Machine translation and "biotranslation": an unlikely marriage

Christine Breyel, Université de Strasbourg
Thierry Grass, Université de Strasbourg

The working environment of professional translators has been constantly evolving and two years ago, the emergence of "deep learning methods" and "neural machine translation" appeared on the market. These programs provide an acceptable translation quality never reached before, so that translation resource providers and professionals have to adapt to a new situation. The translation norm ISO 18587 (2017): "Translation services - Post-editing of machine translation output – Requirements" reflects this evolution. In this increasingly computerized environment, the "biotranslator" as Élisabeth Lavault-Olléon (2016: 2) calls him, is looking for his place and translation teaching faces an urgent necessity to train language specialists in order for them to acquire solid computer skills.

1. Deep learning machine translation: a new paradigm

Artificial intelligence is gradually being introduced into all areas of human knowledge. If we thought that the humanities would be spared, we were wrong. The progress of deep learning in the field of machine translation is now obvious and the arrival of DeepL in 2017 was a real thunderbolt in the world of professional translation. According to Forcada (2017: 291): "The last few years have witnessed a surge in the interest of a new machine translation paradigm: neural machine translation (NMT). Neural machine translation is starting to displace its corpus-based predecessor, statistical machine translation (SMT)." For Philipp Koehn (2017: 10), one of the pioneers of statistical machine translation, research is now completely focused on the neural model: "Within a year or two, the entire research field of machine translation went neural."

1.1. Short historical review

Within 70 years, Machine Translation (MT) has experienced at least 4 different stages of development. The first stage which was based on a word-for-word translation offered no quality at all, one could hardly understand the result. In this model, each word or phrase of the source language was directly linked to the corresponding unit of the target language. In 1954, IBM's Georgetown experiment took place: a machine translation (Russian English) was presented, based on 49 selected proposals in the field of chemistry. The system used a dictionary and only six syntactic rules. Four independent projects were presented, one of which became *Systran* which is still a big player in the world of professional translation.

The systems of the second stage, the so called rule-based system, are built on the model of language inspired by the work of Noam Chomsky[1] and his generative grammar enriched by the interpretative semantics of Fodor and Katz used to restrict context in cases of syntactic polysemy. The rule-based model comes in different sub-models (transfer-based, interlingual and dictionary-based) that we cannot describe here.

The third stage is a real breakthrough. The main idea is not to artificially create a translation using rules, but to reuse chunks of existing translations and put them together. The idea of reusing existing translations is attributed to the Japanese Makoto Nagao (1984). The

---

[1] Chomsky's Aspects of the Theory of Syntax and Syntactic Structures are published 1957 and 1965.

development of large parallel corpora in many different languages through the European institutions has somehow boosted this approach. Again, there are many subsystems and the penultimate one, the statistical model, was used by Google Translate less than five years ago. The fourth stage is the neural model that uses an artificial neural network to predict the likelihood of a sequence of words. The results obtained with such systems are sometimes bluffing and Microsoft claims to have achieved parity with human translation from Chinese into English[2].

1.2. The Advantages of the neural model

The novelty of the model is to proceed with a global contextual analysis: at all levels of analysis, the meaning of words is represented by their environment. Then, the words are grouped into larger semantically homogeneous sets called "word embeddings", a vector representation of each word of the dictionary based on real numbers. Not only the context of a single word is considered, but also the context of the semantically closest words, as mentioned by Poibeau & Reboul (2018):

> The system, by gradually discovering regularities in an incremental way, identifies groups of linguistically related words without the relationships between words being directly and explicitly formalized. The system reconstructs the syntax, in part, by itself.[3]

There is a prerequisite for a good translation according to the neural model: quantitatively representative parallel corpora (containing several millions of words) of good quality. This means that the best results will be achieved in the best endowed languages. For less well-endowed languages, there is therefore an interest in gathering quality parallel corpora to improve the results of neural machine translation. Although the company does not communicate much, it is often heard that the translation quality of DeepL is due to the quality of the Linguee corpus it owns. Neural systems need very abundant data to infer knowledge, which is a problem for languages with less resources that are not as well represented on the Internet, as mentioned by Choi et al. (2018: 900):

> We often try to collect resources to improve machine translation performance. Using the large size of a parallel corpus, it is possible to achieve high-quality machine translation performance. However, there are many cases where resources of language pairs are insufficient. Except for major European languages and some Asian languages, most of the language pairs do not have sufficient resources to develop a neural machine translation (NMT) system.

