Multilinguisme européen et IA entre droit, traduction et didactique des langues

Multilinguismo europeo e IA tra diritto, traduzione e didattica delle lingue

European Multilingualism and Artificial Intelligence: The Impacts on Law, Translation and Language Teaching

Édité par, a cura di, edited by

Rachele Raus, Università di Bologna
Francesca Bisiani, Université Catholique de Lille
Maria Margherita Mattioda, Università di Torino
Michela Tonti, Università di Bergamo

Multilinguisme européen et IA entre droit, traduction et didactique des langues

Multilinguismo europeo e IA tra diritto, traduzione e didattica delle lingue

European Multilingualism and Artificial Intelligence: The Impacts on Law, Translation and Language Teaching

Édité par, a cura di, edited by

Rachele Raus, *Università di Bologna*
Francesca Bisiani, *Université Catholique de Lille*
Maria Margherita Mattioda, *Università di Torino*
Michela Tonti, *Università di Bergamo*

**Artificial Intelligence
for European Integration**
Jean Monnet Centre of Excellence

C P S
CULTURE
POLITICA
SOCIETÀ

In cooperation whit:

TO EU
Centro Studi sull'Europa

Co-funded by
the European Union

Graphics and page layout
Silvio Ortolani, SISHO - fotografia & archivi

Ledizioni
The Innovative LEDIpublishing Company

Deuxième partie : expérimentations pédagogiques

Promoting machine translation literacy: A focus on gender mistranslations and bias in English-Italian NMT

Alessandra Molino

Introduction

This article reports on a survey on Artificial Intelligence (AI) and Machine Translation (MT) conducted among graduate-level learners of English at the University of Turin, Italy, within the framework of the project *Linguistic rights and language varieties of Europe in the age of artificial intelligence*[1]. The participants attended an English-Italian MT module (18 hours, a.y. 2021-2022) aimed at illustrating the functioning of MT, its current affordances and limitations, and its impact on issues such as multilingualism, language variation, and gender representation. The survey consisted of two questionnaires administered at the beginning (Q1) and at the end (Q2) of the module; the objective was to gather information about the students' background, their knowledge of AI tools and their use of MT. Thanks to the inclusion of comparable questions in the two questionnaires, it is possible to assess whether the students' opinions and behaviour shifted after the teaching activities.

The learning outcomes set for the module aimed at the acquisition of elements of *MT literacy* (Bowker and Ciro 2019; O'Brien and Ehrensberger-Dow 2020), namely "knowing how MT works, how it can be useful in a particular context, and what the implications are of using MT for specific communicative needs" (O'Brien and Ehrensberger-Dow 2020: 145). For the students attending the module, that meant gaining an initial understanding of MT "so they can use it intelligently and avoid common pitfalls" (Kenny 2022: iv). The necessity to employ MT in a conscious and critical way is particularly important nowadays when Neural Machine Translation

Alessandra Molino, Università di Torino, alessandra.molino@unito.it

(NMT) has significantly improved output fluency compared with Statistical Machine Translation (SMT). As Guerberof Arenas and Moorkens (2019: 231) remind us, "this fluency can be deceiving" because the content may be wrong or misleading, and the target text may contain forms of bias resulting from the training data.

The students taking part in the survey were enrolled in the second year of the MA degree in either International Communication or International Communication for Tourism, two programmes offered in the Department of Foreign Languages, Literature and Modern Cultures. For this category of users, MT represents one of the digital resources for language learning. Considering their future profession in the world of corporations, international organizations, or tourism, NMT will also be for many of them an instrument to overcome the barriers of multilingual communication, gather information, and write more effective texts in their second language. These students are not trained to become professional translators, but they will likely perform a range of translation tasks for both internal and external communication in their future workplace. Acquiring MT literacy is therefore an important component of their preparation, which has been given little attention so far.

The questionnaires were structured according to best practices in survey research in Applied Linguistics (see Dörnyei 2007). They comprised factual questions aiming to gather information about the students' profile; behavioural questions used to understand what the students knew about AI and whether they regularly used AI-based systems; and attitudinal questions, which solicited opinions on the advantages and disadvantages of AI and MT. During the module, special attention was devoted to mistranslations involving wrong gender coreference and gender bias. Therefore, this paper focuses particularly on the part of the questionnaires related to these topics.

The article is structured as follows. Section 1 presents the students' profile. Section 2 illustrates the questionnaire results in relation to the participants' declared knowledge and use of AI tools. Section 3 concentrates on MT and reports on the students' opinions of its reliability and their experiences in using MT systems. Section 4 focuses on gender mistranslations and gender bias, first describing the teaching activities carried out in class and then examining the Q2 responses related to these topics. Section 5 discusses the results and concludes the paper with some recommendations for effective MT literacy instruction.

1. The students' profile

The module on MT was offered to a large class of students: Q1 obtained 197 responses, whereas Q2 was answered by 132 participants[2]. While the difference between the two questionnaires' response rates is not insignificant (33% fewer participants in Q2), the sample remains large enough to notice trends and be considered quite representative of the target group, the Q2 respondents amounting to 67% of the students participating in the module's activities.

Most of the answers came from female students who constituted 85% of the participants[3]. The age group of most students ranged from 23 to 25 years old. The mother tongue was Italian for 92% of the respondents, while 5% of them reported being bilingual but having Italian as one of the two languages. The rest of the group (3%) declared to speak other first languages, including Romanian, Albanian, Russian and, in the case of Erasmus students, other European languages such as French and Spanish.

In addition to their mother tongue, the participants spoke a minimum of two foreign languages although most of them reported three. Figure 1 shows the four foreign languages most often cited by the students. Almost the totality of the participants stated that they could speak English (98%). Spanish and French followed, mentioned by 62% and 58% of the participants respectively. German was known by 30% of the respondents. Other languages mentioned were, in order of frequency, Russian, Portuguese, Japanese, and Arabic.

The students were asked to indicate their perceived competence in their foreign languages. Table 1 displays their self-reported level of English, i.e., the source language in the MT module under scrutiny in this article. The vast majority reported good or excellent knowledge
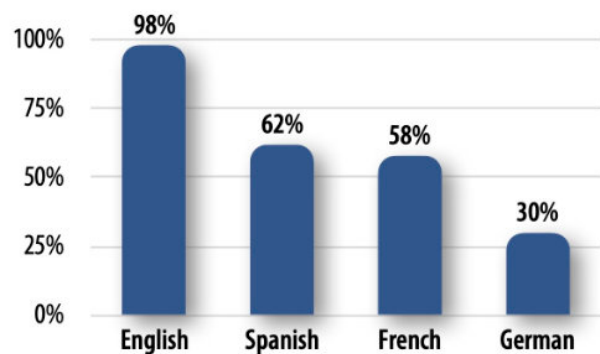
Figure 1: Main foreign languages spoken

[2] This diminished response rate may be related to the fact that attendance was not mandatory and that the number of attendees may decrease towards the end of the semester. In addition, the sanitary situation connected with the Covid-19 pandemic (the module was held in late November 2021) might also have affected participation in the survey.

