
This is the **published version** of the bachelor thesis:

Puigbert Moreno, Ona; Vargas-Urpi, Mireia , dir. Comparación de herramientas de traducción automática e inteligencia artificial para el chino coloquial. 2024. 71 pag. (Grau en Traducció i Interpretació)

This version is available at <https://ddd.uab.cat/record/300059>

under the terms of the  license

FACULTAT DE TRADUCCIÓ I D'INTERPRETACIÓ

GRAU DE TRADUCCIÓ I INTERPRETACIÓ

TREBALL DE FI DE GRAU

Curs 2023-2024

**Comparación de herramientas de traducción
automática e inteligencia artificial para el chino
coloquial**

Ona Puigbert Moreno

TUTORA

Mireia Vargas Urpí

Barcelona, 7, juny 2024

UAB
**Universitat Autònoma
de Barcelona**

Datos del TFG

Título: Comparación de herramientas de traducción automática e inteligencia artificial para el chino coloquial

Autora: Ona Puigbert Moreno

Tutora: Mireia Vargas Urpí

Centro: Facultad de Traducción e Interpretación (UAB)

Estudios: Grado en Traducción e Interpretación

Curso académico: 2023-24

Palabras clave

Traducción automática, inteligencia artificial, jerga de Internet, neologismo, chino coloquial

Traducció automàtica, intel·ligència artificial, argot d'Internet, neologisme, xinès col·loquial

Automatic translation, artificial intelligence, Internet slang, neologism, colloquial Chinese

Resumen del TFG

En los últimos años, las herramientas de traducción automática (TA) y la inteligencia artificial (IA) aplicada a la traducción han mejorado y crecido exponencialmente. En cuanto al par chino-español, a pesar de que el número de estudios siga siendo bajo, hay algunos que tratan todos los factores importantes que hay que tener en cuenta al usar herramientas de TA o IA, como por ejemplo el uso del inglés como lengua pivote o la precisión de las traducciones en los sintagmas más usados. Así pues, aunque los resultados sean mucho más precisos que hace unos años, aún tienen un largo camino de mejora.

Por ende, se seleccionarán términos coloquiales chinos de distintas cuentas de enseñanza en redes sociales para, más tarde, comparar las traducciones ofrecidas por motores de traducción automática e inteligencia artificial como DeepL, Google Traductor, Chat GPT y Gemini (antes conocida como Bard). De esta manera, se ampliará el número de estudios con el par chino-español, se examinará la precisión de las traducciones y se determinará en qué medida estas aplicaciones pueden ser útiles para un traductor profesional.

En els últims anys, les eines de traducció automàtica (TA) i la intel·ligència artificial (IA) aplicada a la traducció han millorat i crescut exponencialment. Quant al par xinès-espanyol, tot i que el nombre d'estudis segueix sent baix, hi ha alguns que tracten tots els factors importants a considerar en utilitzar eines de TA o IA, com per exemple l'ús de l'anglès com a llengua pivot o la precisió de les traduccions en els sintagmes més utilitzats. Així doncs, malgrat que els resultats siguin molt més precisos que fa uns anys, encara tenen un llarg camí de millora.

Per tant, es seleccionaran termes col·loquials xinesos de diferents comptes d'ensenyança en xarxes socials per, més tard, comparar les traduccions ofertes pels motors de traducció automàtica i intel·ligència artificial com DeepL, Google Traductor, Chat GPT i Gemini (abans coneguda com Bard). D'aquesta manera, s'ampliarà el nombre d'estudis amb el par xinès-espanyol, s'examinarà la precisió de les traduccions i es determinarà en quina mesura aquestes aplicacions poden ser útils per a un traductor professional.

In recent years, machine translation (MT) tools and artificial intelligence (AI) applied to translation have improved and grown exponentially. As for the Chinese-Spanish pair, although the number of studies is still low, there are some that deal with all the important elements that must be considered when using MT or AI tools, such as the use of English as a pivot language or the accuracy of translations in the most used phrases. Thus, even though the results are much more precise than a few years ago, they still have a long way to improve.

Therefore, Chinese colloquial terms will be selected from different social media teaching accounts to afterwards compare the translations offered by machine translation and artificial intelligence engines like DeepL, Google Translate, Chat GPT and Gemini (formerly known as Bard). This way, the number of studies with the Chinese-Spanish pair will increase, the accuracy of the translations will be examined, and it will be determined to what extent these applications can be useful for a professional translator.

Tabla de contenido

INTRODUCCIÓN.....	1
Presentación del tema.....	1
Justificación del tema	2
Objetivos del trabajo	3
Estructura del trabajo	4
PRIMERA PARTE: MARCO TEÓRICO.....	5
1. La traducción automática e inteligencia artificial aplicadas a la traducción	5
1.1.La traducción automática	5
1.2.La inteligencia artificial	9
1.3.Estudios previos en el caso del chino-español	11
2. El chino coloquial: su relación con las redes sociales	13
2.1.La aparición de neologismos en las redes sociales	14
2.2.Las redes sociales como plataforma para aprender jerga coloquial	15
SEGUNDA PARTE: MARCO PRÁCTICO.....	19
3. Metodología	19
3.1.Método de recogida de datos.....	19
3.2.Selección de los sistemas de TA e IA	23
3.3.Método de análisis de datos	24
4. Resultados del análisis	26
4.1.Descripción de los ejemplos utilizados	26
4.2.Resultados de las traducciones realizadas por las herramientas	29
4.3.Comparación y discusión de los resultados	34
4.4.Análisis cualitativo de las traducciones por TA e IA en el contexto de las redes sociales	36
CONCLUSIONES.....	46
REFERENCIAS	49
Lista de fuentes bibliográficas.....	49

ANEXOS.....	53
Anexo 1: Capturas de pantalla de ejemplos en redes sociales	53
Anexo 2: Interacciones con los chatbots	65

Abreviaturas	Significado
GLN	Generación de Lenguaje Natural
IA	Inteligencia artificial
LLM	<i>Large Language Models</i>
LM	Lengua meta
LO	Lengua original
NLG	<i>Natural Language Generation</i>
NMT	<i>Neural Machine Translation</i>
RBMT	<i>Rule Based Machine Translation</i>
SMT	<i>Statistical Machine Translation</i>
TA	Traducción automática
TABR	Traducción automática basada en reglas
TM	Texto meta
TO	Texto original

Tabla 1: Abreviaturas recurrentes y su significado. [Elaboración propia].

INTRODUCCIÓN

Presentación del tema

Desde que el ser humano es capaz de hablar y comunicarse, cada grupo social se ha encargado de crear un bagaje idiomático con expresiones y vocablos que en general solo ese conjunto de personas conoce o entiende. En ese aspecto, el idioma chino no es una excepción, y en esta investigación nos centraremos en el lenguaje coloquial usado entre jóvenes y adolescentes chinos u otros usuarios de las redes sociales de China, y en cómo lo traducen traductores automáticos como Traductor de Google o DeepL y otras herramientas digitales como Chat GPT y Gemini.

Cualquier persona que está aprendiendo un idioma extranjero, al empezar a dominarlo un poco más, se aventura a aprender expresiones o palabras más modernas y cotidianas que no enseñarán en el aula. Una manera de aumentar el nivel del idioma es hablando con nativos, por lo que es posible que los jóvenes, por ejemplo, para poder comunicarse con otros jóvenes extranjeros, intenten aprender vocabulario coloquial juvenil. Es cierto que, a veces, el profesorado suele centrarse más en el aprendizaje del chino estándar, y no tanto en el coloquial, ya sea por usar aún los sistemas tradicionales de enseñanza de idiomas extranjeros u otras razones. Por ende, hay mucho alumnado que recurre a Internet, en concreto, a las redes sociales como Instagram o Tik Tok, para aprender neologismos de la lengua extranjera que estén aprendiendo. Estas plataformas están llenas de profesores en línea que, en publicaciones o vídeos cortos, te enseñan alguna que otra expresión coloquial utilizada actualmente.

Hasta este punto, podría parecer que este procedimiento es sencillo; sin embargo, en realidad, existen múltiples consideraciones que se deben tener en cuenta al recurrir a una publicación o vídeo de Internet: ¿Es verdad lo que explican? ¿Cómo podemos saber que es realmente alguien nativo o que conoce bien la lengua que enseña? ¿Es realmente un neologismo del ámbito coloquial usado en la actualidad? Y es que, hoy en día, el problema que tiene el Internet radica en eso: el ruido. Tal y como dice Morduchowicz (2018, p. 5):

La enorme cantidad de información disponible en Internet no trae consigo todas las respuestas. Por el contrario, genera nuevas preguntas. Y estas son más complicadas. Surgen interrogantes que ya no se responden con un dato o con una frase, sino que necesitan reflexión, pensamiento crítico y análisis. En el siglo XXI, más información no significa necesariamente mejores decisiones.

Así pues, cuando un estudiante del idioma chino entra en las redes sociales con la intención de aprender más, muchas veces, para mirar si esos términos aprendidos son correctos, recurrirá a una herramienta universal: el traductor. Dentro de las clases de traducción encontramos la pedagógica, que es cuando los alumnos traducen mental y automáticamente a su lengua nativa un término u oración para poder comprenderlo. Dicho esto, en el caso de la traducción de términos coloquiales chinos, muchas veces el alumno no conseguirá obtener una traducción mental rápida, a causa del nivel de complejidad que suponen, ya que muchas veces no tienen ningún sentido literal. Por lo tanto, recurrirá a los traductores automáticos como Traductor de Google o DeepL o, incluso a inteligencias artificiales, como ChatGPT o Gemini para hacerse una idea de cómo podría usar este término en su vida diaria y de cómo podría traducirse en su idioma nativo.

El problema es que el lenguaje coloquial es muy cambiante y reciente, por lo que, a menudo, al no ser unas herramientas que están actualizadas con tanta frecuencia, los traductores automáticos no los conocen y, por ende, los traducen incorrectamente o de manera poco precisa. Por lo tanto, es justamente esto en lo que se va a centrar esta investigación: en indagar, a partir de una comparativa entre las distintas herramientas de traducción automática, la traducción que ofrecen de estos términos coloquiales y si esto pudiera ser clave para separar y valorar aún más el trabajo de un traductor humano de uno automático.

Justificación del tema

El consumo de las tecnologías es cada vez más y más elevado: los estudiantes ahora recurren al ordenador portátil para tomar apuntes en la universidad y se usa antes la aplicación de Google Maps que un mapa en físico para pasear por una ciudad desconocida. Por ende, los sectores de trabajo se ven forzados a cambiar y a adaptarse a esta nueva realidad, y el mundo de la traducción no es ninguna excepción.

La traducción automática es una herramienta que ya lleva muchos años en el mercado, y en algunos casos es incluso recomendable usarla, asegurándose siempre de recurrir a la postedición ulteriormente. A pesar de ello, hay mucha discrepancia en cuanto a su uso, pues muchos profesionales están a favor y muchos otros en contra. Hacer uso de una máquina traductora no es siempre ser un mal experto, sino saber usar los recursos que uno tiene a su alcance para su beneficio. Esta puede ser también una herramienta para el traductor humano, un puente para facilitar la comprensión de un texto original cuando

uno no sabe por dónde empezar, aun así, evidentemente, esto no funciona en todos los casos. Y es aquí donde entra la investigación pertinente.

El chino coloquial, así como cualquier otra lengua, es un idioma vivo, todavía más que la lengua en sí, pues aunque un idioma esté siempre en construcción y sometido a cambios, el idiolecto de los jóvenes es aún más variable e inconstante. Estos términos coloquiales son considerados neologismos, es decir, tal y como define el *Diccionario de la lengua española* (s.f.), «un vocablo, acepción o giro nuevo en una lengua». La mayoría de las veces nace de la adopción y adaptación de palabras extranjeras, o de la creación de acepciones de palabras ya existentes, y no suelen quedar recogidos en obras de referencia hasta pasado un tiempo. Un ejemplo claro para mostrar el origen peculiar de las expresiones coloquiales es la expresión 芭比 Q (en pinyin: *babiQ*), del inglés *barbecue*, que se sacó de un vídeo en directo en el que el presentador empezó a usar esta palabra cada vez que su personaje se quemaba en el juego. Después de esto, los internautas empezaron a usar esta expresión para indicar que «estaban acabados», al igual que el personaje del presentador, que iba a morir porque se quemaba¹.

Así pues, como estudiante de chino, me pareció muy interesante explorar el mundo del chino coloquial, y como estudiante de traducción, consideré que era esencial relacionarlo con la traducción, y qué mejor que tratar a su vez traducciones automáticas neuronales, como sería el caso de Google Translate y DeepL, o las traducciones propuestas por herramientas de inteligencia artificial, como sería el caso de Gemini y Chat GPT-3.5 y 4. De esta manera, se podrá saber si son capaces de traducir los términos, si podrán ayudar a un estudiante que tan solo quiere saber qué significa en su idioma y si, a la vez, podrían ser una ayuda para un traductor cuando este ya no sabe dónde más encontrar información para traducirlo.

Objetivos del trabajo

Para llevar a cabo esta investigación, es necesario examinar detenidamente los objetivos que se pretenden alcanzar con ella y que constituyen los fundamentos principales que han impulsado la realización de este trabajo. Estos se presentan a continuación:

- Adquirir conocimientos más profundos acerca de la traducción automática neuronal y la traducción mediante inteligencia artificial.

¹ Midi. [@midiforreal]. (2022, 17 de agosto). Just in case u saw these in Chinese social media apps #learningisfun #chineseslang [TikTok]. TikTok. Consultado noviembre 02, 2023, desde https://www.tiktok.com/@midiforreal/video/7132720070060510465?_t=8gmV5FHnVoW&_r=1

- Determinar la utilidad de estas herramientas como recurso de apoyo para estudiantes del idioma chino y profesionales de la traducción.
- Examinar en qué medida estas herramientas —caracterizadas por su constante cambio y evolución— pueden traducir de forma correcta neologismos en chino coloquial, también en cambio constante.
- Realizar una comparación entre cinco herramientas conocidas de traducción automática neuronal y de inteligencia artificial, siendo cuatro de ellas accesibles al público de manera gratuita.

Estructura del trabajo

Para alcanzar los objetivos del trabajo, primero se presentará el marco teórico, que se basa en dos pilares muy marcados y claves para este estudio: por un lado la traducción automática e inteligencia artificial aplicadas a la traducción, donde se esclarecerán los distintos tipos de TA y se indagará de forma breve en los procesos de evolución de esta. Y la descripción de la IA y los chatbots, con una exposición de opiniones a favor y en contra del uso de la inteligencia artificial en el campo de la traducción y la literatura. Por el otro lado, se encontrarán los neologismos en el chino coloquial, que sería la explicación del objeto de estudio. En este punto se describirán los neologismos, en especial los de la jerga juvenil, la formación de estos, etc. Asimismo se tratará el tema del aprendizaje a través del móvil, con una exposición de las ventajas y desventajas que conlleva.

A continuación se describirá el enfoque metodológico de esta investigación, el cual es seguido por la comparación y discusión de los resultados, expuesto todo a partir de gráficos y tablas de Excel. En este apartado se expondrán algunos de los ejemplos de traducción hechos por los motores de traducción automática e inteligencia artificial, así como las propuestas de traducción de la misma autora del TFG. Finalmente se encontrarán las conclusiones de este estudio junto con las referencias y anexos adjuntos.

Es importante remarcar que, a la hora de citar y referenciar autores chinos, para evitar confusiones por homofonía, se ha mantenido tanto el nombre como el apellido, a la vez que se ha respetado el orden usado por los mismos autores en sus respectivos estudios.

PRIMERA PARTE: MARCO TEÓRICO

En este capítulo, en primer lugar se estudiará e investigará diversos trabajos, artículos u otros documentos académicos relacionados con la traducción automática (TA) y la inteligencia artificial (IA) aplicada a este sector. Se podrá ver una descripción a grandes rasgos de los distintos tipos de TA y sus ventajas e inconvenientes. También se abordará de forma superficial la inteligencia artificial y los «chatbots». Al no ser el objeto de estudio principal de este TFG, se tratará este tema desde un enfoque distinto y no tan teórico al presentar el punto de vista de distintos lingüistas, escritores y expertos de la lengua.

A su vez, se expondrá un breve análisis de los estudios previos identificados que relacionan las herramientas de TA e IA aplicada a la traducción con los objetos de estudio de esta investigación, es decir, el chino y el español. En algunos casos, también se ha contado con trabajos en los que se investiga el idioma chino junto con el inglés como lengua pivote para obtener un análisis más amplio.

En segundo lugar, se tratará la relación que hay entre el chino coloquial y las redes sociales. Para ello se indagará sobre el uso y aparición de estos términos informales, también conocidos como neologismos; y se averiguará por qué es tan común que estudiantes y no nativos del idioma chino usen estas plataformas digitales para aprender la jerga de los jóvenes chinos.

Toda esta investigación y análisis ayudará a entender con más profundidad la parte práctica de este trabajo final de grado, en el que se comparan y analizan las traducciones de neologismos no estandarizados del chino coloquial que ofrecen un conjunto de herramientas de TA e IA, como sería Google Traductor, DeepL, Bard/Gemini y, finalmente, Chat GPT-3.5 y 4.

1. La traducción automática e inteligencia artificial aplicadas a la traducción

1.1. La traducción automática

Una definición apropiada de traducción automática es la que escribieron Ginestí y Forcada-Zubizarreta (2009, p. 43) en su estudio, que a pesar de publicarse hace más de diez años, aún prevalece y se adecúa a la TA actual:

La traducción automática es el proceso de traducción mediante un sistema informático, de textos informatizados escritos en la lengua origen a textos informatizados escritos en

la lengua meta. Un texto informatizado es un fichero de un ordenador que contiene un texto en un formato conocido.

