Towards robotic translation?

Anne Baillot[1], Ellen Carter[2], Thierry Grass[3], Pablo Ruiz Fabo[4][1]

[1] Le Mans Université, France

[2,3,4] Université de Strasbourg, France

Corresponding author: Ellen Carter, ecarter@unistra.fr

Post-Edition: Ellen Carter

Translator's note: in the spirit of this special issue, the introduction was translated from French into English using the DeepL neural machine translation service followed by very light post-editing focusing on technical terminology, gender neutrality, and avoiding ambiguity.

INTRODUCTION

The appearance on the web in 2017 of new translation services based on "deep learning" artificial intelligence algorithms [LeCun *et al.*, 2016; Koehn, 2020] such as DeepL and Google Translate [Kutylowski, 2022; Wu *et al.*, 2016] corresponded to a new leap forward in machine translation (MT). These recent systems, like the previous generation of statistical machine translation and factored machine translation [Koehn and Hoang, 2007; Koehn, 2010], rely on large aligned corpora and produce results whose quality is in some cases comparable to some human translations. It follows that to produce added value, the translator must provide something extra compared to the machine. This added value may be inherent in fields where the use of machines is not in itself of much interest because of the essentially aesthetic dimension of translation: this is the case for the translation of certain literary genres. Although much academic translation research refers to this field, it represents only a small fraction of the existing professional translation activity.[2]

As the machine allows productivity gains of 150 to 200% (some translators reach outputs of 6,000 to 8,000 words per day), the post-editing technique is becoming increasingly important in the language industries, as confirmed in 2017 by the introduction of the ISO 18587 standard "Translation services — Post-editing of machine translation output — Requirements" [ISO, 2017]. Post-editing poses a moral dilemma for translators: accepting that the machine, and not them, is the source of their own translation.

Will the profession of translator one day be supplanted by algorithms? Comparing technology with translations provided by amateur translators, some people are already predicting that humans will be overtaken by machines: this is the case of industry experts and academics who, according to a survey initiated by Katja Grace (Future of Humanity Institute, University of Oxford), predict that artificial intelligence will overtake translation by humans who are proficient in both source and target languages but are not professional translators by 2026 [Grace *et al.*, 2018]. For others, the machine's main obstacle remains its lack of knowledge of the world and therefore of understanding and interpreting it: some texts, such as legal or aesthetic texts, remain resistant to machine translation. Trials have been carried out to systematically put into perspective the advances of neural translation and the gap between its results and professional "biotranslation" or human translation on various types of text; Poibeau [2022]

[1] Authors appear in alphabetical order.

[2] Even in the case of poetry, a genre that presents considerable challenges for machine translation, early attempts at neural translation [Chakrabarty *et al.*, 2021; Ghazvininejad *et al.*, 2018] hint at the systems' potential.

Journal of Data Mining and Digital Humanities
eISSN 2416-5999, an open-access journal

http://jdmdh.episciences.org
https://doi.org/10.46298/jdmdh.10446

provides a synthesis. Complex linguistic phenomena, whose translation may require access to context beyond the sentence, remain particularly challenging for MT [Bawden *et al.*, 2018]. However, the tasks of the translator are changing and diversifying: post-editing is supplanting overwriting in many fields, localisation requires intercultural skills, and layout involves mastering desktop publishing tools.

This issue brings together articles that do not aim to take stock, but rather to present reflections and practical experience not only from professional translators but also from academics specialising in teaching and/or researching translation. The rapid development of artificial intelligence is forcing not only professional translators but also translation courses to adapt. The approaches put into dialogue in the framework of the interdisciplinary conference held at the University of Strasbourg in 2021 concern translation, but also software, linguistics, translation studies and language teaching.

I HISTORICAL AND LINGUISTIC APPROACHES

Machine translation is in line with different disciplinary traditions, some of them ancient. This section does not explore these traditions in their entirety, but at least it sheds light on some of their salient dimensions in the recent context of translation studies (Nicolas Froeliger) and in their historical context with an in-depth perspective on the precursors of formal logic (Marc Lebon).

