Applicability and Challenges of Using Machine Translation in Translator Training

Summary

During the last decade, translation as well as translator training have experienced a significant change. This change has been significantly influenced by the development of the Internet and the successive availability of web-based translation resources, such as Google Translate. Their introduction into the translation didactic process and training is no longer a matter of a teacher’s personal preference and IT skills, but a necessity imposed by the ever-swifter advancement of technology. This article presents the experimental results of an ongoing broader research study focusing on the modes and frequency of use of the Internet, Google Translate and Google Translator Toolkit among translation students at the undergraduate level. The preliminary results, presented in this article, are based on a questionnaire which was prepared in relation to the use of Google Translate while considering the latest professional findings. The article concludes with the author’s observations as to the applicability of these resources in translator training and the challenges thereof.

Key words: machine translation, machine translation systems, translator training, translation didactics, Internet, Google Translate

Smotrnost in izzivi uporabe sistemov za strojno prevajanje pri poučevanju prevajalcev

Povzetek

Na področju prevajanja in poučevanja prevajalcev je v zadnjem desetletju prišlo do korenitih sprememb, še zlasti zaradi razvoja svetovnega spleta in posledične splošne dostopnosti spletnih prevajalskih virov, kot je Google Translate. Vpeljava teh virov, v proces poučevanja prevajalcev ni več le svar učiteljevih osebnih preferenc in poznavanja informacijske tehnologije, temveč nujnost, ki jo narekuje vse hitrejši tehnoški razvoj. Pričujoči članek predstavlja eksperimentalne razultate študije, ki še poteka in ki poučuje načine ter pogostnost uporabe svetovnega spleta, spletnega prevajalnika Google Translate in spletnega orodja s prevajalskim spominom Google Translator Toolkit med študenti prevajanja na dodiplomski ravni. Vmesni rezultati, predstavljeni v tem članku, so zbrani na podlagi vprašalnika o uporabi orodja Google Translate, v povezavi z zadnjimi strokovnimi ugotovitvami na tem področju. Članek zaključujejo avtoričine ugotovitve o aplikabilnosti ter izzivih uporabe teh virov pri poučevanju prevajalcev.

Ključne besede: strojno prevajanje, sistemi za strojno prevajanje, poučevanje prevajalcev, prevajalska didaktika, svetovni splet, Google Translate
Applicability and Challenges of Using Machine Translation in Translator Training

1. Introduction

At the beginning of the past decade Frank Austermühl announced that, in the information age, translation requires a completely new strategy with regard to the logistics of information (2001, 1). His claim that a paradigm shift is needed in both methodological and practical approaches to translation in the areas of translation teaching and researching is truer today than ever. Translator training, which in the not so distant past was predicated upon teacher resourcefulness and the (un)availability of parallel texts and translation-relevant material, is almost impossible to carry out without recourse to information and communication technology (ICT). Consequently, freely available online web tools and services are becoming an increasingly important classroom resource.

Contemporary ICT translation tools can be roughly divided along two lines: general-purpose translation software applications and special–purpose translation software, such as terminology management and translation memory systems (Fulford and Zafra 2005). Among the former, machine translation (MT) systems and platforms occupy a very controversial position: while some professional translators still perceive machine translation as a threat to their status and the way they make a living, an increasing number of others and particularly translation agencies report of significant improvement made possible by it. The truth is that MT has already conquered some specific areas in which human translation cannot offer a competitive advantage. As studies (Aiken and Balan 2011, Fulford and Zafra 2005, Vargas and Ramírez Polo 2011, and the sources cited therein) report, MT is already used by many professional translators to aid their work, primarily for the purposes of quick information gisting, or for preparing a rough first draft of a translation.

That said, MT has – stealthily and mostly involuntary on the part of teachers – asserted itself also in contemporary translation classrooms. Students of translation seem to be increasingly resorting to machine translation to support their work, despite their teachers’ attempts to prevent them from using it. The teachers are consequently left with two basic options: to take even harsher measures to exclude MT from classes or to incorporate these resources into the didactic process. The author believes that the latter should be the case and that adjustments need to be made in the way translators are taught and trained in order to accommodate this change in the translation didactic paradigm.

