ORKOPATA
Vol. 5 N° 1 (2026)
pp. 24-38

ORKOPATA
REVISTA DE LINGÜÍSTICA, LITERATURA Y ARTE
ISSN: 2955-8263
https://revistaorkopata.com
Editada por: Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú

ARTÍCULO ORIGINAL

¿Cómo mejorar la traducción literaria del español al chino asistida por la inteligencia artificial?

How can AI-assisted literary translation from Spanish into Chinese be improved?

Comment améliorer la traduction littéraire de l'espagnol vers le chinois à l'aide de l'intelligence artificielle?

Yufei Cao[1]
https://orcid.org/0000-0003-0635-8110
Shanghái International Studies University, Shanghái - Shanghái, China
yufeielisa@qq.com (correspondencia)

Yiwen Ding
https://orcid.org/0000-0003-0358-6767
Shanghái International Studies University, Shanghái - Shanghái, China
dingyiwen0528@163.com

DOI : https://doi.org/10.35622/j.ro.2026.01.002
Recibido: 03-12-2025 **/ Aceptado:** 27-02-2025 **/ Publicado:** 11-03-2026

Resumen

En el marco de la traducción asistida por inteligencia artificial (IA), el uso creciente de los grandes modelos de lenguaje (LLM) abre nuevas posibilidades para la traducción automática. No obstante, su aplicación a textos literarios sigue planteando desafíos relacionados con la preservación del estilo y de los matices culturales. En este contexto, el presente trabajo analiza cómo mejorar la calidad de la traducción literaria del español al chino mediante el uso de LLM y la post-edición. Con este objetivo, se adopta un enfoque mixto secuencial explicativo de alcance exploratorio-descriptivo y carácter comparativo. Para ello, se diseñan cuatro *prompts* de complejidad creciente y se comparan las traducciones generadas por IA con la traducción humana de referencia. Los resultados demuestran que el uso de LLM, combinado con la post-edición humana, puede mejorar la calidad de la traducción literaria del español al chino. Por un lado, esta mejora depende de la redacción de *prompts* detallados que incluyan información contextual, como el rol del traductor, el estilo literario del texto y ejemplos de referencia. Por otro lado, también se debe a una post-edición orientada a naturalizar el lenguaje y adaptar las expresiones a la cultura meta. Se recomienda que la intervención

[1] Profesora y Coordinadora del Departamento de Español de la Escuela de Estudios Europeos y Latinoamericanos de la Universidad de Estudios Internacionales de Shanghái, China.

humana se realice en tres niveles: el ajuste del registro y del ritmo literario, la adecuación cultural mediante el uso de modismos y metáforas propios del chino y la corrección del uso excesivo de signos de puntuación.

Palabras clave: estilo literario, inteligencia artificial, lingüística informática, traducción automática, traducción.

Abstract

In the field of artificial intelligence (AI)-assisted translation, the growing use of large language models (LLMs) opens up new possibilities for machine translation. Nevertheless, their application to literary texts continues to pose challenges related to the preservation of style and cultural nuances. In this context, the present study examines how to improve the quality of literary translation from Spanish into Chinese through the use of LLMs and post-editing. To this end, a sequential explanatory mixed-methods approach was adopted, with an exploratory-descriptive scope and a comparative character. Four *prompts* of increasing complexity were designed, and the translations generated by AI were compared with a human reference translation. The results show that the use of LLMs, combined with human post-editing, can improve the quality of literary translation from Spanish into Chinese. On the one hand, this improvement depends on the formulation of detailed *prompts* that include contextual information, such as the translator's role, the literary style of the text, and reference examples. On the other hand, it also depends on post-editing aimed at making the language sound more natural and adapting expressions to the target culture. Human intervention is recommended at three levels: adjustment of register and literary rhythm, cultural adaptation through the use of idioms and metaphors characteristic of Chinese, and correction of the excessive use of punctuation marks.

Keywords: artificial intelligence, computational linguistics, literary style, machine translation, translation.

Résumé

Dans le domaine de la traduction assistée par l'intelligence artificielle (IA), l'usage croissant des grands modèles de langage (LLM) ouvre de nouvelles possibilités pour la traduction automatique. Néanmoins, leur application aux textes littéraires continue de poser des défis liés à la préservation du style et des nuances culturelles. Dans ce contexte, la présente étude analyse comment améliorer la qualité de la traduction littéraire de l'espagnol vers le chinois au moyen des LLM et de la post-édition. À cette fin, elle adopte une approche mixte séquentielle explicative, à visée exploratoire-descriptive et à caractère comparatif. Pour ce faire, quatre *prompts* de complexité croissante sont conçus, et les traductions générées par l'IA sont comparées à une traduction humaine de référence. Les résultats montrent que l'usage des LLM, combiné à la post-édition humaine, peut améliorer la qualité de la traduction littéraire de l'espagnol vers le chinois. D'une part, cette amélioration dépend de la rédaction de *prompts* détaillés intégrant des informations contextuelles, telles que le rôle du traducteur, le style littéraire du texte et des exemples de référence. D'autre part, elle repose également sur une post-édition visant à naturaliser le langage et à adapter les expressions à la culture cible. Il est recommandé que l'intervention humaine s'effectue à trois niveaux : l'ajustement du registre et du rythme littéraire, l'adaptation culturelle au moyen d'expressions idiomatiques et de métaphores propres au chinois, ainsi que la correction de l'usage excessif des signes de ponctuation.

Mots-clés: intelligence artificielle, linguistique informatique, style littéraire, traduction automatique, traduction.

