

# Machine translation literacy and undergraduate students in applied languages: report on an exploratory study



Rudy Loock  
Sophie Léchauguette



Rudy Loock  
Université de Lille, France,  
and CNRS UMR 8163 -  
STL - Savoirs, Textes,  
Langage  
rudy.loock@univ-lille.fr;  
ORCID:  
[0000-0003-0347-6138](https://orcid.org/0000-0003-0347-6138)



Sophie Léchauguette  
Université de Lille, France,  
and ULR 40 74 - CECILLE  
- Centre d'Études en  
Civilisations Langues et  
Lettres Étrangères  
sophie.lechauguette@univ-  
lille.fr;  
ORCID:  
[0000-0002-3878-8310](https://orcid.org/0000-0002-3878-8310)

## Abstract

This article reports on an exploratory study conducted on applied languages undergraduate students' use of machine translation. Starting from the observation that they make extensive use of free tools available online, our aim was to understand whether they are capable of identifying and correcting machine translation errors, and if so, to what extent.

**Keywords:** (neural) machine translation; MT literacy; applied languages teaching; translation teaching; language acquisition

## Resum

Aquest article mostra el resultat d'un experiment realitzat sobre l'ús de la traducció automàtica per part d'estudiants de grau en llengües aplicades. Partint de la premissa que ells habitualment utilitzen eines gratuïtes disponibles en línia, el nostre objectiu era entendre si realment són capaços d'identificar i corregir errors de TA i, si és així, fins a quin punt.

**Paraules clau:** traducció automàtica (neuronal); competències en TA; ensenyament de llengües aplicades; ensenyament de traducció; aprenentatge de llengües

## Resumen

Este artículo presenta los resultados de un experimento centrado en el uso de la traducción automática por parte de estudiantes de grado de lenguas aplicadas. Tras observar que utilizaban de forma habitual las herramientas gratuitas disponibles en línea, quisimos comprobar si eran capaces de identificar y corregir errores de traducción automática, y hasta qué punto.

**Palabras clave:** traducción automática (neuronal); competencias en TA; enseñanza de lenguas aplicadas; enseñanza de traducción; aprendizaje de lenguas



## Introduction

With the ever-increasing place of machine translation (MT) in our everyday lives, especially since the advent of neural machine translation (NMT), students from all disciplines have been experimenting with the use of online, free MT tools such as Google Translate or DeepL for their homework assignments. As a consequence, they need to develop a new competence, namely “MT literacy”, a concept put forward by Bowker & Buitrago Ciro (2019), for a “principled approach” to such tools (Loock, 2019). MT literacy is defined as users’ capacity to understand how MT systems work and can be used, to evaluate the MT-friendliness of a text, and to modify MT output (Bowker & Buitrago Ciro, 2019: 88). In this article, our focus is on one aspect of MT literacy: our students’ capacity to correct MT output, that is the translation of a text that has been selected and automatically translated for them (students did not choose the text nor the MT tool). Contrary to popular (and student) belief, most of the time MT output is never ready-for-use and often requires post-editing (PE), with both accuracy and fluency errors. In particular, with NMT giving priority to fluency in the target language, sometimes at the expense of accuracy (see e.g. Bojar et al., 2016, Macken et al., 2019), errors are more difficult to identify, both for students and professionals (e.g. Castilho et al., 2017a/b, Yamada, 2019). This means that the notion of “trust” in MT output will be a key factor in users’ capacity to assess the appropriateness of what the machine provides. The question of trust boils down to the more general question of human-machine interaction: how to make sure that the human is *assisted* and not *misled* by the machine because of excessive trust in the information generated by online tools? As far as MT is concerned, this issue has become crucial for professionals and professionals-to-be in the translation industry, but also for any user outside the industry because of the widespread availability of MT technology.

This article reports on an exploratory study carried out with undergraduate students (n=169) in applied languages at the University of Lille, France. The students were in the final year of a three-year programme offering compulsory translation classes each year. Starting from the observation that students use MT tools on a regular basis (see below), we wanted to find out the extent to which they were actually able to “correct” the output provided by an MT system (we are avoiding the term “post-edit” here, since the students receive no post-editing training). An English press article was translated into French with DeepL, a well-known, publicly-available MT tool. Students were provided with both texts and asked to correct the MT output. MT errors covered accuracy issues, as well as grammatical and punctuation errors, wrong collocations, and stylistic issues in the target language. The students’ corrections were then analysed to determine the kinds of errors they were able/unable to identify, as well as the appropriateness of their solutions. Complementarily, students answered a survey on their use and perception of MT tools. Our aim was to evaluate our students’ MT literacy – precisely their capacity to deal with MT errors – and specifically to check both the extent of their trust in the results provided by the machine, and the extent of their need for specific training to use MT tools effectively. As future professionals, students need to optimise their use of information provided by computers; placing them in the role of critical decision makers is a way to raise their awareness of the limitations of MT.

The remainder of the article is organized as follows. The first part explains students' need for a new form of digital literacy, namely MT literacy, in the light of the challenges posed by neural machine translation. The second part provides methodological information on our study, while the third part discusses the results and their implications for the training of students in applied languages, as well as the study's limitations.

### 1. Starting point: the need for MT literacy, a new form of digital literacy

These last few years have witnessed an impressive development of MT tools both inside and outside the translation industry. With the advent of neural machine translation, quality has greatly improved, and thanks to free, easy-to-use, online websites such as DeepL and Google Translate or smartphone applications, machine translation has become increasingly popular. The technology can easily be described as disruptive, since it is changing both the business model and the working processes in the translation industry. The year 2018 was a turning point in Europe since for the first time more than half of European translation companies used MT (source: 2018 European Language Industry Survey Report). The general public is influenced both by the promotion of services like smartphone applications or simultaneous interpretation devices, and by "human parity" claims (see e.g. Hassan et al., 2018), which have been widely circulated by the media (Moorkens, 2018 uses the expression "media hyperbole"). Scientific experiments keep debunking the myth (see e.g. Toral et al., 2018) but fail to reach the general public. As a result, there is now a widespread belief "that MT is now good enough and can replace the work of professional translators" (O'Brien & Ehrensberger-Dow, 2020: 146). Although translation professionals know that this is untrue, easy access to MT tools somehow contributes to reinforcing the layman misconception of what translation really is, and has made translation teaching, and language teaching in general, a challenge, whatever the discipline.