2. Machine Translation and the Translation Classroom

2.1 Machine Translation: Where do we stand?

Machine translation is not a recent addition to the vast array of ICT technologies and has been a subject of research for over a half of a century. It was instigated, as Berner (2003, 5) reports, by the
desire to “remove the language barriers that hinder scientific communication and international understanding”. The challenge was and still is to produce translations that are as good as those produced by humans.

By definition, “machine translation involves the use of computer programmes to translate texts from one natural language into another automatically” (Baker and Saldanha 2008, 165). Like translation done by human translators, MT does not simply involve substituting words in one language for another, but applies complex linguistic knowledge to the text and/or selects the most probable words and sentence sequences from huge corpora of already existing translations.

Today, there are two main approaches to machine translation: MT systems can be broadly categorized as either rule–based or statistical. Rule–based MT systems, the best known among which is Systran, retrieve linguistic data from gigantic bilingual dictionaries and grammars which are then supplemented by sophisticated morphological, syntactic and other rules. Contemporary corpus-based statistical MT systems rely on large parallel corpora of human–engineered translations which are utilised to automatically infer a statistical model of translation. The underlying premise is that for every source language element there are a number of possible translations, and the most adequate translation is assigned the highest probability by the system (Veritas 2009).

There are several freely available web-based MT systems, including the following:

Google Translate ([http://translate.google.com/](http://translate.google.com/))


Bing Translator ([http://www.microsofttranslator.com/](http://www.microsofttranslator.com/))

Yahoo! Babel Fish ([http://babelfish.yahoo.com/](http://babelfish.yahoo.com/))

Google Translate (GT), the most popular among them, is a statistical MT system that currently provides automated translations, directly or via a pivot (i.e. a bridging or intermediary language) between almost 60 natural languages. GT was introduced by the Google Corporation in 2007. The corpora utilised by GT for statistical analysis include translations made for the purposes of e.g. the EU, UN and the European Patent Office, as well as a huge library of books included in the Google Books database, to mention but a few. Slovene was added to the list of GT–supported languages in September 2008.

Google Translate soon proved to be very powerful, albeit with great variation in the degree of accuracy between individual languages – a fact closely tied to the (un)availability of large and qualified parallel corpora. GT’s accuracy also seems to be related to text–type, genre or subject domain (cf. Aiken 2009). There are also reports (Drugan and Babych 2010) that “for some language pairs (e.g. translation between closely–related languages) or in certain narrow subject domains (e.g. software manuals, development documentation), post-editing MT output requires less effort than translating the original text from scratch”.
Among other disadvantages of GT and statistical machine translation systems in general authors highlight users’ complete lack of control over both the translation input (this pertains to the controversial issue of confidentiality of data fed into the system) and output. This is often caused by the generic, surrogate, and unverified data used to produce translations\(^1\) (resulting in unpredictable and inconsistent translations); significant hardware and software infrastructure which is needed to build, manage and maintain large translation models,\(^2\) and last but not least, human assistance before, during and particularly after the machine translation process. On the other hand, the advantages include improved quality due to greater terminological and phraseological consistency, enhanced productivity and speed of translations, cost reduction, lack of bias and general availability of the systems. Neither of positives or negatives is by any means exhaustive.

3. Translator Training: Approaches, Competences, Objectives, Outcomes

In the past, translator training was predicated upon what Kiraly calls the transmissionalist approach, which understood both teaching and learning as “the transmission and reception of knowledge (truth) about the world” (Kiraly 2000, 23ff). The teacher, who presented the students with his/her own “correct” version as a model translation, was in the centre of attention, not the student. Therefore, as both Kiraly and Kelly (Kelly 2005, 97) rightfully observe, this approach was (or maybe sometimes still is – as even today translator training is still embedded in the transmissionalist tradition) frustrating for the students and did not contribute much to the overall development of their translation skills. Furthermore, it postulated the primacy of the translation product over process, leaving the student in a void without proper principles or methods for arriving at an appropriate translation solution.

