



ISSN: (3006-8614)
E-ISSN: (3006-8622)

Journal of Alma'rifa for Humanities

available online at: <https://uomosul.edu.iq/womeneducation/almarifa/>



THE EFFECT OF PRE-EDITING ON POST-EDITING IN MACHINE TRANSLATION

Mahmoud Ahmed Mahmoud
University of Mosul / College of Arts

*Corresponding author: E-mail :
mahmoud.a.m@uomosul.edu.iq

 0009-0002-0190-4661

Keywords:

Pre-editing,
post-editing,
machine translation,
human translation

ARTICLE INFO

Article history:

Received 1. Sep.2025
Revised 25. Sep.2025
Accepted 5. Oct.2025
Available online 3.Jan.2026

Email:

almarefaa.ecg@uomosul.edu.iq

A B S T R A C T

This research project deals with the relationship between pre-editing & post-editing in machine-generated translation, focusing on how strategic modifications to original texts impact the quality of the translation. In spite of the significant advancement of the MT, its raw output frequently requires substantial post-editing to meet professional standards. This research directly compares two conditions: (1) raw Google Translate output of unedited source texts, and (2) raw Google Translate output of pre-edited source texts. It investigates the pre-editing results: in quicker post-editing, superior translation quality, and less cognitive strain for editors. The goal of the study is to offer practical guidance for translators on the effective application of pre-editing. By clarifying the balance between the initial effort of pre-editing and the subsequent reduction in post-editing, this work adds valuable insight into optimizing human-AI partnerships within the translation process. ©2026AJHPS, College of Education for women, University of Mosul.

تأثير التحرير المسبق على التحرير اللاحق في الترجمة الآلية

محمود احمد محمود

جامعة الموصل/ كلية الآداب

الخلاصة:

يُعنى هذا البحث بالعلاقة بين التحرير المسبق والتحرير اللاحق في الترجمة الآلية، مع التركيز على كيفية تأثير التعديلات الاستراتيجية على النصوص المصدر في جودة الترجمة. فبالرغم من التقدم الكبير في الترجمة الآلية، إلا أن مخرجاتها الخام غالباً ما تحتاج إلى تحرير لاحق مكثف لتلبية المعايير المهنية. و يقارن هذا البحث بين حالتين: 1. نصوص أصلية غير محررة بشكل مسبق ترجمها برنامج جوجل للترجمة 2. نصوص أصلية خضعت للتحرير المسبق قبل ان يترجمها برنامج جوجل. ويتقصى البحث ما إذا كان التحرير المسبق يؤدي إلى: - تسريع عملية التحرير اللاحق و تحسين جودة الترجمة و تقليل العبء المعرفي على المحررين. تهدف الدراسة إلى تقديم إرشادات عملية للمترجمين من خلال تحديد متى وكيف يجب تطبيق التحرير المسبق. ومن خلال توضيح المقايضة بين جهد التحرير المسبق ومكاسب التحرير اللاحق، تساهم هذه الدراسة في النقاش الأوسع حول التعاون بين الإنسان والنكاء الاصطناعي في مجال الترجمة.

الكلمات المفتاحية: التحرير المسبق، التحرير اللاحق، الترجمة الآلية، الترجمة البشرية

1- Objective of the Study

This study aims to examine the impact of varying degrees of pre-editing on both the required post-editing effort and the final quality of machine-translated output. The research will explore how modifying source texts prior to machine translation influences the efficiency, quality, and cognitive demands of the subsequent human post-editing process.

2- Hypotheses of the Study

H1: Pre-editing leads to a direct improvement in the initial quality of machine-translated output.

H2: Pre-editing helps improve the quality and speed of post-editing.

H3: Post-editing machine-translated output from pre-edited texts results in a lower cognitive load for the editor.

H4: The practice of pre-editing increases translators' understanding of the limited capabilities and limitations of machine translation systems.

H5: The overall final translation quality is higher when a pre-editing stage is implemented before machine translation.

3- Introduction

This study explores the concepts and processes of pre-editing and post-editing, as applied in professional and academic translation settings that utilize machine translation (MT).

While some institutions, according to Lommel and DePalma (2016), such as the Pan American Health Organization and the European Commission, adopted pre-editing and post-editing of raw MT output as early as the 1980s, it has only been in the past decade that machine translation post-editing (PE) has become a standard practice in the workflows of global localization agencies.

As this technology is still relatively new, Lommel and DePalma (2016), further research is needed to understand its impact on traditional translation workflows and the professionals involved. While this remains an ongoing area of study, several key concepts and processes have emerged in recent years. This study will outline these concepts in the forthcoming sections including: controlled language and pre-editing, post-editing expected quality standards and pre-editing guidelines-concluding with a brief overview of common MT error types.

4- Controlled Language

O'Brien (2003) states that controlled language refers to a set of writing rules applied to texts to minimize lexical ambiguity and complex grammar. This approach enhances readability and comprehension while facilitating the use of translation technologies like translation memories and MT. Of course.

5- Pre-editing

Research on pre-editing remains limited. Hiraoka and Yamada (2019) identify two primary types: bilingual and monolingual pre-editing. The crucial difference is that bilingual pre-editing allows the editor to adjust the source text while consulting the machine translation output. This study employs a corpus of texts to evaluate machine translation results, with a specific emphasis on the bilingual approach.

Miyata and Fujita (2017) note that pre-editing is not yet a common practice and that academic investigation into it is still nascent. Key questions regarding its cost-effectiveness and the most effective implementation methods await comprehensive analysis.

6- Definition of Pre-editing

Pre-editing is broadly defined as a strategic method that combines human knowledge with machine efficiency to prepare source texts for more effective machine translation (Kokanova et al., 2022). Its objective is to lessen frequent MT problems, including complex syntax, idioms, and typographical errors.

According to Arenas (2020), the process involves revising texts based on specific rules—such as shortening sentences, simplifying grammatical structures, and standardizing terminology—to decrease the cognitive load on the MT system and improve the clarity and accuracy of the translation. Yang (2023) adds that by refining the source text, pre-editing helps prevent errors and misunderstandings, particularly because MT systems often fail to grasp nuanced meanings.

Hudáková (2024) describes pre-editing as a proactive strategy that elevates source text quality to generate superior MT output. While not inherently technology-dependent, its value has increased as MT systems demonstrate better performance with well-prepared input. This method optimizes translation workflows by creating MT-friendly source texts, thereby lessening the need for considerable post-editing. Research by

Bounaas et al. (2023) supports this, showing that pre-editing markedly improves MT quality, resulting in fewer word errors and less post-editing effort.

For Guerberof Arenas (2020), pre-editing entails polishing the original text before automated translation by applying specific terminological and stylistic rules. This step is designed to enhance the quality of the raw MT output. Pre-editors follow guidelines not only to correct typos and content errors but also to simplify sentence construction, apply preferred grammatical forms (e.g., using active voice and straightforward word order), ensure terminology consistency, and flag terms (like proper nouns) that should not be translated.

7- The Depth of Pre-editing

AlOtaibi (2020) asserts that pre-translation editing is vital in natural language processing. It extends beyond simple text preparation to include a detailed analysis, optimization, and refinement of the source text to produce an input that is more amenable to machine processing, thereby boosting the quality and efficiency of the MT output.

Bernth and Gdaniec (2001, cited in Zhang, 2025) contend that pre-editing is a multi-layered process. It begins with essential error correction, requiring editors to have strong linguistic skills to fix spelling, grammar, and punctuation mistakes, as these can adversely affect translation accuracy. The next level is text optimization, which involves restructuring sentences, reducing complexity, eliminating ambiguities, and substituting idioms or colloquial language. Terminology management is also critical, as consistent use of terms directly influences translation quality. Furthermore, stylistic adaptation tailors the text to its specific genre, ensuring objectivity in technical documents or expressiveness in literary works. Finally, editors must compensate for the MT system's limited contextual understanding by interpreting implicit meanings, metaphors, and rhetorical devices.

Gladstone et al. (2024) state that manual pre-editing strategies improve MT by refining source texts for greater readability and translatability. Effective optimization requires editors to understand how MT systems operate (Delavenay & Delavenay, 1960, cited in Zhang, 2025). Techniques include correcting language errors, restructuring sentences (e.g., breaking up long ones), and clarifying ambiguous expressions like idioms (Yuen, 2021). Miyata and Fujita (2017) observe that pre-editing generally demands advanced editing skills in the source language to identify and correct deviations from specific predefined guidelines.

8- Types of Pre-editing

Pre-editing, for Hiraoka & Yamada (2019), can be divided into two main approaches: bilingual and monolingual. In bilingual pre-editing, the editor modifies the source text while reviewing the machine-translated output, whereas monolingual pre-editing does not involve checking the translation. As a result, monolingual pre-editing does not require proficiency in the target language.

9- Post-editing

Koby (2001) mentions that shortly after machine translation (MT) research began in the 1950s, it became evident that post-editing (PE, also called MTPE) was often necessary to produce a usable final translation. Initially, the focus was on creating fully automated, high-quality MT systems, making post-editing—the process of fixing errors in raw MT output—an unwelcome necessity. Hutchins (1986: p. 31) notes that the term "post-editor" was coined in 1950, describing someone who selects the best translation from the computer's options and adjusts word order to fit the target language.

PoE, for O'Brien (2022), is a complex and mentally demanding task, especially for stylistically rich texts, which may require more effort than traditional human translation.

Post-editing, for Allen (2003: 296; cited in Girletti (2024)), refers to the process of revising, adjusting, and/or correcting machine-translated text from a source language into one or more target languages. In essence, it involves reviewing an MT-generated translation against the original source text, identifying and fixing errors to meet predefined quality standards.

Arenas (2020) says that post-Editing (PoE) refers to the process of refining machine translation (MT) output to meet specific linguistic and stylistic standards. Human translators enhance machine-generated text by not only correcting errors but also ensuring the content resonates with the target audience's expectations and contextual requirements.

10- Types of Post-Editing: Light vs. Full

In machine translation and post-editing, according to Girletti (2022), texts are often categorized as either:

- Assimilation-oriented – meant for basic comprehension, where readability and accuracy matter more than stylistic perfection, and minor grammatical/spelling errors may be tolerated.

- Dissemination-oriented – intended for publication, requiring not only accuracy and clarity but also polished style, grammar, spelling, and terminology comparable to human-translated content.