INTRODUCCIÓN

En la época en que la inteligencia artificial (IA) avanza a pasos agigantados, la traducción automática en combinación con la post-edición se ha convertido en un proceso común que adoptan muchos traductores a la hora de enfrentar las tareas de traducción, ya que la tecnología puede acelerar la eficiencia operativa (Yang et al., 2023; Calvo-Ferrer, 2023; Abdelhalim et al., 2025). El reciente surgimiento y la rápida evolución de los grandes modelos de lenguaje (LLM, por sus siglas en inglés), como ChatGPT o Gemini, están consolidando aún más este proceso. A diferencia de los sistemas tradicionales de traducción automática que siguen un proceso lineal y carecen de retroalimentación dinámica, estos modelos más generales pueden funcionar como agentes lingüísticos interactivos y flexibles, y, por lo tanto, tienen un mayor potencial en la traducción (Vilar et al., 2023; Xu et al., 2023; Zhu et al., 2024). Esta tendencia resulta igualmente perceptible en tareas complejas, como la traducción literaria (Abdelhalim et al., 2025).

Por un lado, Abdelhalim et al. (2025) indican que los estudiantes de traducción prefieren a ChatGPT frente a Google Translate, especialmente al reconocer sus ventajas en el mantenimiento del estilo original, la fluidez textual y el tratamiento de elementos culturales. Por otro lado, Calvo-Ferrer (2023) analiza la capacidad de los usuarios para distinguir entre subtítulos generados por ChatGPT y por traductores humanos, constatando que la calidad de los primeros es comparable a la de los segundos, ya que los participantes no logran diferenciarlos de forma fiable. Farghal y Haider (2024) consideran que los LLM suponen un competidor válido del traductor humano en la traducción poética. Comparan traducciones de poesía árabe clásica realizadas por un traductor humano, ChatGPT y Gemini. La evaluación muestra que las traducciones humanas y las de ChatGPT obtienen buenos resultados en claridad temática, creatividad y prosodia, mientras que Gemini queda por detrás, especialmente en prosodia y creatividad.

Si bien es verdad que la traducción automática en la era de la IA ya ha sido estudiada desde múltiples enfoques, como los errores comunes en la traducción automática (Ferragud Ferragud, 2023; Yang et al., 2023; Calvo-Ferrer, 2023), la interacción entre el género textual y la traducción automática (Kunst & Bierwiaczonek, 2023; Alwazna, 2024; Gao et al., 2024), o la tipología lingüística y la traducción asistida por IA (Yang et al., 2023), entre otros, son todavía escasos los estudios centrados en las traducciones literarias del español al chino, una tarea desafiante en la que se requieren amplios conocimientos culturales y un dominio sólido del estilo literario y de las dos lenguas tipológicamente distintas. En este trabajo se examina este terreno poco explorado mediante un análisis tanto cuantitativo como cualitativo.

En este contexto, el objetivo del presente estudio es analizar cómo mejorar la calidad de la traducción literaria del español al chino mediante el uso de grandes modelos de lenguaje y la post-edición humana. Para ello se abordan dos preguntas de investigación principales: ¿cómo mejorar la calidad de este tipo de traducción con un buen prompt (instrucción utilizada para interactuar con los modelos de IA, diseñada para que estos procesen el mensaje y generen un resultado)?, y ¿cuáles son las mejoras necesarias que deben realizarse en la post-edición? En este trabajo se evalúan primero diferentes versiones traducidas con distintos prompts para identificar la opción más eficaz y, posteriormente, se compara la traducción automática con la versión humana a fin de elaborar una guía práctica para la post-edición.

El presente artículo se organiza de la siguiente manera. En los dos primeros apartados se realiza una breve revisión bibliográfica y se presentan los métodos de la investigación. En el

apartado 3, se discuten los resultados obtenidos. Finalmente, se sintetizan los hallazgos más relevantes en las conclusiones.

METODOLOGÍA

El presente estudio adopta un enfoque empírico con un análisis comparativo, cuyo objetivo es examinar la efectividad de la traducción automática asistida por *prompts* optimizados en la traducción literaria del español al chino. Desde el punto de vista metodológico, la investigación se sitúa dentro de un diseño mixto secuencial explicativo de alcance exploratorio-descriptivo, que combina análisis cuantitativos y cualitativos para comprender mejor el fenómeno estudiado.

La metodología se organiza en dos fases relacionadas. En la primera se realiza un análisis cuantitativo de las traducciones generadas mediante distintos *prompts*, con el fin de comparar sus resultados e identificar la opción más eficaz. En la segunda se lleva a cabo un análisis cualitativo centrado en aspectos lingüísticos, estilísticos y culturales de las traducciones producidas. Este análisis permite interpretar los resultados obtenidos en la fase anterior y examinar las estrategias de post-edición necesarias para mejorar la calidad de la traducción literaria del español al chino.

Diseño experimental

El diseño metodológico se fundamenta en una secuencia de dos fases experimentales consecutivas. La primera fase tiene carácter cuantitativo y se centra en la optimización de *prompts* mediante la métrica BLEU (*Bilingual Evaluation Understudy*). La segunda fase adopta un enfoque cualitativo, basado en el análisis contrastivo de traducciones para identificar patrones de mejora en la post-edición. Esta combinación de métodos permite obtener una visión holística del proceso de traducción asistida por IA.

Corpus de estudio

El material textual seleccionado proviene de la obra *Tradiciones peruanas* (3ª edición, 2006) de Ricardo Palma (1998; 2020), publicada por la editorial Cátedra. Esta elección se justifica por las siguientes razones. En primer lugar, se trata de un texto literario de reconocido valor cultural que presenta desafíos tipológicos y estilísticos propios de la traducción literaria. En segundo lugar, existe una traducción de referencia de alta calidad realizada por el hispanista Fengsen Bai y publicada en 2020 por la Editorial del Pueblo de Sichuan, reconocida en el ámbito académico chino por su fidelidad textual y adecuación estilística. Por último, la editorial de referencia goza de prestigio en China, lo cual garantiza la calidad del corpus comparativo.

