

GOOGLE TRANSLATE POST-EDITING: THE CASE OF BULGARIAN-ENGLISH TRANSLATIONS

Svetlana NEDELICHEVA

Konstantin Preslavsky University of Shumen, Bulgaria

Abstract: *This article studies some linguistic problems emerging in the translation pair Bulgarian – English using Google Translate. Google Translate is a translator tool that influences students’ learning nowadays, so the application of Google Translate is highly significant because it can help in translating texts from English into Bulgarian and vice versa, making this tool one of the solutions in the field of translation. Therefore, students should know the benefits obtained from that program and it can also be used to make it easier to understand the contents of fiction and non-fiction texts in English. The research method used is descriptive qualitative research, which aims to comment on some common problems with Google Translate most Bulgarian learners are not aware of. The study shows students how suitable post-editing can improve the quality of target texts. Editors need an advanced level of language knowledge to notice the mistakes made by the software and to correct them. Instructors should show their students that Google Translate in some cases provides alternative translations in a popup, and the alternative may be more suitable than the first translation variant. Google translate is an efficient tool, because it is free for its users and its software is under constant development.*

Keywords: *Google Translate, machine translation, Bulgarian, English, post-editing*

Introduction

Computer-assisted translation tools have been in use by translators for many years. Since the development of artificial intelligence, machine translation is in the process of constant improvement. Google Translate (GT) is a good example of this tendency. Among all machine translators, Google’s service, GT, is probably the most well-known and used because of its link to the unanimous search engine. Still no machine translation engine can produce better translations than human translators. Machine translation, however, can enhance and facilitate a translator’s work. Since 2016 Google Translate is a neural machine translation engine which translates sequences of words at a time, rather than just word by word. It uses broader context to figure out the most relevant translation. Neural machine translation is the most advanced form of machine translation (MT). It has resulted in translations that are significantly superior to those produced prior to the upgrade. According to Wu et al. (20), it “approaches the accuracy achieved by average bilingual human translators...” It teaches itself on how to translate by using a large neural

network. Yet, it has its own pitfalls, especially mistranslations while the target text looks very natural - hence it is harder to spot issues when post-editing.

Literature review

Back in 2015 Groves and Mundt (112) predicted that:

“given the acceptance of other digital technology for teaching and learning, it seems likely that machine translation will become a tool students will rely on to complete their assignments in a second language”.

Only seven years later digital technologies have advanced so much that users are able to do a lot of lexical and grammatical research independently.

A few small-case studies of GT have been conducted using a number of languages. For example, one study (Chen et al.) translated a pamphlet with Google from English to Spanish and Chinese. Results showed that there was a more accurate translation from English to Spanish than English to Chinese, and a Spanish human translator did not provide a significantly better translation than GT.

Another study was conducted using the newer version of Google Translate with neural machine translation (cited in Aiken 255) using 20 English phrases: fly out of London, like a bat out of hell, out cold, out of bounds, out of breath, out of curiosity, out of focus, out of his mind, out of milk, out of order, out of pocket, out of steam, out of style, out of the closet, out of the game, out of the office, out of this world, out of time, out of wedlock, and out on the town. These phrases were translated to the 102 non-English languages supported by GT at the time, and fluent speakers of these languages evaluated the translations. The results showed that the translation in Bengali, Haitian Creole, and Tajik failed 100%, and the following languages failed 80% or more: Kurdish, Nepali, Latin, Malaysian, Urdu, Maori, Cebuano, Georgian, Persian, Punjabi, and Uzbek. In 2019 Aiken published an updated research of a comprehensive evaluation of accuracy made in 2011 with Google Translate using 51 languages which showed that many European languages had good results, but several Asian languages performed poorly. Eight years later the reevaluation, which used the same text as the original study showed a 34% improvement based upon BLEU scores. This new study revealed that translations between English and German, Afrikaans, Portuguese, Spanish, Danish, Greek, Polish, Hungarian, Finnish, and Chinese tend to be the most accurate.