The contemporary transformational didactic model endorsed by Kiraly views translator training primarily as a learner–oriented activity. The teacher assumes secondary roles such as guide, assistant, mentor and facilitator and creates a communicative and interactive learning environment for pro–active students engaged in collaborative activities. This model is “centred around principles, methods and procedures, rather than on the translation product” (Gile 1995, 10–1). It is, nevertheless, still very much characterized by learning–by–doing (i.e. hands–on translation), and, in this context, “associated with […] skills\(^3\) needed to produce an acceptable translation” (Pym 2009).

Translation training is closely associated with the notion of translator competence. Translator competence has several components and describes what professionals (or students holding a degree in translation) should know and be able to accomplish as competent translators. As the most comprehensive and all–encompassing categorization of translator competency, Kelly’s (Kelly 2005, 32, 33, 73–8) summarized classification has been used for the purposes of this article:

1\(^{http://works.bepress.com/uwe_muegge/52/}\)
2\(^{http://www.systran.co.uk}\)
3\(^{It shall be understood for the purposes of this article that translation ability is not (only) an innate capacity, but a skill gradually acquired through theoretical learning and practical training.}\)
- **Language competence**, i.e. communicative and textual competence in at least two languages, taking into account the ‘real’ level of language competence of students.

- **Cultural (and intercultural) competence** beyond geographical and other factual knowledge. Translators should acquire competence in perceptions, myths, beliefs, values, stereotypes shared by the members of their working culture.

- **Professional and instrumental competence** which encompasses the use of resources of all kinds, terminological research, information management, and the use of IT tools.

- **Interpersonal competence and networking**, ability to work with other professionals involved in the translation process, ability to work in teams.

- **Subject area competence** in different disciplines, allowing the translator to comprehend the source text, to know where to look for information.

- **Attitudinal competence** in relation to self-confidence and socialization as a professional translator.

Translation competences need to be translated into clear and achievable teaching objectives (i.e. statements capturing the knowledge, skills, attitudes achieved by students), both general and specific, and the subsequent learning outcomes (i.e. results of a period of specified and supported study, identified prior to every translation class). Setting teaching objectives and learning outcomes has the following important advantages: it facilitates communication between teachers and students as well as the choice of the applicable teaching tools; it suggests different learning activities and provides a basis for assessment of the achievements. The author anticipates that the results of her research will help endorse these competences as well as contribute to improved translator training techniques integrating machine translation into the translation training process.

### 4. The Study

#### 4.1 The Setting, Teaching Objectives and Learning Outcomes

The study was carried out by the Department of Translation Studies at the University of Maribor as part of the *Translation 2* tutorial during the winter semester of the 2010/2011 academic year. It was conducted with a group of 3rd year students of the BA Inter-lingual Studies (English) programme. This was the students’ second English–Slovene translation tutorial comprising a total of 45 semester hours, and the students were expected to have already acquired some basic competences and skills in English–Slovene and Slovene–English translation. Apart from that, their previous experience and knowledge included classes on English and Slovene linguistics (phonetics, word–formation, syntax etc.), literature and culture, as well as tutorials in oral and written communication, English language development (focusing on grammar), and introductions to the study of language as well as professional translation and interpretation. Some elements of English–Slovene stylistics have also been covered and during *Translation 2* tutorial, focusing on newspaper style, headlines and official documents.
The didactic model applied in the Translation 1 and Translation 2 tutorials was a traditional one, supplemented by a process-based approach, combining lectures – or rather presentations – on selected subject matter, with out-of-class preparation of authentic texts and the group assessment of translations, whereby special emphasis was placed on the very process of translation. The general focus of teaching objectives of the tutorial was on the acquisition of competences in English–Slovene and Slovene–English translation, based on the use of authentic texts by students at an advanced level. More specifically, the students were to be made acquainted, among other, with different text types, skopos, cultural specifics in English and Slovene, and with translation strategies that would enable them to reflect critically on their translation decisions and solutions.