Es decir, la TA es un procedimiento traductológico que utiliza un sistema informático para producir una traducción de un idioma a otro. A pesar de esto, hay un problema básico con esta herramienta de traducción: la ambigüedad. Y es que, tal y como explica Ginestí y Forcada-Zubizarreta (2009), los idiomas son muy complejos, y existen muchas interpretaciones a nivel léxico o sintáctico que se deben tener en cuenta al traducir. Hay muchos géneros traductológicos que se encuentran con este problema, como por ejemplo en el mundo jurídico, puesto que un mismo concepto probablemente tenga una equivalencia diferente en los distintos sistemas judiciales (Briva-Iglesias, 2021). Otro caso es el de los términos coloquiales que ofrece la lengua. Un traductor automático, en muchos casos, traducirá el término según su traducción acuñada porque no entenderá de qué se trata, ya que mientras que un traductor humano podría usar el contexto y sus conocimientos previos para entender la ambigüedad, un motor de TA tiene aún un recorrido muy largo para poder llegar a seleccionar la opción correcta en casos como estos. Es por eso por lo que actualmente, estos programas de TA tan solo pueden realizar pretraducciones que, en general, luego deben ser revisadas y posteditadas por una persona, basándose en la función, intención, destinatario, etc. (Zhang Xinyu, 2017).

Dicho esto, sí que es cierto que en muchos otros tipos de traducción, como los manuales de instrucciones o textos económicos, sí que es posible recurrir a la TA y confiar en ella sin tener que editar mucho el resultado final. Eso es porque son textos muy parecidos entre ellos que deben seguir una estructura de la oración particular y una terminología muy concreta que no tiende a la ambigüedad, es decir, que no suele tener dobles sentidos.

1.1.1. Tipos de traducción automática

La traducción automática puede dividirse en dos grandes grupos: la basada en reglas y la basada en corpus, que tiene dos tipos principales, la TA estadística y la neuronal.

En primer lugar encontramos el método basado en reglas (TABR), conocido en inglés como *Rule Based Machine Translation* (RBMT). Forcada-Zubizarreta (2022) expone que la primera traducción automática se basó en este sistema y, por lo tanto, fue el enfoque dominante desde su aparición en los años cincuenta hasta los noventa. Consiste

en traducir palabra por palabra para después añadirle reglas que incluyan y tengan en cuenta toda la oración.

Así pues, para hacer el sistema de TABR, encontramos unos expertos que primero deben recopilar los diccionarios de forma electrónica y, a continuación, escribir una serie de reglas para que la máquina analice el texto original (TO), copie las estructuras de la lengua de origen (LO) y las pase a sus estructuras equivalentes en la lengua meta (LM). Después de esto, otros profesionales del campo de la informática desarrollarán programas llamados «motores de TA», los cuales se encargarán de consultar esos diccionarios y aplicar esas reglas formalizadas con anterioridad para traducir las oraciones o los términos que se le pidan (Forcada-Zubizarreta, 2022).

La traducción automática basada en reglas carece de fluidez y naturalidad, y suele destacar por no saber detectar ni resolver ambigüedades tanto léxicas como sintácticas y estructurales. Además, «la personalización de la TABR para obtener un sistema de propósito específico o ajustado a la tarea es bastante costosa, puesto que requiere el trabajo de expertos que editen los diccionarios y las reglas» (Forcada-Zubizarreta, 2022).

En segundo lugar encontramos los sistemas basados en corpus, divididos en dos variantes de TA:

Por un lado está la traducción automática estadística, en inglés *Statistical Machine Translation* (SMT), y fue el primer planteamiento basado en corpus y el más utilizado hasta 2015. Consiste en programar los sistemas de tal modo que «aprenden a traducir», tal y como explican Ginestí y Forcada-Zubizarreta (2009). Estos sistemas aparecieron a finales del siglo XX, cuando empezaron a mejorar los procesadores informáticos y el Internet afloró junto con la popularización del ordenador y, más adelante, el teléfono móvil. Junto con esto se crearon un conjunto de corpus bilingües y monolingües disponibles en línea, hecho que aprovecharon los expertos en la materia para crear nuevos sistemas alejados de las reglas lingüísticas, es decir, del basado en reglas que había hasta ahora. Aparecieron unos sistemas basados en algoritmos matemáticos que conseguían un conocimiento profundo de traducción, lo cual les permitía traducir automáticamente basándose en este corpus ya mencionado que actúa como una memoria de traducción de grandes dimensiones y que incluía millones de oraciones equivalentes en la lengua original y lengua meta (Hearne y Way, 2011).

Forcada-Zubizarreta (2022) afirma que este sistema de traducción automática fue útil y práctico en el instante de su creación, pero actualmente, comparado con los modelos neuronales más actuales, presenta ciertas carencias significativas. La principal es que a

pesar de parecer que traduce de forma muy fluida, puede llegar a ser poco fiel al original al incluir términos adicionales innecesarios u omitir palabras esenciales. Además, la SMT se basa en modelos estadísticos y de probabilidades que tienen como base patrones, por lo que son incapaces de ofrecer nuevos términos de vocabulario fuera de los que rige el corpus, lo que podría facilitar una traducción más rica y compleja. Así pues, sus resultados serán de menor calidad y mucho más inflexibles y restringidos.

Por el otro lado encontramos la traducción automática neuronal, conocida por las siglas NMT, que viene del inglés *Neural Machine Translation*. Esta herramienta dispone de corpus de entrenamiento que incluyen cientos de miles e incluso millones de textos en LO, junto con sus respectivas traducciones en LM (Forcada-Zubizarreta, 2017), y recibe el nombre de neuronal porque consiste en redes neuronales artificiales que se inspiran de forma vaga en la forma en que el cerebro aprende y generaliza, de manera que estas aprenden y generalizan a partir de la observación de los corpus bilingües ya mencionados (Forcada-Zubizarreta, 2017, Casacuberta-Nolla y Peris-Abril, 2017).

«En las redes neuronales artificiales, las neuronas suelen formar *capas*, es decir, grupos de neuronas que reciben señales solo de las neuronas de la capa anterior y envían señales solo a las neuronas de la capa siguiente» (Forcada-Zubizarreta, 2022, p. 12). Este mismo autor manifiesta que se conoce como *aprendizaje profundo*, o *deep learning* en inglés, cuando se juntan una red de neuronas artificiales con otras con la intención de crear un conjunto de capas que ofrecerán una capacidad de aprendizaje mucho más profunda.

Tal y como argumentan Casacuberta-Nolla y Peris-Abril (2017, p. 68), esta alternativa está ofreciendo unos resultados mucho más efectivos y precisos que las otras TA. Esto es porque los términos y las oraciones se representan numéricamente por vectores, al contrario que con las herramientas anteriores, que tenían una representación discreta. Así pues, al ser mostrados de esta forma, es posible usar técnicas mucho más potentes de aprendizaje automático, como es el caso de las redes neuronales. A pesar de todo esto, Forcada-Zubizarreta (2022) expone tres problemas que puede suponer este sistema de TA:

- 1) Es mucho más complicado de rastrear de dónde viene cada término en una oración, puesto que la descodificación parte de representaciones de la oración completa; en el caso de la traducción automática estadística, es fácil hacerlo porque sus traducciones son mucho más rígidas, por lo que siguen al pie de la letra lo que pone en el corpus.

- 2) Es posible que cree textos gramaticalmente correctos que no son fieles al texto original. Esto también pasa en la estadística, pero es más probable verlo en la neuronal por culpa del tercer problema.
- 3) Ofrece muchos más errores semánticos, puesto que si desconoce un término, busca otro con significado similar. Pero si esta técnica no le da resultados, opta por segmentar en unidades subléxicas, lo que puede producir términos inventados. Un ejemplo utilizado por el autor es *ingenieraje*, que viene de *ingeniería* o *reclutación*, que viene de *reclutamiento*.

Teniendo en cuenta esto, es importante que, a la hora de posteditar, el traductor o corrector permanezca muy alerta, ya que los productos que ofrecen estos sistemas son mucho más refinados y los errores serán notablemente más sutiles.

Hoy en día la mayoría de los sistemas de traducción en línea públicamente disponibles usan el motor de TA neuronal: Deepl lleva desde 2017 usando la NMT como motor principal para traducir, y Google empezó en 2016 a usar su propia tecnología basada en esta herramienta, la cual nombraron *Google Neural Machine Translation*, y tal y como explica Wikipedia, usa el aprendizaje profundo para ofrecer traducciones más optimas en los distintos idiomas.

1.2. La inteligencia artificial

La inteligencia artificial, conocida también por las siglas «IA», consiste en el área de estudio que se dedica a desarrollar programas informáticos capaces de realizar funciones parecidas a las del ser humano, como el razonamiento, el aprendizaje o procesamiento del lenguaje (Real Academia Española, 2024, definición 1).

Dentro de este campo tan amplio, se deben mencionar los «chatbots», unos programas informáticos que simulan conversaciones humanas para poder resolver las consultas de los usuarios (Gómez, 2021). Estas aplicaciones son consideradas *Large Language Models* (LLM), puesto que a partir de neuronas artificiales y algoritmos de lenguaje profundo tienen la capacidad de analizar, entender y generar un texto en lenguaje humano (Flores-Borjabad, 2023). Generalmente, esta clase de aplicaciones también usan la tecnología de Generación de Lenguaje Natural («GLN», también conocido en inglés como *Natural Language Generation* o «NLG»), lo que les permite dar una respuesta satisfactoria al usuario mediante cuadros de diálogo (Gómez, 2021).

En este estudio, es interesante destacar dos chatbots conocidos y usados a gran escala: ChatGPT y Google Bard/Gemini. Estas dos aplicaciones son importantes porque más adelante serán el objeto de estudio y análisis junto con otras herramientas de traducción automática neuronal. Aun así, es importante mencionar que la investigación no profundiza mucho en su funcionamiento, ya que queda fuera del alcance del trabajo.

La popularización de estas inteligencias artificiales en los últimos años y su acceso generalizado ha producido una gran discusión sobre las ventajas y desventajas que esto conlleva. Esta investigación trata sobre la traducción, por lo que a continuación se podrán ver una serie de opiniones respecto al uso de la inteligencia artificial en el campo de la traducción y la literatura.

Francesca Novajra (2023), presidenta de la asociación CEATL (*Conseil Européen des Associations de Traducteurs Littéraires*), asegura que las máquinas no traducen, sino que simplemente generan textos a partir de materiales anteriores, por lo que usarlas como herramientas de traducción en obras literarias tan solo provocaría una desmejora del texto, lo que, a su vez, empobrecería el lenguaje y la cultura en general. Esto es porque, al fin y al cabo, la IA siempre se decantará por la primera opción obtenida, en vez de buscar otros sinónimos para enriquecer el texto. Además, Novjara afirma que cada género literario y cada lenguaje es merecedor de una traducción humana, por lo que es muy injusto que algunos textos, por el hecho de ser más fáciles de traducir a través de máquinas, tengan que ser productos de ello y se dejen atrás.

Hansen (2023) aborda un problema muy común en los chatbots: la distorsión de la información. Y es que un gran inconveniente que suponen los LLM es que a veces omiten o alteran el significado de un texto. Muchas veces se habla de «alucinaciones», puesto que estas herramientas de IA son capaces de generar información poco fiable y sin sentido pero propuesta de tal modo que parece verosímil (Ziwei Ji *et al.*, 2023). Esto puede crear grandes problemas en muchos campos de estudio, entre ellos la medicina, donde un error así puede poner en riesgo la vida de un paciente. Además, Hansen (2023) afirma que estos errores son mucho más complicados de ver, sobre todo entre personas que no han sido entrenadas para ello.

Strojin (2023) afirma que se ha conseguido producir más datos en estos dos últimos años que desde el inicio de la humanidad. Es por eso por lo que la IA es importante, puesto que tiene un gran número de información, a pesar de que mucha de ella sea improcedente. Pero para conseguir más datos, se necesita más información, lo que incluye obras, como novelas literarias u otras propiedades intelectuales.

Así pues, otro problema que supone la IA es la violación de los derechos de autor. Desde su aparición, ha surgido una gran preocupación por ello y por el hecho de si, por culpa de esto, podrían reducirse los ingresos de los artistas. Ha habido incluso autores, diseñadores gráficos y programadores que han iniciado demandas a OpenAI (los creadores de ChatGPT) y a otras empresas relacionadas por infringir normas de derechos de autor y propiedad intelectual.

El estudio de McGee (2023) trata los dos problemas mencionados anteriormente. En su artículo, el doctor propuso una serie de preguntas a ChatGPT para ver si conseguía obtener el capítulo entero de una de sus novelas. Más adelante, le pidió más información sobre esta, como el género, argumento, etc. La conclusión del estudio fue que, por ahora, los autores no debían temer a la IA, puesto que las respuestas fueron incorrectas y poco resolutivas. A pesar de ello, no significa que en un futuro no pueda cambiar, por lo que es necesario hacer más estudios para poder llegar a una conclusión definitiva.

Dicho esto, Ganascia (2017) en su libro *Le mythe de la Singularité. Faut-il craindre l'intelligence artificielle?* (que podría traducirse en español como «El mito de la singularidad: ¿se debería temer a la inteligencia artificial?») afirma que no hay comparación entre la inteligencia artificial y la humana, y refuta de forma clara la «singularidad tecnológica», es decir, que en un futuro las máquinas puedan controlar y superar a la humanidad.

1.3. Estudios previos en el caso del chino-español

Los idiomas referenciales de esta obra son el chino y el español, dos de las lenguas más habladas en el mundo², aun así, sorprendentemente, el número de investigaciones sobre la traducción automática con ambos idiomas como lengua pivote no es muy elevado, y muchos usan el inglés como intermediario para facilitar el análisis.

Zhi Chen (2020), que indaga en las opciones de traducción propuestas por *Google Neural Machine Translation* (GNMT) en cuanto a expresiones metafóricas y metonímicas, afirma que cuanto mayor es el uso, mayor es la calidad de las traducciones ofrecidas por la TA. Un ejemplo de ello es «estar en cabeza», en chino 领先 (*ling xian*). El sintagma en español tiene una frecuencia de uso en CORPES muy elevada, por lo que al traducirlo con GNMT, el resultado obtenido ha sido correcto en ambos casos, tanto con el sintagma solo como con una oración entera de contexto. Además, Zhi Chen también

² Fuente: www.ethnologue.org

declara que, al tratarse de una herramienta de aprendizaje profundo, la traducción será más precisa si se le ofrece un mayor contexto, es decir, una oración entera. El diagrama número 7 de su investigación corrobora este comportamiento del GNMT: al ofrecer una oración entera, el porcentaje de corrección es de un 39,90 %, mientras que al ser solo el sintagma, el porcentaje baja a un 15,38 %.

Costa-Jussà et al. (2011) estudiaron el uso del inglés como lengua pivote en la traducción chino-castellano. A través de una comparación entre el sistema directo (ZH-ES) y el pivote (ZH-EN, EN-ES), concluyeron que este último ofrecía mejores traducciones, sobre todo si se usaba el sistema basado en cascada, que tal y como explican los autores «consiste en obtener la traducción de chino-castellano a través de la concatenación de traducciones de chino-inglés e inglés-castellano» (p. 121). Así pues, declaran que el inglés podría ser una lengua óptima como intermediación de la traducción entre el chino y el castellano.

Otra investigación interesante respecto al uso de herramientas de TA para la traducción chino-español es la de Piqueras (2020). La autora afirma que sería muy positivo para el traductor generar un corpus propio bilingüe y monolingüe de un mismo campo temático, ya que a la larga mejoraría y optimizaría su rendimiento laboral. También trata un tema problemático con relación a esto: la segmentación del chino. El idioma mandarín carece de separación de términos, además, es una lengua ideográfica, por lo que un carácter tiene un significado distinto por sí mismo que junto a otro carácter. Dicho esto, a una máquina le planteará una mayor dificultad determinar qué caracteres van juntos o separados, puesto que todo dependerá del contexto y de los conocimientos lingüísticos que sí tiene un traductor humano. Esto determina que «las diferencias lingüísticas de los pares de lenguas con los que trabajamos sí que pueden afectar a los resultados de la traducción» (p. 63).

Así pues, a pesar de que aún sea recurrente el uso del inglés como lengua pivote, poco a poco las herramientas de traducción automática e inteligencia artificial van perfeccionándose para obtener resultados más precisos. Si bien no es recomendable utilizar estas herramientas para traducir textos completos y reemplazar el trabajo del traductor humano, sí son útiles como herramientas de apoyo en la labor del traductor.

2. El chino coloquial: su relación con las redes sociales

Tal y como explica Hernández Alonso (1991, p. 12), «la expresión popular coloquial es la manera de comunicarnos más espontánea, vital, ágil, expresiva e imaginativa de la lengua y, además, es la que con mayor frecuencia utilizamos». A lo largo del día, un joven suele relacionarse con amigos o compañeros suyos del mismo rango de edad, y esto, con el tiempo, hace que se inflencie del léxico usado durante sus conversaciones. De esta forma se crea y extiende la lengua juvenil, «una de las principales manifestaciones de un importante cambio generacional» (Hernández Alonso, 1991, p. 16), y es que cada generación crea su propio léxico asociado con los cambios sociales de esa época: en los años ochenta, el punk, rock y los vídeos de aeróbic hicieron que naciera una jerga juvenil concreta; mientras que, en la actualidad, las nuevas tecnologías y las redes sociales han hecho que nazca otra muy distinta que consigue extenderse aún más rápido gracias a ello.

Curiosamente, además, cuando estos jóvenes que utilizan tal argot se encuentran en ambientes distintos (trabajo, Universidad, administración, etc.), utilizan automáticamente la norma estándar, salvo en algún *tic* lingüístico ya generalizado y admitido. Esto nos dice que ellos mismos tienen conciencia de que su expresión específicamente juvenil no es más que una variante admitida en determinados ambientes y ocasiones. Se trata, pues, con todas las reservas necesarias, de una situación de "diglosia" y no de bilingüismo, que sería su objetivo y anhelo. (Hernández Alonso, 1991, p. 20).

