

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

Páez Mérida, María; Espín García, María del Carmen, dir. La traducción automática en internet : Google Traductor y Baidu Translate. 2019. (0 Estudis d'Àsia Oriental)

This version is available at <https://ddd.uab.cat/record/213041>

under the terms of the  **IN**
COPYRIGHT license

FACULTAD DE TRADUCCIÓN E INTERPRETACIÓN

GRADO EN TRADUCCIÓN E INTERPRETACIÓN

TRABAJO DE FIN DE GRADO

Curso 2018-2019

**La traducción automática en internet:
Google Traductor y Baidu Translate**

María Páez Mérida

1422846

TUTORA

MARÍA DEL CARMEN ESPÍN GARCÍA

Barcelona, 3 de junio de 2019



**Universitat Autònoma
de Barcelona**

Datos del TFG

Título: La traducción automática en internet: Google Traductor y Baidu Translate

Autor/a: María Páez Mérida

Tutor: María del Carmen Espín García

Centro: Universitat Autònoma de Barcelona

Estudios: Grado en Traducción e Interpretación

Curso académico: 2018-2019

Palabras clave

Traducción automática, posesición, Google, Baidu, traducción chino-español

Resumen del TFG

Este trabajo pretende analizar si el resultado de los traductores automáticos Google Traductor y Baidu Translate en el par de lenguas chino-español puede ayudar a los estudiantes de traducción con chino como lengua C a tener acceso a bibliografía escrita en chino para elaborar trabajos académicos. Esta investigación consta de dos partes: por un lado, el marco teórico en el cual se explica en qué consiste la traducción automática, la posesición y cuál es la situación actual de la traducción automática en la combinación chino-español; por otro lado, en la parte práctica compararemos las traducciones de seis textos de temática diversa realizadas por estos dos traductores automáticos para ver cuál es mejor. Asimismo, pasaremos un cuestionario a varios estudiantes de traducción de chino para conocer su experiencia y su opinión sobre la traducción automática del chino al español.

Aviso legal

© María Páez Mérida, Bellaterra, 2019. Todos los derechos reservados.

Ningún contenido de este trabajo puede ser objeto de reproducción, comunicación pública, difusión y/o transformación, de forma parcial o total, sin el permiso o la autorización de su autor/a.

Dades del TFG

Títol: La traducció automàtica a internet: Google Traductor i Baidu Translate

Autor/a: María Páez Mérida

Tutor: María del Carmen Espín García

Centre: Universitat Autònoma de Barcelona

Estudis: Grau en Traducció i Interpretació

Curs acadèmic: 2018-2019

Paraules clau

Traducció automàtica, posesició, Google, Baidu, traducció xinès-espanyol

Resum del TFG

Aquest treball pretén analitzar si el resultat dels traductors automàtics Google Traductor i Baidu Translate en la combinació xinès-espanyol pot ajudar als estudiants de traducció amb xinès com a llengua C a tenir accés a bibliografia escrita en xinès per elaborar treballs acadèmics. Aquesta investigació consta de dues parts: d'una banda, el marc teòric en el qual s'explica en què consisteix la traducció automàtica, la posició i quin és la situació actual de la traducció automàtica en la combinació xinès-espanyol; d'altra, a la part pràctica compararem les traduccions de sis textos de temàtiques diferents realitzades per aquests dos traductors automàtics per veure quin és millor. A més a més, passarem un qüestionari a estudiants de traducció de xinès per conèixer la seva experiència i la seva opinió sobre la traducció automàtica del xinès a l'espanyol.

Avís legal

© María Páez Mérida, Bellaterra, 2019. Tots els drets reservats.

Cap contingut d'aquest treball pot ser objecte de reproducció, comunicació pública, difusió i/o transformació, de forma parcial o total, sense el permís o l'autorització del seu autor/de la seva autora.

Bachelor's Degree Final Project information

Title: Machine translation on the internet: Google Translator and Baidu Translate

Author: María Páez Mérida

Tutor: María del Carmen Espín García

Centre: Universitat Autònoma de Barcelona

Studies: Bachelor's Degree in Translation and Interpreting

Academic year: 2018-2019

Key words

Machine translation, post-edition, Google, Baidu, Chinese-Spanish translation

Bachelor's Degree Final Project abstract

The purpose of this Bachelor's Degree Final Project is to analyse whether Google Translate and Baidu Translate automatic translators can help Chinese-Spanish translation students to clearly understand Chinese-written bibliography for academic work. This project is divided in two parts: the theoretical part, where machine translation and post-editing are defined and explained, and the current situation of Chinese-Spanish machine translation; in the practical part, we will compare the translations of six texts of different topics made by these two machine translation engines to know which one can do it better. Moreover, we will launch a questionnaire to know the experiences and opinions of Chinese translation students about Chinese-Spanish machine translation.

Legal notice

© María Páez Mérida, Bellaterra, 2019. All rights reserved.

None of the content of this academic work may be reproduced, distributed, broadcasted and/or transformed, either in whole or in part, without the express permission or authorization of the author.

Índice

1. INTRODUCCIÓN	8
2. LA TRADUCCIÓN AUTOMÁTICA	13
2.1. Introducción	13
2.2. Historia	15
2.3. Sistemas de traducción automática	17
2.3.1. Sistema basado en reglas	17
2.3.2. Traducción automática estadística	18
2.3.3. Traducción automática neuronal	19
2.4. Criterios para evaluar un sistema de TA	22
2.5. Problemas de la TA	25
2.5.1. Problemas de la traducción automática neuronal	26
2.6. La posesición	28
2.7. La traducción automática	30
2.8. Situación en España	32
2.9. Situación en China	33
3. COMPARACIÓN DE GOOGLE Y BAIDU	34
3.1. Metodología del estudio	34
3.2. Análisis de los textos	37
3.2.1. Grado de calidad del texto	37
3.2.2. Detección y categorización de errores	42
3.3. Análisis de resultados	49
4. PROPUESTA DE POSEDICIÓN	65
5. CUESTIONARIO	72
5.1. Descripción del cuestionario	72
5.2. Análisis de datos	74
5.3. Conclusiones	83
6. CONCLUSIONES	84
7. BIBLIOGRAFÍA	90
8. ANEXOS	96
I. Tabla con las puntuaciones para obtener el grado de calidad de los títulos	96
II. Tabla con las puntuaciones para obtener el grado de calidad de los textos	96
III. Tabla con la suma total de errores de Google	97
IV. Tabla con la suma total de errores de Baidu	97
V. Cuestionario	98

Índice de figuras

Fig. 1. Descripción gráfica de un sistema de traducción automática basado en reglas .	17
Fig. 2. Descripción gráfica de un sistema de traducción automática estadística	18
Fig. 3. Descripción gráfica de un sistema de traducción automática neuronal	20
Fig. 4. Clasificación de errores según MQM (2015).....	22
Fig. 5. Descripción gráfica del funcionamiento de Google Translator Toolkit	29

Índice de gráficos

Gráfico 1. Grado de calidad de los títulos de Google y Baidu	39
Gráfico 2. Grado de calidad de los textos de Google y Baidu.....	40
Gráfico 3. Comparación del grado de calidad de los textos de Google y Baidu	41
Gráfico 4. Número global de errores de cada traductor	49
Gráfico 5. Número total de errores encontrados en cada texto según el traductor.....	49
Gráfico 6. Representación detallada de los errores cometidos por Baidu	51
Gráfico 7. Representación detallada de los errores cometidos por Google.....	51
Gráfico 8. Representación del número de errores de mala traducción de ambos traductores	54
Gráfico 9. Representación del número de errores de tipografía de ambos traductores ..	55
Gráfico 10. Representación del número de errores de ininteligibilidad de ambos traductores	56
Gráfico 11. Representación del número de errores de terminología de ambos traductores	57
Gráfico 12. Representación del número de errores de gramática de ambos traductores	59
Gráfico 13. Representación del número de errores de inconsistencia de ambos traductores	59

Gráfico 14. Representación del número de errores de adición de ambos traductores....	61
Gráfico 15. Representación del número de errores de omisión de ambos traductores...	63
Gráfico 16. Representación de los resultados de la pregunta 1 del cuestionario.....	74
Gráfico 17. Representación de los resultados de la pregunta 2 del cuestionario.....	75
Gráfico 18. Representación de los resultados de la pregunta 3 del cuestionario.....	76
Gráfico 19. Representación de los resultados de la pregunta 4 del cuestionario.....	76
Gráfico 20. Representación de los resultados de la pregunta 5 del cuestionario.....	76
Gráfico 21. Representación de los resultados de la pregunta 6 del cuestionario.....	77
Gráfico 22. Representación de los resultados de la pregunta 7 del cuestionario.....	77
Gráfico 23. Representación de los resultados de la pregunta 8 del cuestionario.....	78
Gráfico 24. Representación de los resultados de la pregunta 9 del cuestionario.....	78
Gráfico 25. Representación de los resultados de la pregunta 10 del cuestionario.....	79
Gráfico 26. Representación de los resultados de la pregunta 11 del cuestionario.....	79
Gráfico 27. Representación de los resultados de la pregunta 12 del cuestionario.....	80
Gráfico 28. Representación de los resultados de la pregunta 13 del cuestionario.....	80
Gráfico 29. Representación de los resultados de la pregunta 14 del cuestionario.....	81
Gráfico 30. Representación de los resultados de la pregunta 15 del cuestionario.....	81
Gráfico 31. Representación de los resultados de la pregunta 16 del cuestionario.....	82

Abreviaturas

GNMT: Google Neural Machine Translation

MQM: Multidimensional Quality Metrics

PE: posesición

STA: sistemas de traducción automática

TO: texto original

TM: texto meta

TA: traducción automática

TABR: traducción automática basada en reglas

TAE: traducción automática estadística

TAN: traducción automática neuronal

TAO: traducción asistida por ordenador

1. INTRODUCCIÓN

1.1. Motivación

Los traductores debemos realizar muchas tareas de documentación previas y durante el proceso de traducción para resolver las dudas que se nos puedan presentar y realizar una traducción correcta. Como muchos otros estudiantes de traducción de chino, a menudo nos hemos topado con el problema de la barrera lingüística, ya que mientras aún estamos estudiando el idioma, el desconocimiento hace que a veces resulte complicado encontrar toda la información que necesitamos para traducir. Este problema se agrava por el hecho de que todavía existe muy poco material traducido del chino al español sobre temas lingüísticos y culturales de China. Cuando tenemos una duda, en clase podemos acudir a un compañero nativo chino o podemos ir a la biblioteca, pero entonces nos encontramos con el problema de la falta de bibliografía sobre China en español, ya que la mayoría de los originales están en chino o traducidos al inglés.

Ante este panorama, muchas veces solo nos queda una opción: Internet. La dificultad reside en que internet es un mundo cada vez más amplio, por lo que hay que saber diferenciar claramente las fuentes fiables de las que no lo son. Para poder hacerlo, debemos comprender casi en su totalidad el texto que estamos leyendo. Otro problema es que, una vez más, en internet tampoco existe mucha bibliografía en español, por lo que debemos recurrir a las fuentes en inglés que muchas veces tampoco logran resolver nuestras dudas. La mejor opción para encontrar una duda en concreto es buscar en la lengua de origen. He aquí la principal dificultad con la que debemos lidiar los estudiantes de traducción con chino como lengua C: no tenemos el suficiente nivel de chino como para poder comprender artículos muy extensos o sobre una temática específica, ya que, a pesar del rápido ritmo de aprendizaje del grado, el chino es una lengua muy lejana y diferente del español que requiere de mucho tiempo para poder lograr un buen nivel.

Para poder superar esta barrera lingüística, una de las soluciones a las que recurrimos es utilizar alguno de los traductores automáticos en línea, como Google Traductor y Baidu Translate. Esta tampoco es la solución ideal, ya que la traducción automática aún está en proceso de desarrollo y la calidad de los resultados que ofrece no es muy alta. Esta es la razón principal que me ha motivado a realizar este trabajo: comprobar si realmente la traducción automática puede considerarse una herramienta más para los estudiantes de traducción de chino como lengua C para poder acceder a bibliografía original en chino.

1.2. Objetivos y preguntas de investigación

El objetivo principal de este trabajo es comprobar si los estudiantes de chino no nativos pueden hacer uso de los traductores automáticos Google Traductor y Baidu Translate para acceder y comprender bibliografía escrita originalmente en chino.

Asimismo, para poder alcanzar el objetivo principal, primero debemos resolver los objetivos secundarios que nos guiarán hasta él. Estos objetivos son los que van a formar la estructura de nuestro trabajo: conocer qué es la traducción automática, tanto el funcionamiento como la eficacia; saber en qué consiste la posesición y cuál es la situación actual de la traducción automática del chino al español; comparar los traductores automáticos de Google y Baidu para después hacer una propuesta de posesición; y por último, conocer la opinión de los estudiantes sobre la traducción automática en la combinación chino-español.

Para poder lograr estos objetivos nos hemos planteado las siguientes cuestiones:

1. ¿Qué es la traducción automática? ¿Cómo surgió? ¿Cómo ha evolucionado a lo largo del tiempo?
2. ¿Cómo funcionan los motores de traducción automática? ¿Qué tipos de TA existen en la actualidad? ¿Cuál es el más usado? ¿Cómo puedo identificar un buen traductor automático de uno que no lo es?
3. ¿Cuáles son los pros y los contras de la TA? ¿Qué problemas puede traer consigo? ¿Cómo se identifican y clasifican los errores de TA? ¿Cuáles son los más frecuentes? ¿Y lo menos frecuentes?
4. ¿Qué es la posesición? ¿Cuándo se lleva a cabo? ¿Por qué es importante? ¿Existen normas de calidad para la posesición? ¿Cómo debe realizarse?
5. ¿Cuáles son los traductores automáticos más conocidos para el par de lenguas chino-español? ¿Qué sistema de TA utilizan? ¿Qué tienen en común? ¿Qué diferencias existen entre ellos? ¿Cuál es la situación actual de la TA en los países de origen de esta combinación lingüística (China y España)?
6. ¿Qué traductor ofrece textos de mayor calidad? ¿Qué traductor automático comete más errores? ¿Y cuál comete menos? ¿A qué se deben estos errores? ¿Cuál de los dos podría resultar más útil para tareas de documentación? ¿Qué texto necesita más posesición? ¿Cuál necesita menos? ¿Por qué?
7. ¿Qué opinan otros estudiantes de chino de la TA? ¿Creen que es útil? ¿Por qué? ¿Ejemplos?
8. ¿Los estudiantes de chino no nativos pueden utilizar la TA para acceder y comprender bibliografía escrita originalmente en chino? ¿Qué traductor es mejor para ello?

1.3. Metodología

Para poder lograr los objetivos y responder a las preguntas planteadas en el apartado anterior, dividiremos el trabajo en dos partes: una teórica y una práctica. Cada uno de los capítulos y apartados de este trabajo pertenece a un objetivo de los mencionados anteriormente y responde a las preguntas de investigación relacionadas con esos objetivos.

Para la parte teórica de este trabajo, realizaremos una recopilación y resumen de las investigaciones más relevantes sobre el tema que ocupa nuestro trabajo. Para tener una base sólida de conocimiento sobre la TA y todo lo relacionado con ella: tipos, problemas y dificultades, etc., consultaremos las investigaciones de Hutchins y Sommers (1992) y las llevadas a cabo por el grupo de investigación Tradumàtica de la Universitat Autònoma de Barcelona, entre otras.

Para la parte práctica, realizaremos un análisis cualitativo y cuantitativo de los traductores automáticos de Google y Baidu. Para ello, seleccionaremos mediante el buscador chino Baidu seis textos originales en chino sobre diversos aspectos de la lengua y cultura chinas, los más comunes entre estudiantes de chino: año nuevo chino, caligrafía, cocina, la Gran Muralla, medicina tradicional china y turismo. A continuación, los traduciremos al español con los motores de traducción de Google y Baidu. Una vez tengamos el corpus traducido, y antes de pasar a la fase de análisis, primero realizaremos una lectura de los textos para ver si el grado de comprensión es el adecuado para poder entender toda la información importante correctamente. Analizaremos, por una parte, el grado de comprensión de los títulos y, por otra, la de los textos en general. Hemos decidido hacer esta distinción porque consideramos que el título es lo primero que el lector ve y lo que le puede proporcionar una primera idea sobre lo que se va a hablar en el texto, por lo que es importante que tanto la traducción como la comprensión del mismo sean lo más adecuadas posible. Una vez tengamos el grado de calidad del texto, procederemos a la fase de análisis de cada texto en la que señalaremos y clasificaremos los diferentes errores que hayan cometido los traductores de Google y de Baidu según la norma de calidad MQM (2015). Cuando tengamos el análisis terminado, podremos pasar a la fase de posedición. Este capítulo se presentará en forma de tabla en la que aparecerán el original chino, las traducciones automáticas de Google y Baidu y nuestra propuesta de posedición según las recomendaciones de TAUS (2016).

Para completar nuestro trabajo, dedicaremos un capítulo a la descripción, el análisis y las conclusiones del cuestionario que pasaremos a estudiantes de traducción de chino como lengua C para conocer cuál es su opinión al respecto.