El corpus analizado está compuesto por la primera sección de cuatro tradiciones de la obra: *Dos millones*, *El corregidor de Tinta*, *El virrey de la adivinanza* y *Un virrey casamentero*. Como la última tradición no presenta subdivisiones internas, se incorporó el texto completo al corpus. El corpus seleccionado contiene en total 2022 palabras. Estas tradiciones se sitúan en distintos contextos históricos (aproximadamente del siglo XVII al XIX), lo cual permite la incorporación de referencias culturales y lingüísticas de diferentes periodos. Desde el punto de vista discursivo, en los fragmentos seleccionados se encuentran diversas formas narrativas, como diálogos, pasajes poéticos y monólogos, lo cual aporta variedad estilística al corpus. Asimismo, la existencia de una traducción china de alta calidad realizada por Fengsen Bai

permite contar con una referencia fiable para la posterior evaluación de la calidad de las traducciones generadas por IA.

Instrumentos de recolección de datos

Para el primer experimento, se diseñaron cuatro *prompts* de complejidad creciente, siguiendo la taxonomía propuesta por Gao et al. (2024):

Prompt 1 (DT): Descripción básica de la tarea.

Prompt 2 (DT+DP): Incorporación de la descripción del papel o rol del traductor.

Prompt 3 (DT+DP+DE): Adición de la descripción del estilo literario.

Prompt 4 (DT+DP+DE+FL): Inclusión de ejemplos de aprendizaje.

Estos *prompts* se introdujeron en ChatGPT para generar las traducciones automáticas candidatas. El sistema se utilizó en diciembre de 2025, empleando la versión del modelo ChatGPT-5.0 en la interfaz. La generación de las traducciones se realizó mediante la interacción textual directa y se mantuvieron los parámetros de configuración por defecto del sistema, sin modificar variables de generación como *temperature*, *top-p* o el límite máximo de tokens.

Para la evaluación cuantitativa, se empleó la métrica BLEU (Papineni et al., 2002), que cuantifica la similitud entre traducciones candidatas y traducciones de referencia mediante el cálculo de una precisión modificada de n-gramas. Un n-grama es una secuencia contigua de *n* elementos (típicamente palabras) extraída de un texto. Por ejemplo, un bigrama (2-grama) es un par de palabras adyacentes. BLEU compara los n-gramas de la traducción candidata con los de la traducción de referencia y calcula la proporción de coincidencias. Se trata de una precisión modificada porque el cálculo no se limita a contar coincidencias simples, sino que restringe el número de veces que un mismo n-grama puede contabilizarse, evitando así que las repeticiones artificiales afecten la puntuación. Además, la métrica incorpora un mecanismo de penalización por brevedad para evitar que las traducciones demasiado cortas obtengan puntuaciones engañosamente altas.

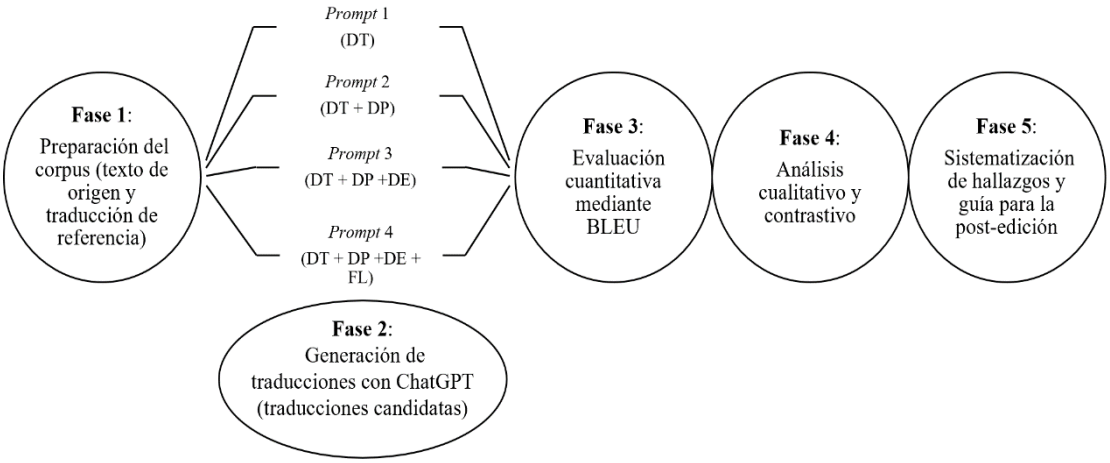
La métrica BLEU es un estándar ampliamente adoptado y reconocido en la evaluación de la traducción automática (Costa-Jussà et al., 2017; Tang et al., 2025), ya que su puntuación se correlaciona estrechamente con los juicios humanos de calidad. Debido a la influencia del tipo textual de origen, de las lenguas meta empleadas y de la cantidad y la calidad de las traducciones de referencia, el valor absoluto de la puntuación BLEU suele carecer de significado, mientras que las diferencias entre puntuaciones comúnmente se utilizan como herramienta para validar la optimización de distintos enfoques técnicos en la traducción automática.

Procedimiento

El procedimiento de la investigación se presenta en la siguiente figura.

Figura 1

Esquema del procedimiento del estudio



Fase 1: Preparación del corpus. Se efectuaron la tokenización y la normalización del texto de origen y la traducción de referencia para asegurar la coherencia en el preprocesamiento textual. La tokenización consiste en segmentar el texto en unidades mínimas comparables (generalmente palabras), lo que permite que el sistema identifique con precisión los n-gramas. La normalización, por su parte, implica aplicar ciertos ajustes formales, como unificar mayúsculas y minúsculas o estandarizar signos de puntuación, con el fin de reducir variaciones superficiales que no afecten al contenido semántico. Después, se realizó la alineación de oraciones del corpus bilingüe español-chino.

