The spread of Brazilian science through the advancement of written proficiency in English: a matter that cannot be postponed

A difusão da ciência brasileira pelo do avanço da proficiência escrita em inglês: uma questão sem adiamento

Ana Luiza Pires de Freitas[1]

A recent bulletin from the Universidade de São Paulo[1] states that 168,546 scientific publications on the topic of COVID-19 have been issued worldwide in 2020. Records from the Dimensions Database Platform refer a mere 4,029 or 2.3% of them as sourced from Brazil. We could wonder what makes the figures of national publications that scanty. Financial cuts for academic research, the possibly excessive bureaucracy concerning clinical trial implementation, and the lack of Brazilian representativeness could also be associated with pressing problems, such as a poor mastery of the English language among Brazilian scientists. Nevertheless, publishing in high-impact journals, and hence forwarding the national research to an elevated pattern, implies a suitable use of the *lingua franca* of science so as to cooperate and liaise with the international scientific community.

Framing scientific investigations through articulate language and circulating them in eligible journals provide both scientific recognition and credibility, two paramount premises for sharing knowledge among members of the international research community. Ultimately, access to valuable scientific updates allows experts to broaden their range of productions and thus expand specialized communication. Science, in turn, takes shape through studies that are organized and disseminated through records of academic writing genres. The scientific community, as an organized group, regulates its activity and discourse, whose norms are followed by its members and acknowledged in the texts used to communicate its production[2], such as academic papers, theses, and dissertations, among others. As a result, we could assume that the ways of saying science also make up science, in a mutually reassuring interrelation. Hence the importance of properly adopting scientific language for those who wish to be accepted as members of a cosmopolitan research community, especially when such access occurs through English as an additional language. Proficient academic writers will, after all, be better skilled to take part in a global communication network and take a stance in the worldwide scientific community. From such framework, making a case for a fine mastery of the passport code for international publications, so as to promote far more inclusion of Brazilian scientific contributions, and sparking a critical discussion on the issue of advancing written English proficiency amid our scientists turns into a matter that cannot be postponed.

As a starting guideline to pursue fluent writing, it should be argued that even though it may appear that text authors count on a seemingly endless range of lexical choices at their text production disposal, scientific language options are actually limited. This limitation is due to the high level of conventionality ruling over academic language, an aspect believed to enhance textual fluidity. In other words, scientific language is oriented by an idiomatic principle[3], implying it works in fixed and/or relatively fixed blocks, and is guided by assumptions greater than the users' free choice.

The notion of academic discourse ranging from a view of language as a probabilistic system has produced an increasing number of studies[4-10] which define fluency as the fluid use of typically frequent word associations. Advances in Applied Linguistics led to the compilation of extensive textual corpora in specialized written genres (such as the English Scientific Text Corpus (SciTex), the British Academic Written English Corpus (BAWE), and the Corpus of Contemporary American English (COCA), with subcategories in health sciences), enabling language researchers to implement empirical studies so as to identify academic language patterns actually based on a description of use. These investigations frequently explore the presence of recurrent word strings, also known as collocations, associations of words from extracting programs, grounded in distributive criteria on the basis of recurrence in a data collection. Accordingly, the frequency of word usage determines the pattern or rather what is accepted as typical of an academic field. Therefore, it is under the guidance of recurrence that typical word strings in medical sciences are available to text producers.

The aid of automated word sequence extracting tools leverages the writing task and potentially the chances of publication by making semantic, pragmatic, sociolinguistic, and even stylistic explorations of aspects of word frequency. From the development of corpus techniques, Jalali et al.[11] tackled the most recurring phraseological sequences in medical research from high-impact journal publications and highlighted their relevance as building blocks to the production of comprehensive and coherent written discourse. Based on their study, the most recurring medical field-specific word associations extracted from 790 research articles in 33 medical subfields were '*in the present study*', '*at the time of*', '*were more likely to*', '*the results of the*', '*of the present study'*, '*as shown in fig*', '*in the control group*', '*in the current study*', '*are shown in table*', and '*this study was to*', from a total of 102 repeated 4-word strings. Such phraseological pattern is the outcome of computer software recuperation in a corpus of 2,420,914 words, following a recurrence frequency of 20 times per million words and an occurrence dispersion of at least 5 different texts in the corpus.

[1]Department of Education and Humanities, Universidade Federal de Ciências da Saúde de Porto Alegre – Porto Alegre (RS), Brazil.

Freitas[12] conducted a contrastive terminological research by compiling a body of investigative texts in health sciences from four major Brazilian public universities — Universidade Federal do Rio Grande do Sul (UFRGS), Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Universidade Federal de Santa Catarina (UFSC), and Universidade Federal de São Paulo (UNIFESP) — against a corpus of international publications from high-impact journals (the American Journal of Clinical Nutrition, Annals of Internal Medicine, and the British Journal of Pharmacology). By adopting collocations as an analytical unit, this study identified both a higher adoption of typically repeated word sequences by the international publishing authors and a richer variability of use by these writers (the most used 4-word string sequences in a corpus of 1,800,00 words under a frequency and distribution criterion of at least 5 recurrences in 5 different texts were '*with the use of*', '*was associated with a'*, '*associated with a lower*', '*had no effect on*', '*the objective was to*', '*of this study was*', '*aim of this study*', '*as well as the*', '*on the basis of*', and '*were associated with a*'.). Moreover, in a pragmatic analysis, while international scientists adopted more high-frequency word associations to embellish longer discussion sections in their papers and thus reinforce the core of their scientific contributions, their Brazilian counterparts tended to use more field typical word sequences to produce longer introductions to their studies. Even though these findings cannot be generalized, they seem to represent a timely opportunity for educational awareness-raising.

