Effects of pre-editing operations on audiovisual translation using TRADOS: an experimental analysis of Saudi students' translations

Efeitos das operações de pré-edição na tradução audiovisual usando o TRADOS: uma análise experimental das traduções de estudantes sauditas

Chaouki Bounaas [iD]*[1], **Bahia Zemni** [iD]†[2], **Fadiah Al Shehri** [iD]‡[2] and **Mimouna Zitouni** [iD]§[2]

[1]*University Mohammed Boudiaf of M'sila, Faculty of Letters and languages, Department of Letters and English Language, M'Sila, Algeria.*
[2]*Princess Nourah bint Abdulrahman University, College of Languages, Translation Department, Riyadh, Saudi Arabia.*

Abstract

Neural Machine Translation (NMT) is a revolutionary innovation that has had a significant impact not only on the translation industry but also on translation studies. This technology is creating new opportunities for shaping identities and generating new knowledge and perspectives. This led researchers to undeniably and increasingly conduct studies to improve NMT performance. These studies involve assessing the quality of the output, determining the extent to which humans can be involved in the machine-based process, and proposing solutions to redress the deficiencies. In this regard, the paper focuses on the impact of pre-editing (PE) operations on the translation of audiovisual children's literary texts (subtitles) from English to Arabic using TRADOS. The work seeks to accomplish a number of pre-editing-related research goals in the context of Arabic NMT systems. Initially, it aims to define and clarify the idea of PE and its function in enhancing Arabic NMT systems' performance. Second, it looks at PE mechanisms and methods and how they might improve the appropriateness and accuracy of Arabic machine translation. Additionally, the study attempts to investigate the value of PE in enhancing the Arabic translation of audiovisual children's literature. The study also seeks to list the difficulties and restrictions related to PE and suggest viable strategies to get over them. Finally, the paper offers suggestions for further investigation into NMT systems using Arabic and PE. To achieve these goals, the researchers conducted a human evaluation of NMT for two animated movie subtitles and determined the necessary PE operations. The findings revealed that PE has the potential to improve the quality of literary translations and enhance their comprehensibility and acceptability.

Keywords: Pre-editing. NMT systems. TRADOS. Literary translation. Audiovisual translation.

Resumo

A Tradução Automática Neural (TAN) é uma inovação revolucionária que teve um impacto significativo não só no setor da tradução, mas também nos estudos de tradução. Essa tecnologia está criando novas oportunidades para moldar identidades e gerar novos conhecimentos e perspectivas. Esse fato levou a que os investigadores realizassem cada vez mais estudos para melhorar o desempenho das TAN. Tais estudos envolvem a avaliação da qualidade dos resultados, a determinação da medida em que os seres humanos podem ser envolvidos no processo baseado na máquina e a proposta de soluções para corrigir as deficiências. Nesse contexto, o artigo centra-se no impacto das operações de pré-edição (PE) na tradução de textos literários audiovisuais infantis (legendas) de inglês para árabe utilizando o TRADOS. O trabalho procura atingir uma série de objetivos de investigação relacionados com a pré-edição no contexto dos sistemas TAN árabes. Em primeiro lugar, pretende-se definir e esclarecer a ideia de PE e a sua função na melhoria do desempenho dos sistemas TAN árabes. Em segundo lugar, analisam-se os mecanismos e métodos de PE e a forma como podem melhorar a adequação e a precisão da tradução automática árabe. Além disso, o estudo procura investigar o valor da PE na melhoria da tradução árabe de literatura infantil audiovisual. O estudo procura também enumerar as

---
*Email: chaouki.bounaas@univ-msila.dz
†Email: baalzemni@pnu.edu.sa
‡Email: fadiah2425@gmail.com
§Email: MBZitouni@pnu.edu.sa

dificuldades e restrições relacionadas com a PE e sugerir estratégias viáveis para as ultrapassar. Por fim, o documento apresenta sugestões para futuras investigações sobre sistemas de TAN que utilizam o árabe e a PE. Para atingir esses objetivos, os investigadores realizaram uma avaliação humana de TAN para legendas de dois filmes de animação e determinaram as operações de PE necessárias. Os resultados revelaram que a PE tem potencial para melhorar a qualidade das traduções literárias e aumentar a sua compreensibilidade e aceitabilidade.

Palavras-chave: Pré-edição. Sistemas de TAN. TRADOS. Tradução literária. Tradução audiovisual.

1 Introduction

Recent progress in Neuronal Machine Translation (NMT) has greatly helped human translators to be more effective and efficient. However, this does not eliminate the human role. A process like PE can significantly assist NMT software programs in approximating human-generated outputs. Pre-editors can modify an ST into a text that can be machine-translated with sufficient quality (MIYATA; FUJITA, 2021). The process may involve a wide range of linguistic and extra-linguistic adjustments at a grammatical or cultural level among others.

The effectiveness of PE has so far been demonstrated in many studies (PYM, 1990). However, the feasibility and possibility of PE for NMT software using Arabic has not been examined extensively. In fact, the current MT systems using Arabic such as, among other, SYSTRAN and Google Translate suffer from serious shortcomings. They produce inappropriate and inconsistent source-target equivalences (JABER, 2015). Some features such as being written from right to left, the no existence of capitalization or gender-neutral pronouns, the existence of duality and diacritics, or the flexibility of word order can explain that (IZWAINI, 2015). In short, while PE can easily be implemented in other languages' NMT cases, what PE is and how it works with Arabic NMT systems remain open questions.

The relationship between PE and Audiovisual Translation (AVT) has not received much attention in the Arab world. Despite the significant growth of audiovisual content, thanks to the increase in satellite channels and the expansion of video content on social media, Arabic translation departments do not cover certain aspects of AVT (GAMAL, 2007, 2014, p. 80, p.1). With the exception of the American University of Cairo initiative two decades ago and some training programs offered by the University of Doha and the University of Balamand in Beirut and a limited number of universities, Arab countries have shown little interest in audio-visual translation. As a result, Arab audio-visual translators often subtitle and dub foreign films and programs without mastering technical issues or software design (JABER, 2015). On the other hand, research in the field of AVT has not evolved and has perhaps been limited to some basic aspects with no consideration of new technologies and their impact on audiovisual translators' workflow. Given this situation, it is essential to explore how PE can assist in the machine translation of audio-visual material into Arabic, ensuring accuracy and appropriateness.