En el último capítulo desarrollaremos las conclusiones a las que hemos llegado tras la realización de este trabajo. Tras responder a nuestras preguntas de investigación, finalmente podremos dar respuesta a nuestro objetivo principal: comprobar si los estudiantes de chino no nativos pueden hacer uso de los traductores automáticos Google y Baidu para acceder y comprender bibliografía escrita originalmente en chino.

1.4. Estructura del trabajo

Para lograr los objetivos expuestos en el apartado 1.2., hemos decidido dividir el trabajo en ocho capítulos: la introducción, la traducción automática, la comparación de Google y Baidu, la propuesta de posesición, el cuestionario, las conclusiones, la bibliografía y los anexos. Estos ocho capítulos podemos dividirlos en tres grandes bloques: parte teórica, parte práctica y conclusiones. La parte teórica está formada por un capítulo y ocho apartados y la parte práctica por dos capítulos y dos apartados. El primer capítulo es simplemente introductorio y en él se explica con detalle toda la información relacionada con este trabajo: motivación personal, objetivos y preguntas de investigación, estructura del trabajo y cuestiones de formato.

En el capítulo 2 empieza la parte teórica. Primero hacemos una pequeña introducción sobre la traducción automática y seguimos con un breve repaso por su historia. El segundo apartado explica los diferentes tipos de traducción automática más comunes: traducción automática basada en reglas, traducción automática estadística y traducción automática neuronal. En el tercer y cuarto apartados, respectivamente, se describen los criterios para evaluar un sistema de traducción automática y los problemas derivados de la traducción automática, haciendo especial hincapié en los de la traducción automática neuronal. En el quinto se presentan la fase de posesición y las normas de calidad recomendadas por los profesionales para realizarla. En los tres últimos apartados se habla sobre la situación de la traducción automática del chino al español y, en concreto, cuál es la situación en España y en China.

En el capítulo 3 empieza la parte práctica. El apartado de análisis se divide en dos: por un lado, el análisis del grado de calidad del texto y la detección y categorización de errores, y por el otro, se hace una descomposición detallada de todos estos resultados y sus posibles causas. El cuarto capítulo también pertenece a la parte práctica, ya que la propuesta de posesición se hace a partir de los resultados obtenidos en el análisis del capítulo 3.

El capítulo 5 está dedicado al cuestionario elaborado para conocer la opinión de otros estudiantes de traducción de chino como lengua C sobre la traducción automática del chino al español. En él se incluyen la descripción, el análisis de resultados y las conclusiones del cuestionario.

En el capítulo 6 se exponen las conclusiones obtenidas y una pequeña reflexión personal a la que hemos llegado tras la realización de este trabajo.

Por último, al final del trabajo encontramos los capítulos de la bibliografía que hemos utilizado y de los anexos.

1.5. Cuestiones de formato

Para gestionar todas las fuentes bibliográficas empleadas para nuestro trabajo hemos decidido utilizar el programa Mendeley, que nos facilitó mucho el trabajo a la hora de citarlas todas.

- a) El modelo que hemos seguido para las citas es el propuesto por las normas APA: Apellido, Nombre (Año). *Título*. Fuente [Fecha de consulta]
- b) Para las citas de fuentes originales en chino: Apellido, Nombre (Año). Título en caracteres (*Pinyin*) [Traducción al español]. Fuente [Fecha de consulta]
- c) Hemos dividido las fuentes en «Consultadas y Citadas» y en «Consultadas y No citadas».
- d) Se añade un apartado para los diccionarios que hemos utilizado, uno para las fuentes de donde han sido extraídas las imágenes y otro para las fuentes de donde hemos extraído los seis textos del corpus utilizado para el trabajo.

Tras una consulta a la Fundéu, hemos optado por escribir sin cursiva todos los nombres de marcas, empresas y programas en inglés, lo que también afecta al nombre de Baidu Translate.

2. LA TRADUCCIÓN AUTOMÁTICA

2.1. Introducción

La rápida incorporación y evolución de las nuevas tecnologías en nuestra vida cotidiana ha llegado también hasta nuestro campo profesional en concreto: la traducción. ¿Cómo han afectado estas nuevas tecnologías a nuestra forma de trabajar? Con la creación de un *software* capaz de traducir por sí solo y en menos tiempo que un traductor humano. La denominada «sociedad de la información» ha traído consigo un importante aumento del contenido traducible. Todos queremos conocer en tiempo real y en nuestra lengua materna qué sucede en otras partes del mundo. Sin embargo, al tratarse de un volumen tan grande de contenido, resulta imposible de traducir para el traductor humano. Esto ha propiciado la creación de lo que se conoce como traducción automática (TA), una herramienta que responde a las necesidades actuales de la traducción y uno de los nuevos retos del siglo XXI.

Para comprender mejor qué es la TA, podemos tomar como ejemplo la definición que hacen Hutchins y Somers (1992:1):

«The development of programs which can produce ‘raw’ translations of texts in a relatively well-defined subject domains, which can be revised to give good-quality translated texts at an economically viable rate or which in their unedited state can be read and understood by specialists in the subject for information purposes».

El principal objetivo de la TA es el mismo que si realizáramos una traducción mediante el proceso tradicional: la comunicación, trasladar el significado del texto original al texto de destino. El uso de la TA se limita actualmente a casos puntuales, como por ejemplo textos de índole científica y técnica (negocios, documentos legales y manuales), es decir, textos de carácter objetivo. Por lo tanto, en la literatura no se utiliza la TA, ya que se trata de textos con un componente subjetivo muy alto, algo que la TA todavía no puede asimilar.

A diferencia del proceso llevado a cabo por los traductores humanos para seleccionar los términos adecuados y así lograr una calidad óptima de la traducción (consultar manuales, diccionarios, enciclopedias, etc.), los traductores automáticos llevan todas estas herramientas incorporadas. Estos sistemas funcionan mediante fórmulas y algoritmos que previamente han introducido profesionales humanos. Existen diferentes sistemas de TA, pero en este trabajo nos centraremos únicamente en tres: sistema basado en reglas, sistema estadístico y sistema neuronal, los cuales detallaremos más adelante.

Es evidente que ambos procesos, tanto el humano como el automático, presentan sus propios retos. Un traductor humano es más lento y puede resultar más costoso en un proceso de traducción; sin embargo, posee una ventaja importante que las máquinas no tienen: el conocimiento de todos los elementos pragmáticos que definen a una comunidad de hablantes. La traducción automática no deja de ser una máquina a la que vamos introduciendo datos y más datos para que produzca un buen resultado y facilite el trabajo de los traductores, por lo que no es capaz de comprender los matices que un ser humano sí, por ejemplo las emociones, los aspectos culturales o la sensibilidad con la que deben tratarse ciertos temas. Los principales problemas que presenta la TA están relacionados con aspectos lingüísticos y pragmáticos, no técnicos: la ambigüedad, la complejidad de la sintaxis y las diferencias de vocabulario entre culturas. No obstante, un traductor automático es capaz de traducir en minutos o incluso segundos lo que un traductor humano podría hacer en semanas, por lo que esto la convierte en una herramienta mucho más rápida y asequible para las empresas (Hutchins y Somers, 1992).

Al mismo tiempo, estas son las mismas razones por las que muchos traductores se niegan a hacer uso de la TA. Por un lado, consideran que la traducción automática no es de calidad, ya que siempre es necesario realizar una posedición y, por otro lado, algunos incluso creen que la evolución de esta forma de traducir podría hacer desaparecer la figura del traductor humano. No obstante, la idea de que la TA pueda sustituir el trabajo de los traductores humanos en el futuro es un mito falso. Son los humanos los que hacen posible la creación de cualquier sistema de TA y la realización de las tareas de posedición y revisión.

«There are no ‘translating machines’ which, at the touch of a few buttons, can take any text in any language and produce a perfect translation in any other language without human intervention or assistance. That is an ideal for the distant future, if it is even achievable in principle, which many doubt (Hutchins y Somers, 1992:1). »

Cada vez son más las empresas que optan por la incorporación de la TA en el flujo de trabajo de sus traductores. Si bien es cierto que la TA no es perfecta, sí resulta una herramienta útil para los traductores. La dificultad reside en saber si la TA puede ayudarnos en nuestro proyecto de traducción, identificar qué sistemas de TA son los que mejor se adaptan a nuestras necesidades y, por supuesto, realizar siempre una posedición de la TA para lograr un TM de calidad (Vilarnau i Dalmau, 2006).

Para entender mejor su evolución y cómo funciona, primero haremos un breve repaso de los eventos y avances más significativos a lo largo de la historia que han definido la TA actual.

2.2. Historia

La capacidad comunicativa es un elemento esencial del ser humano, es lo que nos distingue del resto de seres vivos del planeta. Desde la creación de nuestro mundo el ser humano ha aprendido y mejorado mucho esta capacidad. De hecho, a día de hoy lo seguimos haciendo. Como respuesta a esta necesidad comunicativa nació la traducción, el método que ha servido y sirve de puente entre todas las culturas e idiomas del mundo. Gracias a ella, la cultura de cualquier comunidad puede llegar a otras comunidades en la otra punta del planeta. En 1799 fue hallada la que hasta hoy es considerada la primera traducción de la historia de la humanidad: la piedra Rosetta. En ella aparece por primera vez un texto en tres escrituras diferentes: jeroglíficos egipcios, escritura demótica y griego antiguo, lo que ayudó a descifrar la escritura jeroglífica egipcia. Por esta razón, la piedra se ha convertido en un símbolo para los traductores y expertos de la lengua. Desde este importante hallazgo hasta el presente, la traducción ha evolucionado a pasos de gigante.

Son muchos los profesionales que han estudiado la evolución de la TA a lo largo de la historia y así lo han plasmado en los muchos trabajos que podemos consultar hoy. Este breve repaso es un resumen de las investigaciones realizadas por Jordi Centelles (2013) y John Hutchins y Harold Somers (1992).

La primera idea de automatización apareció en el siglo XVII de la mano de Descartes y Leibniz con su propuesta de código numérico universal para poder crear diccionarios mecanizados. Esta derivó en lo que se conoce como «el movimiento del lenguaje universal» (Hutchins y Somers, 1992). Ya hubo anteriormente otros intentos de crear un lenguaje universal, como por ejemplo el Esperanto. Años más tarde, Warren Weaver y Andrew D. Booth sugirieron el uso de ordenadores para traducir. Sin embargo, estas primeras líneas de investigación sobre la TA no vieron la luz hasta 1949, en el memorándum de Weaver.

El público pudo observar por vez primera la utilidad de la TA en 1954 con el proyecto de Georgetown con IBM, el cual consistió en traducir unas 250 palabras del ruso al inglés. No fue una traducción de gran calidad, pero tuvo una gran repercusión a nivel mundial. La traducción fue recibida por parte del público como el nacimiento de una poderosa herramienta, ya que se abría la posibilidad de eliminar la barrera lingüística a la hora de hacer negocios o compartir información. Durante los siguientes años, se continuó la investigación de la TA con gran optimismo, pero sin grandes resultados.

En 1964 se publicó un informe que marcaría un antes y un después en la historia de la traducción automática: el informe ALPAC. Este informe fue elaborado por el *Automatic Language Processing Advisory Committee* (Comité Consultivo para el Procesamiento

Avanzado del Lenguaje), compuesto por siete científicos que evaluaron el sistema de funcionamiento de la traducción automática. Este informe concluyó, básicamente, que la traducción automática era inútil, ya que era un sistema inexacto y lleno de incoherencias que precisaba de mucha revisión y que, por lo tanto, tampoco era rentable económicamente. El impacto de este informe fue de tal magnitud que se paralizaron las investigaciones en curso sobre la TA y se dejó de hablar de ella durante mucho tiempo en Estados Unidos. A pesar de ello, las investigaciones continuaron en Canadá y en Europa (en aquel entonces Comunidad Económica Europea). De esas investigaciones surgieron los sistemas Météo (1976) y Eurotra Project (finales de la década de los 70). No fue hasta principios de los años 90 que en Estados Unidos se volvió a hablar de TA y se empezaron a instalar los primeros sistemas de traducción automática en ordenadores personales. A partir de entonces, la gente empezó a tomar conciencia de la importancia y utilidad de la TA.

A pesar de las investigaciones y el cada vez más alto interés por la TA, los expertos todavía no han logrado crear un sistema perfecto, por lo que es necesario realizar una fase de posesición de la TA. A veces incluso es necesaria una pre-edición, conocida también como pre-traducción.

Actualmente, este es un campo de investigación muy amplio y que avanza a grandes pasos. La prueba está en que las principales empresas del mundo ya cuentan con su propio sistema de traducción automática, como por ejemplo Google Traductor y Bing Microsoft Translator.

2.3. Sistemas de traducción automática

A lo largo de la historia los sistemas de traducción automática han evolucionado considerablemente. Al principio, se empezó a usar el sistema basado en reglas (TABR). Sin embargo, los expertos lo descartaron porque requería de una constante actualización de los datos, era más lento y no resultaba rentable. Después llegó el sistema de traducción estadístico (TAE), que ha sido el más utilizado hasta hace poco por muchas compañías y, a pesar de que ofrece buenos resultados, la calidad es muy baja en algunas combinaciones lingüísticas. Así se ha llegado al último sistema de TA actual y al que cada vez se suman más empresas y traductores: la traducción automática neuronal (TAN).

2.3.1. Sistema basado en reglas

Los sistemas de traducción basados en reglas « [...] realizan transformaciones a partir del texto original, reemplazando palabras por el equivalente más adecuado en la lengua meta y aplicando las reglas gramaticales y diccionarios propios de cada idioma interviniente en el proceso de traducción (Córdoba; Rico *et al*, 2015)».

Este sistema, como bien dice su nombre, utiliza las reglas gramaticales y los diccionarios de cada lengua para realizar la traducción automática. El proceso realizado por este sistema lo podemos dividir en tres fases: análisis, transferencia y generación. Durante el análisis, el texto original se analiza morfológica y sintácticamente. En segundo lugar, se transfieren las estructuras gramaticales del texto original a la lengua meta (Interlingua). Por último, con estas estructuras gramaticales se genera una frase correcta en la lengua meta. En el siguiente esquema podemos ver reflejado este proceso:

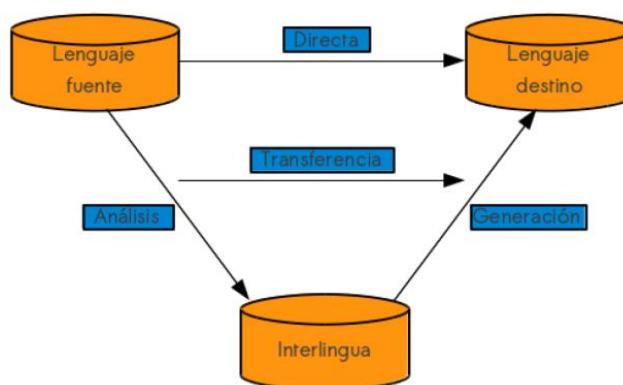


Fig. 1. Descripción gráfica de un sistema de traducción automática basado en reglas

(Useless paperclip, 2017)

Es especialmente útil entre lenguas muy parecidas, como es el caso del catalán y el castellano, e incluso puede llegar a presentar mejores resultados que los sistemas estadísticos. Por ello, programas como Apertium se han convertido en una herramienta

muy útil para el desarrollo de motores con lenguas minoritarias (catalán, gallego, valenciano, etc.) (Parra, 2018).

Este sistema ofrece una traducción correcta sintácticamente, sin embargo, el tono de la traducción final podría resultar demasiado literal. Otro de los principales problemas de las tecnologías basadas en reglas radica en la gran dificultad de llegar a un consenso unánime sobre la normativización del conocimiento lingüístico. Además, este tipo de sistema debe llevar incorporado un gran volumen de normas y combinaciones lingüísticas para que funcione, y esto supone un incremento del coste de producción y mantenimiento. Todo ello hace que esta sea la opción de TA menos rentable, pero aun así hay empresas que continúan utilizando este sistema, como por ejemplo la empresa Systran (Centelles, 2013).