Based upon the intended learning outcomes, on completion of the tutorial, students were supposed to be able to translate texts at an advanced level from English to Slovene and Slovene to English, recognize the text type and its characteristic features, identify the possible skopos, i.e. objective of translation, prepare alternative translations, present translator’s arguments, as well as find parallel texts and use electronic devices and dictionaries.

4.2 Motive and Data Collection

The need for the study has arisen from the author’s own classroom experience. During translation tutorials, students were starting to increasingly use Google Translate to support their homework assignments and the consequences were beginning to have an unnerving effect on classroom workflow. Students were disinterested and unwilling to participate in class discussions because they could not reflect on the translation process or the translation solutions offered by either themselves or their colleagues. With a view to creating an interesting classroom setting which would not exclude but incorporate this resource in the didactic process, a questionnaire (Questionnaire A) was prepared in order to address the use of Google Translate (GT).

Questionnaire A comprised a total of 11 questions: 9 multiple choice and 2 open–ended. The questions were designed to be as clear, concise and as motivating as possible. Thematically, they were divided into three groups: the first group addressed the use of GT in general and the second group the use of GT in relation to different types of text; the third thematic group assessed the students’ perceptions as to the reliability of such tools and their opinions as to their present and future roles. The aim was to make students both aware of pitfalls associated with the use of GT and cognisant of areas that needed their special attention.

4.3 Evaluation

Questionnaire A was answered anonymously by 33 students. Multiple answers were also possible to some questions. For example: What are the main deficiencies of GT, and What post–editing activity demanded the most time.

---

4 Here, skopos is understood as a “purpose, aim, goal or objective of a translation” in line with Vermeer’s skopos rule which postulates: “Human action is determined by its purpose, and it’s therefore a function of its purpose” (Baker and Saldanha,117).
We first asked the students whether they used *GT* and how they learned of it, as well as how often they used it. The majority (74.5%) confirmed the use of *GT* when preparing for translation classes and a further 15.5% reported combining *GT* with *Amebis Presis*, another Slovenian–developed rule–based machine translation system (also freely accessible on the web). Only 10% of students claimed not to use *GT*. For the most part, students learned of *GT* by chance (e.g. by surfing the Internet) while the majority of the remainder were told of it by a colleague or friend (Fig. 1). The introductory lesson at the beginning of the semester also addressed the issue of machine translation and thus some answered that they learned of *GT* in class. Two students offered multiple answers. Most students reported that they used *GT* occasionally; a lesser number stated that they always used it, and some 20% stated that they used it rarely. One student left this question unanswered. (Fig. 2)

![Figure 1. Where have you heard of GT?](image1)