Google translate is based on hundreds of millions of pre-translated words and phrases from the internet. This means that if one variation of a word appears, the program will prefer the more general translation. To avoid this, Google has developed an update that provides masculine and feminine translations for these neutral words. The problem occurs when translating into

a language as Bulgarian where not only nouns but also adjectives have their respective genders such as “beautiful” turning into masculine, feminine and neuter. According to Google, the weaknesses were learned from existing examples and translated online. Fitria discusses gender bias in terms of Indonesian-English translation by using Google Translate. The author suggests that most likely every language has gender-biased sides, including English. In Google translate, the linguistic differences between men and women translated by GT cause the system to be considered gender biased and sexist. Similarly to English, Indonesian is a gender-neutral language. But when Google translates Indonesian into English, the sentence becomes gendered. Nowadays both men and women can have various activities and jobs.

“The Indonesian language in this case seems to have been saved from being sexist because it does not associate a particular profession or activity with any gender. Unlike English, which adjusts personal pronouns based on gender. Google Translate is not always accurate, especially when translating from English to other languages” (Fitria 285).

Gender bias in Google Translate is also studied in the translation pair Spanish-English. The researcher, Lopez Medel, entered for translation a number of gender-invariable Spanish nouns whose referent’s gender was unknown due to the omission of pronouns and other particles. In several cases, an invariable (genderless) adjective was added to the search in order to test for changes. Due to the update Google launched in 2020 based on its neural translation technology that generates a default translation, the software rewrites results that are gendered and checks for accuracy. Therefore, translations between different-gender languages like Spanish and English (as well as Bulgarian and English) should offer the feminine and masculine versions, however, it turned out this is not always the case.

Another GT research, which includes Indonesian, studies the translation of English language materials, so students can switch from using the print dictionary to the Google translate application (Sitorus). The results show that 90% of Class B students in the State Administration Study Program, Faculty of Social Sciences, Manado State University use the Google translate application to help translate English-language material. Among the characteristics, which the students point out as reasons to choose Google translate applications over a printed dictionary, are “fast” and “lightweight”. Although there are deficiencies in this application, students still feel positive about its use.

Other GT studies focus on using it in foreign language teaching, e.g. as a tool to improve L2 writing (Farzi; Tsai; Lee), to devote more time to the content and communicative aspects of language (Raza, Nor), to explore the possible benefits of using GT at various course levels of English for Academic

Purposes (EAP), i.e., to see if the use of GT affects the quantity and quality of student writing (Kol et al.), to analyse the aspect of literalness of Google translate in translating business correspondence (Riadi, Angelina), etc. This research studies a language pair which has not been specifically studied so far, namely Bulgarian and English. The following analysis aims at pointing out some common translation mistakes made by GT, which can help Bulgarian learners using this software in their post-editing practices.

Data Analysis

The analysis in this study is based on examples collected ad hoc during translation classes of *English Studies – Linguistics and Translation* MA program at *Konstantin Preslavsky* University of Shumen. The corpus data consists of non-fiction English and Bulgarian texts, which are excerpts of articles in the field of linguistics, history and ethnography. The examples provided in this analysis are representative for particular types of mistakes relevant to the present day status of GT software. Before focusing on some GT fallacies, however, its significant advantages should be highlighted.

We can point out a number of benefits of Google translate such as:

- a. It is a free online translator, which is user friendly (no registration is required).
- b. It is an online dictionary. (It translates the word automatically but displays several other translation choices of the source word in the target language. Many words that were not found in the dictionary can be found on GT.)
- c. It can be used as an online Thesaurus. (As it provides a number of translation variants, they serve as a variety of word synonym choices, i.e. a reference to the choice of meanings of the same word.)
- d. It is also a spell checker (for words that are spelled incorrectly due to typos).
- e. It is a Word Pronunciation tool. (Foreign language learners who study English on their own can benefit from Google translate as it provides the correct pronunciation of the intended word.)

Having in mind the abovementioned, some of the benefits of Google Translate are not yet well known by all users. Most users only know its main function, which is to translate faster than if they use a paper dictionary. However, some weaknesses are observed when the program is used to translate longer texts. The examples in this analysis are taken from scholarly texts which illustrate the problems a Bulgarian translator can come across. The strong points in the translation are marked in bold in the corresponding texts, while the weaknesses are underlined in the following examples.

