INNOVATION IN LANGUAGE LEARNING International Conference

Do use Google Translate!
Machine Translation as a Tool for Language Learning

Barbara Bakker

Dalarna University, Sweden

Abstract

Machine translation, and specifically Google Translate, is becoming a favourite tool for language learners. As it is freely available across a variety of platforms for both desktop and mobile devices, students tend to rely on it excessively for solving their course assignments, which is ultimately detrimental to their language learning. However, the functionalities of machine translation can be exploited proactively within the curriculum and activities involving its use can be included in a course in order to increase language awareness and ultimately language learning. This paper is a contribution to the research field of machine translation in language pedagogy and reports the results of an empirical study conducted on students of an Arabic beginners course at a Swedish University. The study aimed at making the most of the students' attitude of relying on Google Translate, while at the same time taking advantage of GT's capabilities and limitations as a tool for language learning. A brief introduction covers machine translation and the reasons for GT inconsistencies with translating Arabic. This paper then presents two assignments given to the students at two separate moments of the term, accounting for their scope and execution. An analysis of the results follows, highlighting the students' performance and reflections on the advantages and disadvantages of GT.

Keywords*: Arabic, machine translation, language learning, language pedagogy, language learning tools.*

Introduction

The widespread availability of Google Translate and its ease of use have transformed the way students engage with a foreign language. Machine translation (MT) - the umbrella term used in the research literature to indicate such free online tools like Google Translate, Bing Microsoft Translator, Reverso and DeepL Translator - is freely available across a variety of platforms for both desktop and mobile devices. Language students tend to rely on it excessively, a practice that in most cases is detrimental to their language learning process and that affects, more generally, their behaviour during their academic studies. In fact, not only is MT, and in particular Google Translate, largely employed to solve foreign languages course assignments, but its use has also been observed as a widespread strategy for authoring research assignments, ultimately resulting in various forms and degrees of plagiarism (Ducar & Schocket, 2018 [1]; Mundt & Groves, 2016 [2]). Therefore, rather than to prohibit, with obviously questionable results, the use of MT for language studies, it is becoming more and more necessary to find ways to incorporate MT technology into the curricula as well as to explore instead how students can engage with it proactively and ultimately benefit from it - parallels have been drawn with the introduction of the electronic calculator in the classroom when teaching mathematics (Groves & Mundt, 2015 [3]). Indeed, several studies have proved that exploiting MT tools potential and integrating them into language courses may lead to increased language awareness and ultimately support language learning (see for ex. Clifford, Merschel & Munné 2013 [4] and Hellmich 2021 [5]).
Machine-translated texts have been getting more and more reliable because of the evolution of MT technology, which went from the traditional SMT, or Statistical Machine Translation, "the dominant translation paradigm for decades", to NMT, or Neural Machine Translation, able to "learn directly, in an end-to-end fashion, the mapping from input text to associated output text" (Wu et al. 2016, p. 2 and p. 1 respectively [6]). In other words, NMT systems can provide quite accurate results because they use "artificial intelligence to represent all the sentences in the target language at once, instead of breaking them into small chunks such as phrases or words" (Abdelaal & Alazzawie, 2020 [7]).

However, in the case of Arabic, the translated results are not always as accurate and reliable as with other European languages such as English, French, German, or Spanish, due to a variety of factors that affect language use and that are strictly related to the Arabic language. The first and probably the most relevant of them is the phenomenon of diglossia that characterises the Arabic speaking world. In his seminal article, Ferguson defined diglossia as "a relatively stable language situation in which, in addition to the primary dialects of the language […], there is a very divergent, highly codified […] superposed variety, the vehicle of a large and respected body of written literature […], which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation" (1959, p. 336 [8]). In fact, the written language (Modern Standard Arabic, or MSA), learnt in school and used in literature, news and more formal contexts, is thoroughly regulated within the framework of the Arabic grammar tradition, but it is not used in common spoken contexts. What is used in every day's life is a multitude of so-called dialects, or vernaculars, which vary from country to country and even, within the same country, from region to region - with phonological, morphological, syntactical and even lexical differences (see for example Ryding 2005, pp. 5-8 [9]).