*Figure 1. Where have you heard of GT?*

The second thematic group addressed the types of texts students translated using *Google Translate*, how reliable they believed its output was, and their general opinion of *GT* as a tool. When asking students about different types of text, we differentiated solely between literary and non–literary texts because it was not our intention to delve into the realm of type–specific translation phenomena but rather to generalise. Typology had been addressed in a previous class; therefore, students were familiar with the main characteristic features associated with different types and styles of text. Approximately one half of students (48.5%) did not discern text with regard to text type and reported using *GT* with all translations; another half (48.5%) replied that they used it for non–literary texts only. One student further specified scientific literature and official documents, whilst two students did not supply any answers. No one believed that *GT* was very reliable; approximately one half claimed it was fairly reliable, and another half believed it to be unreliable. Two students added that *GT*’s reliability depended on the text type; one student claimed *GT* was more reliable when translating from Slovene into English than vice versa, whilst another student noted that *GT* was reliable when dealing with set expressions or collocations. One questionnaire remained unanswered (Fig. 3). No students reported that *GT* significantly facilitated their work; the majority, however, stated that it helped them with their translation assignments, but a lot of additional work was still necessary. One questionnaire remained unanswered. (Fig. 4)
The third thematic group was designed to assess the students’ critical attitude to Google Translate output. First, we wanted to assess GT’s perceived deficiencies in terms of unreliable vocabulary, faulty composition of sentences, grammatical issues and other shortcomings. Since multiple answers were allowed, almost one half of the students (15) cited more than one reason – opting mostly for faulty composition of sentences and unreliable grammar – whilst eight students chose all. One student listed other reasons, namely: a disregard of context and a word–for–word translation approach (Fig. 5). All students edit GT’s output using dictionaries (Pons) and other resources, among which they list corpora and terminology databases, specifically Evroterm and Evrocorpus. Two students also reported using their own resources (glossaries) (Fig. 6). Consistent with the above replies were their answers to the next questions, where they said that during the editing process, the most time was spent in formulating sentences (43%), the search for appropriate vocabulary (31.5%) and the correction of grammar (23.5). Important, too, was the apparent awareness of the complexity of these issues, with 16 students choosing more than one reason, and 4 citing all of them.
reported that translation engines were very important in their professional training, while on the other hand only four were convinced they were “not very important”, “contributed little to their education”, were no “substitute for a human translation” and would “merely support idle students”. The majority believed translation engines to be “fairly important”, “useful” “helpful” and that they “facilitated work”; several also added that they speed up work and save time. About one third of students (10) noted the good as well as the bad sides to translation engines, which is, we believe, responsible behaviour on their part.

As to the role of translation engines and – if we may generalize – of machine translation in general in students’ future professional lives, we attained a similar picture. Only a small number of them (3) believed that translation engines will become increasingly important and that they will probably use them a lot. On the other hand, only four students expressed their belief that machine translation will play “no role whatsoever”, that its importance will diminish and that it “will have no proper role”. The majority were again of the opinion that translation engines will be “helpful”, “will facilitate work” and “save time”, but pointed to their limited applicability (lexis, framework translations, certain types of text). Five students also expressed their belief that in the future “better (in terms of quality) translation tools will be available” or that translation will become “increasingly automated”.

5. Results and Implications

On evaluation, Questionnaire A’s results revealed that the vast majority of students were using Google Translate during their preparation for translation classes, which corroborated our experience and expectations. In line with our hypothesis, students reported using GT either occasionally or all the time. The fact that the majority of students learned of this tool by surfing the Internet testifies to their advanced IT and research skills (professional and instrumental competence) and of the pervasive presence of Internet and online translation resources in their daily lives; translation teachers, of course, cannot neglect this fact. The fact that the remainder learned of GT through colleagues and friends also testifies to the close–knit network of relations amongst students, an aspect which also contributes to the development of their interpersonal competences and networking.

The second thematic group of Questionnaire A, addressing the perceived reliability of GT output and its competence and applicability in the translation of different types of text, proved that students – at least reportedly – were not overly reliant on the output of machine translation engines, and are aware that additional editing is necessary. The fact that their effort involved in editing MT output still lags behind what they declare and can conceivably achieve is another area that demands special attention. As noted in the introduction to this article, this study was undertaken particularly because students were unable to critically reflect upon texts translated with the use of GT; therefore, structured activities are needed addressing the post-editing process to obtain an acceptable text. The author identifies here an important area of future research for comparativists and translation teachers.

In the third thematic group, grammar and syntactic issues were identified by the students as deficiencies, followed by unreliable vocabulary. Their answers were consistent with replies
provided to the question as to what editing activity was most laborious; here students stated that correcting grammatical errors and formulating sentences was the most time–consuming. When asked about \textit{GT}'s main deficiencies, multiple answers provided by almost one half of the students confirmed our expectations that they – at least in theory – were aware of numerous pitfalls associated with the use of \textit{GT}. Important, too, is that according to their (multiple) answers students are aware of the complexity of \textit{GT} output editing.

