

Comparing DeepL and ChatGPT in Translating Chinese Tourist Public Signs: Accuracy and Pragmatic Appropriateness

Yang Liu

College of Foreign Studies, Guilin University of Technology, Guilin, Guangxi 541006, China
874013500@qq.com

DOI: <https://doi.org/10.36892/ijlls.v8i3.2640>

APA Citation: Liu, Y. (2026). Comparing DeepL and ChatGPT in Translating Chinese Tourist Public Signs: Accuracy and Pragmatic Appropriateness. *International Journal of Language and Literary Studies*, 8(3), 430-442. <https://doi.org/10.36892/ijlls.v8i3.2640>

Received:
20/04/2026

Accepted:
18/05/2026

Keywords:
AI-assisted translation;
tourist public signs; DeepL;
ChatGPT;
accuracy;
pragmatic appropriateness

Abstract

AI-assisted translation is becoming increasingly visible in tourism communication, yet the translation of Chinese tourist public signs into English remains underexplored. Unlike longer tourism texts, public signs are condensed, function-sensitive, and pragmatically constrained. Many Chinese tourist signs also contain figurative, descriptive, or rapport-oriented wording, which complicates translation. Against this background, this study compares DeepL and ChatGPT in the translation of 25 Chinese tourist public signs into English, focusing on accuracy and pragmatic appropriateness. ChatGPT is examined under two conditions—a minimal-prompt condition and a domain-sensitive prompt condition—whereas DeepL is tested through direct source-text input. Using an exploratory qualitative comparative design, the study identifies recurrent tendencies across formulaic, warning, environmental, and rhetorically marked signs. The analysis suggests that DeepL is relatively stable on conventional signs, whereas domain-sensitive ChatGPT more often produces concise and context-appropriate English signage, albeit sometimes at the cost of reduced rhetorical texture or weakened rapport-related nuance.

1. Introduction

AI-assisted translation has become increasingly visible in tourism communication. Recent research on tourism translation has shown that tourism discourse cannot be evaluated solely in terms of sentence-level correctness, because tourism texts often involve not only information transfer but also rhetorical appeal, cultural mediation, and audience orientation. Chen and Lin (2025) are particularly relevant in this regard. Their comparison of ChatGPT, Google Translate, and DeepL evaluates tourism translation across multiple dimensions and shows that prompt design can influence ChatGPT's performance in domain-specific tasks. More broadly, Hendy et al. (2023) demonstrate that GPT-based systems can perform strongly on machine-translation tasks in some settings, but that their strengths are uneven and should not be conflated with stable, domain-independent superiority over conventional neural machine translation.

Tourist public signs, however, differ in important ways from longer tourism texts such as brochures, websites, destination descriptions, and hotel webpages. They are short, spatially constrained, and function-sensitive. Their communicative effectiveness depends not only on whether the source meaning is conveyed, but also on whether the English version works as an actual public sign in the target environment. Ko (2010) argues that Chinese public signs in tourist areas constitute a special genre because they frequently use descriptive or figurative language to communicate relatively simple messages. This feature creates a distinctive translational difficulty: the translator must decide how much of the source sign's rhetorical surface to retain and how much to normalize into conventional English signage.

The issue becomes even more complex when the source signs are directive or prohibitive. Sun and Chen (2023) show that Chinese directive public signs are not always terse or purely command-oriented. Some employ mitigation, explanation, courtesy markers, or personalized wording in order to manage rapport with the public. Their analysis suggests that such signs may attend to face, sociality rights and obligations, interests, and interactional goals rather than functioning only as behavioral instructions. A related cross-cultural point is raised by Lees (2024), who shows that the translation of politeness in directive public notices is shaped by language-specific conventions of directness, mitigation, and reader positioning. Taken together, these studies suggest that the translation of tourist public signs cannot be reduced to propositional transfer alone.

Broader research on Chinese-English public-sign translation has made considerable progress, but it still leaves clear gaps. Amenador and Wang (2022) show that although public-sign research has addressed definitions, functions, linguistic features, translation problems, principles, and theories, it still suffers from repetition, limited scope, and insufficient diversification. Scenic-spot studies by Gao (2020), Hu et al. (2021), Yao et al. (2021), and Zhan (2021) also confirm that sign translation problems remain highly recurrent in tourist destinations. Yet most of these studies focus either on error correction in human-produced signage or on general translation principles, rather than on AI-assisted translation under explicitly controlled output conditions.

Against this background, the present study examines a narrow yet highly practical question: how do DeepL and ChatGPT perform when translating Chinese tourist public signs into English? More specifically, it compares them across two dimensions: accuracy and pragmatic appropriateness. Accuracy refers to whether the intended meaning and sign function are correctly transferred. Pragmatic appropriateness refers to whether the English output is suitably worded, toned, and conventionalized for public tourist settings. These dimensions are related but not identical, and the present article argues that their divergence becomes especially visible in tourist public signs.

