The Potential of Large Language Models for Translating Tourism Promotional Texts: A Mixed-methods Study

Raghad Alsulami
Centre for Translation Studies (CenTraS)
University College London
Raghad.alsulami.23@ucl.ac.uk

Abstract

This paper reports on a small-scale pilot study examining the potential of a large language model (LLM) for translating tourism promotional texts (TPTs), in comparison with a conventional neural machine translation (NMT) system, from English into Arabic. Four professional translators participated in a post-editing experiment followed by cue-based retrospective interviews. The post-editing task aimed to provide empirical evidence of the effort involved in working with TPTs, while the interviews sought to capture participants' judgments and evaluations of the outputs. Overall, most participants exerted less effort post-editing LLM-generated outputs to a publishable standard compared to NMT outputs. They perceived the LLM outputs to be more creative, with creativity manifested through non-literal translations and aesthetic augmentation, while also noting that the outputs were unpredictable and far from perfect; in contrast, the NMT outputs were generally viewed as more informative yet lacking the promotional appeal required for TPTs. The paper concludes with implications and directions for future research.

1 Introduction

The rise in international travel, driven by advances in globalization and technology, has positioned tourism as a key industry in the global economy. In order to remain competitive in the global tourism market, destinations invest in the production of a wide range of promotional materials to attract visitors. Such materials range from traditional brochures to digital platforms such as social media channels and websites dedicated entirely to promoting tourism. Given that potential tourists represent a wide range of linguistic and cultural backgrounds, a significant amount of translation work is carried out in the tourism industry (Katan, 2011; Kelly, 1997; Sulaiman & Wilson, 2018).

In parallel, recent years have witnessed the rise of generative artificial intelligence (AI) following the emergence of LLMs. This transformative wave has affected various sectors, and the tourism industry is no exception. For instance, a recent report by the European Travel Commission, which surveyed 29 national tourism organizations in Europe, found that marketing departments perceive LLMs to be valuable for copywriting, positioning them as both early adopters and opportunistic users of the technology (Kairos Future, 2025). Research also shows that potential tourists value LLM-generated TPTs as highly as those produced by human marketers (Zhang & Prebensen, 2024).

While translation is crucial for tourism marketing, very little research has examined how LLMs handle this specific task. Translating TPTs often requires a degree of linguistic and cultural adaptation to ensure they resonate with the target audience, making them a demanding yet insightful test case for LLMs. Against this backdrop, this paper aims to explore the potential of an LLM, used with a tailored prompt, in comparison with a conventional NMT system, for translating TPTs from the perspective of professional translators. It adopts a mixed-methods design comprising a post-editing experiment, in which effort data (e.g., time, keystrokes, and pauses) are collected, and semi-structured interviews involving cue-based retrospection. Post-editing as an evaluation method is used because it likely represents the closest real-world scenario for the use of AI-generated outputs when they are intended for dissemination. It also

offers the advantage of eliciting translators' perceptions and evaluations of the outputs via post-hoc interviews, grounded in their actual use of and direct engagement with the outputs.

The decision was made to include NMT in this study both to contextualize the findings and because some studies suggest that NMT is being used with tourism texts (e.g., Janodet & Jurado, 2025). Including NMT as a baseline therefore allows us to empirically examine what the LLM paradigm enables or expands beyond conventional approaches, particularly in ways that were not previously possible or not easily achievable. Ethical approval to conduct the study was obtained from the Research Ethics Committee at University College London, with the reference AH/2025/21.

2 The Language of Tourism

Dann (1996) argues that "the language of tourism attempts to persuade, lure, woo and seduce millions of human beings, and, in so doing, convert them from potential into actual clients" (p. 2). Several verbal techniques can be employed to achieve persuasion. For instance, Dann (1996) highlights strategies such as ego-targeting, which directly addresses the readers to make them stand out from the crowd. Another strategy is languaging, which refers to the striking use of foreign terms to spark readers' interest (Dann, 1996). A further strategy is keying, which refers to the strategic use of evaluative words to enhance the appeal of a tourist destination (Sulaiman & Wilson, 2019). Keying is perhaps the most widely used strategy; it involves using glowing and highly euphoric terms (e.g., breathtaking, vibrant, unique) to highlight the positive aspects of a destination (Sulaiman & Wilson, 2019), often with a tendency toward exaggeration (Malamatidou, 2024).

3 The Functions of TPTs

Reiss (1981) proposed a functional text typology that distinguished between informative texts, which prioritize the transmission of content; operative texts, which aim to influence attitudes or behavior; and expressive texts, which foreground aesthetic and stylistic qualities. Scholars argue that TPTs combine all three functions (Snell-Hornby, 1999), albeit to different degrees. They inform potential visitors about destination highlights and attractions; they also seek to persuade potential tourists to visit, an objective that aligns them closely with advertising (Dann, 1996; Sanning, 2010; Sulaiman, 2016). The dual info-promotional functions are regarded as the predominant ones (Muñoz, 2012). Sanning (2010) argues that the informative elements in TPTs serve as the "premise" for accomplishing the broader purpose of persuading potential tourists (p. 125). Put differently, the promotional purpose can only be achieved if the informative components are in place. The expressive function is also observed in TPTs in the sense that creative language and rhetorical devices are used to enhance engagement and support the intended persuasive effect of the texts (Dann, 1996; Snell-Hornby, 1999).

Considering these functions and the wide range of linguistic and cultural backgrounds of potential tourists, being attentive to cultural practices, values, and norms becomes essential to ensure that translated TPTs resonate with their target audiences (Maci & Spinzi, 2025). This explains why a degree of linguistic and cultural adaptation is often necessary in their translations, both to address the expectations of diverse audiences and to amplify their persuasive impact (Maci & Spinzi, 2025; Malamatidou, 2024; Sulaiman & Wilson, 2019).

4 AI and Translating TPTs

Research into the application of AI for translating creative texts has so far focused on literature, audiovisual texts, and video games (e.g., Brenner, 2024; Guerberof-Arenas et al., 2024; Toral et al., 2018). However, its potential for translating TPTs remains largely underexplored. This is perhaps unsurprising as TPTs typically demand a degree of linguistic and cultural adaptation to achieve their dual functions of informing and persuading potential tourists to visit a given destination. Until late 2022, this posed a considerable challenge for machines; yet, the emergence of LLMs that can be prompted with task-specific requirements is believed to have transformed this landscape (Navarro, 2025; Smartling, 2024). Among the several advantages LLMs can offer are context-awareness and customized translations (Lye et al., 2024), which can be hard to achieve using traditional NMT systems. Customizing translations via prompt engineering techniques forms a major development. Users, for example, can adjust the level of formality, tone, and style. Information on the purpose of the translation and target readers can also be included in the prompt, although there is no guarantee that LLMs will properly adhere to such demands from users.