[3] Male students amounted to 14% and the remaining 1% declared not to identify as either gender. These data presented in this section are the average of the answers obtained in the two questionnaires.

of English. The abilities in which the students felt more confident are the receptive ones — reading and listening — with more than 60% of the participants declaring to possess excellent skills. On the other hand, regarding productive abilities — writing and speaking — the group mainly stated to have good competences.

	Reading	Listening	Writing	Speaking
Excellent	68%	61%	34%	32%
Good	31%	36%	59%	55%
Average	2%	3%	7%	13%
Low	0	0	0	0

Table 1: Self-reported ability in the four language skills in English

Considering that most participants had a previous BA degree in Foreign Languages and Literature and that their declared English language proficiency was overall good or excellent, it may be assumed that these students possessed adequate metalinguistic awareness and sensitivity to language use. This may be regarded as a positive condition for effective instruction in MT (Carré *et alii* 2022: 193). By contrast, the target group was generally lacking previous formal training in the use of MT tools although, as the following sections show, AI-based instruments and NMT systems such as Google Translate (GT) or DeepL (DL) were regularly employed by the students. Hence, the module aimed to engage participants in activities that would encourage reflection on the output received and the performance of the specific tools adopted in a range of different tasks and with a variety of text typologies (i.e., press releases, news articles, and tourist brochures).

2. Knowledge and use of AI

The answers discussed in this section are those provided to the behavioural questions aiming at acquiring information about the students' interest in AI and the tools that they normally used. Overall, the respondents were interested in the development of AI, with 48% manifesting attention and 44% expressing a strong interest. They were then asked to specify what digital tools they used in their personal life and for their studies (Fig. 2). Students could cite multiple instruments, as in the following answer: "social networks, mobile apps, online translators". For quantification purposes, each item on a list was counted individually and then grouped into broader categories. For instance, "Instagram", "social networks", and "Face-

book" were all included in the "Social networks" category. The groupings reported in Figure 2 are those that had at least five mentions.

Tool	Value
social networks (e.g., Instagram, Twitter, Facebook)	172
MT systems, CAT tools, bi/multilingual platforms	112
mobile apps (e.g., WhatsApp, internet banking apps)	108
word processing, presentation and spreadsheet programmes	80
online dictionaries and encyclopedias	39
entertainment platforms (e.g. Netflix, Disney Plus, Prime Video)	34
online conference programmes	17
graphics and photo-editing software (e.g., Photoshop)	14
university apps and e-learning platforms (e.g., Moodle)	12
corpora	9
email clients	6
computer-assisted writing tools	5
voice assistants (e.g., Alexa, Siri)	5

Figure 2: Digital tools most often used by the students

The results outline key identity features of the students under analysis. The most often cited digital instruments are social networks (172 occurrences), but the participants also frequently mentioned mobile applications (108). These are predictable results considering the pervasiveness of these programmes and the young age of the participants. What is more interesting is the participants' evident positioning as university students and foreign language learners. The second most recurrent category is indeed that of MT systems, CAT tools, and bi/multilingual platforms (112). Then, in order of frequency, they cited word processing, presentation, and spreadsheet software (80), online dictionaries and encyclopedias (39), corpora (9), and computer-assisted writing tools (5). The students also mentioned online conferencing programmes (17) and e-learning platforms (12), two instruments that are clearly connected to the pandemic situation which made the use of distance learning devices indispensable.

Comparing what the students declared in Q1 and Q2 about the perceived contribution of AI to the development of their field of study, the percentage of those who believed that AI would be very useful remained stable at 94%. Similarly, they also regarded AI as playing an important role in their future profession, with an approximately similar distribution of answers in both questionnaires (Fig. 3).
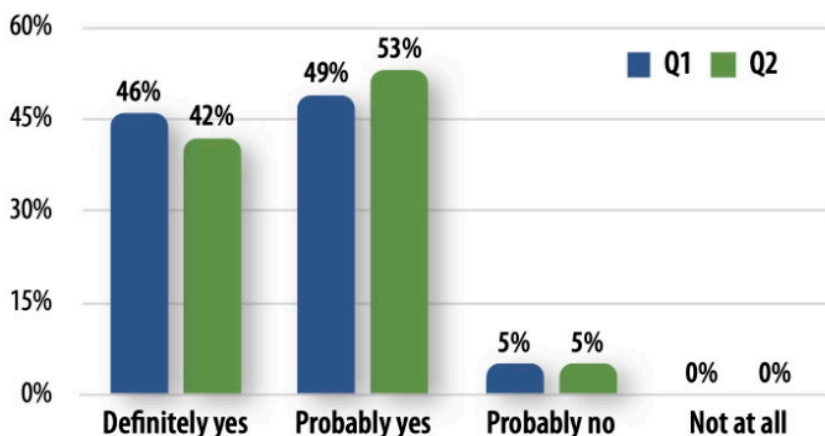
Figure 3: Do you think AI will play a major role in your future profession?

Although in Q2 the students seemed slightly less convinced ('Definitely yes' = 42%; 'Mostly yes' = 53%), overall, the positive tendency remained unaltered, and the picture obtained is that of language learners and trainee language professionals who expect AI to be important if not essential in their education and career development. Therefore, it becomes imperative for degree courses in language-related disciplines to include training in AI tools for natural language processing in such a way as to "'[empower]' students to have a confident interaction with this technology" (Guerberof Arenas, Moorkens 2019: 220).

3. Machine translation: Opinions and experiences

In this section, the focus is on the use of AI tools by the students under investigation. The attitudinal part of the questionnaire was quite diversified, including questions on voice messages, chatbots, and the machine translation of legal terminology. Since these aspects were not discussed in the module, only the answers directly related to the content covered are reported here. The questions on the MT translation of gender will be dealt with separately, in Section 4, because special attention was paid to this topic, hence it deserves more focused treatment.

To begin with, in Q1 the students were asked to declare whether they use MT tools and how often. The answers indicate a strong positive tendency toward their utilisation, with 28% of the respondents recurring to them often and 66% employing MT systems on a daily basis. The students could also specify what instrument they typically utilised. The most often mentioned tool was Reverso[4] (49 occurrences), followed by

---

[4] It is not possible to ascertain whether, by writing 'Reverso', the students meant Reverso Translation or Reverso Context, thus confusing the MT tool and the bilingual platform. Student uncertainty about what counts as MT has been reported in previous studies (e.g., Resende and Way 2021: 72).