The study makes two contributions. Empirically, it compares DeepL with ChatGPT under two prompting conditions in the translation of Chinese tourist public signs, a genre that has received much less attention than longer tourism texts in current AI translation research. Conceptually, it argues that tourist public signs should be evaluated through a dual framework of accuracy and pragmatic appropriateness, because the two dimensions may diverge sharply when AI systems translate softened, poeticized, or rapport-oriented Chinese signage into English.

The article addresses the following questions: (1) How do DeepL, minimally prompted ChatGPT, and domain-sensitive ChatGPT differ in handling the semantic content of Chinese tourist public signs? (2) How do these outputs differ in pragmatic appropriateness when the source signs are directive, prohibitive, warning-oriented, or rhetorically marked? (3) What recurrent translational tendencies can be identified across the three output conditions, and what do they reveal about genre-specific AI translation evaluation?

2. Literature Review

Tourist public signs are not simply ordinary public notices located in tourist spaces. Ko (2010) shows that signs in tourist areas often adopt descriptive, figurative, and affectively softened language even when the underlying message is simple. The genre therefore combines brevity with rhetorical marking. This makes translation especially difficult because the translator must balance informational clarity, stylistic texture, and target-side sign convention. Ko identifies three main translational tendencies in tourist public-sign translation: literal translation, semi-literal or semi-adaptive translation, and free adaptation. Importantly, he argues that none of these strategies is problem-free. Literal translation often produces unnatural or unclear English, whereas free adaptation may flatten stylistic or interpersonal aspects of the

source sign. This observation is directly relevant to the present study, since the three AI output conditions examined here exhibit broadly comparable tendencies. Broader shifts in technology-mediated language learning, including the integration of digital tools for developing pragmatic competence, have been systematically reviewed by Alrefaee and Al-Shammari (2026), who trace the evolution from text-based computer-mediated communication to AI-driven immersive environments across two decades of research.

Chinese-English public-sign translation has developed rapidly since the mid-2000s. Amenador and Wang (2022) show that the field has addressed sign definitions, functions, features, translation problems, principles, strategies, and theories. At the same time, they also highlight recurring weaknesses, including repeated research content, limited investigation scope, and insufficient methodological diversification. This suggests that public-sign translation has become a recognized research area, but many subdomains remain theoretically and methodologically underdeveloped. Scenic-zone case studies reinforce this point. Yao et al. (2021), for example, classify public signs in the Wuxi Qingming Bridge Ancient Grand Canal Scenic Zone by type and analyze errors in terms of coherence rule, fidelity rule, and skopos rule, while proposing borrowing, annotation, and omission as corrective strategies. Other scenic-spot studies likewise identify spelling errors, mistranslation, inconsistency, omission, and Chinglish as common problems (Gao, 2020; Hu et al., 2021; Zhan, 2021). Although such studies are limited in scope, they reinforce an important methodological point: scenic-zone public-sign research often relies on functional classification and representative error analysis rather than on large random corpora alone.

A major limitation of earlier public-sign research is that many studies focus primarily on lexical or grammatical correctness while paying less attention to pragmatic function. Sun and Chen (2023) address this gap by examining Chinese directive public signs through the lens of rapport management. Their study shows that some signs do more than regulate behavior: they also mitigate force, justify restrictions, or express interpersonal care. Such language may attend to face, rights, interests, and interactional goals rather than functioning merely as terse behavioral instruction. Lees (2024) complements this view by demonstrating that politeness in directive notices does not travel straightforwardly across languages. Public notices may become either too blunt or too vague when translated without careful attention to target-side conventions. These insights are central to the present study, because they show why a translation may be semantically accurate yet still pragmatically weak as an English sign.

The present study is also informed by functionalist translation theory. Nord's (2018) account of translation as purposeful activity is especially useful because it foregrounds skopos, translation brief, intratextual coherence, intertextual coherence, adequacy, and equivalence. These concepts are highly relevant to tourist public signs, where the translation must function effectively in a target communicative setting rather than simply mirror source-text form. In a function-sensitive genre such as public signage, target-side adequacy cannot be reduced to formal correspondence alone. Similarly, House (2015) argues that translation quality assessment requires more than an intuitive judgement of whether a translation "sounds right." It requires contextualized comparison, analysis of function, and a distinction between description and evaluation. These ideas are particularly useful when the object of analysis is a short public text with a strong situational function.