While still in its infancy, a body of research exploring the use of AI for translating TPTs is beginning to gain momentum. Giampieri and Harper (2022) compared Italian-English tourism translations produced by NMT with those produced by human translators. Using a promotional text for a beach resort, they found that NMT struggled with collocations, figurative language, and conventions of English tourism writing; however, they noted that it performed well with informative tourism texts. In addition, Li and Li (2025) compared student post-editing (SPE) and GPT-based post-editing (GPTPE) in Chinese-English translations of TPTs. They found that GPTPE excelled in lexical diversity, while SPE showed greater adherence to target-language conventions, reflecting students' strategic application of their translation knowledge. Moreover, Navarro (2025) studied the use of LLMs for the transcreation of TPTs from Portuguese to English. Her analysis showed that LLMs produced impressive results, generating outputs with sophisticated phrasing and enhanced persuasive tone. Lastly, Spinzi (2025) examined how GPT-4 handles the translation of metaphors from Italian to English in texts promoting sustainable tourism. The results showed that metaphorical expressions were at times literally rendered, making the texts sound less natural for their target readers.

The present pilot study, which is part of a larger doctoral research project, aims to contribute to this line of work by conducting an empirical study with professional translators on the use of LLMs relative to NMT for translating TPTs in the English-Arabic language pair. Through a post-editing experiment and cue-based retrospective interviews, the study offers insights into the potential and limitations of AI tools when translating such content. It also seeks to inform tourism boards with limited budgets (e.g., Napu, 2016) and/or an interest in adopting these technologies, especially LLMs, which represent the latest development in machine translation (MT).

5 Methodology

5.1 Participants

Four participants were recruited for the pilot study and are referred to here as T1, T2, T3, and T4. Recruitment was carried out through a public post shared on both LinkedIn and X (formerly known as Twitter), where potential participants were invited to express their interest in taking part in the pilot study. The criteria for selecting participants included being a native Arabic speaker, working as a professional translator, and having experience in translating tourism promotional content.

The group included two males and two females, with an average age of 32 years. Participants self-assessed their English proficiency using the common European framework of reference for languages, rating themselves at the C1 (advanced) level, except for T2, who rated their proficiency as C2 (proficient/native-like). They all held a Bachelor's degree as their highest educational qualification, except for T3, who held a Master's degree. Participants were working on a freelance basis at the time of data collection. Regarding their general professional translation experience, they all had over five years of experience, except for T3, who had between three and five years. In terms of translating tourism promotional content, their experience varied from a one-time project for T3, to one year for T1 and T4, and up to four years for T2. All participants were compensated upon completion of the entire study.

5.2 Source Texts

Considering the general short length of TPTs, it was necessary to select multiple texts to ensure sufficient and meaningful data. The study materials consisted of 14 English TPTs obtained from the VisitBritain website, which promotes tourism in various destinations across the United Kingdom (UK), including cities, coastal areas, and rural regions in England, Wales, Scotland, and Northern Ireland. Each text focused on a distinct location within the UK. Most texts shared a consistent style and logical structure: a standardized title beginning with "why we love [a given place]", followed by an introduction, one or two short paragraphs highlighting key features of the destination, and a conclusion or closing line that guided readers on how to plan their trip. The consistent structure across the selected texts, along with the similar writing style, suggests that they were likely produced by the same marketing team, which helps ensure a degree of comparability between them.

While the website is translated into Arabic, raising the possibility that its translations were used to train the chosen LLM, visual inspection revealed little resemblance between the LLM outputs and the published versions. This was further supported by a pilot comparison using the human-targeted edit rate (HTER), which indicated that the AI-generated outputs and the human translations on the website differed considerably (Snover et al., 2006). Nevertheless, the website was selected for

the appropriate length of its texts compared to other tourism websites and for potential future reception-based studies using the available human translations, as no human translation condition was included in this study.

The texts chosen were similar in length, containing an average of 154 words, and were primarily segmented at the paragraph level in the post-editing tool used (see Appendix A for a sample task). This approach to segmentation was chosen to minimize constraints on the translators, as previous research has shown that working at the sentence or segment level can feel overly restrictive. For instance, participants in earlier studies perceived translating and post-editing creative texts at the segment/sentence level as "too limiting" (Daems, 2022, p. 59) and described it as "translation in the darkness" (Moorkens et al., 2018, p. 253), expressing a preference for handling larger portions of text. Moreover, coordinating sentences is a salient feature of Arabic, even in TPTs (Al-Fahad, 2012), which further supports the use of broader text segments.

5.3 Post-editing Tool

Participants used PET (Aziz et al., 2012), a popular research-based tool that captures data related to the three primary post-editing effort dimensions proposed by Krings (2001): temporal, technical, and cognitive effort. It produces detailed log files for each participant and each task containing information on time, number and types of keystrokes, as well as number and length of pauses. For keystrokes, it records every single key pressed by participants and groups them into separate categories: content keys (e.g., letters, digits, symbols, and whitespace), erase keys, navigation keys, and shortcut keys (e.g., cut, copy, paste, and undo). The tool also helps extract the final post-edited outputs, which were used to calculate HTER by comparing them against the raw outputs. While the number of keystrokes obtained from PET generally provides insight into the technical effort involved, it does not necessarily reflect how much of the outputs was altered in the final translations as HTER does. The PET version used was based on Sarti et al. (2022) who tweaked the tool so that it could support right-to-left visualization, which is necessary for Arabic.

5.4 MT Systems and Prompt

Google Translate was used as a representative of NMT while GPT-4o was used to represent LLMs. At the time of the experiment, GPT-4o was the most recent model released by OpenAI. The model outperformed earlier models in handling text in non-English languages, which was relevant to this study given the use of Arabic (OpenAI, 2024).

In designing the prompt for the LLM, both the communicative goal of the target texts and insights from prior research on effective prompt design were taken into account. Accordingly, the model was assigned a translator persona (He, 2024), and information about the target readers and purpose of the translation (i.e., to encourage readers to visit the destinations) was included (Yamada, 2023). The model was then instructed to adapt the texts into Arabic in a way that aligns with Arabic linguistic norms and meets the expectations and needs of Arabic-speaking potential tourists (see Appendix B for the full prompt used).