Finally, yet importantly, the impact PE would have on the translation of literary texts in general and children's literature, in particular, is no exception. A quick scan of research engines about the topic, Google, for instance, does not show much result, which is rather ironic. There is no doubt that PE operations can help Artificial Intelligence (AI) provide outputs that are more sophisticated. The translation of songs, poems, fairy tales, comics, cartoons, or animated movies necessitates more than a coding-decoding process. Such texts have a distinct style and an elegant language combined in a creative way to provide life-paintings full of scientific, cultural, and religious details (BOUNAAS; BEDJAQUI, 2022). Children's literature not only captures their imagination but also imparts moral values in a literary style that is specific to their age, historical period, and social environment (IBROKHIMOVICH et al., 2022). A translator must consider a child's language abilities to understand and analyze the text, along with other stylistic, syntactic, and cultural aspects (IBROKHIMOVICH et al., 2022). This raises the question of how PE can affect the translation of children's literature and at what level of operation.

To investigate the impact of PE on NMT of audiovisual texts, subtitles in particular, this study

conducts a detailed analysis of human PE practices and their effects on output quality. The paper is structured around examining PE practices and challenges in general, as well as specific issues encountered with audiovisual literary texts. Furthermore, the study aims to demonstrate how PE can enhance the technological competence of Arab translators, particularly Saudi, translators, and students (TOMASZKIEWICZ, 2018, p. 77), by equipping them with the skills to improve machine translation:

- Master the basics of MT and its effects on the standard translation workflow and process (TOMASZKIEWICZ, 2018, p. 78) especially for translation of multimedia and audiovisual content;
- Know when to use the technology, when it is efficient, and also what to focus on during the PE process. (LOOCK, 2020, p.151);
- Help NMT developers improve Arabic software programs performance as MT systems cannot be improved without evaluating the translated output (RIVERA-TRIGUEROS, 2022).

It is important to mention here that a recent survey of 59 Saudi students of translation showed that only 46.7% of them do know PE (mean=2.03), and only 43.3% (mean= 2.25) know how to perform PE operations. 45.8% (mean=1.71) of respondents strongly agree on the importance of PE to ensure translation quality.

2 Methods

To establish the paper's rationale, we collected instances. As subtitles are one or two lines placed near the bottom of the screen (KARKABOU, 2019), the selected corpus includes instances with no more than two lines in length. The instances are YouTube auto-generated scripts for a video including extracts from three animated movies broadcasted on the same platform: *Monsters, Inc., Rise of the Guardians*, and *Robots*[1], which are then translated using the NMT system of TRADOS. It is worth mentioning that TRADOS, besides the fact of being a popular computer-assisted translation (CAT) software with project management functions (WANG, 2016), was chosen for different reasons. First, the university within which the study was carried out provides a TRADOS license key which enables the researchers to collect the needed data. The use of this program is part of a project the University has launched in order to create Translation Memories that can be used later in some didactic contexts. Finally, using CAT tools and TAV are two fertile fields that still need more research and investigation and combining them can reveal interesting and insightful results.

After collecting data, human editing is, then, performed fully or minimally by a group work of three Saudi master's students (after obtaining an approval from the responsible body) at different levels: linguistic, lexical, and cultural. Data are the subject of in-depth analysis (analytically and critically) from three perspectives: the translation of the raw ST, the performed PE operations, and the impact of PE operations on the NMT outputs. The outputs of the MT of the pre-edited texts will be used as didactic support to help and accompany the students in understanding and learning AVT and in mastering its different aspects, especially with regard to subtitles. The research adopts human evaluation to assess the quality of MT output in terms of adequacy and fidelity (AFZAAL et al., 2022, p. 2). In this study, the human evaluation is conducted with Arab native speakers who are also lecturers and practitioners of translation.

3 Literature review

Machine translation (MT) systems are not able to produce translations of human-level quality all the time. As supporting processes to boost operating MT systems, pre-editing, and post-editing seem to be very practical to bridge this gap and produce high-quality translations. To achieve such an objective, many researchers consecrate considerable efforts and time to enlighten on the various aspects of Neural Machine Translation (NMT) in general and the former two concepts in particular. However, it is noticeable that PE has taken less of their attention.

In his paper entitled "How MT errors correlate with post-editing effort: a new ranking of error types", Ke Hu (2020) reported on a preliminary study that examines the correlation between translation

---

[1] Available at: https://www.youtube.com/watch?v=Ruw5l8CumIw

error types and cognitive post-editing effort using empirical data on participants' pauses in the post-editing process. The study revealed that cognitive effort expended on an error is positively correlated with the stretch of the text in which the error is embedded.

In 2021, Miyata and Fujita (2021) tried in their paper entitled "Understanding Pre-Editing for Black-Box Neural Machine Translation" to investigate human PE practices. They first implemented a protocol to incrementally record the minimum edits for each ST and collect 6,652 instances of PE across three translation directions, two MT systems, and four text domains. The researchers analyzed the instances from three perspectives: the characteristics of the pre-edited ST, the diversity of PE operations, and the impact of the PE operations on NMT outputs. The findings include enhancing the explicitness of the meaning of an ST, its syntactic structure is more important for obtaining better translations than making the ST shorter and more straightforward, and although the impact of PE on NMT is generally unpredictable, there are some tendencies of changes in the NMT outputs depending on the editing operation types.

For his part, Taufik (2020) studied the concept in his paper entitled "Pre-Editing of Google Neural Machine Translation" and published it in 2020 to increase the efficiency and effectiveness of MT, especially for the language pair Indonesian-English. The researcher adopted product-oriented research. The results show that in the PE process, the length of the sentence, the conjunctions (subordinative and correlative), and the inappropriate ST words should be the focus of attention. For that reason, the research intended to identify the PE rules required to create a solid basis to translate Indonesian Source Texts (ST) into English Target Texts (TT).

In her paper entitled "Traduction automatique dans la formation des traducteurs: une analyse expérimentale de la post-édition" or "Automatic Translation in Translators Training: an Experimental Analysis of the Post-editing", Tomaszkiewicz (2018) analyzed in 2018 the evolution of the role of the translator in a world dominated by new technologies and the influence of these on the training of future translators. She focuses primarily on machine translation and post-editing. The purpose is illustrated by the results of experiments carried out within the framework of a program belonging to the EMT network and whose objective is to compare the steps of rewording in the target language by post-editing and by a human translation exclusively. These experiments make it possible to draw some conclusions concerning the training of students integrating machine translation and post-editing.

When Translation Goes Digital: Case Studies and Critical Reflections was a book edited by Desjardins, Larsonneur, and Lacour in 2021. The book brings a series of case studies connected by the overarching theme of translation in digital contexts. It examines various facets of the translation landscape and industry and includes works on non-Western languages such as Japanese, Korean, and Arabic.

For his part, Wang (2016) presented in 2016 a modified approach to the translation process management and quality control, especially in terminology and corpus management using computer-assisted translation software TRADOS, on Energetic Materials at Extreme Conditions translation project. The study highlighted the quality of the target text and efficiency of translation by calculations of comprehensive error rate and drawing of the Gantt Table, in the hope of promoting computer-assisted translation and further, establishing a valuable database for future projects.