A number of studies on the subject show that language students, even before the advent of neural MT, have been extensively using online MT tools. For instance, in a survey of 310 university students of Spanish or French as a foreign language, O'Neill (2019: 163) found that 87.7% of them use online translators, whether for non-graded or graded work, even when this is prohibited by the university rules and regulations. Other studies, reported in O'Neill (2019: 155-156), provide similar results. White & Heidrich (2013), who focused on a class of intermediate German university students, found that 67.8% of them admitted to using MT to complete classwork. Clifford, et al. (2013) surveyed 905 Romance language students and 88% of them reported using an online translator at some point. Jolley & Maimone (2015) found that among a population of 128 students enrolled in university Spanish programs, 74.22% admitted to frequent or occasional MT use, for translation and interestingly also for writing tasks, with a majority thinking that MT use has a positive impact on language learning in general. Resende & Way (2021: 72-73) investigated the use of MT tools by 90 Brazilian Portuguese speakers learning English as a second language (beginner to advanced levels) and found that all of them used MT, with a more frequent use at intermediate and advanced proficiency levels and with a preference for the translation of words rather than complete sentences

(use of MT as a bilingual dictionary). As will be detailed below, 83% of the students in our study admit to using an MT tool when preparing their homework assignments. These few results, although they focus on a limited number of languages, clearly illustrate how widespread the use of MT tools in university language courses has become, and we believe it is no use forbidding them. The time has come for specific training to help students use them in an effective manner for translation, writing or any language learning task.

This does not come without challenges, both for students and their trainers: if “[u]sing MT is easy, using it critically requires some thought” (Bowker, 2020: 4). Among various challenges is the need for users to understand that MT output is rarely perfect and systematically requires checking, even if machine-translated texts can sometimes be quite impressive since NMT lays the emphasis on the fluency of the target language. As a consequence, MT errors are now harder to spot (e.g. Castilho et al., 2017a/b; Yamada, 2019) and students might overlook accuracy errors or omissions because they fail to pay enough attention to the source text. Cases of omissions are not rare in NMT: for instance, Loock (2020) shows that the proportion of -ly adverbs that are not translated in the case of English-French machine translation ranges from 3.5% to 10% depending on the MT tool. Another challenge is that when reading the target text, students might be “heavily primed” by the MT output (Carl & Schaeffer, 2017), as is already the case for professional post-editors, who “more easily accept sub-optimal translations which human translators, working from scratch, would otherwise not produce” (Carl & Schaeffer, 2017:44). This has been confirmed by a number of studies like Yamada (2019), which is a follow-up study on students’ use of NMT after Yamada (2014) on statistical machine translation (SMT). The study shows that when post-editing an English text into Japanese, college language learners majoring in English with some PE training still fall short of meeting professional standards, although their results are better than in the case of SMTPE. According to Yamada, this is because NMT makes the same kinds of errors as those found in human translations, errors that are more difficult to identify for students, who need to learn what to focus on during the post-editing/correction of MT output. Finally, yet another challenge is the fact that the widespread use of online MT tools in the language classroom can influence the language learning process. For instance, thanks to a “syntactic priming study” with Brazilian Portuguese learners of English as a second language, Resende & Way (2021) have shown that syntactic constructions provided by Portuguese-to-English MT output has an influence on learners’ production through the reuse of syntactic constructions. The authors conclude that MT has a “robust long-lasting priming effect” (Resende & Way, 2021: 82).

All of these challenges suggest that students might place excessive trust in what the machine has to offer, and might not be aware of the limits of the technology. This kind of attitude is nothing new and is not related to the development of MT tools: students have often been found to tend to ‘overtrust’ translation aids like general bilingual dictionaries (Meyer 1988), translation memories (Bowker 2005), or electronic corpora (Loock 2016).

Hence the need to empower students as users of online MT tools so that they develop a new kind of digital literacy, namely “MT literacy” as defined in Bowker & Buitrago Ciro (2019), which focuses on the use of MT by the scientific community. Digital literacy refers to “all sets of specific skills and competences needed for searching, finding, evaluating and handling information in computerized form” (Shopova, 2014: 27). MT literacy refers to specific skills in relation to the capacity to understand how MT systems work, when they can be used (MT-friendliness of a text and suitability of the MT tool), and when/how to modify MT output (Bowker & Buitrago Ciro, 2019: 88). MT literacy also implies applying a critical eye in order to develop “a healthy level of mistrust in [MT] output” or a kind of “healthy scepticism” (O’Brien & Ehrensberger-Dow, 2020). As a consequence, trainers have been wondering about the best way to introduce MT tools in language classes, whether dedicated to translation or not. In this article, we focus only on one element of MT literacy: students’ capacity to identify and correct MT output in order to provide a translation that is both accurate and fluent.

MT teaching has become the norm in translation programmes: it is now part of the list of competences defined by the European Commission’s European Master’s in Translation Network, and most translation programs now include specific training on the use of MT tools and post-editing. However, language programs have also been wondering about whether to introduce MT tools into the curriculum, a decision neither consensual nor neutral. If it is decided that indeed MT tools should be introduced, the next questions are when, and how? These last few years, many surveys and experiments have been conducted on the use and influence of MT in the language classroom (e.g. Tsai 2019, Niño 2020, Resende & Way 2021 for recent examples – see also the conference organised by the University of Nottingham “Google Translate and Modern Languages in Education”, <https://www.nottingham.ac.uk/conference/fac-arts/clas/google-translate-and-modern-languages-education/index.aspx>). Also relevant for the challenges posed by MT for language teaching are the discussions on how to sensitize students to the limits and risks of the technology, e.g. fluent but inaccurate translations, algorithmic biases, discriminatory language, omissions, lexical variety loss (as found e.g. in Vanmassenhove et al., 2019, De Clercq et al., 2021), the existence of machine-translationese and even post-editese. These shortcomings of the technology can be of particular interest for language learning.