5.5 Design and Setup

This study employed a within-subject design in which each participant post-edited translations produced by the LLM and NMT system. This design enables direct comparison within the same translator, thereby controlling for individual differences (e.g., translation experience, cognitive abilities, typing speed, etc.). Given the repeated-measures design, it was essential to counterbalance and randomize tasks across participants (Saldanha & O'Brien, 2014). To explain, the 14 texts were first assigned random codes from 1-14. Participants were then split into two equal groups. For group 1 (T1 and T2), texts 1-7 were translated by NMT, while texts 8-14 were translated by the LLM. The assignment was reversed for group 2 (T3 and T4), with texts 1-7 translated by the LLM and texts 8-14 by NMT. This counterbalancing ensured that each text was translated an equal number of times by both systems in the dataset, while also preventing any participant from encountering the same text twice. The order of the 14 texts was then randomized for every participant using the Random.org list randomizer to mitigate order effects. Appendix C presents the full tasks for all participants, shown in the same order in which they were completed.

As for the setup, the experiment was conducted online both for practical reasons, since participants were based in different countries, and to enhance ecological validity by allowing them to work in their usual environments on their own computers. This setup avoids the constraints of laboratory settings, such as working with unfamiliar computers, which may impact participants'

behavior (Ehrensberger-Dow, 2014). To perform the experiment, participants used the remote desktop connection software AnyDesk to connect to a remote computer on which PET and other relevant tools were installed.

Before the experiment, participants received a translation brief outlining the purpose of the task and the functions of the materials (i.e., to inform and promote). They were instructed to post-edit the outputs to a publishable standard and to adapt the content as necessary to align with the target readers. Because recruiting translators with first-hand knowledge of all target destinations was difficult, and drawing on recommendations from Sulaiman and Wilson (2019), participants were also given a document listing the 14 destinations along with useful sources for background reading so that they could familiarize themselves.

During the experiment, participants were allowed to consult any online resources they would normally use in professional practice. This was particularly important given the texts used, which included background-specific details and cultural elements that might have required participants to use external resources. The only exception was the use of AI to re-translate entire paragraphs, which was not permitted.

After the experiment, participants took part in a short semi-structured interview, during which part of the discussion involved cue-based retrospection to reflect on some of the outputs they had post-edited. This was made possible by screen-recording the entire post-editing session via FlashBack Express. It should be noted, however, that while cue-based retrospection is often used to investigate cognitive processes (see Tiselius et al., 2025), in this study it served to elicit translators' subjective reflections and evaluations of the raw outputs they worked on in order to gain additional insights into the quantitative data obtained from the experiment. Rather than requesting participants to provide a verbal report for every sample shown, they were guided through targeted questions designed to elicit focused and meaningful responses (for the set of questions, see Appendix D). At least two outputs were shown to each participant, one produced by the LLM and one by the NMT system. The specific outputs differed between participants as selection was informed by real-time observations made during their post-editing process. By drawing on a broader and more diverse set of texts, this approach was expected to yield richer and more insightful qualitative data than using a fixed pre-determined set of samples for all participants.

6 Results and Analysis

In this section, the experimental data are first reported, followed by the interview data. Due to the small number of participants (n = 4), and, consequently, the limited number of data points (n = 255),[2] no statistical modeling was performed. That said, the preliminary results presented here still allow for an initial comparison between the two systems when translating TPTs, based on participants' post-editing effort. The Python script written by Toral et al. (2018) was used to process the log files. Data were processed and summarized using the dplyr package (Wickham et al., 2016) and visualized with the ggplot2 package (Wickham, 2016) in R (R Core Team, 2026).

6.1 Temporal Effort

Temporal effort was operationalized as the time spent on post-editing, measured in seconds and normalized by the number of words in the MT output. For each participant, the time per word was averaged across the seven texts generated by each MT system. Figure 1 shows the average temporal effort for each participant when post-editing the LLM and NMT outputs.

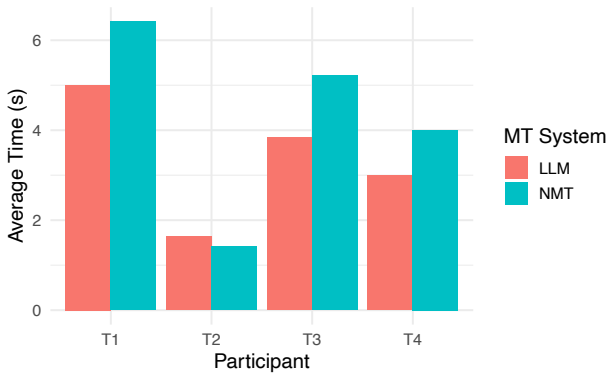

Figure 1: Post-editing time (in seconds) per participant and MT system

Based on Figure 1, there is clear variability across the four translators, which is expected given the natural individual differences in human performance. Nevertheless, a clear trend emerges: participants T1 (LLM: M = 5.00, SD = 5.48; NMT: M = 6.41, SD = 4.77), T3 (LLM: M = 3.84, SD = 3.38; NMT: M = 5.22, SD = 2.80), and T4

[2] The original dataset included 256 observations; however, participant T3 skipped a segment by mistake.

(LLM: $M = 2.99$, $SD = 2.82$; NMT: $M = 4.00$, $SD = 3.09$) spent less time when working with the LLM than with NMT. This may suggest that LLM-generated translations of TPTs were superior to those produced by NMT, requiring less time to post-edit. The only exception was participant T2 (LLM: $M = 1.63$, $SD = 0.63$; NMT: $M = 1.41$, $SD = 0.89$) who represents an outlier, as they spent the least time overall and showed comparable times across both systems. During the interview, T2 reported that they had forgotten to take the guidelines into account while post-editing, which may help explain their minimal temporal effort across both systems.

6.2 Technical Effort

Technical effort was assessed by dividing the number of keystrokes by the number of words in the MT output. As with temporal effort, keystrokes per word were averaged across the seven texts generated by each MT system for every participant. Figure 2 shows the average technical effort per participant for the LLM and NMT outputs, further broken down into content keys, erase keys, and navigation keys.

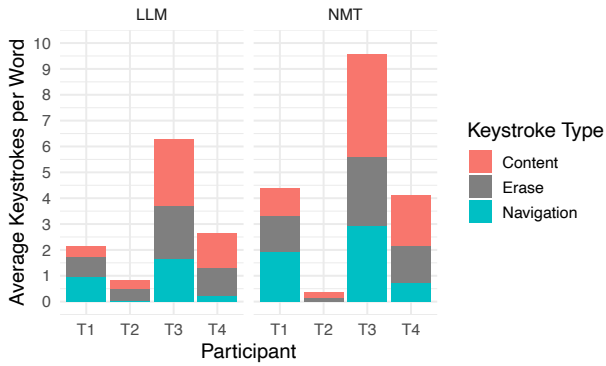

Figure 2: Number and types of keystrokes per participant and MT system

Similar to temporal effort, and as expected, there is clear variability among the four translators in terms of the number of keystrokes produced during post-editing. Despite these individual differences, on average, participants required 2.91 keystrokes per word ($SD = 4.75$) when post-editing the LLM outputs compared with 4.67 keystrokes per word ($SD = 5.47$) for NMT outputs, with T2 again acting as an outlier. In fact, T2 produced more keystrokes with the LLM outputs, with the majority being erase keys. In the interview, they explained that whenever the LLM added appealing elements (see subsection 6.4), they would delete them in an attempt to stay closer to the source texts.