In 2022, Rivera-Trigueros (2022) focused on her study entitled "Machine translation systems and quality assessment: a systematic review" on the specialized literature produced by translation experts, linguists, and specialists in related fields that include the English–Spanish language combination. Research findings show that neural MT is the predominant paradigm in the current MT scenario, with Google Translator being the most used system. Moreover, most of the analyzed works used one type of evaluation – either automatic or human – to assess machine translation and only 22% of the works combined these two types of evaluation. However, more than half of the works included error classification and analysis, an essential aspect for identifying flaws and improving the performance of MT systems.

Last but not least, Ehrensberger-Dow and Massey (2019) confirm in their « Le traducteur et la machine: Mieux travailler ensemble? » or "The translator and the machine: Better to work together?" published in 2019, that the use of technologies has led to an interest in their impact on cognitive

processes and translations. Based on data from professional translators and students, the authors question tool translation as a cognitive and organizational activity. They recommend that translators take greater ownership of translation support tools, from their design to their integration into the organization of the activity.

Overall, it can be said that the feasibility and possibility of PE for neural MT (NMT) has not been examined extensively. That is why the present study invests in clarifying PE tools and strategies and enlightening on potential gains PE can add to NMT.

4 A glance on theory

4.1 MT and NMT

In today's life, translators' workflows know gigantic disruption due to the huge proliferation of contents and the growing need for translating them. MT emerges then as a solution due to its ability to reduce effort and cost. It equips users with great potentials for solving this problem in terms of research and from the perspective of professional settings (RIVERA-TRIGUEROS, 2022, p. 594). MT first adopted different approaches; among them are the Rule-Based Machine Translation (RBMT), and Statistical Machine Translation (SMT). Nevertheless, the last decade witnessed the launch of Neural Machine Translation (NMT).

As for NMT, "revolutionary innovations in the computational architectures made in 2014–2017 have led to dramatic improvements in the quality of machine translation and transformed the field forever" (BALASHOV, 2022, p. 6). Based on the encoder-decoder model, NMT models use artificial neural networks composed of increasingly complex and interconnected layers of basic feed-forward and recurrent units, or "neurons" (RIVERA-TRIGUEROS, 2022). They use distributed, parallel information processing to perform calculations. Inspired by human brain functioning, more specifically human neurons interconnection system, neurons receive signals from the environment from other neurons, process them and send them as input signals for the neurons in their surroundings. Knowledge is stored primarily through the strength of links between individual neurons (BENKOV, 2020, p. 501). In other words, NMT looks at the sentence as a whole and can form associations between phrases even at greater distances in the sentence. The result is then more accurate compared to SMT (BENKOV, 2020, p. 499). In addition, NMT also uses sub-word segmentation algorithms as a solution for rare and unknown words traditionally machine-untranslatable (BALASHOV, 2022).

Accordingly, NMT is considered the finest advancement technology has ever given. That is why, Google, Microsoft, Facebook, Amazon, SDL, Yandex, and many more have deployed NMT in their production of systems (STAHLBERG, 2020, p. 343). Loock (2020, p. 150) asserts that reports such as the one of the 2018 European Language Industry Survey show that for the first time, more than half of European translation companies now use MT. Ironically, this is not the case for Arabic and Arab countries. The lack of investment is for something. Zantout and Guessoum argue that: "research and development of Machine Translation and computational linguistics for Arabic has remained limited with almost no involvement of governmental institutions to support it" (2000, p. 118 apud Jaber (2015, p. 135). For their part, leading firms do not show much interest in using Arabic. Some web applications and social platforms, Instagram for instance, were unable to support right-to-left (RTL) languages for a rather long time; it took it seven years to enable Hebrew, Farsi, and Arabic on the platform (DESJARDINS; LARSONNEUR; LACOUR, 2021, p. 8). The reason may also lay in the nature of the language itself. Habash asserts that Arabic language is characterized by the complexity of its grammar, rich morphism, and lack of corpora. It has a very rich morphology characterized by a combination of templatic and affixational morphemes, complex morphological rules, and a rich feature system (2007, p. 263 apud Izwaini (2015). In addition, in Arabic Grammar Error Correction (AGEC), the largest parallel dataset is 20% less than a version of English parallel corpus. In addition, the total number of unique words in Arabic language is twice the number of words in English language (SOLYMAN, 2021, p. 304). This resulted in MT systems offering flawed outputs and inappropriate and inconsistent translations from and into Arabic.

4.2 Pre-editing

PE is the process of modifying the source text (ST) to be translated in order to obtain better translations by MT (MIYATA; FUJITA, 2021). The translator uses a set of tools to modify his text; he proceeds to a number of operations to correct mistakes, remove errors, simplify structures, or shorten long sentences. Pre-editors may also modify meanings by replacing some ambiguous expressions, revealing the implicit or softening the forbidden. Here one can say that the nature of PE operations depends on the type of the ST. Literary texts, for example, mostly need the pre-editor to take stylistic, semantic, or cultural decisions since ideas are sometimes presented implicitly or with a stylistic device that the reader cannot perceive, and they must be carefully and faithfully transmitted. Nevertheless, this does not eliminate the need for vocabulary and grammar checking. For technical texts, pre-editors are more concerned with controlled language choices (PYM, 1990). In this context, Guerberof Arenas (2020) argues that controlled language refers to using certain rules (vocabulary and grammar) applied when writing technical texts. The objective is to avoid ambiguity and complexity, thus making the task of translation easier for the machine and cheaper for the translator, the client, and the customer. As for controlled language rules, Marzouk and Hansen-Schirra (2019, p. 184) proposed the following:

- Using straight quotes for interface texts.
- Avoiding light-verb construction.
- Formulating conditions as if sentences.
- Using unambiguous pronominal references.
- Avoiding participial constructions.
- Avoiding passives.
- Avoiding superfluous prefixes.
- Avoid omitting parts of the words.

For their part, Hiraoka and Yamada (2019, p. 66) demonstrated the effectiveness of the following three PE rules in improving Japanese-to-English TED Talk subtitles:

- Inserting punctuation;
- Making implied subjects and objects explicit.
- Writing proper nouns in the target language (English).

It should be noted that these rules output better translation with RBMT, statistical MT (SMT), and hybrid systems, but did not have positive effects on the NMT system. Besides, the languages' combinations were German-English in the first case and English-Chinese in the second, this means that the two propositions cannot represent a holistic model applicable to the other languages, especially Arabic. For that matter, the present study is conducted in order to propose rules for Arabic-controlled language checkers.