Tal y como explica Otero (2018), el vocabulario de los jóvenes y adolescentes se crea, en mayor parte, a partir de los neologismos: palabras nuevas empleadas por un grupo social o comunidad de habla. Dicho esto, en el diccionario de la Real Academia Española se puede encontrar cómo se forman. Hay veces que se crean a partir de palabras ya existentes que luego adquieren un nuevo significado; mientras que en otros casos o bien se inventan o bien se crean por derivación o composición. Es evidente que estos términos no siempre son aceptados por la institución principal que regula el uso de la lengua, puesto que son expresiones muy nuevas y que cambian con frecuencia. Por ende, esto puede suponer un problema al traducir, ya que en muchas ocasiones uno no sabe qué término usar, ya que son conceptos no estandarizados que pueden diferir dependiendo de dónde se busque. Esto, a su vez, también complicará los resultados de las herramientas de traducción automática o inteligencia artificial, puesto que en muchos casos estos no encontrarán

ningún resultado en su base de datos, por lo que se decantarán por la acepción estandarizada, que no suele ser la misma que el significado nuevo del neologismo.

2.1. La aparición de neologismos en las redes sociales

Si nos centramos en el estudio principal de esta investigación, es decir, el chino coloquial, es muy acertado hablar de neologismos chinos de Internet. Tal y como afirma Tao Yingnian (2017), estas nuevas expresiones muestran unas connotaciones sociales muy arraigadas a la cultura china y a la juvenil, es por eso por lo que es interesante analizarlas e investigarlas en profundidad. La misma autora les atribuye las siguientes categorías: arbitrariedad, velocidad de transmisión y efectividad a corto plazo, ya que estos neologismos suelen crearse a partir de sucesos casuales y actuales de la cultura popular y, actualmente, con el uso del Internet estos aparecen tan rápido como desaparecen, puesto que se difunden a una velocidad sin precedentes. Además, cada rango de edad suele inventar los suyos propios e ir creando y dejando atrás los menos usados con el paso de los años y la aparición de nuevas tendencias de Internet.

Finalmente, es importante conocer las formaciones que suelen tener estos neologismos. En muchos casos se mezclan números, caracteres chinos y letras en inglés, como sería el ejemplo de 栓 Q, en que se usa la letra Q para conseguir la fonética de *thank you* a partir de la pronunciación del primer carácter en chino y la letra: *shuan qyu*. También hay expresiones formadas por dos o tres caracteres, como por ejemplo 辣眼睛 (*la yanjing*, traducido literalmente como «ojos picantes», aunque la traducción no estandarizada sería «algo feo que no deberías ver»). Otras veces se recogen nuevas definiciones de términos que ya tienen una acepción, un ejemplo de ello es 炒鱿鱼 (*chao youyu*, traducido literalmente como «calamares salteados», aunque la traducción no estandarizada sería «estar despedido»), aunque el término en sí es el nombre de un plato de calamares salteados. Por último se encuentran las expresiones metafóricas, que suelen abundar mucho en la creación de neologismos. El ejemplo perfecto de ello es 菜 (*cai*), que significa verdura, aunque se usa para referirse a una persona se le da muy mal alguna cosa (Tao Yingnian, 2017).

2.2. Las redes sociales como plataforma para aprender jerga coloquial

Hasta el inicio del siglo XXI, en muchas ocasiones, las expresiones coloquiales eran desconocidas para los estudiantes de un idioma extranjero, puesto que la única manera de conocerlas y aprenderlas era hablando, relacionándose con nativos o visitando ese país. A pesar de ello, con la aparición de Internet, ya no es tan difícil investigar sobre las palabras que están de moda entre los jóvenes hoy en día, puesto que las redes sociales sirven como puente entre el estudiante y el país extranjero.

Instagram o Tik Tok, entre otras plataformas, ofrecen la posibilidad de conocer y aprender términos o expresiones coloquiales de otros idiomas gracias a publicaciones o vídeos cortos que cuelgan los usuarios en la misma plataforma. Normalmente, las personas detrás de estas cuentas suelen ser profesores de lengua, a veces nativos y otras veces no.

En este contexto aparece el concepto de *microaprendizaje*, que describe muy bien el estudio por redes sociales. Bruck (2006, p.15), citado en la investigación de Rovira-Esteva y Vargas-Urpí (2023, p. 3), explica esta idea como una manera de ofrecer al estudiante la posibilidad de aprender a partir de una información reducida, tanto en cantidad como en complejidad, así se alivia la presión del alumno y se evita una sobrecarga informativa. De esta forma, también se ofrece la oportunidad de decidir el momento y ritmo de estudio.

Es evidente que aprender un idioma con tan solo el uso de las redes sociales es casi imposible y poco recomendado, puesto que no podrá satisfacer todas las necesidades que requiere el ser fluido en una nueva lengua. No obstante, sí que puede ser de ayuda como ejercicio suplementario a las clases, en muchos casos, para aprender términos que no se suelen enseñar allí, como la jerga de Internet.

Así pues, a continuación se listará una serie de ventajas y desventajas de esta nueva moda, puesto que si la plataforma se usara de forma adecuada, podría suponer una gran ayuda al estudiante con ganas de ampliar su conocimiento lingüístico. Es cierto que los artículos listados a continuación se centran en las redes sociales como puente de estudio de una lengua, pero también incumbe a aquellos que las usan para solo aprender términos coloquiales.

Por una parte, una de las razones por las que plataformas como Tik Tok han ido popularizándose con el tiempo es porque son muy cómodas. El formato de los vídeos,

cómo se organizan y cómo van saliendo en la pantalla del teléfono móvil con tan solo deslizar hace que uno quiera seguir mirando, es decir, engancha a los usuarios a seguir consumiendo contenido. Chuxuan Duan (2023) afirma que hoy en día muchas personas recurren a estas plataformas para aprender idiomas porque lo encuentran un ambiente más agradable, dicho de otra manera, aprecian la informalidad y poca seriedad que no suele haber en un aula.

Además, el foco de estudio y aprendizaje varía dependiendo de la cuenta que uno sigue. Mientras que un profesor te enseña cómo diferenciar los distintos acentos de una lengua, otro te enseña vocabulario o trata puntos de gramática; de esta manera, las publicaciones distintas permiten un rango mucho más amplio de espectadores, ya que se encuentran al alcance del gusto de cualquiera. Pese a ello, estos profesores virtuales suelen seguir un estilo parecido en todas sus publicaciones y suelen centrarse en enseñar una parte específica de la lengua (vocabulario y explicaciones de gramática, tan solo comprensión oral, comprensión escrita y también vocabulario, etc.). Así, el algoritmo de las redes sociales consigue dividir a los usuarios teniendo en cuenta los gustos de cada uno. De esta manera, las cuentas tienen una audiencia concreta que espera ver a menudo vídeos o imágenes de un estilo determinado o con un contenido delimitado. Hay material muy variado, pero normalmente se pueden hallar pequeños diálogos más o menos cómicos para que el estudiante pueda entender mejor el contenido y tenga un mayor contexto, algo fundamental para el aprendizaje a largo plazo. Otras veces se utilizan también fragmentos de programas de televisión o películas en el idioma de aprendizaje para entrenar la comprensión oral con situaciones más realistas y auténticas (Chuxuan Duan, 2023). En general, muchas de estas cuentas suelen incluir pequeños apartados donde tratan términos coloquiales, mientras que otras simplemente introducen las expresiones junto con sus publicaciones diarias de vocabulario general.

Asimismo, el uso de las redes sociales para aprender idiomas permite al usuario crear una estimulación multisensorial. Tal y como afirma Zhang Jingshuo (2022) en su TFM sobre el aprendizaje de chino a través de podcasts, gracias a las aplicaciones móviles como Tik Tok o Instagram es posible que los estudiantes extranjeros entren en contacto con la cultura y el país de la lengua que estudian sin haber visitado el sitio. De esta manera, en el caso del chino, uno puede «sumergirse y absorber la cultura china desde la distancia» (p. 16), y adoptar vocabulario de Internet casi sin darse cuenta.

Por otra parte, según los participantes del estudio de Chuxuan Duan (2023), que usan Tik Tok como herramienta de estudio, las mayores desventajas recaen en el aprendizaje fragmentado y las limitaciones que pueden aparecer al usar estas plataformas.

En general, no hay forma de garantizar que los profesores virtuales tengan títulos de enseñanza de lenguas, puesto que, aunque cada vez sea más común poner en la descripción de la cuenta los años que uno lleva enseñando, o los títulos profesionales de los que dispone, nada te asegura que lo que se haya escrito sea cierto, así pues, al final, el usuario tan solo puede confiar en su propio criterio.

Además, el contenido desordenado con mucho vocabulario nuevo y distintas normas que varían de profesor en profesor también pueden llegar a abrumar al estudiante que acaba de introducirse en el idioma (Figura y Jarvis, 2007). Al fin y al cabo, es muy distinto aprender en línea que en un aula presencial, puesto que los profesores mirarán de estructurar la materia para simplificar el aprendizaje, mientras que la orientación que se conseguirá por Internet será nula en comparación. Dicho esto, Zhang Jingshuo (2022) declara que los podcasts, por ejemplo, sí que pueden ser unos recursos útiles para el estudiante si lo han proporcionado los mismos profesores de chino, ya que garantizarán que el nivel y el contenido sea el adecuado a las necesidades del alumno.

Chuxuan Duan (2023) también pone en la mesa un argumento muy interesante. En su investigación afirma que el aprendizaje a través de las redes sociales no es para todos los estudiantes, y es que al parecer, para estudiantes de niveles básicos, las publicaciones, ya sean imágenes o vídeos cortos, son muy efectivas; mientras que para niveles más avanzados el uso de podcasts es mucho más recomendado, ya que pueden ser un buen material complementario al aprendizaje del aula, al fomentar la habilidad de comprensión y expresión oral (Zhang Jingshuo, 2022).

Así pues, a pesar de que las redes sociales ofrezcan muchas ventajas, también es importante recordar que suponen algunos inconvenientes. Y es que, al fin y al cabo, son recursos muy generales que pueden confundir al estudiante, sobre todo si está iniciándose en el idioma, puesto que no se seguirá un plan concreto adaptado e individualizado. Es por esta razón que, probablemente, las publicaciones sean de gran ayuda para estudiantes que quieren ampliar su conocimiento, sobre todo del lenguaje informal de los nativos, o mejorar alguna destreza concreta, mientras que supondrán un rompecabezas para otros que quieran adentrarse por primera vez en la lengua.

Dicho esto, es importante recordar que el aprendizaje a través del móvil crea un puente entre el mundo real y el virtual, constituye comunidades de estudiantes

entusiasmados, ofrece expertos a la carta y respalda un aprendizaje de por vida (Sharples, Milrad, Arnedillo-Sánchez y Vavoula, 2009: 2.). En palabras de Wang Jing (2012, p. 138), que diez años más tarde fueron traducidas del chino al español por Zhang Jingshuo (2022, p. 15):

El proceso de aprendizaje de la lengua es, de hecho, un proceso de intervención cultural, y a través del proceso de comunicación verbal se alcanza la conciencia cultural. El material lingüístico de la comunicación por Internet es vívido y cercano a las situaciones reales de comunicación, lo que refleja la vida social actual en China de forma más completa y aporta los patrones de expresión del habla más precisos a los estudiantes de chino. Esta formación en comunicación verbal permite contextualizar culturalmente las expresiones verbales en el entorno de Internet, lo que mejora la capacidad de los alumnos chinos para comunicarse entre culturas.

SEGUNDA PARTE: MARCO PRÁCTICO

3. Metodología

La metodología de este trabajo se basa en la realización de un estudio empírico centrado en términos coloquiales encontrados a través de distintas plataformas digitales como Tik Tok, Weibo, Bilibili y YouTube. A partir de los datos extraídos de estas redes sociales, se ha creado un corpus de chino coloquial con la terminología y sus ejemplos de uso. Con esta parte ya hecha, se ha iniciado el segundo paso de este largo proceso de comparación, que consiste en traducir las distintas oraciones de ejemplo por los traductores automáticos de Google Traductor y DeepL, y por las dos inteligencias artificiales más usadas hoy en día del mercado, Chat GPT-3.5 y 4, y Gemini, antes conocida como Bard. Los resultados se han copiado en la tabla hecha por Excel y se han evaluado siguiendo una valoración del uno al cinco que se explicará con más profundidad a lo largo de este apartado.

Gracias a este trabajo se ha podido analizar y evaluar de forma exhaustiva si un traductor automático o una máquina de inteligencia artificial es capaz de traducir correcta y adecuadamente un término según su versión no estandarizada, siempre teniendo en cuenta que son expresiones muy nuevas y actualizadas por lo que es posible que por falta de actualización, estas herramientas no ofrezcan la respuesta que se busca.

Así pues, en este capítulo se ha descrito en profundidad cómo se ha llevado a cabo el proceso práctico de recogida, análisis y evaluación de datos.

3.1. Método de recogida de datos

En primer lugar, se ha hecho una breve investigación para descubrir qué cuentas de Tik Tok pueden ser fiables, ya que de allí se han sacado los datos para el análisis. La selección final consiste en siete cuentas de contenido relacionado con la enseñanza del chino: @midiforreal, @survivalchinese, @jiayouchinese, @torontomandarinschool, @mandarinmelon, @chinesegirlcoco y @chineseewithyan.

Para poder afirmar que su contenido será útil para la investigación, se han tenido en cuenta distintos parámetros mostrados en la siguiente tabla:

Cuentas	Cuenta centrada en la enseñanza del chino	Contenido relacionado con términos coloquiales	Ejemplos de uso
@midiforreal	No	Sí	Sí

@survivalchinese	Sí	Sí	Sí
@jiayouchinese	Sí	Sí	Sí
@torontomandarinschool	Sí	Sí	Sí
@mandarinmelon	Sí	Sí	No
@chinesegirlcoco	Sí	Sí	Sí
@chineseewithyan	Sí	Sí	No

Tabla 2: Parámetros de selección de las cuentas en redes sociales. [Elaboración propia].

Como se puede ver en la tabla superior, todas las cuentas menos @midiforreal se centran en la enseñanza del chino, un hecho crucial, ya que gracias a ello se puede analizar el resto del contenido y ver si la información que proporciona es lícita y correcta. La usuaria que no centra su perfil en la enseñanza del chino es nativa, y suele publicar vídeos relacionados con el idioma mandarín y China, así pues, tal bagaje de conocimiento hacia la cultura china genera confianza en sus habilidades y conocimiento de la lengua.

Asimismo, es importante que todas las cuentas incluyan vídeos explicativos de términos coloquiales. Esto es imprescindible, puesto que, si no, no sería posible llevar a cabo el vaciado de contenido. Además, también se ha observado si los profesores ofrecen ejemplos de uso al mencionar el término o expresión, y así ha sido en cinco de las siete cuentas. Estas oraciones más tarde han sido el objeto de estudio al usar las herramientas de TA e IA aplicadas a la traducción. En el caso de los términos sin oraciones de ejemplo, se ha recurrido a frases reales sacadas de redes sociales chinas como Bilibili o Weibo.

En segundo lugar, al tener ya las cuentas elegidas, se ha llevado a cabo el vaciado de términos y expresiones de los vídeos. Se han escogido treinta vocablos para tener un muestreo cuantitativo de distintos tipos de términos coloquiales: números, siglas, otros neologismos extraídos por derivación o composición, nuevas acepciones de caracteres ya estandarizados, etcétera. De esta manera, los resultados determinarán si se sigue algún patrón en cuanto a los términos coloquiales que una herramienta de traducción automática e inteligencia artificial conoce. A continuación se puede ver la tabla en cuestión con los términos, traducción³, pinyin, cuenta en que se han encontrado y fecha de publicación del vídeo de donde se han extraído.

Nombre cuenta Tik Tok	Expresión	Pinyin	Traducción	Fecha de publicación
--------------------------	-----------	--------	------------	-------------------------

³ Traducción del término como neologismo, no según la definición estándar.

@chineseewithyan	炫	xuan	Comer	28.07.2022
@chineseewithyan	干饭	ganfan	Comer una comida	28.07.2022
@chineseewithyan	栓 Q	shuanQ	Gracias	28.07.2022
@chineseewithyan	emo	emo	Estar triste	28.07.2022
@chineseewithyan	集美	jimei	Hermana / tía	28.07.2022
@chineseewithyan	修勾	xiugou	Perrito	28.07.2022
@chinesegirlcoco	牛逼	niubi	1. Ser guay 2. También en sentido irónico cuando alguien se regodea de cosas malas que le han pasado.	06.04.2022
@chinesegirlcoco	菜	cai	Ser muy malo en algo	06.04.2022
@chinesegirlcoco	土豪	tuhao	1. Alguien rico 2. Entre amigos, para reírse de algún amigo que se ha comprado alguna cosa cara	06.04.2022
@mandarinmelon	250	er bai wu	Persona idiota	29.08.2023
@mandarinmelon	1314	yi san yi si	一生一世: Toda la vida	29.08.2023
@mandarinmelon	484	si ba si	是不是: ¿Sí o no?	29.08.2023
@mandarinmelon	996	jiu jiu liu	De 9 de la mañana a 21 de la noche, 6 días a la semana (referencia cultural de trabajar mucho)	29.08.2023
@mandarinmelon	520	wu er ling	我爱你: Te quiero	11.09.2023
@mandarinmelon	2333	er san san san	哈哈: Ja ja ja	11.09.2023
@mandarinmelon	555	wu wu wu	呜呜呜: sollozar	11.09.2023

@mandarinmelon	9494	<i>jiu si jiu si</i>	就是就是: sí que lo es	11.09.2023
@mandarinmelon	JM	<i>jiemei</i>	姐妹: Hermana/ tía	26.08.2023
@mandarinmelon	OJBK	--	OK	26.08.2023
@mandarinmelon	U1S1	<i>you yi shuo yi</i>	有一说一: Para ser honestos, a decir verdad	26.08.2023
@mandarinmelon	PYQ	<i>pengyou quan</i>	朋友圈: momentos de WeChat (WhatsApp de China)	25.08.2023
@torontomandari nschool	YYDS	<i>yongyuan de shen</i>	El mejor en alguna cosa	30.03.2023
@torontomandari nschool	吃货	<i>chihuo</i>	Persona que le encanta comer. Conocido en inglés como “foodie”.	30.03.2023
@torontomandari nschool	佛系	<i>foxi</i>	Alguien sin pasiones ni ambiciones	30.03.2023
@midiforreal	芭比 Q	<i>babiQ</i>	Me voy a morir, estar acabado	17.08.2022
@midiforreal	XSWL	<i>xiaosi wo le</i>	Mearse de la risa	04.09.2021
@survivalchinese	炒鱿鱼	<i>chao youyu</i>	Estar despedido	29.07.2020
@jiayouchinese	辣眼睛	<i>la yanjing</i>	Feo, algo que no deberías ver	13.10.2023

Tabla 3: Corpus de términos de chino coloquial. [Elaboración propia].