2.3.2. Traducción automática estadística

La traducción automática estadística resurgió en los 90 con modelos estadísticos que sí podían comprender los ordenadores de entonces. Este sistema «aplica un modelo estadístico a partir del análisis de corpus de textos bilingües y paralelos, optimizando recursos y el uso del lenguaje natural (Córdoba; Rico *et al*, 2015)», es decir, utiliza un corpus basado en millones de documentos traducidos de internet en el idioma solicitado para realizar una traducción automática. Entre todos esos millones de documentos busca traducciones que ya se hayan hecho previamente sobre ese grupo de palabras y escoge la opción más utilizada. En el siguiente cuadro podemos observar el funcionamiento de este sistema de TA:

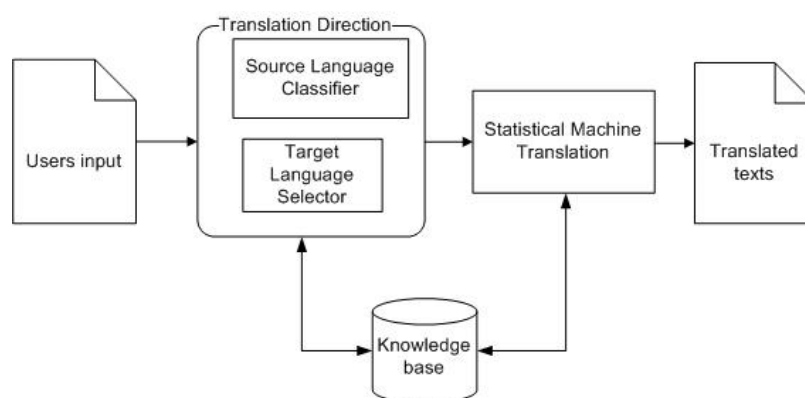


Fig. 2. Descripción gráfica de un sistema de traducción automática estadística

(Wibawa; Nafalski *et al*, 2013)

Matemáticamente, la TAE consiste en encontrar una frase en la LM que resuelva la ecuación que presenta una frase en LO. Esto se consigue mediante el análisis de corpus para que los parámetros de probabilidad de la TAE asignen una probabilidad alta a aquellas frases que son traducciones en la LO y una probabilidad baja o nula a aquellas

que no lo son. Una de las herramientas más populares para la construcción y uso de sistemas de TAE es el sistema Moses (Casacuberta; Peris, 2017).

A pesar de la rapidez y eficacia de la TAE, esta ofrece la mayoría de las veces traducciones muy deficientes, tanto en la gramática como en el léxico, e incluso a veces llegan a ser incomprensibles, especialmente con las lenguas que más se alejan de la estructura del inglés como las lenguas asiáticas. Al utilizar un gran volumen de documentos que se encuentran en la red, resulta el sistema perfecto para realizar traducciones a través de traductores en línea. Hasta hace poco, este era el sistema de traducción automática por excelencia, empleado por empresas tan importantes como Google o Microsoft. No obstante, recientemente las empresas se están sumando al uso de un nuevo sistema de traducción: la traducción neuronal.

2.3.3. Traducción automática neuronal

«La Traducción Automática Neuronal ha llegado y parece que realmente va a barrer con todo lo que había antes. Se está convirtiendo rápidamente en la nueva tecnología de vanguardia», Way (2017).

En los últimos años, los modelos de traducción automática neuronal, también conocidos como inteligencia artificial, están ganando cada vez más terreno en el sector de la traducción.

La principal característica de la traducción neuronal es el uso de redes neuronales cuyo funcionamiento es muy similar al de las redes neuronales biológicas del cerebro humano, por lo que los resultados que ofrece son más precisos y naturales, como si de un traductor humano se tratase. El uso de estas redes neuronales hace que las palabras se representen de forma numérica mediante el uso de vectores. Esto se debe a su funcionamiento a través del denominado *deep learning* o *Deep Neural Networks*, el cual consiste en el uso de algoritmos capaces de aprender automáticamente de sus errores y por lo tanto de autocorregirse a través de la recopilación y almacenamiento de datos.

Por un lado, tenemos la red neuronal conocida como «perceptrón multicapa», la cual se compone de capas de neuronas de forma que las salidas de una capa son las entradas de la capa siguiente. Por otra parte, existen las denominadas «redes neuronales recurrentes», en las que las neuronas se realimentan con sus propias salidas de forma directa o indirecta. Cada palabra traducida depende de las palabras previamente traducidas de una cierta representación de la frase origen y de los parámetros del modelo de traducción» (Casacuberta; Peris, 2017).

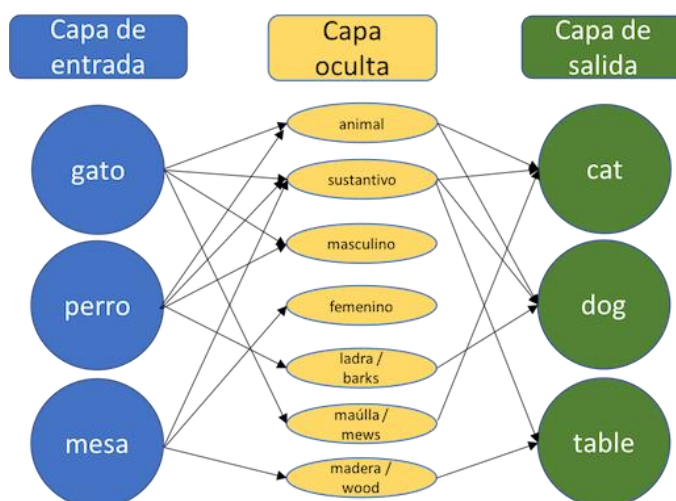


Fig. 3. Descripción gráfica de un sistema de traducción automática neuronal

(Parra, 2018)

Muchos traductores se muestran preocupados por el auge de este sistema de traducción, ya que muchos creen que finalmente sustituirá a la figura del traductor humano. Sin embargo, el ser humano es una pieza clave en la traducción automática, ya que, según la entrevista realizada por Matthew Carrozo al profesor Way (2017): «necesitamos traducciones humanas que sirvan como entrada a nuestros sistemas y así entrenarlos». Una vez ha procesado los datos, el sistema comienza a aprender patrones y a reproducirlos de manera que las traducciones sean de mayor calidad, es decir, se empieza a entrenar el motor de TA. Para el profesor, la clave está en «dejar que la máquina haga lo que sabe hacer realmente bien, y luego dejar que los elementos humanos se encarguen de la validación, la corrección y de la actualización de la fuente original».

Al contrario que la TAE, la traducción neuronal no traduce solo grupos de palabras, sino frases enteras, lo que hace que suene más natural y por lo tanto parezca una traducción humana. Además, la traducción estadística pierde calidad cuando trata de traducir a lenguas diferentes a las occidentales, como es el caso de las lenguas asiáticas.

Algunos ejemplos de *software* que ya utilizan este tipo de traducción son Tensor Flow, Open NMT, Baidu Translate y, el más conocido, Google Neural Machine Translation.

2.3.3.1. Aplicaciones de la TAN

En primer lugar, la TAN se puede aplicar a la traducción asistida por ordenador (TAO), con la que podemos realizar traducciones que son revisadas y corregidas posteriormente por un traductor humano durante la fase de posesición (PE).

Otra de las aplicaciones de la TAN es la traducción interactiva, en la cual el traductor influye en el proceso de traducción y va corrigiendo los errores a medida que van

saliendo de manera que el sistema aprende de ellos y mejora la calidad de la traducción en curso y de las que puedan realizarse en un futuro.

Finalmente, tenemos la traducción multimodal, donde una frase origen se traduce con la ayuda de una imagen relacionada con el texto a traducir (Caglayan, Barrault y Bougares, 2016, citado en Casacuberta; Peris, 2017).

2.3.3.2. Google Neural Machine Translation

En octubre de 2016 varios expertos llevaron a cabo una investigación sobre el funcionamiento del sistema de traducción neuronal creado por Google. Este informe, titulado *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation* (GNMT) (Wu, Schuster *et al*, 2016), recoge los problemas que plantea la TAN y cómo el nuevo sistema diseñado por Google pretende solucionarlos. Entre las conclusiones que obtuvieron, podemos destacar que: los expertos pudieron comprobar que en algunos de los documentos que tomaron como ejemplo la media de precisión y calidad de las traducciones de GNMT era igual a la de un traductor bilingüe, ya que el número de errores de traducción de este sistema se reduce en un 60% en comparación con otros sistemas de TA.

2.4. Criterios para evaluar un sistema de TA

Según Hutchins y Somers (1992), el principal aspecto que debemos tener en cuenta los traductores a la hora de evaluar un sistema de traducción automática es la cantidad de texto que tendremos que pre-editar y/o poseditar para obtener una traducción de calidad. Si para conseguir este objetivo tenemos que realizar muchas modificaciones antes y/o después de pasar el texto por TA, quiere decir que ese sistema no nos facilitará el trabajo, por lo tanto, no cumplirá el objetivo de ser una herramienta útil para los traductores. Para evaluar la calidad lingüística debemos tener en cuenta: la fidelidad, la inteligibilidad del texto y el estilo. Para realizar la comparación de los traductores automáticos de Google y Baidu hemos investigado dos opciones: la que propone eTransFair y las normas de calidad MQM (2015).

Siguiendo el modelo de eTransFair (2016-2019), en primer lugar debemos crear una tabla con diferentes columnas en la que veremos claramente la comparación entre los diferentes sistemas. A continuación, escogemos un fragmento de texto que creamos que puede ocasionar problemas a la hora de traducir y los pasamos por los diferentes motores que hemos seleccionado (en esta ocasión Google y Baidu). Copiamos los resultados de cada traductor automático en la tabla y realizamos la comparación. Los dos criterios a seguir son, por un lado, la adecuación, es decir, valorar qué cantidad del TO está presente en el TM y, por otro lado, indicaremos la tipología de cada uno de los errores que hemos encontrado (adición, omisión, cambio de orden, etc.).

La otra opción que hemos investigado son las normas MQM (*Multidimensional Quality Metrics*, 2015), las cuales hablan sobre los criterios de calidad que debe cumplir la TA. Estas normas ofrecen una definición clara y precisa de lo que es el control de calidad y ofrece una lista de categorización de errores detallada, con la definición y ejemplos de cada tipo. En lo referente a la calidad de la TA, tiene en cuenta los siguientes factores:

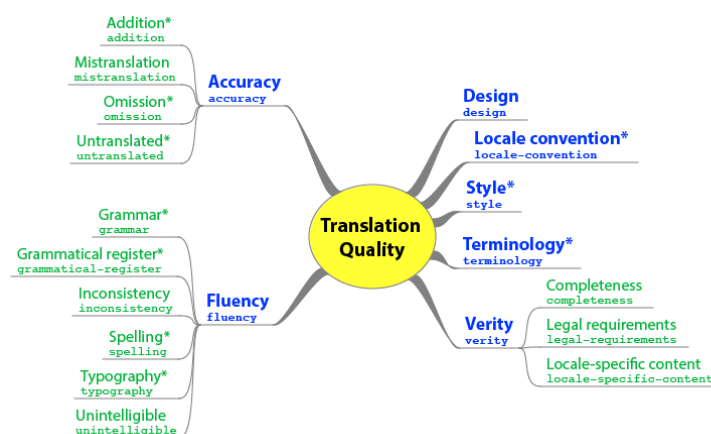


Fig. 4. Clasificación de errores según MQM (2015)

Tras examinar ambas propuestas detenidamente, nos hemos decantado por utilizar las normas MQM, ya que la información que ofrece con respecto a los criterios de calidad, cómo aplicarlos para evaluar un sistema de TA y los errores que esta puede cometer es mucho más amplia y detallada. Además, las definiciones de cada criterio y tipo de error están acompañadas de ejemplos que facilitan su comprensión para después poder identificarlos rápidamente.

Para este trabajo, nos centraremos en tres de esos criterios:

- a) **Fidelidad:** errores que hacen que el texto meta no represente las ideas del texto original.
- b) **Fluidez:** errores que impiden la clara comprensión del texto.
- c) **Terminología:** este tipo de error se da cuando se traduce un término de un ámbito específico con otro de un ámbito distinto.

Dentro de cada uno de estos criterios se tienen en cuenta diferentes tipos de error:

Fidelidad	Adición	El texto meta incluye texto que no está presente en el original.
	Mala traducción	El contenido del texto meta no se corresponde con el original.
	Omisión	En el texto meta se ha omitido contenido presente en original.
	No traducido	El contenido que debería haber sido traducido se ha dejado sin traducir.
Fluidez	Gramática	Errores relacionados con la gramática o la sintaxis del texto.
	Registro gramatical	El texto meta emplea un registro gramatical incorrecto, como por ejemplo hacer uso de pronombres o formas verbales informales cuando el texto requiere un grado de formalidad más alto.
	Inconsistencia	El texto presenta inconsistencias internas.

	Ortografía	Errores relacionados con la ortografía de las palabras.
	Tipografía	Errores relacionados con la presentación del texto. Esta categoría debe ser utilizada para cualquier tipo de error tipográfico.
	Ininteligibilidad	No se puede determinar la causa del error. Indica un error mayor en la fluidez del texto.
Terminología	Un término (palabra específica de un ámbito) se ha traducido con otro diferente al esperado para el ámbito de especialidad en cuestión.	

Algunas empresas deciden realizar una pre-edición o pre-traducción para facilitar las cosas al sistema de TA y así obtener mejores resultados. Esto se hace mediante el uso de lenguaje controlado, cuyo principal objetivo es adaptar las construcciones y vocabulario del TO para que el sistema de TA pueda producir un buen TM.

2.5. Problemas de la TA

En el presente, existen y se continúan desarrollando muchas herramientas de traducción automática para agilizar y economizar el proceso de traducción. No obstante, estas herramientas son un arma de doble filo, ya que nos pueden ser muy útiles, pero también pueden ensombreceer nuestro trabajo. Por eso debemos tener mucho cuidado cuando hagamos uso de ellas y revisar siempre a conciencia la traducción que obtengamos para evitar errores de gravedad. Entre los problemas más comunes de la traducción automática podemos encontrar los siguientes (Tradel Barcelona, 2017):

1. **Polisemia:** una palabra puede tener varios significados según el contexto en el que la encontremos. Este es un problema que los traductores automáticos no pueden resolver correctamente, ya que no tienen la capacidad de entender y diferenciar los contextos. Ejemplo:

Can I search your car? > ¿Puedo buscar tu coche?*

2. **Frases hechas:** cada país tiene sus propias costumbres y tradiciones, así como dichos, refranes y frases hechas. Este tipo de expresiones solo puede comprenderlas un nativo, ya que contienen muchos matices de esa cultura que una máquina no es capaz de descifrar. Ejemplo:

It's raining cats and dogs -> esta lloviendo gatos y perros* (Google Traductor)
> Llueve a cántaros

3. **Falsos amigos:** es un error muy común entre el inglés y el español, ya que las palabras que se parecen entre sí pueden llevar a interpretaciones erróneas. Ejemplo:

Exit > éxito* (Google Traductor)
> salida

4. **Nombres propios:** según la combinación de idiomas que escojas, los errores pueden ser más o menos graves. Los errores suelen ser bastante frecuentes cuando hay que traducir nombres propios, ya que no todos los idiomas saben reconocerlos. Un caso reciente tuvo lugar en Galicia, cuando Google Traductor tradujo:

La fiesta de Grelo > La fiesta del clítoris*

Otros errores famosos de Google Traductor son, por ejemplo, la traducción del nombre de Shakira por «maniática», Britney Spears por «Natalie Imbruglia», casco histórico por «historic helmet» o la traducción del Centro Botín de Santander como «Centro de Pillaje»¹.

¹Estos ejemplos han sido extraídos de: Tradel Barcelona (2017), Dixit (2013) y Justo (2018).

Según Gilbert y Chang (1963), en cuanto a la traducción automática del chino a cualquier lengua, esta trae consigo una dificultad que no presentan otras lenguas: los caracteres, ya que cada carácter representa un concepto o idea. Para poder utilizar un sistema de traducción automática del chino, sería necesario realizar una descripción codificada de todos los caracteres de una oración. Si tenemos en cuenta que en los periódicos chinos se utiliza una media de 4000 caracteres y en el chino literario y técnico entre 8000 y 12000 caracteres diferentes, esta tarea resulta prácticamente imposible, ya que tendría que ser un traductor humano quien introdujera todos los conceptos y ejemplos en la base de datos que después se utilizaría para la TA.

La estructura gramatical de la lengua china también puede provocar complicaciones. En chino, los caracteres que conforman una oración se escriben juntos, sin espacios. Esto supone un problema cuando nos encontramos con una palabra compuesta por más de un carácter, ya que no siempre es fácil identificar qué caracteres forman parte de una palabra y cuáles forman parte de otra. Un humano podría resolverlo fácilmente, pero una máquina no.

2.5.1. Problemas de la traducción automática neuronal

En la actualidad, el uso de la TAN está empezando a extenderse cada vez más, tanto entre los usuarios como en grandes empresas de todo el mundo. Esto se debe a dos factores principales: la rapidez y la eficacia, los cuales hacen que este sistema resulte mucho más rentable que otros sistemas de TA. Sin embargo, no es un sistema perfecto, la TAN tiene todavía mucho camino por recorrer y muchos aspectos mejorables. En la siguiente lista se muestran los principales retos y problemas de la TAN, recogidos en el artículo académico de Koehn y Knowles (2017).

- 1. Grado de especialización.** Muchos de los traductores automáticos actuales ofrecen traducciones fluidas pero que no tienen nada que ver con el original. Esto ocurre principalmente en contextos específicos, como el financiero o el legal, en los que la calidad de las traducciones es más baja. Esto se debe a la falta de entrenamiento de los motores en ámbitos especializados consecuencia de la falta de corpus, ya que cuanto menor es la cantidad de datos, menor es también la calidad de la traducción.
- 2. Calidad del corpus.** Además de por la falta de entrenamiento, otra de las razones por las que los motores de traducción ofrecen malas traducciones es la falta de corpus de calidad extraído de fuentes fiables. Esto es lo que provoca que a veces la TAN ofrezca resultados absurdos o sin sentido.