1.3. Technical principles

Like the statistical model, the neural model uses large parallel corpuses translated by human: it needs data to train (learn) the model parameters. While the "pure" statistical model creates several alternatives called hypotheses and assigns them a score, the neural model – which also

---

[2] https://blogs.microsoft.com/ai/chinese-to-english-translator-milestone/
[3] Original French text: « Un aspect intéressant de cette approche est que le système, en découvrant progressivement des régularités de façon incrémentale, identifie des groupes de mots linguistiquement liés. Autrement dit, sans que la syntaxe – c'est-à-dire les relations entre les mots – soit directement et explicitement formalisée, le système la reconstitue de lui-même en partie. »

uses statistics - predicts one word at a time. The structure is learned by the neural network and does not look like a linguistic structure. The system has the following features:

- It retrieves translation segments (word sequences) from labelled parallel corpora
- It uses a neural network to predict the probability of a word sequence in each context
- Deep learning uses several layers that simultaneously calculate the probability of various possible sequences from vector information on words.

As with statistical machine translation, it appears that the learning capabilities of neural systems are very sophisticated but require large amounts of data: the better the quality of parallel corpora, the better the results.

In contrast to statistical machine translation, neural systems work on a different computer model which requires special hardware (GPUs).

1.4. How does artificial intelligence contribute to machine translation?

As Herbig et al. (2019) from the German Research Center for Artificial Intelligence mention, artificial intelligence contributes to the improvement of a translation in several ways outside the translation process itself:

- It allows quickly finding and correcting errors on initial drafts
- Quality Estimation tools can assess the quality of an automatic translation without access to a reference translation
- The system "learns" by itself and adapts to human corrections avoiding repetitive errors
- The final translation output evolves through incremental re-training with updated linguistic data considering human physiology and behaviour

In the future, the system will be able to correct biases of machine translation models.

2. ISO 18587:2017: post-editing

As laid down in the general quality standard for translations (ISO 17100), the establishment of standards in the field of translation makes it easier to compare services in terms of quality and provides a guarantee that a service corresponds to a verifiable quality standard. However, the descriptions in the standard mainly simplify the definition of requirements for those involved in the project (post-editor - translation service provider – client/customer) and make it possible to use the same language and reach a common understanding of post-editing.

Following the publication of ISO 18587:2017 "Translation services — Post-editing of machine translation output — Requirements", a set of rules describing all the conditions under which translation services must be provided and the applicable requirements has been in existence since the beginning of 2018. However, the standard is limited to the full post-editing performed by a professional translator-reviser. This is a variant of post-editing, whereby the quality of the result comes close to that of a human translation ("comparable to a product obtained by human translation").[4] Light post-editing ("process of post-editing to obtain a

---

[4] ISO: 18587-3.1.5-3.1.6. https://www.iso.org/obp/ui#iso:std:iso:18587:ed-1:v1:en

merely comprehensible text without any attempt to produce a product comparable to a product obtained by human translation") is not taken into account in the standard.

In addition, the ISO 18587 post-editing standard precisely defines the terms and abbreviations used by professionals as well as, of course, the objectives of post-editing and the different stages of the post-editor's work. In paragraph 5, the mandatory qualifications and skills of the post-editor are presented, which correspond to the qualifications of the professional (human) translator as described in ISO 17100 ("Translation services — Requirements for translation services"). Special training in post-editing is also required. Overall, the knowledge and skills of the post-editor are almost identical to those required by the translator as set out in ISO 17100, which describes the requirements for human translation services and demonstrates that the high linguistic and technical qualifications and skills of translators and post-editors are essential prerequisites for quality service delivery.

2.1. Light and full post-editing

Overall, the requirements of this standard reflect the current process for the provision of post-editing, although they do not specify the difference between light post-editing and full post-editing. As mentioned by Nunes Vieira & Alonso (2018):

> Two target quality levels are normally referred to:
> - a lower level where translation products including stylistic issues may be deemed 'good enough' and be fit for purpose (where 'light post-editing' is usually suitable)
> - and a higher level where post-edited products need to be indistinguishable from a human translation carried out from scratch (where 'full post-editing' is often required)

The standard gives no indication as to the author of the superficial post-editing: it could be a non-professional translator, a trainee in a translation agency, as well as any person without any special translation skills.