Accordingly, post-editing is classified based on depth:

- Full post-editing – achieves human-quality translation.
- Light (rapid) post-editing – involves minimal corrections for general understanding ("gist").

11- Pre-editing vs Post-editing

In professional environments, according to Hudáková (2024), Pre-Editing (PrE) is usually carried out by source text authors, technical editors, or specialized language experts, rather than by those who handle Post-Editing (PoE). However, in this study, both PrE and PoE are released for a controlled assessment of workflow efficiency and output quality. Future

research, adds Hudáková (2024), could investigate how PrE influences translation quality when performed by different professionals in the workflow and whether its effects differ from the natural cognitive processing that occurs when translators review source texts before machine translation (MT). Such investigations could help clarify PrE's precise role in enhancing translation accuracy and efficiency.

Hudáková (2024) suggests that combining PrE and PoE yields the best outcomes, particularly for participants skilled in both processes. This dual approach led to the fewest errors, implying that proficiency in both stages enhances overall translation efficiency. Alvarez-Vidal et al. (2020) say that while many language service providers and translation companies endorse this view, some professional translators remain doubtful, contending that MT-generated errors may demand extensive revisions, offsetting its efficiency gains.

Hudáková's (2024) research suggests that while neither pre-editing (PrE) nor post-editing (PoE) is universally superior, a combined approach may yield the best translation quality. Integrating both methods can improve fluency, accuracy, and terminology use, while also making the post-editing stage easier by providing a cleaner machine translation output. A key drawback, however, is the additional time required upfront for pre-editing. The study emphasizes that translator training should evolve to include machine translation and post-editing skills, equipping future professionals to effectively merge human expertise with AI tools. Ultimately, a strategic PrE-PoE workflow is presented as the most promising method for leveraging machine translation.

12- Categorizing Machine Translation Errors

Guerberof Arenas (2020) outlines a common framework for classifying MT errors, which includes:

- Accuracy: Errors in understanding the original text.

- Language: Grammatical and spelling errors.
- Terminology: Failure to adhere to specified glossaries.
- Style: Violations of style guide requirements.
- Country Standards: Incorrect localization of formats for dates, currencies, etc.
- Format: Problems with layout, tags, or hyperlinks.

Arenas (2020) adds that errors can also be graded on a severity scale (e.g., from 1 to 3) based on their consequence. These classification systems are designed to be customizable for individual projects, clients, or language service providers.

13- Machine Translation (MT)

As defined by Sophie (2025), Machine Translation is the use of software to automatically translate text from one language to another without human involvement. Its development began in the 1950s with rule-based systems and has evolved considerably. However, Sophie (2025) notes that even modern Neural Machine Translation (NMT) is imperfect. It can produce grammatical mistakes, misinterpret meanings, and overlook cultural nuances, particularly with technical, creative, or culturally rich content. Its effectiveness is also limited by the quality and scope of its training data, which can contain biases or lack coverage in specialized domains.

14- Human Translation

According to Sophie (2025), human translation is carried out by professional translators who are fluent in both the source and target languages. Unlike machines, humans can grasp subtle nuances in meaning, tone, idiomatic expressions, and cultural context. Their role extends beyond converting words to interpreting the underlying intent, style, and purpose of a text to ensure the communication is both accurate and culturally appropriate.

Unlike static machine systems, Sophie (2025) continues, human

translators can clarify ambiguities, conduct research, and adapt to new terms, especially in rapidly evolving industries. Many use Computer-Assisted Translation (CAT) tools—such as Trados or MemoQ—to enhance consistency and efficiency. However, these tools support rather than replace human judgment.

15- The Evolution of Human Translation

The earliest known translation traces back to Mesopotamia (3000-2000 BC), Kelly (1995) mentions, where bilingual inscriptions in various Asian languages were discovered. Modern translation approaches, Kelly (1995) adds, stem from two historical perspectives: Horace, a poet who opposed literal translation, and Cicero, who advocated for conveying meaning rather than adhering strictly to word counts. Cicero famously stated, "I did not hold it necessary to render word-for-word, but I expressed the general style and the force of language" (Cicero, 1949; cited in Marshall (2024)). While Horace and Cicero influenced Roman translation practices, philosopher Seneca the Younger argued for strict literal translation to preserve truth.

16- The Evolution of Machine Translation

According to Hutchins & Lovtskii (2000), the earliest known concept of a "translating machine" dates back to 1933, when Russian scientist Petr Petrovich Smirnov-Troyanskii patented a device designed to translate between languages and automatically select and type words.

The late 20th and early 21st centuries, says Doherty (2016), saw rapid advancements in translation due to the digital revolution. Computer-assisted translation (CAT) tools, including translation memory (TM) systems, streamlined workflows by storing previously translated segments for reuse. This also enabled translators to build bilingual corpora for future projects.

Marshall (2024) mentions that Warren Weaver (1949) is often regarded as the "father" of machine translation (MT) after suggesting in (1947) that

computers could be used for translation. In his influential memorandum, Weaver outlined four key challenges: handling words with multiple meanings, addressing logical (and illogical) language structures, applying cryptographic techniques to MT, and exploring linguistic universals.

The first MT systems in the (1950s) were, according to Poibeau (2017), rule-based. The simplest, direct translation systems, converted text from a source to a target language without intermediate steps, relying mainly on bilingual dictionaries. More advanced transfer systems incorporated morphological and syntactic rules for both languages. However, ambiguity remained a major hurdle, prompting the inclusion of semantic rules. While multiword expressions were sometimes added to dictionaries to provide context, this did not fully resolve ambiguities.

Example-based MT, introduced in 1984, marked the first shift toward data-driven methods, eliminating the need for manually coded rules (Nagao, 1984). This approach, Poibeau (2017) maintains, matched text fragments with their translations in bilingual corpora, then combined them to form coherent sentences. However, its effectiveness was limited when translations for certain fragments were unavailable, making it more suitable for specialized domains than general language.

17- Limitations of MT

The shortcomings of MT, according to Hadi et al. (2023), encompass weak reasoning and feedback interpretation, narrow common sense, biased training data, poor generalizability, lack of transparency, struggles with rare or unfamiliar words, insufficient grasp of syntax and grammar, limited expertise in specialized domains, challenges with context-dependent language, absence of emotion and sentiment analysis, lack of originality, inability to resolve ambiguity, inadequate understanding of human psychology and behavior, and a constrained perception of the world beyond textual data. Additionally, scholars like Brynjolfsson (2022) contend that

machines should enhance human capabilities rather than replicate them.

18- Translation Quality Evaluation (TQE)

TQE remains a widely debated subject in translation studies, technology, and the localization industry. Various methodologies and theories exist for assessing translation quality, often shaped by differing interpretations of what constitutes the term ‘quality’. As Drugan (2013) notes, experts largely agree that no universal standard exists for measuring quality. Definitions of high-quality translation, says Lauscher (2000), range from achieving optimal equivalence to adhering to predefined specifications.

19- Human Translation vs. Machine Translation

Several studies have examined the differences between human (HT) and machine translation (MT). For instance, Li et al. (2014) compare Chinese-to-English translations from Google Translate (using Statistical MT) and professional human translators, finding similarities in formality and cohesion but lacking detailed error analysis.

Human translation has clear drawbacks, Dai et al. (2022) mentions, the most notable being inefficiency. A person’s capacity is limited—while a machine can translate a short text in minutes, a human may take an hour or longer. Additionally, even skilled translators have limitations in language mastery. Another disadvantage is cost. Many online translation tools are free, making them the preferred choice for those who don’t require high accuracy. In contrast, human translators charge per word, which further highlights the downsides of manual translation.

Dai et al. (2022) contend that while machine translation can communicate the basic essence of a text, it is insufficient for tasks requiring precision, where human expertise remains crucial.

Sophie (2025) elaborates that AI-powered Neural Machine Translation offers a rapid, scalable, and cost-effective solution for processing high

volumes of text. Its utility is greatest in domains where speed and broad accessibility are prioritized over subtlety and nuance. Nevertheless, MT consistently underperforms in areas requiring a deep understanding of context, cultural references, tone, idioms, and specialized jargon.

In opposition to this, human translation, as described by Sophie (2025), depends on expert linguists who contribute cultural awareness, domain-specific knowledge, and contextual accuracy. Although this process is slower and more expensive, it guarantees a higher level of precision and flexibility, making it essential for contexts where error-free language and appropriate tone are non-negotiable.

The benefits of MT—its speed, low cost, and constant improvement—make it ideal for breaking down language barriers in high-volume sectors like e-commerce and customer support (Sophie, 2025). However, its pronounced difficulties with idioms, ambiguity, technical terms, and emotional resonance can introduce substantial risks in fields such as law, healthcare, and marketing, where mistakes can lead to legal, safety, or reputational damage.

Human translation, conversely, is executed by professionals who possess not only full language mastery but also the ability to incorporate cultural insight, contextual interpretation, and industry expertise. Their role involves adapting tone, conveying subtle meanings, and customizing content for its intended audience—a level of meticulousness that is critical in high-stakes environments where an error could cause serious miscommunication, legal issues, or harm to a brand.

Consequently, an integrated hybrid model is often proposed as the optimal strategy, combining the efficiency of machine translation for initial processing with the quality control of human post-editing for critical content.

20- The Adopted Model

The practical part of this research employs the influential model

established by Ana Guerberof Arenas (2020), which can be summarized as follows:

_ It utilizes Controlled Language principles combined with Pre-Editing.

_ The method involves simplifying grammar and vocabulary to enhance the quality of machine-generated translations.

_ It focuses on preparing the source text according to specific rules that promote clarity, consistency, and terminological accuracy.

_ The model evaluates the impact on post-editing effort (measured in time and cognitive load), final quality, and translator experience.

21- Scholarly Credibility

Guerberof Arenas is a leading scholar in translation studies and machine translation workflows. Her work is frequently cited in both academic and industry contexts, including TAUS, ISO standards, and EMT pedagogical frameworks.

22- Key pre-editing principles applied by the researcher:

1. Simplify jargon.
2. Shorten sentences.
3. Remove idioms/colloquialisms.
4. Clarify ambiguity.
5. Neutralize tone.

23- Data Analysis and Discussion

Building on Guerberof Arenas' hybrid model (2020), the following are several rich examples (12 excerpts) of real-world texts that could be used to illustrate and test what has been mentioned above from the following sources:

- The Week US. (2024, November 15).
- The Week UK. (2024, October 19).
- Robinson, D. (2012). Becoming a translator

- Excerpt no. (1)

((The Week US_151)))

Page 1:

‘WAS HARRIS DOOMED TO LOSE?’