It needs to be added, however, that a translation's acceptability at this level of students' translation expertise is mainly confined to lexical, syntactical, and grammatical aspects only. Cohesion and pragmatic equivalence (coherence) are mostly overlooked by them, as are the translation assignment specific deliberations such as \textit{skopos}, translation brief and the like. The use of machine translation further accentuates this focus and diverts students' attention away from translation process and source text analysis. Therefore, in order to come closer to Kiraly's \textit{transformational} approach to translator training (postulating the study of translation principles, methods and procedures), didactic activities should incorporate focused exercises surpassing those concentrating solely on lexical, syntactical or grammatical aspects (such as engaging students to search for unknown words, deleting superfluous translation alternatives, correcting lexical and grammatical errors, and rewriting parts of (or entire) sentences). Examples of suggested activities are identifying source–text characteristics and their rendering in the target text on advanced levels (textual, pragmatic, ethical, etc.), producing an imaginary translation brief, visualising the target audience, back–translation and analysis etc.

The fact that half of the respondent students use \textit{GT} in a non–discriminatory manner in the translation of all texts – literary and non-literary alike – points to another possible problem: highlighting the need for further instruction and structured activities addressing this issue. In this context, students' observations on how the quality of \textit{GT} output crucially rests on the type of text are also important and somewhat contradictory. Some possible activities to consider in this respect include translations of selected texts pertaining to different types/genres/subject domains, individually addressing selected issues of vocabulary and sentence construction but also matters of register, audience, the differences between source– and target–text cultures.

Lastly, when asked about the importance of translation engines in their current role as students of translation, students report that such technology is of limited help and application (useful only for certain types of text as well as drafts, etc). However, in line with reports from professional translators, they highlight their time-saving potential and believe that MT is going to facilitate their work and thus be helpful in their future careers.

The author finally identifies a fundamental change in the roles of translation teachers resulting from the growing use of machine translation among the students. Namely, teachers are forced to resign from their position of sole “knowledge providers” to their students. Consequently, they often find themselves in a situation where they have to compete for their attention with technology which can resort to much more extensive databases and can produce (more or less acceptable) translation solutions much more quickly. To be able to overcome this challenge, they need to admit their limitations and step in primarily as an organizers, managers and evaluators.
of machine produced information by providing advice and instigate guided activity e.g. on the structure and selection of translation solutions. This requires increased time and effort spent on devising new classroom strategies, learning materials, activities, etc. but results in a significant change also on the part of students: they no longer remain passive learners but become pro-active and motivated individuals.

6. Conclusion

The preliminary findings of the study confirm that technology is not a mere option in today’s world of professional translation but a necessity upon which the younger generation of translation students seems to be becoming increasingly reliant. The results show that freely accessible machine translation systems and platforms are used by students of translation to provide first drafts and save labour, and therefore need to be taken into account by translation teachers when designing didactic activities. The results also point out that students seem to be conscious of the pitfalls associated with the use machine translation, but need structured activities to be able to deal with them successfully.

Building upon Kiraly's transformational didactic process which stresses the primacy of translation process over product, the author observes that the use of machine translation, diverts students’ attention away from translation process and the source text analysis, and suggests activities to overcome these deficiencies. She also highlights the welcome change in roles of teachers and students affected by the use of machine translation: while the teachers are forced to resign from their position of knowledge transmitters and are assigned new roles as coaches and managers of information, the students do not simply receive information any more but act proactively and responsibly.

The author is aware of the limitations of this study and continues to improve research methods and approaches to augment her experience. These limitations include the rather narrow scope of the sample, lack of specific examples and utilization of a non-probability (convenience) sampling technique. Such a technique is usually used when it is impossible or impractical to use random sampling techniques and is common in educational research. Its disadvantage is that it cannot be generalised to the desired extent. Nevertheless, in the absence of scientific investigation in this area the author is certain that the translation didactics can nevertheless benefit from her preliminary findings. Further empirical evaluations, supported by a larger sample of specific examples, however, are needed to confirm (or refute) the findings in different contexts and surroundings.

Bibliography