2.2. Post-editing and proofreading

While proofreading means examining the target language content carried out by the human translator and applying corrections, post-editing means reworking machine translation outputs (draft translations). One needs to be familiar with the common errors of the MT-system to perform an effective post-editing. According to Zaretskaya (2017: 119):

> Many linguists expect MTPE [Machine Translation Post-Edition] to be similar to proofreading human translations. However, both the quantity and the nature of MT errors are different from those you can find in a human translation. Consequently, the effort and time necessary to post-edit MT texts is not the same as the one spent on proofreading a text delivered by a human translator.

While the human translator starts from scratch, the post-editor starts from a well-constructed draft. Although the vigilance of the reviewer and post-editor must be identical, one of the main problems with machine translation is that it looks quite perfect, above all if one translates in its B language. In other words, unconsciously, the result of machine translation will influence the

reviser's understanding of the text, and sometimes even distort the perception of his own language. This is particularly the case for texts' contents including idioms and terms (quite usual in the translator's work).

Post-editing is not revision nor review: both revision and review occur at the end of a human translation process. If both post-editor and reviser examine the target language content against source language content for terminological consistency, style and language level (ISO 17100 – 2015) there is one more dimension in review which consists in the "monolingual examination of target language content for its suitability for the agreed purpose" (ISO 17100 – 2015). Of course, these both phases are also part of the work of the post-editing station, but the term "revision" is only used for the correction of a human translation, which is quite different from the correction of a machine translation. While human translators behave unpredictably, machine translation encounters several linguistic problems we will discuss later.

2.3. Implications of post-editing

The publication of an ISO standard for post-editing implies changes in the translation process. More and more, post-editing becomes a standard working method for translators. Post-editing assumes that the professional translator knows the "defects" of machine translation and any serious training in translation must now offer training in post-editing.

The financial pressure on translators is increasing due to the use of many CAT tools to increase productivity: over the last ten years, client expectations have evolved from 2500 to 4000 words per day (300 to 500 words per hour). Some post-editors i.e. in Germany process 6000 to 8000 words per day, a number of words that seemed unthinkable even five years ago.

3. Common errors of machine translation

It is less expensive in terms of working hours to produce algorithms than dictionaries. A machine is less expensive than a translator. Post-editing refers to a cost logic, a "productivist" approach which is compatible with a qualitative approach. This refers also to the Time-Cost-Quality triangle in project management. In this section, we will discuss common and repetitive mistakes made by machine translation engines. In a previous article (Grass 2010), ten years ago, we mentioned the same problems and we were wondering if machine translation was of any use… It must be said that now most of the translation problems that prompted the mockery have been solved, but fortunately for the human [(bio)translator], the question of meaning has not yet been resolved and unconscious algorithms cannot replace the intelligence of the human being. All examples in this section have been translated with DeepL and/or Google Translate.

3.1. Polysemy and homonymy

This is still one of the main issues of machine translation, even though taking context into account in large translation segments has greatly improved quality.

The polysemic adjective "light" has at least three possible translations into French which are more or less equivalent in English.

Ce sac à dos est léger. (a peu de poids)
Le directeur est léger dans son travail. (est négligent)
Marie est une femme légère. (aimant le plaisir sexuel)

This backpack is light. (has little weight)
The director is light in his work. (is negligent)
Marie is a light woman. (loving sexual pleasure) [DeepL]

This backpack is lightweight. (has little weight)
The director is light in his work. (is careless)
Marie is a light woman. (is loving sexual pleasure) [Google translate]

In a segment without context, machine translation offers a translation that human knowledge of the world will consider improbable:

They saw many arms. (example by Christian Boitet)
Ils ont vu beaucoup de bras. [DeepL]
Ils ont vu beaucoup de bras. [Google translate]

In an idiomatic formulation, the restitution does not always live up to expectations. In this example, the whole novel has been translated automatically and the context therefore considered:

Je vous ai rendu là un rude service, allez. (Maupassant, Au printemps)