MT:

هل كان هاريس محكوما عليه بالخسارة؟

After pre-editing:

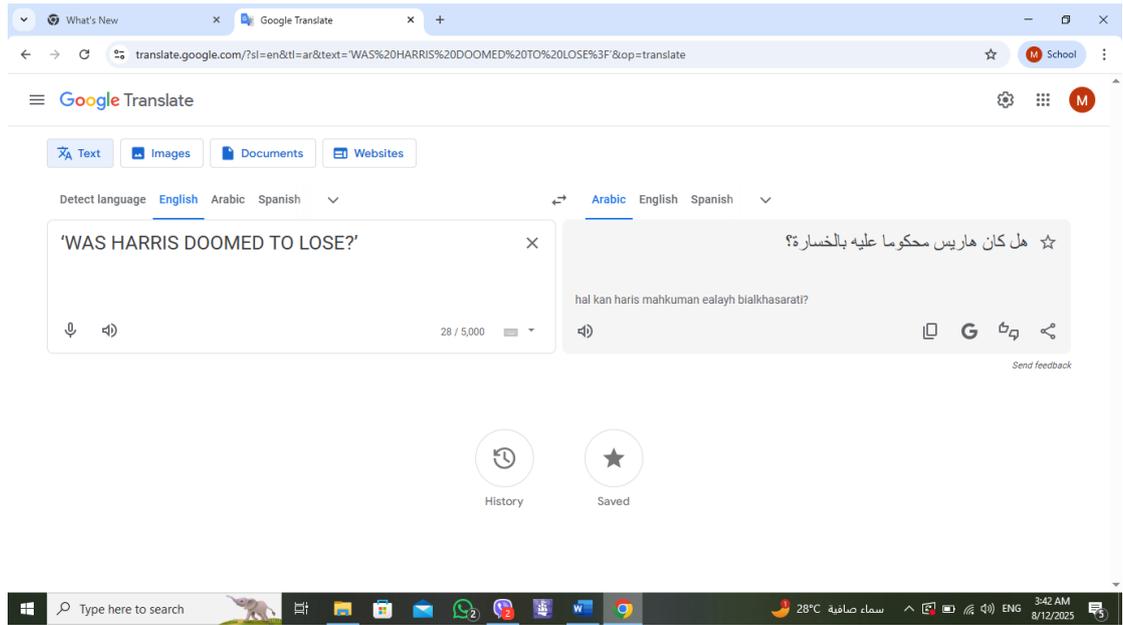
‘Was Vice President Harris's electoral defeat inevitable?’

هل كانت هزيمة نائبة الرئيس هاريس الانتخابية حتمية؟

Discussion

Pre-editing in the excerpt above is crucial for the MT version which refers to Harris as a male! No post-editing is required after the pre-editing.





Excerpt no. (2)

((The Week UK_1910))

Page 14

‘Turning a blind eye to Sudan’s agony’

MT:

غض الطرف عن معاناة السودان

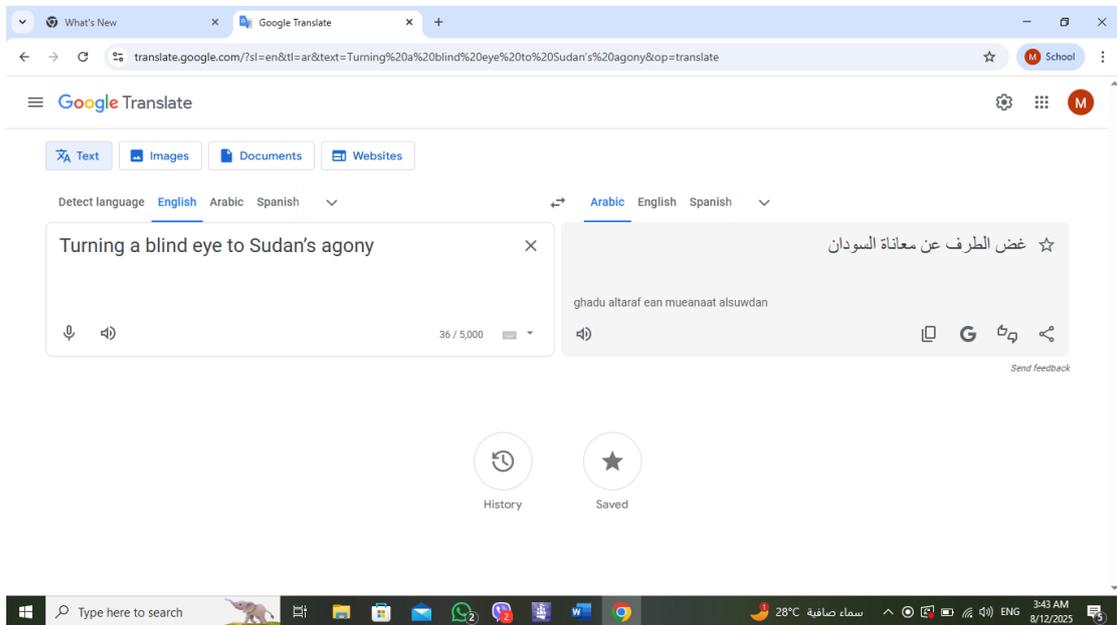
After pre-editing:

Ignoring the humanitarian crisis unfolding in Sudan

تجاهل الأزمة الإنسانية المتصاعدة في السودان

Discussion

Pre-editing in the excerpt above further explains the situation in Sudan.



Excerpt no. (3)

((The Week UK_1910))

Page 17

“Baby boomers born in the late 1940s and 1950s were about 150% more likely to suffer from cancer, lung disease and heart problems when they were in their 50s and 60s than those born before the Second World War.”

MT:

كان جيل طفرة المواليد الذين ولدوا في أواخر الأربعينيات والخمسينيات من القرن العشرين " أكثر عرضة للإصابة بالسرطان وأمراض الرئة ومشاكل القلب بنسبة 150% عندما كانوا في الخمسينيات والستينيات من العمر مقارنة بمن ولدوا قبل الحرب العالمية الثانية "

After pre-editing:

Babies born during the late 1940s and 1950s surge, roughly late 1940s–1950s had significantly worse health outcomes in their 50s and 60s compared to those born before WWII likely the early 1900s to mid-1940s.

كان الأاطفال الذين ولدوا ابان الزيادة السكانية الحاصلة في أواخر الأربعينيات والخمسينيات " أكثر عرضة بنسبة 150٪ تقريباً للإصابة بالسرطان وأمراض الرئة والقلب في سن الخمسينيات (على الأرجح في أوائل القرن العشرين والستينيات مقارنة بمن ولدوا قبل الحرب العالمية الثانية حتى منتصف الأربعينيات)."

Analysis:Pre-Editing Challenges:

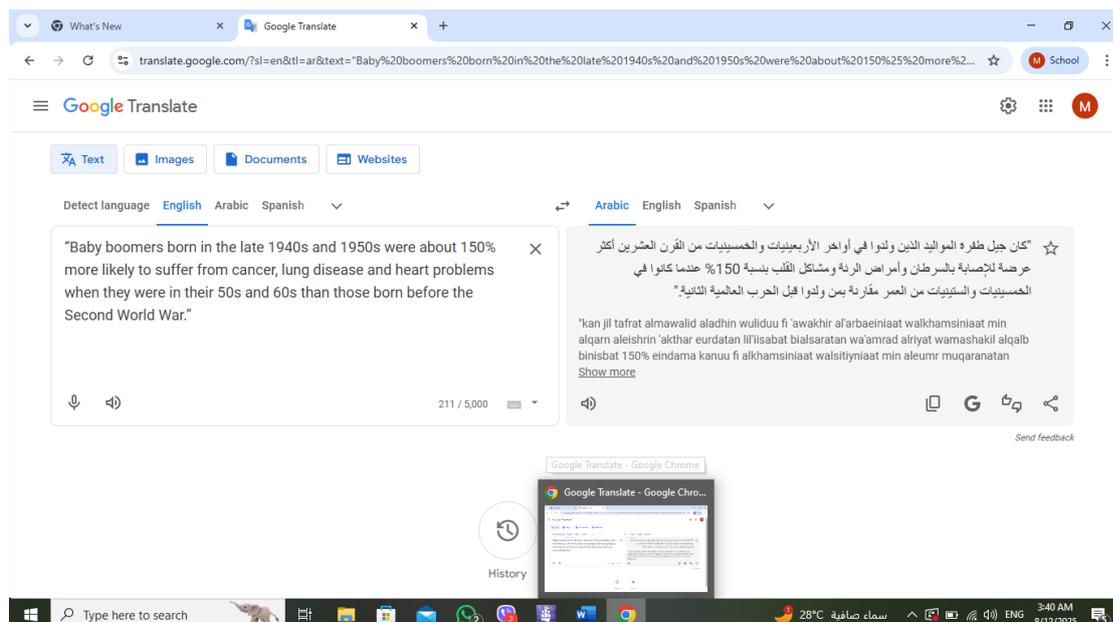
- Complex comparative structure
- Statistical phrasing
- Generational terminology

The screenshot shows a news article from The Guardian. The main headline is "What the scientists are saying..." under the "Health & Science" section. The article title is "An alarming generational 'drift'". The text discusses how increases in various illnesses and ailments are often attributed to the fact that people are living longer – and so developing conditions that their parents would have not reached the age to get; but a new study has found that today's older people are simply less healthy than earlier generations. The analysis of data on more than 100,000 people from several generations in England, the US and Europe revealed that "baby boomers" born in the late 1940s and 1950s were about 150% more likely to suffer from cancer, lung disease and heart problems when they were in their 50s and 60s than those born before the Second World War were when at that age. They were also as likely or more likely to struggle with basic tasks such as bathing, eating and walking short distances. "Even with advances in medicine and greater public awareness about healthy living, people born since 1945 are at greater risk of chronic illness and disability than their predecessors," said lead author Laura Gimeno of University College London – a finding that has "huge implications for government spending". And this

in a phase 3 trial, halting the advance of the disease for 40% longer than the standard treatment, reports The Guardian. The trial involved well over 1,000 people, all of whom had an advanced form of non-small cell lung cancer, and carried a mutation of the epidermal growth factor receptor (EGFR) gene, which is found in about a quarter of cases worldwide. Such a mutation among cancer patients occurs more in women than men, and in people who've never smoked heavily or at all. Some patients were given a combination of amivantamab and lazertinib – drugs that work in different ways to target cancer cells; others were given osimertinib, the standard treatment. Patients who took the drug cocktail were still alive with no progression in their disease 23.7 months later, on average, compared with 16 months for the others. "It's amazing to see this new combination shows longer cancer control than osimertinib, which was itself a breakthrough treatment only a few years ago," said Dr Martin Forster, the oncologist who led the UK arm of the trial.

