Machine Translation Literacy as a Social Responsibility

Lynne Bowker
School of Translation and Interpretation, University of Ottawa
70 Laurier Ave East, Hamelin Building (401), Ottawa, ON, K1N 6N5, Canada
lbowker@uottawa.ca

Abstract

Machine translation is easily accessible and easy to use, but this doesn't mean that everyone uses it in an informed way. We suggest that translators have a social responsibility for helping people outside the language professions to become informed users of machine translation, and that partnering with libraries could provide a means of reaching and educating a broad cross-section of citizens. We briefly summarize key elements of a machine translation literacy workshop that we piloted with two academic libraries, and we outline our plans for the next phase of the project with a public library.

Keywords: machine translation, machine translation literacy, social responsibility, ethic of care, libraries

Résumé

La traduction automatique est facilement accessible et facile à utiliser, mais cela ne signifie pas que tout le monde adopte une approche raisonnée. Nous suggérons que les traducteurs ont une responsabilité sociale de fournir une aide à ceux qui ne font pas partie des professions langagières pour qu'ils puissent devenir des utilisateurs critiques de la traduction automatique. De plus, nous suggérons qu'un partenariat avec les bibliothèques pourrait fournir un moyen d'éduquer un large éventail de citoyens. Nous résumons brièvement les éléments clés d'un atelier qui présente une approche raisonnée de la traduction automatique que nous avons mise à l'essai avec deux bibliothèques universitaires, et nous décrivons nos plans pour la prochaine étape du projet avec une bibliothèque publique.

1. Introduction

Although machine translation is nearly ubiquitous, not everyone is a critical user. We advocate for an ethic of care where translators can demonstrate social responsibility by helping those outside the language professions to become informed users of machine translation. We explain how partnering with libraries could allow translators to educate a broad cross-section of society, and we outline key elements in a machine translation literacy workshop that we piloted with two academic libraries. Finally, we briefly note our plans for the next phase of the project with a public library.

2. Machine Translation: An Evolving Landscape

The context of machine translation use has evolved considerably since the first tools were developed in the late 1940s and early 1950s. This includes a change in the types of user, in how the tools work, and in the type of education needed to ensure critical use.

2.1 Machine Translation is "In the Wild"

Until relatively recently, machine translation tools could be found primarily in the hands of researchers or language professionals. Then in 2006, Google launched a free, online machine translation system called Google Translate. Other companies followed suit with their own browser-based machine translation tools, including Bing Microsoft Translator, DeepL Translator, Baidu Translator, and Yandex.Translate, to name a few. In addition, some of these translation engines provide built-in translation options in tools such as the Google Chrome browser, or for platforms such as Facebook or Twitter. We could say that machine translation is now "in the wild", meaning that these tools are no longer restricted to language professionals but are available to everyone with an internet connection. Non-language professionals are using machine translation to assist with tasks such as conducting genealogical research (Vestal, 2016) or searching for international patents (Nurminen, 2019), among other uses.

2.2 Machine Translation is Easy to Use

Some translation technologies, such as translation memory systems, are still used primarily by language professionals. These tools can be quite sophisticated and require specialized training to learn *how* to use them. In other words, they can be complicated to use from a technical standpoint – knowing which file to open, which option to select, which filter to apply, and so on. In contrast, browser-based machine translation or built-in machine translation tools are very simple to use. Sometimes it takes just one click! However, the effortlessness with which we can employ these tools means that it is very easy to use them in an unthinking or non-critical way, which could lead to problems.

2.3 Machine Translation is Undergoing a(nother) Paradigm Shift

Another change that has occurred is that the underlying approach to machine translation has changed. Machine translation research began just after the Second World War. For approximately 50 years, the main approach to machine translation was known as Rule-Based Machine Translation (RBMT) (Hutchins and Somers, 1992). With RBMT, developers approached machine translation in a way that was similar to how linguists study language – through grammar rules and bilingual lexicons. These systems had limited success, and a common problem was that of "translationese", where the translated text would be awkward or overly literal. Around the turn of the millennium, researchers began to adopt a corpus-based or data-driven approach to machine translation, where statistics rather than linguistics took centre stage (Koehn,

2010). Statistical machine translation (SMT) approaches allowed computers to do what they excel at: number crunching and pattern matching. With SMT, translation quality got noticeably better, and it was during this period that the previously mentioned free, online machine translation systems first began to appear.