I have done you a great service, you must acknowledge. (Human translation, from now on HT)
I did you a big favor there, come on. [DeepL]
I made you a hard service, go. [Google translate]

In another example, the word "singe" (usually translated by *monkey* or *ape*) used in a familiar way means the head, the director, the boss in a company:

Mon désir de liberté grandit tellement, que, malgré ma répugnance, j'allai trouver mon singe. (Maupassant, Au printemps)

My yearning for freedom grew so intense, that, in spite of my repugnance, I went to see my chief [HT]
My desire for freedom grew so much that, despite my repugnance, I went to find my monkey. [DeepL]
My desire for freedom grew so much that, despite my reluctance, I went to find my monkey. [Google translate]

3.2. Synonymy

As with polysemy, the question of synonymy is still unresolved in machine translation. The human translator always chooses the best synonym. DeepL offers in its web interface a function to display the synonyms for each translated word and to correct the result of its translation.

Two British Free Tibet campaigners are in custody in China after unfurling a Tibetan flag and banner outside the Olympic stadium. (freetibet.net)

14

Deux militants britanniques du Tibet libre sont détenus en Chine après avoir déployé un drapeau tibétain et une bannière devant le stade olympique. [Google translate]
Deux militants britanniques de la campagne du Tibet libre sont en détention en Chine après avoir déployé un drapeau tibétain et une bannière à l'extérieur du stade olympique. [DeepL]
Deux militants britanniques de la campagne du Tibet libre sont en détention en Chine après avoir déployé un drapeau tibétain et une banderole à l'extérieur du stade olympique. [PE]

They are three possible translation of the word *banner* into French, the best one would be *banderole*.

bannière : étendard d'une confrérie, d'une société
étendard : signe de ralliement d'une cause, d'un parti
banderole : grande bande de tissu qui porte une inscription (en signe de protestation)

The translation made by the MT-system seems to be correct, but there is a better synonym. Because it is not really an error, the post-editor could accept the version as such. He must therefore be particularly vigilant regarding this type of error.

3.3. Semantic ambiguity

Two compounds may have the same syntactic structure but a different semantic structure:

to fly and *to clean* are both transitive verbs

Cleaning fluids can be dangerous (possible: cleaning fluids but not *to clean fluids)
Flying gliders can be dangerous (possible: flying gliders et to fly gliders)

Les liquides de nettoyage peuvent être dangereux.
Les planeurs volants peuvent être dangereux. [DeepL]
Les liquides de nettoyage peuvent être dangereux.
Les planeurs peuvent être dangereux. [Post editing, from now on PE]

3.4. Referential ambiguity

MT does not solve the referential ambiguity that results from a false anaphora recognition, note that translation students may have also problems with this! In the following example, the pronoun "him" refers to the president and "it" to the number:

Le président staliniste birman Ne Win avait un numéro préféré. (...) En 1987, il a aboli d'un seul coup tous les billets de banque qui n'étaient pas divisibles par lui.

Burmese Stalinist president Ne Win had a favorite number. (...) In 1987, he abolished all banknotes that were not divisible by him at once. [DeepL]
Burmese Stalinist President Ne Win had a favorite number. (...) In 1987, he abolished in one fell swoop all banknotes that were not divisible by him. [Google translate]
Burmese Stalinist president Ne Win had a favorite number. (...) In 1987, he abolished in one fell swoop all banknotes that were not divisible by it. [PE]

3.5. Modal constructions

Modal verbs are often used to lessen the impact of an utterance due to constraints on the interaction between the speaker and addressee. They are mainly used in the context of oral communication. As they are subject to interpretation, these modal constructions are difficult to translate automatically.

I'm not an expert but you might want to try restarting your computer.

Je ne suis pas un expert, mais vous voudrez peut-être essayer de redémarrer votre ordinateur. [DeepL]
Je ne suis pas un expert, mais vous voudrez peut-être essayer de redémarrer votre ordinateur. [Google translate]
Je ne suis pas un expert, mais vous devriez peut-être essayer de redémarrer votre ordinateur. [PE]

3.6. Idiomatic expressions

Not all idiomatic expressions are identified as such by machine translation engines and some errors may result i.e. *courir ventre à terre = to run like the wind*

Rappelons qu'il est inutile de courir ventre à terre pour maigrir. (http://www.courseapied.net)