En tercer lugar, se han buscado los ejemplos de uso de todas aquellas expresiones que no contaban ya con oración de ejemplo. Para ello se ha buscado en plataformas chinas como Weibo y Bilibili comentarios o publicaciones que contengan dichos términos. Las capturas de pantalla de las oraciones usadas se pueden encontrar en el anexo 1.

Para finalizar, después de recoger toda la jerga coloquial junto con sus oraciones de ejemplo, se ha creado una hoja de cálculo en Excel para plasmar este vaciado. Aquí se puede encontrar una tabla con los términos, contextos y fuentes, y un espacio para cada herramienta de traducción automática e inteligencia artificial para poder añadir más adelante las traducciones y evaluaciones de estas.

3.2. Selección de los sistemas de TA e IA

Para poder explicar el proceso de análisis de datos, es importante primero saber la razón por la cual se han elegido estos motores de traducción automática e inteligencia artificial concretos.

Por un lado, se han seleccionado Google Traductor y DeepL porque mientras que el traductor de Google es actualmente la máquina de TA más utilizada del mercado (Yulianto y Supriatnaningsih, 2021); DeepL pasó a ser de las más fiables y exactas ya desde su lanzamiento en 2017 (Rescigno, *et al.*, 2020).

Por el otro lado, se han escogido Chat GPT y Gemini por el repentino auge que ha surgido respecto a estos sistemas de IA desde su aparición en el mercado en los años 2022 y 2023 respectivamente. De esta manera, se amplía el número de investigaciones sobre estos modelos de lenguaje en torno a la traducción y, en especial, a la combinación de idiomas chino-español.

Teniendo presente este pretexto, a continuación se describirá de forma breve cada uno de estos motores de TA e IA:

Google Traductor es un sistema multilingüe de traducción automática lanzado en 2006 por la empresa de tecnología Google. Durante los primeros años, el motor se basaba en reglas, más tarde pasó a ser estadístico y en 2016 se implementó el sistema de TA neuronal. Uno de sus servicios más interesante y, en su momento, innovador es el de traducir textos de imágenes o incluso transcribir y traducir audios de voz (Fitria, 2023).

DeepL es un motor de traducción automática neuronal difundido en 2017 por la empresa DeepL SE. La memoria que utiliza se recoge a partir de la base de datos de Linguee y la interfaz es muy parecida a la de Google Traductor (Rescigno, *et al.*, 2020). Como ya se ha dicho, Fitria (2023) afirma que este servicio ofrece traducciones más precisas y naturales de los idiomas con los que trabaja, lo que ha hecho que se diferencie del resto. Aun así, se debe mencionar que presenta una variedad lingüística más limitada y cuenta con una versión de pago para traducir sin restricciones.

Chat GPT es una aplicación de chatbot de inteligencia artificial desarrollada en 2022 por la empresa de investigación OpenAI. Cuenta con dos modelos, el Chat GPT-3.5 y el Chat GPT-4⁴, este segundo es de pago. Es un modelo de lenguaje basado en el aprendizaje supervisado y por refuerzo, es decir, que entrenadores humanos se encargan

⁴ El 13 de mayo de 2024 salió una nueva versión llamada Chat GPT-4o, el modelo más avanzado de Chat GPT. Está disponible para todos los usuarios de forma gratuita.

de revisar y mejorar su rendimiento. En cuanto a la traducción, Kalla y Smith (2023, p. 830) afirman lo siguiente:

Chat GPT puede traducir textos en distintos idiomas, lo que lo convierte en una herramienta valiosa para la comunicación global. El sistema traduce textos de forma precisa e instantánea, ofreciendo una manera eficiente e impecable para que los usuarios se comuniquen más allá de las barreras lingüísticas.⁵

Gemini es sistema de chatbot de inteligencia artificial creado y divulgado por la empresa de Google en 2023 para competir con Chat GPT. Al ser bastante reciente, cuenta con un número muy limitado de investigaciones. En principio se llamaba Bard, pero más tarde lo renombraron como «Gemini», a la vez que lanzaban la versión de pago del producto, «Gemini Advanced», y se producían algunas mejoras significativas como una mayor capacidad de respuesta, de proceso de la información, flexibilidad, etcétera.

3.3. Método de análisis de datos

Después de elegir las herramientas adecuadas para este estudio, se ha iniciado el proceso de análisis de datos. Para poder evaluar de forma correcta y uniforme las traducciones de estos motores de traducción automática e inteligencia artificial, se ha creado una rúbrica de evaluación que puntúa del uno al cinco, siendo el número uno la nota más baja y el cinco, la más alta.

Puntuación 1-5	
1	No se ha traducido el término, se ha dejado tal y como estaba escrito en el idioma original o directamente se ha omitido.
2	Se ha intentado traducir el término, pero no tiene nada que ver con el significado real.
3	Se ha traducido el término correctamente, pero no hay ni coherencia ni cohesión con el contexto.
4	Se ha traducido el término correctamente, tiene coherencia la oración, pero se puede mejorar.

⁵ Chat GPT can translate text between different languages, making it a valuable tool for global communication. The system can accurately translate text in real time, providing a seamless and efficient way for users to communicate across language barriers. [Original en inglés, traducido por la propia autora].

5	Se ha traducido el término exitosamente, con el contexto adecuado y coherencia en la oración.
---	---

Tabla 4: Rúbrica de evaluación para las traducciones. [Elaboración propia].

Al evaluar las traducciones, tan solo se ha tenido en cuenta que el neologismo concreto estuviera bien traducido y tuviera lógica con el contexto. Esto no significa que la oración en general haya sido traducida de la forma más natural y coherente posible, puesto que en muchos casos no es así. Pero como el objeto de estudio de esta investigación es tan solo los neologismos de Internet, no se han analizado en más profundidad las traducciones, puesto que, de lo contrario, se ampliaría demasiado el análisis y estudio de este TFG.

Dicho esto, la rúbrica de evaluación ayuda a justificar cada traducción de la misma forma y así simplificar el proceso de análisis. En algunos casos, se ha añadido algún que otro comentario más explicativo y descriptivo para mostrar el proceso traductológico que han llevado a cabo los traductores automáticos o inteligencias artificiales a la hora de ofrecer una solución al problema que suponen estos términos coloquiales.

Así pues, el proceso de análisis de datos ha consistido en traducir cada una de las oraciones de ejemplo por los cuatro motores de traducción automática y sistemas de inteligencia artificial ya descritos anteriormente: DeepL, Google Traductor, Gemini y Chat GPT-3.5 y 4. En el caso de Chat GPT, tan solo se han traducido por el modelo Chat GPT-4 las oraciones que han sido evaluadas con menos de un dos. Esto es porque se quiere observar si con un modelo más actualizado y de pago, las traducciones ofrecidas son más correctas y adecuadas al contexto. Se podrá ver que el resultado sí que ha sido más satisfactorio en estas nuevas traducciones.

Toda la información obtenida a partir de estas traducciones se ha añadido en la hoja de cálculo de Excel. Menos en el caso de Gemini, que ofrecía una explicación de la traducción y tres versiones distintas, con el resto de las herramientas tan solo se ha pegado la opción en español y, al lado, se ha comentado y evaluado el resultado obtenido. En la misma tabla también se encuentra una columna con una propuesta de traducción de la propia autora de este TFG para comparar de forma breve la diferencia entre una traducción hecha por TA o IA y una hecha por un traductor humano.

4. Resultados del análisis

A lo largo del siguiente capítulo, se describirá y analizará una selección de ejemplos variados a modo de estudio cualitativo. Más adelante se mostrarán los resultados de las traducciones realizadas a partir de herramientas de traducción automática e inteligencia artificial con gráficos. Asimismo, estos se compararán y analizarán para así conseguir un análisis que sirva a futuras investigaciones de TA e IA aplicada a la traducción en contextos de términos sacados de las redes sociales.

4.1. Descripción de los ejemplos utilizados

A continuación se explicarán cinco de los veintinueve términos coloquiales usados como objeto de análisis en las traducciones por TA e IA. A su vez, se mostrará la traducción ofrecida por la misma autora de este TFG, para poder ver una traducción más acorde con el significado del neologismo, así como más natural y lógica según el contexto en el que aparecen. El resto se podrá ver en el Excel adjunto a este trabajo para poder tener una perspectiva mucho más amplia de la investigación. En este se incluye las propuestas de traducción automática, las valoraciones, los comentarios y una sugerencia de traducción de la misma autora.

El primer término es 炫 (*xuan*), un verbo que según su forma no estandarizada significa «comer», mientras que si se busca en el diccionario bilingüe 小马词典 (*Xiaoma Cidian*) aparecen tan solo las acepciones de «presumir» o «deslumbrar». La oración sacada de un vídeo de Bilibili es: 一人独享 ! 整只豉油鸡 🐔 鲜嫩爆汁快来一起炫 (véase en contexto en el anexo 1, figura 20). La propuesta de traducción de la autora del TFG es: «¡A disfrutarlo solito! Un pollo entero en salsa de soja, tierno y con un jugo muy sabroso. 🐔 Ven a probarlo conmigo».

Tal y como se puede ver, se ha optado por cambiar «comerlo» por «probarlo», puesto que por el contexto es mucho más natural invitar primero a probar un bocado. Es cierto que no sería incorrecto usar el verbo «comer», pero «probar» es un verbo mucho más común para ofrecer un trozo de tu plato a otra persona.

El segundo término es 520 (*wu er ling*), un neologismo formado por tres números que, pronunciados en chino, tienen una fonética muy parecida a 我爱你 (*wo ai ni*), que significa «te quiero». La oración de ejemplo es «给我转 520», y está sacada de una pegatina usada en los mensajes de WeChat. La propuesta de traducción de la autora del

TFG es: «Mándame un te quiero». Otra opción interesante es la siguiente: «Mándame un <3». En este caso se mantendría el juego de números, pero aplicado al español. Aunque en esta lengua no se usen tanto los neologismos con números como en chino, sí que se usa mucho esta mezcla de símbolos: «<3».

Más adelante se podrá ver que en la gran mayoría de las traducciones automáticas de neologismos con números, estos se han mantenido igual, un ejemplo de ello es la traducción de Gemini de esta pegatina: «Pásame 520». Es evidente que no se puede dejar el neologismo chino, ya que en español no tiene el mismo significado, y para un lector español tan solo serán tres números, lo que, en este caso, incluso podría tener el significado erróneo de pedir dinero. Así pues, la traducción más adecuada sería conservar la traducción del término sin el juego de números, y al ser acompañado de la imagen de la pegatina, la informalidad del mensaje se mantendría.

El tercer término es JM (*jiemei*), otro neologismo formado por las dos primeras letras de los caracteres en chino del término 姐妹 (*jiemei*), que traducido al español significaría «hermana» o «tía», como apodo cariñoso e informal dirigido a las amigas. La oración de ejemplo ha sido sacada de un comentario de un usuario en Weibo: «jm 这真的很崩溃从想主题开始然后疯狂讨论删删改改写策划废时将近三个小时这简直nt⁶».

Esta palabra es interesante porque ninguna de las traducciones automáticas ha sido capaz de distinguirla como un neologismo del ámbito coloquial y, por lo tanto, traducirla. En todos los casos, al usarse como un apelativo, se ha dejado «jm», seguramente pensando que era un nombre. La propuesta de traducción de la autora es: «Tías, esto es muy frustrante. He malgastado casi tres horas pensando el tema y discutiéndolo sin parar. He borrado, cambiado y modificado tanto... En serio, qué locura». Se ha decidido añadir el plural porque la usuaria se dirige a sus amigos o seguidores de la red, por lo que es más probable que el comentario sea para más de una persona. Asimismo, «tías» es un término de parentesco, por lo que también ayuda a mantener esta cercanía hacia el usuario al que se dirige el mensaje.

El cuarto término es 佛系 (*foxi*), un neologismo que describe a alguien sin pasiones ni ambiciones. Suele referirse a jóvenes sin objetivos ni metas para el futuro. El término viene del budismo, es por eso por lo que las traducciones automáticas lo han

⁶ El neologismo chino «nt» es un término ofensivo para decir que alguien no está en su sano juicio. En este contexto se utiliza para maldecir. Al traducirlo, se ha optado por usar: «qué locura», un término menos ofensivo pero que expresa el estrés que siente el usuario del comentario.

traducido como «un joven budista». La oración de ejemplo, sacada del mismo vídeo de Tik Tok donde explicaban el neologismo, dice: «我不是佛系青年，我很努力复习». La autora de esta investigación lo traduciría como: «No soy un joven apalancado, me esfuerzo mucho al estudiar». Chat GPT 3.5 ha usado el término «joven despreocupado». También sería correcto, pero uno puede mirar de no preocuparse mucho por los eventos futuros, pero aun así tener ambiciones para construirse un buen futuro. Así pues, «apalancado» parece más adecuado en este contexto.

El quinto término es 炒鱿鱼 (*chao youyu*), un complemento del nombre compuesto por un adjetivo y un sustantivo (炒, que significa «salteado»; y 鱿鱼, que significa «calamares»). Originalmente se refería a un plato de calamares salteados, pero en los últimos años se ha empezado a utilizar esta expresión no estandarizada para reflejar que te han despedido del trabajo.

La frase de ejemplo, cogida del comentario de un usuario en Weibo, muestra muy bien estas dos acepciones tan distintas: «刷到个视频标题: 工作被炒鱿鱼了; 于是我内心: 我想吃烤鱿鱼了……没有吃饭是真得饿». Literalmente la oración dice: «Me he encontrado con un vídeo que se titula: “Me han despedido del trabajo”, y lo primero que he pensado ha sido: “Ahora tengo ganas de comer calamares salteados...” No he comido nada y tengo un hambre». Los motores de traducción automática y los chatbots lo han traducido así, pero no sería una oración muy lógica para un lector español, puesto que no entendería la relación que tiene el haber sido despedido con que uno quiera comer calamares salteados. Así pues, la solución sería reformular la oración usando la técnica de la creación discursiva, omitiendo completamente el plato de comida: «Me he encontrado con un vídeo que se titula: “Me han chutado del trabajo”, y lo primero que he pensado ha sido: “Ahora tengo ganas de ir a jugar al fútbol y chutar unos penaltis...” Hace tanto tiempo que no toco el campo...». Con esta oración, se mantiene el juego de palabras a la vez que se expresa el mismo significado del título del vídeo, que es el neologismo que queremos mantener. Es importante remarcar que, en este contexto, la expresión no va acompañada de ningún elemento de vídeo o imagen que muestre el plato de calamares (véase anexo 1, figura 17), lo que facilita la traducción, ya que deja más margen a la creatividad. En otros casos en que incluyera un recurso audiovisual, se debería de buscar una solución que mantuviera el significado de la acepción estandarizada.

4.2. Resultados de las traducciones realizadas por las herramientas

En el siguiente punto se mostrará una serie de gráficos que describen el porcentaje de puntuaciones en aciertos y errores (según la rúbrica presentada en el capítulo 3.3.) que ha tenido cada herramienta de traducción automática e inteligencia artificial a la hora de traducir términos coloquiales chinos. Asimismo, se comentarán de forma breve los resultados debajo de los gráficos. Este apartado está intrínsecamente relacionado con el siguiente, que compara en mayor profundidad las evaluaciones de las herramientas para poder ver con facilidad cuáles son las que mejor traducen términos coloquiales chinos.

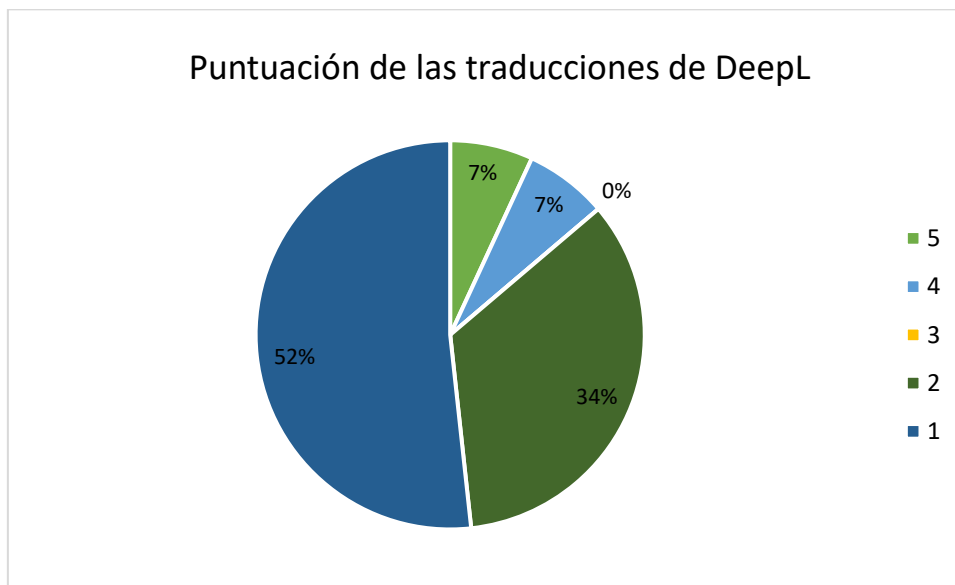


Gráfico 1: Puntuación de las traducciones de DeepL. (5 = Puntuación máxima).