DeepL (24) and, quite surprisingly in the third position, Google Translate (11). When asked about their perceived reliability of MT tools, in both questionnaires approximately 80% of the participants declared that MT output is 'quite reliable'[5]. Similarly, the students' answers to the question "When MT output does not seem correct, what might be the reason for the mistake?" remained rather stable: before and after the module, most of the students believed that the cause of mistranslations was the idiomaticity or complexity of the source language expression although a small increase in the percentage of responses indicating corpora or algorithms in Q2 can be noticed (Fig. 4).
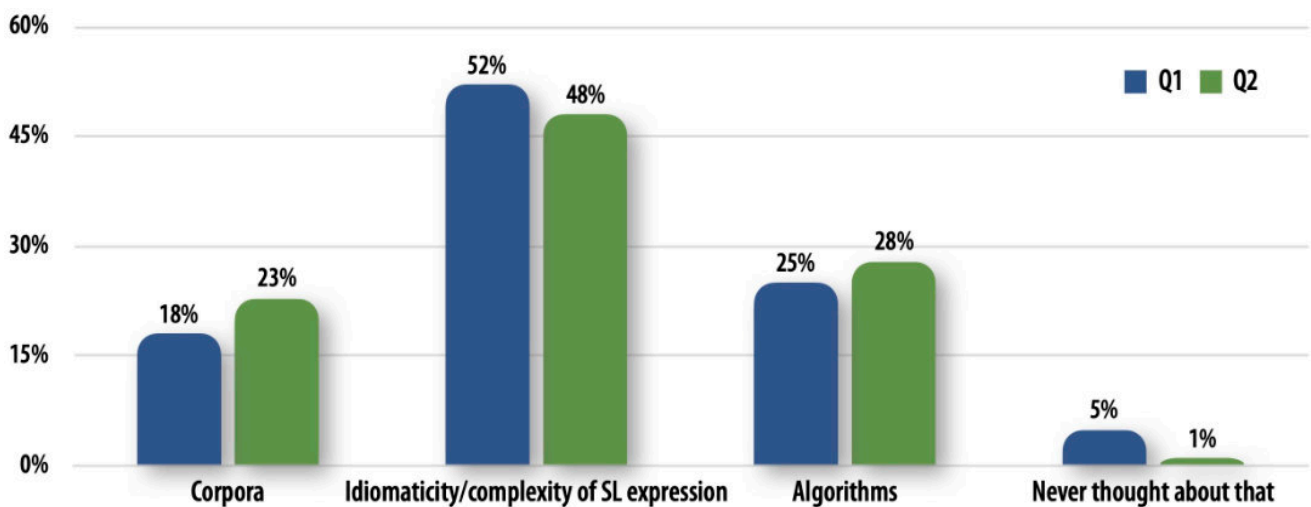
Figure 4: When MT output does not seem correct, what might be the reason for the mistake?

Moving to bi/multilingual platforms (e.g., Reverso Context, Linguee), in Q1 the participants were asked whether they employed such tools for lexical and terminological searches. Again, answers were overall very positive: 45% of the respondents chose the option 'Definitely yes' and 41% selected 'Mostly yes'. In relation to these devices, Figure 5 reveals a quite marked shift in the students' approach to the output after the module. While in Q1 only 19% of the students stated that they always verified the source of the proposed translation equivalent, in Q2 the percentage increased to 45%. On the other hand, the amount of those who never checked the source dropped from 31% to 8%.

---

[5] The other answers were distributed as follows: 'very reliable' (Q1=4%; Q2=0%), 'little reliable' (Q1=15%; Q2=19%), 'not reliable at all' (Q1=0%; Q2=1%), 'I don't know' (Q1=0%; Q2=0%).
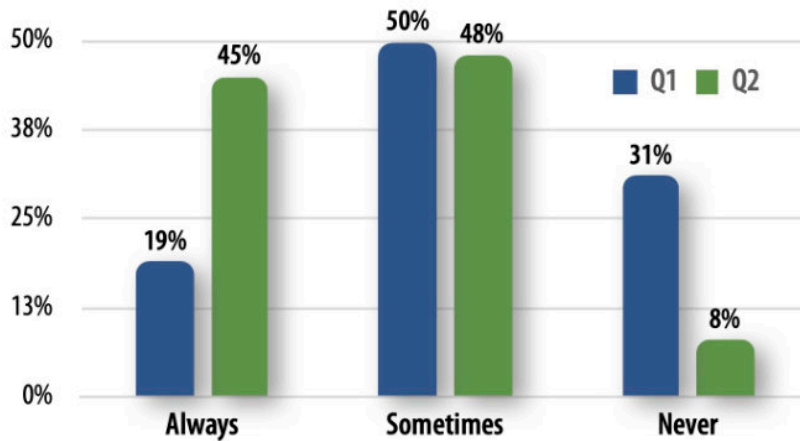
Figure 5: Do you check the source from which the result was extracted?

The fact that the overall perception of reliability of MT systems remained quite high after the teaching activities is not entirely surprising. This result is in line with both subjective and automatic evaluations of NMT, documenting ever-increasing degrees of output fluency (for an overview of relevant studies, see Moorkens 2018). On the other hand, the students' critical attitude toward the use of bi/multilingual platforms increased noticeably. These instruments were no longer considered unquestionable resources, but tools whose results need to be assessed every time according to the text and its communicative purpose. Indeed, when asked about the criterion adopted for the choice of a translation equivalent, as many as 96% of the participants responded that they take the context of use into account before selecting an equivalent. Despite the impression that MT systems offer quite dependable results, especially with general-purpose lexis, the students were also aware that some areas of language still pose challenges to (English-Italian) NMT, such as specialised termin-
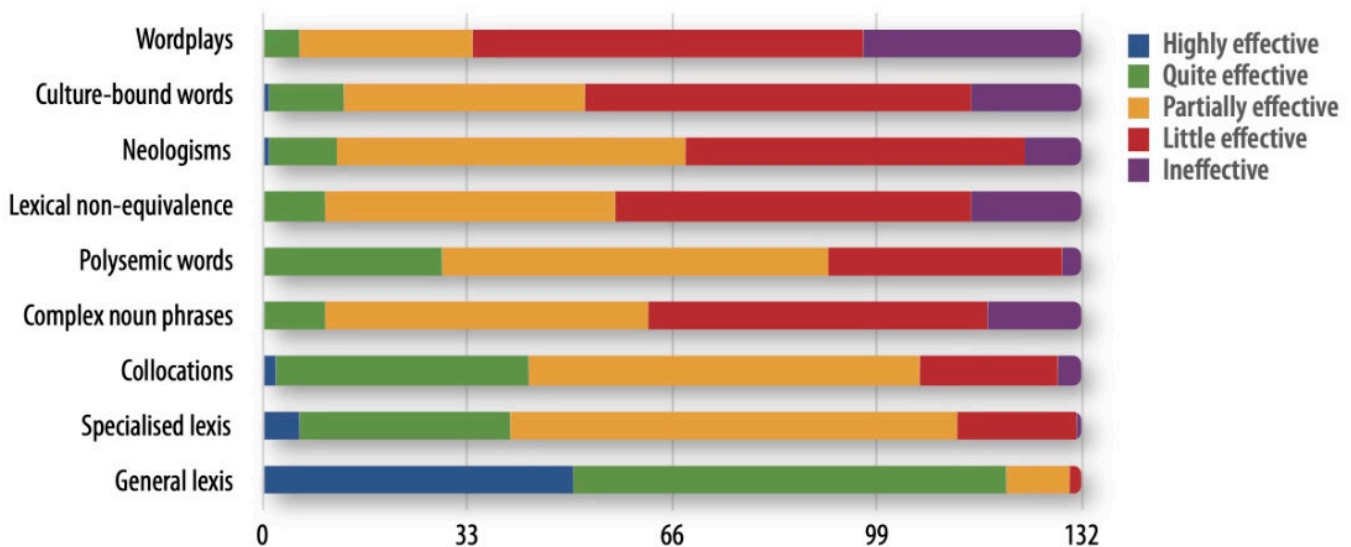
Figure 6: After attending the course, how effective do you think MT is in dealing with these phenomena?

ology, neologisms, word plays, collocations, complex noun phrases, lexical non-equivalence and culture-bound words (Fig. 6). In addition, as will be shown in Section 4, the students' consciousness of the problems connected with the translation of gender was heightened by the course.