In late 2016, the underlying approach to machine translation changed again. Still data-driven, today's state-of-the-art machine translation systems use artificial neural networks, coupled with a technique known as machine learning (Forcada, 2017; Way, 2019). Developers "train" neural machine translation (NMT) systems by feeding them enormous parallel corpora that contain millions of pages of previously translated texts. NMT systems use these examples to "learn" how to translate new texts. With this latest paradigm shift, the quality of machine translation output has further improved. If the texts produced by RBMT systems were often laughable, the output of NMT systems, though not perfect, may be quite usable for many purposes. However, users must show good judgement. For instance, Castilho et al. (2017) found that found that NMT systems often produce text that is more fluent and contains fewer telltale errors such as incorrect word order or other forms of "translationese". However, just because the NMT output reads well doesn't always mean that it's accurate or right for a user's needs.

3. Machine Translation Literacy

Just because machine translation is easily accessible, easy to use, and produces a quality of output that is reasonable for some purposes, this doesn't mean that we instinctively know how to optimize it or even to use it wisely in a given context. The need for a new type of digital literacy is emerging, which we refer to as machine translation literacy (Bowker & Buitrago Ciro, 2019).

Martin (2006) describes digital literacy as the awareness, attitude and ability of individuals to appropriately use digital tools to identify, access, manage, integrate, evaluate, analyze and synthesize digital resources, construct new knowledge, communicate with others, and to reflect upon this process. This definition emphasizes that critical thinking, rather than technical competence, is the core skill of digital literacy. Like digital literacy, machine translation literacy is primarily a cognitive issue, rather than a techno-procedural one. Using machine translation is easy; using it *critically* requires some thought. When faced with free, online machine translation, the important question is not *how to* but rather *whether*, *when*, and *why* to use this technology. With regard to *how*, we could more usefully frame this as 'how can users interact with this tool in order to improve the usefulness of its output?' By asking ourselves such questions, we can become informed and critical users of machine translation tools, rather than being people who simply copy, paste, or click without a second thought.

4. Machine Translation Literacy as a Social Responsibility

Translator education programs typically incorporate translation technology training into their curricula, and professional translators' associations also offer options for technology-related professional development. Therefore, we can be hopeful that language professionals are (becoming) machine translation literate, or at least have the means to do so. However, it is not clear how the many people outside the language professions can learn to become savvy machine translation users. We would like to suggest that translators have a social responsibility in this regard.

Drugan and Tipton (2017) recently observed that relatively little attention has been paid to the question of social responsibility in relation to translation, prompting them to propose a Special Issue of the journal *The Translator* on the topic of translation, ethics and social responsibility. In it, Drugan (2017: 128) notes "we understand social responsibility as individuals' responsibility to the wider society in which they live; that is, interpreters' and translators' responsibility to the broader social context beyond the immediate translated encounter".

In addition to this special issue journal, we can observe some other ways in which the language professions are beginning to engage with social responsibility. For instance, the translation profession is generally regarded as being a caring profession where volunteerism is widespread. As described by Federici and O'Brien (2019), translation can play a key role in reducing risk in crisis situations (e.g. following disasters such as cyclones or earthquakes, or during the spread of infectious diseases), and there are several organizations (e.g. Translators Without Borders, *Solidarités International*) that use volunteer translators to help address humanitarian needs in collaboration with various non-governmental organizations (NGOs).

Meanwhile, Cheung (2017) proposes that *plain language* can be used for social good and indicates that technical communicators who use plain language are exercising social responsibility. She argues that marginalized populations (i.e., people who are oppressed for any reason) have a lot of worries to occupy their minds. The greater stress and mental burden that marginalized populations experience can leave less working memory available for tasks such as reading and learning. Cheung (2017: 448) states "Using plain language to reduce cognitive load can be considered a political act that increases marginalized populations' opportunities to understand." She thus presents the use of plain language by technical communicators as an ethical imperative.

4.1 Towards an Ethic of Care

We suggest that an ethic of care, such as that put forward by Noddings (1984; 2002), presents a good framework for encouraging members of the language professions to promote machine translation literacy to those outside these professions. To date, if professional translators discuss machine translation with non-translators, it has tended to take the form of warning them off using this technology. For example, as outlined in Bowker (2019) the website of the Canadian Translators, Terminologists and Interpreters Council (CTTIC) contained a message actively dissuading people from using machine translation and warning of the dangers of relying on machine translated output. The message on the CTTIC site noted:

> As part of their mandate, CTTIC's member organizations have a duty to ensure the protection of the public. As such, CTTIC and its members urge users to exercise the highest degree of caution, and to call upon a certified translator for all their translation requirements.

