics, particularly with respect to plagiarism. A lack of understanding of what constitutes plagiarism and the use of a linguistic support strategy known as patchwriting can lead to inadvertent misuse of source material by non-native speakers writing in English as well as to unfounded accusations of intentional scientific misconduct on the part of these authors. A rational and well-informed dialogue about this issue is needed among both native English speaking and non-native English speaking writers, editors, educators, and administrators. Recommendations for educating and training are provided.

Scientific publication has been dominated by the English language for decades and is rapidly moving towards near complete hegemony of English, while the majority of the world's publishing scientists are not native English speakers¹. National Science Foundation (NSF) statistics for global science and engineering article output for 2007 show that the U.S. share is just 28%, decreasing from 34% since 1995. Of this 28%, it is safe to assume that a good number are written by scientists whose first language is not English (L2s): by 2003, 40% of doctoral level workers in U.S. science and engineering jobs were foreign-born, and since 1994, over 50% of postdocs in biosciences have been visa holders.² Using nationality as an imperfect proxy for native language, it is clear that scientists who speak languages other than English are not a minority--they *are* the scientific workforce. This imbalance has important implications for training in and enforcement of publication ethics, particularly with respect to plagiarism.

Correspondence should be addressed to Dr. Cameron, Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, P.O. Box 301402, Mail Code 1365 Houston TX 77230-3721; phone: (713) 794-1476; fax: (713) 563-9203; (ccameron@mdanderson.org).

Other disclosures: None

Ethical approval: Not applicable

Defining Plagiarism: The Act and the Concept

Plagiarism is usually defined as using another author's material without proper attribution. A tour of college and editorial webpages on plagiarism quickly reveals a few key themes: 1) plagiarism is so egregious an infraction that words such as 'crime' and 'stealing' appear regularly; 2) absence of intent to plagiarize is not a mitigating factor; and 3) even trivial errors in citation form or muddled wording can be construed as plagiarism. Precise characterizations of plagiarism are difficult to find, however, and several studies have shown that identification of and value judgments about source text repetition vary markedly across individuals, let alone across languages and cultures^{3–10}. Given the gravity of the issue and its enmeshed relationship with culture and language, it is hard to imagine how the plagiarism question could be other than problematic for L2 writers. Whether plagiarism or perceived plagiarism is more prevalent among L2 writers than native-language (L1) writers is not known. But L2 writers' sometimes inappropriate use of source material (by U. S. standards) merits a deeper examination of how we teach and talk about publication ethics.

The Impetus for Patchwriting

The additional burden on those who work in a second language is significant. In 2006 an anonymous informal survey was conducted at The University of Texas MD Anderson Cancer Center of 41 L2 postdoctoral fellows taking a course in English for Academic Purposes. The survey, administered by the instructor, indicated that over 75% (31) believed it takes them at least twice as long to write a research communication in English as in their native language; some responded that it takes them 3 or 4 times as long or longer. 81% (34) strongly agreed or agreed that they had "missed professional opportunities" due to their L2 status. Under this kind of pressure, it is not unreasonable that researchers will develop compensatory strategies.

The term 'patchwriting', originally coined by Rebecca Howard¹¹ to describe writing patterns of primarily L1 undergraduates, denotes the process of weaving together both original and borrowed text. Howard characterizes patchwriting as a scaffolding tool that students use to progress from 'school' writing to academic expository prose. Patchwriting is considered by many to be a developmental strategy^{12–16}, distinct from plagiarism in that there is no deliberate intent to misrepresent the source of the ideas; it is used by L2 writers in academia with similar motivations as those of novice L1 writers. Its widespread use among L2s is suggested by a 2006 study of a corpus of 17 predoctoral L2 texts from various disciplines¹⁷ which found that 41% of the writing in the corpus consisted of unattributed repetition of source language; of the disciplines represented in the corpus, biology writers were more likely than engineering, humanities, or social sciences writers to rely on this strategy. Given the workload of the average trainee and the pressure to publish, it is plausible that an L2 writer would seek a way to both reduce the time needed to complete an article and to elevate the quality of its prose. Use of patchwriting and/or plagiarism has been associated with cultural heritages of L2 speakers^{18,19} that prize emulation of traditional scholarship over individuality. In this view, it would be considered presumptuous for a novice to impose his or her own voice over the cited scholar's. But the cultural values explanation is built on unexamined assumptions. A 2009 study²⁰ on attitudes toward plagiarism among 77 Japanese undergraduates suggests that, in Japan—one of the Confucian-influenced cultures often said to encourage imitation of canonical scholars—students are aware of what plagiarism is, can identify it when they see it, and reject it as poor scholarship (although not necessarily as a form of theft). Cultural values may play a role in academic writing in some instances, but relying on them as an explanation for patchwriting leads us to a dead end: either we sacrifice scholarly integrity in favor of