The second, but not less relevant, factor is related to MSA morphology. MSA is based on consonantal skeletons, the elements of which are called radicals, or roots; the roots work together with sets of prefixes and infixes forming specific patterns, as well as with diacritics representing short vowels – which are normally not marked because they are naturally inferred by native speakers. This entails that both a context and a deeper knowledge of Arabic morphology are necessary in order to translate a word, as it is usually spelt only with its consonants. A simple example is *mustamiʿ* 'listener', compared to *mustamaʿ* 'listened to', active and passive participles respectively of the verb *istamaʿa* 'to listen': in standard texts, i.e. non-religious and/or not aimed at learning MSA, both participles are spelt in the same way, that is only with their consonants *m s t m ʿ*, as the differentiating vowels *i* and *a* are diacritics usually not written down and inferred from the context. For students of MSA, this precise and complex root system also entails learning how to look up words in Arabic bilingual dictionaries, as they are typically alphabetically ordered by root: the same verb *istamaʿa* 'to listen' would not be found under the letter '*i*' as it is a derived form (constituted by the three basic roots *s mʿ*, preceded by the prefix *ist*), and is therefore found under the first root, i.e. the letter '*s*'.

The many spoken varieties of Arabic, together with their differences when compared to MSA as well as MSA's morphological system, create a number of inconsistencies in machine-translated texts - for a major survey of Arabic MT issues, see Hadj Ameur et al. (2020 [10]). In addition, several studies have accounted for specific MT issues: for example, both Ali (2020 [11]) and Al-Khresheh & Almaaytah (2018 [12]) compared the effectiveness of MT applications in translating English into Arabic, while Bin Damash (2020 [13]) investigated the attitudes towards MT of Arabic students of English. In addition, it is necessary to mention that, as Harrat, Meftouh and Smaili specifically point out, Arabic dialects are "under-resourced languages" and they lack basic Natural Language Processing tools: "these dialects are not enough studied regarding to NLP area. Most MSA resources and tools are not adapted to them and do not take into account their features" (2019, p. 263 [14]).

GT as language learning tool: two assignments

The empirical study presented in this paper was conducted within the frame of a MSA course for beginners, taught at Dalarna University, Sweden. The course is assessed continually in weekly mandatory classes and through weekly homework assignments, as well as with one mid-term take home exam and online final written and oral exams at the end of the term.

The theoretical framework for this study loosely draws on the three-levels model of activity as presented by Engeström and Miettinen (2012, p. 4 [15]), with the top level being "driven by an object-related motive" (i.e. the reason for doing an activity), the middle level "driven by a goal" (i.e. what is being achieved with the activity) and the lowest level "driven by the conditions and tools of action at hand" (i.e. what means are employed in order to carry out such activity). Within a language learning environment, these three hierarchically ordered levels of an activity may rather be viewed as concurrent, i.e. three converging / combining aspects (Case 2015 [16]) and may be exemplified by any kind of course assignment. While the top and the middle levels are inherently implied in the course curricula (e.g. learning objectives and pedagogical implications of the assignment completion), the third aspect may consists in laying the focus on the employment of a specific tool for carrying out the activity.

During spring term 2022, the students were given two homework assignments, hereafter termed GT1 and GT2, where they were expressly requested to work with Google Translate. The GT1 instructions specifically explained the reasons for the unreliability of GT when translating into Arabic, thereby, in terms of levels of activity, also relating to the second level or aspect (what is being achieved). The GT2 instructions encouraged the use of GT as a language learning tool, therefore also providing a reason for employing MT (top level of the activity). In terms of timing, GT1 had to be handed in before the mid-term take home exam, i.e. after six weeks study. At this moment in the course, the students have just learnt how to read and write Arabic letters and their knowledge of Arabic is limited to very basic grammar rules, such as personal pronouns, gender and number agreement, as well as nominal sentences (i.e. simple sentences with only the copula in the present tense, which is normally not expressed in MSA) and a vocabulary of approx. 130 lemmas. The assignment consisted in translating, from English into Arabic, 12 short sentences comprising the grammar and the vocabulary issues covered during the first six weeks of the course, and then in comparing own translations with the translations provided by GT. The sentences were tested *a priori* with GT and purposely formulated so that GT would provide either grammatically incorrect translations, unknown vocabulary or dialectal formulations, or even advanced grammatical constructions, suitable for more formal MSA contexts but not covered yet in the course. Seven of the 12 sentences, when translated with GT, resulted in at least one grammatical mistake each: three of them in an incorrect demonstrative pronoun (proximal instead of distal), three showing incorrect gender (of a conjugated and as yet unknown verb, of the noun or the pronoun) and one resolving a predicate of a nominal sentence through a dialectal grammatical construction, i.e. not MSA. Moreover, a total of seven sentences translated by GT contained vocabulary and/or grammatical constructions that had not been encountered yet in the course - with, among them, three constructions formulated in a grammatically incorrect way. For this GT1 assignment, students were also asked to find, and to the best of their knowledge explain, any differences and/or mistakes, as well as identify words or expressions used by GT and not yet met during the course.