As concluding remarks, it is crucial to reinforce that catering for the writing challenges briefly stated along this reasoning is a feasible enterprise; an endeavor to be pursued through exposure to good textual models with the aid of text-editing tools. These resources can help writers find words that go well together in English so that they can access them as they write. Specialized texts, all in all, do not bear universal characteristics and are prone to situational variation[13], in view of their conditions of publication and their subfields of knowledge. The traits that can be recognized as the most sustained ones, however, determine what is most relevant to the user and what, thereby, should be taught as a priority.

Finally, to be fully accepted in an academic community or legitimized by its expert users, research authors' language work ought to be properly organized according to the ways of saying science. To achieve discursive proficiency, therefore, an author should produce texts in which each element or lexical item represents a value to form a whole of meaning. Fluency, in this case, is related both to broad discursive aspects from the account of textual functionality and to lexical-grammatical choices. Namely, the fundamental assertion of this argument is that learning to associate words properly helps writing researchers in health sciences to make appropriate language choices and develop references about the type of language expected from members of an international scientific community. That said, the intent is leaving an open door to further expand this discussion by exploring learning resources in a forthcoming opportunity.

REFERENCES

1. Bernardes J. USP está entre as 20 instituições que mais publicam no mundo. Jornal da USP [Internet]. 2020 [accessed on Nov. 20, 2020]. Available from: https://jornal.usp.br/ciencias/usp-esta-entre-as-20-instituicoes-que-mais-publicam-sobre-covid-no-mundo/
2. Swales J. English as Tyrannosaurus rex. World Englishes [Internet]. 1997 [accessed on Nov. 20, 2020];16(3):373-82. Available from: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/74462/1467971X.00071.pdf;jsessionid=371A2D781D244A7596240E4A6A0FFFA6?sequence=1
3. Sinclair J. Corpus, Concordance, Collocation. Oxford: OUP; 1991.
4. Biber D, Conrad S. Lexical Bundles in Conversations and Academic Prose. In: Hasselgard H, Oksefjell S (eds.). Out of corpora: studies in honour of Stig Johansson. Amsterdam: Rodopi; 1999. p. 181-90.
5. Biber D, Barbieri F. Lexical bundles in university spoken and written registers. English Specific Purp. 2007;26(3):263-86. https://doi.org/10.1016/j.esp.2006.08.003
6. Biber D, Conrad S, Cortes V. If you look at...: Lexical bundles in university teaching and textbooks. Applied Linguistics. 2004;25(3):371-405. https://doi.org/10.1093/applin/25.3.371
7. Biber D, Johansson S, Leech G, Conrad S, Finegan E. The Longman Grammar of Spoken and Written English. London: Longman; 1999.
8. Chen Y-H, Baker P. Lexical Bundles in L1 and L2 Academic Writing. Lang Learn Technol [Internet]. 2010 [accessed on Nov. 20, 2020];14(2):30-49. Available at: https://www.researchgate.net/publication/45681690_Lexical_Bundles_in_L1_and_L2_Academic_Writing
9. Hyland K. As can be seen: Lexical bundles and disciplinary variation. English Specific Purposes. 2008;27(1):4-21. https://doi.org/10.1016/j.esp.2007.06.001
10. Pan F, Reppen R, Biber D. Methodological issues in contrastive lexical bundle research. Int J Corpus Linguist. 2020;25(2):215-29. https://doi.org/10.1075/ijcl.19063.pan
11. Jalali SZ, Moini MR, Arani MA. Structural and Functional Analysis of Lexical Bundles in Medical Research Articles: A Corpus-Based Study. Int J Inf Sci Manag [Internet]. 2015 [accessed on Nov. 25, 2020];13(1):51-69. Available from: https://www.researchgate.net/publication/281242979_Structural_and_Functional_Analysis_of_Lexical_Bundles_in_Medical_Research_Articles_A_Corpus-Based_Study
12. Freitas ALP. Proficiência Escrita em Inglês Especializado: Estudo de Corpus [thesis on the Internet]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2016 [accessed on Nov. 28th, 2020]. Available from: https://lume.ufrgs.br/handle/10183/148955
13. Biber D, Conrad S. Register, genre, and style. Cambridge: Cambridge University Press; 2009.

Address for correspondence:
ANA LUIZA PIRES DE FREITAS
Departamento de Educação e Humanidades, Universidade Federal de Ciências da Saúde de Porto Alegre
Rua Sarmento Leite, 245, 4º andar – Centro Histórico
Porto Alegre (RS), Brazil
CEP: 90050-170
E-mail: analuizaf@ufcspa.edu.br

Received on: 11.30.2020
Approved on: 12.02.2020