We are personally convinced that failing to integrate the use of MT tools in language and/or translation teaching is counter-productive, as it leaves students alone to understand the strengths and limitations of the technology. In the same way, prohibiting MT use is also problematic, as students will use online MT tools anyway (see O’Neill, 2019). This is the reason why we have conducted the study described in this article: How effectively do language students use MT tools for translation tasks? To what extent are they capable of spotting and correcting MT errors, whether accuracy or fluency errors? What kind of training should be considered to empower students as critical users and develop their MT literacy so that they do not allow the machine to mislead them? We now turn to the description of our study.

## 2. Our study

### 2.1. The students' profile

Our exploratory study was conducted on students enrolled in a three-year applied languages university programme at the University of Lille in 2019-2020. Such students attend economics, law, management, and communication classes, and specialise in two foreign languages: English and another language to be chosen from among Chinese, Dutch, Italian, German, Polish, Portuguese, Russian, Spanish, Swedish, or French for non-native speakers. Specific translation classes, where students translate from the two foreign languages into French and vice versa, are mandatory from the second semester of the programme. It is important to stress that in our context – French universities – applied languages students do not study linguistics or applied linguistics. Also, translation classes consist of what we call “pedagogical translation” (*traduction pédagogique*), that is translation exercises which aim to develop language skills. This is very different from training in professional translation. The students receive no training on the use of computer-assisted translation (CAT) tools or professional post-editing of MT output. Also, most of these students work in the tourism industry or in the international relations sector after graduating, with only a minority of them going on to enrol in a translation master's programme; for instance, out of the 169 students who took part in our study, only 7 enrolled in our translation master's programme the following year – 4%.

For our study, we tested students in their final semester, which means that when they took part in the study, they had attended 5 semesters of (language-learning) translation classes, in which they translate different kinds of texts: press articles, tourism leaflets, extracts from websites, but no literary texts. In the third year, exercises comparing human translation and MT output are introduced to help students understand the strengths and limitations of online MT tools, but no training on professional use of MT or post-editing is provided. As the MT output to be corrected was a translation from English into French, international students, whose mother tongue is not French, have been removed from our results.

### 2.2. The correction of the machine-translated text

The study was conducted online in April 2020 with 169 students, who (i) took a test consisting in the correction of an English press text translated into French with the free and online version of DeepL ([www.deepl.com](http://www.deepl.com)), and (ii) filled in a questionnaire on their use of machine translation tools for translation classes. Both tasks were performed remotely, with students free to use all material and tools which they thought to be relevant. While the correction exercise was part of their evaluation for the semester and therefore compulsory for credit attribution, the questionnaire was filled in anonymously by students on a voluntary basis (89 students answered the survey, a little more than half of those who took the test).

For the correction exercise, students were provided with a 408-word English source text, extracted from a recent press article published in *The New York Times*, “Coronavirus Empties European Cities of Chinese Tourists”. The article was provided alongside its

(unmodified) machine translation into French obtained via DeepL, by copying/pasting the English text on the website. Both texts are provided in Appendix 1. The test was organised online on 3 April 2020 and lasted 2 hours. Students were instructed to read carefully both texts and to modify anything they thought was a translation error, either because of meaning issues (accuracy) or because of problems with the French language in the target text (correctness and fluency). Students were not asked to justify their decisions; they were requested to modify the text directly and send it back without any specific comments. Errors included wrong use of moods/tenses, sentence structure based on source text, inaccurate translations, wrong choice of terms, punctuation misuse, wrong collocations or stylistically awkward sentences. Twenty errors were identified before the test by the teachers in charge of the course (n=2). The errors were important enough to need correction according to them, and diverse enough to demonstrate what kinds of errors students are more able to identify and/or correct. The twenty errors are listed in Table 1.

	MT error	Type of error	Correction(s)
1	vide les villes européennes de touristes chinois	determination	vide les villes européennes de leurs touristes chinois
2	compter sur les visiteurs	vocabulary	dépendre des visiteurs
3	visiteurs	vocabulary	touristes
4	était à peine plus longue que celle de Paris	accuracy	n'avait rien de la file d'attente parisienne typique/ne ressemblait pas à une file d'attente selon les normes parisiennes
5	tous	grammar (pronoun/antecedent)	toutes les personnes
6	dictait le mandarin	grammar (verbal complementation)	dictait quelque chose en mandarin
7	parfois la queue a été encore plus courte récemment	style	ces derniers temps, la file d'attente a parfois été plus courte
8	"D'habitude..."	punctuation (unnecessary quotation mark)	D'habitude...
9	un matin récent	style	il y a quelques matinées de cela

10	dans le dos	accuracy	l'arrière (du magasin)
11	une étape incontournable dans l'une des destinations	grammar (structure)	une étape incontournable dans liste des destinations
12	la ligne qui y est tracée	vocabulary	la file d'attente que l'on y trouve
13	preuve éclatante	collocation	preuve évidente/indéniable/indiscutable
14	a eu	grammar (tense)	a
15	après qu'un touriste... soit mort	grammar (mood)	après qu'un touriste... est mort/après la mort d'un touriste
16	qui approvisionnent le marché chinois	accuracy	qui visent un/le //s'adressent à un/au marché chinois
17	dans les destinations	grammar (choice of preposition)	pour les destinations
18	l'impact a été limité jusqu'à présent	style	l'impact est resté jusqu'à présent limité
19	que les Chinois ont visité en deuxième position après les Américains	style	que les Chinois ont été les plus nombreux à visiter derrière les Américains
20	a déclaré	style (repetition of the same verb)	a expliqué/ajouté or selon...

*Table 1. List of 20 MT errors to be corrected by the students*

### 2.3. The questionnaire

After taking the online test, students were sent an online questionnaire via a Google Form, but had no obligation to answer it. The answers were completely anonymous. Students who submitted a positive answer to the first question, “Do you use any machine translation tool for your homework assignments?”, had to answer all the following questions. In total, 89 students volunteered and sent their answers, which was meant to shed light on their use of MT tools: which websites, method (translation of full text, isolated sentences, chunks of sentences, words), specific aims (help with source text



understanding, lexical choices, sentence structure, or target language correctness), trust in and satisfaction with MT tools' results, self-perception of their capacity to identify and correct MT errors and use of MT tools outside university. The 11 questions of the questionnaire are provided in Appendix 2. As the answers to the questionnaire were anonymous, we have no information on the profile of the students who volunteered to fill it in; also, no correlation could be investigated between their answers and their capacity to identify MT errors. This is a limitation of our study (see section 3.3), but we wanted students to answer the questionnaire as freely as possible.