Remember that it is useless to run belly down to lose weight. [DeepL]
Remember that it is useless to run to the ground to lose weight. [Google translate, May 2018]
Remember that it is useless to run belly to lose weight. [Google translate, February 2020]
Remember that it is useless to run like the wind to lose weight [PE]

3.7. Neologisms and archaisms

Both are difficult to translate and machine translation will tend to privilege word-for-word translation. Examples of neologisms:

Are you an Ego surfer? Or a Google stalker? Find out (rediff.com)

Êtes-vous un surfeur de l'ego ? Ou un harceleur de Google ? Découvrez-le [DeepL]
Êtes-vous un surfeur de l'ego? Ou un harceleur de Google? Trouver [Google translate]
Etes-vous de ceux qui googlisent leur nom ? Ou googlisez-vous les autres ? [PE]

Example on an archaism: "crinkum-crankum" = elaborate decoration or detail

I had no notion of being lost in so much light; but I had wander'd out of the main streets, and was got into the crinkum-crankum parts of London, where there are turnings and windings on every side. (The Sporting Magazine, 1802)

Je n'avais aucune idée d'être perdu dans tant de lumière ; mais j'avais erré hors des rues principales, et je me suis retrouvé dans les quartiers de Londres où il y a des tournants et des enroulements de chaque côté. [DeepL]

16

Je n'avais aucune notion d'être perdu dans autant de lumière ; mais je m'étais égaré dans les rues principales et je suis entré dans les quartiers de crinkum-crankum de Londres, où il y a des tournants et des enroulements de chaque côté. [Google translate]
Je ne me considérais pas comme perdu dans tant de lumière ; mais j'avais erré hors des rues principales et me retrouvais dans les quartiers enchevêtrés de Londres, caractérisés par des tournants et détours de toutes parts. [PE]

3.8. Proper nouns

Proper nouns are no longer so much of a problem in machine translation except for the adaptation of titles and transliteration.

Dr David James is the consultant at Leeds City Hospital.

Dr David James est le consultant à l'hôpital de Leeds City. [Google translate]
Le Dr David James est le consultant de l'hôpital de la ville de Leeds. [DeepL]
Le docteur David James est le consultant de l'hôpital de la ville de Leeds.

Transliteration from Russian varies from one language to another because the rules of transliteration are not identical, resulting in inaccurate translation of proper names at the spelling level.

Иван Тургенев (RU) = Ivan Turgenev (EN) = Ivan Tourgueniev (FR)

Ivan Turgenev was a novelist, poet and playwright, known for his detailed descriptions of everyday life in 19th century Russia. (russiapedia.rt.com)

Ivan Turgenev était un romancier, poète et dramaturge, connu pour ses descriptions détaillées de la vie quotidienne dans la Russie du 19ème siècle. [Google translate]
Ivan Turgenev était un romancier, poète et dramaturge, connu pour ses descriptions détaillées de la vie quotidienne en Russie au XIXe siècle. [DeepL]
Ivan Tourgueniev était un romancier, poète et dramaturge, connu pour ses descriptions détaillées de la vie quotidienne en Russie au XIXe siècle. [PE]

3.9. Words of foreign origin, borrowings and literal translations

Borrowings of foreign origin sometimes differ between the original language and the target language. German, for example, has words that look like they are borrowed from English but are in fact words that have a special meaning in German (i.e. *Handy* for *cell phone* or *Oldtimer* for *vintage car*). In French, these "false borrowings" also exist.

slow dance (EN) = *slow* (FR)

Quatre ans après leur premier slow, Fabienne et Nicolas sont mariés et attendent leur premier bébé. (ladepeche.fr)

Four years after their first slow, Fabienne and Nicolas are married and are expecting their first baby. [Google translate]

Four years after their first slow dance, Fabienne and Nicolas are married and are expecting their first baby. [PE]

3.10. Slang

It is not only in literary translation that slang expressions have to be translated. Slang is also frequent in chats and social networks.

Yo, dude wassup? = Hey man, what's up?