En cuanto a DeepL, se puede ver que no es un motor de TA muy fiable a la hora de traducir neologismos del ámbito coloquial, ya que un 86 % de las expresiones han sido puntuadas con 1 o un 2. En muchos casos, el traductor automático ha dejado el término tal y como estaba en chino o lo ha traducido según la versión estandarizada. En cuanto a puntuaciones positivas, solo un 14 % de las traducciones han sido evaluadas con un 4 o un 5, lo que no ofrece mucha confianza al traductor humano que quiera recurrir a las opciones del traductor automático para su propia traducción.

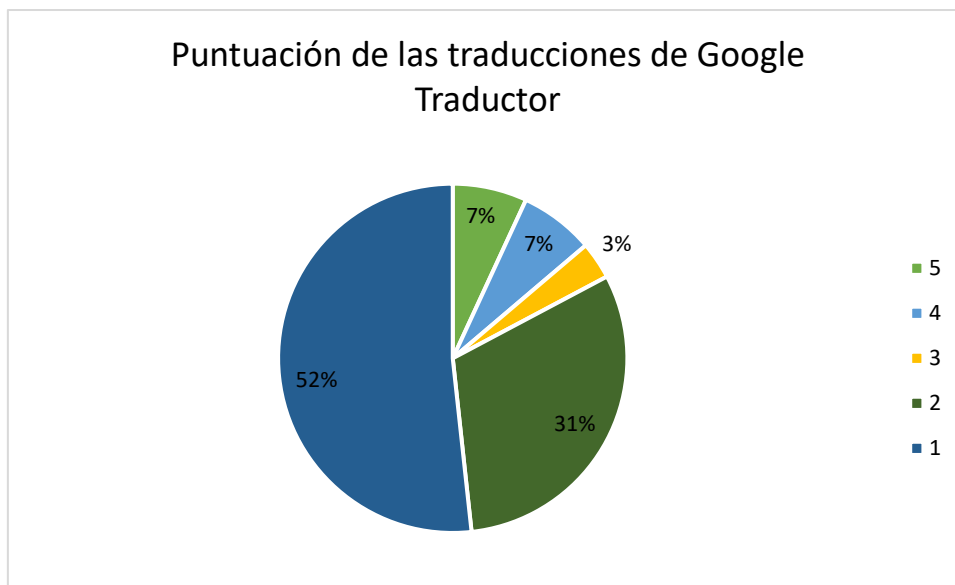


Gráfico 2: Puntuación de las traducciones de Google Traductor. (5 = Puntuación máxima).

Google Traductor tampoco se podría considerar un motor de TA muy fiable a la hora de traducir neologismos del ámbito coloquial en chino. En este caso, más de la mitad de las traducciones han sido evaluadas con una puntuación de un 1, en concreto, un 52 %. En la mayoría de los casos el sistema de TA ha traducido las expresiones siguiendo su acepción estandarizada o ha dejado el término en chino sin traducir. Solamente un 17 % se podría considerar una traducción buena y adecuada al contexto.

El porcentaje de aciertos es muy parecido al de DeepL y se puede afirmar que cuatro de los cinco neologismos traducidos correctamente (seis en el caso de Google Traductor) han coincidido en los dos motores de traducción automática, es decir, ambos traductores han traducido casi los mismos términos siguiendo la acepción no estandarizada.

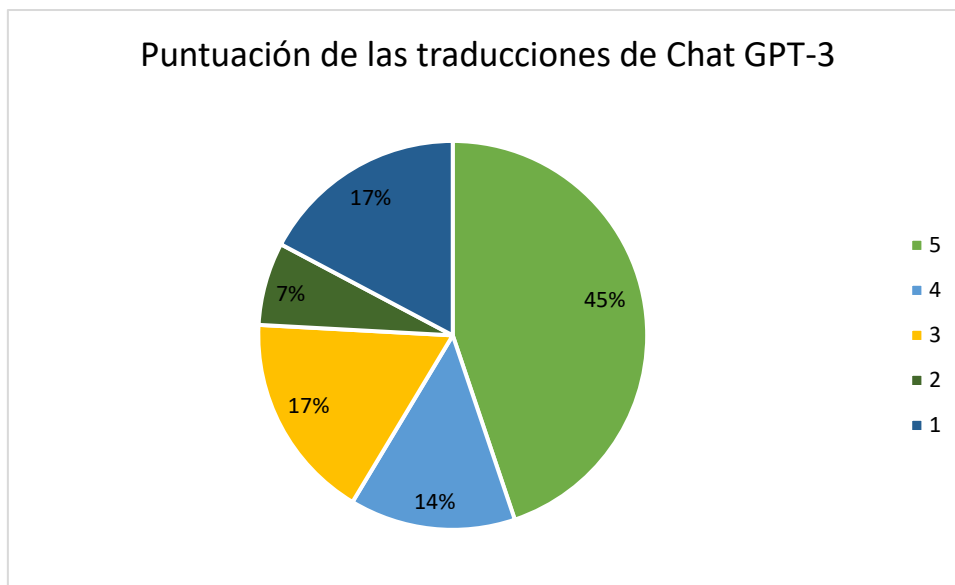


Gráfico 3: Puntuación de las traducciones de Chat GPT-3. (5 = Puntuación máxima).

Chat GPT-3, que se basa en la inteligencia artificial generativa, ofrece un alto porcentaje de aciertos: un 59 % de los términos han sido evaluados con un 5 o un 4, y un 17 % de ellos se han puntuado con un 3 ya sea porque se han traducido correctamente, pero carecen de adecuación al contexto, o porque se han traducido incorrectamente, pero se han explicado en los comentarios complementarios que ofrece este chatbot. Asimismo, tan solo un 24 % de las expresiones se han evaluado con un 1 o un 2, bien porque no han sido traducidas, bien porque lo han sido, pero de forma incorrecta.

Es importante remarcar que Chat GPT, al igual que Gemini, ofrece explicaciones de los términos. Esto es muy útil para el traductor humano ya que, si la traducción es errónea, igualmente podrá observar la explicación del término, pues en algunas ocasiones la definición difiere de la traducción ofrecida. Esto se puede entender bien con el siguiente ejemplo sacado de la respuesta de Gemini al pedirle la traducción del término «佛系»:

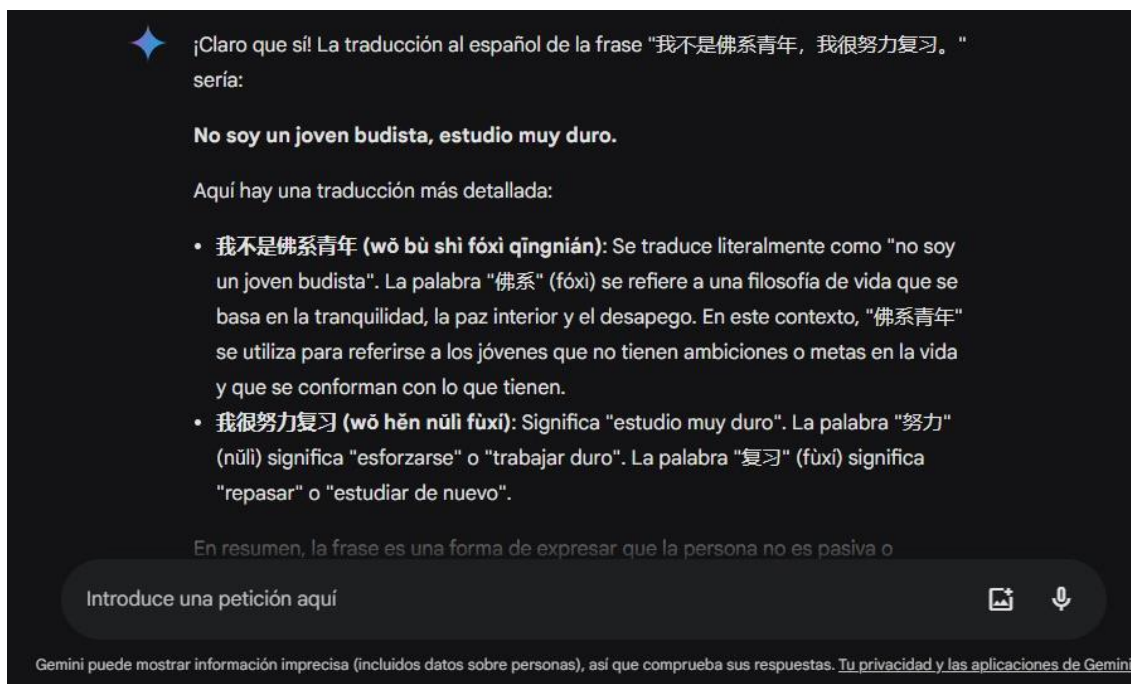


Figura 35: Traducción y explicación de Gemini del término «佛系»⁷.

Así pues, puede ayudar al traductor a buscar una opción más adecuada. Aun así, no se aconseja confiar plenamente en los chatbots, ya que tal y como se ha visto en la parte teórica de este trabajo de final de grado, a veces estos pueden inventarse la información que dan, por lo que siempre es importante contrastar los datos que proporcionan.

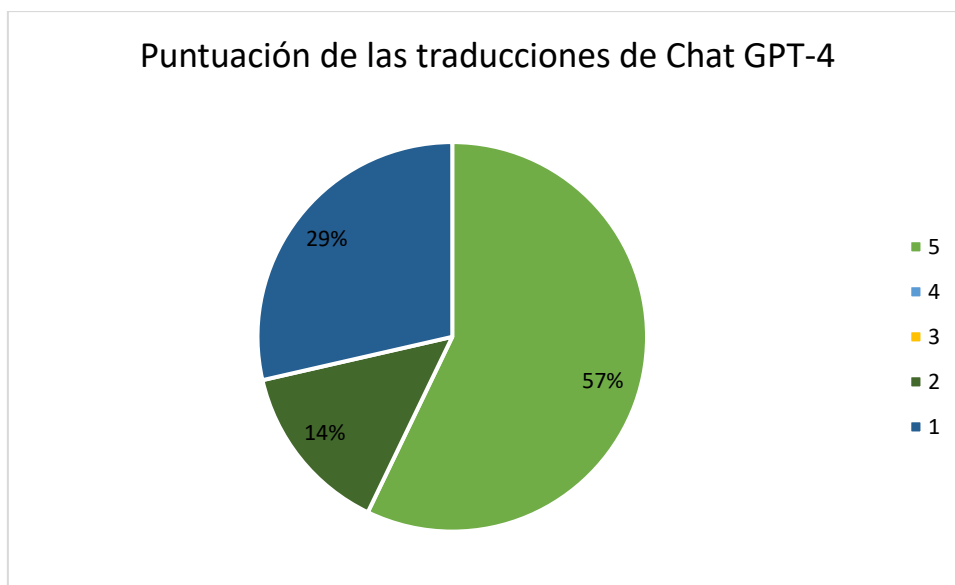


Gráfico 4: Puntuación de las traducciones de Chat GPT-4. (5 = Puntuación máxima).

⁷ Las figuras expuestas en el cuerpo del trabajo también se pueden encontrar en el anexo 2.

En el caso de Chat GPT-4, tan solo se han traducido los términos con una puntuación menor a 2 en las traducciones de Chat GPT-3.5, en este caso, han sido 7 expresiones coloquiales. De esta manera se puede ver si con una versión más actualizada y de pago, el resultado es más satisfactorio.

Tal y como puede verse en el gráfico, las traducciones han diferido mucho. Por un lado, un 57 % de los términos han sido corregidos y mejorados, y lo que era una traducción incorrecta o sin lógica ha pasado a ser una traducción siguiendo la acepción no estandarizada y cohesionada con la oración. Por el otro lado, un 43 % de las traducciones han continuado sin ser acertadas, en algunos casos incluso se ha inventado una explicación del término no estandarizado.

En un caso, se ha hecho la prueba de pedirle al chatbot una explicación de por qué ha dejado el término sin traducir. En su respuesta se ha autocorregido y ha dado el resultado deseado. Se puede ver la interacción en el anexo 2.

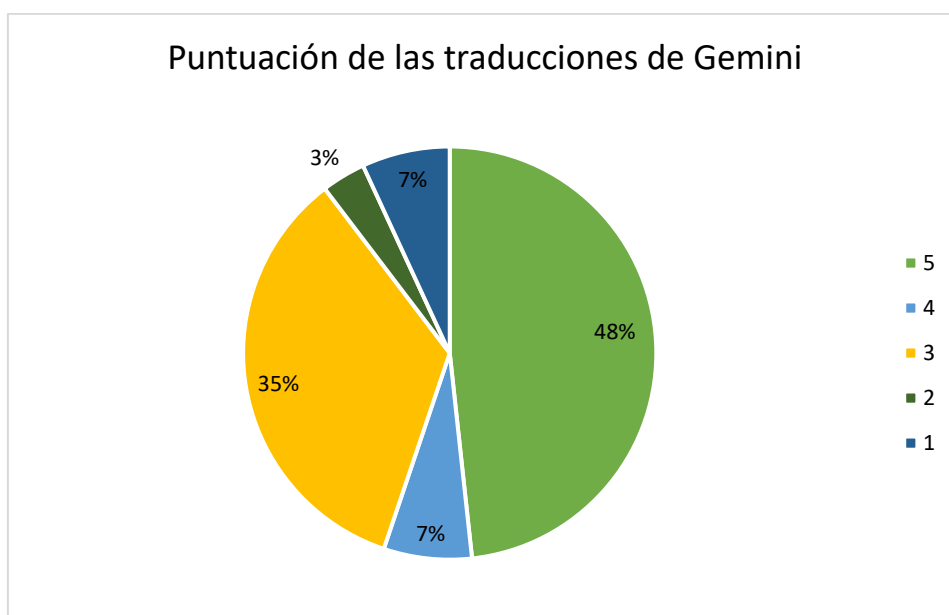


Gráfico 5: Puntuación de las traducciones de Gemini. (5 = Puntuación máxima).

El sistema de inteligencia artificial Gemini también ha ofrecido unos resultados muy satisfactorios. Se puede ver que un 55 % de los términos han sido traducidos de forma correcta y adecuada. Asimismo se puede observar que un 35 % de las traducciones del chatbot de Google se han puntuado con un 3, esto es porque en muchos casos las traducciones son incorrectas, pero las explicaciones sí son correctas y adecuadas al significado real de la acepción no estandarizada.

Cabe recordar que Gemini ofrece tres respuestas a una misma pregunta, lo que complica la evaluación, ya que hay veces que una o dos de las respuestas son correctas,

pero la otra es menos adecuada al contexto. En estos casos, se ha seguido a la mayoría, por lo que si dos eran correctas y una incorrecta, se ha puntuado igualmente con un 5. También ha habido veces en que, a pesar de dar tres versiones distintas, la traducción de la oración o del término ha sido la misma, lo que ha simplificado la evaluación. Estas explicaciones y versiones distintas podrían ayudar al traductor, siempre que este luego contrastara las respuestas con otras fuentes, ya que, como ya ha sido mencionado, a veces el chatbot puede inventarse el origen o la definición del neologismo chino del ámbito coloquial. A continuación se puede observar cómo se visualizan estas tres opciones que ofrece Gemini en la interfaz del programa:

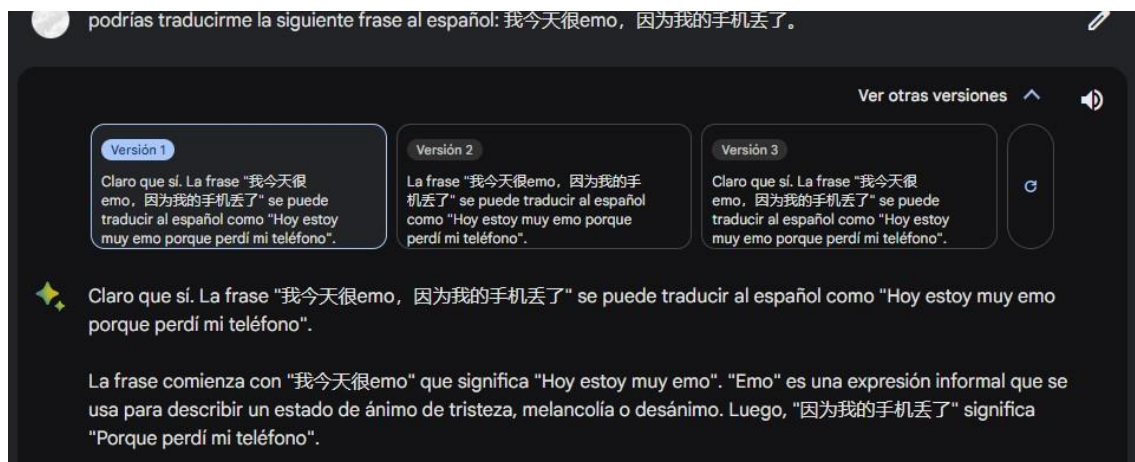


Figura 34: Interfaz del programa Gemini con tres versiones de respuesta a una misma pregunta.

4.3. Comparación y discusión de los resultados

En el siguiente capítulo se compararán las puntuaciones en las traducciones de las herramientas de traducción automática e inteligencia artificial. Para poder visualizar mejor estas diferencias o similitudes se usará el gráfico de Excel siguiente:

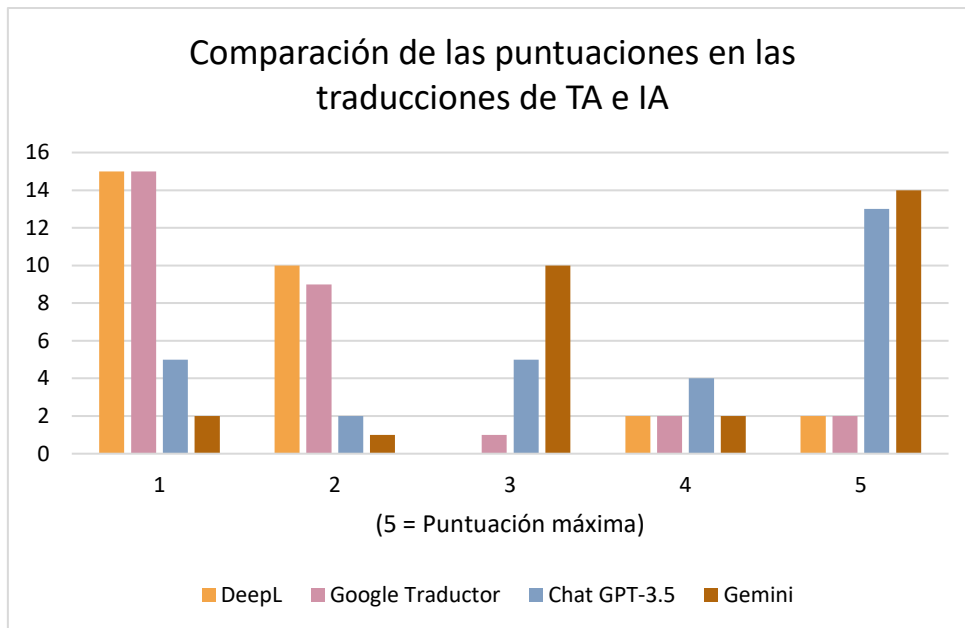


Gráfico 6: Comparación de las puntuaciones en las traducciones por TA e IA.