Fase 2: Generación de traducciones automáticas. Se obtuvieron cuatro versiones traducidas mediante la interacción con ChatGPT, cada una correspondiente a uno de los *prompts* diseñados. Las traducciones generadas se sometieron al mismo proceso de tokenización, normalización y alineación descrito en el paso anterior.

Fase 3: Evaluación cuantitativa. Se calculó la métrica BLEU para cada versión generada.

Fase 4: Análisis cualitativo. Se seleccionó la versión con la puntuación BLEU más alta para realizar un análisis contrastivo con la traducción humana. El análisis se centró en tres dimensiones: (a) registro y ritmo del lenguaje literario; (b) adecuación cultural; y (c) uso de signos de puntuación.

Fase 5: Sistematización de hallazgos. Se elaboraron tablas comparativas y se identificaron patrones sistemáticos de mejora para la post-edición.

Análisis de datos

Los datos cuantitativos (puntuaciones BLEU) se visualizaron mediante un gráfico de evolución. Los datos cualitativos se analizaron mediante el método de análisis de contenido, y se categorizaron las diferencias entre la traducción automática y la traducción humana según los tres ejes anteriormente mencionados. En la siguiente sección del artículo, los ejemplos seleccionados se presentan en tablas trilingües (español del texto original, chino generado por LLM, chino traducido por el ser humano), acompañados de transcripciones fonéticas (pinyin) y glosas interpretativas para facilitar la comprensión del análisis.

RESULTADOS Y DISCUSIÓN

Análisis cuantitativo

En el primer experimento diseñamos cuatro *prompts* con complejidad progresiva para interactuar con LLM, siguiendo el paradigma experimental de Gao et al. (2024). Los *prompts* se redactaron en inglés, porque el rendimiento de estos modelos depende en gran medida de la cantidad de datos disponibles en el corpus de entrenamiento para cada lengua y presentan un rendimiento significativamente superior en inglés (Martínez-Murillo et al., 2025; Li et al., 2024). Veamos la siguiente tabla:

Tabla 1
Prompts con complejidad creciente

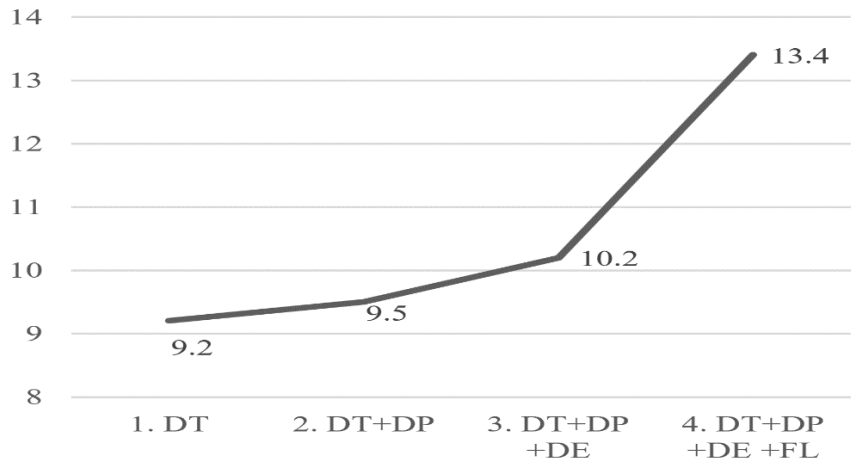
	Descripción	***Prompt***
1	DT (Descripción de la tarea)	This is a Spanish to Chinese translation task. Provide the Chinese translation for the following sentences: [texto de origen]
2	DT **+ DP** (Descripción del papel)	**You are an expert translator.** Your task is to translate the following text from Spanish to Chinese. Provide the Chinese translation for the following sentences: [texto de origen]
3	DT + DP **+ DE** (Descripción del estilo)	You are an expert translator. Your task is to translate the following text from Spanish to Chinese. **The following story is a Peruvian folk tale.** Provide the Chinese translation for the following sentences: [texto de origen]
4	DT + DP + DE **+ FL** (*Few-shot Learning*, aprendizaje con ejemplos)	You are an expert translator. Your task is to translate the following text from Spanish to Chinese. The following story is a Peruvian folk tale. **Below are a few examples of excellent translations of two excerpts from Peruvian folk tale from Spanish to Chinese. Please study their style and approach:** **[Example 1]** <Source Text in Spanish>: Por los años de 1817 un joven escocés, de aire bravo y simpático, se presentó a las autoridades de Valparaíso solicitando un puesto en la marina de Chile, y comprobando que había servido como aspirante en la armada real de Inglaterra. <Expert Translation in Chinese>: 大约一八一七年，一个仪表堂堂、讨人喜欢的苏格兰青年到了瓦尔帕莱索当局，请求在智利海军中谋职，并出示文件证明自己曾是英国皇家舰队候补队员。**[Example 2]** <Source Text in Spanish>: Era Robertson valiente hasta el heroísmo, de mediana estatura, rojizos cabellos y penetrante mirada. Su carácter fogoso y apasionado lo arrastraba a ser feroz. < Expert Translation in Chinese>:

	罗伯逊中等身材，头发微红，目光锐利，胆大妄为甚至有点冒险，性情火暴甚至有点残酷。 Now, translate the following text in the same spirit and quality as the examples above: [texto de origen]

Para evaluar la calidad de las traducciones generadas con los diferentes *prompts* presentados en la Tabla 1, se utilizó la métrica BLEU. De acuerdo con Papineni et al. (2002), la implementación de esta métrica permite una comparación objetiva y reproducible. Se trata de una de las herramientas de evaluación más utilizadas en la investigación de la traducción automática debido a su facilidad de cálculo, su gran operatividad y su alto grado de automatización (Li & Zhan, 2025). En comparación con otros métodos de evaluación de la traducción automática, se acerca más a los resultados de la evaluación humana (Hadla et al., 2015). Las puntuaciones BLEU obtenidas se presentan en la Figura 2.