A related point emerges from tourism translation quality research. Pierini (2007) argues that translation quality in tourism discourse cannot be restricted to linguistic correctness, but must also account for pragmatic effect, cultural transfer, target-text acceptability, and genre-specific norms. Although tourism websites differ from public signs in length and medium, the underlying principle remains relevant: tourism translation succeeds only when it works effectively for its intended users in context. Chen and Lin (2025) update this concern in the context of AI translation. Their findings suggest that machine-translated tourism texts should be assessed through multiple dimensions rather than through fluency alone. This broader

perspective supports the present study's decision to distinguish between accuracy and pragmatic appropriateness rather than treating translation quality as a single, undifferentiated construct.

Recent research on LLM-based translation further supports the design of the present study. Peng et al. (2023) show that ChatGPT's translation performance can be influenced by task-specific and domain-specific prompting, while He (2024) argues that prompt design in translation should be treated cautiously and examined empirically rather than assumed to be automatically beneficial. These studies suggest that prompt formulation is not a neutral wrapper around the model; it changes how the model construes the translation task. This is particularly important in the case of tourist public signs. Because the genre is brief, function-bound, and rhetorically uneven, prompt-induced changes in translational priorities are likely to be especially visible. The present study therefore treats prompt design as an analytical variable rather than as an invisible background condition.

3. Materials and Methods

3.1 Research Design

The present study is designed as an exploratory qualitative comparative study rather than a rating-based evaluation. Its purpose is not to establish statistically generalizable rankings across AI systems, but to identify recurrent translational tendencies in a genre-sensitive mini-corpus of Chinese tourist public signs. The analysis therefore focuses on close comparison of three output sets—DeepL, minimally prompted ChatGPT, and domain-sensitive ChatGPT—and examines how they differ in semantic transfer, sign convention, and pragmatic effectiveness. This design is appropriate because the object of study is a short, function-sensitive, and rhetorically marked text type whose translational problems are often most clearly revealed through pattern-based qualitative analysis.

3.2 Corpus Construction and Data Sources

The study uses a purposefully selected mini-corpus of 25 Chinese tourist public signs. The corpus is not intended to represent the full quantitative distribution of signage in Chinese tourist destinations. Instead, it was constructed to include functionally and rhetorically representative cases that are especially revealing for AI-assisted translation comparison. Most items were selected from representative tourist-sign examples discussed in prior scholarship on Chinese-English public-sign translation and Chinese directive public signs, especially Ko's (2010) discussion of tourist public signs as a special genre and Sun and Chen's (2023) analysis of Chinese directive signage as rapport-oriented public discourse. This selection logic is also consistent with scenic-zone case studies that classify public signs by function and recurrent translation problems rather than relying exclusively on large-scale random sampling (Yao et al., 2021).

3.3 Functional Classification of the Corpus

Because tourist public signs often perform more than one communicative function at once, the 25 items were classified according to their dominant function rather than treated as strictly exclusive categories. The corpus covers four analytically central groupings: formulaic directive or restrictive signs, warning or caution signs, environmental-protection signs, and rhetorically marked or personified signs. This classification reflects earlier public-sign research, which emphasizes that sign function, rhetorical form, and target-reader effect are central analytical variables in tourist signage (Amenador & Wang, 2022; Ko, 2010). For ease of reading, Table 1 orders the corpus by dominant function, moving from more conventional sign types to more rhetorically marked ones. In this study, a sign is treated as rhetorically marked if it contains overt personification, poeticized expression, softened directive force, affective address, or other non-neutral rhetorical features that go beyond straightforward informational or regulatory wording.

3.4 Operationalizing Rapport in Tourist Public Signs

In this study, rapport refers to the interpersonal work performed by a public sign when it seeks not only to regulate behavior but also to maintain a socially acceptable relationship with the reader. In Chinese tourist public signs, rapport is operationalized through observable linguistic features such as personification, mitigation, affective address, beneficiary-oriented justification, care-oriented wording, and poeticized appeals to shared responsibility. These features are treated as rapport markers because they soften directive force, invite voluntary compliance, or frame the requested behavior as part of a considerate relationship between the sign-maker, the environment, and the visitor. Rapport preservation, however, is not equally important across all sign types. In warning or safety-critical signs, clarity, brevity, and directive force may legitimately take priority over interpersonal nuance. By contrast, in environmental-protection, civility-oriented, and visitor-care signs, rapport markers are often central to how the sign encourages cooperation without relying on blunt prohibition.