Es importante remarcar que no se compararan los resultados de Chat GPT-4 porque tan solo se utilizó con 7 oraciones de ejemplo, por lo que el cotejo no sería justo.

A simple vista se puede ver que los motores de traducción automática e inteligencia artificial traducen los términos coloquiales chinos de forma muy distinta: mientras que Gemini y Chat GPT destacan por tener un porcentaje de aciertos más elevado que de errores; DeepL y Google Traductor es el contrario, con un porcentaje mayor de errores.

En primer lugar, cabe destacar la diferencia entre los dos traductores automáticos neuronales. De manera superficial se puede afirmar que ambas TA tienen un nivel muy bajo de conocimiento de neologismos de Internet chinos, eso ha hecho que recurrieran a la traducción según la acepción estandarizada o que, directamente, no tradujeran la palabra y la dejaran en chino. Ambos casos han tenido el mismo porcentaje de evaluaciones con un 1, es decir, que el término no ha sido ni traducido. A pesar de esto, parece ser que, en caso de no saber el significado del neologismo del ámbito coloquial, DeepL prefiere traducirlo, aunque sea de forma incorrecta e incluso, a veces, contraria a la acepción real del término no estandarizado.

En segundo lugar, las dos inteligencias artificiales, Chat GPT y Gemini, tampoco difieren mucho entre ellas, pero, en este caso, en cuanto al número de aciertos. Aun así, sí que hay una leve diferencia entre ambas en el porcentaje de errores.

Por un lado, si nos fijamos en Chat GPT-3.5, se puede ver que ha tenido un gran porcentaje de aciertos puntuados con un 5, y un porcentaje alto bastante parecido de evaluaciones con un 3 y un 4:

En el primer caso, se han valorado con un 3 porque el chatbot ha dejado el término en chino, pero luego lo ha explicado correctamente más abajo. Esto ha sido muy común en las traducciones de expresiones numéricas. En otros casos, ha pasado lo contrario, con una traducción correcta pero una descripción incorrecta.

En el segundo caso, se han valorado con un 4 porque las traducciones son correctas y tienen lógica con la oración, pero se podrían traducir con un término mucho más adecuado a la acepción no estandarizada.

Por otro lado, Gemini destaca por tener un gran porcentaje de aciertos con un 5 y un 3. Esto posiblemente sea porque, al ofrecer tres versiones, tiene más posibilidades de dar una traducción adecuada al contexto y acorde con el significado de la acepción no estandarizada. En las evaluaciones con un 3, ha pasado lo mismo que con Chat GPT-3.5, y es que los términos no han sido traducidos pero sí explicados a continuación. Esto también ha pasado en su mayoría con las expresiones numéricas.

Este chatbot ha sido el traductor con menor cantidad de errores, un hecho que se debería de tener en consideración al escoger una herramienta para traducir los términos coloquiales chinos. Aun así, también cabe remarcar que Chat GPT-3.5 destaca por ofrecer unos resultados más regulares, es decir, que tiene un porcentaje más alto de traducciones evaluadas con un 4 o un 5, al contrario que Gemini. Teniendo esto en cuenta, si fuera posible, sería mejor utilizar la versión de pago, es decir, Chat GPT-4, puesto que, tal y como se puede ver en el gráfico 4, ha traducido correctamente más de la mitad de las traducciones valoradas con menos de un 2 en la versión gratuita, así pues, seguramente ofrezca unos resultados más satisfactorios.

4.4. Análisis cualitativo de las traducciones por TA e IA en el contexto de las redes sociales

A continuación se hará un análisis cualitativo de las traducciones ofrecidas por los motores de traducción automática e inteligencia artificial en el contexto de las redes sociales.

El objetivo es justificar las evaluaciones, para luego ser publicadas en las mismas plataformas de donde han sido extraídas, es decir, las redes sociales. Este contexto es

clave para definir si una traducción es buena o mala. Se debe tener en cuenta que un traductor humano, al traducir, sí lo tendrá en cuenta; mientras que las herramientas de TA e IA no. Se mostrará también un caso en que se ha ofrecido cierto contexto a un chatbot. A continuación se verá si esto ha afectado al resultado.

Es importante recordar que tan solo se ha valorado que el neologismo de Internet concreto estuviera bien traducido y tuviera lógica con el contexto. Esto no significa que la oración en general se haya traducido de manera natural y lógica, ya que en muchos casos no es así.

Como ya se ha mencionado anteriormente, se utilizará la tabla número 4, mostrada en el capítulo 3.3.

En primer lugar, se ha valorado con un 1, la puntuación mínima, cualquier término que no ha sido traducido, es decir, que se ha dejado tal y como estaba escrito en el idioma original. También los casos en que se ha omitido el neologismo del ámbito coloquial. Un ejemplo de ello ha sido «JM» (en la oración de ejemplo, escrito en minúsculas: «jm»), traducido como «hermanas» o «tías», en el sentido de amigas. Ninguna herramienta de TA o IA lo ha traducido de manera satisfactoria.

Una de las razones podría ser que el término se encuentra al principio de la oración, por lo que es fácil que un traductor automático lo considere un nombre propio: «jm 这真的很崩溃从想主题开始然后疯狂讨论删删改改写策划废时将近三个小时这简直 nt». La oración se ha traducido más o menos bien dependiendo de la herramienta, pero siempre ha empezado igual:

Ejemplo 1: JM. Traducción de Google Traductor:

jm Esto es realmente frustrante. Comencé a pensar en el tema y luego comencé a discutir locamente sobre eliminar, eliminar, modificar, reescribir y planificar. Fue una pérdida de casi tres horas. Esto es simplemente no.

Ejemplo 1: JM. Traducción de Chat GPT-4

jm, esto es realmente frustrante, desde pensar en el tema, luego discutirlo locamente, borrar y editar, hasta escribir y planificar, he desperdiciado casi tres horas, esto es simplemente nt.

La única versión que ha diferido de forma leve es la de Chat GPT-3.5, que al considerar el término como un nombre propio, lo ha incluido dentro de la oración. Esto la ha hecho

más lógica, pero a la vez más incorrecta, ya que «JM» no es un complemento del nombre, sino una apelación a sus seguidores y amigos. De esta manera se ha creado un contrasentido:

Ejemplo 1: JM. Traducción de Chat GPT-3.5.

La situación con JM realmente es desesperante. Empezamos desde el tema, y luego discutimos y editamos frenéticamente el plan durante casi tres horas. Es increíble, ¡qué pérdida de tiempo!

La propuesta de la autora de este TFG es la siguiente:

Ejemplo 1: JM. Traducción de la autora del TFG.

Tías, esto es muy frustrante. He malgastado casi tres horas pensando el tema y discutiéndolo sin parar. He borrado, cambiado y modificado tanto... En serio, qué locura.

Es cierto que se podría traducir como «hermanas» o «amigas», pero para mantener la informalidad, el término «tías» también sería adecuado al contexto.

En segundo lugar, se ha valorado con un 2 cualquier intento incorrecto de traducción del neologismo del ámbito coloquial. Un ejemplo de ello es «栓 Q», traducido como «gracias». Se encuentra en la oración de ejemplo siguiente: «电脑拿回家加班, 结果都没打开一下 真的栓 Q».

Google Traductor ha intentado traducir la expresión, pero el resultado ha sido erróneo y sin lógica. Parece ser que ha seguido una de las acepciones del término «栓», que sería «rayo», el cual es rápido.

Ejemplo 2: «栓 Q». Traducción de Google Traductor.

Llevé mi computadora a casa para trabajar horas extras, pero ni siquiera la encendí. Muy rápido.

Chat GPT-3.5 también ha intentado traducir el término, pero parece ser que, al no entender su significado, ha buscado una expresión que encajara con el contexto, aunque no tenga nada que ver con «栓 Q». La versión de pago de Chat GPT ha hecho un error similar.

Ejemplo 2: «栓 Q». Traducción de Chat GPT-3.5.

Llevé la computadora a casa para trabajar horas extras, pero al final ni siquiera la encendí.
¡Qué desperdicio!

Ejemplo 2: «栓 Q». Traducción de Chat GPT-4.

Llevé la computadora a casa para trabajar horas extra, pero al final no la encendí ni una vez. Es realmente frustrante.

La propuesta de la autora de este TFG es la siguiente:

Ejemplo 2: «栓 Q». Traducción de la autora del TFG:

Me llevé el portátil a casa para hacer unas horas extras, pero al final no lo abrí ni un segundo. Gracias a Dios.

Ya que el significado de «栓 Q» es «gracias», se ha intentado mantener aún más la informalidad con una exageración: «gracias a Dios».

En tercer lugar, se ha valorado con un 3 los términos traducidos correctamente, pero que carecen de coherencia y cohesión con el contexto. También las expresiones que las máquinas de inteligencia artificial han traducido mal o, directamente, dejado en chino, pero que después sí han descrito en la explicación. Un ejemplo de ello es «辣眼睛», traducido como «feo u algo que uno no debería de ver». Se encuentra en la siguiente oración «这张照片太辣眼睛了吧！快删了！».

Chat GPT-3.5. ha traducido el término de forma bastante literal. A pesar de esto, sí que mantiene el sentido correcto, aunque no sea la expresión más acorde con el significado de la acepción no estandarizada y carezca de lógica con el contexto.

Ejemplo 3: «辣眼睛». Traducción de Chat GPT-3.5:

Esta foto es demasiado molesta para los ojos, ¿verdad? ¡Bórrala rápido!

Gemini ha sido aún más literal y ha traducido el término según el significado original. A pesar de ello, parece que la IA es consciente de la acepción no estandarizada y, por ello, más abajo ha dado la explicación correcta y otras opciones de traducción más adecuadas.

Ejemplo 3: «辣眼睛». Traducción y explicación de Gemini en la versión 1:

¡Esta foto es demasiado picante para mis ojos! ¡Bórrala ya!

太辣眼睛了 (tài là yǎnjīng le)⁸: Es una expresión china que significa "demasiado picante para los ojos", que se utiliza para describir algo que es visualmente desagradable, perturbador o incluso obsceno.

En resumen, la oración es una exclamación de disgusto ante una foto que se considera demasiado picante o desagradable para la vista. Se le pide a la persona que la ha publicado que la elimine de inmediato.

Aquí te dejo algunas otras formas de traducir la oración:

- ¡Esta foto es demasiado fuerte! ¡Quítala de inmediato!
- ¡No puedo creer lo que estoy viendo! ¡Elimina esta foto ahora mismo!
- ¡Esta imagen es demasiado obscena! ¡Bórrala de una vez!

Ejemplo 3: «辣眼睛». Traducción y explicación de Gemini en la versión 2:

¡Esta foto es demasiado picante para mis ojos! ¡Bórrala rápido!

太辣眼睛了 (tài là yǎnjīng le): Es una expresión china que se utiliza para describir algo que es visualmente desagradable o chocante. La palabra "辣" (là) significa "picante", pero en este contexto se usa en sentido figurado para referirse a algo que causa irritación o incomodidad. "眼睛" (yǎnjīng) significa "ojos".

En resumen, la oración es una exclamación que expresa desagrado por una foto y una solicitud para que se elimine de inmediato.

Aquí te dejo algunas otras formas de traducir la oración:

- ¡Esta foto es horrible! ¡Quítala de aquí ya mismo!
- ¡Qué foto tan espantosa! ¡No la quiero ver más!
- Esta imagen es demasiado fuerte para mi vista! ¡Eliminala de inmediato!

Ejemplo 3: «辣眼睛». Traducción y explicación de Gemini en la versión 3:

¡Esta foto es demasiado picante para mis ojos! ¡Bórrala rápido!

太辣眼睛了 (tài là yǎnjīng le): Es una expresión idiomática china que significa "ser demasiado picante para los ojos". Se utiliza para describir algo que es visualmente desagradable o chocante.

En resumen, la oración es una exclamación que expresa el disgusto del hablante por una foto que considera demasiado picante o desagradable. Se le pide a la persona que la publicó que la borre de inmediato.

Aquí te dejo algunas otras formas de traducir la oración:

⁸ Se ha mantenido en redonda el pinyin de los neologismos que se encuentran dentro de la explicación del chatbot Gemini, ya que esta se ha copiado sin modificar ni añadir nada.

- ¡Esta foto es demasiado fuerte! ¡Quítala de inmediato!
- ¡Esta imagen es demasiado obscena! ¡Eliminala ya!
- ¡No puedo ver esta foto! ¡No la publiques más!

Este es un claro ejemplo de la importancia de ver más allá de la traducción que te muestra al principio, ya que aunque la primera opción no sea la que concuerde más con el significado del neologismo, puede que las otras versiones que da más abajo, a pesar de alejarse de la literalidad, aporten también más naturalidad y sentido a la oración. También está bien mirar cómo lo describe y, tras contrastarlo con otras fuentes, ver si puede ser útil para tu propia traducción.

La propuesta de la autora de este TFG es la siguiente:

Ejemplo 3: «辣眼睛». Traducción de la autora del TFG:

¡Esta foto es más fea que un dolor! ¡Bórrala ya!

Para traducir este término, se ha buscado una expresión coloquial en español que reflejara la misma fuerza y énfasis que presenta «辣眼睛». Es cierto que algunas de las opciones de traducción ofrecidas por Gemini están acorde con el significado de la acepción no estandarizada, pero la expresión «más feo/a que un dolor» también mantiene esa informalidad. Además, con una rápida búsqueda en Internet se puede ver que este dicho se usa bastante en la actualidad, sobre todo en redes sociales como Twitter.

En cuarto lugar, se ha valorado con una puntuación de 4 los términos traducidos correctamente que tienen coherencia y cohesión con la oración, pero que se podrían mejorar o que cuentan con una opción que concuerda mejor con el significado de la acepción no estandarizada. Un ejemplo de ello ha sido «炫», traducido por «comer». La expresión se encuentra en la oración siguiente: «一人独享 ! 整只豉油鸡 🍗 鲜嫩爆汁 快来一起炫».

Es interesante observar la traducción de Chat GPT-3.5. Al principio, al pedirle que tradujera la oración, ha dado la siguiente versión:

Ejemplo 4: «炫». Traducción de Chat GPT-3.5.

¡Disfruta solo! ¡Todo el pollo al vapor con salsa de soja está tierno y jugoso! Ven y disfruta juntos.

Esta ha sido valorada con 4 puntos, ya que a pesar de no traducirlo como «comamos juntos», ha utilizado el verbo «disfrutar», que también se adecúa al contexto y está acorde con el significado de la acepción no estandarizada. Aun así, se podría haber mantenido el verbo «comer».

A continuación, se le ha vuelto a pedir que tradujera la oración, pero esta vez también se ha añadido una pequeña oración que explicaba el contexto de donde se ha encontrado la frase de ejemplo. Es interesante ver que en este caso, ha traducido la expresión coloquial siguiendo la acepción estandarizada del término: «presumir». Esta traducción se ha valorado con un 2. Se pueden ver ambas interacciones en el anexo 2, concretamente en las figuras 32 y 33.

Ejemplo 4: «炫». Traducción de Chat GPT-3.5.

¡Disfrute individual! Pollo entero con salsa de soja 🐔, tierno y jugoso. ¡Ven y presume conmigo!

Para ver la propuesta de la autora junto con su justificación, véase el capítulo 4.1.

En quinto y último lugar, se ha valorado con una puntuación de un 5, la puntuación máxima, los términos traducidos exitosamente, con coherencia y cohesión al contexto. Un ejemplo de ello ha sido la primera acepción no estandarizada del término «牛逼», traducido como «ser guay». Se puede encontrar en la oración de ejemplo siguiente: «周也的骨相真的好牛逼».

Tanto los motores de traducción automática como las inteligencias artificiales han traducido correctamente este neologismo del ámbito coloquial. Se podría argumentar que es por la ubicación del término en la oración, ya que al ser un verbo adjetival y una oración de sujeto-verbo, no es muy complicada ni compleja para un traductor automático; pero también es por la antigüedad del neologismo. Esta acepción ya lleva más de 10 años en uso, y no es difícil encontrar foros de Internet de la década pasada en donde se menciona y explica. Esto ha hecho que se convierta en una expresión tan conocida que incluso se ha añadido en algunos diccionarios monolingües y bilingües, a diferencia de otros términos coloquiales.

A continuación se observarán un par de traducciones, puesto que la mayoría han sido similares.

Google Traductor ha usado el adjetivo «asombroso», el cual funciona ya que es correcto y está acorde con el significado de la acepción no estandarizada, además de tener lógica con el contexto.

Ejemplo 5: «牛逼». Traducción de Google Traductor.

El físico de Zhou Ye es realmente asombroso.

El chatbot Gemini ha ofrecido tres versiones, dos de las cuales usan el mismo adjetivo: «increíble». Se puede observar que el neologismo del ámbito coloquial también ha sido traducido de forma adecuada al contexto.

Ejemplo 5: «牛逼». Traducción y explicación de Gemini en la versión 1.

La estructura facial de Zhou Ye es realmente increíble.

La expresión "真的好牛逼" significa "es realmente increíble".