This approach does not seem particularly helpful. It is not realistic to think that all people who seek translation services can afford to hire professional translators, nor is it likely that all translation jobs require a professional level of quality. If translators truly see themselves as "having a duty to ensure the protection of the public", perhaps they could carry out a greater service to society by helping potential users to become machine translation literate instead of trying to convince them to steer clear of machine translation altogether.

In her early work on the ethics of care, Noddings (1984) distinguishes between 'caring for' and 'caring about', but she initially brushes aside 'caring about', noting that it is too easy and involves a sort of benign neglect:

> I can 'care about' the starving children of Cambodia, send five dollars to hunger relief, and feel somewhat satisfied. I do not even know if my money went for food, or guns, or a new Cadillac for some politician. This is a poor second-cousin to caring. 'Caring about' always involves a certain benign neglect. One is attentive just so far. One assents with just so much enthusiasm. One acknowledges. One affirms. One contributes five dollars and goes on to other things. (Noddings, 1984: 112)

However, she later revisited this decision, noting that while the basic distinction between 'caring for' and 'caring about' remains important, the concept of 'caring about' actually does warrant more attention. Indeed, her later work, she puts forward the idea that 'caring about' provides a link between caring and justice:

> … we learn first what it means to be cared for. Then, gradually, we learn both to care for and, by extension, to care about others. This caring-about is almost certainly the foundation for our sense of justice. (Noddings, 2002: 22)

Noddings (2002) explains that 'caring about' moves us from the face-to-face world into the wider public world, where we are moved by compassion for others' suffering, we regret that they do not experience being cared for, and we are outraged when they are exploited. In cases where we cannot directly care for others, we express our care in other ways, such as by donating to charities, supporting certain social groups, or voting.

Noddings (2002) is careful to point out that 'caring about' presents some inherent flaws. For instance, at its worst, it can become self-righteous or politically correct, it can encourage dependence on abstractions, and it can elevate itself above 'caring for' others. Nonetheless, Noddings (2002) believes that 'caring about' (i.e., a sense of justice) is instrumental in in establishing the conditions under which 'caring for' can flourish. In other words, although the preferred form of caring is 'caring for', 'caring about' can help to establish, maintain and enhance it.

In this vein, if translators care about their fellow citizens, they could show this by using their expertise to help others become more informed about machine translation. In this way, people will be in a position to decide whether or not this technology meets their needs for a given task, and if so, how they can use it effectively in a critical way.

5. Resources and Infrastructure

Since translators or their associations may lack a suitable infrastructure and resources to deliver machine translation literacy instruction, an option may be to form partnerships with different types of libraries. One reason for partnering with libraries is that they are typically cross-cutting units that reach a wide cross-section of the populations that they serve. For instance, an academic library serves the entire range of disciplines covered by its host institution, and it offer services to students, staff and faculty alike. Meanwhile, a public library cuts across socio-economic classes, offering services to all members of the public. In addition, both academic and public librarians are already charged with delivering other types of literacy instruction, including information literacy, media literacy and digital literacy (e.g., Julien, 2005). This experience makes librarians well equipped to partner with language professionals in order to offer machine translation literacy training as part of their programming.

5.1 Machine Translation Literacy Instruction in Academic Libraries

In autumn 2019, we conducted a pilot project where we delivered machine translation literacy workshops to international university students, faculty, and staff in collaboration with two university libraries in Canada: Concordia University Library in Montreal and the University of Ottawa Library in Ottawa. We ran three workshops at Concordia and two in Ottawa with a combined total of over 100 participants.

Examples of the type of information that was shared with workshop participants include suggestions such as:

a) **Don't enter sensitive material into an online machine translation system.** Information that you type or paste into a free online machine translation system doesn't simply "disappear" when you close the window. Instead, the companies that own the machine translation system (e.g. Google, Microsoft) could keep the data and use it for other purposes.

b) **Be sure to cite and reference ideas, even if you translate the words.** Academic integrity must be respected even when using machine translation tools to translate ideas into another language.

c) **Try more than one machine translation system.** Today's state-of-the-art neural machine translation systems use large corpora of previously translated texts as examples to "learn" how to translate new texts. Keep in mind that each machine translation system is trained using different texts, so each system might "learn"

different things. If one system doesn't provide helpful information, then try another one. Also, remember that these machine translation systems are constantly learning. If a particular system doesn't meet your needs today, try it again next month and you could get different results.

d) **Consider the purpose of the translation.** Machine translation may be more useful or less useful for different types of tasks or texts. If you are using the translation simply to help you *understand* a text that has been written in another language, such as reading a research article as part of a literature survey for your thesis, then, a machine translation system can probably be quite useful for helping you to get the gist or the main message of that text. However, if you're planning to use machine translation to help you *write* a text (e.g., a term paper or an article for publication), then be aware that unedited machine translated text is *not* likely going to be of a high enough quality for such purposes. The machine translation output will need to be edited to improve the quality.

e) **Improve the output by changing the input.** You might have heard the expression "garbage in, garbage out"? Well, if you want to use a machine translation system to help you produce a good translation, the best thing that you can do is to write the input text in a clear and easy-to-read way. We call this "translation-friendly writing", and it includes things such as using short sentences, avoiding humour, idiomatic expressions or culture-bound references, and using full forms instead of abbreviated forms.