### 3. Results and discussion

#### 3.1. The identification and correction of MT errors

Our results show that in general, students fail to identify most of the errors made by the MT tool. Out of the twenty errors identified as requiring correction (see Table 1), on average less than 8 errors out of 20 were spotted (37.9%) by each student, and only two thirds of them were appropriately corrected (see Figure 1). This means that students managed to spot AND correct approximately one MT error out of four (26.45%) only. On average students left unmodified 12 MT errors out of 20 (62.10%).

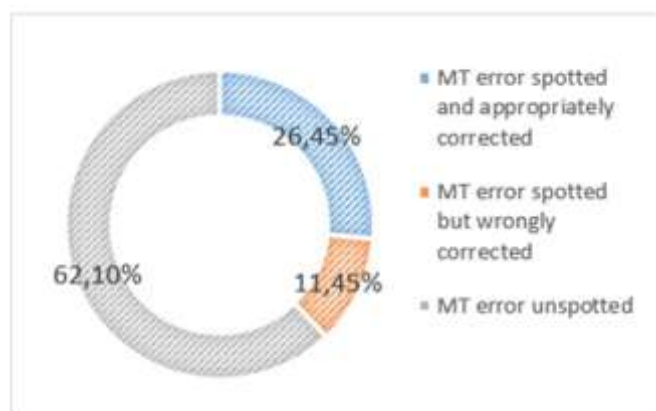


Figure 1. Overall results: (un)spotted and appropriate/wrong corrections

In particular, students seem to have been “heavily primed” by the MT output (Carl & Schaeffer, 2017, see above), and unable to identify fluency-related errors, for example in the case of “dictait le mandarin” (< “dictating Mandarin”) or “dans les destinations les plus populaires” (< “in the most popular destinations”). Only in cases where the target sentence clearly lacks in fluency did students consider the MT output to be incorrect, e.g. “que les Chinois ont visité en deuxième position après les Américains” and “parfois la queue a été encore plus courte récemment”: 106 students (62.7%) and 95 students (56.2%) provided alternative translations for “which Chinese were second only to Americans in visiting last year” and “[s]ometimes, the line’s been even shorter recently”. In the same way, 109 students (64.5%) corrected the determination error in the title with “empties (...) of Chinese tourists” being literally translated with “vider de touristes chinois”, a clear fluency issue.

Also, many students failed to identify errors related to the correctness of the target language. Only 44 students (26%) corrected the use of a subjunctive after *après que* (“après qu’un touriste chinois de 80 ans soit mort du virus dans un hôpital de Paris”) instead of the indicative mood. This is quite a common mistake in French, which explains why DeepL provided such MT output (the use of *après que* followed by the subjunctive must appear in the data used by the MT engine). Additionally, only 22 students (13%) spotted the misuse of punctuation, with unnecessary quotation marks in French, as opposed to English. Finally, only 21 students (12.4%) understood that the present perfect in “the growing economic impact that the coronavirus (...) has had on tourism” would be better translated with a present tense in French since the situation is not over at the moment of writing, a mistake based on the identical morphological structure of what is a present tense in English (present perfect) and a past tense in French (passé composé).

Other types of errors, however, were more easily dealt with by students. Accuracy problems in particular seem to have been easier for them to identify, although a correct solution was not always suggested. For instance, 133 students out of 169 (78.7%) considered that “était à peine plus longue que celle de Paris” (lit. was barely longer than the one in Paris) was a mistranslation for “was barely a line by Paris standards”. However, only 81 of them provided a relevant solution, e.g. “était presque inexistante par rapport à d’habitude” (lit. was almost inexistent compared to the usual) or “ne correspondait pas à ce que l’on voit traditionnellement à Paris” (lit. did not match what is traditionally seen in Paris). Additionally, the translation of “it snakes around the back”, where “back” refers to the back of the store and not the back of a body, required correction for 119 students (70%). However, only 58 of them managed to provide a suitable equivalent. On the other hand, only 18 students (a little more than 1 in 10) noticed that translating “catering to the ever-growing Chinese market” with “qui approvisionnent le marché chinois” was not a fit-for-purpose equivalent since the text mentions Chinese tourists outside China while the expression “approvisionner le marché” would refer to the market in China, a subtler yet problematic accuracy issue.

As far as lexical issues are concerned (wrong term or wrong collocation), results varied importantly. Most students (124, that is 3 out of 4) identified the mistranslation of “line” with “ligne” instead of “queue” or “file d’attente”. However, only 75 (44.4%) considered that the use of “compter sur” to translate “rely on” in the expression “have come to rely on visitors from China” required correction. In the same way, only 68 students (40.2%) decided to modify “preuve éclatante”, which is not a natural collocation in the context, as opposed to “preuve évidente”, “preuve manifeste”, or “preuve indéniable” for example. Additionally, only 41 students corrected “visiteurs” into “touristes”, a more appropriate term in French.

Appendix 3 provides the detailed results for each of the 20 items requiring correction.

### *3.2. Results of the questionnaire*

The follow-up questionnaire allowed us to gain some insight as to students’ use of and trust in MT tools. Unsurprisingly, 83% of students confirmed their use of online MT tools for their homework assignment. This result is clearly in line with the studies mentioned

in the first part of the article, which strengthens the idea that students use MT on a very regular basis, even when they have not received any specific training. One third of them use MT tools sometimes, another third often, while 17% rarely use them and 10% systematically. The most popular tools are DeepL (used by 75.9% of our respondents), Reverso Traduction (32.9%), and Google Translate (29.1%). It was interesting to learn how they use these tools: most of them (38.7%) actually type words, parts of sentences in order to integrate the translation of the isolated parts in their translation. Only 5% of the students copy and paste the whole source text before correcting the translation provided by the MT tool, which is what was required from them in our study, while 10% copy and paste complete sentences separately. 20% of our respondents say that they first translate the text on their own, and then compare their translations with the result of an MT tool. Only 1 student in 3 regularly uses MT tools for personal purposes outside university.