HTER was also used to measure the extent of changes made to the outputs. The closer the score is to 0, the fewer changes are made. Table 1 shows the overall HTER scores per participant and MT system, obtained using the MATEO platform (Vanroy et al., 2023).

	LLM (%)	NMT (%)
T1	15.8	34.5
T2	11.7	6.4
T3	32.3	52.8
T4	13	23.3
Average	18.2	29.25

Table 1: HTER scores per participant and MT system

Looking at how participants T1, T3, and T4 performed, it is clear that they made fewer edits to the LLM outputs to render them publishable. In contrast, T2 exhibited the opposite pattern, making more extensive edits to the LLM outputs. This behavior helps explain the translation style they adopted: again, when the LLM produced additional wording to enhance the appeal of the output, they omitted such additions and adhered closely to the source texts. Considering the average HTER score for the LLM, the result suggests that overall, post-editing LLM-generated translations of TPTs can be regarded as relatively efficient in the sense that only about one-fifth (18.2%) of the outputs required modification to reach a publishable standard. However, this should be treated with caution as individual variation is clearly notable; for instance, T3's higher score (32.3%) indicates a substantially greater number of edits.

6.3 Cognitive Effort

Cognitive effort in post-editing refers to "hidden cognitive processes such as reading, understanding, comparing source language meaning to that of the MT output, decision making, while taking into account the guidelines and expectations, and monitoring the text as it is revised" (O'Brien, 2022, p. 116). While eye-tracking could have provided more precise insights into the cognitive effort involved, it was not feasible given the remote nature of the study. Therefore, three other widely used measures were employed to approximate cognitive effort: average pause ratio (Lacruz et al., 2012), pause-to-word ratio (Lacruz & Shreve, 2014), and subjective cognitive effort ratings (Hart & Staveland, 1988). For space constraints, only pause-based metrics are reported in this section. The threshold for a pause to be considered meaningful was set at 300 ms (Lacruz et al., 2014).

The first metric, the average pause ratio, is calculated as the average time per pause in a segment divided by the average time per word in the same segment. As Lacruz and Shreve (2014) explain, higher cognitive effort is associated with a lower average pause ratio. Figure 3 shows the average scores per participant and MT system.

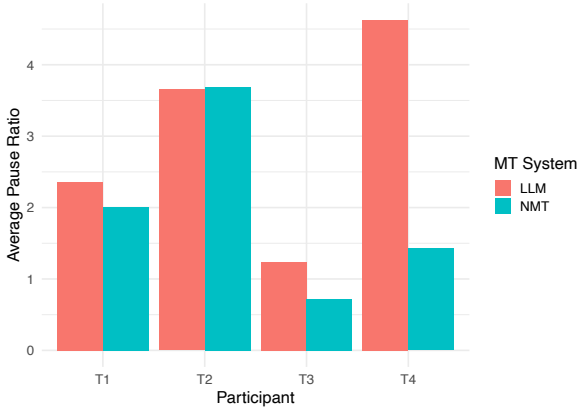

Figure 3: Average pause ratio per participant and MT system

Despite variability among translators, the LLM outputs were associated with higher average pause ratios for most of them, indicating lower cognitive effort in rendering those outputs publishable: T1 (LLM: $M = 2.35$, $SD = 2.32$; NMT: $M = 2.00$, $SD = 4.48$), T3 (LLM: $M = 1.23$, $SD = 1.19$; NMT: $M = 0.72$, $SD = 0.89$), and T4 (LLM: $M = 4.63$, $SD = 6.75$; NMT: $M = 1.43$, $SD = 1.46$). T2 (LLM: $M = 3.65$, $SD = 3.91$; NMT: $M = 3.69$, $SD = 2.96$) was an exception, with both the LLM and NMT outputs requiring comparable levels of cognitive effort.

The second metric, the pause-to-word ratio, is calculated by dividing the number of pauses in a given MT segment by the number of words in the same segment. The higher the pause-to-word ratio, the more cognitive effort is exerted by translators (Lacruz & Shreve, 2014). Figure 4 visualizes the average scores for every participant when working with both systems.

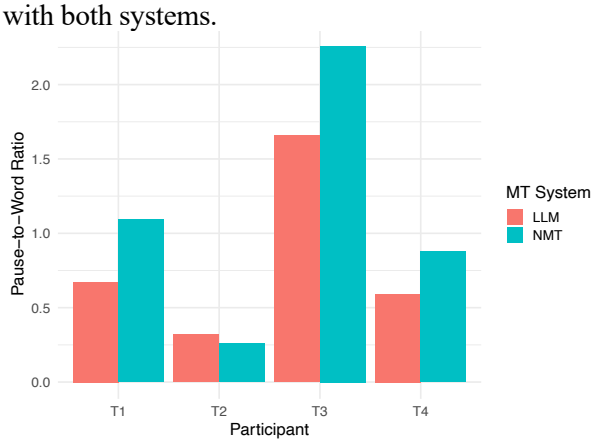

Figure 4: Pause-to-word ratio per participant and MT system

The results replicate those obtained from the first metric. LLM-generated outputs were cognitively less demanding to post-edit for most participants, as reflected in average pause-to-word ratios: T1 (LLM: $M = 0.67$, $SD = 0.57$; NMT: $M = 1.09$, $SD = 0.82$), T3 (LLM: $M = 1.66$, $SD = 1.99$; NMT: $M = 2.26$, $SD = 1.58$), and T4 (LLM: $M = 0.59$, $SD = 0.75$; NMT: $M = 0.88$, $SD = 0.80$). T2 (LLM: $M = 0.32$, $SD = 0.22$; NMT: $M = 0.26$, $SD = 0.23$) was the only exception, exhibiting a slightly opposite pattern.

It is important to note that pause-based metrics depend on physical interaction with the outputs, specifically the presence of keystrokes. In this study, a total of 33 data points with no keystrokes and thus no recorded pauses, were excluded from the cognitive effort analysis. This exclusion does not imply an absence of cognitive effort, but rather reflects a limitation of the metrics, which cannot capture effort without observable interaction.

6.4 Cue-based Retrospective Interviews

The previous subsections focused on analyzing the experimental data, providing an empirical basis for the effort translators exerted when post-editing and improving TPTs translated by an LLM relative to a conventional NMT system. This subsection builds on that by presenting and analyzing translators' subjective reflections and evaluations of the raw outputs they worked on, offering additional insights that complement the quantitative data from the experiment. The interviews were analyzed in Microsoft Word using thematic analysis following Braun and Clarke's (2006) approach. The analysis focused on the segment involving participants' direct reflections on and evaluations of the outputs, which comprised approximately 4,000 words.