4. Gender mistranslations and gender bias in NMT

To better understand the Q2 results related to the NMT of gender, Section 4.1. describes the activities conducted in class to raise the students' awareness of gender coreference mistakes and gender bias. Section 4.2 illustrates the Q2 answers referred to these topics.

4.1. Classroom experiences

The topic of gender representation in MT was approached through a combination of deductive and inductive teaching methods[6]. Lessons comprised class or group discussions about the students' own experiences with MT, their post-editing choices, and controversial topics such as the manifestation of sexism in language. Practical translation tasks included two main types of activities: 1) controlled language exercises (see Doherty, Moorkens 2013; Guerberof Arenas, Moorkens 2019) involving the translation of single words, full sentences, or short passages; and 2) and the translation and post-editing of full-length, authentic texts, followed by an oral or written commentary on the translation process and product.

The activities started by encouraging students to describe and recognise the morphological differences between English and Italian in terms of gender marking. Considering their background knowledge (see Section 1), these students did not need specific instruction about such differences but inductive stimuli to activate their metalinguistic knowledge and identify translation problems. One task required them to type a range of simple English sentences in different free online NMT systems, namely GT, DL, and Reverso Translation (RT), and examine the output in Italian. The students were asked to describe how gender is encoded in the two languages at the morphological level and what problems may arise in English-Italian NMT in general, and with the specific MT system chosen in

---

[6] With deductive pedagogy, students learn by first familiarising themselves with theoretical concepts and principles, and by applying them in tasks or assignments; with inductive pedagogy, students are presented with a set of observations to explain and from which they infer rules and theory.

particular. For instance, the sentence *Let me introduce a long-standing friend of mine. Her name is Sarah* was translated into Italian by GT (2021) as *Permettetemi di presentarvi un mio amico di vecchia data. Il suo nome è Sara*. Here, GT could not create the right coreference between *friend* and the pronoun *her*, thus opting for the masculine *amico* (rather than *amica*).

Building on the observations made by the students on multiple examples and programmes, relevant theoretical notions were formally introduced to consolidate the students' understanding and equip them with the terminology needed. Hence, the distinction between natural gender languages (English) and grammatical gender ones (Italian) was made, according to which pronouns are the main forms of gender marking in the former, while it is word-endings in nouns and dependent parts-of-speech that indicate gender in the latter. The NMT mistakes observed in English-Italian translations were described, considering not only their word-level implications, such as the introduction of referential mistakes, but also the textual level ones, such as the interruption of cohesive chains.

Different sentence structures were tested, using hypotaxis and parataxis. According to Guerberof Arenas and Moorkens (2019: 224), "controlled and pre-editing exercises [...] where the students change the source language using a set of controlled language rules to see how the engine behaves with each change" are of great value as they help trainees "learn how to spot error patterns [...] and anticipate the errors produced by a given engine by just looking at the source text, thereby allowing them to work faster". By testing different sentence patterns, the students

	EN	GT (2021)	DL (2021)
1	*The writer finished the first draft. Then she started proofreading.*	**Lo scrittore ha terminato la prima bozza. Poi ha iniziato a correggere le bozze.*	*La scrittrice ha terminato la prima stesura. Poi ha iniziato a correggere le bozze.*
2	*The writer finished the first draft, and then she started proofreading.*	*La scrittrice ha terminato la prima bozza e poi ha iniziato a correggere le bozze.*	*La scrittrice ha terminato la prima stesura e ha iniziato a correggere le bozze.*
3	*The writer who started proofreading had finished her first draft.*	*La scrittrice che ha iniziato a correggere le bozze aveva terminato la sua prima bozza.*	*La scrittrice che ha iniziato a correggere le bozze ha terminato la sua prima stesura.*
4	*The writer started proofreading because she finished her first draft.*	*La scrittrice ha iniziato a correggere le bozze perché ha terminato la sua prima bozza.*	*La scrittrice ha iniziato a correggere le bozze perché ha finito la prima stesura.*
5	*The writer started proofreading. The reason was that she had finished her first draft.*	**Lo scrittore iniziò a correggere le bozze. Il motivo era che aveva finito la sua prima bozza.*	*La scrittrice iniziò a correggere le bozze. Il motivo era che aveva terminato la prima stesura.*

Table 2: The translation of natural gender words into grammatical gender ones in different syntactic patterns

could observe that syntax plays a role in helping NMT systems create correct cohesive connections between words of the same grammatical gender. For instance, in the examples in Table 2, GT succeeded in creating the right feminine coreference only when the two sentences were joined through coordination or subordination (lines 2, 3, 4). On the other hand, DL provided correct output regardless of the sentence structure. This result was noticed in several different examples, suggesting that DL (2021) possessed a more efficient coreference resolution system for the English-Italian language pair and direction.

Thanks to this activity the students were trained to recognize some of the problems that may arise in the translation of full texts, and they gained awareness of possible strategies to pre-edit a text for NMT. In addition, by comparing the output produced by distinct systems, students learned that the choice of the programme may have consequences in terms of the amount of post-editing required, a factor affecting speed and productivity.

Other activities were aimed at highlighting the risks of gender bias, i.e., the replication of stereotyped views of gender roles in society due to the sources employed for the training of NMT tools. These sources are large corpora or multilingual translation memories, often based on the web (e.g., *Wikipedia*) or made available by international organizations (e.g., EU institutions), which reflect the way society talks and writes about men and women. Hence, by learning language rules and combinations from these datasets, NMT systems inevitably reproduce already existing forms of gender bias.

Students needed to be sensitised to the manifestation and implications of sexism in language and the risks posed by the unaware use of MT. To this end, both theoretical and hands-on classes were held. Among the practical activities, a list of job titles in English was given to the students to type in different NMT systems. Since job titles in English tend not to be marked for gender, it was interesting to examine whether masculine defaults were offered as equivalents in Italian; whether feminine forms were chosen, instead, for typically female professions (e.g., *nurse*, see Prates *et alii* 2019); or whether both masculine and feminine equivalents were provided in parallel, thus assigning them the same importance. Table 3 reports some of the titles analysed and their translation.

EN	IT (GT 2021)	IT (DL 2021)	IT (RT 2021)
Psychologist	Psicologa Psicologo	Psicologo	Psicologo Psicologa
Surgeon	Chirurga Chirurgo	Chirurgo	Chirurgo
Chief Executive Officer (CEO)	Amministratore delegato (CEO)	Amministratore delegato (CEO)/Direttore generale (CEO)	Amministratore Delegato (CEO)/Chief Executive Officer (CEO)
Member of parliament	Membro del parlamento/deputato	Membro del parlamento/deputato	Membro del parlamento/deputato/parlamentare
Nurse	Infermiera Infermiere	Infermiera	Infermiera Infermiere

Table 3: Example of professional titles in English and their translation in Italian by GT (2021), DL (2021), and RT (2021)

The students could observe the adoption of different corporate policies across NMT systems. For example, GT (2021) and RT (2021) tended to offer female and male equivalents for single-word queries. By contrast, DL (2021) mostly reported male default translations providing the female version in a 'Dictionary' section which users could access by scrolling down the output page. On the other hand, the students could also notice similarities across the systems: titles for apical jobs, especially in strategic decision-making sectors, tended to be presented only in the masculine form (see *CEO* and *Member of parliament*) even though the Italian language possesses feminine equivalents (*amministratrice delegata, direttrice, deputata*).