What about their satisfaction? Half of the respondents (52.5%) say that they are often satisfied with the translations provided, with 38.8% being sometimes satisfied and only 7.5% being rarely satisfied. It is interesting to note that only 1 student claimed that s/he was always satisfied. When using MT tools, students are searching for help in relation to lexical choices (60%), sentence organisation (57.7%), meaning of the source text (51.2%), and correctness (spelling, grammar) of the target language (21.3%).

In contradiction with the results obtained for our study, 93% of students think they are able to spot MT errors, including 18.8% without any problem, and only 31.5% found the exercise difficult. However, 65.1% of our respondents disagree with the idea that MT results are sufficiently good and people no longer need to translate themselves.

### *3.3. Discussion*

By comparing students' perception and performance, our study provides thought-provoking information on the capacity of professionals-to-be, most of whom will actually work outside the translation industry, to use MT tools for multilingual communication. While the majority of students think that MT tools cannot replace translators and that they are capable of identifying the errors in the MT output, our results show that students are actually quite far from having developed a sufficient critical use of MT. It turns out that students seem able to identify accuracy errors, but have a tendency to overlook fluency errors, including grammar mistakes and typography issues. Now that the neural version of MT no longer "makes" blatant grammar and spelling mistakes as SMT did, this is clearly an issue. We interpret these results as being the reflection of excessive trust from our students: they expect the information shown on their screens to be as reliable and truthful as the results of a calculator, for instance. As a result, they seem to be heavily primed by the MT output (even in the case of very awkward direct calques), although they are better at spotting accuracy problems. Another possibility, not exclusive of the first one, is a poor command of the target language writing system, in our case French, even though it is their mother tongue.

Our conclusion is therefore clear: students, even in their final year of an undergraduate programme in applied languages, fail to use MT tools effectively. Supplementary, specific

training is therefore necessary for them to improve their MT literacy and their critical use of this technology later in the professional world. Extra training on target language proficiency and writing skills might also prove to be necessary. Their critical analysis of information provided by computers needs to be improved, and students need to be placed in the role of decision makers in order to avoid any priming effect of the MT output.

This naturally raises the question of how MT tools should be introduced in language programmes. Our results cast doubt on the use of MT output as a “bad model”, defined by Yamada (2020: 190) as the use of “errors and mistakes in the raw MT output (...) to highlight language differences between the L1 and the L2 in order to reinforce learners’ understanding of correct grammar and style.” This method is now quite common in L2 teaching and has proved successful for language learning in spite of limits and drawbacks (see Yamada, 2020 for a discussion and a series of references): in particular, “focusing on poor translation containing ungrammatical structures, or ‘negative evidence’, may reinforce unwanted language habits in learners” (Yamada, 2020: 191). It is important to note here that our students had to correct a text written in their mother tongue (L1), not an L2 text, and that our aim was to test their translation skills not their language skills. It is also important to note that most of the references quoted in Yamada (2020) date back to the pre-NMT era (-2016), whereas NMT might mislead students more than SMT because of the focus on fluency. This might mean that the correction, or post-editing to use a more professional term, of a text translated with an NMT tool into L1 might be a case where the use of a bad model to be corrected does not help students improve their translation skills. This requires further investigation, by conducting the same kind of study with a French text translated into English, and is left open for future research. Note however that this hypothesis is in line with the claim that machine-translated texts show errors that are becoming more difficult to identify because they are more human-like.

The contradiction between student’s self-perception and their actual ability to spot and correct MT errors means that there is still room for improvement to “demystify” machine translation tools (Moorkens, 2018: 2). Students still seem to trust what they see on their screens too much, and more sensitization to the limits of the technology is needed as part of the development of relevant MT literacy, and as part of principled machine-human interaction. In order to reach such a goal, a complementary approach could be comparing a human translation and a machine translation of the same text, aimed at showing the added value of human translation and raising students’ awareness of the limits of translating with an online MT tool. Another possibility, which would place students in the role of decision makers, would be to use an online, dynamic interface rather than provide them with a ‘static’ machine-translated text to be corrected. For instance, the DeepL website now provides the possibility to select any word in the machine-translated text in order to visualize more translation equivalents (synonyms generally). Users then see a kind of drop-down menu, from which they need to make a selection. Using the online website might help students realize that what they see on the screen is only one possibility among others and might help them make better choices. In addition, using the glossary feature introduced by DeepL in 2020 could be an additional

way to introduce them to a more professional use of the tool. Finally, exercises focused on comprehension through paraphrasing of the source text before correcting the MT output could be yet another approach.

As a final note to this discussion, we would like to stress that it is not only students who may be challenged or misled by the quality improvement in MT output. Professionals also find it hard to evaluate the quality of MT output as well as the relevance to use MT for their translation projects. For instance, López Pereira (2019) has shown that there is a discrepancy between professionals' perceptions and reality: presented with translations produced by an SMT and an NMT tool, they selected the NMT output in 70% of cases as being the most "productive", that is, requiring less post-editing. However, even though the post-editing distance is smaller with NMT, the study shows that the time spent on post-editing is actually longer. MT is a challenge for everyone these days, and it is not surprising that language students should find it hard to acquire advanced MT literacy. The more plausible the translated text looks, the harder it is to correct.

### *3.4. Limitations*

Our study has some methodological limitations. Twenty errors were identified as needing correction, and our analysis has focused on how our students treated these twenty errors. This does not mean that they did not modify other parts of the text, sometimes actually adding errors, but these were not analysed. Also, not all students answered the questionnaire, and as the answers were completely anonymous (we did not want students to feel judged on their use of MT), it is impossible to see whether there is a link between their use of MT and their capacity to identify and correct MT errors. It is also impossible to know whether the students who answered the questionnaire were also those who achieved a better performance. Finally, our study relies on the analysis of how the MT output of one single text was corrected. Although we repeated exactly the same test with the new cohort of students in 2020-2021 with roughly similar results, it would be interesting to set up an experiment with a control group, in order to compare students' performance when they correct MT output and when they translate with other translation aids. An exercise on the other translation direction with a French text translated into English, i.e., from students' first language into a foreign language, could also bring complementary interesting results. This is left open for future research.