The analysis adopted a hybrid thematic approach: deductive coding was informed by the interview guide, while inductive coding allowed for new themes to emerge from the data. To enhance credibility, an additional Arabic-speaking researcher with experience in thematic analysis reviewed a sample of the coded dataset to ensure consistency and alignment with participants' responses. Analysis of the participants' evaluations yielded 40 initial codes. Through iterative refinement, the most salient codes were synthesized into three primary themes (see Table 2), ensuring the analysis focused on the most prevalent patterns in the data. Due to space constraints, the main text includes only back-translated segments for ease of reading. The Arabic MT renderings are available in Appendix E, where

they are presented alongside their corresponding back-translations and source text segments.

Main Themes	Subthemes
LLMs manifest creative liberties in the translated outputs	Creativity as non-literal translation Creativity as aesthetic augmentation
LLM outputs are unpredictable and far from perfect	
NMT outputs are informative but lack promotional appeal	

Table 2: Main themes and subthemes

1. LLMs manifest creative liberties in the translated outputs

Participants perceived the LLM outputs to exhibit a degree of creative liberty which was manifested through 1) non-literal translations and 2) aesthetic augmentation. First, they frequently noted departures from literal renderings as creative acts that were highly appropriate in this context. For instance, the LLM rendered "*But scratch further beneath the surface, and you'll uncover a nation bursting with adventure*" as "*But natural beauty is only the beginning; Wales hides in its folds unforgettable adventures*" (back-translated from Arabic). Reflecting on this sentence, T1 noted: "Here it did good, it added some creativity, it did not go for word for word". In a similar vein, the LLM rendered "*Welcome to nature's playground*" as "*Welcome to the paradise of nature*" (back-translated from Arabic). T2 mentioned here: "It is not translated literally…but it is more attractive to say [paradise] than the literal translation of playground, it adds a sense of attractiveness". Another example is the translation of "*…not to mention some of Britain's premier hiking trails*" as "*…and we do not forget also some of the most wonderful walking paths in Britain*" (back-translated from Arabic); T2 noted that while translators might not typically choose this phrasing, the LLM's approach was, in their view, "creative and good". These observations on perceived creative acts align with established dimensions of creativity in translation studies, particularly novelty and acceptability (e.g., Guerberof-Arenas & Toral, 2022). Participants recognized creativity in renderings that deviated from the source texts, provided these deviations were also acceptable and appropriate; they generally appreciated such choices by the LLM for enhancing the attractiveness and communicative effectiveness of translated TPTs.

Expanding on this, creativity in the LLM outputs was also evident through aesthetic augmentation. This is understood here as the addition of linguistic descriptors or elements not present in the specific source segments being translated, which serve to enhance the aesthetic quality of the outputs and, in turn, their promotional appeal. In most cases, rather than viewing them as hallucinations, participants appreciated these additions and perceived them as creative. For instance, the source segment "*Regardless of what you want to do when you visit…*" was rendered by the LLM as "*Whether you are a fan of adventure, history, or relaxation in the midst of nature…*" (back-translated from Arabic). In reflecting on this, T1 noted that "the machine is very creative. It is not faithful to the original … I do not know how it will be if there is no creativity, but it is good. It is natural. It does offer all these [additions]… yes, we read that before, so it did not add things that are not real. I think it considered the paragraphs before. It took some ideas". This notion aligns with Navarro's (2025) analysis of LLM outputs in the tourism domain, where she observes that such "additions are vague enough to the point of bringing some color to the text without introducing any factual inaccuracies" (p. 124). Interestingly, even T2, whose post-editing behavior differed from the other participants as they often deleted LLM-generated additions, observed that these additions contributed a sense of "originality" to the outputs: it is "good to have this sense of originality, the feeling that this is not a translated text". Overall, these findings suggest that LLM outputs exercise creative liberty through aesthetic augmentation which can enhance the promotional appeal of translated TPTs.

2. LLM outputs are unpredictable and far from perfect

While the LLM demonstrated creative strengths, participants identified several shortcomings, most notably the literal renderings of some grammatical structures and phrases. In several instances, the LLM mirrored the source text's syntax, a choice T3 perceived negatively. For example, when translating the sentence "*Bath is the city that has it all*", the LLM retained the English-style nominal structure by starting with the subject "Bath". T3 criticized this approach, noting that in Arabic, it is preferred and more natural to start sentences with verbs rather than nouns. In addition, the literal rendering of some phrases by the LLM was

perceived as creating a communicative mismatch. Take the phrase "*rugged coastline*" as an example which was used for promoting tourism in Wales. T1 observed that the word "rugged" carries a negative connotation in this context: "If someone says to me go to the rugged coastline, why would I go to something rugged? It is not a pleasant experience. It might be pleasant, but we can use more appealing terms"; this reasoning explains their decision to change this phrase to "*coasts surrounded by rocks*" (back-translated from Arabic) in an attempt to create a more evocative and alluring destination image. More critically, the LLM's literal rendering of some phrases extended to references that are culturally sensitive. A clear example emerged in a text promoting Wales as a filming location for popular series such as Sex Education. T3 was quick to point out that the output failed to anticipate how this specific title might be received in an Arabic-speaking context. They explained their decision to replace it with a different show after searching online for alternative series filmed in Wales, noting: "It [the LLM] does not understand the Arabic culture… the name itself is not appropriate… to be safe, I looked up another show". For T3, this was not a minor error but rather clear evidence of the system's difficulty in handling deep-rooted cultural nuances that human translators navigate instinctively. It necessitated intervention on their part, driven by their awareness that such references can be counterproductive to the purpose of attracting this particular group of tourists. Overall, these examples illustrate that while LLMs can produce creative outputs as seen in the first theme, their translations remain unpredictable and prone to a range of issues, highlighting that they are far from perfect.