It is worth mentioning that Hiraoka and Yamada (2019) distinguished PE into different types. It can be bilingual when the pre-editor uses the ST while correcting the MT output; or monolingual in the contrary case. PE can also be manual or automatic PE. The first type fits RBMT since this technology is more predictable and controllable. The second aims at simplifying sentences to make them more machine-translatable (MIYATA; FUJITA, 2021). After an overview of the different aspects of PE and MT, we will proceed in the following section to the analysis of some translated examples and the verification of the different operations necessary to well render the meanings.

5 Results

For reminder reasons, the paper highlights how PE operations affect the translation of audiovisual children's literary texts from English to Arabic using TRADOS. To demonstrate the effects, we focus in this section on the translation of 18 examples to be examined.

In the present analysis of the examples, the data is coded as follows:

- RST (Raw Source Text)
- TRST (Translation of Raw Source Text)
- P-EST (Pre-Edited Source Text)
- T P-E ST (Translation of Pre-Edited Source Text)

- **Example 1 (Rise of the Guardians: 22:54)**
- **RST**: you think I need help to beat a bunny
- **TRST**: تعتقد أنني أحتاج إلى مساعدة للتغلب على الأرنب
- **P-EST**: do you think I need help to beat a bunny
- **TP-EST**: هل تعتقد أنني بحاجة إلى مساعدة للتغلب على الأرنب

In this analysis, we can see that the sentence was subjected to PE, which involved fixing a grammatical error and changing the sentence structure to reflect the meaning that was intended in the original language. In this instance, the auxiliary verb "do", which is required to form a question, was omitted. This omission is often observed in informal language, particularly in American English. This error could potentially cause confusion for the target audience, who may not be able to understand the intended meaning of the sentence. This is reasonable given that the sentence is part of a dialogue involving a monster and his companions. Nonetheless, this omission had an impact on the machine's output. The intended meaning of the sentence in the source language was that of a question, but the machine's output presented it as a statement of confirmation. To address this issue, the student added the missing auxiliary verb, resulting in the inclusion of the interrogative marker "هل" or *"hel"* in the pre-edited version of the sentence, which typically elicits a "yes" or "no" response. The PE operation performed in this example is grammatical and semantic in nature, as it corrects the grammatical error while also removing the ellipsis. Therefore, the addition of the interrogative marker not only corrected the grammatical error but also made the sentence clearer and more appropriate for the intended meaning.

- **Example 2 (Rise of the Guardians: 27:00)**
- **RST**: Easter is new beginnings
- **TRST**: عيد الفصح هو بدايات جديد
- **P-EST**: The feast is new beginnings
- **TP-EST**: العيد هو بداية جديدة

In the second example, the term "عيد الفصح" which is equivalent to Easter in Christian culture, was produced by the machine translation, but it may not have been obvious or appropriate for the target audience, which in this case was presumed to be Muslim. It is widely known that Easter is a Christian religious holiday that commemorates the resurrection of Jesus Christ, as documented in the Christian New Testament. In contrast, Islamic culture maintains the belief that Jesus never actually died, but was instead lifted up to heaven by Allah and will return to Earth prior to the end of time. The pre-editor, therefore, had to take into account the cultural and ideological differences between the source and target languages to ensure accurate and appropriate communication. In other words, while the translation is technically correct, it presents a cultural and ideological problem. To address this issue, the student replaced "Easter" with "Feast" in the source text. While the term "Feast" may not fully capture the meaning of "*Easter*" in the source text, it should be noted that "*Feast*" may still carry cultural significance in the target culture, as Muslims observe two sacred feasts annually, Eid al-Fitr and Eid al-Adha. The pre-editor's decision to use the more general term "Feast" instead of "Easter" helps to avoid cultural and religious insensitivity that may result in miscommunication and offense. Additionally, the PE process in this instance was performed on a cultural level, which entails resolving not just grammatical and semantic issues but also cultural and ideological disparities. To make sure that the message delivered in the target language is correct, relevant, and culturally sensitive to the target audience, PE at a cultural level is crucial. In conclusion, our analysis emphasizes the significance of PE in addressing the issues with machine translation and the requirement for culturally relevant PE to promote successful cross-cultural communication.

- **Example 3 (Rise of the Guardians: 23:10)**
- **RST**: no stop that's the easter bunny
- **TRST**: لا توقف ذلك هو أرنب عيد الفصح
- **P-EST**: No... stop... that is the sheep of the feast
- **TP-EST**: لا.توقف... هذه هي خراف العيد

Similar to the previous case, the student conducted a cultural PE operation by replacing the term "Easter bunny" with "sheep of the feast" to demonstrate respect for the target culture, while also

taking into account the cultural modifications made previously. The concept of the "Easter bunny" is not universal and may not be familiar to the target culture. Therefore, the student replaced it with "sheep of the feast", a term that is more relevant to the cultural context of the target audience. Moreover, the PE operation also involves the grammatical aspect, as the student added ellipses to accurately reflect the speaker's hesitation and uncertainty. The use of ellipses in this context also helps convey the intended meaning of the original English text. It is important to note that the NMT system of TRADOS has translated "sheep" in its plural form since it has an irregular plural. The machine translation could not determine whether the intended meaning was singular or plural. Additionally, the student corrected the grammatical mistake in the output, which translated "هو" to "that". The student changed it to "هذه", which is the feminine form of "this" in Arabic and is more appropriate in this context. Overall, the PE operation carried out by the student is a noteworthy example of how to address cultural and linguistic differences in audio-visual translation, ensuring that the message is communicated accurately and respectfully to the target audience.

- **Example 4 (Monsters, Inc.: 0:09)**
- **RST**: sorry michael I didn't see you
- **TRST**: عذرا ، لم أراك هذا
- **P-EST**: sorry michael I didn't see you
- **TP-EST**: آسف مايكل لم أرك

The student observed that the machine has a tendency to overlook or preserve non-capitalized proper nouns written in the Latin alphabet. However, there are some cases where the absence of capitalization does not have any impact on the translation. In the example cited above, the proper noun "*michael*" was excluded from the TRST. However, when the student capitalized the first letter, the machine was able to generate the appropriate Arabic equivalent. This highlights the need for capitalization (grammatical pre-editing operation) to be performed before using TRADOS.

It should be noted that machine translation should not be depended on solely because it does not always precisely capture the intended meaning and subtleties of the source text. Human translators and editors are still required to achieve the greatest level of translation quality, particularly when dealing with highly specialized or sensitive material.