Figura 2

Puntuaciones BLEU de las traducciones generadas con los cuatro prompts



Las puntuaciones BLEU sirven para cuantificar la similitud entre las traducciones candidatas y la traducción de referencia. Una puntuación más alta indica mayor proximidad al texto de referencia y, por consiguiente, mejor calidad de traducción. Como muestra la Figura 2, la incorporación de más detalles al *prompt*, tales como el papel que juega el modelo, el estilo literario de la obra por traducir, así como algunos ejemplos de traducción, ha mejorado bastante la calidad de las traducciones generadas.

Análisis cualitativo

El análisis cualitativo consiste en comparar la traducción automática de LLM con la puntuación BLEU más alta y la traducción humana de Bai para detectar los puntos aún por mejorar en la post-edición. A continuación, analizaremos, a través de una lectura comparativa minuciosa, las diferencias más destacadas, ilustradas con ejemplos representativos.

En primer lugar, se observa que la versión generada por ChatGPT se acerca más a una traducción literal, mientras que en la versión de Bai se encuentra un lenguaje más literario y rítmico. Veamos algunos ejemplos.

Tabla 2
Comparación del registro y ritmo del lenguaje literario

Texto de origen	Traducción de ChatGPT	Traducción de Bai
(1) El 16 de julio de 1826 fue día **de gran agitación** en Lima y el Callao.	一八二六年七月十六日，利马和卡亚俄一带**一片骚动**(*yī piàn sāodòng*, 'conmoción general')。	一八二六年七月十六日那一天，利马和卡亚俄**全城骚动，群情激愤** (*quánchéng sāodòng, qún qíng jīfèn*, 'ciudad en alboroto y multitudes indignadas')。
(2) [**El buque**] **fue abordado por una lancha** con trece hombres.	一艘载有十三名男子的**小艇靠船登临**(*xiǎo tǐng kào chuán dēnglín*, 'la lancha realiza el abordaje')。	一只**小艇靠近大船**(*xiǎo tǐng kàojìn dàchuán*, 'la lancha se acerca al buque')，艇上十三个人。
(3) [...] procedieron **con tal cautela y rapidez**, que la ronda del resguardo no pudo advertir lo que acontecía.	[...]行事**极为谨慎而迅捷** (*jíwéi jǐnshèn ér xùnjié*, 'con extremada precaución y celeridad')，以至于海关巡逻毫无察觉。	[...]行动**诡秘而迅疾**(*guǐmì ér xùnjí*, 'sigilosa y velozmente')，岸上巡逻队未能发现出了什么事。

En el primer ejemplo presentado, ChatGPT utiliza "一片骚动" (*yī piàn sāodòng*, 'conmoción general') para traducir literalmente la frase original "gran agitación". Aunque el texto traducido se acerca al texto de origen, la traducción automática resulta plana y sin ritmo. En cambio, en la traducción humana, el traductor emplea dos frases con estructura similar "全城骚动，群情激愤" (*quánchéng sāodòng, qúnqíng jīfèn*, 'ciudad en alboroto y multitudes indignadas') para expresar la misma idea. Estas expresiones paralelas y simétricas, que suenan contundentes, consiguen transmitir mejor la intensa emoción colectiva que conlleva el texto original. La estructura tetrasilábica de patrón 2+2, es decir, una secuencia de cuatro caracteres organizada en dos unidades bisilábicas paralelas, constituye una de las unidades rítmicas más estables y productivas del chino, y representa una importante tendencia prosódica del chino moderno (Jing, 2011; Feng, 2019). En el ejemplo «全城骚动，群情激愤» (*quánchéng sāodòng, qúnqíng jīfèn*, 'ciudad en alboroto y multitudes indignadas'), cada expresión se compone de dos segmentos de dos caracteres (*quánchéng* / *sāodòng* y *qúnqíng* / *jīfèn*), lo cual genera un ritmo equilibrado y enfático. Además, en el contexto cultural chino, frecuentemente caracterizado como de alto contexto, la aparición paralela de estructuras tetrasilábicas puede ofrecer al lector una experiencia perceptiva más rica; a su vez, estas unidades presentan una mayor flexibilidad semántica y funcional, lo que facilita su adaptación a la interpretación global del contexto (Wang, 2008; Wang & Gao, 2019). Farghal y Haider (2024) sostienen que ChatGPT puede reproducir la prosodia de la poesía árabe de manera comparable a la de los traductores humanos. Sin embargo, los resultados del presente estudio sugieren que, en el caso del chino, el modelo todavía muestra limitaciones en la reproducción de patrones rítmicos característicos, especialmente en el uso natural de estructuras tetrasilábicas, donde aún no alcanza el nivel de naturalidad logrado por los traductores humanos.

En el segundo ejemplo, ChatGPT utiliza una expresión muy formal "小艇靠船登临" (*xiǎo tǐng kào chuán dēnglín*, 'la lancha realiza el abordaje') para traducir la frase original. Aunque logra transmitir la idea del texto en español, el empleo de esta frase con tanta formalidad no resulta adecuado en este contexto narrativo. En contraste con ello, el traductor Bai utiliza en este caso una expresión sencilla "小艇靠近大船" (*xiǎo tǐng kàojìn dàchuán*, 'la lancha se acerca al buque'), la cual suena más natural y es, por consiguiente, más apropiada para esta ocasión.