En este contexto, la frase se puede interpretar como una expresión de admiración por la belleza de la estructura facial de Zhou Ye. La persona que escribió la frase está diciendo que la actriz tiene una estructura facial muy armoniosa y atractiva.

También podría ser una expresión de asombro por el talento de Zhou Ye como actriz. La persona que escribió la frase podría estar diciendo que la actriz tiene un gran potencial para el éxito en su carrera.

En última instancia, el significado de la frase depende del contexto en el que se use. Sin embargo, en general, se puede interpretar como una expresión de admiración, asombro o elogio por la estructura facial de Zhou Ye.

Aquí hay algunas otras formas de traducir la frase:

"La estructura facial de Zhou Ye es realmente impresionante"

"La estructura facial de Zhou Ye es realmente asombrosa"

"¡La estructura facial de Zhou Ye es realmente increíble!"

Ejemplo 5: «牛逼». Traducción y explicación de Gemini en la versión 2.

La estructura ósea de Zhou Ye es realmente increíble.

"牛逼" es una expresión coloquial que significa "increíble" o "asombroso".

En este contexto, la frase se puede interpretar como una expresión de admiración por la belleza natural de Zhou Ye. La estructura ósea de su rostro se considera ideal, con rasgos simétricos y armoniosos. Esto le da una apariencia elegante y sofisticada, que la ha convertido en una de las actrices más populares de China.

También podría ser una expresión de orgullo por la cultura china. La belleza natural se considera una cualidad importante en la cultura china, y la estructura ósea de Zhou Ye es un ejemplo perfecto de ello.

En última instancia, el significado de la frase depende del contexto en el que se use. Sin embargo, en general, se puede interpretar como una expresión de admiración por la belleza natural de Zhou Ye o por la cultura china.

Aquí hay algunas otras formas de traducir la frase:

"La cara de Zhou Ye tiene una estructura ósea realmente increíble"

"La belleza natural de Zhou Ye es realmente admirable"

"Zhou Ye es una belleza natural con una estructura ósea perfecta"

Ejemplo 5: «牛逼». Traducción y explicación de Gemini en la versión 3.

La estructura ósea de Zhou Ye es realmente asombrosa.

La expresión "真的好牛逼" significa "es realmente asombroso".

En este contexto, la frase se puede interpretar como una expresión de admiración por la belleza y el atractivo de Zhou Ye. La persona que dice la frase está elogiando la estructura ósea de la actriz, que se considera un factor importante en la belleza facial china.

También podría ser una expresión de sorpresa o asombro por la belleza de Zhou Ye. La persona que dice la frase está diciendo que no puede creer lo bonita que es la actriz.

En última instancia, el significado de la frase depende del contexto en el que se use. Sin embargo, en general, se puede interpretar como una expresión de admiración por la belleza de Zhou Ye.

Aquí hay algunas otras formas de traducir la frase:

"La estructura ósea de Zhou Ye es realmente impresionante"

"La estructura ósea de Zhou Ye es realmente envidiable"

"¡La estructura ósea de Zhou Ye es increíble!"

La propuesta de la autora de este TFG es la siguiente:

Ejemplo 3: «牛逼». Traducción de la autora del TFG:

La cara de Zhou Ye es tan perfecta.

Se ha preferido usar el adjetivo «perfecta» por el contexto. El usuario menciona la esplendor y perfección de la forma facial de la famosa actriz china Zhou Ye, así pues, se ha preferido usar un adjetivo mucho más común para describir un aspecto físico, como sería «perfecto». Llamarlo «asombroso» o «increíble» no es incorrecto, pero hay otros

adjetivos que funcionan mejor. Uno de ellos también podría ser «envidiable», que se encuentra entre las propuestas de Gemini.

CONCLUSIONES

En este trabajo de fin de grado se ha evaluado la precisión y utilidad de las traducciones al español de neologismos chinos del ámbito coloquial por diversas herramientas de traducción automática e inteligencia artificial. Los resultados obtenidos muestran que, si bien estos motores han mejorado significativamente en los últimos años, aún presentan limitaciones para una traducción fluida, precisa y adecuada de esta clase de lenguaje.

Se pueden sacar dos conclusiones clave de este estudio. Por un lado, es seguro afirmar que las TA como Google Traductor o DeepL aún tienen un largo camino por recorrer, puesto que tal y como se puede ver en el apartado práctico, han tenido un mayor número de errores que de aciertos, y sus traducciones son incorrectas y poco precisas. Por el otro lado, las IA como Chat GPT y Gemini sí que ofrecen unos resultados mucho más satisfactorios. Además, añaden explicaciones junto con las traducciones, lo que puede ser de gran ayuda para el traductor, sobre todo si este no encuentra en ninguna otra plataforma el significado del término o expresión. A pesar de esto, aunque pueda ayudar a crear una primera traducción, es importante contrastar tal información. Asimismo, estas explicaciones también pueden ayudar a determinar la fiabilidad de las traducciones ofrecidas.

El trabajo no parte de una hipótesis concreta, sino de una serie de objetivos que se han conseguido cumplir. El estudio sobre la TA y IA aplicada a la traducción ha ofrecido una visión que, *a posteriori*, ayuda a entender mejor las respuestas que ofrecen las herramientas seleccionadas y su evaluación subsiguiente. Asimismo, gracias a la exhaustiva investigación práctica de comparación y valoración de las traducciones donde se explica con detalle en qué medida traducen correcta e incorrectamente los neologismos, se consigue afirmar lo siguiente: esta clase de traducción sí puede ser de ayuda para un estudiante o traductor de chino, aunque este siempre debe de trabajar con cautela y verificar cualquier información que le dé la IA.

Es importante resaltar que este estudio ha contado con ciertas limitaciones. Como por ejemplo que tan solo se ha valorado la traducción de los términos, no de todo el texto de partida, puesto que sino se extendería mucho la investigación y el análisis. Tampoco se han incorporado otras inteligencias artificiales conocidas del mercado, como Le Chat de Mistral o 文心一言 (*Wenxin Yiyen*), por falta de tiempo.

Asimismo, algunas de las herramientas cuentan con la versión de pago, por lo que se ha tenido que decidir si contar con ellas o no: finalmente se ha optado por no incluir

DeepL Pro, pero sí Chat GPT-4. Aun así, aparecieron algunos problemas al utilizarlo, ya que, al ser de pago, se tuvo que usar la cuenta de la universidad y hacer las pruebas *in situ*, en la misma facultad. En su momento, esto complicó un poco la situación ya que se tenía que acordar una fecha en que tanto la tutora como la alumna del TFG tuvieran disponibilidad durante un buen rato del día, puesto que el proceso era largo. Así pues, se decidió reducir el número de pruebas, traduciendo tan solo los términos que en la versión gratuita hubieran obtenido una nota baja, es decir, un 1 o un 2. De esta manera se vería si los resultados son más satisfactorios y fieles con una versión mejorada y actualizada.

Dicho esto, el presente trabajo abre distintas vías de investigaciones futuras, ya sea en el campo de la traducción aplicada a la educación, como en el análisis de los neologismos de Internet chinos o en el estudio de las herramientas de TA e IA. Sería muy interesante, por ejemplo, aprovechar las limitaciones que han aparecido en este TFG para futuros estudios, como hacer una investigación a más gran escala de las herramientas de TA e IA aplicadas a la traducción de neologismos coloquiales chinos, puesto que tanto los programas como los términos de Internet están en constante cambio. También es importante hacer más hincapié en los chatbots, ya que al llevar poco tiempo en el mercado, aún no hay un gran número de estudios que traten con profundidad el impacto que pueden tener en el mundo de la traducción. De igual forma, los neologismos de Internet chinos también son un campo que necesita más investigación en español, por lo que se podría llegar a crear un glosario o corpus chino-español de estos términos y expresiones.

Por último, personalmente, como estudiante de 4.º de traducción, debo agradecer haber escogido este trabajo, ya que aparte de haber sido muy valioso para mi formación profesional y académica, también me ha abierto las puertas a un área profesional que no había pensado en dedicarme, y esta es la vía de la investigación. A partir de este trabajo y la motivación de la tutora, la Dra. Mireia Vargas Urpí, continuaré mi carrera profesional en el campo de la traducción, educación e investigación traductológica, centrándome en el idioma chino. De esta manera también podré continuar algunos de los posibles estudios mencionados anteriormente.

En cuanto al tema de este trabajo de fin de grado, en mi día a día, sí que hago uso de las herramientas de TA e IA disponibles en el mercado. Gracias al estudio pertinente, he empezado a usar más Gemini, el cual normalmente queda eclipsado por la creciente popularidad de Chat GPT, a pesar de ofrecer unos resultados más satisfactorios. Asimismo, he dejado de usar tanto DeepL, uno de los traductores automáticos más usados del mercado, puesto que, aunque en algunos casos puede ayudar a entender mejor una

oración en chino cuando uno está muy perdido; no es nada útil para términos más específicos, como sería el caso de los neologismos de Internet.

REFERENCIAS

Lista de fuentes bibliográficas

- Briva-Iglesias, V. (2021). Traducción humana vs. traducción automática: análisis contrastivo e implicaciones para la aplicación de la traducción automática en traducción jurídica. *Mutatis Mutandis. Revista Latinoamericana de Traducción*, 14(2), 571-600.
- Bruck, P. A. (2006). Microlearning as strategic research field: An invitation to collaborate. T. Hug, M. Lindner, & Peter A. Bruck (eds.). *Microlearning: Emerging Concepts, Practices and Technologies after e-Learning. Proceedings of Microlearning 2005. Learning and Working in New Media*. Innsbruck, Innsbruck University Press, 13-18.
- Casacuberta Nolla, F. y Peris Abril, Á. (2017). Traducció automàtica neuronal. *Tradumàtica*, 15, 66-74. <https://doi.org/10.5565/rev/tradumatica.203>.
- Costa-Jussà, M. R., Henríquez, C. y Banchs, R.E. (2011). Evaluación de estrategias para la traducción automática estadística de chino a castellano con el inglés como lengua pivote. *Procesamiento del lenguaje natural*, 47, 119-126. <http://hdl.handle.net/10045/18519>.
- Duan Chuxuan (2023). Tik Tok: A New Way of English Learning. *Journal of Education, Humanities and Social Sciences*, 8, 127-133. <https://doi.org/10.54097/ehss.v8i.4237>.
- Figura, K. y Jarvis, H. (2007). Computer-based materials: A study of learner autonomy and strategies. *System*, 35, 4, 448-468. <https://doi.org/10.1016/j.system.2007.07.001>.
- Fitria. T. N. (2023). Performance of Google Translate, Microsoft Translator, and DeepL Translator: Error Analysis of Translation Result. *Al-Lisan: Jurnal Bahasa (e-Journal)*, 8(2), 115-138. <https://doi.org/10.30603/al.v8i2.3442>.
- Flores-Borjabad, S. A. (2023). La intersección entre la adquisición del lenguaje y la inteligencia artificial: explorando el potencial de los modelos de lenguaje natural. *Amazonia Investiga*, 12 (62), 7-9. <https://doi.org/10.34069/AI/2023.62.02.0>.
- Forcada-Zubizarreta, M. L. (2017). Making sense of neural machine translation. *Translation spaces*, 6(2), 291-309. <https://doi.org/10.1075/ts.6.2.06for>.
- Forcada-Zubizarreta, M. L. (2022). Automática (Traducción). *ENTI (Enciclopedia de traducción e interpretación)*. AIETI. <https://doi.org/10.5281/ZENODO.6634083>.

- Ganascia, J. G. (2017). *Le Mythe de la Singularité. Faut-il craindre l'intelligence artificielle?* Le Seuil, Science ouverte.
- Ginestí, M. y Forcada-Zubizarreta, M. L. (2009). La traducció automàtica en la pràctica: aplicacions, dificultats i estratègies de desenvolupament. *Caplletra. Revista Internacional de Filologia*, (46), 43-60.
- Gómez, L. M. (2021). *Inteligencia artificial y pymes: Chatbots*. [Trabajo de fin de grado, Universidad Politécnica de Cartagena]. Repositorio Digital de la Universidad Politécnica de Cartagena. <https://repositorio.upct.es/handle/10317/10188>.
- Hansen, D. (2023). The ambivalence of machine translation and AI: Resouce or replacement? *Counterpoint*, 10, 10-12.
- Hearne, M. y Way, A. (2011). Statistical Machine Translation: A guide for linguists and translators. *Language and Linguistics Compass*, 5(5), 205-226. <https://doi.org/10.1111/j.1749-818X.2011.00274.x>.
- Hernández Alonso, C. (1991). El lenguaje coloquial juvenil. *Revista de la Asociación Europea de Profesores de Español Año XXI-XXII*, 38-39, 3.
- Ji Ziwei, Yu Tiezheng, Xu Yan, Lee, N., Etsuko, I. y Fung, P. (2023). Towards Mitigating Hallucination in Large Language Models via Self-Reflection. *Findings of the Association for Computational Linguistics: EMNLP*, 1827–1843. <https://arxiv.org/abs/2310.06271>.
- Kalla, D. y Smith, N. (2023). Study and analysis of chat GPT and its impact on different fields of study. *International Journal of Innovative Science and Research Technology*, 8(3), 827-833.
- McGee, R. W. (2023). *ChatGPT and Copyright Infringement: An Exploratory Study*. [Working paper, Fayetteville State University]. <http://dx.doi.org/10.2139/ssrn.4578430>.
- Mogas Ribas, M. (2016). *La traducción automática: ¿una amenaza para el traductor del siglo XXI?* [Trabajo de fin de grado, Universidad Autónoma de Barcelona]. Depósito digital de documentos de la UAB. <https://ddd.uab.cat/record/160589>.
- Morduchowicz, R. (2018). *Ruidos en la web: Cómo se informan los adolescentes en la era digital*. Buenos Aires, Ediciones B, 192. <https://doi.org/10.32870/Pk.a9n16.378>.
- Nitzke, J. (2019). *Problem solving activities in post-editing and translation from scratch*. Language Science Press. <https://doi.org/10.5281/ZENODO.2546446>.

- Novajra, F. (2023). No one, no language and no book left behind: CEATL takes action on AI. *Counterpoint*, 10, 7-8.
- Otero, M. (2018, 19 de marzo). Los más jóvenes, grandes creadores de palabras. *LaVoz*. <https://www.lavoz.com.ar/ciudadanos/los-mas-jovenes-grandes-creadores-de-palabras/>.
- Piqueras, A. (2020). *Estudio comparativo entre los procesos de traducción humana y traducción automática estadística con la combinación de lenguas chino-español*. [Trabajo fin de máster, Universidad Autónoma de Barcelona]. Depósito digital de documentos de la UAB. (p. 63). <https://ddd.uab.cat/record/249933>.
- Real Academia Española. (s.f.). Inteligencia artificial. En el *Diccionario de la lengua española*. Recuperado el 22 de enero, 2024, en <https://dle.rae.es/inteligencia>.
- Real Academia Española. (s.f.). Neologismo. En el *Diccionario de la lengua española*. Recuperado el 12 de diciembre, 2023, en <https://dle.rae.es/neologismo>.
- Rescigno, A. A., Eva, V., Monti, J. y Andy, W. (2020). A Case Study of Natural Gender Phenomena in Translation-A Comparison of Google Translate, Bing Microsoft Translator and DeepL for English to Italian, French and Spanish. *Workshop on the Impact of Machine Translation*, 62-90.
- Rovira-Esteva, S. y Vargas-Urpí, M. (2023). Del aula tradicional al microaprendizaje móvil: análisis del potencial de Instagram para aprender chino. *Sinología Hispánica*, 16(1), 1-26. <https://ddd.uab.cat/record/275333>.
- Sharples, M., Arnedillo-Sánchez, I., Milrad, M. y Vavoula, G. (2009). Mobile Learning: Small devices, Big Issues. N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder y S. Barnes (Eds.), *Technology Enhanced Learning: Principles and Products* (p. 233-249), Heidelberg, Springer.
- Strojin, G. (2023). Copyright in light of generative AI. *Counterpoint*, 10, 10-12.
- Tao Yingnian (2017). An Investigation Into Chinese Internet Neologisms. *Canadian Social Science*, 13(12), 65-70. <https://eprints.lancs.ac.uk/id/eprint/142671/>.
- Viver-Sorolla, P. (2018). *La evaluación de las herramientas de traducción automática (TA) desde la perspectiva del traductor: Google Translate, Bing, Babylon y Systran*. [Trabajo de fin de grado, Universidad de Valladolid]. <https://uvadoc.uva.es/handle/10324/33981>.
- Wang Jing (2012). Lun xin meiti zai duiwai hanyu kecheng jiaoxue zhong deyingyong (论新媒体在对外汉语课程教学中的应用) [On the Application of New Media in Teaching Chinese as a Foreign Language Courses]. *Qinghai Shifan Daxue*

- Xuebao (*Zhexue Shehui Kexue Ban*) (青海师范大学学报 (哲学社会科学版)), J. Qinghai Normal University, 34(4), 137-139.
- Xiaoma Cidian (小马词典). (s.f.) 炫 (*xuan*). En *Xiaoma Cidian* (小马词典). Recuperado el 10 de mayo, 2024, en <https://www.xiaoma.info/hanzi.php?hz=%E7%82%AB&fhz=%E7%82%AB&fpy=qiao>.
- Yulianto, A. y Supriatnaningsih, R. (2021). Google Translate vs. DeepL: A quantitative evaluation of close-language pair translation (French to English). *AJELP: Asian Journal of English Language and Pedagogy*, 9(2), 109-127. <https://doi.org/10.37134/ajelp.vol9.2.9.2021>.
- Zhang Jingshuo (2022). *El Aprendizaje de Chino con podcasts*. [Trabajo final de postgrado, Universidad Autónoma de Barcelona]. Depósito digital de documentos de la UAB. <https://ddd.uab.cat/record/265400>.
- Zhang Xinyu (2017). *Traducción de los nombres de los platos chinos y su relación con herramientas de traducción automática*. [Trabajo de fin de máster, Universidad Autónoma de Barcelona]. Depósito digital de documentos de la UAB. <https://ddd.uab.cat/record/188589>.
- Zhi Chen (2020). Un estudio empírico del alcance de la traducción automática del español al chino. Caso de estudio de GNMT sobre las expresiones metafóricas y metonímicas. *Círculo de Lingüística Aplicada a la Comunicación*, (83), 1-25. <http://dx.doi.org/10.5209/clac.70560>.