3.5 Prompting Procedure and Model Information

Three output conditions were analyzed. DeepL was used through direct source-text input via the DeepL Translator mobile app. ChatGPT was tested under two prompt conditions. The minimal-prompt condition, labelled ChatGPT-CP1, used the following prompt: “Please provide the English translation for the following text.” The domain-sensitive condition, labelled ChatGPT-CP2, used the following full prompt: “Translate the following Chinese tourist public sign into concise, natural, and context-appropriate English for international visitors. Preserve the intended function of the sign and avoid word-for-word translation.” The ChatGPT outputs were generated through the ChatGPT interface using GPT-5.5 Thinking in May 2026. The DeepL outputs were generated through the DeepL Translator mobile app in May 2026 by directly entering each Chinese source sign. No prompt, glossary, terminology setting, or post-editing was applied to the DeepL outputs. No post-editing was applied to the ChatGPT outputs before analysis.

The two ChatGPT conditions are analytically useful because they approximate two different translational orientations. The minimal prompt asks for translation in a generic sense and therefore tends to preserve more of the source-text surface. The domain-sensitive prompt frames the task as one of producing usable English tourist signage and therefore steers the model toward function-oriented normalization. The contrast does not establish a hierarchy in advance. Rather, it makes visible how prompt formulation changes the model’s translational priorities. In this respect, the present design is not merely comparative but also methodological: it tests how far prompt conditioning affects the negotiation between source-text rhetoric and target-side sign convention.

3.6 Analytical Procedure

The analysis is guided by two dimensions: accuracy and pragmatic appropriateness. Accuracy refers to whether the translation correctly conveys the source sign’s intended meaning and basic function. Pragmatic appropriateness refers to the degree to which the output works as English public signage in a tourist setting, including wording, register, directness, and contextual fit. This dual framework is theoretically justified by translation quality assessment, tourism translation, and public-sign pragmatics, all of which indicate that translation quality in public tourism communication cannot be reduced to sentence-level correctness alone (House, 2015; Pierini, 2007; Sun & Chen, 2023).

The comparison focuses on recurrent translational tendencies rather than isolated sentence-level judgments. In particular, the analysis tracks four patterns: formulaic-sign stability, source-surface retention, target-oriented normalization, and rapport loss under optimization. This procedure allows the study to move beyond simple tool ranking and toward a more genre-sensitive account of how AI systems negotiate semantic transfer, sign convention, and interpersonal function. The aim is therefore not to produce an abstract ranking of outputs, but

to describe how different systems respond to different communicative pressures within the same genre.

Table 1
Profile of the 25-Sign Mini-Corpus

No.	Chinese Sign	Category	Rhetorically Marked?	Source
1	禁止吸烟	Formulaic	No	Ko (2010)
2	非工作人员请勿入内	Formulaic	No	Ko (2010)
3	持半票者，出示证件	Formulaic	No	Ko (2010)
4	请您照看好自己的物品。	Formulaic	No	Sun & Chen (2023)
5	讲究秩序，注意安全	Formulaic	No	Ko (2010)
6	禁止设摊，禁止停车， 违者取缔	Formulaic	No	Ko (2010)
7	为了您和家人的健康， 请勿吸烟	Formulaic	Yes	Sun & Chen (2023)
8	上车莫匆匆，最紧要轻 松	Formulaic	Yes	Ko (2010)
9	小心台阶	Warning	No	Ko (2010)
10	请勿进入，钙化地带	Warning	No	Ko (2010)
11	人间仙境，不食人间“烟 火”	Warning	Yes	Ko (2010)
12	美景当前留英姿？请注 意安全！	Warning	Yes	Ko (2010)
13	水是生命的源泉，请节 约用水	Environmental	Yes	Sun & Chen (2023)
14	除了脚印什么都别留下 ，除了记忆什么都不要 带走。	Environmental	Yes	Ko (2010)
15	绕行三五步，留得芳草 绿	Environmental	Yes	Ko (2010)
16	小草有生命，足下多留“ 青”	Environmental	Yes	Ko (2010)
17	青青小草，不踩多好	Environmental	Yes	Ko (2010)
18	小草也在长，踏入想一 想	Environmental	Yes	Ko (2010)
19	小草微微笑，请您旁边 绕	Rhetorical	Yes	Sun & Chen (2023)
20	让您的行为和这风景一 样美！	Rhetorical	Yes	Ko (2010)

Comparing DeepL and ChatGPT in Translating Chinese Tourist Public Signs: Accuracy and Pragmatic Appropriateness

21	请别摸我，您怕污水， 我怕污染！	Rhetorical	Yes	Ko (2010)
22	一花一草皆生命，一枝 一叶总关情	Rhetorical	Yes	Ko (2010)
23	大地之绿与您同在	Rhetorical	Yes	Ko (2010)
24	我很美丽，但我很脆弱 ，需要您的呵护！	Rhetorical	Yes	Ko (2010)
25	亲爱的，别踩我！	Rhetorical	Yes	Ko (2010)

Note. “Rhetorically marked” refers to signs containing overt personification, poeticized wording, softened directive force, affective address, or other non-neutral rhetorical features.