The focus on the English-French language pair means that our results are not meant to illustrate students' MT literacy for all language pairs: with other languages for which training data is not as available and MT tools not as powerful, less deceptive MT results might be easier to correct for students.

### **Conclusion**

In this article, we have reported on an exploratory study on the use of machine translation by advanced students in applied languages. The aim was to assess their MT literacy as defined by Bowker & Buitrago-Ciro (2019), in particular their capacity to identify and/or correct MT errors. Our results show that our students, who are regular users of online

MT tools but have received no specific training in professional post-editing, failed to correct a machine-translated text by removing all adequacy and fluency errors. While our students did manage to identify accuracy errors more easily than fluency errors, they were not always able to provide relevant corrections. In particular, students tend to overlook fluency errors, including grammar mistakes in the target language, which could be the result of several factors: excessive trust in what the computer provides, the influence of the MT output preventing them from considering alternative solutions (priming), or a lack of proficiency in the target language, in this instance the students' mother tongue.

This means that the approach to MT tools in our translation classes should be refined. Several possibilities have been mentioned in the discussion section, for example using the online dynamic interface of an MT tool rather than providing a static version of the text, or devising a test where students need to compare two translations of the same text, one by a human translator and one by an MT tool. Our aim should however remain the same: demystify MT technology so that students understand that human intervention remains necessary, even for translation purposes outside the translation industry. Since today MT tools are unavoidable in language classes, and not only for translation tasks, we need to find ways to empower students with the necessary skills to use what machines have to offer critically and professionally.

## References

- Bojar, O. *et al.* (2016). Findings of the 2016 conference on machine translation (WMT16). In: *Proceedings of the 1st conference on machine translation*, vol 2. Berlin, Germany, August 2016, p. 131-198. <<https://www.aclweb.org/anthology/W16-2301/>>. [Accessed: 20211105].
- Bowker, L. (2005). Productivity vs quality? A pilot study on the impact of translation memory systems. *Localisation Focus*, v. 4, n. 1, pp. 13-20.
- Bowker, L. (2020). Machine translation literacy instruction for international business students and business English instructors. *Journal of Business & Finance Librarianship*, v. 25, n. 1-2, pp. 25-43. <<https://doi.org/10.1080/08963568.2020.1794739>>. [Accessed: 20211105].
- Bowker, L.; Buitrago-Ciro, J. (2019). *Machine Translation and global research: Towards improved machine translation literacy in the scholarly community*. Bingley: Emerald Publishing.
- Carl, M.; Schaeffer, M. (2017). Why translation is difficult: A corpus-based study of non-literality in post-editing and from-scratch translation. *HERMES: Journal of Language and Communication in Business*, n. 56, pp. 43-57. <<https://doi.org/10.7146/hjlc.v0i56.97201>>. [Accessed: 20211105].
- Castilho, S. *et al.* (2017a). Is neural machine translation the new state of the art? *The Prague Bulletin of Mathematical Linguistics*, n. 108, pp. 109-120. <<https://doi.org/10.1515/pralin-2017-0013>>. [Accessed: 20211105].

- Castilho, S. *et al.* (2017b). A comparative quality evaluation of PBSMT and NMT using professional translators. In: *Proceedings of the Machine Translation Summit XVI*, Nagoya, Japan, September 2017, v. 1, pp. 116–131.  
<[https://www.researchgate.net/publication/320016264\\_A\\_Comparative\\_Quality\\_Evaluation\\_of\\_PBSMT\\_and\\_NMT\\_using\\_Professional\\_Translators](https://www.researchgate.net/publication/320016264_A_Comparative_Quality_Evaluation_of_PBSMT_and_NMT_using_Professional_Translators)>. [Accessed: 20211107].
- Clifford, J. *et al.* (2013). Surveying the landscape: What is the role of machine translation in language learning? *@tic revista d'innovació educativa*, n. 10, pp. 108–121.
- De Clercq, O. *et al.* (2021). Uncovering machine translationese using corpus analysis techniques to distinguish between original and machine-translated French. Special issue of *Translation Quarterly*, n. 101, pp. 21–45.
- Hassan, H. *et al.* (2018). Achieving human parity on automatic Chinese to English news translation. <<https://arxiv.org/abs/1803.05567>>. [Accessed: 20211105].
- Jolley, J. R.; Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors, In: A. J. Moeller (ed.). *Learn Languages, Explore Cultures, Transform Lives* Minneapolis: 2015 Central States Conference on the Teaching of Foreign Languages, pp. 181–200.
- Loock, R. (2016). L'utilisation des corpus électroniques chez le traducteur professionnel: quand? comment? pour quoi faire? *ILCEA*, n. 27.  
<<https://doi.org/10.4000/ilcea.3835>>. [Accessed: 20211105].
- Loock, R. (2019). La plus-value de la biotraduction face à la machine. *Traduire*, n. 241, pp. 54–65. <<https://doi.org/10.4000/traduire.1848>>. [Accessed: 20211105].
- Loock, R. (2020). No more rage against the machine: How the corpus-based identification of machine-translationese can lead to student empowerment. *The Journal of Specialised Translation*, n. 34, pp. 150–170.
- López Pereira, A. (2019). Traducción automática neuronal y traducción automática estadística: percepción y productividad. *Revista Tradumàtica: traducció i tecnologies de la informació i la comunicació*, n. 17, pp. 1–19.  
<<https://doi.org/10.5565/rev/tradumatica.235>>. [Accessed: 20211105].
- Macken, L. *et al.* (2019). NMT's wonderland where people turn into rabbits: A study on the comprehensibility of newly invented words in NMT output. *Computational Linguistics in the Netherlands Journal*, n. 9, pp. 67–80.
- Meyer, I. (1988). The general bilingual dictionary as a working tool in thème. *Meta: Translators' Journal*, v. 33, n. 3, pp. 368–376. <<https://doi.org/10.7202/003645ar>>. [Accessed: 20211105].
- Moorkens, J. (2018). What to expect from Neural Machine Translation: A practical in-class translation evaluation exercise. *The Interpreter and Translator Trainer*, v. 12, n. 4, pp. 375–387. <<https://doi.org/10.1080/1750399X.2018.1501639>>. [Accessed: 20211109].