3. **Palabras poco comunes.** La TAN no es capaz de traducir palabras poco frecuentes, como por ejemplo los nombres propios de lugares o empresas poco conocidos que existen solo en algunas zonas del mundo. Esto es consecuencia de la falta de vocabulario de los motores de traducción automática.
4. **Frases largas.** La TAN encuentra dificultades tanto a la hora de analizarlas como a la hora de producirlas, lo que perjudica bastante el resultado final de las traducciones. Un claro ejemplo es la traducción jurídica, ya que es un ámbito donde predominan las frases largas y complejas.
5. **Orden de las palabras.** En ocasiones la TAN puede modificar el orden de las palabras de una oración, lo que podría ocasionar el cambio del sujeto por el objeto, cambios de sentido, etc.
6. **Control de calidad.** Los motores de traducción neuronal todavía no están suficientemente preparados para ofrecer y garantizar un buen control de calidad. Prueba de ello es la cantidad de errores que pueden encontrarse en cualquier TAN.

Traducción automática neuronal VS traducción automática estadística

A pesar del rápido avance y expansión de la TAN, muchos motores de traducción siguen trabajando con TAE. Pero, ¿cuál es el mejor de los dos? ¿De cuál podemos fiarnos más? La empresa báltica Tilde realizó un estudio comparativo en 2017 entre sus propios sistemas de TAE y TAN, Tilde Statistical MT (SMT) y Tilde Neural MT (NMT). En él, los traductores debían identificar errores de orden de las palabras, morfología, omisiones o partes no traducidas, adiciones y errores de léxico. Los resultados fueron los siguientes:

- a) La TAN ofrece mejores resultados en lo que a orden de las palabras, morfología, sintaxis y concordancia se refiere. Esto quiere decir que sus traducciones son más fluidas y precisas que los sistemas estadísticos. No obstante, la TAN comete más errores de léxico que la TAE.
- b) El porcentaje final de frases traducidas completamente correctas, sin ningún tipo de error, es del 35% en el caso de la TAN y del 25% en la TAE. Por lo tanto, y aunque aún necesita mejorar mucho, la TAN es el sistema de traducción automática que actualmente ofrece mejores resultados.
- c) Resaltan la importancia de la evaluación humana a la hora de comparar diferentes motores de traducción para así poder mejorar en el futuro.
- d) Asimismo, señalan las ventajas que posee la TAN: traducciones más fluidas, más humanas; mayor calidad y adecuación; tiene en cuenta el contexto; y es la mejor opción para trabajar con lenguas cuyo corpus es escaso y de baja calidad.

2.6. La posesición

La posesición consiste en la revisión y corrección de una traducción realizada por un sistema de traducción automática. Esta es la última fase del proceso llevado a cabo por la TA y es una fase imprescindible y demandada por la propia TA, ya que hoy en día la calidad de la mayoría de los traductores automáticos es muy deficiente.

TAUS (2016) trabajó en conjunto con otros profesionales de la traducción con la finalidad de proponer una serie de criterios para lograr una posesición de calidad. Antes de empezar a poseer, proponen tener en cuenta las siguientes recomendaciones:

1. Comprobar que el sistema se adecúa a las necesidades lingüísticas, es decir, debemos asegurarnos de que el sistema escogido para traducir de forma automática (TABR, TAE, TAN) es el ideal para lograr nuestros objetivos y no supondrá un obstáculo que dificulte nuestro trabajo.
2. Verificar que la calidad del texto original sea óptima y, si es posible, adaptarlo con técnicas de lenguaje controlado al sistema de TA que vamos a usar.
3. Integrar la gestión terminológica en todo el proceso.
4. Formar con antelación a los poseedores.
5. Evaluar el resultado de la traducción automática «en bruto» antes de negociar el precio de la posesición.
6. Definir la calidad final del texto que se va a poseer, basándose en el público al que va dirigido y en la aceptación por parte de éste.
7. Remunerar los comentarios estructurados de los poseedores con la finalidad de mejorar el sistema con el paso del tiempo.

Una vez se han aplicado las recomendaciones anteriores, TAUS prosigue con las pautas para conseguir una posesición de calidad, que estará determinada, en primer lugar, por la versión «en bruto» de la TA y, en segundo lugar, por la calidad esperada del contenido final. Según la calidad que deseemos alcanzar en nuestra posesición, podemos diferenciar entre las pautas a seguir para lograr una calidad «suficientemente buena» y las pautas para conseguir una calidad «similar o igual a la de una traducción humana». No obstante, y como se ha mencionado anteriormente, el uso de unas u otras pautas vendrá determinado por la calidad de la TA en bruto. Es decir, si la calidad de la TA es muy pobre, no podremos realizar una posesición rápida; en cambio, si la calidad de la TA es buena bastará con una simple revisión.

Si lo que deseamos es un resultado «suficientemente bueno», es decir, un texto final que sea comprensible y preciso pero inferior a nivel estilístico, las pautas a seguir son:

1. Intentar obtener una traducción correcta semánticamente.
2. Comprobar que no hay omisiones ni adiciones.

3. Suprimir todo tipo de contenido que pueda considerarse como ofensivo para el público al que va dirigido el mensaje.
4. Utilizar tanto como sea posible la traducción automática en bruto.
5. Aplicar las reglas básicas de ortografía.
6. No realizar correcciones de estilo.
7. No reestructurar oraciones para mejorar el flujo natural del texto.

Si por el contrario deseamos un resultado «similar o igual a la traducción humana» que sea comprensible, preciso y estilísticamente correcto, aunque no tan bueno como el que podría realizar un nativo, las pautas que debemos seguir son:

1. Intentar obtener una traducción correcta a nivel gramatical, sintáctico y semántico.
2. Comprobar que la terminología clave está traducida correctamente.
3. Comprobar que no hay omisiones ni adiciones.
4. Suprimir todo tipo de contenido que pueda considerarse como ofensivo para el público al que va dirigido el mensaje.
5. Utilizar tanto como sea posible la traducción automática en bruto.
6. Aplicar las reglas básicas de ortografía, puntuación y guiones.
7. Comprobar que el formato es correcto.

Estas indicaciones son las que seguiremos a la hora de realizar la versión poseditada de los textos en el capítulo 4 «Propuesta de posesición» para obtener un buen resultado final.

Algunos de los programas que podemos usar para poseditar son: MemoQ Translator Pro, SDL Trados Studio, Apertium Advanced Web Interface o Google Translation Toolkit, entre otros. Esta última es gratuita, permite a los traductores organizar sus trabajos gracias a las herramientas que posee, junto con el uso compartido de traducciones, glosarios y memorias de traducción. Se pueden subir y traducir documentos de Microsoft Word, OpenOffice, RTF, HTML, texto, artículos de Wikipedia y hasta hace poco Knols. Este sistema se retroalimenta de Google Traductor.

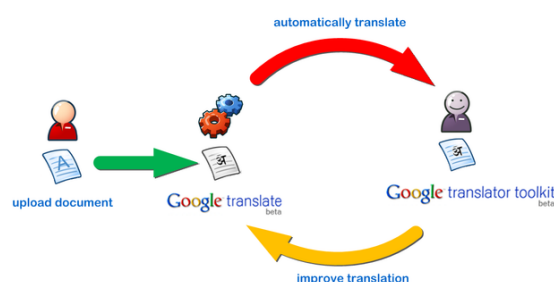


Fig. 5. Descripción gráfica del funcionamiento de Google Translator Toolkit
(Herramientas para traductores, 2012)

2.7. La traducción automática

En las últimas décadas, y gracias a las numerosas investigaciones llevadas a cabo en el campo de la traducción automática, el número de traductores automáticos en internet ha aumentado considerablemente. Podríamos elaborar una larga lista con todos ellos, no obstante, en este trabajo vamos a centrarnos en dos de los grandes: el Traductor de Google (líder a nivel mundial) y Baidu Translate (gigante asiático).

El Traductor de Google es, según la definición más reciente de Wikipedia (2018), «un sistema multilingüe de traducción automática, desarrollado y proporcionado por Google, para traducir texto, voz, imágenes o video en tiempo real de un idioma a otro». Lanzado en octubre de 2007 y con más de 200 millones de personas registradas, es un servicio multilingüe y gratuito, capaz de traducir a más de 100 idiomas diferentes. Hasta hace poco empleaba sistemas de traducción automática estadísticos, pero en 2016 incorporó el sistema de traducción automática neuronal. Por el momento, la TAN de Google es capaz de traducir los idiomas inglés, francés, alemán, portugués, español, chino, japonés y turco. Asimismo, puede funcionar también como diccionario multilingüe y es capaz de corregir sus propios errores. Este traductor también está disponible en forma de aplicación para sistemas iOS y Android.

No obstante, y a pesar de ser una herramienta que ha mejorado mucho y que cada vez utilizan más usuarios, Google Traductor también tiene limitaciones. La mayoría de las veces no es capaz de producir traducciones de calidad, ya que no cuenta con un corpus terminológico completo y tampoco sigue ninguna norma gramatical, motivo por el que muchas veces encontramos sinsentidos en sus traducciones. También sufre más o menos limitaciones según el idioma que traduce, es decir, los resultados pueden ser de una calidad mucho más alta cuando utiliza el inglés, ya que es el idioma que utiliza la Unión Europea, de donde proceden la mayoría de datos que almacena. Sin embargo, no ocurre lo mismo con lenguas más lejanas, por ejemplo, las lenguas asiáticas.

Al otro lado del mundo encontramos Baidu Translate. Baidu es el buscador más grande de la República Popular China, por lo que podríamos llamarlo el «Google chino», y en 2011 lanzó su propio traductor automático: Baidu Translate. Según podemos leer en Baidu Baike (2018), este traductor incluye hasta 28 idiomas, incluido el chino. El sistema que utilizaba Baidu hasta ahora era el método Hybrid machine translation (HMT), el cual incluía la traducción automática basada en reglas (TBR), la traducción automática basada en ejemplos y la traducción automática estadística (TAE) para aprovechar las ventajas de cada sistema y mejorar la calidad de la traducción final.

Tal y como hizo Google, Baidu también incorporó en 2015 la traducción automática neuronal (TAN) y dispone de una aplicación para dispositivos móviles de la que

podemos hacer uso en cualquier momento, incluso si no tenemos conexión a internet. Igual que Google, Baidu Translate es uno de los primeros STA en línea a gran escala basado en el *deep learning*. Cabe destacar también que en 2018 la empresa Baidu creó su primer traductor simultáneo de bolsillo. Su eficacia quedó probada en la presentación durante la conferencia EmTech Digital en San Francisco, Estados Unidos, en la que este traductor tradujo a la perfección una conversación entre Wu y Will Knight, es decir, la combinación lingüística chino-inglés, inglés-chino (Temple, 2018).

2.8. Situación en España

En 2016, el equipo de investigadores de ProjeTA, profesionales de varias universidades españolas, realizó un informe sobre «El uso de traducción automática y posesición en las empresas de servicios lingüísticos españolas». Este informe pretendía identificar el grado de implantación de la TA en empresas españolas, conocer la opinión de los traductores con respecto a la TA y las ventajas o inconvenientes que esta podía conllevar para así luego elaborar una serie de recomendaciones a la hora de incorporar la TA en el flujo de trabajo.

Este informe reveló que casi la mitad de las empresas encuestadas han implementado la TA en su flujo de trabajo, pero hacen muy poco uso de ella y muy pocas empresas disponen de un sistema de traducción automática propio. En cuanto al sistema empleado por cada empresa, sistema basado en reglas o estadístico, la preferencia es la traducción estadística. Más de la mitad de las empresas encuestadas (es decir, la mayoría) no utiliza nunca los servicios de TA, esto puede deberse principalmente a: la falta de confianza en la TA, ya que es un campo poco explorado por las empresas a pesar de su rápido desarrollo en los últimos años, y la reducción de las tarifas de traducción, ya que los traductores perciben un menor beneficio económico al tener parte del trabajo ya hecho. Otro motivo para no utilizarla es el tiempo que hay que dedicar a la posesición, es decir, la baja calidad resultante de la TA obliga a hacer siempre una revisión de la traducción y esto también aumenta los costes. En resumen, «implementar la TA [...] requiere sin duda una inversión en recursos económicos, tecnológicos y humanos, pero probablemente requiere también un cambio de mentalidad (2016:28)».

Lo mismo nos revela el «Estudio de viabilidad para la implantación de la traducción automática en la empresa VITAE (2015)», el cual reafirma la idea de que la traducción automática no consiste solo en implantar una herramienta más al proceso de traducción, sino que conlleva cambios en todo el proceso. Cada empresa es diferente, tanto en el tamaño, la cantidad de clientes, facturación, etc., por lo tanto, cada empresa deberá crear su propio sistema de traducción automática para adaptarlo a sus necesidades específicas. Se deben analizar una por una todas las ventajas e inconvenientes que un determinado sistema de TA podría tener en la productividad y calidad de una empresa. Este estudio resalta tres puntos clave en la implantación de la TA al proceso de traducción:

1. la traducción automática permite mejorar la productividad;
2. es necesario personalizar el sistema de TA (sistema basado en reglas o sistema estadístico); y
3. es imprescindible realizar una integración de la traducción automática con el resto de herramientas TAO para aumentar el rendimiento.

2.9. Situación en China

El concepto de traducción automática surgió en China en 1956, época de grandes proyectos tras la devastación de la guerra. China se unió a la cola de los países que ya investigaban esta cuestión: Estados Unidos, Reino Unido y la Unión Soviética. El primer sistema de traducción automática (1959) recogía la combinación ruso-chino y llevaba incorporados un diccionario con 2030 acepciones y 29 grupos de normas (Zhongjun, 2015). A continuación, se llevó a cabo el mismo procedimiento con la combinación chino-inglés. El primer sistema de TA comercial (TranStar) salió a la venta en 1988. Desde entonces, son muchos los nuevos sistemas de TA que han salido al mercado, todos con la combinación inglés-chino incluida (Fu, 1999).

En sus inicios, la investigación de la TA estaba subvencionada por el gobierno. Fue a partir de los 80 cuando las empresas empezaron a invertir en ella. La colaboración entre las instituciones y las empresas ha hecho que la TA avance rápidamente, así como la cooperación internacional. Esto ha posibilitado la creación de asociaciones, como la Chinese Information Association of China (CIAC), y seminarios sobre el tema, como el seminario anual China Workshop on Machine Translation (CWMT), en el cual los expertos se reúnen y analizan los últimos avances en TA desde el año 2005. Además de sistemas de TA, también se desarrollaron al mismo tiempo sistemas de evaluación que comprobaban el léxico, la morfología, la sintaxis y la semántica (Fu, 1999).

Los sistemas de TA más conocidos y utilizados en China son Yaxin CAT, diseñado por Beijing Orient Yaxin Software Technology Co., Ltd.; Transmate, diseñado por empresas chinas y basado en Trados, y Snowman CAT, diseñado por el sector privado de Foshan (Guangdong) (Liang, Wu y He, 2003).

Actualmente, la traducción automática en China es un mercado en expansión. No obstante, aún son necesarios muchos más estudios y pruebas de los sistemas de TA para que estos puedan satisfacer las demandas de los usuarios. Una de las líneas de investigación que se está siguiendo es el de la traducción automática neuronal, ya presente en los motores de traducción de muchas empresas a nivel internacional y nacional, como Baidu.

A pesar de que el español y el chino son dos de las lenguas más habladas en el mundo, en la actualidad no existen muchos estudios sobre esta combinación de lenguas en TA. Es más, si queremos realizar una TA del chino al español, muy probablemente se utilice una lengua pivote: el inglés. No obstante, empiezan a aparecer poco a poco los primeros estudios sobre TAN entre chino y español, como por ejemplo el llevado a cabo por Marta R. Costa-Jussà, David Aldón y José A. R. Fonollosa (2017).

3. COMPARACIÓN DE GOOGLE Y BAIDU

Tras investigar brevemente la historia y evolución de la traducción automática, y los casos español y chino en particular, en el siguiente capítulo vamos a comprobar qué motor de traducción de los dos más utilizados en esta combinación puede resultar de más utilidad para los estudiantes de chino como lengua C. Explicaremos detalladamente el procedimiento que vamos a seguir para analizar el corpus seleccionado y el grado de comprensión de los textos, así como para clasificar los errores encontrados. Por último, realizaremos la comparación final entre ambos motores a partir de los datos obtenidos.

3.1. Metodología del estudio

La metodología escogida para la realización de este estudio ha estado inspirada en la empleada anteriormente en el TFG de Viver (2018).

3.1.1. OBJETIVOS

Antes de empezar con el análisis es imprescindible definir con precisión todos los elementos que se van a utilizar, así como los objetivos que se desean alcanzar. Estos objetivos son los siguientes:

- Objetivo principal: ver qué motor de traducción automática (Google o Baidu) resulta más útil en la combinación chino-español.
- Identificar el grado de calidad de los textos una vez traducidos con TA, ya que es lo que determinará el grado de posesición.
- Detectar y categorizar los errores de TA según las normas MQM.
- Comprobar la terminología específica de cada texto.