Yo, mec wassup? [DeepL]
Yo, mec wassup? [Google translate]
Salut mec, quoi de neuf ? [HT]

Wesh c est troo bien niska c est le boss du 91 (youtube.com)
Wesh it's troo bien niska it's the boss of 91 [DeepL]
Wesh it's too good niska it's the boss of 91 [Google translate]

Yeah buddy, that's great, Niska he's the boss of the 91' department

3.11. Segmentation

MT never replaces a dot with a comma, the boundaries of a sentence are strictly respected, as in following example where there are two sentences in the original text and only one sentence in the human translation:

Mon désir de liberté grandit tellement, que, malgré ma répugnance, j'allai trouver mon singe. C'était un petit grincheux toujours en colère. (Maupassant, Au printemps)

My desire for freedom grew so much that, despite my repugnance, I went to find my monkey. He was a grumpy little guy who was always angry. [DeepL]
My desire for freedom grew so much that, despite my reluctance, I went to find my monkey. It was a little grumpy always angry. [Google translate]
My yearning for freedom grew so intense, that, in spite of my repugnance, I went to see my chief, who was a short, bad-tempered man, who was always in a rage. [HT]

3.12. Abbreviations and acronyms

As with proper names, machine translation has made great strides in the recognition of acronyms, but some acronyms still need to be clarified or they may not be understood in the target language.

Si vous comparez l'affaire DSK à celle de Mike Tyson, le boxeur a passé plusieurs années en prison sur la base d'une accusation bien plus faible. (Wikipédia)

If you compare the DSK case to Mike Tyson's, the boxer spent several years in prison on a much lower charge. [DeepL]

If you compare the DSK case to that of Mike Tyson, the boxer spent several years in prison on a much lower charge. [Google translate]
If you compare the case of Dominique Strauss-Kahn to that of Mike Tyson, the boxer spent several years in prison on a much lower charge. [PE]

The use of the Internet has seen the development of a new writing technique using "brachygraphy", a term used in French linguistics for a system of writing using abbreviations or special characters. Some of them are not decrypted by machine translation systems.

Mdr son Manager il a du l'insulter dans toutes les langues possibles meme les langues mortes (tweeter.com)
Mdr / Ptdr = « mort de rire » et « pété de rire » / equivalent to English LOL "laughing out loud"

Mdr his Manager he had to insult her in all possible languages even dead languages [DeepL]
Mdr his Manager he had to insult him in all the possible languages even the dead languages [Google translate]
LOL his Manager had to insult him in every possible language, even the dead ones. [PE]

3.13.   Spelling

Spelling is not a real problem in machine translation if the source texts are written in accordance with the spelling conventions of the source language. It can be an issue when translating comments from commercial websites, forums, social networks, chats, etc.

100 personnes détienne la motier de l'argent de la planet c'est ouf (youtube.com)
100 personnes détiennent la moitié de l'argent de la planète, c'est fou (formulation correcte)

100 people hold the money motive of the planet it's phew. [DeepL]
100 people hold half the money in the world [Google Translate]
100 people hold half the money in the world. That's crazy. [PE]

3.14.   Interjections and exclamations

Interjections, although common in most languages, are not always recognized by machine translation systems. The same applies to slang expressions.

Ce cocorico, c'est l'Agence bio qui le pousse. (liberation.fr)
cocorico means literally ky-ky-ri-ký in Czech or cock-a-doodle-do in English. In French, it is also exclamation of victory which expresses one's patriotism.

This cocorico is being pushed by the organic agency. [DeepL]
This cocorico is the organic agency that pushes it. [Google translate]
It's the Bio Agency's cry of victory. [PE]

4. Consequences on teaching translation

The global translation market is estimated at 43 billion USD: Such a market attracts big players and freelance translators must share the leftovers. According to scientific and industry experts, artificial intelligence (AI) will be able to perform almost all human tasks between 2040 and 2050.

As Elisabeth Lavault-Olléon (2016: 2) points out, the acquisition of automatic language processing techniques has had a strong influence on today's translation studies:

> Or ce sont justement ces aspects appliqués qui méritent de plus en plus l'attention des chercheurs : le renouveau de la pédagogie dans le cadre de la formation professionnelle des traducteurs et des recherches sur les compétences ; la prise en compte des outils issus du traitement automatique des langues (TAL) et leur impact sur l'activité du traducteur humain devenu « biotraducteur »[5]

For teaching, the EMT model (European Master's in Translation) has become a guiding principle for training of translators Europe-wide, as summarized by Behr (2018: 9):

> Its aim is to support and steer the development of the following competences among future translators: competence in translation service provision, language competence, intercultural competence, competence in information mining, thematic competence, and technological competence. Professional translators, and translation service providers in particular, are encouraged to follow the international standard ISO 17100.