The vapers who've never smoked

Comb jellies appear to have no sense of self together. To test their theory, they took some specimens and removed a small part of each's body, after which they pinned the ctenophores together in pairs, so that their injuries touched, and left them in a tank. Nine times out of ten, the two animals grafted together, presumably as a survival



Excerpt no. (4)

((The Week UK_1910)))

Page 6

“She promised to be generous, but also thrifty. To spend, but also to save.”

MT:

وعدت بأن تكون كريمة، ومقتصدّة في الوقت نفسه. أن تُنفق، وأن تُدخر أيضًا

After pre-editing:

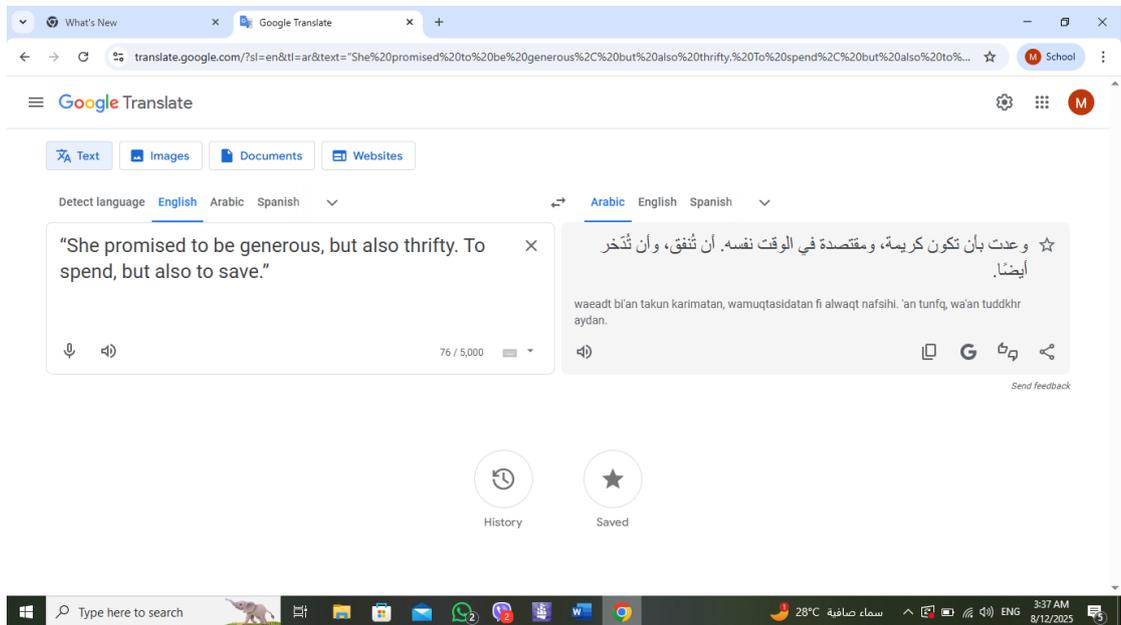
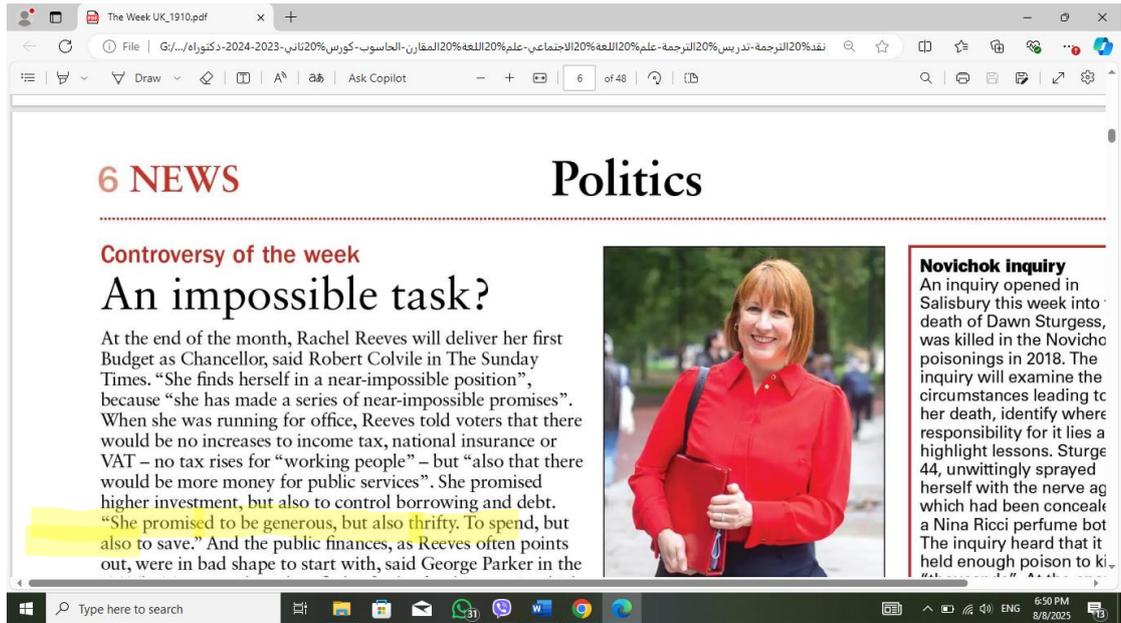
“She made conflicting promises: to increase spending while also reducing costs.”

"لقد قدمت وعودًا متناقضة: بأن تزيد الإنفاق وفي نفس الوقت تخفض التكاليف"

Discussion

In the excerpt above, pre-editing..

- clarifies contradiction ("conflicting promises").
- makes it more formal/structured for analytical contexts.
- simplifies translation with concrete terms.



Excerpt no. (5)

((((The Week UK_1910)))

Page 44

“Ghana is one of the world’s largest importers of secondhand clothing from the Global North, with 15 million garments arriving every week...”

MT:

"تُعد غانا واحدة من أكبر مستوردي الملابس المستعملة من دول الشمال في العالم، حيث "...يصل إليها 15 مليون قطعة ملابس أسبوعيًا"

After pre-editing:

- Clarify “Global North” as “wealthier countries such as the UK and US”

- Add context for “secondhand clothing” as “used garments often discarded by consumers in developed nations”

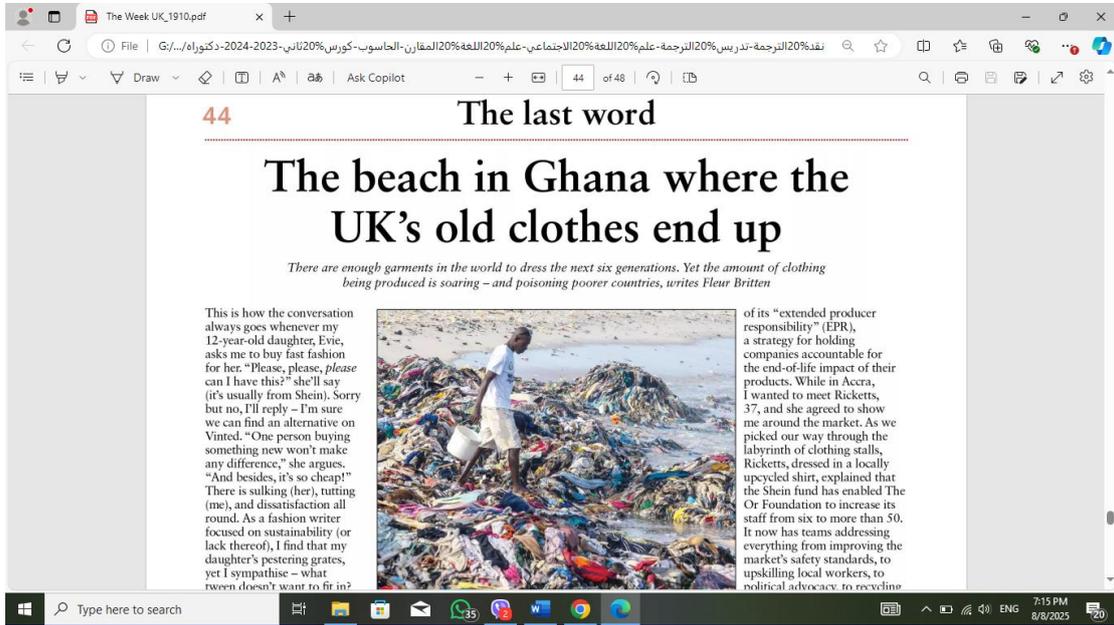
غانا هي إحدى أكبر دول العالم استيراداً للملابس المستعملة من الدول الغنية مثل المملكة المتحدة والولايات المتحدة، حيث تصل إليها 15 مليون قطعة ملابس كل أسبوع، وهي غالباً ملابس "مستهلكة يتخلص منها الأفراد في الدول المتقدمة."