<p>1. In this paper I argue that the lexeme time constitutes a lexical category of <u>distinct senses instantiated in semantic memory</u>.</p>	<p>В тази статия твърдя, че лексемата време представлява лексикална категория от <u>различни сетива, инстанцирани в семантичната памет</u>.</p>
<p>2. <u>The array of distinct senses</u> constitutes a motivated semantic network organised with respect to <u>a central sense termed the SANCTIONING SENSE</u>.</p>	<p><u>Масивът от различни сетива</u> представлява мотивирана семантична мрежа, организирана по отношение на <u>централно сетиво, наречено САНКЦИОННО ЧУВСТВО</u>.</p>

The translation of *argue* in the first example shows that literalness is sometimes overcome by Google translate software. *Твърдя* is not among the frequent translation equivalents of *argue* in general, but it is the only option in research papers. However, the rest of the sentence is literal translation, in which a wrong equivalent of *senses* is chosen, probably because the presence of *semantic memory* in the same clause is not taken into consideration. On the other hand, the collocation of *instantiate* and *memory* appears quite often in computer-related texts and this is reflected in the translation as the borrowing *инстанциране* in Bulgarian is only used in computer programming. The association with computer discourse is repeated in the second example where *array* is translated as *масив* which is only possible in the collocation *масив от данни* (data array). The wrong translation of *sense* persists in this example although the term *semantic network* should be indicative of the synonymy of *sense* and *meaning*. What is more, *sanctioning sense* is translated literally as *санкционно чувство*, which means nothing in Bulgarian. It is an abstract term based on a metaphor and should be translated as *определящо значение*.

<p>3. The <u>senses</u> associated with time are derived by virtue of the interaction between the Sanctioning Sense, conceptual processing and structuring, and context.</p>	<p><u>Сетивата, свързани с времето, се извличат по силата на взаимодействието между санкциониращото чувство, концептуалната обработка и структуриране и контекста</u>.</p>
<p>4. Hence, semantic representations, cognitive mechanisms, and situated language use <u>are appealed to in accounting for the polysemy</u> associated with time. The model adduced is termed PRINCIPLED POLYSEMY.</p>	<p>Следователно семантичните репрезентации, когнитивните механизми и ситуираната езикова употреба <u>са призовани за отчитане на полисемията, свързана с времето</u>. Посоченият модел се</p>

	нарича ПРИНЦИПНА ПОЛИСЕМИЯ .
--	-------------------------------------

Example 3 continues the tendency, previously set in the text, of translating *senses* as *сетива*, despite the fact that the phrase *сетивата, свързани с времето* makes no sense. Literal translation is observed in the combination of the verb *derive* and the complex preposition *by virtue of* (an acceptable variant would be *произлизат от*). Prepositions appear to be the most difficult part of speech to tackle in machine translation. Literalness is a problem in example 4 as well, because *to be appealed to* is usually used with animate creatures but here it is combined with abstract notions. The complex preposition *in accounting for* should be translated as a set phrase (*за да се обясни*); it has nothing to do with *accounting* as *auditing* and *calculating*. In the case of the term *principled polysemy* literal translation works as the corresponding term is used in Bulgarian.

5. The conclusion which emerges, in keeping with recent studies in lexical semantics, most notably <u>Lakoff (1987), Pustejovsky (1995), Tyler & Evans (2003) and Evans (2004)</u> , is that the lexicon is not an arbitrary repository of unrelated lexemes;	Заклучението, което се появява, в съответствие с последните изследвания в областта на лексикалната семантика, най-вече <u>Lakoff (1987), Pustejovsky (1995), Tyler & Evans (2003) и Evans (2004)</u> , е, че лексиконът не е произволно хранилище на несвързани лексеми;
--	---

In example 5, the translation of *in keeping with* as a set phrase clearly shows that phraseological units and idioms are included in the algorithm of Google translate. The complication seems to be that the software cannot always “decide” whether to choose the literal or the metaphorical meaning in a particular context (something which will not be difficult for a human translator). The second part of the sentence shows a problem with proper nouns, more specifically surnames which appear in the target language as they are in the source language. It sometimes happens when Bulgarian is the target language but not when it is the source language (see ex. 12 below).

Some inconsistencies are observed not only with human names but also titles (see ex. 6).