Nicolas Froeliger begins by outlining the history of the positions taken by translation studies, mainly in France, regarding machine translation, since the end of the 1950s. He distinguishes three stages: that of a prejudicial objection in which machine translation served above all as a deterrent and as an exercise in thinking about what translation cannot be; that of popularisation, as the world and the market become aware of technological advances; and finally, today, that of convergence, with strong interdisciplinary content. As Nicolas Froeliger points out, the challenge in this hybridization is to maintain the scientific, professional, and societal relevance of translation studies.

In his contribution, Marc Lebon goes back much further in time: he looks at the attempts to encode and automate languages as they have been developed since the 17th century, in particular by Athanasius Kircher, John Wilkins and Gottfried Leibniz. He analyses on the one hand the principles underlying the development of their machines for making equivalences between languages, and on the other hand the mechanisms that may have led to their failure. His contribution sheds light on the conceptual proximity between these computational devices and the machines developed since the end of the twentieth century for the purpose of automating translation, pointing out that the debates we may have around these processes are perhaps older than we imagine.

II PEDAGOGICAL PRACTICES

The three articles in this section investigate the pedagogical use of machine translation and post-editing for undergraduate and graduate translation students in Spain (Julio de Los Reyes Lozano), the Netherlands (Joop Bindels and Mark Pluymaekers) and Switzerland (Lise Volkart *et al.*). The issues examined include translation quality for dubbing with and without machine translation, or under the influence of different post-editing strategies, as well as the spontaneous use of machine translation by students.

In the first contribution, Julio de Los Reyes Lozano presents a teaching experiment combining machine translation and audiovisual translation with translation students in Spain. After translating an excerpt from a television series for dubbing, the participants post-edited the same text and then evaluated the time spent and the typology of errors generated by the machine translation. While the results show difficulties in inserting machine translation into a dubbing project, a reduction in overall time was observed, which encourages further research in this area.

Secondly, Joop Bindels and Mark Pluymaekers' study investigates the use of machine translation by undergraduate students for different learning tasks. The results suggest that most students use machine translation regularly, although the frequency of use decreases as they progress in their studies. The

http://jdmdh.episciences.org
https://doi.org/10.46298/jdmdh.10446

results argue for a comprehensive approach to the development of machine translation knowledge and post-editing skills.

In their contribution, Lise Volkart *et al.* conducted an experiment with translation students to evaluate the influence of two different post-editing strategies on three aspects: post-editing time, the ratio of corrected errors and the number of optional edits per word. Their results show that the chosen strategy has no influence on the post-editing time, nor on the ratio of corrected errors. However, it does have an influence on the number of optional changes per word. Two other observations also stand out: firstly, the ratio of errors corrected shows that students only correct half of the machine translation errors, which highlights the need for post-editing practice. Second, when students are not required to read the source segment first, they tend to do monolingual post-editing.

III BIOTRANSLATION VS. MACHINE TRANSLATION

Through post-editing, which has become essential in many fields, human translation or "biotranslation" is increasingly adapting to the requirements of machine translation. For the time being, human translators are still irreplaceable in the translation process, even if they intervene at the end of the chain as post-editor. Adapting to machine translation requires that biotranslators are not only aware of the biases of machine translation, but also that their linguistic skills in the source and target languages are optimal compared to the increasingly powerful systems.

Maryam Alrasheed's article examines the positive aspect of human subjectivity as a tool for comprehension and contextualisation compared to the subjectivity of the machine, which could reside in the set of subjectivities present in the corpora used to train it.

For Françoise Bacquelaine, despite the progress of neural machine translation, artificial intelligence still does not allow the machine to avoid all the pitfalls of translation, especially those related to lexical ambiguity, phraseology, syntax, and semantics [Koehn, 2020]. She studied the translation into French of two Portuguese polylexical phraseological units with a medium degree of structural and/or lexical stability thus having the characteristics of "pre-fabricated construction units" described by Schmale [2013]. Therefore, they fall within the scope of phraseology in the broad sense and need to be translated as a block. The block translation of these pre-fabricated construction units is a challenge for the machine, especially because of the syntactic properties of splitting and inversion of elements on the syntagmatic axis. A sample of 168 occurrences of these units in phrasal context was taken from a Portuguese journalistic corpus. This sample was translated into French by DeepL and Google Translate in 2019 and 2021. The raw machine translations were compared with a biotranslation model (a set of linguistic strategies for biotranslation) developed based on analysis of parallel or aligned Portuguese-French corpora and analysed according to two general criteria (non-literality and acceptability) and some challenges specific to each unit. The analysis evaluates the evolution of these two machine translation systems in the face of phraseological ambiguity and draws conclusions about the threat of extinction for biotranslation and the implications of these powerful tools on the training of future language service providers.