3. NMT outputs are informative but lack promotional appeal

While NMT outputs were consistently described as informative, participants noted a significant deficit in the promotional appeal required for TPTs. For instance, T2 observed that the NMT output was "reflective of the source text", "informative", with a structure that is both "good" and "clear" yet lacking creativity. Similarly, T4 noted that it was "approachable" and "easy to understand for regular people" but then added "I do not think it is creative". While NMT outputs may fulfill the informative function of TPTs, they may fall short with respect to their promotional function, which requires language capable of exciting and persuading prospective travelers. This limitation is illustrated by NMT's verbatim rendering of "*Northern Ireland has everything from …*" into Arabic, which T3 criticized on the grounds that "these generic expressions do not sell. We do not say 'everything'. What is 'everything'?".[3] The literal rendering of some phrases by NMT can also create a communicative mismatch. Take the phrase "*adrenaline addicts*" as an example which was rendered as-is into Arabic. Rather than evoking an appealing image for adventure tourists, T3 noted how this rendering "gives a very negative connotation" in Arabic, with T4 describing it as "unacceptable".[4] T1 also noted on NMT outputs that while the meaning remains clear—stating "I would understand what it means"—the mechanical nature of the translation is evident: "I would say definitely it is a machine translation, and that maybe will affect the attractiveness of the text itself". Taken together, these observations show that NMT outputs, while informative, do not fulfill the promotional demands of TPTs.

7 General Discussion and Conclusion

This small-scale pilot study examined the potential of an LLM in translating info-promotional tourism texts, relative to a traditional NMT system, using a mixed-methods design that combined a post-editing experiment with semi-structured cue-based retrospective interviews.

Overall, the post-editing results revealed that the majority of participants exerted less effort in rendering the LLM outputs publishable compared to those produced by a conventional NMT system. This likely stems from the broader contextual capabilities of LLMs and/or the specific prompt used. The interview data also align with these results: the LLM outputs were generally perceived as more creative and attractive, and that creativity was manifested through non-literal translations and aesthetic augmentation. By starting with a stronger initial draft, the LLM likely reduced the effort on the translators, allowing them to focus on refining

[3] For the purpose of comparison, this was rendered by the LLM as "*Northern Ireland brings together the splendor of nature and the charm of history, …*" (back-translated from Arabic). The LLM enhanced the sentence through aesthetic augmentation, incorporating contextual descriptors drawn from the rest of the sentence.

[4] For the purpose of comparison, this phrase was rendered by the LLM as "*thrill lovers*" (back-translated from Arabic), which can be considered more appealing.

the output rather than reconstructing it. In contrast, NMT was found to be less effective: it required greater post-editing effort and tended to produce translations that were perceived to be more informative than promotional. This is in line with Giampieri and Harper (2022) who found that while NMT performs relatively well with informative tourism texts, it struggles to meet the demands of those that are of a promotional nature which require more appellative and euphoric language.

The results also indicate that LLMs, even when prompted with information about who the target readers are, remain insufficiently sensitive to cross-cultural differences and may fail to handle deep-rooted cultural nuances in their translations. This finding aligns with Guo et al. (2025) who argue that despite the progress made so far, "LLMs are still some distance away from reaching a truly nuanced cross-cultural understanding" (p. 1). Given that translating TPTs is inherently an intercultural activity (Sulaiman & Wilson, 2018), the limitations of LLMs in this regard underscore the continued need for professional translators who make culturally informed decisions that ensure the final translations resonate with the target readers and successfully fulfill their goal of attracting potential tourists. Awareness of such limitations, along with the others mentioned earlier, can also help reduce the hype around LLMs and foster more realistic expectations. It can also alleviate concerns and potentially challenge the widely held claim that AI is signalling the end of human translation (Pym, 2023), including post-editing.

In this regard, some industry reports claim that translating promotional content like TPTs can now be automated by LLMs but with "human oversight" (Smartling, 2024, p. 9). However, the use of the term "oversight" is quite troubling as it may understate the level of professional agency and expertise required in post-editing and improving LLM outputs. Oversight may imply a passive monitoring role that could, in principle, be performed by anyone. Yet, the findings of this study suggest instead that LLM outputs necessitate active expert intervention, involving the identification and resolution of subtle linguistic, cultural, and pragmatic deficiencies that a non-expert reviewer might easily overlook. In this sense, translators function as intercultural mediators engaged in "an active process that requires sensitivity, awareness, and adaptability" (Maci & Spinzi, 2025, p. 41).

Taken together, the preliminary findings from this pilot study show that, overall, LLMs can serve as useful tools for translating info-promotional content like TPTs; with the support of professional translators in improving and enhancing the outputs, LLMs may offer a potential opportunity for tourism boards that have limited budgets (e.g., Napu, 2016) or wish to take advantage of the technology.

8 Limitations and Future Work

This pilot study has a number of limitations. The small number of participants means that the results may be strongly influenced by individual variability and therefore may not be generalizable. Moreover, the results reported here are based on a single relatively well-resourced language pair. Given that LLMs are resource-intensive, these findings may not be replicable for low-resourced language pairs. Importantly, LLM outputs can be inconsistent across different prompts or iterations. This means that experiments conducted with LLMs may not be fully replicable, and any conclusions drawn from a specific set of outputs remain tentative as future iterations could produce outcomes that either support or contradict the initial findings (Lee, 2023). However, as Lee notes, "this need not discourage us from making preliminary observations based on the technology at the time of writing. Of interest to us is what [LLMs] can potentially do as compared to dedicated translation programs and also human translators, irrespective of the iterations" (p. 6).

As an initial step in exploring the potential of LLMs for translating TPTs, this study focused solely on professional translators. Future research could broaden this scope of work by including other key stakeholders such as end readers (i.e., potential tourists). Research suggests that potential tourists are unable to distinguish between TPTs produced by tourism marketers and those generated by LLMs, and that LLM-generated TPTs are perceived as equally fluent and appealing as human-produced ones (Zhang & Prebensen, 2024). It would therefore be worthwhile to examine whether similar perceptions apply to TPTs fully translated by LLMs, to those produced through hybrid workflows involving human post-editing, and to TPTs translated entirely by professional translators. One possible way to approach this is through models from advertising and marketing research, such as the AIEDA model (Attention, Interest, Evaluation, Desire, and Action; Weng et al., 2021), following a line of inquiry similar to that adopted by Carvalho et al. (2025).

References

Al-Fahad, Saleem. 2012. Stylistic analysis of Arabic and English translated tourist brochures: A contrastive study. *Diyala Journal for Human Research*, 56(1): 554–578.

Aziz, Wilker, Sheila Castilho, and Lucia Specia. 2012. PET: A tool for post-editing and assessing machine translation. *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, pages 3982–3987, European Language Resources Association.

Braun, Virginia, and Clarke Victoria. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101.

Brenner, Judith. 2024. The MTxGames project: Creative video games and machine translation – different post-editing methods in the translation process. *Proceedings of the 25th Annual Conference of the European Association for Machine Translation*, 47–48, European Association for Machine Translation.

Carvalho, Inês, Sandra Loureiro, Stanislav Ivanov, Peter Björk, and Faruk Seyitoğlu. 2025. Beyond human touch: Evaluating the effectiveness of AI, human, and hybrid-generated tourism promotional texts. *Journal of Hospitality and Tourism Insights*, 8(10): 3804–3824.