6. The most comprehensive statement of the theory resides in the hulking two-volume mass called Foundations of Cognitive	Най-изчерпателното изложение на теорията се намира в огромната двутомна маса, наречена Основи на
---	---

<p>Grammar (Langacker 1987, 1991). More accessible – or easier to lift at any rate – is <u>Concept Image and Symbol</u> (Langacker 1990), a collection of articles tailored as a single text. A second collection of this sort is Grammar and Conceptualization (Langacker 1999a)...</p>	<p>когнитивната граматика (Langacker 1987, 1991). По-достъпно – или по-лесно за повдигане – е <u>Concept Image and Symbol</u> (Langacker 1990), колекция от статии, съобразени като един текст. Втора колекция от този вид е Граматика и концептуализация (Langacker 1999a)...</p>
--	--

At first glance, there is no satisfactory explanation why two of the titles are rendered into Bulgarian by Google translate and the other one is not. A closer look shows missing commas between the words in the title, the original one being *Concept, Image, and Symbol*. This makes punctuation a very important factor for the correct translation by this software, especially when long sentences with a number of subordinate clauses need to be translated.

<p>7. In order to adduce what constitutes a distinct sense, I introduce three criteria: (1) a meaning criterion, (2) a concept elaboration criterion and (3) a grammatical criterion. A further claim is that the lexicon exhibits <u>significant redundancy</u>.</p>	<p>За да предложи какво представлява различен смисъл, въвеждам три критерия: (1) критерий за значение, (2) критерий за изработване на концепцията и (3) граматически критерий. Друго твърдение е, че лексиконът показва <u>значителна излишество</u>.</p>
<p>8. This position is at odds with <u>SINGLE-MEANING APPROACHES</u> to polysemy, which posit <u>highly underspecified</u> lexical META-ENTRIES, such as the generative approach of Pustejovsky (1995) or the monosemy position of Ruhl (1989).</p>	<p>Тази позиция е в противоречие с <u>ЕДНОЗНАЧИТЕЛНИ ПОДХОДИ</u> към полисемията, които постулират <u>силно недостатъчно</u> специфицирани лексикални МЕТА-ЗАПИСВАНИЯ, като генеративния подход на Pustejovsky (1995) или позицията на моносемия на Ruhl (1989).</p>

Although so far sense was wrongly translated as *сетиво* or *чувство*, in example 7 it is translated correctly as *смисъл* for the first time. Key words in the context which lead to this translation are *meaning, concept, grammatical criterion*. Sometimes when the source language has fewer grammatical markers than the target language, there are cases of lack of concord, especially in collocations with abstract nouns, e.g. *значителна излишество*. Bulgarian

rich morphology is the reason for the mistake in the translation of *single-meaning* (ex. 8). *Еднозначителен* is confused with *еднозначен*, which derives from *едно+значение*. There is no adjective *еднозначителен* in standard Bulgarian. It is created analogically to *многозначителен* (significant, telling). *Силно* (literal translation of *highly*) is incompatible as an intensifier of *недостатъчно*. More acceptable variants would be *изключително* and *много*.

<p>9. This necessitates a set of criteria for determining what counts as a distinct sense <u>without deriving a proliferation of unwarranted senses</u>, a criticism which has been levelled at some studies of word-meaning in cognitive linguistics (e.g. Lakoff 1987).</p>	<p>Това налага набор от критерии за определяне на това, което се счита за отделен смисъл, <u>без да се извлича разпространението на неоправдани сетива</u>, критика, която е отправена към някои изследвания на значението на думите в когнитивната лингвистика (напр. Lakoff 1987).</p>
--	---

Example 9 shows how the software translates differently the same word, *sense(s)*, within one sentence. Initially, it is correctly rendered as *отделен смисъл* and immediately after this collocation a number of words are translated literally, *deriving, proliferation, unwarranted, senses*, which results in a loss of meaning. In this case a human translator can be more analytical and eventually more successful in developing logical relations between the words in order to make the text coherent. Consequently, the debatable non-finite clause *without deriving a proliferation of unwarranted senses* could be translated as *без да води до появата на необосновани значения*.