relativistic broad-mindedness or we sacrifice the legitimate scientific contributions of a significant number of investigators in favor of absolutist ethics. Either way, science loses.

A thoughtful examination of why writers patchwrite can lead to better understanding as well as more productive solutions to the dilemma. L2 writers are faced with competing priorities. Simply informing them that patchwriting is unacceptable does not address the motivations for using it. Patchwriting, for L2s, must first be understood as a strategy solving a linguistic problem that is experienced by novice L1 writers¹³, but that is especially acute for L2s^{21–26}. Skillful synthesis of source material requires discernment of both the denotation (what is meant) and connotation (what is suggested) of words, phrases, idioms, and grammatical constructions. The expression “It has come to my attention,” carries a complex set of associations different than, for example, “Somebody said,” or “They let me know.” L2 writers realize that unknown connotations present risks for them. Even in cases where the differences in connotation may be minor, the L2 writer is often reluctant to take a chance; it is safer to use the original words than to attempt a synonym.

Proper integration of source material requires skill in paraphrasing and summarizing, which require not only accurate interpretation of the source language, but considerable linguistic agility and range in order to recreate the meaning faithfully. These skills are difficult to master in a second language, where the writer may struggle to find one way of expressing an idea, much less alternative ways to do so.

Thus L2 writers must choose between producing work rapidly with less sophistication, or delaying publication in order to meet L1 linguistic standards. Moreover, L2 writers are likely to have been educated in environments where plagiarism and especially patchwriting are neither vigorously condemned nor openly condoned—they simply aren’t discussed as frequently as they are in the U.S., where middle and high school writing classes teach proper attribution of sources, use strict grading criteria on proper citation style, and punish those who don’t conform to standards. L2 writers are thus less likely to have received such in-depth training in proper source attribution.

The Risks and Implications of Patchwriting

The difficulty of producing sophisticated academic prose in a second language together with a lesser degree of socialization regarding U.S.-style attribution standards may result in the L2 writer both overestimating the benefits and underestimating the risks of patchwriting. The following anecdote was provided by an Asian L2 postdoctoral fellow working in Texas in 2009; it recounts a discussion among three or four fellows and a scientific writing instructor about plagiarism after their one-hour class:

My friend who is a non-English speaker was caught in plagiarism in her term paper when she was a master degree student. We were all shocked since what she did was copying sentences or paragraph from papers journals into her term paper. This was quite common for us non-English speakers to do since we usually felt the language in the published paper is more precise and correct than our own words, and on the other hand, it took a long time to rewrite or paraphrase other's work; this is especially common if we are under time pressure. But we usually cite those sentences or change a few words. My friend probably didn't cite or she copied too much... But she got disciplined by the course instructor, she almost got kicked out of the school. It was so scary.

In scientific writing talks at our institution, patchwriting is defined and the risks of engaging in it are explained. Many of the international audience, whether from Germany, China, or Colombia, state that they know essentially what intentional plagiarism is; many, however,

state that they did *not* realize that patchwriting as a textual strategy is very likely to be equated with plagiarism, nor are they aware of the intense condemnation that perceived plagiarism elicits. In an interview in 2010 with one of the authors (CC), one South American L2 postdoctoral fellow explains:

They tell us about plagiarism a lot, you hear about it all over the place when you come here. But they assume that you already know since undergrad how to write... So you take those [English] sentences and you use the thesaurus to change it a little bit and then you cite the guy. We don't really connect that with what they told us about plagiarism since we're not stealing the scientific ideas, we're just imitating some sentences. You copy the model. The language is a model as well. As a foreigner, they don't really teach you in graduate school how not to do that.