In the GT2 assignment, to be handed in around the end of the course, the students had to choose a news article from a provided list of Arabic online newspapers and translate it with GT into English or Swedish. In particular, the assignment consisted of three parts. Firstly, they had to account for their course of action (for example copy the whole text directly into GT vs copying&pasting small chunks, adjusting possible spelling complications, etc) as well as discuss the correctness and the intelligibility of the translation. Secondly, they had to briefly sum up (max 30 words) the contents of the article, to prove that they had understood it correctly. Thirdly, they were requested to identify three words, understood as keywords and therefore relevant for the topic of the article, and analyse them in terms of word class, roots and already known words, if any, sharing the same roots. This task was so construed in order to show a number of reasons for using GT. First of all, it demonstrated GT's usefulness when it comes to roughly comprehend and quickly overview the contents of any incomprehensible text in any foreign language - in fact, although almost at the end of the course, the students are still at a beginners' level and do not know enough MSA in order to read, understand and summarise a news article by themselves. It also aimed at showing GT's potentiality within the framework of vocabulary learning strategies - use the dictionaries, find synonyms, set words in context, etc.

Both assignments' instructions were accompanied by information on the research project, including purpose, research responsibility and personal data management. Students were asked to express their consent if they agreed that their assignments would be used for this research study and they were informed that they could withdraw their consent at any time. Students were also guaranteed total anonymity and notified that their assignments would be graded regardless their giving or not giving their consent.

Results and discussion

Only a total of 28 GT1 and 20 GT2 assignments were available for the analysis, as several students did not expressly consent to their assignments being used for research purposes - whether this was intentional or just simple oversight on their part goes beyond the scope of this paper.

The GT1 assignments were analysed both quantitatively and qualitatively, on the basis of the students' comparisons between their own translations and GT's. Of 28 students, 13 students reported and explained differences in all the sentences, 7 in 11 of the 12 sentences and 4 students found

discrepancies in 10 sentences. This equals to 86% of the students able to identify find at least 83% of the differences - an extremely positive result, considering their very limited knowledge of Arabic grammar at this point of the course. Several students also questioned GT's ability to translate gender agreement correctly and reported how GT translations changed when adjusting the English text, for example writing "my female friends" or "my girlfriends" instead of "my friends (f)", as was instructed in the assignment. Comments like "Google Translate is not good at understanding masculine or feminine" point at the students' awareness and comprehension of the gender agreement rules in Arabic and the exceptions discussed in class. Some students also speculated on GT's choice of demonstrative pronouns, as GT used proximal instead of distal demonstrative pronouns in all instances. A few students reflected on the reason of the discrepancies and wondered if they were caused by dialectal variation, different vocabulary choices or higher language register. Three students reported some advantages with GT, as the GT translations made them aware of a number of spelling mistakes and pointed at their own grammar mistakes: "Google reminds me of number agreement, which I had missed in this sentence".

The GT2 assignments were analysed qualitatively. Of 20 students, 14 translated an article into Swedish and 6 into English. All of them expressed their surprise at the accuracy of GT translation, as the translation was intelligible and comprehensible enough to let them understand the contents of the article in a very good way. Two students compared a double translation of their chosen article, i.e. first into Swedish and then into English, and reported that the English one was more grammatically correct and formulated in a slightly more comprehensible way. Several students reported some difficulties or encountered issues in connection with the procedure of the translation, for example the need to break down a longer text into smaller chunks, in order to allow GT to "make sense of the context", and GT's somewhat weird word choice sometimes, arguing that some nouns and/or verbs were not entirely proper in the context, although quite close synonyms. A couple of students pointed at GT's inconsistency in terms of rendering the spelling of proper names of persons or products, as they were spelt differently within the same article translation. Interestingly, three students reflected on the GT's features of automatic transcription of the Arabic words into Latin letters as well as GT's audio rendering of the Arabic: while they noted that both features may be useful and helpful in a language learning context, they argued that there was no way to know if the transcription and the sound were correct, because the copied and pasted Arabic text was not vocalised. Finally, two students noted that GT is definitely better at translating from, rather than into, Arabic. All the students satisfactorily summarised the contents of the article of their choice, which proves that GT had fulfilled its purpose in terms of enabling the students to comprehend a text otherwise for them impossible to grasp at this stage of the course. All of them also identified three relevant keywords for the article of their choice, providing correct details about their grammatical and morphological features. As for identifying the roots of the words of their choice, only approx half of the students were able to correctly deduce them and consequently find the words in the dictionary: as mentioned before, the Arabic morphological structure is a complex system and the skills necessary to know how to "go-back-to-the-roots" are taught and developed throughout several terms of Arabic studies. However, in several instances students reflected on the procedures necessary to identify them and reported their findings, asking for explanations. Eventually, the GT2 assignment also encouraged the students to work with Arabic authentic texts despite their limited knowledge of Arabic and made them get acquainted with Arabic websites and Arabic sources, which also, ultimately, may enhance their critical thinking.