<p>10. Research in CG began in 1976, and the basic framework of the theory has now existed for over a quarter century. <u>Under the rubric “space grammar”</u>, it was first extensively described in Langacker 1982, whose numerous and unfortunately rather crudely drawn diagrams must have startled and dismayed the readers of Language.</p>	<p>Изследванията в областта на CG започнаха през 1976 г., а основната рамка на теорията съществува вече повече от четвърт век. <u>Под рубриката „космическа граматика“</u> тя е описана за първи път подробно в Langacker 1982, чиито многобройни и за съжаление доста грубо начертани диаграми сигурно са стреснали и уплашиха читателите на Language.</p>
--	--

In example 10 we come across literalness again as regards the adverbial *under the rubric* “space grammar”. In Bulgarian *rubric* denotes not only *рубрика* but also as *заглавие* or *пасаж*, *написан с отличаващи се букви*. We can guess the algorithm here has chosen the equivalent considered as the most general in meaning. This explanation, however, does not apply to *space grammar*. No logical relation can be found between outer space and grammar. What is more, the main meanings of *space* are¹ 1. an empty area that is available to be used. 2. the area around everything that exists, continuing in all directions (*пространство, разстояние, място*). Therefore, an acceptable translation would be *под заглавието “пространствена граматика”*. As a positive feature in this example could be considered the fact that the journal is rendered with its original title in the translation. Thus, it is recognizable to Bulgarian academic readership, otherwise it can cause confusion if the name coincides with that of a Bulgarian journal.

A number of issues need to be discussed concerning Bulgarian-English translation by Google translate. If we start with proper nouns, we should say this translation program is quite good with names as it uses the resources of the internet.

<p>11. В началото на войната <u>Добруджа</u> е в периферията на интересите на Германия, най-вече с оглед на геополитическите ѝ цели за влияние в Дунавския регион и концепцията за създаването на „Средна Европа“. <u>Добруджанският</u> край е второстепенен и за политическите среди в София и Букурещ.</p>	<p>At the beginning of the war, <u>Dobrogea</u> was on the periphery of Germany's interests, especially in view of its geopolitical goals of influence in the Danube region and the concept of creating a “Central Europe”. The Dobrudzha region is also secondary to the political circles in Sofia and Bucharest.</p>
---	---

However, in some cases (see ex. 11) more than one version of a name can be found: Dobrogea, Dobrudzha, Dobruja or Dobrudja. Such instances should be corrected in the post-editing process and as *Dobrudzha* is the spelling accepted as transliteration of *Добруджа* in Bulgarian, it should be made consistent. The translation of *Средна Европа* as *Central Europe* is a good example of the fact that GT tries to avoid literalness by searching the internet for corresponding collocations.

<p>12. Анализът установява, че митологията на заселването е</p>	<p>The analysis finds that the mythology of the settlement is</p>
---	---

¹ Cambridge English Dictionary <https://dictionary.cambridge.org/dictionary/english/space>

героична, <u>в нея централна е ролята на великия княз Витаутас</u> , а предците са представени като благородни и лоялни войни.	heroic, <u>the central role of Grand Duke Vytautas is central</u> , and the ancestors are presented as noble and loyal warriors.
--	--

Example 12 reveals that the software has found the correct English variant of *великия княз Витаутас*, nevertheless, proper nouns should always be checked during post-editing. The specific syntax used in the subordinate clause *в нея централна е ролята на великия княз Витаутас* leads to confusion and the program renders a clause in which *central* is mentioned twice. This is a mistake that cannot be made by a human translator but in the machine translation can be easily corrected in the editing process.

13. Хуморът, както е известно, създава една от най-непреодолимите межкултурни бариери. <u>Кой, кога и на какво се смее</u> – това въпрос отдавна се задават психолози, антропозии, философи и социолози, лингвисти и литературоведи.	Humor is known to create one of the most insurmountable intercultural barriers. <u>Who, when and what laughs</u> - this question has long been asked by psychologists, anthropologists, philosophers and sociologists, linguists and literary critics.
--	--

Talking about complicated syntax, we should mention combining several questions in one. In Bulgarian it is not difficult because they follow the same word order, but in English it is not the case. Questions with *who* have straight word order: Who laughs? Questions asking about time have inverted word order: When do they laugh? And questions which inquire about things to laugh at need a preposition: What do they laugh at? This grammatical problem cannot be solved if we just arrange the question words one after the other. In this case we need the whole questions in English: *Who laughs, when do they laugh and what do they laugh at – these questions have long been asked...*

Impersonal constructions also lead to ambiguity and syntactic errors (see ex. 14 and 15).