The stigma surrounding plagiarism in the U.S. is strong enough that it can be difficult to have an open and constructive discussion of L2 patchwriting. Yet a rational and well-informed dialogue about this issue affecting such a large portion of the research community is critical. Use of plagiarism detection software such as *Turn it In*, *SafeAssign*, and *Déjà Vu* is bringing patchwriting cases to the fore with greater frequency than ever. An article in *Science News Focus*²⁷ reported on an instance of L2 patchwriting in which 95% of the text of a cancer research article was duplicated, the authors changing little more than the disease site and the figures. The *Science* item notes that the authors responded to editorial inquiries by stating that the quality of their English was insufficient and that they acknowledged their lapse in judgment in publishing the paper without using their own language. The authors eventually retracted the paper. In another instance, an English-speaking teaching assistant for a public health course in environmental science wrote:

As a teaching assistant in a required course for all graduate students in the School of Public Health, I was astonished at the number of students who used multiple un-referenced sentences in their submitted assignments. These cases did not include the instances where I felt it was lack of knowledge on how to properly cite a reference.... When I suspected egregious plagiarism, I merely copied the sentences into a Google search and would easily find the primary source. Even though students are given a lecture on plagiarism and our course required students to take a quiz on plagiarism, every semester would bring new cases. Because of this, my school instituted a program called SAFE ASSIGN, which allows students to check their work for potentially plagiarized text before submitting the assignment. The school even provides a course on how to use it. That said, the professor who still teaches the course, even two years after SAFE ASSIGN was introduced, still gets new cases of egregious plagiarism every semester. I find it astonishing that there are still individuals who do not understand what constitutes plagiarism. I often hear it is "cultural" and "acceptable in other countries"; however, I personally feel that any student, regardless of country of origin, who chooses to get an academic degree or attain an academic position, must follow the rules of their institution and be held to the same academic and ethical standards.

Towards a Win-Win Solution

Increased understanding of patchwriting is needed among both L1 and L2 writers, editors, educators, and administrators. While excessive duplication of source material is unacceptable, categorical condemnation of patchwriting as willful scientific misconduct both misses the point and fosters an adversarial climate. Rather than further marginalizing L2 writers, we should encourage appropriate education and support at various points of their careers. In order to facilitate such a relationship the following should be taken into consideration by all parties::

- Thoughtfully address patchwriting and its significance in English-language publication in scientific writing and advanced ESL courses, as well as provide training in linguistic skills such as paraphrase and summary. Support can also be provided by scientific editors and English instructors in cases where mentors or co-authors may be unsure of how to help.
- Enhance Responsible Conduct of Research (RCR) units in publication and authorship ethics. “RCR Seminars: a Postdoc Toolbox,” a lecture series on research ethics conducted at M. D. Anderson, includes a session on publication ethics that addresses patchwriting explicitly. Other opportunities for education and training include journal clubs, lab meetings, and career development seminars.
- Encourage mentors to provide direct, specific guidance to trainees on avoiding patchwriting while modeling skillful attribution. Except in cases where the writer inserts a passage of dramatically more sophisticated language than the rest of the manuscript, patchwriting can be difficult to detect, especially if mentors are unaware of the likelihood of it, while trainees who are unaware of the implications of patchwriting may not think it an issue worth bringing up. Both mentors and mentees need to be aware that patchwriting is an understandable yet unacceptable practice. An open and trusting dialogue is critical.

Awareness, education, and communication about this prevalent and poorly understood issue will serve science and scientists alike as international collaboration and integration continue to grow.

Acknowledgments

The authors wish to thank Professor Shine Chang, MD Anderson Cancer Center, for her support of the Responsible Conduct of Research initiative and valuable insights on mentoring of trainees, and Kathryn Carnes and Sue Moreau (editors in MD Anderson’s Department of Scientific Publications).

Funding/Support:

This work is supported in part by the National Institutes of Health through MD Anderson’s Cancer Center Support Grant CA016672; by National Cancer Institute grant R25T CA57730, Shine Chang, Ph.D., Principal Investigator; by National Cancer Institute grant R25E CA56452, Shine Chang, Ph.D., Principal Investigator; and by a National Postdoctoral Association Bring Responsible Conduct of Research Home Seed Grant.