Participants were surveyed about the workshops. The vast majority of participants indicated that they had learned new things. Most said that they plan to integrate machine translation more regularly into their scholarly work, and that they now feel equipped to do so more effectively. More than half the respondents replied that they would be interested in taking a more advanced follow-up workshop.

5.2 Next Steps: Working with Public Libraries

Following on from the success of introducing machine translation literacy workshops in academic libraries, we now plan to expand this type of training to a broader public. To this end, we are currently working with the Ottawa Public Library to explore how machine translation literacy instruction can be usefully adapted for different types of public library patrons. In particular, members of the newcomer or immigrant community are interested in learning how to become more informed and critical users of machine translation. Other possible target audiences include teens and seniors. We aim to pilot workshops that have been adapted for these groups in 2020.

6. Conclusion

Machine translation is being increasingly used in our society, where it has the potential to help if used critically, but to harm if used carelessly. Translators bear some responsibility for helping those outside the language professions to become informed users of this technology. Partnering with libraries can provide a means of reaching and educating a wide range of machine translation users.

7. Acknowledgements

This project was first developed as part of the Researcher-in-Residence program at Concordia University Library.

8. Bibliographical References

Bowker, L. (2019). Fit-for-Purpose Translation. In M. O'Hagan, editor, *Routledge Handbook of Translation and Technology*, pages 453–468, London, Routledge.

Bowker, L. and Buitrago Ciro, J. (2019). *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Emerald Publishers, Bingley, UK.

Canadian Translators, Terminologists and Interpreters Council (CTTIC). Homepage. Available online: www.cttic.org/ [last accessed 3 October 2018]

Castilho, S. et al. (2017). A Comparative Quality Evaluation of PBSMT and NMT Using Professional Translators. In *Proceedings of Machine Translation Summit XVI*, vol. 1, pages 116–131, Nagoya, Sept.

Cheung, I. (2017). Plain Language to Minimize Cognitive Load: A Social Justice Perspective. *IEEE Transactions on Professional Communication* 60(4): 448–457.

Drugan, J. (2017). Ethics and social responsibility in practice: interpreters and translators engaging with an beyond the professions. *The Translator* 23(2): 126–142.

Drugan, J. and Tipton, R. (2017). Translation, ethics and social responsibility. *The Translator* 23(2): 119–125.

Federici, F. M. and O'Brien, S. (2019). Cascading Crises: Translation as Risk Reduction. In F. M. Federici and S. O'Brien, editors, *Translation in Cascading Crises*, pages 1–22, Routledge, London.

Forcada, M. L. (2017). Making Sense of Neural Machine Translation. *Translation Spaces* 6(2): 291–309.

Hutchins, W. J. and Somers, H.L. (1992). An Introduction to Machine Translation. Academic Press, London.

Julien, H. (2005). Education for Information Literacy Instruction: A Global Perspective. *Journal of Education for Library and Information Science* 46(3): 210–216.

Koehn, P. (2010). Statistical Machine Translation, Cambridge University Press, Cambridge.

Martin, A. (2006). Literacies for the digital age. In A. Martin and D. Madigan, editors, *Digital Literacies for Learning*, pages 3-25, Facet Publications, London.

Noddings, N. (1984). *Caring: A Feminine Approach to Ethics and Moral Education*. University of California Press, Berkeley.

Noddings, N. (2002). *Starting at Home: Caring and Social Policy*. University of California Press, Berkeley.

Nurminen, M. (2019). Decision-making, Risk and Gist Machine Translation in the Work of Patent Professionals. *Proceedings of the 8th Workshop on Patent and Scientific Literature Translation*, pages 32–42, Dublin, August.

Vestal, P. (2016) Why I love Google Translate. *Assoc. of Professional Genealogists Quarterly* 31: 84–86.

Way, A. (2019). Machine Translation: Where are we at today? In E. Angelone, M. Ehrensberger-Dow and G. Massey, editors, *The Bloomsbury Companion to Language Industry Studies*, pages 311–332, Bloomsbury, London.