3.1.2. SELECCIÓN DE LOS MOTORES DE TA

Para este trabajo hemos seleccionado los motores de Google Traductor y Baidu Translate, ya que la combinación lingüística escogida es chino-español y estos son los motores más importantes que trabajan con esta opción lingüística, Google a nivel internacional y Baidu en China. Tal y como se ha explicado en el capítulo anterior, tanto Google como Baidu trabajan ya con traducción automática neuronal. Uno de los puntos comunes que tienen ambos motores es que utilizan el inglés como lengua pivote para traducir del chino al español y viceversa, ya que las dos lenguas son muy diferentes entre sí y aún no se han desarrollado sistemas que hagan una traducción directa. Dada su presencia mundial, Google suele proporcionar resultados más fiables, pues está presente en casi todos los países del mundo. Sin embargo, Baidu solo trabaja en China, por lo tanto está más limitado.

3.1.3. CORPUS

El corpus que vamos a utilizar para comprobar la utilidad de estos motores de traducción está compuesto por un total de seis textos de temáticas diversas (año nuevo chino, caligrafía, cocina, la Gran Muralla y turismo), las que actualmente suscitan más interés tanto a estudiantes como a turistas extranjeros. De cada uno de los textos se ha seleccionado un fragmento que vamos a utilizar en este trabajo.

Para poder definir el grado de especialidad del corpus, es importante saber diferenciar qué es una temática especializada y una temática general. Para ello, nos va a resultar de gran ayuda la obra de Cabré (1992) sobre los lenguajes de especialidad.

Según Cabré, hablamos de temáticas no especializadas cuando el tema es común a todos los hablantes de una lengua y de mensajes especializados cuando se requiere de conocimiento específico previo para poder comprenderlo. Esto quiere decir que si el texto es comprensible para el público general, estamos ante un texto no especializado. En cambio, si es necesario que el público domine ciertos conocimientos para comprender un texto, entonces nos encontramos ante un texto especializado. Por esta razón, los principales usuarios de los lenguajes de especialidad son los especialistas, ya que poseen el conocimiento necesario y favorecen el intercambio comunicativo sobre un saber científico.

Por lo tanto, en nuestro caso podemos afirmar que los textos seleccionados no son especializados, ya que en todos los casos el registro y el vocabulario empleado son de fácil comprensión para cualquier tipo de público. No obstante, sí podemos afirmar que algunas temáticas son especializadas, como por ejemplo la caligrafía o la medicina tradicional china, ya que en ocasiones aparece vocabulario específico de ese ámbito.

3.1.4. CRITERIOS DE ANÁLISIS

Es muy importante definir cuáles van a ser los parámetros de análisis que van a determinar la calidad de las traducciones automáticas que obtengamos. Tal y como hemos explicado en el capítulo anterior, los criterios seleccionados para este trabajo se basan en la norma de calidad MQM (2015). Esta norma no solo nos proporcionará los criterios de calidad para una buena traducción automática, sino que también nos ayudará a clasificar los errores encontrados en las traducciones.

Con el objetivo de realizar un análisis visual en el que se vea claramente dónde se encuentran los errores, seguiremos una leyenda de colores en la que cada tipo de error tendrá un color distinto:

Fidelidad	Adición
	Mala traducción
	Omisión
	No traducido
Fluidez	Gramática
	Registro gramatical
	Inconsistencia
	Ortografía
	Tipografía
	Ininteligibilidad
Terminología	

3.1.5. ANÁLISIS DE LOS RESULTADOS

Estos van a ser los pasos principales a seguir en este capítulo. Una vez obtengamos los resultados de la TA, se realizará un trabajo más minucioso en relación a la clasificación de errores de TA y a la posterior comparación de ambos sistemas de TA.

Empezaremos con el análisis del grado de calidad del texto, en el cual se tendrán en cuenta la fidelidad, la fluidez y la terminología. Para ello, puntuaremos cada uno de estos factores en el fragmento seleccionado con una puntuación que oscilará entre 1 y 4 puntos, siendo 1 la puntuación más baja y 4 la puntuación más alta. De esta manera podremos comprobar con qué motor de TA podemos comprender más porcentaje de texto. Los resultados obtenidos se representarán en forma de gráficos (para ver las puntuaciones en forma de tabla, véase anexos I y II).

Una vez tengamos el grado de calidad, realizaremos un análisis más exhaustivo mediante la identificación y categorización de errores. Cada uno de los errores estará identificado con un color (véase subapartado anterior), lo que hará que los resultados sean más claros. Una vez analizados todos los errores de los seis textos, se representarán los resultados también por medio de gráficos (para ver el recuento total de errores en forma de tabla, véase anexos III y IV).

Una vez finalizada esta comparación, en el siguiente capítulo se realizará una propuesta de posesición de los textos. Esta se presentará en forma de tabla en la que podremos ver, además del original chino, las versiones dadas por Google y Baidu y la propuesta posesitada. La posesición estará basada en las normas de calidad propuestas por TAUS (2016), explicadas en el capítulo 2, apartado 2.5 («La posesición»).

3.2. Análisis de los textos

En este apartado procederemos con el análisis de los textos seleccionados para el trabajo. En primer lugar, veremos cuál es el grado de calidad del corpus según los criterios que propone Hurtado (1995:49-74) y a continuación realizaremos la categorización detallada de los errores encontrados según las normas MQM (2015).

3.2.1. Grado de calidad del texto

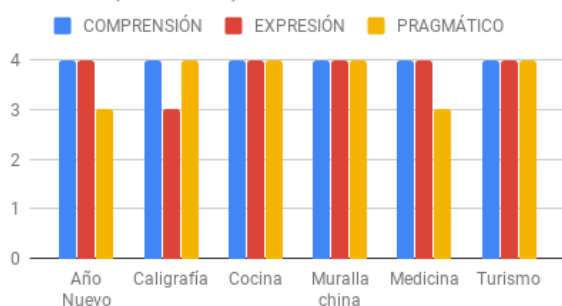
Para verificar la calidad de la TA primero debemos analizar qué porcentaje del texto podemos comprender y si nos servirá para poder entender el texto completo de forma correcta. Para analizar la calidad de las traducciones vamos a utilizar el baremo propuesto por Hurtado (1995:49-74), el cual se basa en la comprensión, la expresión y los elementos pragmáticos. Asignaremos una puntuación de entre 1 (nivel más bajo) y 4 puntos (nivel más alto) a los textos para realizar una comparación entre ellos.

	Nivel 1	Nivel 2	Nivel 3	Nivel 4
C O M P R E N S I Ó N	<p>No se entiende el texto original:</p> <p>Numerosos contrasentidos, distintos sentidos (matiz no reproducido, exageración, reducción, ambigüedad, poca precisión, connotaciones negativas), adiciones innecesarias y omisiones importantes.</p>	<p>Ambigüedad, matices no reproducidos y adiciones y omisiones reiteradas.</p>	<p>Contrasentidos poco comunes o inexistentes.</p> <p>Adiciones y omisiones ocasionales y poco importantes.</p>	<p>Sin contrasentidos y adiciones y omisiones que no repercuten negativamente en la comprensión del texto.</p>
E X P R E S I Ó N	<p>La expresión no es natural en la lengua de llegada, falta de riqueza expresiva.</p> <p>Errores de reformulación, ortográficos, de puntuación, gramaticales, sintácticos, morfológicos y léxicos (barbarismos, calcos, etc.).</p> <p>Repeticiones, falta de coherencia y cohesión, encadenamiento discursivo incorrecto y errores ortotipográficos.</p>	<p>Errores morfológicos y sintácticos.</p> <p>Repeticiones y calcos reiterados.</p>	<p>Errores de reformulación, gramaticales y léxicos poco comunes. Calcos esporádicos.</p>	<p>Buen encadenamiento discursivo, texto con coherencia y cohesión.</p> <p>Errores ortotipográficos ocasionales.</p>

P R A G M Á T I C O	Problemas derivados del contexto (intencionalidad, ironía, connotaciones), del destinatario, del encargo. Referencias culturales mal solucionadas, problemas metalingüísticos, inadecuación de variedad lingüística, dialecto y registro y descuidos.	Connotaciones diferentes y problemas metalingüísticos importantes.	Inexistencia de problemas derivados del contexto. Problemas metalingüísticos leves.	Adaptación adecuada al contexto, a los destinatarios, teniendo en cuenta la cultura de los mismos y la intención del texto.
--	---	--	--	---

Comenzaremos con el análisis de los títulos. Hemos decidido empezar el análisis por aquí porque el título de un texto es lo primero que leemos cuando nos entregan un texto, es el primer elemento que nos introduce de qué se va a hablar. Según el título que leamos, nuestra curiosidad puede aumentar o disminuir, por lo que es importante que esté bien traducido.

TÍTULOS (GOOGLE)



TÍTULOS (BAIDU)

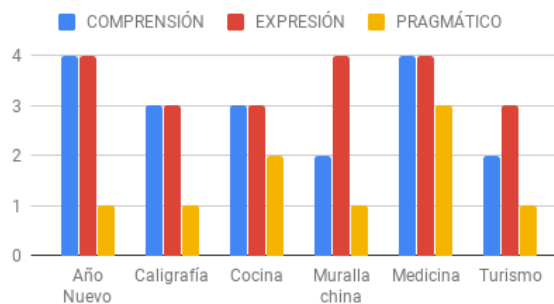


Gráfico 1. Grado de calidad de los títulos de Google y Baidu

Como podemos observar, el motor que resulta ser más fiable con solo leer el título es Google, ya que es el que mantiene los tres factores prácticamente en su totalidad. Con Baidu, además de que la calidad es visiblemente más baja, vemos que tiene más

dificultades a la hora de traducir los elementos pragmáticos del texto, lo cual es muy importante ya que entonces quiere decir que el entrenamiento de su motor de TA no es suficiente para poder transmitir los elementos culturales del TO al TM. Los textos en los que la calidad es más deficiente son los de caligrafía, cocina y turismo.

Después de analizar los títulos, vamos a proceder a hacer lo mismo con los fragmentos de los textos escogidos. En los siguientes gráficos aparecen las puntuaciones asignadas a los fragmentos traducidos primero con Google y después con Baidu:

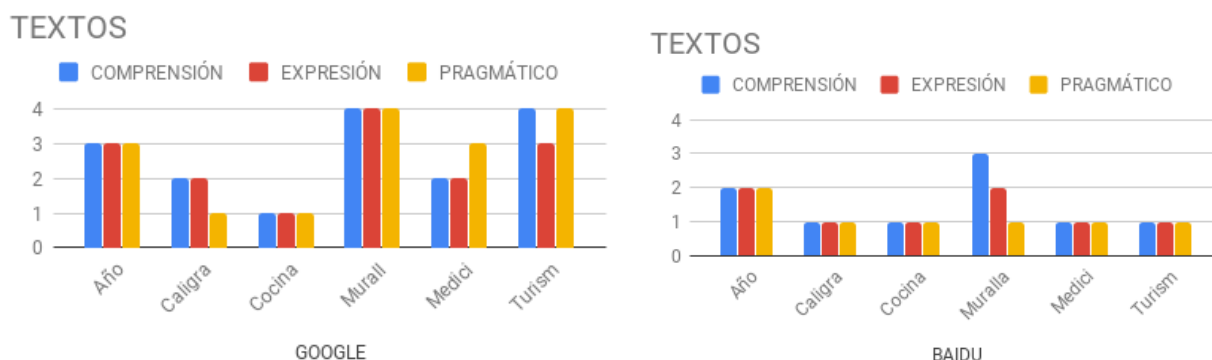


Gráfico 2. Grado de calidad de los textos de Google y Baidu

Los gráficos muestran claramente que las traducciones cuyo porcentaje de comprensión es más alto son las que proporciona Google. Hay temáticas más difíciles para ambos traductores, como la caligrafía y la cocina china. Esto se debe a que existe poco conocimiento (y por ende, material) sobre estas temáticas en español, por lo que los motores de traducción están poco entrenados y esto deriva en que la terminología específica no se traduce correctamente y el texto pierde todo sentido. Vemos también que el resultado en el texto del Año Nuevo Chino es más o menos el mismo en ambos casos. Sin duda, esta es una celebración muy conocida que cada año se celebra en muchas partes del mundo, no solo en China. En España, por ejemplo, el Año Nuevo Chino se celebra en varias ciudades importantes, como Madrid y Barcelona. Esto provoca que haya mucha más información sobre esta fiesta en muchos idiomas, así que el entrenamiento de los motores también es mayor.

Por lo tanto, si trasladamos los gráficos anteriores a una visión global, la media del grado de calidad de los textos de ambos motores es la siguiente:

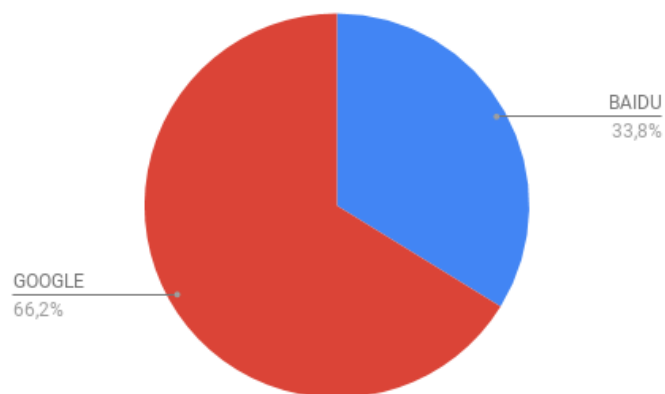


Gráfico 3. Comparación del grado de calidad de los textos de Google y Baidu

Tal y como habíamos visto hasta ahora, Google es el que ofrece un mayor grado de calidad en los textos que traduce. Así, ahora sabemos que Baidu dificultará mucho más nuestro trabajo si deseamos incorporar la TA como una herramienta más en nuestro flujo de trabajo.

3.2.2. Detección y categorización de errores

En la siguiente tabla se exponen las versiones ofrecidas por los traductores automáticos Google Traductor y Baidu Translate, así como la versión original en chino, para poder hacer una comparación global con las tres opciones a mano. Primero aparece el título del texto traducido y después el fragmento del texto seleccionado con los errores marcados.

TEXTO	ORIGINAL	GOOGLE	BAIDU
Año Nuevo (春节)	春节的历史	La historia del Festival de Primavera.	Historia de la primavera
ERRORES		Terminología	Mala traducción
	<p>年的传说</p> <p>除了对神明和祖先的庆祝，春节的庆祝也与传说和神话故事有着很大的关联。最受欢迎的神话故事便是怪兽年以及一年的结束。按照中国的传说，年是一个猛兽，会在一年中的每一天吞噬动物与人类。后来人们发现，这个年复一年的猛兽可以被红色驱赶，所以便在新年的前两个星期大量使用红色装饰。现在，红色被认为是吉祥的颜色，在</p>	<p>Leyenda del año</p> <p>Además de la celebración de los dioses y antepasados, la celebración del Festival de Primavera también está estrechamente relacionada con las leyendas y los mitos. El mito más popular es el año del monstruo y el fin del año. Según la leyenda china, el año es una bestia que engullirá animales y humanos todos los días del año. Más tarde, se descubrió que las bestias del año y el año podían ser expulsadas por el rojo, de modo que las decoraciones rojas se usaron en grandes cantidades durante las dos primeras semanas del Año Nuevo. El rojo ahora se considera de color auspicioso y se usa para</p>	<p>La leyenda del año</p> <p>Además de las celebraciones de los dioses y los ancestros, la celebración de la primavera está muy relacionada con las historias de mitos y mitos. La historia del mito más popular es el año de la bestia y el final de un año. Según la leyenda de China, el año es una bestia que consume animales y seres humanos cada uno de los días de un año. Posteriormente, se descubrió que la bestia de un año de edad podía ser expulsada de rojo y, por lo tanto, en las dos primeras semanas del año nuevo se utilizó un gran número de decoraciones rojas. Ahora, el rojo se considera un color de la mascota que se utiliza en el Festival de la primavera para la ropa y los fuegos artificiales. Se considera que el rojo puede expulsar a la bestia y</p>

	春节期间用于服饰和烟花爆竹。红色被视为可以驱赶怪兽年，并为即将到来的新的一年带来好运。	disfraces y fuegos artificiales durante el Festival de Primavera. El rojo es visto como una forma de ahuyentar el Año de los Monstruos y traer buena suerte al próximo año nuevo.	traer suerte para el próximo año que viene.
ERRORES		Mala traducción; Terminología; Gramática; Inconsistencia; Adición;	Mala traducción; Tipografía; Gramática; Inconsistencia; Terminología
Caligrafía (书法)	中国书法造诣最深的十大家，不可不知	Las diez personas más famosas de la caligrafía china, debemos saber.	Los (diez calígrafos) más profundos de la historia de China.
ERRORES		Gramática; Tipografía	Omisión; Terminología; Tipografía
Caligrafía (书法)	东晋著名书法家 “书圣”——王羲之 王羲之是中国最著名的、成就最高的书法家，他广采众长，自成一家，不仅作为中国书法史上一座极具象征性的丰碑深远影响着中国千年书法，还留下了大量的丰富多样的书法精品。书法兼善隶、草、楷、行各体，精研体势，心摹手追，广采众长，备精诸体，冶于一炉，摆脱了汉魏笔风，自成一家，	Famoso calígrafo de la dinastía Jin del este "Libro de San" - Wang Xizhi Wang Xizhi es el calígrafo más famoso y consumado de China. Tiene una larga historia y es una familia autónoma. No solo tiene un monumento simbólico en la historia de la caligrafía china, sino que también afecta profundamente a la caligrafía china de (hace) mil años. También dejó una gran cantidad de caligrafía rica y variada. Boutique Caligrafía y buenos ligamentos, pasto, grilletos y cuerpos diversos, investigación intensiva, confesión, persecución, a largo plazo, bien preparados,	La famosa escritora (de la dinastía Jin del este). "Santo" (Wang Xizhi) El rey, que es el mejor libro de la historia de China, es el más famoso y el más alto de los éxitos de China. La ley y la buena disposición, la hierba, el pasto, las líneas, los cuerpos, el poder de investigación, la persiguiendo, la gran audiencia, la preparación de los cuerpos, la metalurgia del horno, se ha librado de la letra de hinwei y se ha convertido en una familia de gran alcance.

