

TTR

Traduction, terminologie, rédaction



Lynne Bowker and Jairo Buitrago Ciro. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley, Emerald Publishing, 2019, 111 p.

Brian Mossop

Volume 32, numéro 2, 2e semestre 2019

URI : <https://id.erudit.org/iderudit/1068912ar>

DOI : <https://doi.org/10.7202/1068912ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Association canadienne de traductologie

ISSN

0835-8443 (imprimé)

1708-2188 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Mossop, B. (2019). Compte rendu de [Lynne Bowker and Jairo Buitrago Ciro. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley, Emerald Publishing, 2019, 111 p.] *TTR*, 32(2), 250–255. <https://doi.org/10.7202/1068912ar>

Tous droits réservés © Brian Mossop, 2020

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

érudit

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

<https://www.erudit.org/fr/>

Lynne Bowker and Jairo Buitrago Ciro. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley, Emerald Publishing, 2019, 111 p.

This book, by a translation scholar and a library scientist, aims to promote the intriguing notion of “machine translation literacy” (MT literacy) among researchers in all fields (not specifically Translation Studies), and also among librarians (academic, school and public), professors of library and information science, abstractors and indexers, peer reviewers, journal editors, publishers, and translators/editors who deal with research. The authors define MT literacy (p. 88) as understanding the basics of how MT works and how it can be used to find, read and write scholarly publications; understanding the wider implications of using MT; and having the abilities to evaluate how MT-friendly a text is, to write or modify a text so as to make it MT-friendly, and to edit MT outputs to improve their accuracy and readability.

Chapter 1 looks at the rise of English as an international language of scholarly communication, and the various options open to non-native readers and writers of that language who need to publish in English. The authors talk about non-native *speakers*. However, there are many scholars who *speak* English as an additional language quite well but can’t *write* scholarly articles in that language well enough for acceptance. In this regard, I was not sure why they mention the colloquialization of academic writing in English over the past 30 years (pp. 60-61). This may well make reading easier for native readers, but probably makes it harder for those non-native readers whose mastery of the spoken language is weaker than their mastery of the written form. Both conditions exist: strong writing/weak speech and the opposite.

Options (not all of which may be available to a given researcher) include publishing in a language other than English (but then the article will not be as widely read and may not count as much for career advancement purposes), improving one’s English (very time-consuming), recourse to professional translators (very expensive) or editors (fairly expensive), asking a language teacher at one’s university to translate or edit (this will be problematic if the colleague lacks subject-matter knowledge), asking a colleague in the discipline to translate or edit (the colleague may want to be credited as an author and, I would add, may not be very good at translating/editing), and

approaching an online translation service in the gig economy (very cheap but the quality is likely to be poor, often just lightly edited MT output).

The final option is machine translation into English of a researcher's writing in their own language. The machine output can then be edited by a professional 'post-editor' or by someone in the discipline, either the researcher or an Anglophone colleague (who, according to some evidence, does not need to be familiar with the source language). The authors focus on self-post-editing by the researcher. They report (p. 26) a study in which researchers having little experience writing in English, and not very confident in that language, obtained (in their own opinion) better results when they self-post-edited machine translations into English from their own language than when they wrote directly in English. In that study, a professional reviser did not have to make more changes in the self-post-edited texts than in the texts written in English. However another study showed that not all necessary changes were made by self-post-editors who had no training in post-editing.

Next, the authors look at the use of MT during the literature search at the outset of a research project. The first problem is finding relevant material in English. The authors conducted an experiment (pp. 28-29) in which they used MT to translate Spanish keywords in a Spanish-only library science journal into English. They then used the English translations to conduct a subject search in *Library, Information Science & Technology Abstracts*. Of the 71 translated keywords, 37 returned articles on a topic similar to the corresponding Spanish article.

Once relevant English materials have been found, the second problem is understanding them. The authors look at several recent studies which confirm the old idea that MT can be useful for getting the gist of a text. I did wonder however at the underlying assumption here that if MT users are field experts, they will be able to detect significant errors in the MT output. Has this ever been tested? I also wondered whether author Buitrago Ciro, who is I presume a non-native writer of English, had tried out any of the procedures suggested here or elsewhere in the book.