En el tercer ejemplo, en comparación con la versión generada por ChatGPT "极为谨慎而迅捷" (*jíwéi jǐnshèn ér xùnjié*, 'con extremada precaución y celeridad'), la traducción humana "诡秘而迅疾" (*guǐmì ér xùnjí*, 'sigilosa y velozmente') emplea una frase más rítmica y destaca con más claridad la forma propia de actuación de los piratas.

Estos dos ejemplos ponen de manifiesto que, además de los ajustes necesarios en el plano rítmico, las traducciones generadas por LLM también requieren la intervención de traductores humanos para garantizar la precisión léxica. Este resultado coincide con lo señalado por Yang et al. (2023), quienes observan que, en los procesos de post-edición, los errores relacionados con la adecuación superan a aquellos vinculados con la fluidez. En este sentido, se sugiere que, si bien los modelos de lenguaje pueden producir traducciones relativamente fluidas, la adecuación semántica y la selección léxica continúan siendo aspectos que requieren una supervisión humana cuidadosa.

En segundo lugar, como el español y el chino son dos lenguas tipológicamente y culturalmente muy distantes, es necesario utilizar más frases hechas o modismos construidos en la cultura china para que la traducción resulte más familiar para los lectores de destino.

Tabla 3
Comparación de la adecuación cultural

Texto de origen	Traducción de ChatGPT	Traducción de Bai
(4) **A casar gente**, se ha dicho.	于是，他下定决心：**给人们配对成婚** (*gěi rénmen pèiduì chénghūn*, 'emparejar y casar a la gente')。	所以人们早就有言在先：**男大当婚，女大当嫁** (*nán dà dāng hūn, nǚ dà xū jià*, 'cuando los hijos se hacen mayores, deben formar sus propias familias')。
(5) [...] que si lo que ha menester es posición y dinero, yo me ocupo de **cambiar su suerte de mala en venturosa**.	若你所缺的是地位与财富，我自会替你**把坏运气改成好运气** (*bǎ huài yùnqì gǎi chéng hǎo yùnqì*, 'convertir la mala suerte en la buena suerte')。	如果您需要的是地位和金钱，我来负责改变您的命运；保您**祸去福来** (*huò qù fú lái*, 'se va la mala suerte y viene la buena ventura')。
(6) No, señor; que tengo a doña Beatriz por **un pino de oro**.	不，阁下；在我眼中，贝阿特丽斯如同**金松** (*jīn sōng*, 'pino de oro')一般。	不，先生，我看她像**兰花玉立** (*lánhuā yù lì*, 'elegante y erguida como una orquídea')。

En los ejemplos (4), (5) y (6) presentados, aunque las frases en negrita generadas por ChatGPT corresponden literalmente a los textos de origen y transmiten las mismas ideas contenidas en ellos, suenan poco naturales para los lectores chinos. En cambio, para efectuar las mismas tareas, el traductor Bai recurre a modismos chinos de cuatro caracteres: "男大当婚，女大当嫁" (*nán dà dāng hūn, nǚ dà xū jià*, 'cuando los hijos se hacen mayores, deben formar sus propias familias'); "祸去福来" (*huò qù fú lái*, 'se va la mala suerte y viene la buena ventura'); "兰花玉立" (*lánhuā yù lì*, 'elegante y erguida como una orquídea'). Estas expresiones, además de ser breves y concisas, son de uso frecuente en China, y, por consiguiente, resultan más familiares para los lectores de destino. En el último caso, en la versión original en español se emplea la expresión "pino de oro" para describir a una señora elegante, mientras que el traductor Bai opta por una expresión metafórica más cercana a la tradición china, es decir, usa la "orquídea" para referirse a la misma señora erguida con mucha elegancia.

La domesticación y la extranjerización son dos estrategias opuestas pero complementarias en la práctica de la traducción (Venuti, 2008; Al-Rikaby et al., 2018). En la práctica de la traducción literaria, el uso excesivo o exclusivo de estrategias de extranjerización para tratar los elementos culturalmente cargados puede hacer que el texto traducido resulte rígido, difícil de comprender para los lectores de la cultura meta o incluso susceptible de generar malentendidos culturales. En este sentido, la domesticación responde a la necesidad de tener en cuenta el trasfondo cultural y los hábitos lingüísticos del público receptor, con el objetivo de *"hacer que la traducción sea más comprensible para el público destinatario"* (Yang, 2023).

En tercer lugar, en el texto traducido por ChatGPT, se observa, en numerosos casos, el uso excesivo de signos de puntuación. Veamos a continuación dos ejemplos representativos.

Tabla 4
Comparación del uso de signos de puntuación

Texto de origen	Traducción de ChatGPT	Traducción humana
(7) [...] sino **la noticia de** haber desaparecido del fondeadero el bergantín inglés Peruvian, cargado con dos millones de pesos en oro, barras de plata y moneda sellada.	[...] 而是**一则消息：** (*yī zé xiāoxi*:, 'una noticia:')停泊在锚地的英国双桅帆船"秘鲁号"不翼而飞。该船装载着价值两百万比索的黄金、银条和铸币。	[...] 而是**有消息说** (*yǒu xiāoxī shuō*, 'hay una noticia que dice')，装载着价值二百万比索黄金、银锭和蜡封钱币的英国造双桅帆船"秘鲁人号"，在停泊地点失踪了。
(8) Ocupaba aquella mañana la cabecera de la mesa, teniendo a su izquierda a un descendiente de los Incas, **llamado don José Gabriel Tupac-Amaru**, y a su derecha a doña Micaela Bastidas, esposa del cacique.	那天上午，他坐在宴席主位，左侧坐着一位印加皇族后裔**——何塞·加夫列尔·图帕克·阿马鲁阁下** (*- hé sài · jiā fú liè ěr · tú pà kè · ā mǎ lǔ géxià*, '- don José Gabriel Tupac-Amaru')，右侧则是酋长夫人米卡埃拉·巴斯蒂达斯夫人。	那天，他坐在宴席的首位，左侧是一位印卡王后裔，**名叫堂何塞·加夫列尔·图帕克-阿马鲁** (*míng jiào táng hé sài · jiā fú liè ěr · tú pà kè-ā mǎ lǔ*, 'llamado don José Gabriel Tupac-Amaru')，右侧是这位酋长的妻子堂娜米卡埃拉·巴斯蒂达斯。