This study has obviously no claims in terms of the possibility of generalising its results, mostly because of the small amount of data available, but also because GT is a NMT system and GT translations will therefore change (and possibly improve?) over time, despite the complications arisen from diglossia mentioned above. Moreover, it is also important to notice that the validity and the reliability of the results of the two assignments may also affected by a number of factors beyond the teacher's control, for example the fact that some students may already know (some) Arabic and/or may have asked an Arabic speaker for help in order to complete the assignments. Evidence of such instances is for example the fact that in GT1 a few students were able to correct those GT's wrongly formulated grammatical constructions that were "new", i.e. not covered yet in the first 6 weeks of the course.

The main scope of this paper is to provide an example of how MT may be employed to both the students' and the language course's benefits, in order to enhance language learning, and specifically grammar and vocabulary.

References

1. Ducar, C. & Schocket, D. H., “Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google Translate”, Foreign Language Annals, 2018, 51, pp. 779-795.
2. Mundt, K. & Groves, M., “A double-edged sword: the merits and the policy implications of Google Translate in higher education”, European Journal of Higher Education, 2016, 6:4, pp. 387-401.
3. Groves, M. & Mundt, K., “Friend or foe? Google Translate in language for academic purposes”, English for Specific Purposes, 2015, 37, pp. 112-121.
4. Clifford, J., Merschel, L. & Munné, J., “Surveying the landscape: What is the role of machine translation in language learning?”, @tic Revista d’innovació educativa, Valencia, Universitat de Valencia, 2013, pp. 108-121.
5. Hellmich, E., “Machine translation in foreign language writing: Student use to guide pedagogical practice”, ALSIC Apprentissage des Langues et Systèmes d’Information et de Communication, 2021, 24:1, https://doi.org/10.4000/alsic.5705
6. Wu, Y. & al., “Google’s neural machine translation system: Bridging the gap between human and machine translation”, Cornell University, arXiv:1609.08144, Cornell University, https://doi.org/10.48550/arXiv.1609.08144.
7. Abdelaal, N. M. & Alazzawie, A., “Machine translation: The case of Arabic-English translation of news texts”, Theory and Practice in Language Studies, 2020, 10:4, pp. 408-418.
8. Ferguson, C., “Diglossia”, Word, 1959, 15:2, pp. 325-340, https://doi.org/10.1080/00437956.1959.11659702
9. Ryding, K. C. “A reference grammar of Modern Standard Arabic”, Cambridge University Press, 2005.
10. Hadj Ameur, M. S. & al, “Arabic machine translation: A survey of the latest trends and challenges”, Computer Science Review, 2020, 38, pp. 100305, https://doi.org/10.1016/j.cosrev.2020.100305
11. Ali, M. A., “Quality and machine translation: An evaluation of online machine translation of English into Arabic texts”, Open Journal of Modern Linguistics, 2020, 10, pp. 524-548.
12. Al-Khresheh, M. H. & Almaaytah, S. A. “English proverbs into Arabic through machine translation”, International Journal of Applied Linguistics & English Literature, 2018, 7:5, pp. 158-166.
13. Bin Damash, N., “‘I can’t live without Google Translate’: A close look at the use of Google Translate app by second language learners in Saudi Arabia”, Arab World English Journal, 2020, 11:3, pp. 226-240.
14. Harrat, S., Meftouh, K. & Smaili, K., “Machine translation for Arabic dialects (survey)”, Information Processing and Management, 2019, 56, pp. 262-273.
15. Engeström, Y. & Miettinen, R., “Introduction”, Perspectives on Activity Theory, Cambridge University Press, 2012, pp. 1-16, doi:10.1017/CBO9780511812774.
16. Case, M., “Machine translation and the disruption of foreign language learning activities”, eLearning Papers, 2015, 45, pp- 4-16.