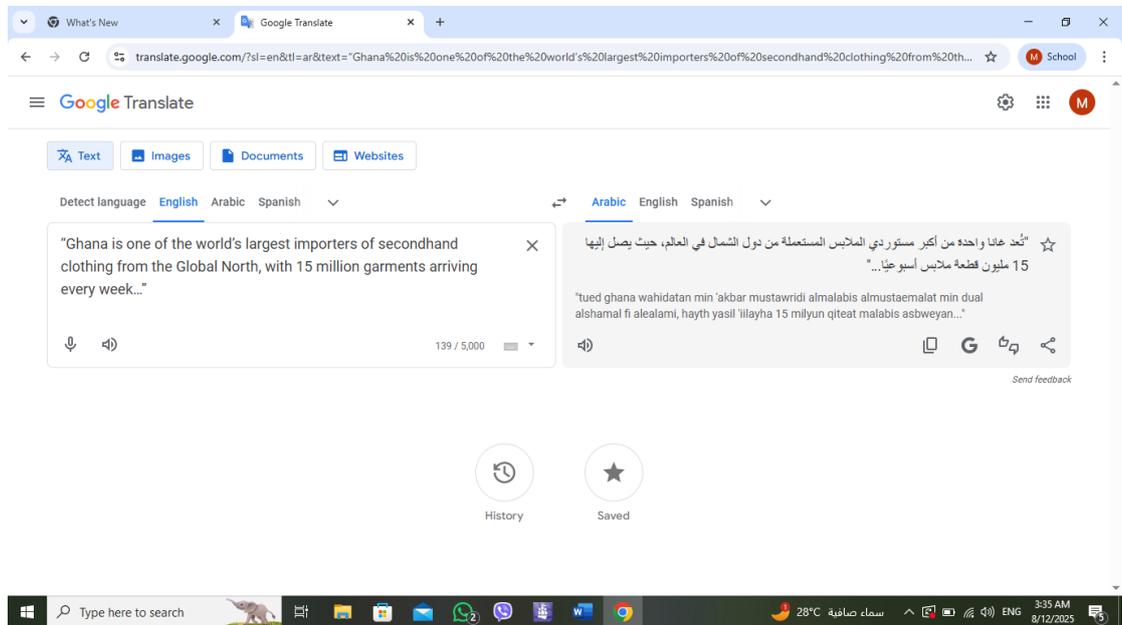
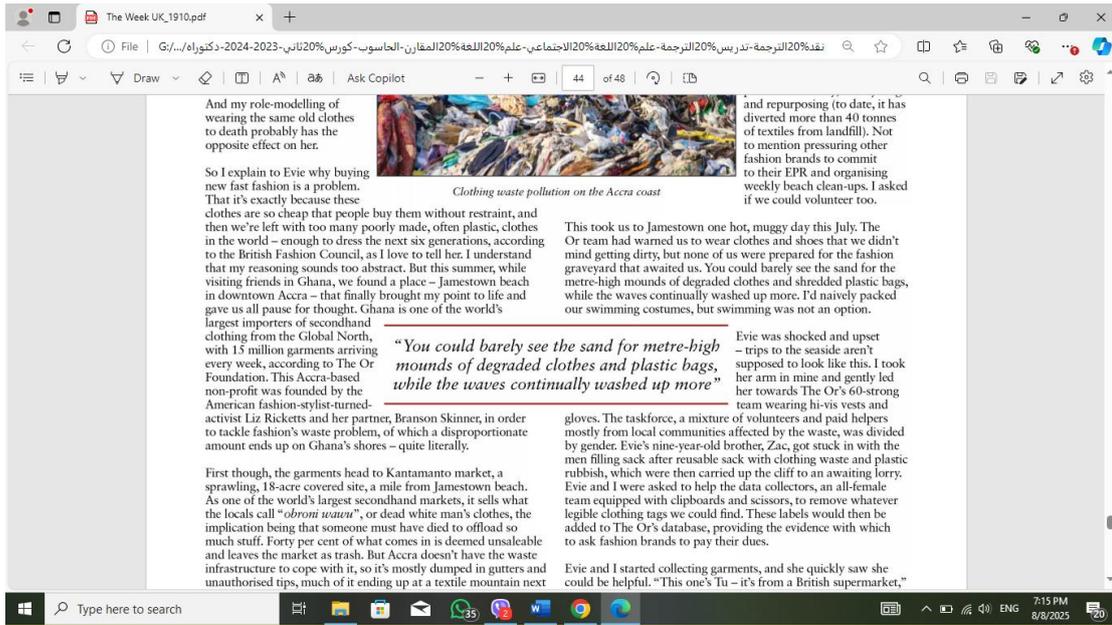
Discussion

_"Global North" was clarified to "wealthier countries like the UK and US," making it instantly understandable.

_ Context was added to explain that the clothes are "often discarded by consumers in developed nations," highlighting the global power dynamic and nature of the trade.

_ The final Arabic text is more informative, accessible, and impactful than the original MT, transforming a simple fact into a story.





Excerpt no. (6)

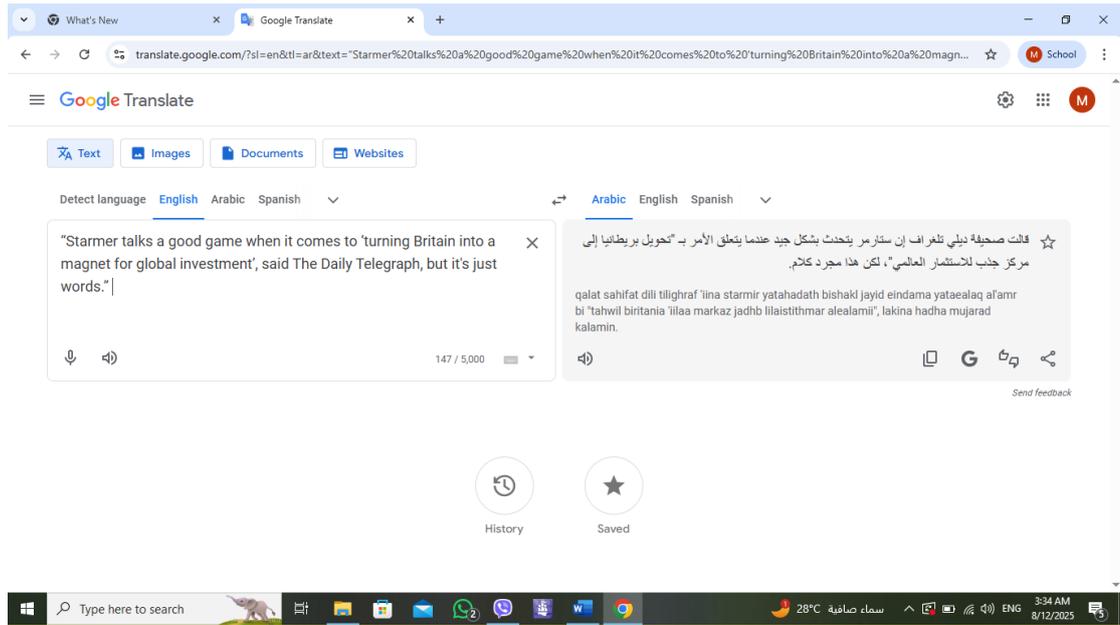
((The Week UK_1910))

Page 4

“Starmer talks a good game when it comes to ‘turning Britain into a magnet for global investment’, said The Daily Telegraph, but it's just words.”

MT:

قالت صحيفة "ذا ديلي تلغراف" إن "ستارمر يتحدث بلغة طيبة عندما يتعلق الأمر بتحويل بريطانيا إلى مغناطيس للاستثمار العالمي، لكنها مجرد كلمات".



Excerpt no. (7)

((The Week UK_1910))

Page 17

“Antarctica is turning from dazzling white to a dirty green at an alarming rate..

MT:

"تتحول القارة القطبية الجنوبية من اللون الأبيض المبهر إلى اللون الأخضر القذر بمعدل مثير للقلق.

After pre-editing:

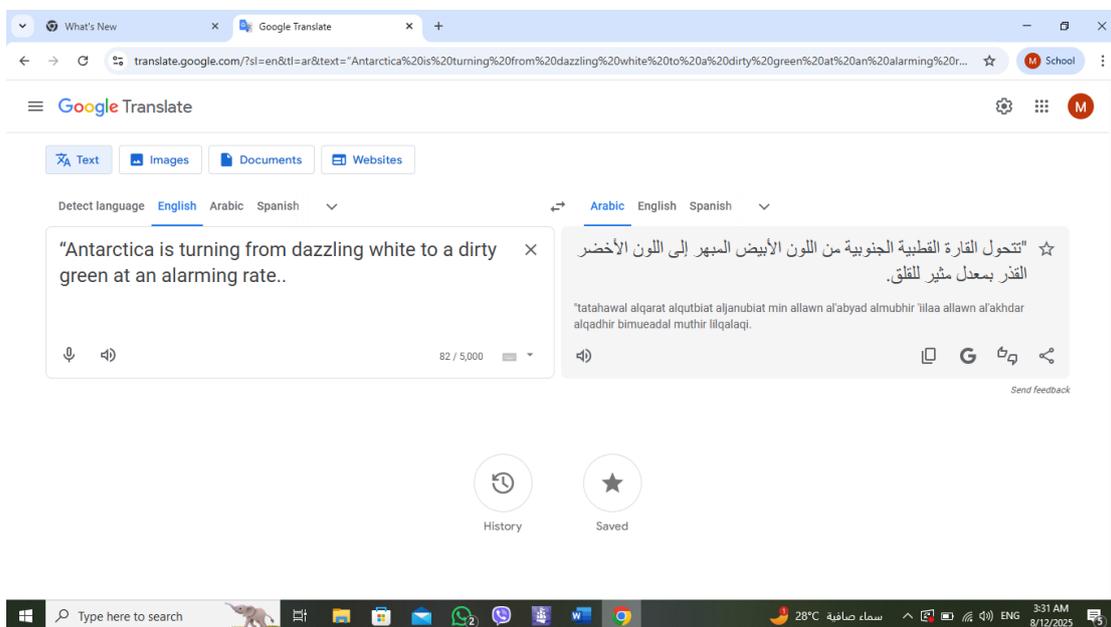
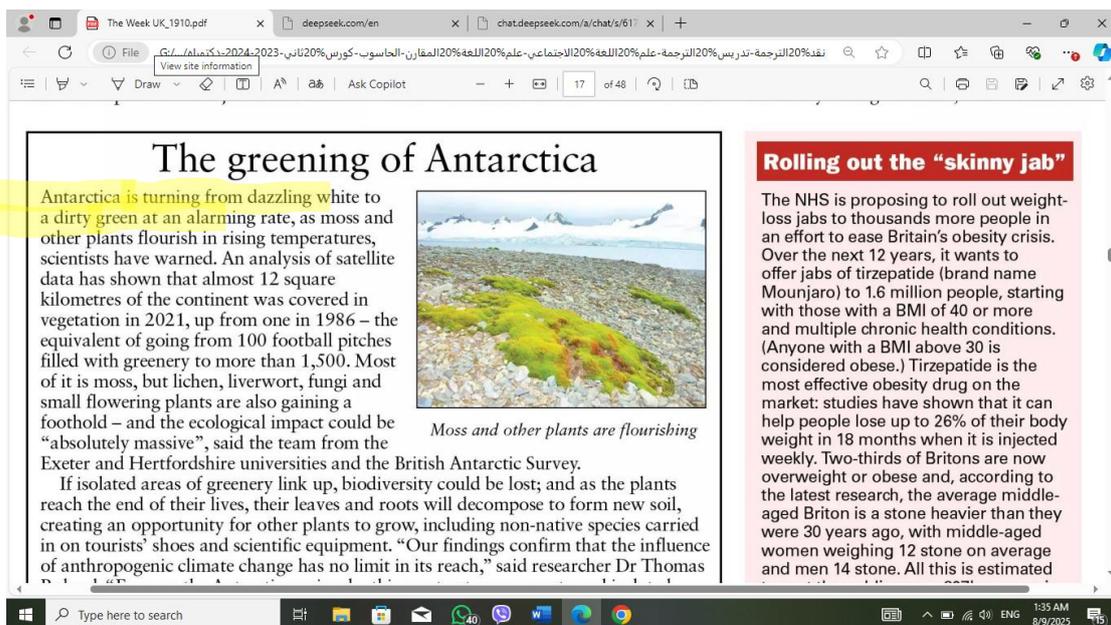
“Antarctica’s pristine white landscape is rapidly transforming into murky green due to algal growth.”