14. <u>Често се цитира по този повод Хегел</u> , който казва, че не съществува нищо по-несъвместимо от нещата, на които хората се смеят...	<u>It is often quoted on this one on the occasion of Hegel</u> , who says that there is nothing more incompatible than the things that people laugh at...
15. <u>Говори се и за различни „смехови култури“</u> , за различни „смехови светове“, за „смеха като	<u>There is also talk of different “laughing cultures”</u> , of different “laughing worlds”, of “laughing as

мироглед“, <u>говори се дори за</u> „ <u>смехови нрави и обичаи</u> “.	a worldview”, <u>there is even talk of</u> “ <u>laughing manners and customs</u> ”.
---	--

These constructions are widely used in Bulgarian academic discourse and cause troubles even to human translators let alone computer programs. The best way to avoid the trouble with the impersonal constructions is to transform them into personal ones, i.e. in example 14, the sentence should start with an agentive subject: instead of *It is often quoted...*, it should be *Hegel is often quoted on this occasion...* In 15 the program follows the rule that the equivalent to the Bulgarian impersonal constructions is the English existential constructions: *there is/ are*. However, the collocation *there is talk* is infrequent in English (only 380 hits), according to COCA corpus. A better solution of this translation problem would be to use an agent as a subject: *People also talk of different “laughing cultures”... they even talk of “laughing manners and customs”*.

16. <u>Видими са</u> тенденциите на постепенно отмиране на редица традиционни елементи на народна култура и оформянето на новия градски бит.	The tendencies of gradual extinction of a number of traditional elements of folk culture and the formation of the new urban way of life <u>are visible</u> .
17. В хода на проучването са открити отделните етапи, <u>през които преминава процесът</u> на изграждане на представата за новия тип градско облекло и неговото усвояване от българските граждани.	In the course of the research the separate stages are highlighted, <u>through which the process of</u> building the idea of the new type of urban clothing and its assimilation by the Bulgarian citizens <u>passes</u> .
18. <u>Създават се условия</u> за различни нагласи и поведение, които променят всекидневния език на модерността.	<u>Conditions are created</u> for different attitudes and behaviors that change the everyday language of modernity.

The translation of 16 shows a typical sentence structure chosen by GT when rendering impersonal constructions. The subject of the clause is so long that the relation between the main word in the noun phrase and the main verb is lost. Such cases provide a suitable context for the abovementioned *there is/ are* construction, e.g. *There are clear tendencies of gradual extinction of a number of traditional elements of folk culture and the formation of the new urban way of life*. Similarly in 17, a gap opens up between the head noun in the heavy noun phrase, *the process*, and the verb *passes*. The translation made by GT could be edited in two steps; the first one would be to change the passive

voice of the main clause into active, and the second to change the place of the verb *passes* or omit it altogether if it is possible. Here is a suggestion of post-editing: *In the course of the research we highlight the separate stages in the process of building the idea of the new type of urban clothing and its assimilation by the Bulgarian citizens.*

As for example 18, no such gap exists between the subject and the verb, but the literal translation, *conditions are created*, suggested by GT, is very limited in English (only 10 hits in COCA corpus). A possible solution would be to rearrange the beginning of the sentence: *The new conditions create different attitudes and behavior that change the everyday language of modernity.*

Bulgarian is gender-sensitive language, not only as far as nouns are concerned, but adjectives are also marked for gender. Sometimes in a complex sentence, GT fails in the concord of the words (see 19).

<p>19. Независимо от обстоятелството, че още първите изследвачи на българската градска култура, включват в кръга на своите научни занимания и въпросите на градското облекло, <u>неговото изследване има по-скоро публикаторски характер.</u></p>	<p>Despite the fact that the first researchers of Bulgarian urban culture included in the scope of their research activities and issues of urban clothing, <u>his research is more of a publishing nature.</u></p>
---	--

Example 19 has two subordinate clauses, which makes it difficult for GT to keep the agreement between the parts of speech. In this case the program offers literal translation of *неговото изследване има по-скоро публикаторски характер*, which is rendered as *his research is more of a publishing nature*. In Bulgarian *неговото изследване* refers to *градското облекло* (*urban clothing*), *clothing* is neuter, therefore, the translation should be *its research is meant to be published*.

When post-editing, we should be cautious about homonyms in the source text, because it is difficult for GT to distinguish between such words only on the basis of context (see 20).