- Niño, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning Technology*, v. 28.  
<<https://doi.org/10.25304/rlt.v28.2402>>. [Accessed: 20211109].
- O'Brien, S.; Ehrensberger-Dow, M. (2020). MT Literacy: A cognitive view. *Translation, Cognition & Behavior*, v. 3, n. 2, pp 145–164.  
<<https://doi.org/10.1075/tcb.00038.obr>>. [Accessed: 20211105].
- O'Neill, Errol M. (2019). Online translator: dictionary, and search engine use among L2 students. *CALL-EJ: Computer-Assisted Language Learning–Electronic Journal*, v. 20, n. 1, pp. 154–177.
- Resende, N.; Way, A. (2021). Can Google Translate rewire your L2 English processing? *Digital*, v. 1, n. 1, pp. 66–85. <<https://doi.org/10.3390/digital1010006>>. [Accessed: 20211105].
- Shopova, T. (2014). Digital literacy of students and its improvement at the university. *Journal on Efficiency and Responsibility in Education and Science*, v. 7, n. 2, pp. 26–32. <<https://doi.org/10.7160/eriesj.2014.070201>>. [Accessed: 20211103].
- Tsai, S-C. (2019). Using Google Translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, v. 32, n. 5-6, pp. 510–526.  
<<https://doi.org/10.1080/09588221.2018.1527361>>. [Accessed: 20211105].
- Toral, A. *et al.* (2018). Attaining the unattainable? Reassessing claims of human parity in Neural Machine Translation. In: *Proceedings of the 3rd conference on machine translation*, Brussels, Belgium, October 2018, pp. 113–123.  
<<https://www.aclweb.org/anthology/W18-6312/>>,  
<<https://doi.org/https://doi.org/10.18653/v1/W18-6312>> [Accessed: 20211105].
- Vanmassenhove, E. *et al.* (2019). Lost in translation: Loss and decay of linguistic richness in machine translation. In: *Proceedings of Machine Translation Summit XVII v. 1: Research Track*, Dublin, Ireland, August 2019, pp. 222–232.  
<<https://www.aclweb.org/anthology/W19-6622/>>. [Accessed: 20211109].
- White, K. D.; Heidrich, E. (2013). Our policies, their text: German language students' strategies with and beliefs about web-based machine translation. *Die Unterrichtspraxis*, v. 46, n. 2, pp. 230–250. <<https://doi.org/10.1111/tger.10143>>. [Accessed: 20211108].
- Yamada, M. (2014). Can college students be post-editors? An investigation into employing language learners in machine translation plus post-editing. *Machine Translation*, v. 29, n. 1, pp. 49–67. <<https://doi.org/10.1007/s10590-014-9167-7>>. [Accessed: 20211105].
- Yamada, M. (2019). The impact of Google Neural Machine Translation on post-editing by student translators. *The Journal of Specialised Translation*, n. 31, pp. 87–106. <[https://jostrans.org/issue31/art\\_yamada.php](https://jostrans.org/issue31/art_yamada.php)>. [Accessed: 20211105].
- Yamada, M. (2020). Language learners and non-professional translators as users. In: M. O'Hagan (ed). *The Routledge Handbook of Translation Technology*. London:



Routledge, pp. 183–199. <<https://doi.org/10.4324/9781315311258>>. [Accessed: 20211106].

## Appendix 1. English source text and its translation with DeepL into French

<b>Coronavirus Empties European Cities of Chinese Tourists</b>	<b>Le coronavirus vide les villes européennes de touristes chinois</b>
<p>Hotels, stores and destinations across the Continent have come to rely on visitors from China. Quarantine measures have changed all that.</p> <p><i>New York Times</i> Feb 17, 2020</p> <p>PARIS — The line in front of the Louis Vuitton store was barely a line by Paris standards: only 10 people. All were Asian and many spoke in Chinese, with one couple dictating Mandarin into a smartphone and waiting for the answers in French.</p> <p>“Sometimes, the line’s been even shorter recently,” Yasmine Ben, who works at a kiosk directly facing the store, said on a recent morning. “Usually, it’s wider, much, much longer, and it snakes around the back.”</p> <p>Louis Vuitton, in the Galeries Lafayette department store in central Paris, is a favorite stop inside one of the favorite shopping destinations of Chinese tourists to France. And the line there is prime evidence of the growing economic impact that the coronavirus, which broke out in Wuhan, China, late last year, has had on tourism in Paris and elsewhere across Europe.</p> <p>Though it is too soon to quantify it precisely, the potential economic impact of the coronavirus is evident nearly everywhere. Fears were heightened over the weekend after an 80-year-old Chinese tourist died of the virus at a hospital in Paris — the first fatality outside Asia since the start of the outbreak.</p> <p>The effects, especially on businesses catering to the ever-growing Chinese market, have been immediate. Last week, the Italian government considered allocating assistance to hard-hit tour operators.</p> <p>In the most popular destinations, like Paris, already overwhelmed with tourists, the overall economic impact has been limited so far, officials said. The Louvre Museum, which Chinese were second only to Americans in visiting last year, has not suffered a decrease in visitors since the start of the outbreak, said a spokeswoman, Sophie Grange.</p> <p>The Chinese account for 3 percent of visitors to Paris — about 800,000 visitors a year, compared</p>	<p>Les hôtels, les magasins et les destinations de tout le continent en sont venus à compter sur les visiteurs de Chine. Les mesures de quarantaine ont changé tout cela.</p> <p><i>New York Times</i> 17 février 2020</p> <p>PARIS - La file d'attente devant le magasin Louis Vuitton était à peine plus longue que celle de Paris : seulement 10 personnes. Tous étaient asiatiques et beaucoup parlaient en chinois. Un couple dictait le mandarin dans un smartphone et attendait les réponses en français.</p> <p>"Parfois, la queue a été encore plus courte récemment", a déclaré Yasmine Ben, qui travaille dans un kiosque situé directement en face du magasin, un matin récent. "D'habitude, elle est plus large, beaucoup, beaucoup plus longue, et elle serpente dans le dos.</p> <p>Louis Vuitton, dans le grand magasin des Galeries Lafayette au centre de Paris, est une étape incontournable dans l'une des destinations de shopping préférées des touristes chinois en France. Et la ligne qui y est tracée est une preuve éclatante de l'impact économique croissant que le coronavirus, qui s'est déclaré à Wuhan, en Chine, à la fin de l'année dernière, a eu sur le tourisme à Paris et ailleurs en Europe.</p> <p>Bien qu'il soit trop tôt pour le quantifier avec précision, l'impact économique potentiel du coronavirus est évident presque partout. Les craintes se sont accrues au cours du week-end après qu'un touriste chinois de 80 ans soit mort du virus dans un hôpital de Paris - le premier décès hors d'Asie depuis le début de l'épidémie.</p> <p>Les effets, en particulier sur les entreprises qui approvisionnent le marché chinois en pleine expansion, ont été immédiats. La semaine dernière, le gouvernement italien a envisagé d'allouer une aide aux voyageurs durement touchés.</p> <p>Dans les destinations les plus populaires, comme Paris, déjà submergée de touristes, l'impact économique global a été limité jusqu'à présent, ont déclaré les responsables. Le musée du Louvre, que</p>