These controlled language exercises were preparatory to the application of NMT to full-length texts and their post-editing. The module focused particularly on press releases, newspaper articles, and tourist brochures, three genres that were chosen in accordance with the interests of the degree courses in which the module was offered (International Communication and International Communication for Tourism). The topics selected for press releases and news articles mainly revolved around the role of women in corporate and political decision-making. Students were required not only to post-edit the NMT output but also to produce a commentary on the translation mistakes identified and the strategies used to arrive at a satisfactory target text (see Baker 2018; Laviosa-Braithwaite 2001). This activity was conducted partly in class in small groups and partly at home as homework. The work was then discussed orally by the students, and feedback was given to them on both the translation and the commentary.

At the end of the module, the students were required to repeat the exercise individually, and the result was assessed as part of their coursework[7].

Examining the marks obtained by the students, it can be affirmed that the majority of them produced generally good work although some errors, unnoticed phenomena, improvable translation choices, and knowledge gaps were present (50% = grade C on the ECTS scale); approximately one-third of the students obtained a very good evaluation producing work with only minor inaccuracies or mistakes (31% = grade B); and 4% of the students were assigned an excellent or outstanding evaluation (grade A)[8]. The sum of these percentages indicates that 85% of the students could perform the task well enough according to the assessment criteria illustrated in class. Therefore, the level of student achievement was overall satisfactory. Examining the data collected in Q2 and described in the following section, it is possible to further assess learning outcomes and the degree of student awareness of gender issues in NMT.

4.2. Impact of the teaching activities

The end-of-course questionnaire contained a series of questions entirely devoted to the topics of gender variation and bias in NMT. The first question (Fig. 7) required the students to indicate whether, in their view, NMT output took gender variation into account. The sweeping majority stated that this feature was not sufficiently catered for in NMT, with 68% of the respondents answering 'Mostly not' and 17% 'Definitely not' (Fig. 7).
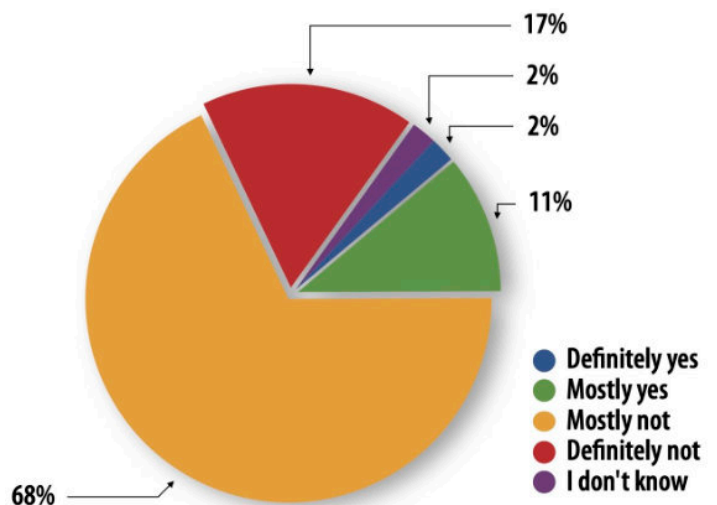
Figure 7: Do you think that MT systems take gender variation (male, female) into account?

---

[7] It should be pointed out that the students were expected to focus not only on gender bias and gender coreference mistakes, but also on other aspects dealt with during the module, such as issues of word order, collocations, forms of address, degree of formality, idiomatic expressions, anglicisms in Italian, product names, and unnecessary omissions.
[8] The rest of the marks were distributed as follows: 13% = grade D (fair work but with significant shortcomings) and 2% grade E (the performance just meets the minimum criteria).

The subsequent question aimed to better understand the students' opinions and awareness, by asking them to recognize the potential reasons for this situation. This question was not mandatory, as it was only relevant to those who had answered negatively or partially so to the previous one. The students could choose among four pre-given alternatives, selecting more than one, and could also write additional answers by flagging the 'other' option. With 84% of the mentions, the inability of NMT systems to effectively resolve gender coreference tasks was the most chosen answer. The second most frequent reason (31% of all mentions) was the specific characteristics of the language pair in question, which make the translation of gender challenging (of course many students had the English-Italian language pair in mind). The other possible answers were chosen by a minority ("the masculine is used as a 'neutral' form comprising the feminine" = 9%; "I don't know" = 2%). The students providing other options referred to the reproduction of stereotypes and prejudices in NMT output (3%) and the insufficient degree of gender variation in the corpora used for machine learning (1%). Overall, the answers obtained are encouraging and indicate that the course helped students gain an initial understanding of the current limits of NMT technology and recognize that gender marking, gender coreference, and gender bias are aspects worthy of attention in pre/post-editing and for their societal implications.

The students were then asked to indicate whether they had ever noticed any differences in the translation of gender variation depending on the languages involved. In this case, there were no clear-cut answers. Overall, collating the positive responses ('Mostly yes' = 38% and 'Definitely yes' = 18%), the largest portion of the sample seemed to have observed differences, but 25% answered 'Mostly not' and 18% said that they had no ideas or were unsure ('I don't know'). This result is probably related to the focus of the module on just the English-Italian language pair and direction. Indeed, looking at the subsequent open question, where the students could elaborate on the likely causes for potential differences, those who responded mainly referred to the English-Italian combination (ex. 1), and references were made to the morphological characteristics of Romance languages (ex. 2) or the use of masculine generics in MT output (ex. 3).

1. Per esempio nelle traduzioni da inglese a italiano [For instance in the translations from English into Italian]

2. Lingue romanze o che hanno la differenziazione dei generi. [Romance languages or those that differentiate between the two grammatical genders]

3. La predispozione è ad utilizzare il maschile come genere neutro [The tendency is the use of the masculine as a neuter gender].

Moving to the problem of the reproduction of stereotypes, the students were required to say whether they had ever noticed forms of gender bias in NMT. As many as 73% of the respondents believed that NMT systems are not unbiased tools: specifically, 48% of the students chose 'Mostly yes', and 25% opted for 'Definitely yes' (Fig. 8).

The open answers elaborating on the reasons for the potential dissemination of stereotypes mentioned the unbalanced role of gender in contemporary society (ex. 4), a situation which is reflected in the corpora utilised for NMT training (ex. 5). In addition, some answers also identified the source of the problem in gender detection algorithms (ex. 6) (see Stanovsky *et alii* 2019).

Figure 8: Have you noticed a different treatment in the translation of feminine/ masculine forms (i.e., the presence of stereotypes)?