En el ejemplo (7), para introducir el contenido de la noticia, en la traducción automática se emplea la expresión con el signo de dos puntos "一则消息：" (*yī zé xiāoxi*:, 'una noticia:'), la cual no coincide con los hábitos de lectura del público chino. En esta ocasión, la versión traducida por el señor Bai "有消息说" (*yǒu xiāoxī shuō*, 'hay una noticia que dice') es mucho más natural para los lectores chinos. En el ejemplo (8), en la traducción de ChatGPT el empleo de la raya para indicar el nombre del descendiente de los Incas tampoco es natural para los nativos chinos, la versión más aceptable es la del señor Bai, quien usa "名叫" (*míng jiào*, 'llamado') para efectuar la misma tarea.

Los ejemplos presentados indican que, una vez obtenida la traducción realizada por LLM, es necesario realizar una serie de mejoras a diferentes niveles, las cuales se recogen en la siguiente tabla.

Tabla 5
Categorización de intervenciones humanas

Categoría	Subcategoría	Ejemplo
Nivel estilístico-rítmico	Ajuste al registro literario y a paralelismos	Uso del registro adecuado y reestructuración de secuencias binarias
Nivel pragmático-cultural	Adaptación de fraseología y uso de metáforas adecuadas	Sustitución de equivalencias literarias por modismos y empleo de recursos metafóricos apropiados
Nivel ortográfico	Puntuación	Corrección del uso excesivo de signos de puntuación

La tabla anterior propone una guía práctica para la post-edición de traducciones literarias del español al chino asistida por IA: la intervención humana en la post-edición debe realizarse en tres niveles. Al nivel estilístico-rítmico, se requiere el ajuste al registro literario y a paralelismos; al nivel pragmático-cultural, la adecuación cultural mediante el uso de modismos y metáforas propios del chino; y al nivel ortográfico, la corrección del uso excesivo de signos de puntuación.

CONCLUSIONES

En conclusión, con el avance de la inteligencia artificial, la traducción asistida por LLM en combinación con la post-edición se ha consolidado como un método cada vez más relevante en el ámbito de la traducción literaria. En este proceso, para obtener una traducción de mayor calidad, es necesario redactar primero un buen *prompt*, en el que, además de la descripción de la tarea, se incluya información detallada relacionada con la traducción, tales como el papel que juega el modelo, el estilo literario del texto por traducir y algunos ejemplos a seguir.

Una vez obtenida la versión traducida por LLM, se sugiere realizar una serie de mejoras en la etapa de post-edición. En primer lugar, conviene pulir el lenguaje para que resulte más natural y más armónico en su ritmo. En segundo lugar, puede resultar conveniente emplear con mayor frecuencia modismos y metáforas arraigados en la cultura china. Por último, se recomienda evitar el uso excesivo de los signos de puntuación, tales como los dos puntos o las rayas.

De cara a la continuidad de esta línea de investigación, resulta pertinente ampliar el corpus y diversificar traducciones de referencia para validar la efectividad de los *prompts* propuestos en

contextos más diversos. También será útil incluir más métricas complementarias tales como COMET o BERTScore para obtener una evaluación más precisa de la calidad de traducción. Finalmente, se propone explorar, con modelos más específicos, mecanismos de adaptación interlingüística específicos del par español-chino.

Conflicto de intereses / Competing interests:

Los autores declaran que no incurrieron en conflictos de intereses.

Rol de los autores /Authors Roles:

Yufei Cao: Conceptualización, metodología, software, validación, análisis formal, investigación, recursos, escritura – borrador original, escritura – revisión y edición, supervisión, administración del proyecto, adquisición de fondos.

Yiwen Ding: Conceptualización, metodología, software, curación de datos, análisis formal, investigación, recursos, escritura – borrador original, escritura – revisión y edición, visualización.

Fuentes de financiamiento / Funding:

Los autores declaran que este artículo no ha sido financiado.

Aspectos éticos / legales; Ethics / legals:

Los autores declaran no haber incurrido en aspectos antiéticos, ni haber omitido aspectos legales en la realización de la investigación.

REFERENCIAS

Abdelhalim, S. M., Alsahil, A. A., & Alsuhaibani, Z. A. (2025). Artificial intelligence tools and literary translation: a comparative investigation of ChatGPT and Google Translate from novice and advanced EFL student translators' perspectives. *Cogent Arts & Humanities*, *12*(1), 2508031. https://doi.org/10.1080/23311983.2025.2508031

Al-Rikaby, A. B. M., Mahadi, T. S. T., & Lin, D. T. A. (2018). Domestication and foreignization strategies in two Arabic translations of Marlowe's Doctor Faustus: Culture-bound terms and proper names. *Pertanika Journal of Social Sciences and Humanities, 26*(2), 1019–1033.