As Table 1 shows, the corpus was deliberately structured rather than merely accumulated. The reordered numbering aligns the data table, the analytical discussion, and the representative examples, so that each sign can be traced consistently across the article. This organization is methodologically important because the most revealing differences among DeepL, ChatGPT-CP1, and ChatGPT-CP2 do not arise on routine signs alone, but on signs where brevity coexists with personification, mitigation, or poetic compression.

4. Results and Discussion

4.1 General Patterns Across the Three Output Sets

The comparison reveals not a simple hierarchy of system quality, but a patterned distribution of strengths and weaknesses across sign types. DeepL is relatively stable when a Chinese sign corresponds to an established English signage formula. Minimally prompted ChatGPT often remains closer to the source-text surface and therefore preserves more rhetorical texture. Domain-sensitive ChatGPT more frequently produces concise and context-appropriate English public signage, but often at the cost of reduced stylistic and rapport-related nuance. These differences suggest that the translation of tourist public signs is shaped not only by system capability, but also by genre conventions and prompt-conditioned translational priorities (He, 2024; Ko, 2010).

A key finding of the dataset is that the greatest divergence does not emerge on standardized and formulaic signs, but on rhetorically marked ones. Signs with explicit prohibition or straightforward instruction often produce considerable convergence among the three outputs. By contrast, signs involving personification, poetic compression, mitigation, or rapport-building language produce the strongest divergence between direct NMT output, minimally prompted LLM output, and prompt-optimized LLM output. This indicates that the central challenge of tourist public-sign translation lies not in sentence-level transfer alone, but in the interaction between sign function, target-side convention, and source-side rhetorical design. Table 2 presents representative cases from across the corpus in a side-by-side comparative format.

Table 2

Representative Comparative Outputs across the Three Systems

No.	Chinese Sign	DeepL	ChatGPT-CP1	ChatGPT-CP2
1	禁止吸烟	No smoking.	No Smoking.	No Smoking.
10	请勿进入，钙化地带	No Entry: Karst Area.	No Entry. Calcified Area.	No Entry: Calcified Area.

11	人间仙境，不食人间“烟火”	A paradise on earth, untouched by the “smoke and fire” of the mundane world.	This fairyland on earth admits no smoke or fire.	No Smoking or Open Flames in This Scenic Area.
18	小草也在长，踏入想一想	The grass is growing—think twice before stepping on it.	The grass is growing too. Think before stepping on it.	Please Keep Off the Grass.
23	大地之绿与您同在	The green of the earth is with you.	May the green of the earth be with you.	Help Keep the Earth Green.
25	亲爱的，别踩我！	Dear, please don't step on me!	Dear, don't step on me.	Please Do Not Step on the Grass.

4.2 Formulaic-Sign Stability

The first clear pattern is formulaic-sign stability. DeepL performs relatively well on signs whose English counterparts are already conventionalized in public-sign usage. Examples include “禁止吸烟,” “请您照看好自己的物品,” “小心台阶,” “持半票者，出示证件,” and “非工作人员请勿入内.” In these cases, DeepL often generates concise and institutionally natural outputs such as “No smoking,” “Please keep an eye on your belongings,” “Watch your step,” “Half-price ticket holders must present valid ID,” and “Authorized personnel only.” These versions are pragmatically effective because they align with established English public-sign conventions rather than preserving source-language structure too closely. Ko (2010) notes that many serious mistranslations in tourist signage arise when translators fail to activate existing English public-sign formulas, as in cases like “Carefully step” or “Show your certificate for half-ticket.” The present comparison suggests that both DeepL and ChatGPT now handle such routine formulas more successfully than many older human-produced signs did, but DeepL remains particularly stable where a standard English equivalent is readily available. In such cases, the translational problem narrows from genre negotiation to formula retrieval.

ChatGPT-CP1 and ChatGPT-CP2 also perform adequately on formulaic items, but they do not always surpass DeepL. For example, DeepL’s “Authorized personnel only” is more compact and idiomatic than more literal staff-based renderings. Likewise, “Karst Area” is preferable to “calcified area,” because it better matches scenic-site geological terminology. These examples show that prompt optimization does not automatically guarantee better performance: where a sign already has a stable target-language equivalent, DeepL’s default NMT behavior may be entirely sufficient.