4. *Avviene perché si basano su forme "più diffuse", quindi se un determinato lavoro, ad esempio, è "tipico" più delle donne, viene automaticamente tradotto al femminile e viceversa.* [This happens because they are based on the "most widespread" forms; hence if a given job, for instance, is more "typical" of women, it is automatically translated in the feminine form and vice versa]

5. *Gli stereotipi di genere fanno parte della nostra società, di conseguenza sono presenti nei corpora su cui si basano le MT* [Gender stereotypes are part of our society; as a consequence they are present in the corpora on which MT is based]

6. *Algoritmi predisposti nella preferenza del maschile.* [Algorithms that tend to prefer the masculine form].

The two following questions regarded a very specific topic: the use of graphic conventions in the target language output by NMT programmes. In Italian, when referring to mixed groups, the masculine generic is regarded as the norm in most cases, and it is used as an unmarked form referring to both genders (Marcantonio, Pretto 2001: 334-338). However, speakers may diverge in their perception of markedness, with some language users seeing the masculine generic as conveying specific semantic traits, particularly when it comes to social roles or professions (Thornton 2016: 15-18). Hence, in accordance with growing sensitivity in relation to gender equality and inclusiveness, alternative proposals have been made when referring to mixed groups, especially in public communication. Examples in writing are the simultaneous use of the feminine and the masculine forms separated by a slash or the conjunctions *and/or* (*gli studenti/le studentesse* [students]). In some contexts, mainly when writing with a computer keyboard (e.g., e-mails), symbols such as the asterisk or, less frequently, the Schwa (ə) are now being used to replace the masculine generic suffix. For example, the expression *Dear all* addressed to a mixed mailing list could be translated as *Car* tutt**, where the asterisks imply both the masculine ending *-i* and the feminine *-e*. The Schwa symbol, which in the IPA (International Phonetic Alphabet) represents the central vowel sound of many languages and some Italian dialects, is used for the same purposes, especially with adjectives (e.g., *buonə* including *buoni* and *buone*, but also for the singular *buono/buona*). This convention is utilised particularly as a form of non-binary gender representation (for an overview of these uses see Iacona 2022).

The adoption of inclusive language strategies is quite controversial among the public as well as governmental institutions. It is thus not surprising, albeit somewhat discomforting, that in July 2022 the Italian Senate rejected a proposal for gender equality in institutional language, voting against the introduction of feminine forms for political roles in official documents and communications (e.g., *deputata* (f.) alongside *deputato* (m.) [member of parliament]). The *Accademia della Crusca,* a prestigious four-century-old institution studying and monitoring the Italian language and spreading "critical awareness of its current evolution" ["consapevolezza critica della sua evoluzione attuale", my translation][9], has manifested its position in favour of language inclusivity on multiple occa-

---

[9] *Accademia della* Crusca (8th September 2011). "Presentazione". Retrieved at https://accademiadellacrusca.it/it/contenuti/presentazione/6938 (last access: 22nd August 2022).

sions. In particular, it expressed its support for the parallel use of masculine and feminine forms in writing and speaking (e.g., *Cari tutti/Care tutte*). By contrast, with reference to the asterisk and the Schwa, its position is less favourable. The asterisk is considered acceptable in written, computer-mediated "private, professional and union communication", when intended for silent reading and within (ideologically) homogenous groups[10]. On the other hand, the Schwa is not considered necessary or appropriate in Italian because not only would it be difficult to be read aloud, but it would also be challenging for dyslexic people. Being employed beyond ways of address, that is, also for adjectives, the Schwa creates further difficulties in words written in capital letters, as it does not have a capitalised version. Furthermore, unlike the asterisk, the Schwa is not part of the inventory of symbols normally used in Italian.

These points were discussed with the students. No right or wrong position was proposed; rather, students were encouraged to make informed choices in their translations or post-editing work, justifying them in the light of the translation skopos and the ideological impact that gender-related language choices might have on the target readers.

The Q2 questions related to the topic of graphic conventions are interesting because they reveal awareness of the absence of these conventions in NMT output in Italian, but also the students' stance on these innovative proposals. When asked whether they had ever encountered the use of graphic conventions in NMT output, the quasi-totality of the sample answered negatively ('Definitely not' = 78%; 'Mostly not' = 14%) (Fig. 9).
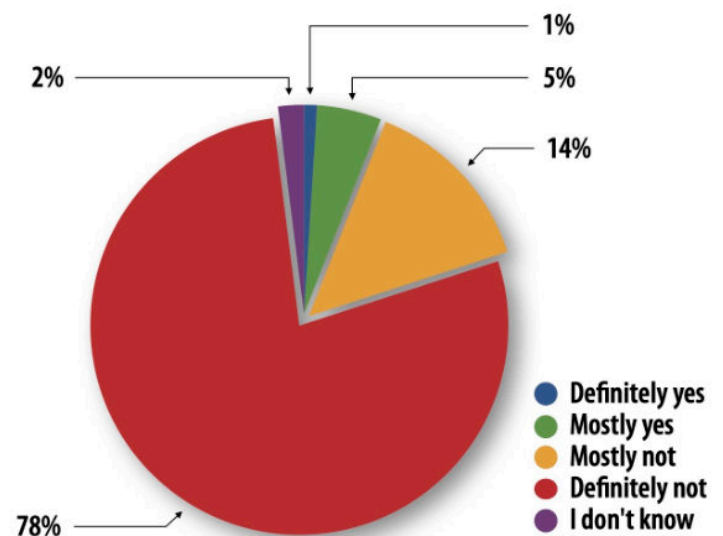
Figure 9: Have you noticed the use of gender-inclusive graphic conventions in MT (e.g., Car* tutt* [Dear all])?

---

[10] "[...] comunicazioni scritte o trasmesse che sono destinate unicamente alla lettura silenziosa e che hanno carattere privato, professionale o sindacale all'interno di gruppi omogenei (spesso anche sul piano ideologico)". Extract taken from D'Achille Paolo (24 September 2021). "Un asterisco sul genere". Retrieved at https://accademiadellacrusca.it/it/consulenza/un-asterisco-sul-genere/4018 (last access: 22nd August 2022).

Despite this and the cautious position of the *Academia della Crusca*, most of the students would appreciate the inclusion of inclusive graphic conventions in NMT results. Specifically, 38% answered 'Mostly yes' and 24% chose 'Definitely yes'. On the other hand, 20% of the respondents were not sure about this proposal and answered 'Mostly not', while 13% had a strong negative opinion about using inclusive graphic conventions in NMT (Fig. 10).

5%
13%
24%
20%
38%

- Definitely yes
- Mostly yes
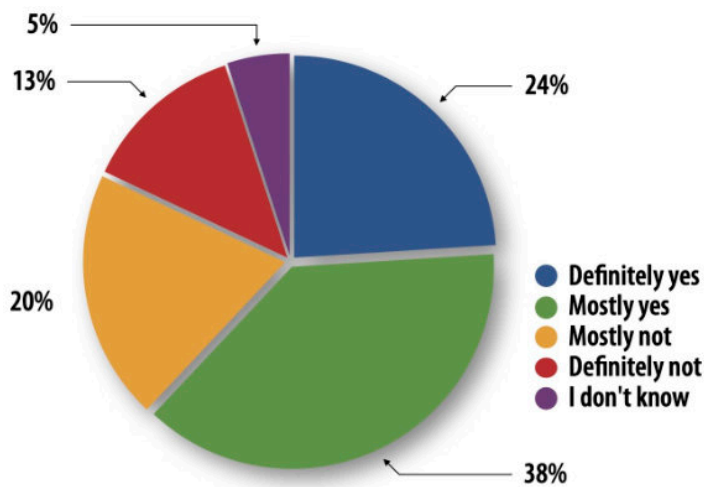- Mostly not
- Definitely not
- I don't know

Figure 10: If not, do you think that MT systems should use inclusive graphic conventions (e.g., * or the schwa symbol)?