Dann, Graham. 1996. *The Language of Tourism: A Sociolinguistic Perspective*. CAB International.

Daems, Joke. 2022. Dutch literary translators' use and perceived usefulness of technology: The role of awareness and attitude. In Hadley, James, Kristiina Taivalkoski-Shilov, Carlos Teixeira, and Antonio Toral (Eds.), *Using Technologies for Creative-text Translation* (pp. 40–65). Routledge.

Ehrensberger-Dow, Maueen. 2014. Challenges of translation process research at the workplace. *MonTI,* 355–383.

Giampieri, Patrizia, and Martin Harper. 2022. Tourism translation: From corpus to machine translation (and back). *Umanistica Digitale*, (14): 119–135.

Guerberof-Arenas, Ana, and Antonio Toral. 2022. Creativity in translation: Machine translation as a constraint for literary texts. *Translation Spaces*, 11(2): 184–212.

Guerberof-Arenas, Ana, Joss Moorkens, and David Orrego-Carmona. 2024. "A Spanish version of EastEnders": A reception study of a telenovela subtitled using MT. *The Journal of Specialised Translation*, 41: 230–254

Guo, Shiwei, Sihang Jiang, Qianxi He, Yanghua Xiao, Jiaqing Liang, Bi Yude, Minggui He, Shimin Tao, and Li Zhang. 2025. Do large language models truly understand cross-cultural differences? arXiv:2512.07075

Hart, Sandra, and Lowell Staveland. 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In Hancock, Peter, and Najmedin Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). Elsevier Science & Technology.

He, Sui. 2024. Prompting ChatGPT for translation: A comparative analysis of translation brief and persona prompts. *Proceedings of the 25th Annual Conference of the European Association for Machine Translation*, pages 314–324, European Association for Machine Translation.

Janodet, Francisco, and Manuela Jurado. 2025. Landscape of tourism translation: Trends and future challenges. MonTI, 8: 40–71.

Kairos Future. 2025. Artificial intelligence (AI) in tourism: Assessing and supporting NTO's research & marketing operations. European Travel Commission. https://etc-corporate.org/reports/artificial-intelligence-ai-in-tourism-assessing-and-supporting-ntos-research-marketing-operations/

Katan, David. 2011. Occupation or profession: A survey of the translators' world. In Sela-Sheffy, Rakefet, and Miriam Shlesinger (Eds.), *Identity and Status in the Translational Professions* (pp. 65–88). John Benjamins Publishing Company.

Kelly, Dorothy. 1997. The translation of texts from the tourist sector: Textual conventions, cultural distance and other constraints. *Trans: Revista de Traductología*, 2: 33–42.

Krings, Hans. 2001. *Repairing Texts: Empirical Investigations of Machine Translation Post-editing Processes*. Kent State University Press.

Lacruz, Isabel, and Gregory Shreve. 2014. Pauses and cognitive effort in post-editing. In O'Brien, Sharon, Laura Balling, Michael Carl, Michel Simard, and Lucia Specia (Eds.), *Post-editing of Machine Translation: Processes and Applications* (pp. 246–272). Cambridge Scholars Publishing.

Lacruz, Isabel, Gregory Shreven, and Erik Angelone. 2012. Average pause ratio as an indicator of cognitive effort in post-editing: A case study. *Proceedings of the AMTA 2012 Workshop on Post-Editing Technology and Practice*, pages 29–38, Association for Machine Translation in the Americas.

Lacruz, Isabel, Michael Denkowski, and Alon Lavie. 2014. Cognitive demand and cognitive effort in post-editing. *Proceedings of the 11th Conference of the Association for Machine Translation in the*

Americas, pages 73–84, Association for Machine Translation in the Americas.

Lee, Tong. 2024. Artificial intelligence and posthumanist translation: ChatGPT versus the translator. *Applied Linguistics Review*, 15(6): 2351–2372.

Li, Menglu, and Dechao Li. 2025. Human expertise vs AI efficiency: A comparative analysis of student and ChatGPT post-editing. In Sun, Sanjun, Kanglong Liu, and Riccardo Moratto (Eds.), *Translation Studies in the Age of Artificial Intelligence* (pp. 151–171). Routledge.

Lyu, Chenyang, Zefeng Du, Jitao Xu, Yitao Duan, Minghao Wu, Teresa Lynn, Alham Aji, Derek Wong, and Longyue Wang. 2024. A paradigm shift: The future of machine translation lies with large language models. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation,* pages 1339–1352, ELRA and ICCL.

Maci, Stefania, and Cinzia Spinzi. 2025. *Translating Tourism*. Routledge.

Malamatidou, Sofia. 2024. *Translating Tourism: Cross-linguistic Differences of Alternative Worldviews*. Palgrave Macmillan.

Moorkens, Joss, Antonio Toral, Sheila Castilho, and Andy Way. 2018. Translators' perceptions of literary post-editing using statistical and neural machine translation. *Translation Spaces*, 7(2): 240–262.

Muñoz, Isabel. 2012. Analysing common mistakes in translations of tourist texts (Spanish, English and German). *Onomázein*, 26(2): 335–349.

Napu, Novriyanto. 2016. *Translation in Tourism: Understanding the Quality of Translation Across Multiple Perspectives* [Doctoral dissertation]. University of South Australia.

Navarro, Sandra. 2025. Transcreation in the tourism domain: Is artificial intelligence up to the task? In Walter, Katharina, and Marco Agnetta (Eds.), *Applying Artificial Intelligence in Translation: Possibilities, Processes and Phenomena* (pp. 118–131). Routledge.

O'Brien, Sharon. 2022. How to deal with errors in machine translation: Post-editing. In Kenny, Dorothy (Ed.), *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence* (pp. 105–120). Language Science Press.

OpenAI. 2024. Hello GPT-4o. https://openai.com/index/hello-gpt-4o/

Pym, Anthony. 2023. *Exploring Translation Theories*. Routledge.

R Core Team. 2026. R: A language and environment for statistical computing.

Reiss, Katharina. 1981. Type, kind and individuality of text: Decision making in translation. *Poetics Today*, 2(4): 121–131.

Saldanha, Gabriela, and Sharon O'Brien. 2014. *Research Methodologies in Translation Studies*. Routledge.

Sanning, He. 2010. Lost and found in translating tourist texts: Domesticating, foreignising or neutralising approach. *The Journal of Specialised Translation*, 13: 124–137.

Sarti, Gabriele, Arianna Bisazza, Ana Guerberof Arenas, and Antonio Toral. 2022. DivEMT: Neural machine translation post-editing effort across typologically diverse languages. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 7795–7816, Association for Computational Linguistics.