- **Example 5 (Monsters, Inc.: 0:15)**
- **RST**: no biggie bailey
- TRST: لا يوجد بايلي biggie
- P-EST: It's nothing bailey
- TP-EST: لا شيئ بيلي

The student's PE operation in this case highlights the importance of considering the cultural and linguistic context of the target audience. Idiomatic expressions and slang often do not have a direct equivalent in the target language and can be difficult for machine translation to accurately capture. Therefore, PE operations that simplify or clarify the meaning of the source text can be useful in producing a more accurate translation. A good case in point is this particular example. The idiomatic expression (slang) "*no biggie*" – meaning "not a big deal" – was not translated in the TRST. The NMT system preserved the expression as it was written in the source language. To address this issue, the student carried out a PE operation on the source text by replacing the expression with a simpler phrase "*It's nothing*", which carries the same meaning. This PE operation proved to be effective in producing an accurate translation.

- **Example 6 (Monsters, Inc.: 5:30)**
- **RST**: fellas
- **TRST**: فلما
- **P-EST**: Hey guys
- **TP-EST**: يا رفاق

The program was unable to recognize the informal urban term "*fellas*", which is a colloquial substitute for "*fellow*". Instead, it translated the term to "فلما" or "*felma*", which has no meaning in Arabic, causing a cultural and linguistic mismatch. To address this issue, the student conducted a PE operation focused on both cultural and lexical considerations by replacing "*fellas*" with the

more commonly used expression "*hey guys*", which is a more formal and widely used expression. The cultural aspect was taken into account by selecting a more common expression, while the lexical aspect was addressed by selecting an expression that would have a more accurate translation. The resulting translation "يا رفاق" (ya rifaq) appears to be accurate and appropriate in conveying the intended meaning. "Ya rafaq" is a common Arabic expression used to address a group of friends or colleagues and has a similar informal and friendly tone as the English phrase "hey guys". The PE operation was successful in this case as it produced a culturally and linguistically appropriate translation that conveyed the intended meaning of the source text.

- **Example 7 (Monsters, Inc.: 5:35)**
- **RST**: guys anybody home
- **TRST**: الرجال أي شخص في المنزل
- **P-EST**: guys... anybody home
- **TP-EST**: يا رفاق . أي شخص في المنزل

The RST in this example presents a situation where the speaker is attempting to make contact with someone or a group of people in a location and is inquiring if anyone is present. In this case, the RST consists of two separate instances that occurred with a brief time lapse in between. However, YouTube generated them as a single sentence, which resulted in the machine being unable to translate them accurately. That is, it generated the translation "الرجال" or "*the men*", which does not accurately convey the meaning of the original sentence. This highlights a common issue with machine translation, where the lack of context and inability to understand idiomatic expressions can lead to inaccurate translations. To address this issue, the student performed a grammatical PE operation by using an ellipsis "..." to separate the two instances of the RST. This PE operation helped to provide more context to TRADOS, allowing it to generate a more accurate translation: "يا رفاق .. أي شخص في المنزل" or "Hey guys ... anybody home". This translation more accurately conveys the meaning of the original sentence and is more appropriate in a conversational context. As a result of this PE operation, the outputted text was much more accurate in conveying the intended meaning.

- **Example 8 (Monsters, Inc.: 5:38)**
- **RST**: to the uzma kappa brotherhood
- **TRST**: إلى كابا كابا أخوة
- **P-EST**: To the uzma kappa fraternity
- **TP-EST**: إلى أخوية أوزما كابا

It should be noted that translating idiomatic expressions and context-specific language, such as the use of "brotherhood" in the context of a college fraternity is difficult. In this case, the term "*brotherhood*" was used to refer to a group or organization formed for a specific purpose. However, the NMT system of TRADOS translated it based on its other meaning, which is the quality or condition of being a brother or brothers. In terms of the linguistic features involved in this translation, the error can be attributed to TRADOS's inability to recognize the contextual meaning of the term "brotherhood" in the sentence. The system translated it literally as "إلى كابا كابا أخوة" or "To the Uzma Kappa Brothers". This not only missed the point of the sentence but also used the wrong grammatical form of the word "brotherhood". This resulted in a mistranslation that did not take into account the context of the sentence, which described a monster joining a college where several such groups, including Uzma Kappa, existed. To address this issue, the student opted for a lexical switch, replacing "*brotherhood*" with "*fraternity*". This change was necessary to ensure that the translated text conveyed the intended meaning of the original sentence accurately. By doing so, the student was able to overcome the limitations of the machine translation system and produce a clearer and more accurate translation.

- **Example 9 (Monsters, Inc.: 8:10)**
- **RST**: love that trick never gets old
- **TRST**: الحب تلك الخدعة لا يحصل أبدا قديمة
- **P-EST**: I love that trick, it is never boring
- **TP-EST**: أنا أحب هذه الخدعة، فهي ليست مملة أبدا

In this passage, the first example of the script had an omission of the pronouns "*I*" and "*it*", which

is a common feature of informal language in the US. This omission caused confusion for the machine translation, resulting in the words "love" and "get" being translated into the Arabic noun "حب" or "*the love*" and the verb "يحصل" "*to have*" without considering the co-text. Additionally, the rest of the sentence was translated in a disordered and unstructured way due to the omission. To fix this, the student added the missing pronouns and comma, which helped the machine produce a correct translation.

- **Example 10 (Monsters, Inc.: 11:43)**
- **RST**: what's going on, someone broke into the door lamp
- **TRST**: ماذا يكون يذهب على أحد ما [بروك] داخل المصباح الباب
- **P-EST**: what's going on, someone broke into the lamp door
- **TP-EST**: ما يحدث، شخص ما اقتحم باب المصباح

In this case, the expression "*door lamp*" affected the NMT output. The machine considered the word "*broke*", which is the past of to break, as a proper noun and borrow it into Arabic. The structure came flawed and the meaning unclear. When the student watched the scene, she noticed a door with a red lamp through which monsters can pass into the world of humans. Therefore, she proceeded to a syntactic change, inversed the order of words in the expression, and proposed "*lamp door*". The resulting sentence was "what's going on, someone broke into the lamp door", which was then translated by the NMT system of TRADOS to "ما يحدث، شخص ما اقتحم باب المصباح" which means "what's happening, someone broke into the lamp door". This translation accurately conveyed the intended meaning of the original sentence and removed any ambiguity caused by the unclear expression "door lamp". This example highlights how a syntactic change to the input sentence can help to address translation errors caused by unclear or uncommon phrases. By identifying the intended meaning of the expression and making a syntactic change to the input, the student was able to produce a more accurate and meaningful translation.

- **Example 11 (Robots: 44:03)**
- **RST**: after all that work and toil
- **TRST**: بعد كل هذا العمل و أنا فقط
- **P-EST**: after all that hard work
- **TP-EST**: بعد كل هذا العمل الشاق

"*Toil*", or commonly known as "*time off in lieu*", refers to a physically demanding and exhausting work, as defined by Merriam-Webster dictionary. However, the NMT system of TRADOS only translated the term "work" without capturing its full meaning. To address this issue, the student incorporated the adjective "*hard*" to convey the intended semantic meaning from the source text to the target language, resulting in an appropriate translation.