with 2.4 million from the United States, according to the city's Convention and Visitors Bureau.

China represents "a small fraction of our tourism economy," said Corinne Menegaux, the bureau's director. "If it were the United States, it would be something else."

Still, Ms. Menegaux said that because Chinese tourists often traveled in groups and tended to use specific hotels and stores, some businesses have been disproportionately affected. Sales at some duty-free stores and other businesses, where Chinese account for 80 percent of the clientele, have plummeted, she said.

les Chinois ont visité en deuxième position après les Américains l'année dernière, n'a pas subi de baisse de fréquentation depuis le début de l'épidémie, a déclaré une porte-parole, Sophie Grange.

Les Chinois représentent 3 % des visiteurs de Paris - environ 800 000 visiteurs par an, contre 2,4 millions pour les États-Unis, selon l'Office du tourisme et des congrès de la ville.

La Chine représente "une petite fraction de notre économie touristique", a déclaré Corinne Menegaux, directrice de l'office. "Si c'était les États-Unis, ce serait autre chose".

Néanmoins, Mme Menegaux a déclaré que les touristes chinois voyageant souvent en groupe et ayant tendance à utiliser des hôtels et des magasins spécifiques, certaines entreprises ont été touchées de manière disproportionnée. Les ventes de certains magasins hors taxes et autres commerces, où les Chinois représentent 80 % de la clientèle, ont chuté, a-t-elle dit.

Traduit avec [www.DeepL.com/Translator](http://www.DeepL.com/Translator) (version gratuite)

## Appendix 2. Follow-up questionnaire

1. For your homework preparation for the English-French translation course, do you use machine translation (that is to say internet websites that provide you with translations like Google Translate or DeepL, NOT online dictionaries like Linguee). [If the answer is no, then you can directly go to the last two questions]
  - yes
  - no
2. If the answer is yes, would you say that you use MT:
  - on a regular basis
  - sometimes
  - often
  - systematically
3. Which website(s) do you visit (several answers possible)?
  - Google Translate
  - DeepL
  - Reverso Traduction
  - Systran Translate
  - Other
4. How do you generally use machine translation (only 1 answer, select your most frequent use)?
  - I copy/paste the whole text, and then I 'correct' the translation
  - I copy paste whole sentences, and then I 'correct' the translation
  - I type parts of sentences, or even words, and then I insert the suggestions in my translation
  - I copy/paste the whole text or parts of the text to have a general idea of the

- meaning of the text, and then I translate myself
- I first translate on my own, and then I compare my translation with the translation suggested by the online tool(s)
  - Other
5. I am happy with what MT tools provide:
- always
  - often
  - sometimes
  - rarely
  - never
6. Machine translation helps you to:
- make lexical choices (vocabulary)
  - improve the structure of sentences
  - improve the quality of your French (check spelling and grammar)
  - check the meaning of the text
  - other
7. To me, machine translation is:
- a very useful tool, thanks to which I do not need to translate
  - a tool that is useful to understand the text, but useless for translation
  - a translation tool, but I have the feeling I don't always know how to use it properly
  - a tool that helps me improve my translations and that I use with confidence
  - other
8. I have the feeling I can identify machine translation errors:
- yes, no problem at all
  - yes, quite
  - not really
  - not all all
  - machines do not make errors, so I trust the results
9. According to me, machine translation tools provide results that are sufficiently good and people no longer need to translate themselves:
- I strongly agree
  - I agree
  - I only partially agree
  - I don't really agree
  - I strongly disagree
10. I found the online test on April 3rd 2020:
- very easy
  - quite easy
  - neither easy nor difficult
  - quite difficult
  - very difficult
  - I don't know
11. Outside university, I use machine translation tools:
- never
  - rarely (several times a month)

- regularly (several times a week)
- often (several times a day)

### Appendix 3. Detailed results

	MT error	Number of students who identified and corrected the MT error correctly		Number of students who identified the MT error correctly but did not manage to provide a relevant correction	
1	vide les villes européennes de touristes chinois	90	53.25%	19	11.24%
2	compter sur les visiteurs	64	37.87%	11	6.51%
3	visiteurs	40	23.67%	1	0.59%
4	était à peine plus longue que celle de Paris	81	47.93%	52	30.77%
5	Tous	22	13.02%	3	1.78%
6	dictait le mandarin	18	10.65%	36	21.30%
7	parfois la queue a été encore plus courte récemment	81	47.93%	14	8.28%
8	"D'habitude..."	22	13.02%	0	0.00%
9	un matin récent	30	17.75%	47	27.81%
10	dans le dos	58	34.32%	61	36.09%
11	une étape incontournable dans l'une des destinations	37	21.89%	15	8.88%
12	la ligne qui y est tracée	105	62.13%	19	11.24%
13	preuve éclatante	47	27.81%	21	12.43%
14	a eu	21	12.43%	0	0.00%
15	après qu'un touriste... soit mort	40	23.67%	4	2.37%
16	qui approvisionnent le marché chinois	9	5.33%	9	5.33%
17	dans les destinations	16	9.47%	10	5.92%

18	l'impact a été limité jusqu'à présent	19	11.24%	10	5.92%
19	que les Chinois ont visité en deuxième position après les Américains	53	31.36%	53	31.36%
20	a déclaré	41	24.26%	2	1.18%