<p>20. Основната цел на този труд е да се въведе в научен оборот един почти неизползван и слабо проучен <u>масив от информация</u>, който се съхранява във фондовете на Регионален исторически музей – Шумен(...) Част от този <u>масив</u>,</p>	<p>The main goal of this work is to introduce into scientific circulation an almost unused and poorly studied <u>array of information</u>, which is stored in the funds of the Regional Historical Museum – Shumen(...) Part of this <u>massif</u>,</p>
--	---

<p>предимно снимков материал и документи с научно-историческо значение за развитието на Шумен, се пази във фондовете на Държавен архив - Шумен.</p>	<p>mainly photographic material and documents of scientific and historical significance for the development of Shumen, is kept in the funds of the State Archives - Shumen.</p>
---	---

In 20, the collocation *масив от информация* suggests that GT should choose the set phrase *array of information*. However, in the next sentence, *масив* is used without a modifier, which leads to ambiguity for the machine translation program and it wrongly singles out the geographical term *massif* that has nothing to do with information.

When translating terms, there is usually one-to-one correspondence between English and Bulgarian. But, in some rare cases, we come across “false friends” (see 21).

<p>21. В основните си елементи (сюжет, герои) татарската митография следва траекторията на <u>доминиращия национален наратив</u> като подчертава героичното и достойно татарско участие в него.</p>	<p>In its main elements (plot, characters) Tatar mythography follows the trajectory of <u>the dominant national narrative</u>, emphasizing the heroic and dignified Tatar participation in it.</p>
---	--

Narrative is a term used in the literary theory and criticism, as well as in philosophy, semiotics and aesthetics. In the Bulgarian language it is a borrowing that has preserved the older meaning of *narrative*, which in its most general sense is a cover term and refers to “all those literary works which are distinguished by two characteristics: the presence of a story and a story-teller” (Scholes, Phelan, Kellogg 4). *Narrative* is used differently in French and English literary terminology nowadays - in French as an adjective meaning “narrative”, and in English as a text characterized by an objective description of “a series of facts or events and the establishing of some connection between them” (Childs, Fowler 148), i.e. a story or a description of a series of events². In Bulgarian literature the term *нарация* is used, which is closer to the English *narrative* (*повествование*).

<p>22. В повечето от случаите направените у нас проучвания на облеклото обхващат различни</p>	<p>In most of the cases the researches of the clothes made in our country cover different time periods,</p>
---	---

² Cambridge English Dictionary <https://dictionary.cambridge.org/dictionary/english/narrative>

отрязъци от време, в зависимост най-вече от <u>наличния документален и веществен материал</u> , с който разполагат отделните музейни фондове.	depending mostly on <u>the available documentary and material material</u> , which the separate museum funds have.
23. Освен утилитарното ѝ предназначение, тя има <u>особено знаково значение</u> .	Apart from its utilitarian purpose, it has <u>a particularly significant significance</u> .

Examples 22 and 23 show cases of tautology, whose elimination is obviously not part of the program’s algorithm. This redundancy, therefore, should be corrected in the process of post-editing. In 22 *the available documentary and material material* can be paraphrased as *the available documentary and material evidence*, while in 23 *a particularly significant significance* can be reworded as *considerable significance*. Yet, repetition in the source text is sometimes avoided in the target text simply by omission (see ex. 24).

24. В подкрепа на тези разсъждения са разглежданите колекции от костюми на градското облекло съхранени във фондовете на Регионален исторически музей - Шумен, които налагат редица важни изводи и заключения .	In support of these considerations are the examined collections of costumes of urban clothing stored in the funds of the Regional Historical Museum - Shumen, which impose a number of important conclusions .
25. Модата налага нов вид обувки, които могат да се закупят от „кундуржийските заведения“ (местните работилници-магазини) в града.	Fashion imposes a new type of footwear that can be purchased from the “kunduri establishments” (local workshops) in the city.

In 24 the collocation *важни изводи и заключения*, in which *изводи* and *заключения* are synonyms, is translated only as *important conclusions*, because it stands for both *изводи* and *заключения*. Similarly, in 25 the noun phrase *местните работилници-магазини* is rendered by GT as *local workshops* to prevent the repetition *workshops-shops*. However, in this case the American word *store* can be used instead of *shop* to express the same meaning.

Conclusion

Google translate is an efficient tool for quick translation. It is a handy tool, because it is free and readily available for understanding a text with unknown vocabulary items. The software has improved significantly in the last five years and it is under constant development. Although it cannot replace human translators, it requires less and less post-editing. However, the EFL teacher has to be cautious about the appropriate use of Google translate to avoid students' overuse of this tool to an extent which would hamper language learning and academic development. Language instructors should ensure an analytical approach to the target text asking their students to comprehend, associate and compare the linguistic features of the source and target language. They do not need to be intimidated by technical tools such as Google translate. "As long as we accept this technology and try to work with it, it has the potential to make the teaching... much more exploratory and critical activity" (Groves and Mundt 120).