4.3 Source-Surface Retention

The second major pattern is source-surface retention, which is most visible in DeepL and, to a lesser extent, in ChatGPT-CP1. This pattern appears when the output preserves source-text imagery, personification, or poetic phrasing more directly than English public-sign convention would normally allow. The result is often semantically traceable and sometimes stylistically attractive, but it risks sounding less like English signage and more like translated discourse. This tendency is especially visible in signs such as “大地之绿与您同在,” “我很美丽，但我很脆弱，需要您的呵护！” and “亲爱的，别踩我！” DeepL renders these as “The green of the earth is with you,” “I am beautiful, but I am fragile—I need your care!” and “Dear, please don't step on me!” ChatGPT-CP1 produces similar outputs, though usually with slightly smoother syntax. In all three cases, the outputs preserve the Chinese rhetorical surface, including personification and affective address. Yet they do not fully align with the conventions

Comparing DeepL and ChatGPT in Translating Chinese Tourist Public Signs: Accuracy and Pragmatic Appropriateness

of English tourist signage, which generally favors clarity and immediate function over overt poeticization or interpersonal playfulness.

Ko's (2010) observations are directly relevant here. He argues that tourist public signs often use descriptive, figurative, or friendly wording to convey simple messages, and that literal or semi-literal transfer tends to produce English that is meaningful only in a limited sense. The English reader may understand the wording, but not experience it as a natural sign. The present dataset confirms this tendency. Outputs such as "The green of the earth is with you" or "Dear, please don't step on me!" preserve rhetorical flavor but remain weak in target-side sign conventionality. The same issue appears in "请别摸我，您怕污水，我怕污染！" DeepL and ChatGPT-CP1 both preserve the personified voice and explanatory contrast. Although the translation is understandable, it sounds stylistically marked and pragmatically odd in English public-sign usage. The output is therefore relatively strong in source-surface fidelity but limited in pragmatic appropriateness.

4.4 Target-Oriented Normalization

The third pattern is target-oriented normalization, and it is most strongly associated with ChatGPT-CP2. Under the domain-sensitive prompt, ChatGPT more frequently resolves softened, poeticized, or personalized Chinese signs into concise English signage formulations. In many cases, this produces output that is more natural in the target environment. This tendency is especially visible in signs such as "人间仙境，不食人间烟火，" "小草有生命，足下多留青，" "青青小草，不踩多好，" and "小草也在长，踏入想一想。" DeepL and ChatGPT-CP1 preserve more of the Chinese wording, whereas ChatGPT-CP2 often moves toward formulations like "No Smoking or Open Flames in This Scenic Area" or "Please Keep Off the Grass." These renderings are less rhetorically rich, but they perform more effectively as English public signs.

The importance of this pattern lies in what it shows about prompt sensitivity. He (2024) suggests that prompt design should not be treated as a neutral wrapper around the model. It changes how the model interprets the task. In the present study, the CP2 prompt explicitly instructs ChatGPT to prioritize concise, natural, and context-appropriate English signage over word-for-word correspondence. The output reflects that instruction. This makes CP2 not "truer" in an absolute sense, but more explicitly aligned with target-side public-sign function. This is also where Nord's (2018) functionalist framework becomes especially useful. From a Skopos-oriented perspective, the adequacy of a translation depends on whether it fulfills the intended function in the target communicative setting. If the target purpose is to produce effective English signage for international visitors, then a normalized output may be preferable even when it reduces source-side rhetorical complexity. In this sense, CP2 can be understood as moving toward function-oriented adequacy rather than simply "translating better."

4.5 Rapport Loss Under Optimization

The fourth pattern is rapport loss under optimization. This is the most theoretically significant pattern in the dataset. It emerges when a domain-sensitive prompt improves target-side signage naturalness but suppresses or removes the interpersonal work performed by the Chinese source sign. This is particularly clear in signs such as "为了您和家人的健康，请勿吸烟，" "水是生命的源泉，请节约用水，" and "请别摸我，您怕污水，我怕污染！" The Chinese originals are not purely directive. They justify, soften, or personify the message in order to reduce direct force and engage the reader more relationally. Sun and Chen's (2023) analysis of Chinese directive public signs shows that this is not incidental. Such wording may serve rapport-maintaining or rapport-enhancing functions by attending to face, rights, interests, and interactional goals.

When ChatGPT-CP2 turns such signs into concise English formulas, it often improves pragmatic appropriateness in one sense while weakening it in another. A target-oriented rendering may be more natural as English signage, but it may no longer preserve the source sign's relational logic. This is why the present study treats pragmatic appropriateness as distinct from fluency. A translation can be target-natural and yet still reduce the interpersonal orientation of the original. This finding complicates any simple "optimized prompt is better" narrative. Prompt optimization does improve functional signage output in many cases, but it also narrows the range of interpersonal meanings that survive in translation. The trade-off is therefore not between good and bad translation, but between different priorities: source-side rhetorical and rapport-oriented texture on the one hand, and target-side signage efficiency on the other.