ANEXOS

Anexo 1: Capturas de pantalla de ejemplos en redes sociales



Figura 1: 250.



Figura 2: 484.



Figura 3: 520.



Figura 4: 555.



Figura 5: 996.



Figura 6: 1314.

闲调·烟月 🍵
1小时前 来自 iPhone 13 Pro Max

最近的两大新宠2333



Figura 7: 2333.

KN_烤奶
10-18 12:15 来自 iPhone客户端

9494 我宝从小到大一直都是古灵精怪嘞🤡🤡//@樱桃刻录:我宝从小到大一直都是那么鲜活的哦

@欢天喜锯之风吱磕犊天: 小香的精神状态惹到谁啦? 我宝宝从小就很活泼爱搞怪好不好.....活跃气氛一把手 小元气猫打气筒 有分寸的搞怪且可爱 不了解严浩翔的人别瞎bb。。。什么变疯了模仿别人 小香一直都是开心果小霸王可爱鬼搞怪王~可不是只会大叫张牙舞爪的疯子🤡反思你们自己吧!



Figura 8: 9494.

手作淡奶 🍵
昨天 18:04 来自 聪律cp超话

💡 聪律cp 其实从看到pyq的那瞬间就希望老公把视频给删了吧。。因为彼此在乎互相都爽了的两个人。。“那你下播能把你pyq发的健身视频给删了吗”-“看你表现了”-“我给你逼疯了！” “你不是说你天天健身状态可好吗” · 📺 手作淡奶的微博视频

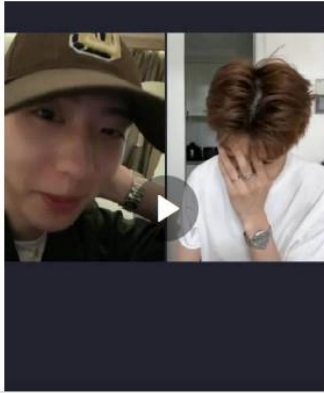


Figura 11: PYQ.

小曦345
45分钟前 来自 iPhone客户端

u1s1晚上下雨的古镇挺吓人的👹👹 突如其来的大风好猝不及防

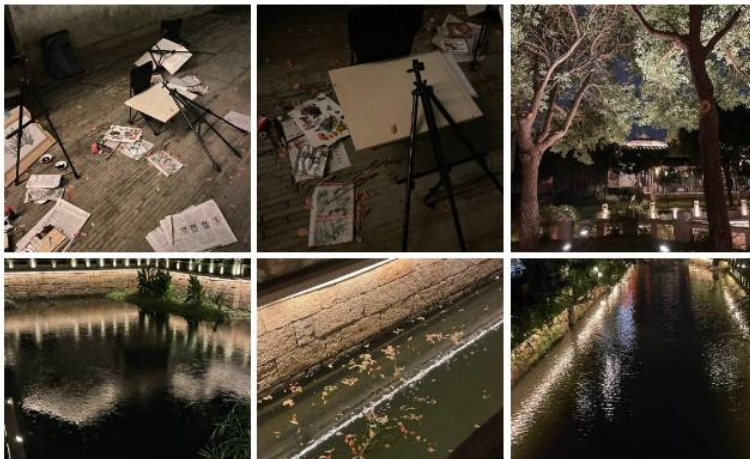


Figura 12: U1S1.

悲伤鳗鱼饭-见过朱志鑫版 🍷
19分钟前 来自 iPhone 14 Pro

xswl, 我爸妈的朋友都认识了朱志鑫 🍷🍷

🔄 转发

💬 评论

👍 赞

Figura 13: XSWL.



Figura 14: 修勾.

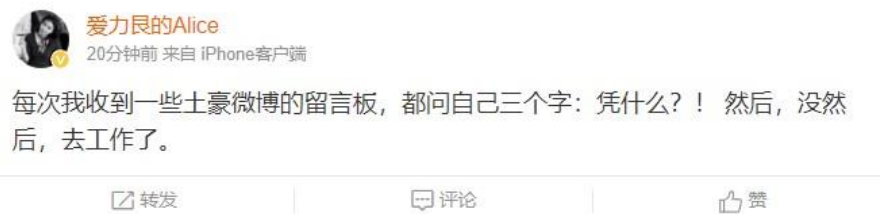


Figura 15: 土豪.



Figura 16: 栓 Q.

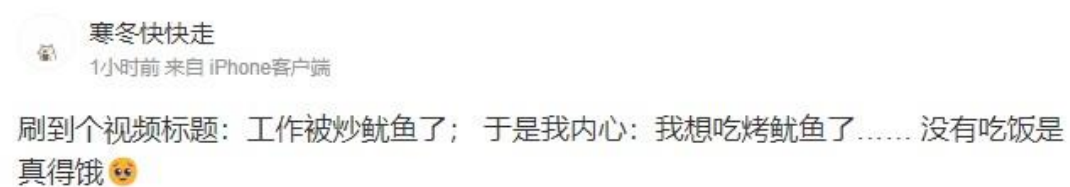


Figura 17: 炒鱿鱼.



Figura 18: 牛逼.



Figura 19: 集美.



Figura 20: 炫.



Figura 21: 干饭.



Figura 22: 辣眼睛.



Figura 23: 芭比 Q.



Figura 24: emo.



Figura 25: YYDS.



Figura 26: 吃货.



Figura 27: 佛系.



Figura 28: 牛逼.



Figura 29: 菜.

Anexo 2: Interacciones con los chatbots

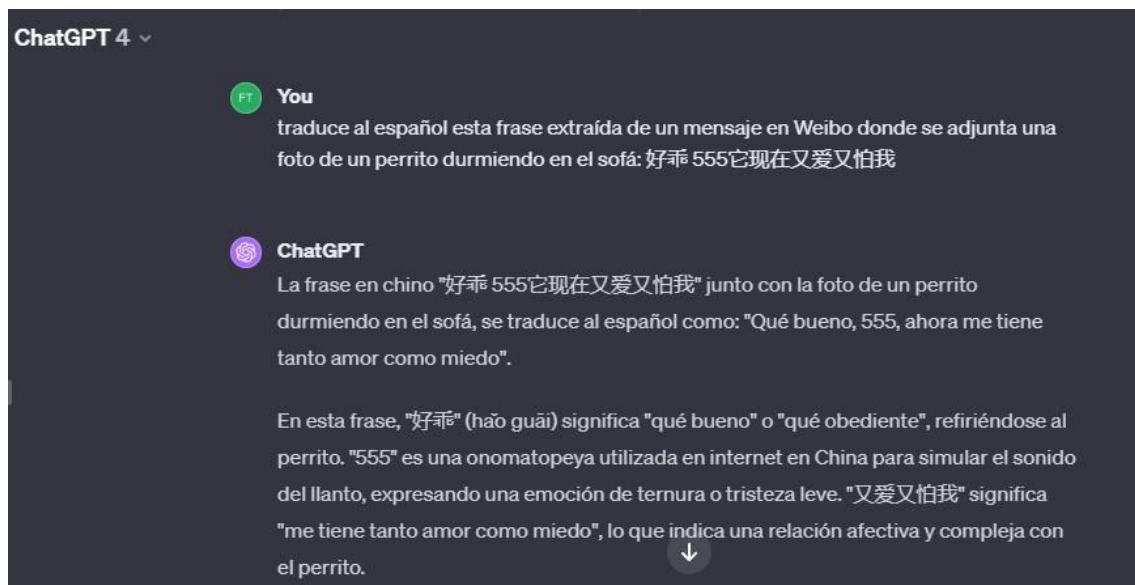


Figura 30: Traducción de Chat GPT-4 del término 555.

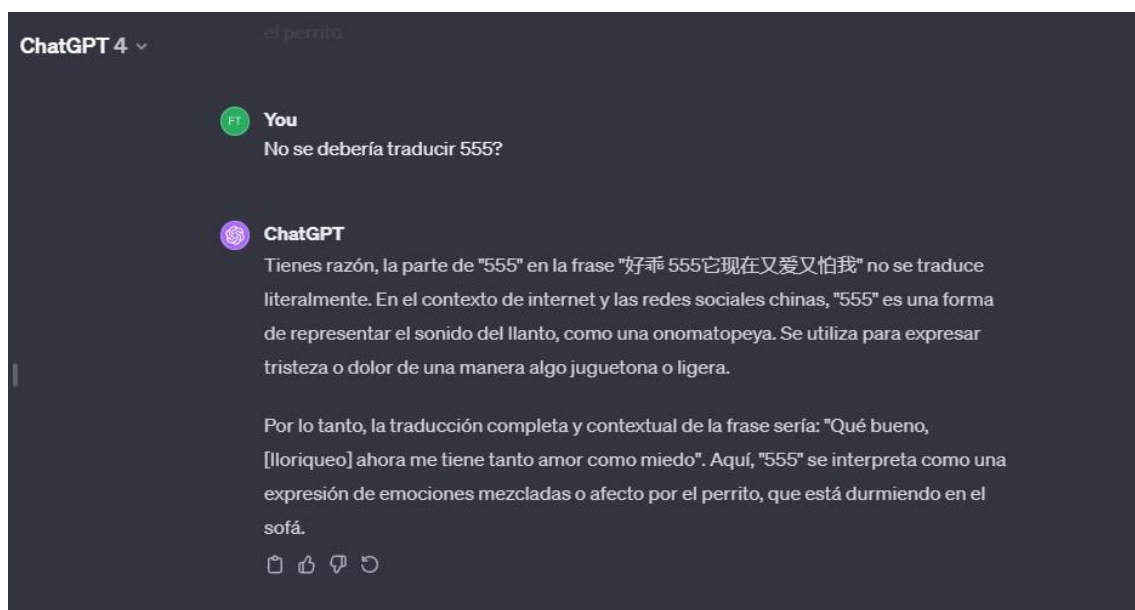


Figura 31: Réplica de Chat GPT-4 a la traducción del término 555.

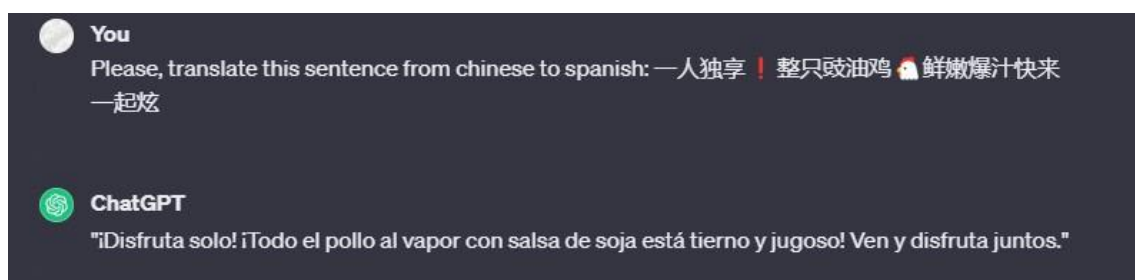


Figura 32: Traducción de Chat GPT-3.5 del término «炫» sin el contexto de la oración.

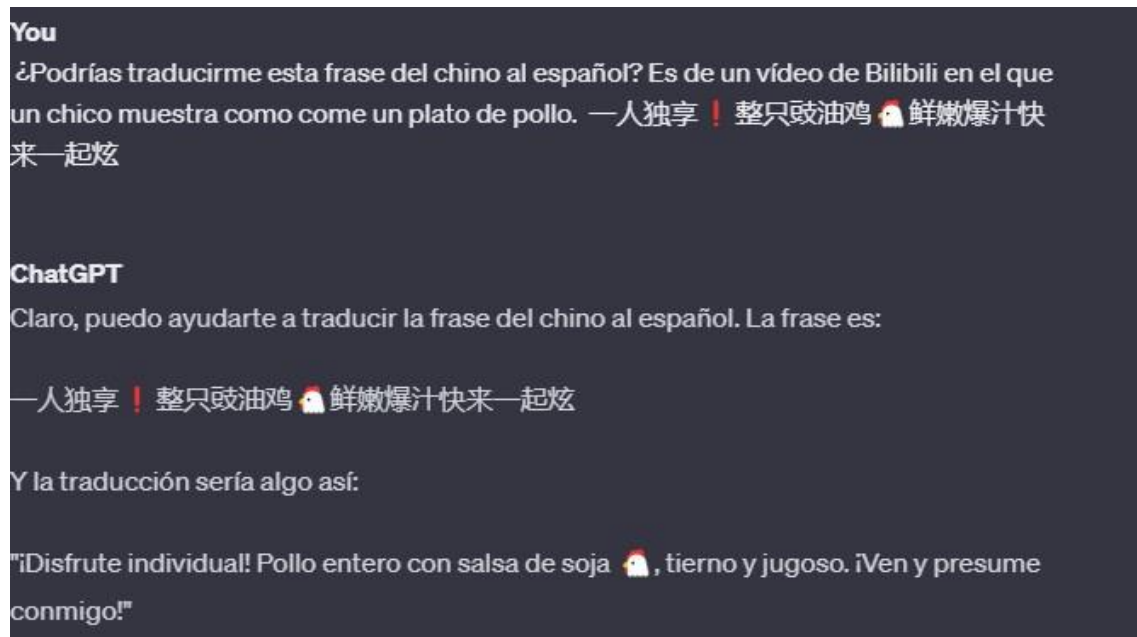


Figura 33: Traducción de Chat GPT-3.5 del término «炫» con el contexto de la oración.

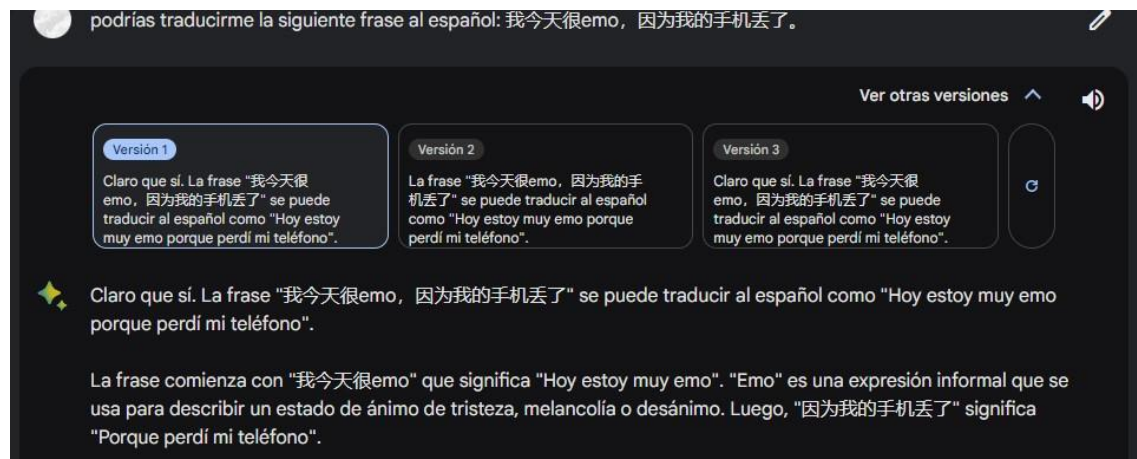


Figura 34: Interfaz del programa Gemini con tres versiones de respuesta a una misma pregunta.

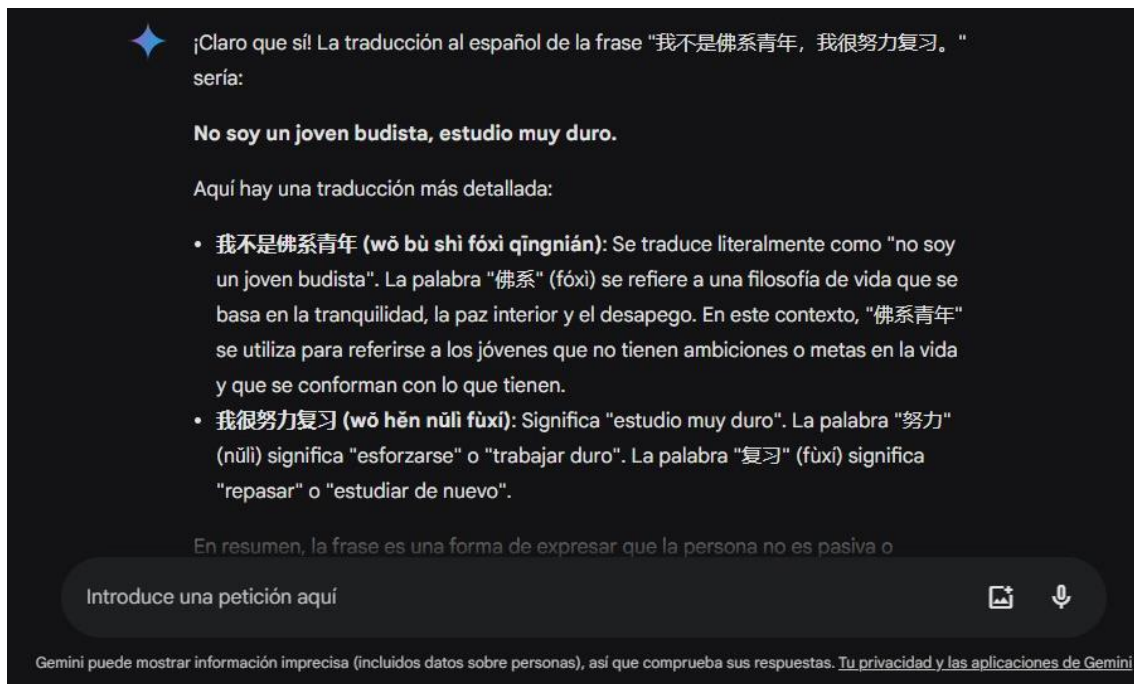


Figura 35: Traducción y explicación de Gemini del término «佛系».