We suggest that GT could be beneficial to human translators, provided the GT's output is approached critically and is post-edited, i.e., editors need an advanced level of language knowledge to notice the mistakes made by the machine and to correct them. This study shows some common errors students can come across when using GT such as literalness, semantic and grammatical mistakes, homonyms and synonyms, false friends and tautology, etc. Teachers should show their students how to use GT effectively rather than forbidding its use. Students need to be able to check the target text for accuracy, cohesion and quality of translation. Even lower level students can use it if they enter single lexical items.

With idiomatic expressions, GT may provide either the literal or the idiomatic translation according to its algorithm and the data it uses as resources. Advanced students can use GT for sentences as well as words and phrases. Instructors should show their students that GT in some cases provides alternative translations in a popup, and the alternative may be more suitable than the first translation variant.

Teachers may prepare a correction task, in which students can work either individually, in pairs or in groups on the same target text and edit it. Then, the edited translation can be discussed, as well as the kinds of mistakes and how to correct them. This activity can constitute preparation for the use of GT in translation classes.

Works Cited

- Aiken, Milam. An Updated Evaluation of Google Translate Accuracy. *Studies in Linguistics and Literature*, Vol. 3, No. 3, 2019: 253-260. p-ISSN 2573-6434.
- Childs, Peter, Roger Fowler. *The Routledge Dictionary of Literary Terms*. London and New York: Routledge Taylor & Francis Group, 2006.

- COCA corpus <https://www.english-corpora.org/coca/>
- Groves, Michael, Klaus Mundt. “Friend or foe? Google translate in language for academic purposes.” *English for Specific purposes*, 37, 2015: 112-121.
- Farzi, Reza. *Taming translation technology for L2 writing: documenting the use of free online translation tools by ESL students in a writing course*. Doctoral dissertation, University of Ottawa, Canada, 2016.
- Fitria, Tira Nur. “Gender Bias in Translation Using Google Translate: Problems and Solution.” *Language circle: Journal of Language and Literature*, 15(2) April 2021: 285-292. p-ISSN1858-0165
- Kol, Sara, Miriam Scholnik, Elana Spector-Cohen. “Google Translate in Academic Writing Courses?” *The EUROCALL Review*, Volume 26 (2), 2018: 50-57.
- Lee, Sangmin-Michelle. “The impact of using machine translation on EFL students’ writing.” *Computer Assisted Language Learning*, 33, 2020: 157-175.
- Lopez Medel, Maria. “Gender bias in machine translation: an analysis of Google Translate in English and Spanish.” *Academia Letters*, Article 2288, 2021: 1-8.
- Newmark, Peter. *A Text Book of Translation*. New York: Prentice Hall, 1988.
- Prates, Marcelo O. R., Pedro H. C. Avelar, Luís C. Lamb. Assessing gender bias in machine translation—a case study with Google Translate. ArXiv:1809.02208 [Cs], 2019.
- Raza, Muhammad, Faizah Nor. “Google Translate in an EFL Classroom.” *International journal of translation*, Vol. 30, No. 1, 2018: 7-21.
- Riadi, Agus, Yovita Angelina. An analysis of literalness aspect of Google Translate in translating business correspondence. *English Language Studies and Applied Linguistics Journal*, 1(1), 2020: 1-14.
- Scholes, Robert, James Phelan, Robert Kellogg. *The Nature of Narrative*. Oxford: Oxford University Press, 1966.
- Sitorus, Thiar YT. Utilization of Google Translate as a Translation Media for English Language Materials, 2020.
- Stanovsky, Gabriel, Noah A. Smith, Luke Zettlemoyer. “Evaluating Gender Bias in Machine Translation.” *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics, 2019: 1679–1684.
- Tsai, Shu-Chiao. “Using Google Translate in EFL drafts: A preliminary investigation.” *Computer Assisted Language Learning*, 32, 2019: 510-526.
- Wu, Yonghui, et al. Google’s neural machine translation system: Bridging the gap between human and machine translation, 2016.