This does not mean that rapport preservation should always override target-side clarity. In safety-critical or strongly prohibitive signs, the loss of rapport-oriented wording may be justified if it improves immediate comprehensibility and reduces ambiguity. The problem becomes more significant in environmental-protection, civility-oriented, and visitor-care signs, where the source text often relies precisely on softened, affective, or personified wording to encourage voluntary cooperation.

4.6 Representative Cases from the Expanded Corpus

The five additional signs included in the expanded corpus strengthen the analysis because they make the four patterns more visible. Sign 5, “讲究秩序, 注意安全,” confirms formulaic-sign stability. DeepL's “Keep things orderly and stay safe” is slightly less conventional than a native signage formula such as “Please keep order and stay safe,” but it remains clear and usable. Sign 23, “大地之绿与您同在,” illustrates source-surface retention. DeepL's “The green of the earth is with you” preserves the wording but does not sound like English signage. ChatGPT-CP2's “Help Keep the Earth Green” is much more effective in target-side public-sign terms, but it removes the quasi-poetic, blessing-like tone of the original. Sign 18, “小草也在长, 踏入想一想,” is especially revealing because DeepL's version, “The grass is growing—think twice before stepping on it,” remains more source-oriented yet still relatively acceptable in English. This shows that source-surface retention is not always a failure; sometimes it occupies a productive middle ground between literalness and full normalization. Signs 24 and 25, “我很美丽, 但我很脆弱, 需要您的呵护!” and “亲爱的, 别踩我!” are among the clearest cases of rapport loss under optimization. DeepL retains personification and affective appeal. CP1 behaves similarly. CP2, however, shifts toward conventional signage phrasing. These cases demonstrate that the genre problem is not confined to message transfer. It extends to the translation of public intimacy, personalization, and soft control.

The implications of these patterns are both theoretical and practical. Theoretically, they support a genre-sensitive approach to AI-assisted translation evaluation. Tourist public signs are an especially revealing test case because their brevity makes close comparison possible, while their rhetorical and interpersonal density exposes tensions that are easier to ignore in longer texts. Practically, the results suggest that public institutions and tourist sites should not rely on raw AI output without review. DeepL may be sufficient for many formulaic notices, but rhetorically marked signs require more careful intervention. ChatGPT with a domain-sensitive prompt may improve usability in English, but its outputs still need human judgment where source-side tone, politeness, or local communicative style matters.

These findings also clarify the relationship between AI assistance and translation expertise. The question is not whether AI can translate tourist public signs at all. It clearly can, at least to a workable degree in many cases. The more important question is which aspects of the source text each system tends to privilege and what gets lost in the process. In that sense, tourist public signs reveal a broader truth about AI-assisted translation: translation quality is not a single

property but the outcome of competing priorities. Once those priorities are made visible, the evaluation of AI output becomes more precise and more useful for both researchers and practitioners.

5. Conclusions

This study has compared DeepL with ChatGPT under two prompt conditions in the translation of 25 Chinese tourist public signs into English. The comparison shows that the three output conditions do not simply vary in overall quality. Rather, they display different translational tendencies across sign types. DeepL is relatively stable on formulaic signs with established English equivalents. Minimally prompted ChatGPT often preserves more of the source-text surface and therefore retains more rhetorical texture. Domain-sensitive ChatGPT more frequently produces concise and context-appropriate English signage, but it also tends to reduce stylistic and rapport-related nuance.

The main contribution of the study is therefore not the claim that one system is universally superior to another. Rather, it shows that tourist public signs expose the inadequacy of one-dimensional AI translation evaluation. In this genre, semantic correctness alone is insufficient because a translation may preserve lexical meaning while failing as English public signage. At the same time, target-oriented fluency is also insufficient because a translation may function smoothly in English while suppressing source-text rhetorical texture or rapport-building work. Tourist public signs therefore require a dual framework of accuracy and pragmatic appropriateness, because the two dimensions do not always coincide in AI-generated translation.

The study also contributes to current discussions of prompt-sensitive AI translation. It shows that prompt design changes more than wording. It changes translational priorities. Under optimization, ChatGPT becomes more target-oriented and more effective as English public signage, but this improvement may come at the cost of source-side rhetorical and rapport-related density. Prompt engineering should therefore be treated not as a hidden technical detail, but as part of the translation method itself. The analysis further suggests that rapport preservation should be evaluated in relation to sign type: it is less central in safety-critical signs, but more consequential in environmental-protection and civility-oriented signs where interpersonal appeal is part of the sign's persuasive force.