References

1. Swales JM. English as Tyrannosaurus rex. *World Englishes*. 1997; 16(3):373–382.
2. National Science Board. Science and Engineering Indicators 2010. National Science Foundation (NSB 10-01); Arlington, VA: 2010. <http://www.nsf.gov/statistics/seind10/> [Accessed 2010 February 7, 2010]
3. Crocker J, Shaw P. Research student and supervisor evaluation of intertextuality practice. *Hermes*. 2002; 28:39–58.
4. Jude C, Appleton J. Towards consistent penalty decisions for breaches of academic regulations in one UK university. *International Journal for Educational Integrity*. 2005; 1(1)
5. Yeo S, Chien R. Evaluation of a process and proforma for making consistent decisions about the seriousness of plagiarism incidents. *Quality in Higher Education*. 2007; 13(2):187–204.
6. Pecorari, D. Plagiarism and international students: how the English-speaking university responds. In: Belcher, D.; Hirvela, A., editors. *Linking literacies: Perspectives on L2 reading-writing connections*. Ann Arbor MI: University of Michigan Press; 2001. p. 229-245.
7. Pecorari D. Visible and occluded citation features in postgraduate second-language writing. *English for Specific Purposes*. 2006; 25:4–29.

8. Sutherland-Smith W. Pandora's box: Academic perceptions of student plagiarism in writing. *Journal of English for Academic Purposes*. 2005; 4(1):83–95.
9. Pecorari, D.; Sutherland-Smith, W. 4th International Plagiarism Conference. Newcastle upon Tyne. England: Northumbria University; 2010. Policy and practice in two academic settings : how the administrative structures of Australian and Swedish universities serve a culture of honesty; p. 1-16.
10. Flint A, Clegg s, Macdonald R. Exploring staff perceptions of student plagiarism. *Journal of Further and Higher Education*. 2006; 30:145–156.
11. Howard R. A plagiarism pentimento. *J of Teaching Writing*. 1992; 11:233–245.
12. Howard, R. Standing in the shadow of giants: plagiarists, authors, collaborators. perspectives on writing: theory, research, practice. Volume 2. Stamford, CT: Ablex Publishing/JAI Press Inc; 1999.
13. Howard RM. Plagiarisms, Authorships, and the Academic Death Penalty. *College English*. 1995; 57(7):788–806.
14. Jones A, Freeman T. Imitation, copying, and the use of models: Report writing in an introductory physics course. *IEEE Transactions on Professional Communication*. 2003; 46(3):168–184.
15. Whitaker EE. A pedagogy to address plagiarism. *College Composition and Communication*. 1993; 44(4):509–514.
16. Hull G, Rose M. Rethinking remediation: Toward a social-cognitive understanding of problematic reading and writing. *Written Communication*. 1989; 6(2):139–154.
17. Pecorari D. Visible and occluded citation features in postgraduate second-language writing. *English for Specific Purposes*. 2006; 25(1):4–29.
18. Bloch, J. Plagiarism and the ESL student: From printed to electronic texts. In: Belcher, D.; Hirvela, A., editors. *Linking literacies: Perspectives on L2 reading-writing connections*. Ann Arbor, MI: The University of Michigan Press; 2001. p. 209-228.
19. Pennycook A. Borrowing Others' Words: Text, Ownership, Memory, and Plagiarism. *TESOL Quarterly*. 1996; 30(2):201–230.
20. Wheeler G. Plagiarism in the Japanese universities: Truly a cultural matter? *Journal of Second Language Writing*. 2009; 18(1):17–29.
21. Abasi AR, Akbari N. Are we encouraging patchwriting? Reconsidering the role of the pedagogical context in ESL student writers' transgressive intertextuality. *English for Specific Purposes*. 2008; 27(3):267–284.
22. Pecorari D. Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*. 2003; 12(4):317–345.
23. Shi L. Textual appropriation and citing behaviors of university undergraduates. *Applied Linguistics*. 2010; 31(1):1–24.
24. Flowerdew J. Discourse Community, Legitimate Peripheral Participation, and the Nonnative-English-Speaking Scholar. *TESOL Quarterly*. 2000; 34(1):127–150.
25. Keck C. The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*. 2006; 15(4):261–278.
26. Swales, J. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge, UP; 1990.
27. Couzin-Frankel J, Grom J. Plagiarism Sleuths. *Science*. 2009 May 22; 324(5930):1004–1007. 2009. [PubMed: 19460978]