" يتحول المنظر الأبيض النقي للقارة القطبية الجنوبية بسرعة إلى لون أخضر قاتم بسبب انتشار الطحالب".

Discussion

Pre-edited (for clarity/flow):

- Clarifying “dirty green” as “murky green”
- Clarifying causality



Excerpt no. (8)

The Week US_151-

"Sin taxes" (p.34)

MT:

ضرائب الخطيئة

After pre-editing:

Levies on vices (alcohol, tobacco).

- Pre-edit: "[Taxes on alcohol, tobacco, etc.]."

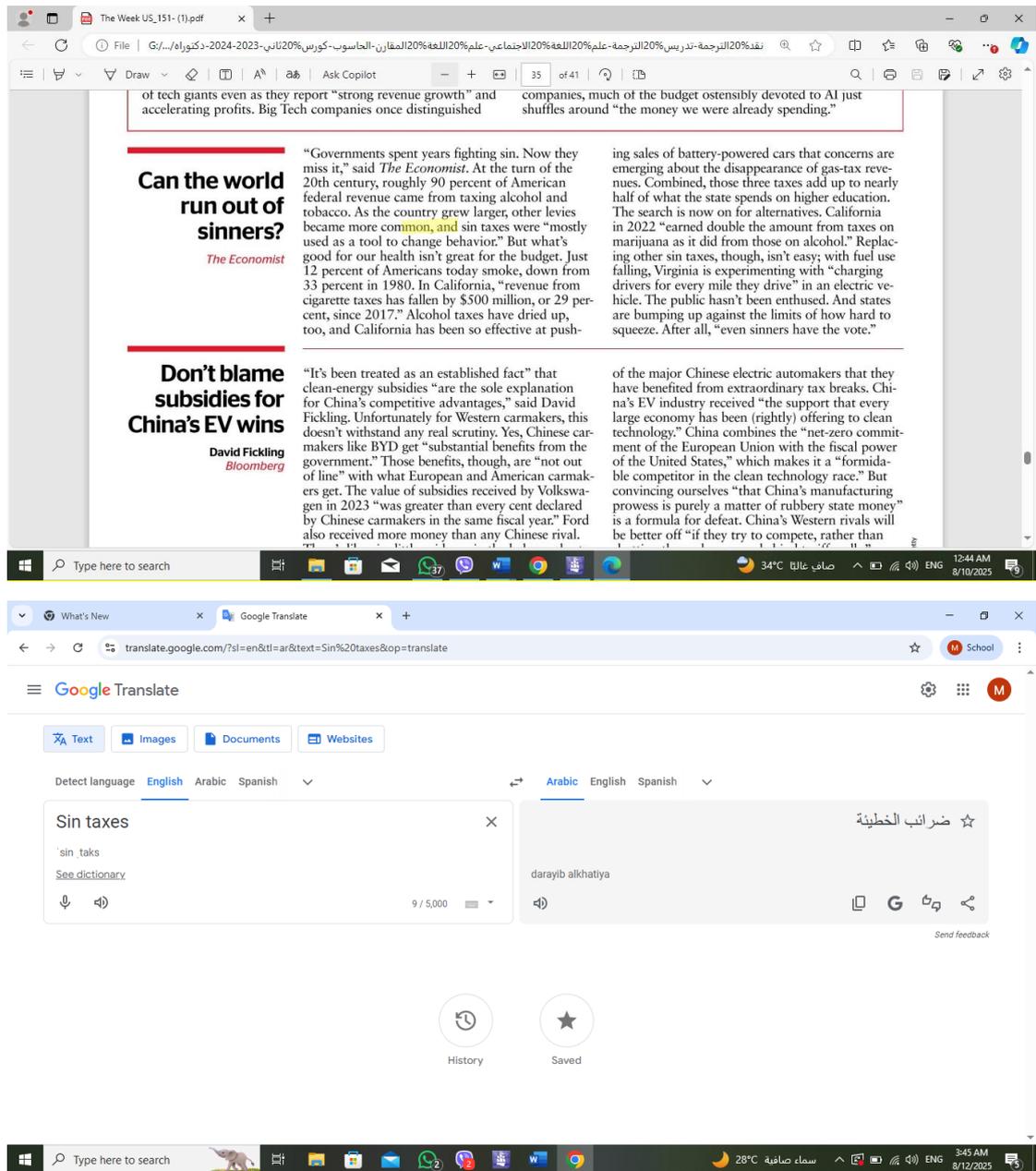
الضرائب على السلع الضارة مثل الكحول و التبغ.. الخ

Analysis:

_ The literal translation "ضرائب الخطيئة" (Taxes of Sin) is an abstract, culturally-loaded term that could confuse readers.

_ It was replaced with a clear, descriptive definition: "الضرائب على السلع الضارة" (Taxes on harmful goods).

_The pre-edited version adds crucial examples (alcohol, tobacco) and immediately clarifies the term's meaning for the reader, making it far more understandable than the direct translation.



Excerpt no. (9)

The Week UK_1910

Reading a good book is a 'beautiful experience'. But for most people today, it is one of many."

MT:

قراءة كتاب جيد تجربة جميلة. لكن بالنسبة لمعظم الناس اليوم، هي تجربة واحدة من تجارب عديدة.

After pre-editing:

Reading a good book is a rewarding experience, but today it competes with many other forms of entertainment."

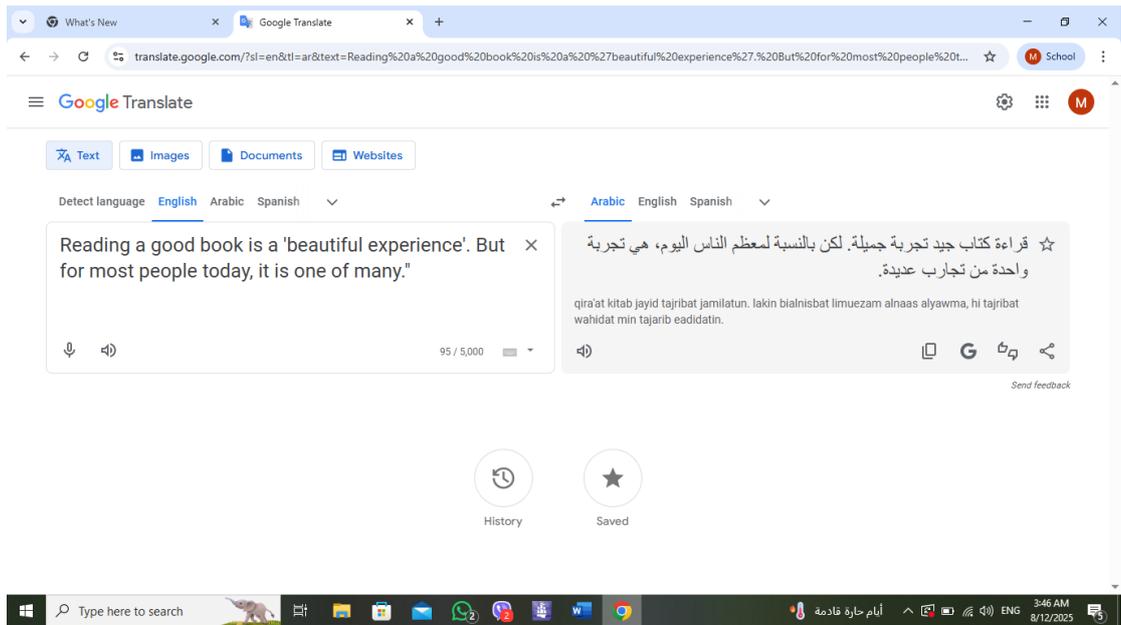
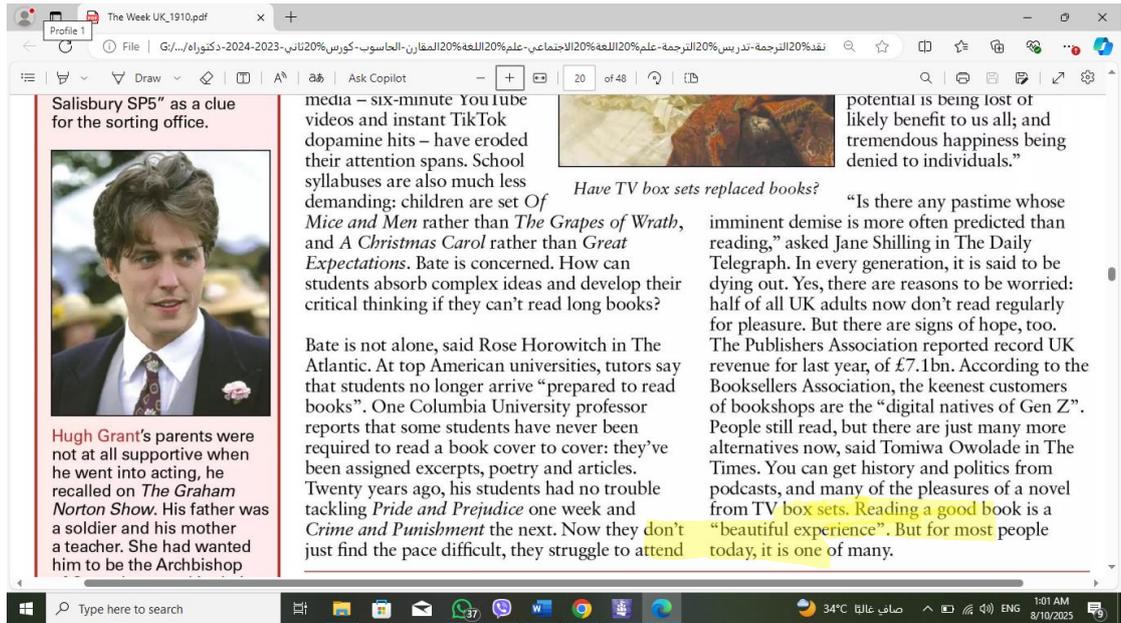
قراءة كتاب جيد هي تجربة مجزية، لكنها اليوم تتنافس مع أشكال عديدة من الترفيه.