The study remains limited in scope. The corpus is intentionally small, and the present analysis is qualitative rather than rating-based. Even so, this does not weaken the article's core value. On the contrary, the small corpus makes close genre-sensitive comparison possible, while the qualitative design allows the study to identify recurrent patterns that might be obscured by a purely score-driven approach. Future work may enlarge the corpus, add reception-based evidence, or compare other subtypes of tourism signage. For the present study, however, the results already show that tourist public signs are an analytically valuable site for investigating how AI systems negotiate meaning, convention, and interpersonal effect in public, culture-sensitive communication.

In practical terms, the article suggests a cautious but constructive role for AI in tourism communication. It would be simplistic either to celebrate AI output as sufficient or to dismiss it as unusable. A more productive approach is to treat AI as a drafting resource whose usefulness varies by sign type. Formulaic notices may require only light checking, whereas poeticized or rapport-oriented signs call for more substantial human intervention. This distinction is modest, but it is directly relevant to institutions, translators, and editors who are increasingly using AI tools in public-facing multilingual environments.

References

- Alrefaee, Y. M. H. A., & Al-Shammari, A. H. (2026). Digital innovations in second language pragmatics: A systematic review (2006-2024). *Social Sciences & Humanities Open*, 13, 103004. <https://doi.org/10.1016/j.ssaho.2026.103004>
- Amenador, K. B., & Wang, Z. (2022). A systematic review of the Chinese-English translation of public signs. *SAGE Open*, 12(2), 1–16. <https://doi.org/10.1177/21582440221094611>
- Chen, S., & Lin, Y. (2025). A multidimensional comparison of ChatGPT, Google Translate, and DeepL in Chinese tourism texts translation: Fidelity, fluency, cultural sensitivity, and persuasiveness. *Frontiers in Artificial Intelligence*, 8, Article 1619489. <https://doi.org/10.3389/frai.2025.1619489>
- Gao, Y. (2020). Normative research on English translation of tourist attractions in Inner Mongolia. *The Frontiers of Society, Science and Technology*, 2(11), 1–7. <https://doi.org/10.25236/FSST.2020.021101>
- He, S. (2024). Prompting ChatGPT for translation: A comparative analysis of translation brief and persona prompts. In *Proceedings of the 25th Annual Conference of the European Association for Machine Translation* (pp. 316–326).
- Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., Kim, Y. J., Afify, M., & Awadalla, H. H. (2023). *How good are GPT models at machine translation? A comprehensive evaluation* (arXiv Preprint No. arXiv:2302.09210).
- House, J. (2015). *Translation quality assessment: Past and present*. Routledge.
- Hu, X., Niu, Y., & Zhang, X. (2021). Analysis of errors in English translation of public signs in Nanjing red tourist attractions and their countermeasures—Take Yuhuatai Scenic Spot as an example. *International Journal of Frontiers in Sociology*, 3(14), 24–28. <https://doi.org/10.25236/IJFS.2021.031406>
- Ko, L. (2010). Chinese-English translation of public signs for tourism. *The Journal of Specialised Translation*, 13, 111–127.
- Lees, C. (2024). Translating politeness on public notices with a directive function in Thessaloniki: A cross-cultural perspective. *Pragmatics*, 34(4), 534–564.
- Nord, C. (2018). *Translating as a purposeful activity: Functionalist approaches explained* (2nd ed.). Routledge.
- Peng, K., Ding, L., Zhong, Q., Shen, L., Liu, X., Zhang, M., Ouyang, Y., & Tao, D. (2023). Towards making the most of ChatGPT for machine translation. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 5622–5633).
- Pierini, P. (2007). Quality in Web translation: An investigation into UK and Italian tourism Web sites. *The Journal of Specialised Translation*, 8, 85–103.
- Sun, X., & Chen, X. (2023). Rapport management in Chinese directive public signs. *Pragmatics*, 33(4), 618–640.
- Yao, M., Ruan, F., Zhao, S., & Yang, J. (2021). Research on Chinese-English translation of public signs in Wuxi Qingming Bridge Ancient Grand Canal Scenic Zone from the perspective of Skopos theory. *Art and Performance Letters*, 2(6), 36–39. <https://doi.org/10.23977/artpl.2021.020608>

Comparing DeepL and ChatGPT in Translating Chinese Tourist Public Signs: Accuracy and Pragmatic Appropriateness

Zhan, Y. (2021). An analysis of the English translation of public signs in scenic spots: With Ancient Longzhong as an exemplar. *Academic Journal of Humanities & Social Sciences*, 4(6), 39–43. <https://doi.org/10.25236/AJHSS.2021.040608>

Author's Bio

Yang Liu is an MTI student in the College of Foreign Studies at Guilin University of Technology, Guilin, Guangxi, China. Her research interests include translation studies and AI-assisted translation, and Chinese-English public-sign translation. This paper forms part of her ongoing work examining how machine translation and large language models negotiate genre conventions and pragmatic function in specialized public communication contexts.