1%
4%
13%
40%
42%

- The course was crucial
- The course allowed me to start a refection on gender variation in language
- The course further heightened my awareness of gender-related issues
- I already had previous knowledge; it was not the course that made me aware of these issues
- All the answers above

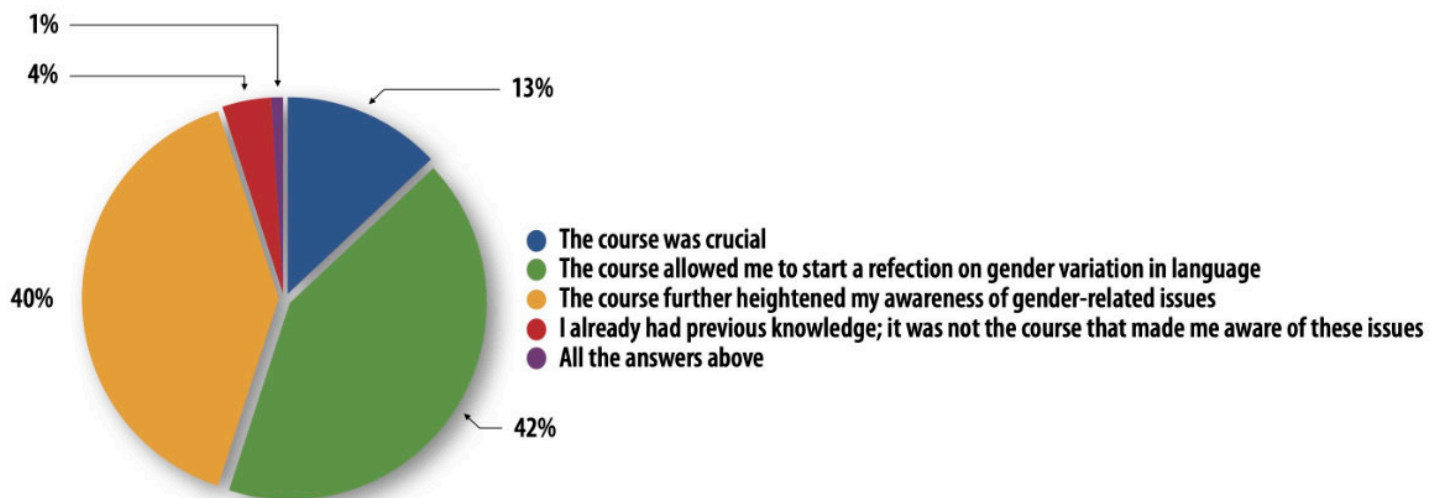Figure 11: How useful do you think the course was for you in terms of grasping aspects of gender variation in MT?

For 40% of the participants, the module increased their knowledge of gender-related problems in NMT, although they were already partly conscious of the issues discussed. On the other hand, for 42% of the students, the module was useful for starting a reflection on these topics, and

for 13% of them, the course was decisive. If we sum up these last two answers, we obtain about 55% of the students who had seldom or never considered the problems of gender variation and bias in NMT before. It is, therefore, important to underscore the role of higher education in generating awareness of such important implications.

5. Discussion and conclusion

This paper reported the results of two questionnaires administered among MA-level students at the beginning and end of an 18-hour module on AI and MT. The students taking part in the activities were not trainee professional translators, but English as foreign language learners aspiring to work in the corporate, institutional, or tourism sector. They will be required to produce texts in English, but also to perform a range of language mediation tasks, including translations from and into this language. Hence, their learning objectives included translation abilities but not necessarily a deep theoretical understanding of translation as a phenomenon, unlike in translation degrees.

MT systems are among the digital tools that university students increasingly adopt for language learning. The results of the questionnaires confirmed that the participants regularly used a range of AI-based instruments, particularly free online NMT systems and bi/multilingual platforms. Not surprisingly, the students manifested attention to and sometimes a strong interest in AI and MT advances. Despite this, MT literacy instruction was still in its infancy in the context analysed, and the module held within the project *Artificial Intelligence for European Integration (AI4EI)*[11] represented a pilot experience.

Among the reasons why MT had not received much attention are its controversial role in foreign language learning and its ethical repercussions in academic settings. As argued by Mundt and Groves (2016), if not used properly and according to an established code of practice, MT runs the risk of being a short-lived remedy to lack of language proficiency and an encouragement of cheating and plagiarism. It is precisely to mitigate these risks that MT literacy must be included in higher education also beyond translator training. Its introduction should promote conscious practices, according to norms of conduct mutually understood and sub-

---

[11] https://www.jmcoe.unito.it/home

scribed to by teachers and students, because "whether someone is cheating or not depends not on the technology they are using, but on the rules of the game" (Carré *et alii* 2022: 189).

Considering the goal of raising (initial) critical awareness among students of the strengths and limitations of AI and MT, to what extent was this objective met? The results of the questionnaires show that after the module, the students demonstrated a more watchful attitude when using bi/multilingual platforms with an increased percentage of participants always checking the source of a proposed equivalent. By contrast, their impression of the quality of NMT output was not significantly affected by the learning experience: before and after the course, the systems commonly available on the web were perceived as producing quite reliable results. This perception may reflect the high degree of fluency reached by NMT, which constantly improves its performance, to the point that the issues discussed with one cohort of students may no longer be relevant in the subsequent academic year (Enkin, Mejías-Bikandi 2016: 145). However, interpreting this general perception of reliability in a clear-cut way is difficult. On the one hand, the answers obtained may point to the need for more awareness-raising activities. It is indeed possible that the impression of dependability can be mitigated by extending the number of contact hours in future editions of the module (e.g., 36 or even 48 hours) and by including proper lab sessions to furnish supplementary occasions for active learning (see Doherty, Moorkens 2013). On the other hand, by examining more specific questions, it emerges that the students were overall capable of recognising that effectiveness may vary depending on the type of language feature, with collocations, neologisms, complex noun phrases and culture-bound lexis still posing great challenges to NMT. In addition, the result obtained would need to be examined further through additional questions going beyond the lexical level of accuracy and including the dimensions of text and discourse.

During the module's activities, NMT output was often found less that satisfactory when it comes to phenomena such as register, coherence and cohesion. The examples of gender coreference mistakes discussed in class are a case in point. The students were mostly conscious of these mistakes because they are very evident and can be amended through attentive proofreading. By contrast, identifying gender bias in NMT output was less obvious and somewhat more controversial, as perceptions of what

constitutes gender stereotypes are in part subjective and depend on the context of use of language forms. It is however important to stimulate discussion and reflection on the topic of inclusiveness, not to prescribe language behaviour but to equip students with tools for critical decision-making. It is significant that for more than half of the participants, the module represented the first occasion to reflect on gender-related issues in MT, indicating that digital literacy needs to include a focus on the socio-cultural implications of its use.