The contribution of Katell Hernández Morin and Franck Barbin presents the OPTIMICE project (optimising machine translation of metadata and its integration into the editorial chain). This project aims to design a method combining neural machine translation and human post-editing to improve the quality of article metadata translated from French to English in the editorial process of scientific journals and formulates recommendations for writing and translating metadata.

Hanna Martikainen presents an experiment in adaptive machine translation. As part of a course on machine translation and post-editing, second-year Master's students carried out group projects on the Lilt platform. Hanna Martikainen analyses the students' views on the machine translation engine, focusing on their interaction with the technology. While the students recognise the potential of adaptive machine translation to enhance the role of the human in the loop, the quality of the machine translation and the usability of computer-assisted translation in general seem to have a greater influence on usability than the interaction with the machine.

http://jdmdh.episciences.org
https://doi.org/10.46298/jdmdh.10446

IV CHALLENGES FOR PROFESSIONAL TRANSLATION

For professional translators, the emergence of machine translation is leading to considerable shifts in practice. This section looks at two important dimensions of these practices: the skills required and the ways in which the result is evaluated. A third article completes these analyses by presenting a survey that sheds light on the field of medical translation.

Ralph Krüger's article focuses on the skills needed by translation professionals in the context of increasing process automation. In particular, he makes the link with the European Union's Digital Competence Framework, and with existing data literacy training initiatives. The competency-based approach allows him to conclude by making the link with professionalizing pedagogical practices developed at the TH Köln – University of Applied Sciences.

Eric Poirier, for his part, looks not at upstream skills, but at the evaluation of the result downstream. His analysis focuses on the methodology for evaluating translations and revisions and shows that digital methods can also make a substantial contribution to this dimension of translation work. The notion of information volume is at the heart of the suggested measures, defined here as a function of the number of characters and lexical words in the aligned segments to be compared. Heteromorphic segments (with different information volumes) are more likely to contain translation errors or to warrant revision. Alongside other methods, those suggested by Eric Poirier are applicable in the context of professional translation to assess how close a translation is to a predefined quality objective, and to what extent correction or revision is required.

The last contribution in this section, presented by Magali Vidrequin, offers practical insights into professional translation practices. She presents a survey of freelance translators working in the medical field. Magali Vidrequin analyses the different profiles and practices of professionals in terms of post-editing. The results she highlights underline the integration of machine translation in the translation environment of these professionals.

V THE CONTRIBUTION OF CORPORA

The so-called "empirical" approach to natural language processing (NLP), which includes statistical machine translation and neural machine translation, consists of inferring useful representations for carrying out a language task from corpora; in the case of machine translation, it is a matter of inferring models capable of translating based on aligned corpora [see Poibeau, 2017, ch. 7 and 9]. When large corpora are available for a language pair, it is mainly neural translation that has led to impressive gains in translation quality for a variety of text types. However, an inherent danger of corpus-based methods is the reproduction by machine learning models of the societal biases from the context in which these corpora are produced, as biases are reflected in the selection of documents included in the corpora and in the content of these documents [Bolukbasi *et al.*, 2016; Kumar *et al.*, 2020; Lu *et al.*, 2020; Sun *et al.*, 2019], be they gender, racial or other biases.

The first paper in this section by Damien Hansen *et al.* illustrates the use of corpora in machine translation in a rather original way, focusing on training models from literary corpora. Next, Isabelle Rivas-Ginel and Sarah Theroine examine gender bias in machine translations of video games, bringing new results to existing analyses of this type of bias in machine translation [Prates *et al.*, 2020; Saunders *et al.*, 2020; Savoldi *et al.*, 2021).