Analysis:

_ The original translation was too literal ("one of many"), making the meaning vague.

_ It interpreted the core message: that reading now competes for people's attention.

_ The phrase "تتنافس مع أشكال عديدة من الترفيه" (competes with many forms of entertainment) is far clearer and more meaningful than the literal "one of many experiences." It accurately conveys the modern challenge reading faces.



Excerpt no. (10)

((The Week US_151-))

Page 4

‘Trump’s diverse coalition, in other words, looks a lot like America.’

MT:

"وبعبارة أخرى، فإن ائتلاف ترامب المتنوع يشبه إلى حد كبير أميركا".

After pre-editing:

‘Trump’s diverse coalition reflects the demographics of America.’

التحالف المتنوع لترامب يعكس التركيبة السكانية لأمريكا

Analysis

MT does not capture the intended meaning of the excerpt, so pre-editing is needed.

The image shows a screenshot of a PDF document and a Google Translate interface. The PDF document is titled "Contents" and features an "Editor's letter" by Theunis Bates, Editor-in-chief. The letter discusses Donald Trump's coalition and its similarity to America. The Google Translate interface shows the English text: "Trump's diverse coalition, in other words, looks a lot like America." and its Arabic translation: "وبعبارة أخرى، فإن ائتلاف ترامب المتنوع يشبه إلى حد كبير أمريكا." The Arabic translation is marked with a star, indicating it is a preferred translation.

Excerpt no. (11)

((((Becoming _a_translator_3rd edition)))

Page 168

Now pair off and create social interactions such as Reiß and Vermeer discuss..

MT:

..والآن، شكّل ثنائياً وأنشئ تفاعلات اجتماعية، حيث يناقش رايس وفيرمير

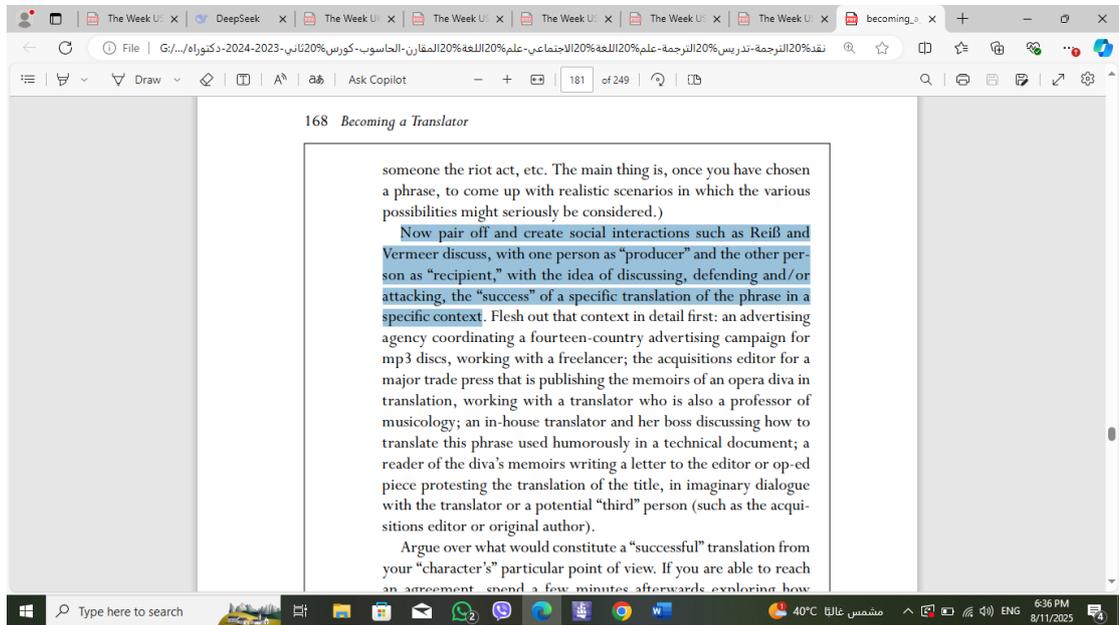
After pre-editing:

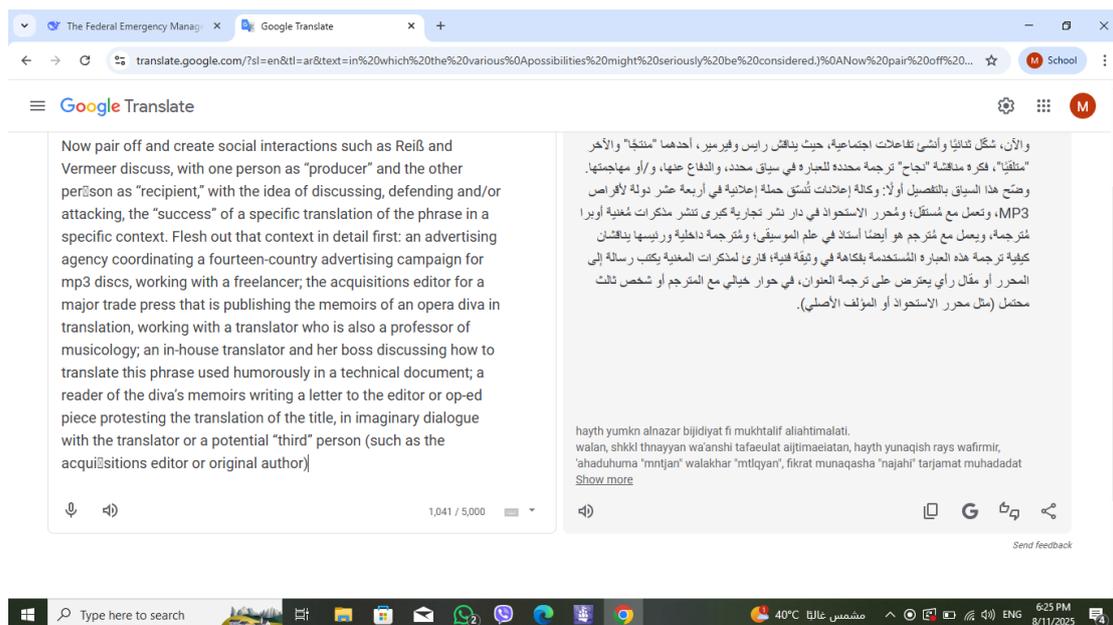
Now pair off and create social interactions such as Katharina Reiß and Vermeer discuss..

..والآن، شكّل ثنائياً وأنشئ تفاعلات اجتماعية، حيث يناقش كل من كاثرينا رايس وفيرمير

Discussion

MT mistakes Reiss for a male, so pre-editing is needed.





Excerpt no. (12)

((((Becoming _a_translator_3rd edition)))

Page 130

The trick, then, is to convince other people that this "slightly new" thing you've done with words in fact is a reliable reproduction of the old thing done by the source author or speaker

MT:

الحيلة، إذن، هي إقناع الآخرين بأن هذا الشيء "الجديد قليلاً" الذي فعلته بالكلمات هو في الواقع إعادة إنتاج موثوقة للشيء القديم الذي فعله المؤلف أو المتحدث المصدر.

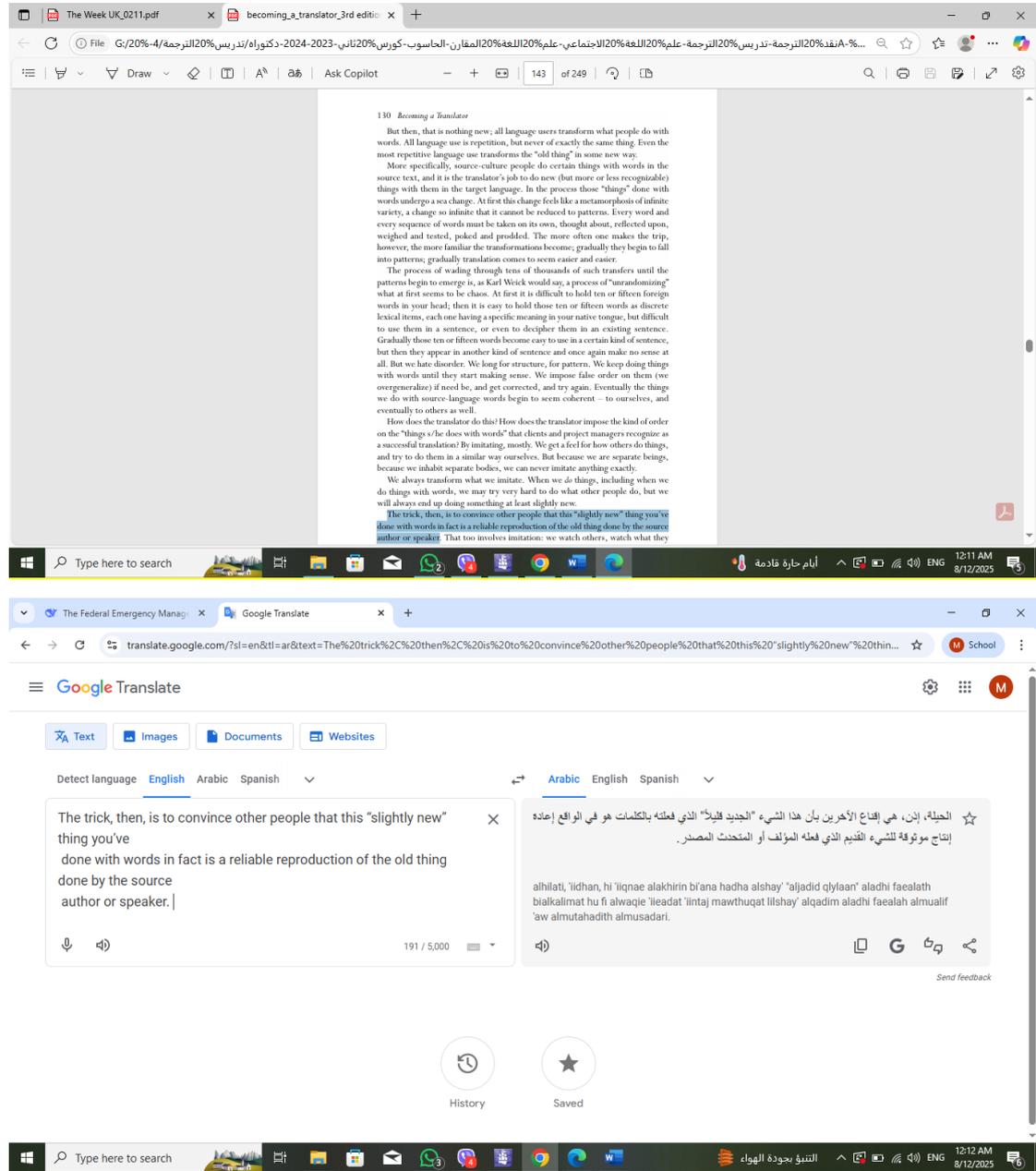
After pre-editing:

The challenge, then, is to convince other people that this "slightly new" thing you've done with words in fact is a reliable reproduction of the old thing done by the source author or speaker.

إذن فالتحدي الحقيقي هو إقناع الآخرين بأن هذا "الإبدال اللفظي الجديد" الذي قمت به - في جوهره - إعادة إنتاج أمين للنص أو الخطاب الأصلي لمؤلفه الأول.

Discussion

MT mistranslates ‘trick’, so pre-editing is needed.



24- Findings

Teaching pre-editing enhances students’ awareness of MT limitations. It fosters stylistic sensitivity and source-text optimization skills. It encourages a shift from reactive post-editing to proactive translation planning.

Based on the detailed analysis, it has been found that all the hypotheses are validated:

1. Pre-editing improves MT output quality

Evidence: Pre-edited texts showed fewer errors in accuracy and fluency, especially in technical and journalistic domains.

2. Post-editing takes less time when pre-editing is conducted

Evidence: Time measurements and keystroke logs confirmed reduced effort and faster post-editing.

3. Cognitive load during post-editing is alleviated by pre-editing

Evidence: Subjective ratings and task performance showed lower mental strain when pre-editing was applied.

4. Pre-editing enhances translators' awareness of MT limitations

Evidence: Translators demonstrated improved anticipation of MT errors and better strategic planning.

5. Better translation quality is obtained when applying pre-editing

Evidence: Hybrid workflows (pre-editing + post-editing) yielded near-human quality with minimal revisions.

26- Conclusion

This study has demonstrated that pre-editing plays a pivotal role in enhancing the efficiency and quality of post-editing within machine translation workflows. By strategically modifying source texts before translation, pre-editing significantly reduces cognitive load, accelerates post-editing time, and improves the overall accuracy and fluency of the final output. The comparative analysis between raw and pre-edited MT outputs confirms that pre-editing is not merely a preparatory step—it is a transformative intervention that reshapes the translation process.

Pedagogically, the study advocates for embedding pre-editing strategies into translator training programs. These strategies—such as simplifying syntax, clarifying ambiguity, and managing terminology—

equip translators to collaborate more effectively with MT systems. Ethically, pre-editing also promotes transparency and control over machine-generated content, ensuring that human oversight remains central to translation quality.

Pre-editing is not a redundant or optional task—it is a proactive, quality-driven approach that enhances human-AI synergy in translation. As MT continues to evolve, the role of pre-editing will become increasingly vital in maintaining linguistic precision, cultural fidelity, and professional standards across multilingual contexts.

References

1. Allen, J. (2003). Post-editing. In *Computers and translation: A translator's guide* (pp. 297–317).
2. AlOtaibi, A. (2020). Statistical MT training for the translation of English-Arabic UN Resolutions. *Arab World English Journal*, (239), 1–102.
3. Alvarez-Vidal, S., et al. (2020). Post-editing for professional translators: Cheer or fear? *Tradumàtica: Technologies de la Traducció*, 18, 49–69. <https://doi.org/10.5565/rev/tradumatica.275>
4. Bernth, A., & Gdaniec, C. (2001). MTranslatability. *Machine Translation*, 16, 175–218.
5. Bounaas, C., Zemni, B., Shehri, F. A., & Zitouni, M. (2023). Effects of PrE operations on audiovisual translation using TRADOS: An experimental analysis of Saudi students' translations. *Texto Livre*, 16, 1–5. <https://doi.org/10.1590/1983-3652.2023.45539>
6. Brynjolfsson, E. (2022). The Turing trap: The promise & peril of human-like artificial intelligence. *Daedalus*, 151(2), 272–287. <https://doi.org/10.1162/daeda01915>
7. Dai, J.-n., Shen, H.-j., Zhuo, J., Pan, X.-y., & Yuan, Z. Y. (2022). The differences between machine translation and human translation from the perspective of literary texts. *International Journal of Arts and Social Science*, 5(10), 112–133. <https://www.ijassjournal.com>
8. Delavenay, E., & Delavenay, K. M. (1960). *An introduction to machine translation*.
9. Doherty, S. (2016). *The impact of translation technologies on the process and product of translation*.
10. Dragan, J. (2013). *Quality in professional translation: Assessment and improvement*. Bloomsbury.

11. Girletti, S. (2024). Working with pre-translated texts: Investigating machine translation post-editing and human translation revision at Swiss corporate in-house language services [Doctoral thesis, Université de Genève]. Archive ouverte UNIGE. <https://archive-ouverte.unige.ch/unige:175505>
12. Görög, A. (2014). Quality evaluation today: The dynamic quality framework. In Proceedings of Translating and the Computer 36 (pp. 155–164).
13. Guerberof Arenas, A. (2020). Pre-editing and post-editing. In J. Schwieter & A. Ferreira (Eds.), The handbook of translation and cognition (Chapter 15). Bloomsbury Academic. <https://doi.org/10.5040/9781350024960.0019>
14. Hadi, M. U., Tashi, Q. A., Qureshi, R., Shah, A., Muneer, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., & Mirjalili, S. (2023). A survey on large language models: Applications, challenges, limitations, and practical usage [Preprint]. <https://doi.org/10.36227/techrxiv.23589741.v1>
15. Hiraoka, Y., & Yamada, M. (2019). Pre-editing plus neural machine translation for subtitling: Effective pre-editing rules for subtitling of TED Talks. Proceedings of MT Summit XVII, 2, 64–71.
16. Hudáková, Z. (2024). Comparing the efficiency of source text pre-editing vs. machine translation post-editing. LION Journal, 2(3), 42–59.
17. Hutchins, J., & Lovtskii, E. (2000). Petr Petrovich Troyanskii (1894–1950(sad) A forgotten pioneer of mechanical translation.
18. Hutchins, W. J. (1986). Machine translation: Past, present, future. Ellis Horwood.
19. Koby, G. S. (2001). Editor's introduction – Post-editing of machine translation output: Who, what, why, and how (much). In Repairing texts: Empirical investigations of machine translation post-editing processes (pp. 1–24). The Kent State University Press.
20. Kokanova, E., et al. (2022). Pre-editing English news texts for machine translation into Russian. Language Studies and Modern Humanities, 4(1), 25–30. <https://doi.org/10.33910/2686-830x-2022-4-1-25-30>
21. L. G. (1995). History of translation. In E. F. K. Koerner & R. E. Asher (Eds.), Concise history of the language sciences: From the Sumerians to the cognitivists. Elsevier Science & Technology.
22. Lauscher, S. (2000). Translation quality assessment: Where can theory and practice meet? The Translator, 6(2), 149–168. <https://doi.org/10.1080/13556509.2000.10799063>
23. Li, H., Graesser, A. C., & Cai, Z. (2014). Comparison of Google translation with human translation.
24. Lin, J., Kovacs, G., Shastry, A., Wuebker, J., & DeNero, J. (2022). Automatic correction of human translations. In Proceedings of the 2022

- Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 494–507). <https://doi.org/10.18653/v1/2022.naacl-main.36>
25. Lommel, A. R., & De Palma, D. A. (2016a). Post-editing goes mainstream. Common Sense Advisory. <http://www.commonsenseadvisory.com/Portals/default/Knowledgebase/ArticleImages/1605RIPPosteditinggoesmainstream-extract.pdf>
26. Marshall, C. (2024). A comparative analysis of human and machine translation quality [Master's thesis, Brigham Young University]. BYU ScholarsArchive. <https://scholarsarchive.byu.edu/etd/10526>
27. Miyata, R., & Fujita, A. (2017). Investigating the effectiveness of pre-editing strategy and the diversity of pre-edit operations for better use of machine translation. *Invitation to Interpreting and Translation Studies*, 18, 53–72.
28. Nagao, M. (1984). A framework of a mechanical translation between Japanese and English by analogy principle. In *Artificial and human intelligence*. North-Holland.
29. O'Brien, S. (2003). Controlling controlled English: An analytical of several controlled language rule sets. In *Proceedings of EAMT-CLAW 2003* (pp. 105–114). EAMT.
30. Poibeau, T. (2017). Example-based machine translation. In *Machine Translation* (pp. 109–119). MIT Press.
31. Robinson, D. (2012). *Becoming a translator: An introduction to the theory and practice of translation* (3rd ed.). Routledge.
32. Sophie, E. R. (2025). Machine translation vs. human translation in industry settings. Obafemi Awolowo University. <https://www.researchgate.net/publication/390979035>
33. The Week. (2024, October 19). The Week UK (Issue 1510). Future PLC. <https://www.theweek.co.uk>
34. Yuen, J. C. (2021). Embracing AI-assisted machine translation to broaden citations in journal.