	影响深远。	listos, fundidos en un horno, deshacerse del estilo Han y Wei, autocontenido y de gran alcance.	
ERRORES		Mala traducción; Tipografía; Ininteligibilidad; Omisión; Terminología	Inconsistencia; Mala traducción; Tipografía; Ininteligibilidad; Omisión
Cocina (中国菜)	中国的四大菜系分别是哪些?	¿Cuáles son las cuatro principales cocinas chinas?	¿Cuáles son los cuatro menú de China?
ERRORES			Terminología
Cocina (中国菜)	鲁 菜 也称山东菜，中国四大菜系之一。鲁菜取材广泛，选料精细，讲究丰满实惠，烹调方法全面，精于制汤，善以葱调味。鲁菜在烹制海鲜上有独到之处，尤其对海珍和小海味的烹制，堪称一绝。代表有：葱烧海参、烩乌鱼蛋、蟹黄鱼翅、德州扒鸡、奶汤核桃肉等。	Lu Cai También conocida como la cocina de Shandong, (es) una de las cuatro principales cocinas de China. Lu Cai toma una amplia gama de materiales , una excelente selección de materiales , presta atención a la plenitud y los beneficios , a los métodos de cocción completos , a la refinada sopa, al buen condimento con cebollas. La cocina Lu tiene su propio sabor único en la cocina de mariscos, especialmente para la (cocina de Haizhen y Xiaohaiwei) . Los representantes incluyen: cebolla, pepino de mar, calamares, calamares, faisán de Texas, faisán de Texas, sopa de leche, carne de nuez, etc.	Remolacha China es uno de los cuatro menú de China. La remolacha se basa en una amplia gama de materiales selectivos , la buena cocina, la buena cocina, la precisión de la sopa y el sabor a la cebolla . La comida de los mariscos para cocinar, especialmente para la cocina de Jen y la pequeña sabor de mar, es una buena cocinera. Los representantes son los siguientes: alambre de cebolla, huevos de risotto, aletas de cangrejo, carteristas de Texas, sopa de nueces de nueces, etc.

ERRORES		Terminología; Mala traducción; Gramática, Omisión; Ininteligibilidad; No traducido	Terminología; Ininteligibilidad; Gramática; Tipografía; Mala traducción; Inconsistencia
Historia (长城)	世界文化遗产——长城	Patrimonio Cultural de la Humanidad – (La) Gran Muralla	Patrimonio Mundial (Cultural de la Humanidad)- la gran ciudad
ERRORES		Omisión	Mala traducción; Terminología; Omisión
Historia (长城)	<p>约公元前 220 年，一统天下的秦始皇，将修建于早些时候的一些断续的防御工事连接成一个完整的防御系统，用以抵抗来自北方的侵略。在明代（公元 1368～1644 年），又继续加以修筑，使长城成为世界上最长的军事设施。它在文化艺术上的价值，足以与其在历史和战略上的重要性相媲美。</p> <p>概况</p> <p>长城是中国也是世界上修建时间最长、工程量最大的一项古代防御工程。自公元前 七八世纪开始，延续不断修筑了 2000 多年，分布</p>	<p>Alrededor del 220 a. C., Qin Shihuang, quien dominó el mundo, conectó algunas de las fortificaciones intermitentes construidas anteriormente en un completo sistema de defensa para resistir la agresión del norte. En la dinastía Ming (1368-1644 dC) se continuó construyendo para hacer de la Gran Muralla la instalación militar más larga del mundo. Su valor en la cultura y el arte es comparable a su importancia histórica y estratégica.</p> <p>Visión general</p> <p>La Gran Muralla es un antiguo proyecto de defensa con el tiempo de construcción más largo y la mayor cantidad de ingeniería en China (y del mundo entero). Desde los siglos VII y VIII aC, se ha construido de forma continua durante más de 2.000 años.</p>	<p>En los 220 años anteriores a la Yu, el emperador Qin, el emperador Qin, construirá una serie de fortificaciones intermitentes que se conectarán a un sistema de defensa completo para resistir a la agresión del Norte. La construcción de la gran ciudad se ha convertido en la más larga de las instalaciones militares del mundo en la generación mañana (1368 a 1644 d. C.). Su valor en el arte y el arte culturales es comparable a su importancia histórica y estratégica.</p> <p>Generalidades</p> <p>La gran ciudad es una de las obras más antiguas de defensa de China, la mayor de las cuales se construyó en el mundo. A partir del siglo VII (y VIII a. C.) antes del dólar de los Estados Unidos, la construcción de más de 2.000 años, que se remonta a más de 50.000 kilómetros de tierra en el norte y el Centro de China, se ha mantenido en un total de más de 50.000 kilómetros, lo que se denomina más de 2.000 años y</p>

	于中国北部和中部的广大土地上，总计 长度达 50000 多千米，被称之为“上下两千多年，纵横十万余里”。如此浩大的工程不仅 在中国就是在世界上，也是绝无仅有的，因而在几百年前就与罗马斗兽场、比萨斜塔等列 为中古世界七大奇迹之一。	Se distribuye en vastas tierras en el norte y centro de China, con una longitud total de más de 50.000 kilómetros. Se denomina dos mil años arriba y abajo, y tiene más de 100.000 millas . Un proyecto tan vasto no solo está en China sino también en el mundo, y es único. Por lo tanto, fue catalogado como una de las Siete Maravillas de la Edad Media con el Coliseo y la Torre Inclinada de Pisa hace cientos de años.	más de más de 100.000 kilómetros . Este gran proyecto, no sólo en China, sino también en el mundo, es uno de los siete milagros del mundo central , que hace cientos de años se enmarcó en el campo de los arcos romanos, la torre de pizza, etc.
ERRORES		Tipografía; Mala traducción; Terminología; Ininteligibilidad; Gramática; Omisión	Tipografía; Terminología; Gramática, Ininteligibilidad; Mala traducción; Omisión; Adición
Medicina (中医)	中医的历史	Historia de la medicina china	Historia de la medicina china
ERRORES		Terminología	Terminología
Medicina (中医)	在 3000 多年前的殷商甲骨文中，中国已经有关于医疗卫生以及十多种疾病的记载。周代已经使用望、闻、问、切等诊病方法和药物、针灸、手术等治疗方法。秦汉时期，形成了《黄帝内经》这样具有系统理论的著作。此书是现存最早的一	En los huesos del oráculo del (Yin Shang) , hace más de 3.000 años, China ha registrado atención médica y de salud y más de 10 enfermedades. La dinastía Zhou ha utilizado (los cuatro) métodos de diagnóstico, tratamiento, acupuntura y cirugía (como método de cura) . En las dinastías Qin y Han se formó el libro Sistema del Emperador Amarillo . Este libro es el clásico teórico	En China, hace más de 3.000 años, se ha documentado la atención de la salud y más de una docena de enfermedades. En el transcurso de las semanas se han utilizado métodos de tratamiento, tratamiento y tratamiento (como método de cura) , así como tratamiento médico, acupuntura, cirugía, etc. La época de qinhan fue una obra de teorías sistemáticas como el menú del rey Huang . El libro es el primero de los libros clásicos teóricos de la medicina china. La

	部中医理论性经典著作。张仲景所著的《伤寒杂病论》，专门论述了多种杂病的辨证诊断、治疗原则，为后世的临床医学奠定了发展的基础。汉代外科学已具有较高水平。据《三国志》记载，名医华佗已开始使用全身麻醉剂“麻沸散”进行各种外科手术。	más antiguo de la medicina china. El "Tratado sobre las enfermedades febriles" de Zhang Zhongjing analiza específicamente los principios del diagnóstico (y cura) de diferenciación del síndrome y el tratamiento de varias enfermedades diversas, lo que sentó las bases para el desarrollo de la medicina clínica posterior. La cirugía de la dinastía Han tiene un alto nivel. De acuerdo con los registros de los "Tres Reinos", el famoso doctor Hua Tuo ha comenzado a usar el anestésico general "Ma Bo San" para varias operaciones quirúrgicas.	teoría del diagnóstico y tratamiento de múltiples enfermedades (de Zhang Zhong Jing), que se centra en el diagnóstico y el tratamiento de múltiples enfermedades (misceláneas), ha sentado las bases para el desarrollo de la medicina clínica en las generaciones venideras. Las cirujanas de generación Han han alcanzado niveles elevados. De acuerdo con los tres países, el doctor Wallace ha comenzado a realizar diversas operaciones quirúrgicas de "ebullición" de anestésicos de todo el cuerpo.
ERRORES		Omisión; Ininteligibilidad; Tipografía; Mala traducción; Gramática; Terminología	Omisión; Inconsistencia; Ininteligibilidad; Mala traducción; Gramática; Terminología; Tipografía
Turismo (旅游业)	三亚后海：探寻隐世渔村的浪漫情调	Houhai en Sanya: explorando el ambiente romántico del pueblo de pescadores escondido	El romanticismo en la búsqueda de una aldea de pesca invisible
ERRORES			Omisión; Tipografía; Mala traducción
Turismo (旅游业)	后海村，又称藤海渔村，藏于山海之间，就连许多三亚本地人都未曾踏足此地。位于该村东北侧的蜈支洲岛，每年接待着上百万名来自世	El pueblo de Houhai, también conocido como el pueblo de pescadores de Tenghai, está escondido entre las montañas y el mar, y muchas personas locales en Sanya nunca han pisado este lugar. Ubicada en el lado	La aldea de atrás, conocida también como la aldea de pescadores del Mar de Kato, se esconde entre los mares de montaña, y ni siquiera muchos de los tres subnacionales de la población local han pisado el lugar. La isla del CICAA, situada al nordeste de la

	<p>界各地的游客，与渔村的静谧安逸形成鲜明对比。后海村虽默默无闻，但其绝美风光曾一度通过电视画面走进千家万户——中央电视台2000 年拍摄千禧年第一缕阳光的地方是三亚后海。</p>	<p>noreste del pueblo, la isla de Wuzhizhou recibe a millones de visitantes de todo el mundo, en marcado contraste con la tranquilidad del pueblo de pescadores. Aunque se desconoce el pueblo de Houhai, su hermoso paisaje una vez ingresó a las casas a través de las pantallas de televisión. La filmación del primer rayo de sol del milenio por parte de CCTV en el 2000 es conveniente para Sanya Houhai.</p>	<p>aldea, recibe cada año a millones de turistas de todo el mundo, lo que contrasta con el silencio de las aldeas de pesca. En la televisión central, en el año 2000, en los primeros días de la luz del Milenio, fue el mar de atrás.</p>
ERRORES		Terminología; Gramática	Mala traducción; Terminología; Tipografía; Ininteligibilidad; Inconsistencia

3.3. Análisis de resultados

Tras la identificación y categorización de los errores de TA encontrados en el corpus, en el siguiente apartado analizaremos con detalle las posibles causas y cómo ha influido la TAN en estos errores según las dificultades que plantean Koehn y Knowles (2017), mencionadas en el subapartado 2.4.1 del capítulo 2. Como soporte a nuestro análisis, hemos utilizado las hojas de cálculo de la plataforma Google Drive para recopilar y exponer de forma gráfica estos resultados. Para crear los gráficos, hemos realizado un recuento de todos los errores, el cual puede encontrarse en forma de tabla en los anexos III y IV.

En primer lugar, veamos qué traductor ha cometido más errores y en qué textos se han cometido. Este dato nos servirá como punto de partida de nuestro análisis de resultados. Podemos observar claramente la diferencia de errores entre ambos traductores: Google ha cometido un total de 106 (43,8%) y Baidu 136 (56,2%); Google ha cometido más errores en los textos de cocina y medicina tradicional china y Baidu en los textos de caligrafía y medicina tradicional china.



Gráfico 4. Número global de errores de cada traductor

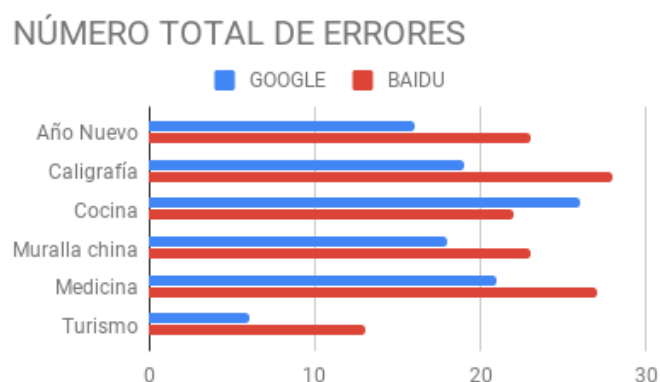


Gráfico 5. Número total de errores encontrados en cada texto según el traductor

Ambos traductores coinciden a la hora de cometer errores en el texto sobre medicina tradicional china. Esto podría deberse a dos motivos principales: por un lado, el desconocimiento sobre el tema aquí en España (hecho que lo convierte en más específico aún) y, por otro, la lejanía. El desconocimiento entre ambas lenguas y culturas (chino-español) hace que exista poca bibliografía en español y, si a esto sumamos las deficiencias de la TAN a la hora de extraer corpus de calidad de la red, de mantener el grado de especialización y de traducir vocabulario poco frecuente, el resultado final de las traducciones es muy bajo.

Por ejemplo, ¿quién podría conocer en España el libro «Canon de medicina interna del emperador Amarillo» sobre medicina tradicional china o la historia del anestésico *Mafeisan*? Muy pocos. Por esta razón, tanto Google como Baidu han ofrecido propuestas erróneas de traducción:

ORIGINAL	GOOGLE	BAIDU
《黄帝内经》	el libro " <u>Sistema del Emperador Amarillo</u> ".	obra de <u>teorías sistemáticas como el menú del rey Huang.</u>
麻沸散	anestésico general " <u>Ma Bo San</u> "	operaciones quirúrgicas de " <u>ebullición</u> " de anestésicos de todo el cuerpo.

Por el contrario, podemos observar que el texto en el que se han cometido menos errores es el de turismo. Se trata de un texto descriptivo sobre un pueblo costero de China que lo que pretende es atraer turistas, por lo que no es un texto con grandes dificultades para el lector chino. Esta es la principal razón de que se hayan cometido tan pocos errores: cuanto más sencillo y estándar es el original, más fácil lo tienen los motores de traducción para ofrecer mejores traducciones. Otro detalle que nos ha llamado la atención es que hemos encontrado varias palabras que forman parte del uso del español de Latinoamérica, en concreto en la versión de Google. En este caso, el fallo ha sido el control de calidad, ya que la versión que ofrece del español no distingue entre España y América Latina. Esto podría ser una consecuencia del hecho de que existen más hablantes latinos que españoles, por lo que la cantidad de input español es menor que la latina.

ORIGINAL	GOOGLE	BAIDU
位于该村东北侧的蜈支洲岛	<u>Ubicada</u> en el lado noreste del pueblo, la isla de Wuzhizhou [...]	La isla del CICAA , situada al nordeste de la aldea, [...]
中央电视台2000年拍摄千禧年第一缕阳光的地方是三亚后海。	La <u>filmación</u> del primer rayo de sol del milenio por parte de CCTV en el 2000 es conveniente para Sanya Houhai.	En la televisión central, en el año 2000, en los primeros días de la luz del Milenio, fue el mar de atrás.