Translation training in the post-editing era must therefore focus on new aspects:

- Provide recommendations on how to be more efficient when post-editing (through focusing on common MT errors)
- Focus on the intelligibility of the target text, be aware that a text is being translated for humans (only humans really understand a source text and a target text and AI is substituting unconscious algorithms for conscious humans)
- Focus on standardization of translation procedures and quality as described in the ISO 17100 (2015) and ISO 18587 (2017) standards
- Remember that traditional translation skills remain crucial: a good post-editor is above all a good translator

In a survey of the French Society of Translators (SFT) conducted in 2019 and not yet published, we found that there is a fear for translators of losing their jobs or seeing the price of their work drop. There is also a reluctance of professional translators towards post-editing. The two main reasons are that post-editing is considered an alienating activity devoid of creativity, which is confirmed by Scheepers & Schulz (2017: 4): "CAT tools mostly use post-editing features, and translators in general find post-editing an extremely boring, tedious and unrewarding chore".

---

[5] However, it is precisely these applied aspects that increasingly deserve the attention of researchers: the renewal of pedagogy within the framework of professional training for translators and research on skills; the consideration of tools resulting from Natural Language Processing (NLP) and their impact on the activity of the human translator who has become a "biotranslator"; and the development of new tools to improve the quality of translation.

The second reason for translators' antipathy to machine translation has to do with copyright issues: through machine translation and post-editing, the translated text no longer belongs to the translator.

Nevertheless, translators are aware of the need to adapt to the evolution of their profession, which means focusing on new tasks with higher added value: there are new challenges in teaching translation: anticipating, fostering creativity, communication, teamwork, commitment to quality management and translation training programs must be diversified (writing, technical writing, transcreation, interpreting, communicative interpretation, etc.)

References

Behr, Dorothée. 2018. Translating questionnaires for cross-national surveys: A description of a genre and its particularities based on the ISO 17100 categorization of translator competences. *Translation & Interpreting* 10(2). 5–20.

Choi, Gyu-Hyeon & Shin, Jong-Hun & Kim, Young-Kil. 2018. Improving a Multi-Source Neural Machine Translation Model with Corpus Extension for Low-Resource Languages. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*. 900–904.

Forcada, Mikel L. 2017. Making sense of neural machine translation. *Translation Spaces* 6(2) 291–309.

Grass, Thierry. 2010. À quoi sert encore la traduction automatique ?. *Les Cahiers du GEPE* 2. (http://cahiersdugepe.fr/index.php?id=1367)

Herbig, Nico & Pal, Santanu & van Genabith, Josef & Krüger, Antonio. 2019. *Integrating Artificial and Human Intelligence for Efficient Translation*. (https://www.dfki.de/fileadmin/user_upload/import/10277_Integrating_Artificial_And_Human_Intellgence.pdf)

Koehn, Philipp. 2017. *Neural Machine Translation*. (3rd draft.) (http://mt-class.org/jhu/assets/nmt-book.pdf)

Lavault-Olleon, Élisabeth. 2016. Traducteurs à l'œuvre : une perspective ergonomique en traductologie appliquée. *ILCEA* 27. (http://ilcea.revues.org/4051)

Poibeau, Thierry & Reboul, Marianne. 2018. La traduction automatique passe à l'apprentissage profond. *La Recherche*. (https://www.larecherche.fr/technologie/la-traduction-automatique-passe-%C3%A0-l%E2%80%99apprentissage-profond)

Scheepers, Thijs & Schulz, Philip. 2017. *Interactive Neural Translation Assistance for Human Translators*. (https://thijs.ai/papers/scheepers-schulz-interactive-neural-translation.pdf)

Zaretskaya, Ana. 2017. Machine Translation PostEditing at TransPerfect – the 'Human' Side of the Process. *Revista Tradumàtica* 15. 116–123.

Christine Breyel
Université de Strasbourg
UR 1339 LiLPa
breyel@unistra.fr

Thierry Grass
Université de Strasbourg
UR 1339 LiLPa
tgrass@ unistra.fr