The authors emphasize the need for MT users to be trained in how to get the most out of MT, and they wonder whether that is just a technical matter or also a matter of critical thinking about the

appropriate use of MT (p. 33). They also assert that academic librarians are best placed to provide such training, once they themselves have been trained (p. 34).

The authors say that MT is easy to use (just copy and paste into the source language box), but that is true only if you have a machine-readable version of your source text. Suppose you have only a paper photocopy. You can scan it to a pdf file but that yields only an image of a text. You still have to convert it to machine-readable text in the script of your source language, and that requires access to conversion software such as the paid version of Adobe Acrobat. Even if researchers have free access to such software through their university, they may not be aware of that. And I've discovered that not all reference librarians confronted with a photocopy of a Russian text are able to provide the required procedural knowledge.

Chapter 2 begins by explaining the main models of MT (rule-based, example-based, statistical, neural and hybrid), though I did not find the explanation of neural MT (p. 45) enlightening. Readers have to wait till Chapter 3 for a brief discussion of the problems with specific language pairs (pp. 76-77), and there is just a passing reference in that chapter, and again in Chapter 5, to the fact that different kinds of mistakes arise not just with different language pairs but with different MT models, different sets of training texts, and indeed different brands within a model (which the authors do not mention). The focus of Chapter 2 is on the central problem with MT: the difficulty of resolving ambiguity in the absence of real-world knowledge (pp. 46-49). There is no sign that MT systems are going to be able to apply such world knowledge.

The authors discuss the controlled languages (with very limited vocabulary and sentence structures) that some companies have used to prepare technical documentation. These reduce ambiguity, but there is no prospect of training academic researchers to write in such languages. However, Chapter 3 shows ways in which scholars can nevertheless make their writing more likely to yield useful MT output. It sets out 10 tips for writing in English that will make texts easier to read by those with some knowledge of the language and, as a by-product, will make the texts more MT-friendly (so that non-native researchers can get better translations of English reference material in their own language). Some of the tips are of the kind found in writing handbooks and plain language advisors addressed to native writers of English: avoid very long sentences, the passive voice and

long strings of nouns. The others are more related to the needs of non-native readers and MT: use nouns instead of personal pronouns, use terminology consistently, choose words that are unambiguous in context, avoid abbreviations, and avoid idioms, humour and cultural references.

The authors claim that “general principles” for making writing easier to read apply to “any language,” not just English (p. 91), but they do not set out these principles. As far as I can see, only two of their specific tips would be universally applicable (choose words that are unambiguous in context, and avoid humour and cultural references). Since we can assume that the two authors do not know all of the commonest languages of scholars in all fields, and the related rhetorical practices, their claim seems to have no basis. I doubt that the concept of “plain language writing” that has been present in the English-speaking world for the past few decades exists in all linguistic cultures. Indeed, the concept ‘plain,’ like ‘clear’ is not at all... clear.

The advice to use specialized terms consistently is no doubt useful for MT purposes (the authors replace “harmonize” with “align” when they edit a text where these two words are used as synonyms (p. 74), but a general elimination of synonyms will create inauthentic writing. Native readers who are subject-matter experts (which will be the case for professors though perhaps not junior graduate students) can recognize synonyms, and synonymizing is natural: people tend to write the word that comes to mind first, or simply want to vary their language (in some linguistic cultures, such as French, failure to use synonyms is generally seen as poor style). A text lacking synonyms may seem childish to native readers, and in this connection, the authors do mention in passing (p. 63) that accessible (‘plain’) writing may be boring. However, they counter that in research articles, what is important is getting content across.