The work of Hansen and his co-authors presents a system trained particularly for the translation of literary works, specifically heroic fantasy, with the English-French language pair. Literary machine translation in the narrative domain with this language pair had not yet been tackled with neural methods, although experiments with statistical machine translation exist. Two neural models were tested, specifically transformers[3], which are at the heart of recent advances in natural language processing and machine translation. The quantitative evaluation metrics of the machine translations obtained for the

[3] LSTM: Long short-term memory. One of the models suggested to help manage long-distance linguistic dependencies (e.g., subject away from its verb).

http://jdmdh.episciences.org
https://doi.org/10.46298/jdmdh.10446

literary corpus are, not surprisingly, lower than the scores achieved in other domains. However, the experiments carried out show the extent to which it is possible to adapt a model to the literature and style of a specific author, or even to the style of a specific translator, in contrast to the results obtained with generic neural systems, not fine-tuned on literary corpora, which erase these stylistic features. The potential usefulness of these literature-specific models for computer-assisted literary translation is also discussed.

Studying the localisation of video games, María Isabel Rivas Ginel and Sarah Theroine analyse the percentage of gender bias resulting from the use of Google Translate, DeepL and SmartCat when translating content from English into French in the case of video games. Their study focuses on the three games *DeltaRune*, *The Devil's Womb* and *The Faces of the Forest* due to the presence of non-binary characters, non-sexualised characters, and female protagonists. By creating a parallel corpus, they compare the results of the three tools to identify and analyse differences in gender-related errors, to visualise the semantic and grammatical directions of words, and to extract collocations and concordance lines that represent gender identity by analysing the patterns in the source language.

VI FEEDBACK FROM PROFESSIONAL TRANSLATORS

One of the aims of the conference was to bring together research on machine translation with the everyday experience of translation professionals. This section sheds light on the experience of professional translators in industry and academia. It also shows how neural translation influences translation activities and the pedagogical practices of translation teachers.

Alain Volclair's article focuses on the specific case of legal translation. His paper raises translation questions and provides some answers through the prism of the legal translation of lawyers' letters from Italian into French. He also ponders good post-editing practices and the transfer of performativity through machine translation. Finally, he wonders whether it is desirable to let artificial intelligence resolve civil law litigation.

In a different vein, Dominique Defert humorously shares his experience of the recent contribution of neural translation to his work as a literary translator since 1984, specialising in authors such as Dan Brown, Patricia Cornwell, and John Grisham. He shows how (even neural) machine translation cannot, of course, translate humour, that which is left unsaid, or textual or rhetorical devices aimed at "intriguing, surprising, charming" readers. He thus compares his translations of bestsellers already on the market with what a neural translation would return for the same source text. However, he stresses that, in his literary translation practice, the usefulness of recent neural translation tools lies elsewhere: they are valuable to him as a reference tool and help him find information quickly, "without taking his hands off the keyboard".

Finally, Jean-Yves Bassole explores the implications of free machine translation tools for teaching translation. Translation errors made by seven free online machine translation tools are examined, based on nine extracts from literary texts (by Louis-Ferdinand Céline), containing spoken and colloquial language. The article addresses the concern of translation teachers who note a lack of critical reflection in the use of these tools, but also their pedagogical usefulness. It also shows how translation teachers try to make students understand various translation challenges based on translation errors made by machine translators.

References

Bawden R., Sennrich R., Birch A. and Haddow B. Evaluating Discourse Phenomena in Neural Machine Translation. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. 2018: 1304-1313. https://doi.org/10.18653/v1/N18-1118

Bolukbasi T., Chang K.-W., Zou J. Y., Saligrama V. and Kalai A. T. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. *Advances in Neural Information Processing Systems*. 2016; 29. https://proceedings.neurips.cc/paper/2016/hash/a486cd07e4ac3d270571622f4f316ec5-Abstract.html

Chakrabarty T., Saakyan A. and Muresan S. *Don't Go Far Off: An Empirical Study on Neural Poetry Translation* (arXiv:2109.02972). arXiv, 2021. http://arxiv.org/abs/2109.02972