Smartling. 2024. Using large language models for translation. Smartling, Inc. https://go.smartling.com/hubfs/3Q240613_Using_Large_Language_Models_for_Translation.pdf

Snell-Hornby, Mary. 1999. The "ultimate confort": Word, text and the translation of tourist brochures. In Anderman, Gunilla, and Margaret Rogers (Eds.), *Word, Text, Translation: Liber Amicorum for Peter Newmark* (pp. 95–103). Multilingual Matters.

Snover, Matthew, Bonnie Dorr, Rich Schwartz, Linnea Micciulla, and John Makhoul. 2006. A study of translation edit rate with targeted human annotation. *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas*, pages 223–231, Association for Machine Translation in the Americas.

Spinzi, Cinzia. 2025. Translating the intangible: The role of artificial intelligence in the translation of metaphors in tourism discourse. *International Journal of Language Studies*, 19(2): 161–177.

Sulaiman, Z. Mohamed. 2016. The misunderstood concept of translation in tourism promotion. The *International Journal of Translation and Interpreting Research*, 8(1): 53–68.

Sulaiman, Z. Mohamed, and Rita Wilson. 2018. Translating tourism promotional materials: A cultural-conceptual model. *Perspectives*, 26(5): 629–645.

Sulaiman, Z. Mohamed, and Rita Wilson. 2019. *Translation and Tourism: Strategies for Effective Cross-cultural Promotion.* Springer.

Tiselius, Elisabet, John Schwieter, Igor Lourenço da Silva, and Gary Massey. 2025. Cued retrospection. In Rojo López, Ana, and Ricardo Muñoz Martín

(Eds.), *Research Methods in Cognitive Translation and Interpreting Studies* (pp. 92–107). John Benjamins Publishing Company.

Toral, Antonio, Martijn Wieling and Andy Way. 2018. Post-editing effort of a novel with statistical and neural machine translation. *Frontiers in Digital Humanities*, 5.

Vanroy, Bram, Arda Tezcan, and Lieve Macken. 2023. MATEO: Machine translation evaluation online. *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, pages 499–500, European Association for Machine Translation.

Yamada, Masaru. 2023. Optimizing machine translation through prompt engineering: An investigation into ChatGPT's customizability. *Proceedings of Machine Translation Summit XIX: Users Track*, pages 195–204, Asia-Pacific Association for Machine Translation.

Weng, Lisheng, Zhuowei Huang and Jigang Bao. 2021. A model of tourism advertising effects. *Tourism Management,* 85, 104278.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. dplyr: A grammar of data manipulation. https://cran.r-project.org/web/packages/dplyr/index.html

Zhang, Yaozhi, and Nina Prebensen. 2024. Co-creating with ChatGPT for tourism marketing materials. *Annals of Tourism Research Empirical Insights*, 5(1).

Appendix A. Sample Task.

Appendix B. LLM Full Prompt.

You are a professional translator working in the tourism marketing domain. Your task is to adapt this English tourism promotional text into Arabic. Adhere to the linguistic norms of Arabic and the expectations and needs of Arabic-speaking tourists. The purpose of the translation is to encourage readers to visit the destination.

Appendix C. Full Post-editing Tasks in Completion Order.

	T1	T2	T3	T4
1	S9 by LLMs	S14 by LLMs	S7 by LLMs	S9 by NMT
2	S6 by NMT	S7 by NMT	S14 by NMT	S10 by NMT
3	S5 by NMT	S11 by LLMs	S9 by NMT	S7 by LLMs
4	S14 by LLMs	S9 by LLMs	S11 by NMT	S11 by NMT
5	S7 by NMT	S12 by LLMs	S5 by LLMs	S1 by LLMs
6	S1 by NMT	S1 by NMT	S13 by NMT	S8 by NMT
7	S2 by NMT	S4 by NMT	S6 by LLMs	S6 by LLMs
8	S10 by LLMs	S3 by NMT	S1 by LLMs	S3 by LLMs
9	S11 by LLMs	S10 by LLMs	S10 by NMT	S5 by LLMs
10	S8 by LLMs	S5 by NMT	S2 by LLMs	S14 by NMT
11	S4 by NMT	S13 by LLMs	S3 by LLMs	S4 by LLMs
12	S3 by NMT	S2 by NMT	S12 by NMT	S2 by LLMs
13	S12 by LLMs	S8 by LLMs	S8 by NMT	S13 by NMT
14	S13 by LLMs	S6 by NMT	S4 by LLMs	S12 by NMT

*S stands for source text.

Appendix D. Questions from the Cue-based Retrospection Segment of the Interview.

1. What do you think of the quality of this translation given the nature of the text type used?
2. Was there anything you liked or did not like about this translation?
3. How would you describe the attractiveness of this translation?
4. In your opinion, how were cultural references handled in this translation?
5. Why did you change this? Or alternatively, why did you leave it unchanged?[5]

[5] This question was asked when certain edits, or lack thereof, stood out to me while observing the participants' post-editing behavior. For instance, if a participant omitted a cultural reference, changed a cultural theme, or chose to leave elements that may not be compatible with the target readers untouched, I asked them to explain that decision. Rather than commenting on every edit, which can be very time-consuming and cognitively demanding, I focused only on

Appendix E. Arabic MT Renderings.

English Source	Arabic MT	Back-translation
But scratch further beneath the surface, and you'll uncover a nation bursting with adventure.	LLM: لكن الجمال الطبيعي ليس سوى البداية؛ فويلز تخفي في طياتها مغامرات لا تُنسى.	But natural beauty is only the beginning; Wales hides in its folds unforgettable adventures.
Welcome to nature's playground	LLM: مرحبًا بكم في جنة الطبيعة	Welcome to the paradise of nature
…not to mention some of Britain's premier hiking trails	LLM: ولا ننسى أيضًا بعضًا من أروع مسارات المشي في بريطانيا	… and we do not forget also some of the most wonderful walking paths in Britain
Regardless of what you want to do when you visit…	LLM: سواء كنت من عشاق المغامرة، أو التاريخ، أو الاسترخاء وسط الطبيعة...	Whether you are a fan of adventure, history, or relaxation in the midst of nature…
Bath is the city that has it all.	LLM: مدينة باث تجمع كل ما يحلم به المسافر.	Bath city has everything a traveler could dream of.
Northern Ireland has everything from…	NMT: تضم أيرلندا الشمالية كل شيء… LLM: تجمع أيرلندا الشمالية بين روعة الطبيعة وسحر التاريخ…	Northern Ireland has everything from… Northern Ireland brings together the splendor of nature and the charm of history…
Adrenaline addicts	NMT: مدمني الأدرينالين LLM: عشاق الإثارة	Adrenaline addicts Thrill lovers

those that seemed particularly interesting or unexpected from my perspective as a researcher.