The authors admit that it is difficult to change one’s writing habits. They interestingly suggest starting with abstracts. Rather than writing a translation-friendly abstract using the tips, they take one of their own abstracts, which they judge to be less than translation-friendly, and edit it in accordance with the tips. However they devote only a paragraph to an informal test of the relative quality of the French and Spanish translations of the unedited and pre-edited versions of the source English. (They do refer in passing to their own previous study of pre-editing for the purpose of MT (Bowker and Buitrago Ciro, 2018), as well as several studies by other researchers that compare MT

outputs for pre-edited and unedited source texts.) Chapter 3 ends with a discussion of post-editing. While it's true, as the authors say, that no general tips can be given (as already mentioned, the errors requiring correction by post-editors vary with the MT model, the language pair and the field of the text), they could have given some examples of light and full post-editing of raw MT translation into English. The book devotes much more space to pre-editing than to post-editing, though the authors do address the latter topic in Bowker and Buitrago-Ciro (2015).

Chapter 4 discusses some of the problems with using MT. First, if MT into English is used for dissemination of research, this bolsters the use of English as a *lingua franca* for research, adding to the already existing pressure to write in English. MT may eliminate the scientific rhetoric of the source-language culture when there is no equivalent in English (p. 81). Also, prestige may sometimes be attached to publishing in the local language rather than in English, so that a researcher has to choose between local prestige with its related rewards on the one hand and international recognition on the other. The authors even suggest that native English writers could consider using MT into other languages to help redress the trend to overwhelming English predominance (p. 82). A second issue is that the contribution of translators to the statistical and neural models of MT is practically never mentioned. Neither model could work at all without human translators providing the data (p. 83). A third issue is that providers of free online MT systems tend to keep the source text and can use it for various purposes (p. 85).

Chapter 5 sets out four modules for a half-day MT literacy workshop to be led by a librarian. The module titles are "Why think about machine translation in the context of scholarly communication?," "Overview of MT systems," "Translation-friendly writing and editing" and "Self-post-editing machine translation output." Modules 3 and 4 include hands-on exercises. I did wonder whether it was realistic to allow only 60 minutes for module 3, bearing in mind that the participants will very often be researchers (say chemists or biologists) who are not used to metalinguistic talk or linguistic exercises that involve things like rewriting English noun strings. Also, participants are asked to bring their own abstracts, which can be in their own language (p. 90). It is hard to see how this could work if there are speakers of several languages at a session. The sessions would, I suspect, need to be language-specific (a German-speaking librarian with

German-speaking researchers), and for the exercises, it seems to me that language teachers would need to be present. The authors say (p. 93) that librarians “may” want to collaborate with such teachers, but I suspect it would be a necessity: how many librarians can provide instruction in editing (e.g., assist German speakers having difficulty rewriting German compound nouns in an MT-friendly way)?

The book struck me as both very interesting and very well written, but then I am not a member of any of its intended audiences. It will be of special interest to non-native readers and writers of English, including Translation Studies scholars. Also, given the authors’ view that MT literacy training should be given by librarians, the reception of the book in library science journals will be crucial to the fulfilment of its purpose.

References

- Bowker, Lynne and Jairo Buitrago Ciro (2015). “Investigating the usefulness of machine translation for newcomers at the public library.” *Translation and Interpreting Studies*, 10, 2, pp. 165-186.
- Bowker, Lynne and Jairo Buitrago Ciro (2018). “Localizing websites using machine translation. Exploring the connection between user experience and translatability.” In S. Chan, ed. *The Human Factor in Machine Translation*. London, Routledge, pp. 138-151.

BRIAN MOSSOP
YORK UNIVERSITY, TORONTO

Jean-Marc Gouanvic. *Hard-boiled fiction et Série noire: les métamorphoses du roman policier anglo-américain en français (1945-1960)*. Paris, Classiques Garnier, 2018, 281 p.

Cet ouvrage s’inscrit dans la foulée de plusieurs autres publiés par Jean-Marc Gouanvic : *Sociologie de la traduction* (1999), sur la science-fiction américaine dans l’espace culturel français des années 1950; *Pratique sociale de la traduction* (2007), qui présente une étude du roman réaliste américain dans le champ littéraire français entre 1920 et 1960; et *Sociologie de l’adaptation et de la traduction* (2014), qui se penche sur le roman d’aventures anglo-américain dans l’espace littéraire français pour les jeunes jusqu’en 1960. Dans tous ces ouvrages, Gouanvic applique les théories sociologiques de Pierre Bourdieu à l’étude de corpus de livres traduits.