- **Example 12 (Robots: 45:46)**
- **RST**: hi mom oh I'm doing fine
- **TRST**: مرحبا أمي أوه أنا أعمل جيدا
- **P-EST**: Hi mom I'm fine
- **TP-EST**: مرحبا أمي أنا بخير

The error can be attributed to the use of the word "*doing*", which the machine translated literally as "أعمل" or "*I work*", instead of its intended meaning of "*do fine*". The student addressed this issue by simplifying the sentence and opting for a more easily translatable construction. This process involved a lexical PE operation.

Literal translation by NMT is frequent mainly in linguistic structures with indirect meaning. The student's lexical PE operation to fit the meaning avoided misunderstanding and made the translation appropriate to the context.

- **Example 13 (Monsters, Inc.: 1:37)**
- **RST**: I'm here to say that registration is that away okay officially a college
- **TRST**: أنا هنا لأقول ذلك هو أن بعيدا [أوكي] رسميا كلية
- **P-EST**: I'm here to say that registration is from here… okay… officially at college
- **TP-EST**: أنا هنا لأقول أن التسجيل من هنا ... حسنا.. رسميا في الجامعة

Once again, the script on YouTube was not properly segmented. Specifically, the segment con-

tained dialogue between two monsters, with a time lapse between their speeches. The automatic transcription software used by YouTube generated a confusing sentence, as evident in the RST above. Consequently, the NMT system of TRADOS produced a translation with an omission and a miswording of the sentence "that away" as "that is far away" instead of "from here". To address this issue, the student added ellipses "…" to indicate the time lapse, which enabled the machine to translate the segment accurately.

- **Example 14 (Monsters, Inc.: 2:05)**
- **RST**: hey there I'm your roomie name's randy
- **TRST**: مرحبا، أنا راندي اسم الغرفة الخاص بك
- **P-EST**: Hey there I'm your roommate, my name is randy
- **TP-EST**: أنا زميلك في الغرفة، اسمي راندي مرحبا

The translated text highlights a common challenge faced by machine translation systems in understanding informal language. The MT's interpretation of the sentence was affected by the absence of the possessive adjective "*my*". The MT considered "*Randy*" to be the room name and did not recognize that "roomie" is a US informal abbreviation for "roommate", which is commonly used by university students. The MT linked it to the closest word, which was "room". In this case, the abbreviation "roomie" was not recognized by the MT system, and the absence of the possessive pronoun "my" led to a mistranslation of the sentence. To remedy this, the student performed two PE operations. First, she inserted the possessive pronoun "my", which yielded a satisfactory result, except for "roomie". Therefore, she performed a second PE operation by replacing the abbreviation with the more formal word "roommate". These modifications helped the MT system to produce a more accurate translation of the sentence.

- **Example 15 (Monsters, Inc.: 4:03)**
- **RST**: zombie snob dominant silverback gorilla
- **TRST**: Zombie nob الغوريلا الغريلية الغليدية المسيطرة
- **P-EST**: a zombie… a strong silver back gorilla
- **TP-EST**: زومبي .. الغوريلا الفضية الخلفية القوية

Two main issues are highlighted in this example: context and punctuation. The absence of proper punctuation to mark dialogue in the source text resulted in a flawed translation by TRADOS, as it treated the entire instance as one sentence, leading to an incomprehensible translation. Furthermore, the expression "*Zombie snob*", which refers to a fan of old zombie movies, was mistranslated by the NMT system as it did not comprehend the meaning and borrowed the word "*snob*" with the 's' dropped. To improve the translation, the student watched the scene and noted that it featured a monster helping his inexperienced friend to be scary. The student decided to omit the term "snob" as it had no impact on the meaning of the sentence, resulting in a satisfactory translation. It is worth noting that the TP-EST contained a grammatical mistake, as the adjective "فضية" or "*silver*" was made definite, which is incorrect in Arabic since only the noun should be definite when an adjective precedes it.

- **Example 16 (Monsters, Inc.: 13:34)**
- **RST**: it is the gosh darndest thing
- **TRST**: هو [جوش] أكثر شيء
- **P-EST**: It is the most wonderful thing
- **T P-EST**: هذا هو الشيء الأكثر روعة

It is crucial to note that slang and informal language can frequently be difficult for machine translation systems to understand because AI models are often trained on formal and standardized language. Therefore, when encountering slang or colloquial expressions, such as the informal phrase "*gosh darndest*", which means "*best*" the machine translation was unable to interpret this slang and mistakenly borrowed the word "*gosh*" as if it were a proper noun. The translation improved when the student opted for a cultural adaptation by replacing "*gosh darndest*" with a more easily understandable phrase. In such a situation, the student's decision to replace the slang term with a more common phrase not only enhanced the translation's correctness, but also made it more accessible to a larger audience that might not be familiar with the slang term. This emphasizes the significance

of understanding the intended audience and context while doing machine translation, as well as the continual need for MT system advancements to properly handle informal language.

- **Example 17 (Monsters, Inc.: 13:36)**
- **RST**: scrabble's letting us into the scare program
- **TRST**: تخبرنا لعبة "التدافع" في الفزع
- **P-EST**: Harddscrabble is permitting us to join the scar program
- **T P-EST**: يسمح لنا هاردسكرابل بالانضمام إلى برنامج الخوف

Due to YouTube's speech recognition system, some words in the subtitles may be omitted. In this instance, the word "*hard*" was omitted, causing the machine to translate the word "scrabble" as its Arabic equivalent. Upon watching the video, the student discovered that "hardscrabble" was actually the name of a teacher at the monsters' college. In her first attempt to rectify the translation, the student added "*hard*" to "scrabble", but TRADOS still produced a literal translation. In a second attempt, the student added a "d" to "hard," resulting in the machine correctly identifying "Harddscrabble" as a proper noun and producing a suitable translation. All in all, by recognizing that "hard" was likely the omitted word, the student was able to make the necessary adjustment to produce a more accurate translation.

- **Example 18 (Robots: 38:01)**
- **RST**: who doesn't seem to be here gee
- **TRST**: من لا يبدو هنا لا عصيدة
- **P-EST**: He is not here apparently … what!
- **T P-EST**: يبدو أنه ليس هنا.. ماذا !

