MACHINE TRANSLATION TECHNOLOGY: EFFECTS ON SCHOLARLY COMMUNICATION[1]

Marileide Dias ESQUEDA[*]

The goal of the book of Lynne Bowker and Jairo Buitrago Ciro (2019), researchers from University of Ottawa, Canada, is to inform about the ways machine translation is employed in the context of scholarly communication, and to enlighten how this technology can be used more effectively.

According to Bowker and Buitrago Ciro, the book *Machine Translation and Global Research* will be of interest to different audience, such as:

- Literacy brokers, or text mediators, who may apply all the different kinds of direct or indirect intervention, other than named authors, in the production of texts. Literacy brokers include translators, publishers, journal editors, and librarians, who may support authors during their academic writing tasks; some of these literacy brokers can help authors to acquire some basic knowledge on machine translation;
- Researchers who have English as an additional language, and who want to properly use machine translation to provide the effective dissemination of their research results to a wider scientific community;
- Native English-speaking researchers who may disseminate their research results in their own language. These researchers can be aware of how machine translation affects their texts, and how they can facilitate specialized communication among scholars.

For Bowker and Buitrago Ciro, by developing machine translation literacy, i.e., by gaining a deeper understanding of how to work effectively with this type of tool, literacy brokers and researchers will be better able to carry out the scholarly communication process, and to ensure maximum participation from researchers around the world, especially considering the global impacts of English as the language of science.

The book is written in English, and it is composed by an introduction, five chapters, a 12-line concluding section, and 140 references, among which are five references from the authors themselves, attesting their experience in the field of translation technologies, including machine translation. The book draws upon Bowker and Buitrago Ciro's experiences with pragmatic texts (commonly technical and scientific texts) that, used

[*] Universidade Federal de Uberlândia (UFU). Uberlândia – MG - Brazil. marileide.esqueda@ufu.br. ORCID: 0000-0002-6941-7926.

[1] BOWKER, Lynne; BUITRAGO CIRO, Jairo. **Machine translation and global research**: towards improved machine translation literacy in the scholarly community. United Kingdom: Emerald Publishing Limited, 2019. 127 p. ISBN: 978-1-78756-722-1 / doi: 10.1108/978-1-78756-721-420191002.

as examples, were partially submitted to machine translation engines such as Google Translate, Bing Translator (from Microsoft), and DeepL, and then analyzed by the authors.

The introduction poses a general question: can machine translation help people, especially researchers, to better communicate? The answer, for Bowker and Buitrago Ciro, is not clear-cut, but the question is certainly worthy of investigation. According to them:

> If we want the best and the brightest minds on the planet working together to solve problems such as climate change, cancer, and energy crises, then we need to make sure that they can effectively share their research findings with one another. (BOWKER; BUITRAGO CIRO, 2019, p.1).

They also affirm in the Introduction of the book that English has emerged as the international language of scholarly communication – particularly in the domains of science and technology – despite the fact that only roughly 6% of the world's population speak English as a native language. So, they raise the question: "What does this mean for the other 94% who do not speak English?" (BOWKER; BUITRAGO CIRO, 2019, p.1). In their opinion, it means that scholarly communication needs tools and techniques, and researchers need training in machine translation to help them engage with and contribute to the scientific literature in their fields for society's sake.

Chapter 1, titled "Scholarly Communication", provides an overview of the state of the art of scholarly communication and the evolution of English as the international language of research dissemination. It starts defining scholarly communication as the process by which academics, scholars, graduate students and other researchers share and publish their findings so that they are available to the wider research community, and society. The chapter also reveals that recent bibliometric analyses point out the essential growth in the development of science, leading to the increase of scientific productions and the inevitable linguistic challenges associated with such progress. Bowker and Buitrago Ciro highlight the rise of English as the international language of scientific communication, and that the Internet also favors this language to a high degree, once websites dealing with research products like pre-print archives, institutional repositories and online journals are typically presented in English or are at least translated and localized into this language. This chapter also presents a number of possible strategies available to scholars who have English as an additional language and who wish to publish the results of their research in this language.

"Machine Translation" is the title of Chapter 2, which provides readers with a more detailed look at the world of machine translation, beginning with a brief history of the field. Starting with an epigraph from the work of Hutchins and Somers (1992), who affirm that machine translation is one of the most challenging of research activities involving the application of complex theoretical knowledge to the building of systems, the chapter introduces the main approaches applied to machine translation technologies.