En los siguientes gráficos aparecen representados los datos sobre la cantidad de errores de cada motor de traducción según la clasificación MQM:

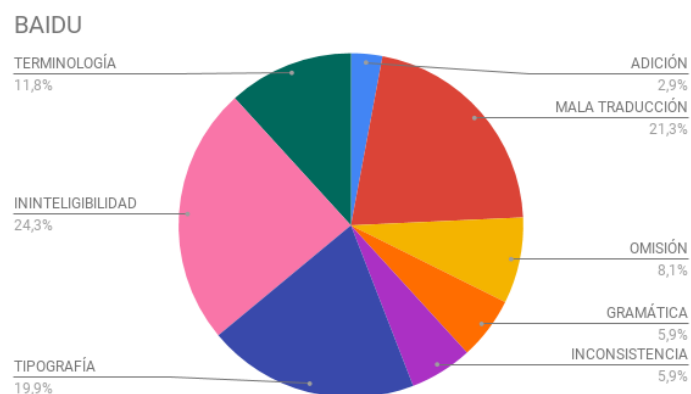


Gráfico 6. Representación detallada de los errores cometidos por Baidu

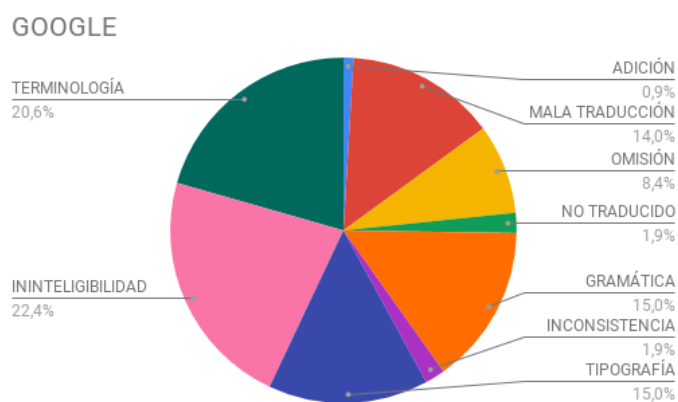


Gráfico 7. Representación detallada de los errores cometidos por Google

El error que más ha cometido Baidu Translate ha sido el de «ininteligibilidad», lo que significa que el 24,3% de sus traducciones era del todo incomprensible. En segundo lugar está el de «mala traducción», con un 21,3%. Si sumamos ambos porcentajes, podemos determinar que casi la mitad de las traducciones (45,6%) no hemos podido aprovecharlas y han tenido que poseditarse. En tercer y cuarto lugar encontramos los errores de tipografía (19,9%) y de terminología (11,8%).

Los errores más cometidos por Google son también el de «ininteligibilidad», con casi el mismo porcentaje (22,4%), y el de «mala traducción» (14%), un porcentaje bastante más bajo que el de Baidu. Si los sumamos, descubrimos que el porcentaje de traducciones que hemos tenido que poseditar es del 36,4%, una cifra muy por debajo de la de Baidu.

La principal causa de esta gran diferencia entre ambos la encontramos en la falta de entrenamiento del traductor automático de Baidu. El alcance de Baidu está limitado al continente asiático, mientras que Google Traductor llega a cualquier parte del mundo. Esto hace que su corpus sea mucho más extenso y esté actualizado, lo que significa que el traductor se sigue entrenando continuamente.

A continuación, comentaremos con más detalle cada tipo de error y sus causas. Comenzaremos con los errores más frecuentes (mala traducción, tipografía, ininteligibilidad y terminología) y después pasaremos a los menos frecuentes (adicción, omisión, no traducido, gramática, registro gramatical, inconsistencia y ortografía):

a) Mala traducción

Las malas traducciones son una consecuencia de las diferencias tanto en aspectos lingüísticos como culturales entre China y España. Pueden estar provocadas por muchas razones, desde la estructura del texto hasta el léxico. En este caso, podemos ver claramente que Baidu es el que ha obtenido peores resultados (65,9%, ver gráfico 8), lo que podría estar motivado, como ya hemos dicho antes, por la falta de entrenamiento de su motor de traducción en la combinación chino-español. Un ejemplo muy frecuente de mala traducción se da con la transcripción de los nombres propios chinos al español. Aunque este problema se está solucionando poco a poco, todavía encontramos casos en los que o bien no se ha transcrito el nombre correctamente o directamente el traductor automático no lo identifica como nombre propio y lo traduce palabra por palabra, provocando cambios de sentido o sinsentidos. Así, cuando hablamos de la «Leyenda de Nian» (texto 1), en referencia a la historia del año nuevo chino, tanto Google como Baidu lo traducen como «La leyenda del año», ya que ambos traducen literalmente la palabra *Nian* (pinyin del carácter 年, que significa «año») y no interpretan que en este caso se refiere a un nombre propio, por lo que deberíamos dejar el pinyin *Nian*. Lo

mismo ocurre con la Gran Muralla, la cual Baidu traduce como «la gran ciudad», y con *Lu Cai*, la cocina típica de Shandong que Baidu traduce como «remolacha». Estos ejemplos son una clara muestra de que Baidu no es capaz de identificar los nombres propios y casi siempre opta por traducirlos literalmente, por lo que el resultado muchas veces no tiene sentido.

En las malas traducciones también influyen las deficiencias de la TAN (véase capítulo 2, subapartado 2.4.1). Estas pueden provocar que una frase, expresión o término se traduzcan incorrectamente. En el siguiente ejemplo, vemos un caso de mala traducción de Baidu que incluso cambia el sentido del original.

ORIGINAL	GOOGLE	BAIDU
春节的历史	La historia del Festival de Primavera	<u>Historia de la primavera</u>

En el siguiente párrafo encontramos una expresión conocida únicamente en China para referirse a la Gran Muralla, por lo que, si para cualquier traductor no nativo es un problema difícil de resolver, para un traductor automático la dificultad aumenta considerablemente, ya que al no existir traducción al español es el traductor quién debe hacer una propuesta. Este ha sido el resultado de ambos traductores:

ORIGINAL	GOOGLE	BAIDU
[...] 总计长度达50000多千米，被称之为“ <u>上下两千多年，纵横十万余里</u> ”。	[...] con una longitud total de más de 50.000 kilómetros. Se denomina " <u>dos mil años arriba y abajo, y tiene más de 100.000 millas</u> ".	[...] se ha mantenido en un total de más de 50.000 kilómetros, lo que se denomina " <u>más de 2.000 años y más de más de 100.000 kilómetros</u> ".

Las frases largas también propician la aparición de malas traducciones, ya que estas son uno de los grandes retos de los traductores automáticos neuronales. Aquí podemos observar que Baidu, por ejemplo, ha omitido prácticamente toda la información. Google, a pesar de cometer otro tipo de error, mantiene toda la información.

ORIGINAL	GOOGLE	BAIDU
王羲之是中国最著名的、成就最高的书法家，他广采众长，自	Wang Xizhi es el calígrafo más famoso y consumado de China. Tiene una larga historia y es una familia	El rey, que es el mejor libro de la historia de China, es el más famoso y el más alto de los éxitos de

成一家，不仅作为中国书法史上一座极具象征性的丰碑深远影响着中国千年书法，还留下了大量的丰富多样的书法精品。	autónoma. No solo tiene un monumento simbólico en la historia de la caligrafía china, sino que también afecta profundamente a la caligrafía china de mil años. También dejó una gran cantidad de caligrafía rica y variada.	China.
---	---	--------

MALA TRADUCCIÓN

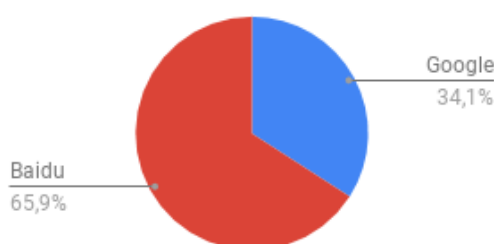


Gráfico 8. Representación del número de errores de mala traducción de ambos traductores

b) Tipografía

La puntuación también es diferente en chino y en español, pero es algo que no siempre detecta la TA. Mientras analizábamos los textos del corpus, hemos observado que Google sí ha sabido adaptar la puntuación empleada en el original chino a la versión española, pero Baidu en la mayoría de los casos mantiene la china. Por ejemplo, en chino no añadimos un espacio después de punto o coma, ya que este aparece directamente. Al traducir al español, Google sí ha añadido los espacios necesarios después de poner punto o coma, pero Baidu no. Otro ejemplo es el uso de las comillas: tanto Google como Baidu mantienen las comillas inglesas en vez de emplear las comillas latinas.

ORIGINAL	GOOGLE	BAIDU
春节的庆祝也与传说和神话故事有着很大的关联。最受欢迎的神话故事便是怪兽年以及一年的结束。按照中国的传说，年是一个猛兽，会在一年	[...] mitos. El mito más popular es el año del monstruo y el fin del año. Según la leyenda china, el año es una bestia que engullirá animales y humanos todos los días del año. Más tarde, [...]	[...] mitos. La historia del mito más popular es el año de la bestia y el final de un año. Según la leyenda de China, el año es una bestia que consume animales y seres humanos cada uno de los días de un año. Posteriormente, [...]

中的每一天吞噬动物与人类。后来人们发现, [...]		
“上下两千多年, 纵横十万余里”。	"dos mil años arriba y abajo, y tiene más de 100.000 millas".	"más de 2.000 años y más de más de 100.000 kilómetros".

Los errores tipográficos son fallos que con una buena revisión final quedan solucionados, por lo que, si vemos que no es así, quiere decir que el control de calidad no es bueno. Este es otro de los grandes problemas de la TAN, y prueba de ello es la cantidad de errores de tipografía encontrados, sobre todo en el caso de Baidu (80%).

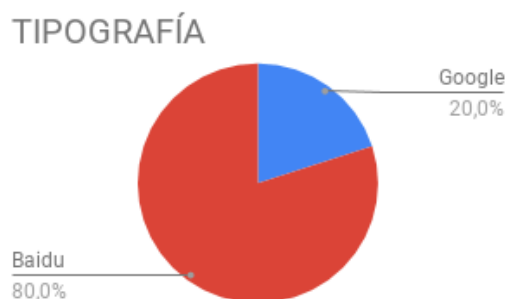


Gráfico 9. Representación del número de errores de tipografía de ambos traductores

c) Ininteligibilidad

El traductor de Google está mejor entrenado como traductor automático que Baidu. Además, al ser un motor utilizado en prácticamente todo el mundo, el corpus que posee es mucho más amplio que el de Baidu. Esto hace que Google disponga de muchas más traducciones anteriores de expresiones, palabras y frases, lo que disminuye la probabilidad de que cometa errores. Con Baidu ocurre lo contrario: su falta de entrenamiento como traductor y su reducido corpus hacen que la mayoría de las veces sus traducciones sean casi del todo incomprensibles. Así lo demuestra el gráfico (véase gráfico 10), aunque el porcentaje de Google también es bastante alto. Este dato revela que, aunque las traducciones de Google funcionan mejor, la TA del chino al español sigue siendo todo un reto. Esta falta de comprensión de las traducciones también es una muestra de las dificultades de la TAN a la hora de traducir frases largas, tanto analizarlas como producirlas, palabras propias de un ámbito de especialidad (como la cocina) y palabras poco comunes (como los nombres de lugares poco conocidos). Todo

ello hace que las traducciones sean del todo incoherentes y absurdas, como las siguientes:

ORIGINAL	GOOGLE	BAIDU
自公元前 七八世纪开始，延续不断修筑了 2000 多年，分布于中国北部和中部的广大土地上，总计 长度达 50000 多千米，被称之为“上下两千多年，纵横十万余里”。	Desde los siglos VII y VIII aC, se ha construido de forma continua durante más de 2.000 años. Se distribuye en vastas tierras en el norte y centro de China, con una longitud total de más de 50.000 kilómetros. <u>Se denomina "dos mil años arriba y abajo, y tiene más de 100.000 millas".</u> ".	A partir del siglo VII <u>antes del dólar de los Estados Unidos</u> , la construcción de más de 2.000 años, <u>que se remonta a más de 50.000 kilómetros de tierra en el norte</u> y el Centro de China, se ha mantenido en un total de más de 50.000 kilómetros, lo que se denomina " <u>más de 2.000 años y más de más de 100.000 kilómetros</u> ".
代表有:葱烧海参、烩乌鱼蛋、蟹黄鱼翅、德州扒鸡、奶汤核桃肉等。	Los <u>representantes</u> incluyen: <u>cebolla, pepino de mar, calamares, calamares, faisán de Texas, faisán de Texas, sopa de leche, carne de nuez, etc.</u>	Los <u>representantes</u> son los siguientes: <u>alambre de cebolla, huevos de risotto, aletas de cangrejo, carteristas de Texas, sopa de nueces de nueces, etc.</u>
三亚后海	Sanya Houhai	<u>mar de atrás</u>

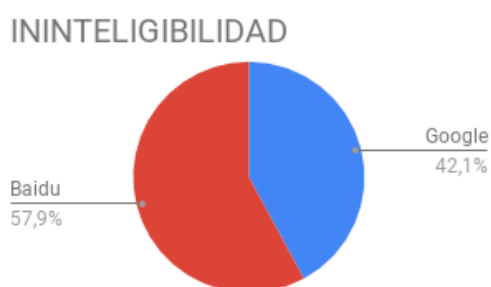


Gráfico 10. Representación del número de errores de ininteligibilidad de ambos traductores

d) Terminología

En lo que a terminología respecta, hemos podido comprobar que, sorprendentemente, el porcentaje de error de Google es más alto que el de Baidu. Google está mucho más avanzado que Baidu, con más experiencia y mucho más entrenamiento, y, al disponer de más opciones de traducción para una palabra, puede que sea más fácil que se equivoque en la selección. Asimismo, los errores terminológicos también están relacionados con los de ininteligibilidad. Muchos de los errores de terminología están incluidos en el

porcentaje de ininteligibilidad de Baidu, ya que en algunos casos hay párrafos enteros en los que la terminología específica es del todo incomprensible. Esto también hace que su porcentaje en este apartado sea menor. Lo que está claro es que ambos traductores tienen dificultades a la hora de encontrar los términos correctos. La TAN no es capaz de identificar y analizar adecuadamente los términos más específicos o las palabras poco frecuentes (lo que las convierte en tipo específico de palabras), por lo que muchas veces las traduce mal o crea sinsentidos. En la siguiente tabla aparecen varios ejemplos significativos:

ORIGINAL	GOOGLE	BAIDU
书法兼善隶、草、楷、行各体，精研体势，心摹手追，广采众长，备精诸体，冶于一炉，摆脱了汉魏笔风，自成一派，影响深远。	Boutique Caligrafía y buenos ligamentos, pasto, grilletos y cuerpos diversos, investigación intensiva, confesión, persecución, a largo plazo, bien preparados, listos, fundidos en un horno, deshacerse del estilo Han y Wei, autocontenido y de gran alcance.	La ley y la buena disposición, la hierba, el pasto, las líneas, los cuerpos, el poder de investigación, la persiguiendo, la gran audiencia, la preparación de los cuerpos, la metalurgia del horno, se ha librado de la letra de hinwei y se ha convertido en una familia de gran alcance.
中医的历史	Historia de la medicina china	Historia de la medicina china
三亚本地人	Personas locales en Sanya	Subnacionales de la población local

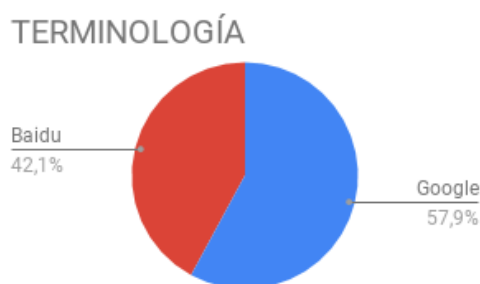


Gráfico 11. Representación del número de errores de terminología de ambos traductores

Por otro lado, también consideramos que es importante resaltar cuáles han sido los errores menos frecuentes y por qué ha sido así. Estos ayudarán a entender y reafirmar los resultados anteriores de mala traducción, tipografía, ininteligibilidad y terminología:

a) Gramática

Los 24 errores de gramática encontrados se reparten en 16 de Google y 8 de Baidu. Esta cifra de errores de gramática también nos ha sorprendido. Primero, porque la cifra de Google dobla la de Baidu y, segundo, porque esperábamos que la gramática fuera uno de los puntos más problemáticos de la TA, ya que si hay fallos en la estructura de las oraciones es muy probable que el texto no se comprenda. Así como en el apartado de errores de terminología, también creemos que muchos de los fallos gramaticales están ya recogidos en el apartado de ininteligibilidad, ya que, si la gramática no es la adecuada, además del léxico, aparecen errores como los que se muestran a continuación. El orden correcto de las palabras dentro de una oración es también uno de los problemas propios de la TAN, ya que en ocasiones puede modificar el orden de todos los elementos y provocar incoherencias y cambios de sentido. He aquí algunos ejemplos:

ORIGINAL	GOOGLE	BAIDU
中国书法造诣最深的十大家，不可不知	Las diez personas más famosas de la caligrafía china, debemos saber.	Los más profundos de la historia de China.