In this example, the original sentence includes the interjection, "gee", a slang word that conveys surprise or disbelief, which was linked to the rest of the sentence with no spacing. Hence, this lack of spacing between "here" and "gee" caused the MT to misinterpret the word as "congee", which led to an inaccurate translation. In other words, the NMT system of TRADOS did not produce a suitable translation for "gee" and instead interpreted it as "Congee", a type of rice porridge or gruel eaten in Asian countries. To improve the translation, the student attempted to separate the interjection from the rest of the sentence, but this did not resolve the issue. In a second attempt, the student replaced the interjection with "what!" which better conveyed the intended meaning of surprise or disbelief according to the context of the movie, and the resulting translation was better.

6   Discussion

From the early findings, it can be pre-concluded that the pre-edited versions are better in terms of correctness, appropriateness, and acceptability. One major finding is on the addition of what is omitted such as personal pronouns (it and I) or to do (to construct an interrogative sentence). While this may not be problematic in human communication, it can create confusion for machine translation systems like the one included in TRADOS. In such cases, a grammatical PE operation was found to be an efficient tool to avoid confusion and help the machine understand what is omitted.

The study found that the robots, YouTube auto-generating robots in our case, often drop sounds, especially when the speaker talks quickly or has a specific intonation as for the US natives. Another interesting finding from the Human Evaluation is that the YouTube-generating robot does not divide instances especially if the speakers are holding a very quick conversation or are talking at the same time. The study also highlights the importance of TRADOS's corpus in automatically translating terminologies, high-frequency phrases, and sentences (WANG, 2016, p. 255-256) i.e. the translation memory (TM) included in TRADOS allows to translate scientific terms, frequent expressions, and even complete sentences easily. However, if the unit is unclear or makes no sense, the machine may decode it wrongly since the engine analyses the source text input, encodes it into vectors, then decodes them into target text (RIVERA-TRIGUEROS, 2022). Therefore, PE can help to ensure that the machine correctly encodes and decodes inputs, leading to more accurate translations. The pre-editor may interfere by dividing the instances, using punctuation, or by introducing them separately to the machine. This methodical operation helps the machine define its vectors correctly.

The researchers also noticed that the NMT system of TRADOS has difficulties when dealing with

cultural markers. It struggles with cultural-specific expressions and requires PE to produce accurate translations. Students had to replace, modify or explain some idiomatic expressions (slangs, urban words, etc.) so that the machine can translate them.

The students also needed to replace some socio-cultural notions in order to avoid offending the reader's culture and religion. Translating a literary text remains a difficult task since literary writing is a dynamic equation that needs a bunch of stylistic and cultural elements, a wave of emotions meant to be aroused in the reader (ZEMNI; LABIB, 2022). For that, NMT in general, and TRADOS, in particular, need a competence other than language to understand and then reformulate a literary text. Idiomatic expressions, for instance, can be challenging at some point and not so many online translation software programs succeed in rendering them into the target language (ZEMNI; AWWAD; BOUNAAS, 2020, p. 2). All in all, only a human, until proven otherwise, can carry out a careful analysis of idiomatic expressions and socio-cultural factors and replace, omit in some cases, them when necessary. Additionally, the researchers have remarked that, in some instances, need to watch the video to understand the situation and then know the needed PE operation. Indeed, the audiovisual text has an image and a sound, colors and shapes and their translation must pass through two simultaneous and complementary channels audio and visual (KARKABOU, 2019). As a last remark, the researchers found out that capitalization has almost no impact on TRADOS outputs. Except for one situation, not all the words set in lowercase affected the translation; however, the researcher recommends this grammatical operation to avoid any inefficacy.

Overall, when using the NMT system of TRADOS, the researchers recommend dividing long instances with respect to the nature of the speech (dialogue, conversation …) and the sequences shown on the screen. They also recommend the use of synonyms and simple phrases instead of idiomatic expressions especially uncommon ones. Finally, yet importantly, the researcher believes that PE is an efficient tool to culturally accommodate translations to their readers.

7 Conclusion

The study investigated the impact of PE operations on the translation of child literary audiovisual texts (subtitles) from English into Arabic using TRADOS. The results revealed that Arabic remains a challenging language for NMT and PE is necessary to improve its outputs. PE operations involved making lexical changes to adapt inadequate notions and untangling ambiguous wordings to enhance the comprehensibility and interpretability of the translation. Furthermore, the analysis of the examples highlighted a potential shortcoming of TRADOS in translating informal language. The pre-editor's modification such as inserting what is missing, reordering the word order, or using punctuation to make the text more formal, improved the effectiveness and appropriateness of the outputs. The outcomes of this research suggest that PE can give rise to a better literary translation quality, especially for sensitive readers as children. The results obtained from this study may have significant implications for various parties involved. Firstly, researchers interested in conducting corpus-based investigations could find value in the approach and corpus presented in this paper. Additionally, Arab translators, and Saudi translation students, who use Arabic in their translation process may greatly benefit from the findings of this research.

8 Acknowledgment

References

AFZAAL, Muhammad et al. Automated and Human Interaction in Written Discourse: A Contrastive Parallel Corpus-based Investigation of Metadiscourse Features in Machine-Human Translations. en. *SAGE Open*, v. 12, n. 4, p. 215824402211422, Oct. 2022. ISSN 2158-2440, 2158-2440. DOI: 10.1177/21582440221142210. Available from: http://journals.sagepub.com/doi/10.1177/21582440221142210. Visited on: 10 July 2023.

BALASHOV, Yuri. The boundaries of meaning: a case study in neural machine translation. *Inquiry*, p. 1–34, Sept. 2022. ISSN 0020-174X, 1502-3923. DOI: 10.1080/0020174X.2022.2113429. Available from: https://www.tandfonline.com/doi/full/10.1080/0020174X.2022.2113429.

BENKOV, Lucia. Neural Machine Translation as a Novel Approach to Machine Translation. In: TURČÁNI, Milan (Ed.). *DIVAI 2020: The 13th International Scientific Conference on Distance Learning in Applied Informatics*. Štúrovo, SK: Wolters Kluwer, 2020. p. 499–508.

BOUNAAS, Chaouki; BEDJAQUI, Wafa. البحث التوثيقي في الترجمة الأدبية: دراسة تطبيقية على ممارسات بعض المترجمين (al bahth tawthiqi fi al tarjma al adabia: dirasat tatbiqia ala momarasat baad al motarjimin Documentary Research in Literary Translation: An Applied Study of Selected Translators' Practices). *Maalim*, v. 13, n. 1, p. 91–111, 2022.

DESJARDINS, Renée; LARSONNEUR, Claire; LACOUR, Philippe (Eds.). *When Translation Goes Digital: Case Studies and Critical Reflections*. Cham: Springer International Publishing, 2021. DOI: 10.1007/978-3-030-51761-8.