As noted above, the activities conducted during the MT module were experimental in many respects. For most students, the course represented the first formal learning experience through MT; additionally, although interest in MT literacy is on the rise, the number of studies describing best practices remains somewhat limited (see, for instance, Carré *et alii* 2022; Guerberof, Moorkens 2019; Loock, Léchauguette 2021). Nevertheless, combining the insights gained through the teaching experience, the questionnaire results, and reflections from the literature, it is possible to make initial suggestions for the effective implementation of MT literacy instruction in higher education.

It is crucial to distinguish between different use-case scenarios (e.g., professional translation, corporate communication, academic writing, see O'Brien, Ehrensberger-Dow 2020; Mundt, Groves 2016) and categories of MT users (e.g., professional translators, translator trainees, second language learners, see Monti 2019), identifying learning goals accordingly. For foreign language learners, the focus may be on MT as a tool to support the development of metalinguistic awareness (Enkin, Mejías-Bikandi 2016). Loock and Léchauguette (2021: 215) point out that "students need to be placed in the role of decision-makers in order to avoid any priming effect of the MT output". This may be done in several ways. For instance, they suggest choosing online NMT programmes with a dynamic interface (e.g., DeepL) where every word in the output can be replaced with a range of synonyms presented in a drop-down list: each newly chosen word leads to ensuing modifications to the text because items come with their typical collocational and phraseological profiles. To support decision-making processes, it may also be useful to compare different translations of the same word, passage, or text by different MT systems, as suggested in this paper (Section 4.1).

Decision-making relates to creativity, a skill that is sometimes regarded as being negatively affected by MT because users may feel con-

strained by the output (Moorkens *et alii* 2018) or they may be induced to concentrate on it, losing sight of the source text. However, if appropriately trained, students can employ MT to generate alternative translation options (Bernardini 2021), "treating the MT output as one possible version of a draft" (O'Brien, Ehrensberger-Dow 2020: 149), a strategy that could stimulate the creative process.

While for language learners MT technology should support the acquisition of the target language, complementing tools such as dictionaries and corpora, for translation trainees, it is essential to develop pre- and post-editing abilities and learn how and when MT can be integrated into the workflow. In either case, helping students acquire analytical skills is of paramount importance and the role of human judgment should be at the core. As future language professionals, they will act as "intercultural mediators [who should be] trained to recognise and deal with cultural differences, appropriateness of register and voice, contextual acceptability, potential ambiguity, terminological inconsistencies as well as conceptual and lexical gaps as they transfer meaning from one language to another [...], issues that MT systems cannot currently resolve" (O'Brien, Ehrensberger-Dow 2020: 147).

References

All the online resources in the References were last accessed on 15th June 2023.

Baker Mona (2018). *In Other Words. A Coursebook on Translation* (3rd edition). New York: Routledge.

Bernardini Silvia (2021). "Sul ruolo dirompente della traduzione automatica nella formazione dei traduttori". *L'intelligenza artificiale per la traduzione: verso una nuova progettazione didattica? #TranslatingEurope Workshop,* 3 December, 2021. https://youtu.be/P5IcV-GSmkQ

Bowker Lynne, Ciro Jairo Buitrago (2019). *Machine Translation and Global Research*. Bingley: Emerald Publishing.

Carré Alice, Kenny Dorothy, Rossia Caroline, Sánchez-Gijónc Pilar, Torres-Hostench Olga (2022). "Machine translation for language learners". In: Dorothy Kenny (ed). *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence.* Berlin: Language Science Press, 187-207.

Doherty Stephen, Moorkens Joss (2013). "Investigating the experience of translation technology labs: Pedagogical implications." *The Journal of Specialised Translation* 19, 122-136.

Dörnyei Zoltán (2007). *Research Methods in Applied Linguistics. Quantitative, Qualitative and Mixed Methodologies*. Oxford: Oxford University Press.

Enkin Elisabeth, Mejías-Bikandi Errapel (2016). "Using online translators in the second language classroom: Ideas for advanced-level Spanish". *LACLIL*, 9/1, 138-158. doi:10.5294/laclil.2016.9.1.6

Guerberof Arenas Ana, Moorkens Joss (2019). "Machine Translation and Post-Editing Training as Part of a Master's Programme". *The Journal of Specialised Translation,* 31, 217-238.

Iacona Andrea (8 January 2022). "Cari tutti". https://accademiadellacrusca.it/it/contenuti/titolo/19528

Kenny Dorothy (2022). "Introduction". In: Dorothy Kenny (ed). *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence.* Berlin: Language Science Press, v-viii.

Laviosa-Braithwaite Sara (2001). "Universals of Translation". In: Mona Baker (ed). *Routledge Encyclopedia of Translation Studies*. London & New York: Routledge, 288-291.

Loock Rudy, Léchauguette Sophie (2021). "Machine Translation Literacy and Undergraduate Students in Applied Languages: Report on an Exploratory Study".

Revista Tradumàtica: tecnologies de la traducció, 19, 204–225.

Marcantonio Angela, Pretto Anna Maria (2001). "Il nome". In: Lorenzo Renzi, Giampaolo Salvi, Anna Cardinaletti (eds). *Grande grammatica italiana di consultazione* (2nd edition). Bologna: il Mulino, 329-346.

Monti Johanna (2019). *Dalla Zairja alla traduzione automatica. Riflessioni sulla traduzione nell'era digitale*. Naples: Paolo Loffredo Editore.

Moorkens Joss (2018). "What to expect from Neural Machine Translation: A practical in-class translation evaluation exercise". *The Interpreter and Translator Trainer*, 12/4, 375-387,

Moorkens Joss, Toral Antonio, Castilho Sheila, Way Andy (2018). "Translators' Perceptions of Literary Post-editing Using Statistical and Neural Machine Translation". *Translation Spaces*, 7/2: 240–262. URL: https://doi.org/10.1075/ts.18014.moo

Mundt Klaus, Groves Michael (2016). "A double-edged sword: The merits and the policy implications of Google Translate in higher education". *European Journal of Higher Education*, 6/4, 387-401.

O'Brien Sharon, Ehrensberger-Dow Maureen (2020). "MT Literacy—A cognitive view". *Translation, Cognition & Behavior*, 3/2, 145-164.

Prates Marcelo, Avelar Pedro, Lamb Luis C. (2019). "Assessing Gender Bias in Machine Translation: A Case Study with Google Translate". *Neural Computing and Applications*, 32/10, 6363–6381.

Resende Natália, Way Andy (2021). "Can Google Translate Rewire Your L2 English Processing?". *Digital*, 1, 66–85.

Stanovsky Gabriel, Smith Noah A., Zettlemoyer Luke (2019). "Evaluating gender bias in machine translation". In Anna Korhonen, David Traum, Lluís Màrquez (eds). *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics, 1679–1684.

Thornton Anna M. (2016). "Designare le donne: preferenze, raccomandazioni e grammatica". In: Fabio Corbisiero, Pietro Maturi, Elisabetta Ruspini (eds). *Genere e linguaggio*. Milan: FrancoAngeli, 15–33.