Posing that their aim is not to drive the content of the chapter to computational linguists, Bowker and Buitrago Ciro show the advantages and disadvantages of machine translation for users in general. They describe the limitations of various approaches such as rule-based methods (where researchers try to program computers to process language using grammar rules), statistical methods (where computers are trained using parallel *corpora* and making substantial use of probability calculations), *corpus*-based approaches (that use pattern-matching and number-crunching techniques), and the very recent neural networks (information processing system that is inspired by the way biological nervous systems process information, such as the brain. The neural network machine translation system finds patterns, such as contextual clues around the source phrase). Regardless the complexity of these approaches and their refinements throughout history, the authors affirm that most defies to machine translation technologies remain to translate homonymy and polysemy; word category ambiguity (or homography), structural ambiguity, noun stacking (like in the English expression "liquid oxygen tank", where there is a risk that machine translation will output the Portuguese translation, for instance, *tanque líquido de oxigênio* instead of *tanque de oxigênio líquido*, depending on the system approach), anaphora, idioms and so on.

Introducing the notion of writing for translation purposes, and, in particular, writing a text with machine translation in mind, Bowker and Buitrago Ciro explain, in Chapter 3, titled "Expanding the Reach of Knowledge Through Translation-Friendly Writing", that if native English-speaking researchers recognize how machine translation is likely to be used by researchers who have English as an additional language, they can write in such a way as to improve the translatability of their abstracts and make it easier for everyone to understand the machine translated contents. They present 10 guidelines for writing texts in a machine translation-friendly way, with emphasis to research abstracts, even though they recognize that the idea of writing in a clear and easily understandable language is not new, mostly emerged in the Linguistics literature with concepts such as International English, Standard English, Common English, Global English, Globish, Basic English, Plain English, and more. They based their guidelines in simple questions surrounding the concept of translation-friendly writing (or pre-editing): 1. for whom researchers are writing for? 2. How readers might be accessing the text? 3. What they want these readers to take away from the text? The chapter also provides some concepts and definitions on post-editing, which is referred by the authors as the process of fixing up a text that has been translated by a machine translation system in order to correct any errors and make the text sounds more natural. However, Bowker and Buitrago Ciro alert that it is difficult to present generic tips for post-editing because machine translation errors are dependent on the language combination, text type, subject field, and machine translation system and approach used.

Chapter 4, titled "Some Wider Implications of Using Machine Translation for Scholarly Communication", steps away from the details of how machine translation systems work to discuss how users can interact with them more effectively. The chapter also emphasizes that free online machine translation can now be easily accessed by

researchers around the world "with the simple press of a button". However, the use of machine translation can affect others, such as human translators. Researchers who choose to use machine translation systems have some ethical obligations toward those whose intellectual production is re-consumed in the process, even if this is simply recognizing the fact that machine translation has not eliminated the need for human translators but is instead highly dependent upon it. Bowker and Buitrago Ciro also advert that users may imagine that the data entered into a free online machine translation service simply disappear once the translation process is completed, but this is not true. On the contrary, machine translation service providers are typically interested in keeping these data and in possibly reusing it in the future (e.g., as training data for feeding systems).

"Towards a Framework for Machine Translation Literacy" is the title of Chapter 5, which introduces a framework for machine translation literacy training that could be used by literacy brokers, especially librarians, to design and promote effective instruction in machine translation. They start the chapter defining "literacy" as a competence or knowledge in a specified area, even though they explain that is it difficult to present an unequivocal definition of this term (literacy) once areas of both scholarly communication and machine translation are currently in a state of flux, because they are each evolving as the world around us – and our relationship with it – evolves too. They propose an effective instruction on machine translation in a workshop format that could be offered by universities where researchers could acquire the following skills:

- to comprehend the basics of how machine translation systems process texts;
- to understand how machine translation systems are or can be used (by oneself or by other scholars) to find, read, and/or produce scholarly publications;
- to appreciate the wider implications associated with the use of machine translation;
- evaluate how a (machine) translation-friendly scholarly text is;
- create or modify a scholarly text so that it could be translated more easily by a machine translation system; and
- modify the output of a machine translation system to improve its accuracy and readability.

Bowker and Buitrago Ciro also affirm that "machine translation" is not a single entity either, once this technology may be based on different underlying approaches (such as the aforementioned rule-based, statistic-based, *corpus*-based, or neural approaches) or on a combination of them. For the authors, different tools are therefore likely to produce different translations, and a single system will not perform equally well for all language pairs or directions. Bowker and Buitrago Ciro highlight that a given system may have a large volume of training data available for English and French but a considerably smaller set of training data for English and Icelandic, for instance, and that systems may perform differently when dealing with texts from different subject fields.