Alwazna, R. Y. (2024). The use of automation in the rendition of certain articles of the Saudi Commercial Law into English: a post-editing-based comparison of five machine translation systems. *Frontiers in Artificial Intelligence*, *6*, 1282020. https://doi.org/10.3389/frai.2023.1282020

Calvo-Ferrer, J. R. (2023). Can you tell the difference? A study of human vs machine-translated subtitles. *Perspectives*, *32*(6), 1115-1132. https://doi.org/10.1080/0907676x.2023.2268149

Costa-Jussà, M. R., Aldón, D., & Fonollosa, J. A. (2017). Chinese–Spanish neural machine translation enhanced with character and word bitmap fonts. *Machine Translation*, *31*(1), 35-47. https://doi.org/10.1007/s10590-017-9196-0

Farghal, M., & Haider, A. S. (2024). Translating classical Arabic verse: human translation vs. AI large language models (Gemini and ChatGPT). *Cogent Social Sciences*, *10*(1), 2410998. https://doi.org/10.1080/23311886.2024.2410998

Feng, S. (2019). *Prosodic syntax in Chinese: Theory and facts*. Routledge.

Ferragud Ferragud, M. (2023). La traducció automàtica literària: Anàlisi d'errors de la traducció automàtica i les traduccions d'estudiants del mateix text original. *Tradumàtica Tecnologies de la Traducció*, *21*, 184–232. https://doi.org/10.5565/rev/tradumatica.334

Gao, Y., Wang, R., & Hou, F. (2024). *How to design translation prompts for ChatGPT: An empirical study*. In R. Wang, Z. Wang, J. Liu, A. Del Bimbo, J. Zhou, A. Basu, & M. Xu (Eds.), *Proceedings of the 6th ACM International Conference on Multimedia in Asia, Workshops, MMAsia 2024* (pp. 1–7). Association for Computing Machinery. https://doi.org/10.1145/3700410.3702123

Hadla, L. S., Hailat, T. M., & Al-Kabi, M. N. (2015). Comparative study between METEOR and BLEU methods of MT: Arabic into English translation as a case study. *International Journal of Advanced Computer Science and Applications, 6*(11), 215–221. http://dx.doi.org/10.14569/ijacsa.2015.061128

Jing, S. (2011). 自拟四字格与英汉文学翻译 [Self-created four-character structures and English-Chinese literary translation]. *Bianji Zhiyou*, (10), 88–90. https://doi.org/10.13786/j.cnki.cn14-1066/g2.2011.10.001

Kunst, J. R., & Bierwiaczonek, K. (2023). Utilizing AI questionnaire translations in cross-cultural and intercultural research: Insights and recommendations. *International Journal of Intercultural Relations*, *97*, 101888. https://doi.org/10.1016/j.ijintrel.2023.101888

Li, H., & Zhan, J. (2025). 基于 BLEU 指标的机器翻译译文差异对比——以 China Daily 新闻为例 [A comparison of machine translation output differences based on the BLEU metric: A case study of China Daily news]. *Modern Linguistics, 13*(10), 92–96. https://doi.org/10.12677/ml.2025.13101030

Li, Z., Shi, Y., Liu, Z., Yang, F., Liu, N., & Du, M. (2024). *Quantifying multilingual performance of large language models across languages* [preprint]. arXiv. https://arxiv.org/html/2404.11553v1

Martínez-Murillo, I., Lloret, E., Moreda, P., & Gatt, A. (2025). *Do LLMs exhibit the same commonsense capabilities across languages?* [preprint]. arXiv. https://arxiv.org/pdf/2509.06401v1

Palma, R. (1998). *Tradiciones Peruanas: Selección*. Ediciones Cátedra.

Palma, R. (2020). *Tradiciones peruanas* (F. Bai, Trad.). Sichuan People's Publishing House.

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). *BLEU: A method for automatic evaluation of machine translation*. In P. Isabelle, E. Charniak, & D. Lin (Eds.), *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (pp. 311–318). Association for Computational Linguistics. https://doi.org/10.3115/1073083.1073135

Tang, L., Qin, J., Ye, W., Tan, H., & Yang, Z. (2025). *Adaptive few-shot prompting for machine translation with pre-trained language models* [preprint]. arXiv. http://arxiv.org/abs/2501.01679

Venuti, L. (2008). *The Translator's Invisibility: A History of Translation*. Routledge.

Vilar, D., Freitag, M., Cherry, C., Luo, J., Ratnakar, V., & Foster, G. (2023). *Prompting PaLM for translation: Assessing strategies and performance*. In A. Rogers, J. Boyd-Graber, & N. Okazaki (Eds.), *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 15406–15427). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.859

Wang, J. (2008). A Cross-cultural Study of Daily Communication between Chinese and American--From the Perspective of High Context and Low Context. *Asian Social Science*, *4*(10), 151-154. https://doi.org/10.5539/ass.v4n10p151

Wang, W., & Gao, J. (2019). *On the chunkiness and discreteness of Chinese four-character idioms. Journal of Beijing International Studies University, 41*(2), 3–20. https://doi.org/10.12002/j.bisu.203

Xu, H., Kim, Y. J., Sharaf, A., & Awadalla, H. H. (2023). *A paradigm shift in machine translation: Boosting translation performance of Large Language Models* [preprint]. arXiv. http://arxiv.org/abs/2309.11674

Yang, J. (2023). Hybrids in literary translation: Binary translation strategies in Howard Goldblatt's English translation of Mo Yan's Frog. *Studia Metrica et Poetica, 10*(1), 108–125. https://doi.org/10.12697/smp.2023.10.1.05

Yang, Y., Liu, R., Qian, X., & Ni, J. (2023). Performance and perception: machine translation post-editing in Chinese-English news translation by novice translators. *Humanities and Social Sciences Communications*, *10*(1), 1-8. https://doi.org/10.1057/s41599-023-02285-7

Zhu, W., Liu, H., Dong, Q., Xu, J., Huang, S., Kong, L., Chen, J., & Li, L. (2024). *Multilingual machine translation with large language models: Empirical results and analysis*. In K. Duh, H. Gomez, & S. Bethard (Eds.), *Findings of the Association for Computational Linguistics: NAACL 2024* (pp. 2765–2781). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-naacl.176