EHRENSBERGER-DOW, Maureen; MASSEY, Gary. Le traducteur et la machine: mieux travailler ensemble? *Des mots aux actes*, v. 8, p. 47–62, 2019. DOI: 10.15122/ISBN.978-2-406-09779-2. Available from: https://classiques-garnier.com/doi/garnier?filename=DmaMS03.

GAMAL, Muhammad. Audiovisual translation in the Arab world: A changing scene. *Translation Watch Quarterly*, v. 3, n. 2, p. 78–95, 2007.

GAMAL, Muhammad. Audiovisual translation in the Arab world: Mapping the Field. *Arab Media & Society*, n. 19, p. 1–12, 2014.

GUERBEROF ARENAS, Ana. Pre-Editing and Post-Editing. In: ANGELONE, Erik; EHRENSBERGER-DOW, Maureen; MASSEY, Gary (Eds.). *The Bloomsbury companion to language industry studies*. London: Bloomsbury Academic, 2020. (Bloomsbury companions). p. 333–360.

HIRAOKA, Yusuke; YAMADA, Masaru. Pre-editing plus neural machine translation for subtitling: effective pre-editing rules for subtitling of TED Talks. In: FORCADA, Mikel et al. (Eds.). *Proceedings of Machine Translation Summit XVII: Research Track*. Dublin: IE: European Association for Machine Translation, 2019. p. 64–72. Available from: https://aclanthology.org/W19-6710.

IBROKHIMOVICH, Fozilov Jakhongir et al. The Importance of Mother Tongue and Children's Literature in Primary School. *Eurasian Journal of Learning and Academic Teaching*, v. 5, p. 1–3, 2022.

IZWAINI, Sattar. Machine translation in the Arab world: Saudi Arabia as a case study. *Trans-Kom. Wissenschaftliche Zeitschrift Für Translation Und Kommunikation*, v. 8, n. 2, p. 382–414, 2015.

JABER, Fadi. The landscape of translation movement in the Arab world: From the 7th Century until the beginning of the 21st Century. *Arab World English Journal (AWEJ) Vol*, v. 6, n. 4, p. 128–140, 2015.

KARKABOU, Souad. دبلجة الأفلام الموجهة للأطفال من اللغة الإنجليزية إلى اللغة العربية (dabljat al aflam al mowjaha lilatfal mina al logha al inglizia ila al logha al arabia Children's movies dubbing from English to Arabic Children's movies dubbing from English to Arabic). *Insaniyat*, p. 83–84, June 2019. ISSN 1111-2050, 2253-0738. DOI: 10.4000/insaniyat.20826. Available from: http://journals.openedition.org/insaniyat/20826.

KE HU, Kevin. How MT errors correlate with postediting effort: a new ranking of error types. *Asia Pacific Translation and Intercultural Studies*, v. 7, n. 3, p. 299–309, Sept. 2020. DOI: 10.1080/23306343.2020.1809763.

LOOCK, Rudy. No more rage against the machine: how the corpus-based identification of machine-translationese can lead to student empowerment. *The Journal of specialised translation*, v. 34, p. 150–170, 2020.

MARZOUK, Shaimaa; HANSEN-SCHIRRA, Silvia. Evaluation of the impact of controlled language on neural machine translation compared to other MT architectures. *Machine Translation*, v. 3, n. 33, p. 179–203, June 2019. DOI: 10.1007/s10590-019-09233-w.

MIYATA, Rei; FUJITA, Atsushi. Understanding Pre-Editing for Black-Box Neural Machine Translation. In: MERLO, Paola; TIEDEMANN, Jorg; TSARFATY, Reut (Eds.). *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. Online: Association for Computational Linguistics, 2021. p. 1539–1550. DOI: 10.18653/v1/2021.eacl-main.132. Available from: https://www.virtual2021.eacl.org/paper_main.636.html.

PYM, Peter. Pre-editing and the use of simplified writing for MT. In: MAYORCAS, Pamela (Ed.). *Translating and the Computer 10: The Translation Environment 10 Years on*. London: Aslib, 1990. p. 80–95.

RIVERA-TRIGUEROS, Irene. Machine translation systems and quality assessment: a systematic review. *Language Resources and Evaluation*, v. 56, n. 2, p. 593–619, June 2022. DOI: 10.1007/s10579-021-09537-5. Available from: https://link.springer.com/10.1007/s10579-021-09537-5.

SOLYMAN, Aiman. Synthetic data with neural machine translation for automatic correction in arabic grammar. *Egyptian Informatics Journal*, v. 22, n. 3, p. 303–315, Sept. 2021. DOI: 10.1016/j.eij.2020.12.001. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1110866520301602.

STAHLBERG, Felix. Neural Machine Translation: A Review. *Journal of Artificial Intelligence Research*, v. 69, p. 343–418, 2020. DOI: 10.1613/jair.1.12007. Available from: https://jair.org/index.php/jair/article/view/12007. Visited on: 10 July 2023.

TAUFIK, Alvin. Pre-Editing of Google Neural Machine Translation. *Journal of English Language and Culture*, v. 10, n. 2, p. 64–74, 2020. DOI: 10.30813/jelc.v10i2.2137. Available from: https://journal.ubm.ac.id/index.php/english-language-culture/article/view/2137.

TOMASZKIEWICZ, Teresa. Traduction automatique dans la formation des traducteurs: une analyse expérimentale de la post-édition. *Studia Romanica Posnaniensia*, v. 45, n. 4, p. 75–89, 2018.

WANG, Yahui. Translation Process Management and Quality Control Based on SDL Trados–A Case Study of Energetic Materials at Extreme Conditions Translation Project. In: PROCEEDINGS of the 2016 International Conference on Education, E-learning and Management Technology. Xi'an, China: Atlantis Press, 2016. p. 255–261. DOI: 10.2991/iceemt-16.2016.50. Available from: http://www.atlantis-press.com/php/paper-details.php?id=25860017.

ZEMNI, Bahia; AWWAD, Wiam; BOUNAAS, Chaouki. Audiovisual translation and contextual dictionaries: an exploratory comparative study of Reverso Context and Almaany uses. *Asian EFL Journal Research Articles*, v. 27, n. 5.1, p. 274–309, 2020.

ZEMNI, Bahia; LABIB, Mona Abdelghani. Transfert des culturèmes religieux dans la traduction française de la Trilogie de Naguib Mahfouz. *Kervan. International Journal of African and Asian Studies*, v. 26, Special Issue, p. 123–142, 2022.

Author contributions

Chaouki Bounaas: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft, Writing – review and editing; **Bahia Zemni**: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft, Writing – review and editing; **Fadiah Alshehri**: Investigation, Methodology, Project administration, Resources, Software, Supervision, Funding acquisition, Writing – original draft, Writing – review and editing; **Mimouna Zitouni**: Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Funding acquisition, Writing – review and editing.