Ghazvininejad M., Choi Y. and Knight K. Neural Poetry Translation. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*. 2018: 67-71. https://doi.org/10.18653/v1/N18-2011

Grace K., Salvatier J., Dafoe A., Zhang B. and Evans O. When Will AI Exceed Human Performance? Evidence from AI Experts. *Journal of Artificial Intelligence Research*. 2018; 62: 729-754. https://doi.org/10.1613/jair.1.11222

ISO. ISO 18587: 2017. Services de traduction – Post-édition d'un texte résultant d'une traduction automatique – Exigences. 2017. https://www.iso.org/fr/standard/62970.html

Koehn P. *Neural Machine Translation*. Cambridge University Press (Cambridge), 2020.

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

Koehn P. and Hoang H. Factored translation models. *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*. 2007.

Kumar V., Bhotia T. S., Kumar V. and Chakraborty T. Nurse is Closer to Woman than Surgeon? Mitigating Gender-Biased Proximities in Word Embeddings. *Transactions of the Association for Computational Linguistics*. 2020; 8: 486-503. https://doi.org/10.1162/tacl_a_00327

Kutylowski J. DeepL - Press Information. DeepL, 2022. https://www.deepl.com/press.html

LeCun Y., Bengio Y. and Hinton G. Deep learning. *Nature*. 2015; 521(7553): 436-444. https://doi.org/10.1038/nature14539

Lu K., Mardziel P., Wu F., Amancharla P. and Datta A. Gender Bias in Neural Natural Language Processing. In Nigam V., Ban Kirigin T., Talcott C., Guttman J., Kuznetsov S., Thau Loo B. and Okada M. (eds). *Logic, Language, and Security: Essays Dedicated to Andre Scedrov on the Occasion of His 65th Birthday*. Springer International Publishing (New York), 2020: 189-202. https://doi.org/10.1007/978-3-030-62077-6_14

Poibeau T. On "Human Parity" and "Super Human Performance" in Machine Translation Evaluation. *Language Resource and Evaluation Conference*. 2022. https://hal.archives-ouvertes.fr/hal-03738720

Poibeau T. *Machine Translation*. MIT Press (Cambridge, Ma.), 2017.

Poirier É. A. What Can Euclidean Distance Do for Translation Evaluations? In Bisiada M. (ed.). *Empirical Studies in Translation and Discourse*. Language Science Press (Berlin), 2021: 165-198.

Prates M. O. R., Avelar P. H. and Lamb L. C. Assessing gender bias in machine translation: A case study with Google Translate. *Neural Computing and Applications*. 2020; 32(10): 6363-6381. https://doi.org/10.1007/s00521-019-04144-6

Saunders D., Sallis R. and Byrne B. Neural Machine Translation Doesn't Translate Gender Coreference Right Unless You Make It. *Proceedings of the Second Workshop on Gender Bias in Natural Language Processing*. 2020: 35-43. https://aclanthology.org/2020.gebnlp-1.4

Savoldi B., Gaido M., Bentivogli L., Negri M. and Turchi M. Gender bias in machine translation. *TACL – Transactions of the ACL*. 2021: 845-874. https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00401/106991/Gender-Bias-in-Machine-Translation

Schmale G. Qu'est-ce qui est préfabriqué dans la langue ? Réflexions au sujet d'une définition élargie de la préformation langagière. *Langages*. 2013; 189(1): 27-45. https://doi.org/10.3917/lang.189.0027

Sun T., Gaut A., Tang S., Huang Y., ElSherief M., Zhao J., Mirza D., Belding E., Chang K.-W. and Wang W. Y. Mitigating Gender Bias in Natural Language Processing: Literature Review. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2019: 1630-1640. https://doi.org/10.18653/v1/P19-1159

Wu Y., Schuster M., Chen Z., Le Q. V., Norouzi M., Macherey W., Krikun M., Cao Y., Gao Q., Macherey K., Klingner J., Shah A., Johnson M., Liu X., Kaiser Ł., Gouws S., Kato Y., Kudo T., Kazawa,H., … and Dean J. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. *ArXiv:1609.08144 [Cs]*, 2016. http://arxiv.org/abs/1609.08144