In their 12-line closing section, the authors conclude that writing the book has offered them a very satisfying opportunity to read many pages on research, communication, languages, literacy, and technologies, and to reflect on and to compare their own experiences as researchers and members of the scholarly community. In the authors' words, "[…] producing the book has been both an eye-opening and a rewarding experience which has led us to new reflections on machine translation and its applications in the research process." (BOWKER; BUITRAGO, 2019, p.95).

In my view, including training reflections and guidelines on the production of texts for scholarly communication with machine translation literacy in mind is the most important feature of the book, since machine translation history and approaches have been already widely explored by many others authors, as highlighted by Bowker and Buitrago Ciro themselves.

Although it has not been specially written to teachers, graduate or undergraduate students of Translation[2], Bowker and Buitrago Ciro's book inspire Translation teaching programs to optimize the training of machine translation in the translator's education, especially when the authors discuss the results of empirical surveys showing that the rate of scientific publication has increased tremendously since the end of World War II. Explored by the authors in Chapter 2, these surveys also show that professional translations performed by human translators have a high quality, like the ones delivered by scientific research agencies such as Elsevier, Taylor & Francis, Wiley, American Journal Experts, among others, while machine translation outputs, despite their advances and new methods, has yet a lower quality. In this sense, if scientific publications grow, machine translation teaching has to emerge as a priority in Translation programs. In my opinion, which was consolidated after reading Bowker and Buitrago Ciro's book, training translators on how to use machine translation technology is an element to consider, not only for them to make better use of this technology in their academic or future professional works, but to make them aware of its limitations, enabling them to better guide their upcoming clients and the scholarly community in general. By way of illustration, a very brief proposal of how to train undergraduate Translation students on machine translation can be found in the book *Tecnologias da Tradução: teoria, prática e ensino* (Translation Technologies: theory, practice and education) written by me and a colleague (ESQUEDA; STUPIELLO, 2019). The works of Gaspari (2001), Garcia (2011), O'Brien (2002, 2010, 2014) and Qun and Xiaojun (2015) have similar teaching proposals.

To sum up, I will certainly use Bowker and Buitrago Ciro's book to design translation technology courses I am responsible for at Federal University of Uberlandia, Minas Gerais, Brazil. Alike any other translation teacher, I am searching for inspirations to enhance my students' instrumental competence which nowadays is pretty much steered by technological impacts like the ones promoted by machine translation engines.

---

[2] "Translation" with "T" in uppercase refers to translation as a discipline or field. When used in lowercase, "translation" refers to the process of re-expressing texts from one language to another.

REFERENCES

BOWKER, L.; BUITRAGO CIRO, J. **Machine translation and global research**: towards improved machine translation literacy in the Scholarly Community. United Kingdom: Emerald Publishing Limited, 2019.

ESQUEDA, M. D.; STUPIELLO, E. N. A. **Tecnologias da tradução***:* teoria, prática e ensino. São José do Rio Preto: Ed. Transitiva, 2019.

GARCIA, I. Translating by post-editing: is it the way forward? **Machine Translation**, Dordrecht, n.25, p.217-237, 2011.

GASPARI, F. Teaching machine translation to trainee translators: a survey of their knowledge and opinions*. In:* FORCADA, M. L.; PÉREZ-ORTIZ, F.; LEWIS, D. R. (ed*.).* WORKSHOP TEACHING MACHINE TRANSLATION, 8., 2001, Santiago de Compostela. **Proceedings** [...]. Santiago de Compostela: [*s.n.*], 2001. p.35-44.

HUTCHINS, W. J.; SOMERS, H. L. **An introduction to machine translation**. London: Academic Press, 1992.

O'BRIEN, S. Teaching post-editing: a proposal for course content. *In*: EAMT WORKSHOP TEACHING MACHINE TRANSLATION, 6., 2002, Manchester. **Proceedings** [...]. Manchester: EAMT, 2002, p.99-106. Disponível em: https://www.researchgate.net/publication/228787033_Teaching_Postiting_A_Proposal_for_Course_Content'. Acesso em: 27 jan. 2020.

O'BRIEN, S. **Introduction to post-editing**: who, what, how and where to next. [*S.l.*]: Association for Machine Translation in the Americas AMTA, 2010.

O'BRIEN, S. *et al*. (ed.). **Post-editing of machine translation:** processes and applications. Cambridge: Cambridge Scholars Publishing, 2014.

QUN, L.; XIAOJUN, Z. Machine translation: general. *In:* SIN-WAI, C. (ed.). **The Routledge encyclopedia of translation technology**. London: New York: Routledge, 2015. p.105-119.

Received on February 19, 2020

Approved on July 2, 2020