Cambios en el tiempo verbal:

ORIGINAL	GOOGLE	BAIDU
按照中国的传说，年是一个猛兽，会在一年中的每一天吞噬动物与人类。	Según la leyenda china, el año <u>es</u> una bestia que <u>engullirá</u> animales y humanos todos los días del año	Según la leyenda de China, el año <u>es</u> una bestia que <u>consume</u> animales y seres humanos cada uno de los días de un año.
中央电视台 2000 年拍摄千禧年第一缕阳光的地方是三亚后海。	La filmación del primer rayo de sol del milenio por parte de CCTV en el 2000 <u>es conveniente</u> para Sanya Houhai.	En la televisión central, en el año 2000, en los primeros días de la luz del Milenio, fue el mar de atrás.
约公元前 220 年，一统天下的秦始皇，将修建于早些时候的一些断续的防御工事连接成一个完整的防御系统，用以抵抗来自	Alrededor del 220 a. C., Qin Shihuang, quien dominó el mundo, conectó algunas de las fortificaciones intermitentes construidas anteriormente en un	En los 220 años anteriores a la Yu, el emperador Qin, el emperador Qin, <u>construirá</u> una serie de fortificaciones intermitentes que <u>se conectarán</u> a un sistema de

北方的侵略。	completo sistema de defensa para resistir la agresión del norte.	defensa completo para resistir a la agresión del Norte.
--------	--	---

GRAMÁTICA

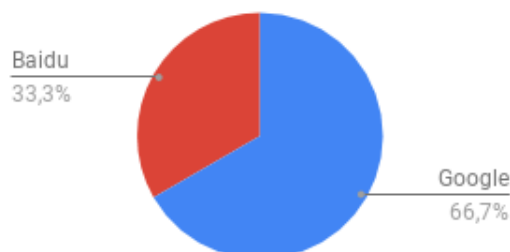


Gráfico 12. Representación del número de errores de gramática de ambos traductores

b) Inconsistencia

Es evidente que las traducciones de Google y Baidu en algunos casos han sido del todo incomprensibles por varios motivos, por lo que nos asombra mucho no haber encontrado más errores de este tipo. En total hemos identificado 10 (8 de Baidu y 2 de Google).

INCONSISTENCIA



Gráfico 13. Representación del número de errores de inconsistencia de ambos traductores

La principal razón de la falta de inconsistencias es que, por una parte, el grado de ininteligibilidad ha hecho imposible poder identificarlas todas y que, por otra, y a pesar de la cantidad de errores que hay, no hemos encontrado grandes incoherencias o contradicciones entre las ideas de los originales y las de las traducciones. Este error viene determinado por la calidad del corpus del que disponen los traductores automáticos neuronales. Algunas de las que hemos encontrado se recogen en las siguientes tablas:

En este ejemplo, vemos que el texto original habla de un escritor en masculino. No obstante, Baidu se refiere a él como mujer. Este es un problema frecuente del chino:

para las lenguas extranjeras es difícil identificar los nombres propios y, además, saber si se trata de un hombre o de una mujer.

ORIGINAL	GOOGLE	BAIDU
东晋著名书法家 “书圣”——王羲之	Famoso calígrafo de la dinastía Jin del este "Libro de San" - Wang Xizhi	La <u>famosa escritora</u> "Santo".

Aquí, podemos ver que el original no habla en ningún momento de la edad del monstruo, pero la interpretación incorrecta de Baidu hace que traduzca el nombre de la bestia de esta manera. Además, ambos traductores analizan y traducen de forma errónea el modo en el que la bestia puede ser ahuyentada, en especial Baidu.

ORIGINAL	GOOGLE	BAIDU
后来人们发现，这个 年复一年的猛兽可以 被红色驱赶，所以便 在新年的前两个星期 大量使用红色装饰。	Más tarde, se descubrió que las bestias del año y el año podían ser <u>expulsadas</u> <u>por el rojo</u> , de modo que las decoraciones rojas se usaron en grandes cantidades durante las dos primeras semanas del Año Nuevo.	Posteriormente, se descubrió que la <u>bestia de</u> <u>un año de edad</u> podía ser <u>expulsada de rojo</u> y, por lo tanto, en las dos primeras semanas del año nuevo se utilizó un gran número de decoraciones rojas.

c) Adición

En total hemos contado 5 errores de adición (1 de Google y 4 de Baidu) y en ningún caso han sido de gravedad. Esto podría deberse a que los traductores automáticos no tienen la capacidad suficiente para dar otro significado a un texto o a una frase o hacer una interpretación diferente, ya que su funcionamiento consiste en buscar traducciones existentes en el corpus del que disponen. Esto evita que el traductor pueda añadir elementos por sí solo que no tienen nada que ver con el texto original. Algunos de los errores corresponden a casos en que el traductor automático no ha sido capaz de encontrar una solución y ha decidido añadir el mismo término repetido:

ORIGINAL	GOOGLE	BAIDU
约公元前 220 年，一 统天下的秦始皇，将	Alrededor del 220 a. C., Qin Shihuang, quien	En los 220 años anteriores a la Yu, el emperador Qin,

修建于早些时候的一些断续的防御工事连接成一个完整的防御系统，用以抵抗来自北方的侵略。	dominó el mundo, conectó algunas de las fortificaciones intermitentes construidas anteriormente en un completo sistema de defensa para resistir la agresión del norte.	el <u>emperador Qin</u> , construirá una serie de fortificaciones intermitentes que se conectarán a un sistema de defensa completo para resistir a la agresión del Norte.
自公元前七八世纪开始，延续不断修筑了2000多年，分布于中国北部和中部的广大土地上，总计长度达50000多千米，被称之为“上下两千多年，纵横十万余里”。	Desde los siglos VII y VIII aC, se ha construido de forma continua durante más de 2.000 años. Se distribuye en vastas tierras en el norte y centro de China, con una longitud total de más de 50.000 kilómetros. Se denomina "dos mil años arriba y abajo, y tiene más de 100.000 millas". "	A partir del siglo VII <u>antes del dólar de los Estados Unidos</u> , la construcción de más de 2.000 años, que se remonta a más de 50.000 kilómetros de tierra en el norte y el Centro de China, <u>se ha mantenido en un total de más de 50.000 kilómetros</u> , lo que se denomina "más de 2.000 años y más de más de 100.000 kilómetros".

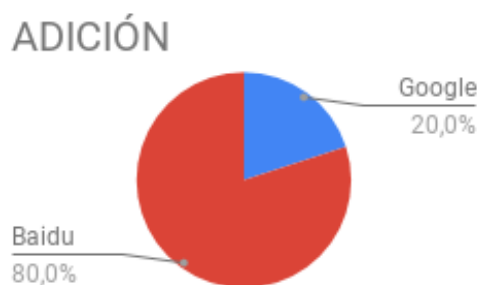


Gráfico 14. Representación del número de errores de adición de ambos traductores

d) Omisión

Los traductores automáticos se rigen por una serie de algoritmos y fórmulas que hacen que se identifiquen todas las partes de un texto o frase y, por lo tanto, no quede nada sin traducir. No obstante, a veces se producen omisiones, creemos que porque la falta de traducciones hace que los motores de traducción no puedan encontrar una buena solución y decidan omitir el contenido. En nuestro análisis hemos encontrado 20 omisiones (9 de Google y 11 de Baidu). Algunas de ellas consideramos que son

omisiones graves, ya que se pierde información importante del texto original, como por ejemplo:

En los siguientes ejemplos, Baidu omite información importante del texto original. En el primer caso omite información del autor, e incluso el nombre de este, y en el segundo omite la especificación de los años (d. C.). Errores como estos podrían provocar confusiones o malos entendidos entre los lectores. La información omitida se presenta entre paréntesis y subrayada.

ORIGINAL	GOOGLE	BAIDU
东晋著名书法家 “书圣”——王羲之	Famoso calígrafo de la dinastía Jin del este "Libro de San" - Wang Xizhi	La famosa escritora (<u>de la dinastía Jin del este</u>). "Santo". (<u>Wang Xizhi</u>)
在明代（公元1368~1644年），	En la dinastía Ming (1368-1644 dC),	La construcción de la gran ciudad se ha convertido en la más larga de las instalaciones militares del mundo en la generación mañana (1368 a 1644 (<u>d. C.</u>)).

En este caso, ambos traductores han omitido mucha información importante, aunque el resultado de Baidu es notablemente peor. En este caso hemos decidido añadir la versión poseditada para dejar más claras estas omisiones.

ORIGINAL	周代已经使用望、闻、问、切等诊病方法和药物、针灸、手术等治疗方法。
GOOGLE	La dinastía Zhou ha utilizado (<u>los cuatro</u>) métodos de diagnóstico, tratamiento, acupuntura y cirugía (<u>como método de cura</u>).
BAIDU	En el transcurso de las semanas (<u>durante la dinastía Zhou</u>) se han utilizado métodos de tratamiento, tratamiento y tratamiento (<u>como método de cura</u>), así como tratamiento médico, acupuntura, cirugía, etc.
POSEDICIÓN	Durante la dinastía Zhou ya se utilizaban los cuatro métodos de diagnóstico (observación, auscultación y olfacción, interrogatorio y toma del pulso) y tratamientos, así como la acupuntura, la moxibustión y la cirugía como métodos de cura.

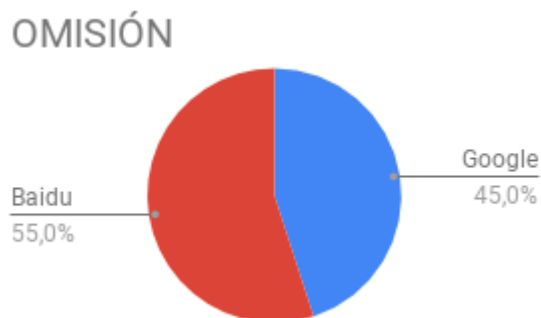


Gráfico 15. Representación del número de errores de omisión de ambos traductores

e) No traducido

En este caso, ocurre lo mismo que con las omisiones y adiciones. Los algoritmos que utiliza la TA hacen que se identifiquen todas las partes de una oración y así se evita que haya información diferente a la del original. El único elemento no traducido que hemos localizado es la denominación popular de la comida de Shandong *Lu Cai* (texto 3). A diferencia de Baidu, Google no lo ha traducido y esa es la opción correcta, ya que, como hemos dicho, se trata del nombre por el que es conocida la cocina típica de un lugar, por lo que mantendríamos el original en pinyin y en cursiva.

f) Registro gramatical

El registro gramatical se ha mantenido intacto en todos los textos. Esto se debe a que los textos seleccionados emplean un registro estándar, fácil de comprender para cualquier tipo de usuario. Este es el nivel con el que trabajan los traductores automáticos, por lo que es normal que no haya errores de este tipo. En caso de haber trabajado con diálogos, quizá hubiéramos encontrado más problemas por los acentos geográficos o los dialectos, ya que a día de hoy los traductores automáticos no son capaces de identificar y traducir bien aspectos lingüísticos como esos.

g) Ortografía

Tampoco se han encontrado errores de ortografía. Las palabras estaban escritas correctamente, así como las tildes en los casos que así lo requerían, por lo que, a pesar de que la gramática falle, hemos podido comprobar que los traductores automáticos respetan la ortografía de las palabras:

GOOGLE	BAIDU
<p>El pueblo de Houhai, también conocido como el pueblo de pescadores de Tenghai, está escondido entre las montañas y el mar, y muchas personas locales en Sanya nunca han pisado este lugar. Ubicada en el lado noreste del pueblo, la isla de Wuzhizhou recibe a millones de visitantes de todo el mundo, en marcado contraste con la tranquilidad del pueblo de pescadores.</p>	<p>Además de las celebraciones de los dioses y los ancestros, la celebración de la primavera está muy relacionada con las historias de mitos y mitos. La historia del mito más popular es el año de la bestia y el final de un año. Según la leyenda de China, el año es una bestia que consume animales y seres humanos cada uno de los días de un año. Posteriormente, se descubrió que la bestia de un año de edad podía ser expulsada de rojo y, por lo tanto, en las dos primeras semanas del año nuevo se utilizó un gran número de decoraciones rojas.</p>

4. PROPUESTA DE POSEDICIÓN

A continuación, se presenta la propuesta de posesición de los textos según las normas TAUS (2016) mencionadas anteriormente.

Original	Google	Baidu	Propuesta de posesición
Texto 1			
<p>春节的历史</p> <p>年的传说</p> <p>除了对神明和祖先的庆祝，春节的庆祝也与传说和神话故事有着很大的关联。最受欢迎的神话故事便是怪兽年以及一年的结束。按照中国的传说，年是一个猛兽，会在一年中的每一天吞噬动物与人类。后来人们发现，这个年复一年的猛兽可以被红色驱赶，所以便在新年的前两个星期大量使用红色装饰。现在，红色被认为是吉</p>	<p>La historia del Festival de Primavera.</p> <p>Leyenda del año</p> <p>Además de la celebración de los dioses y antepasados, la celebración del Festival de Primavera también está estrechamente relacionada con las leyendas y los mitos. El mito más popular es el año del monstruo y el fin del año. Según la leyenda china, el año es una bestia que engullirá animales y humanos todos los días del año. Más tarde, se descubrió que las bestias del año y el año podían ser expulsadas por el rojo, de modo que las decoraciones rojas se usaron en grandes cantidades durante las dos primeras semanas del Año Nuevo.</p>	<p>Historia de la primavera</p> <p>La leyenda del año.</p> <p>Además de las celebraciones de los dioses y los ancestros, la celebración de la primavera está muy relacionada con las historias de mitos y mitos. La historia del mito más popular es el año de la bestia y el final de un año. Según la leyenda de China, el año es una bestia que consume animales y seres humanos cada uno de los días de un año. Posteriormente, se descubrió que la bestia de un año de edad podía ser expulsada de rojo y, por lo tanto, en las dos primeras semanas del año nuevo se utilizó un gran número de decoraciones rojas. Ahora, el rojo se considera un color de la mascota que</p>	<p>La historia de la Fiesta de la Primavera</p> <p>La leyenda de Nian</p> <p>Además de honrar a los dioses y antepasados, la celebración de la Fiesta de la Primavera también está estrechamente relacionada con las leyendas y los mitos. El mito más popular es el del monstruo Nian y el fin de año. Según la leyenda china, Nian era una bestia que cada día a lo largo del año engullía animales y seres humanos. Más tarde, se descubrió que la bestia temía al color rojo, por lo que durante las dos primeras semanas del Año Nuevo se empezó a hacer uso de una gran cantidad de decoraciones rojas. Ahora el rojo está considerado como color de la buena suerte y se usa para disfraces y fuegos artificiales durante el Año Nuevo Chino. Este color es una forma de ahuyentar al monstruo Nian y de</p>

祥的颜色，在春节期间用于服饰和烟花爆竹。红色被视为可以驱赶怪兽年，并为即将到来的新的一年带来好运。		se utiliza en el Festival de la primavera para la ropa y los fuegos artificiales. Se considera que el rojo puede expulsar a la bestia y traer suerte para el próximo año que viene.	traer buena suerte al próximo Año Nuevo.
Texto 2			
中国书法造诣最深的十大家，不可不知 东晋著名书法家 “书圣”——王羲之 王羲之是中国最著名的、成就最高的书法家，他广采众长，自成一家，不仅作为中国书法史上一座极具象征性的丰碑深远影响着中国千年书法，还留下了大量的丰富多样的书法精品。书法兼善隶、草、楷、行各体，精研	Las diez personas más famosas de la caligrafía china, debemos saber. Famoso calígrafo de la dinastía Jin del este "Libro de San" - Wang Xizhi Wang Xizhi es el calígrafo más famoso y consumado de China. Tiene una larga historia y es una familia autónoma. No solo tiene un monumento simbólico en la historia de la caligrafía china, sino que también afecta profundamente a la caligrafía china de mil años. También dejó una gran cantidad de caligrafía rica y variada. Boutique Caligrafía y buenos ligamentos, pasto, grilletos y cuerpos	Los (diez calígrafos) más profundos de la historia de China. La famosa escritora. "Santo". (Wang Xizhi) El rey, que es el mejor libro de la historia de China, es el más famoso y el más alto de los éxitos de China. La ley y la buena disposición, la hierba, el pasto, las líneas, los cuerpos, el poder de investigación, la persiguiendo, la gran audiencia, la preparación de los cuerpos, la metalurgia del horno, se ha librado de la letra de hinwei y se ha convertido en una familia de gran alcance.	Los diez calígrafos más importantes de China que debes conocer Famoso calígrafo de la dinastía Jin del este «El sabio de la caligrafía» - Wang Xizhi Wang Xizhi es el calígrafo más famoso y consumado de China. La mayoría de sus obras son muy buenas, con estilo propio. No solo creó la obra monumental y más emblemática de la historia de la caligrafía china que ha servido de modelo a otros calígrafos del milenio, sino que también tuvo una gran influencia en la caligrafía china de hace mil años. Además, ha dejado un numeroso y variado legado artístico de obras de caligrafía. Su caligrafía también es buena en estilo

体势，心摹手追，广采众长，备精诸体，冶于一炉，摆脱了汉魏笔风，自成一家，影响深远。	diversos, investigación intensiva, confesión, persecución, a largo plazo, bien preparados, listos, fundidos en un horno, deshacerse del estilo Han y Wei, autocontenido y de gran alcance.		oficial, estilo hierba, estilo regular, estilo corriente, hace un estudio detallado sobre las características e intenta copiar el trazo de la mano. Lo mejor de su caligrafía es que, con diferentes estilos, se alejó de los estilos de las dinastías Han y Wei y creó su propio estilo, cuya influencia se ha mantenido hasta el presente.
Texto 3			
中国的四大菜系分别是哪些？ 鲁 菜 也称山东菜，中国四大菜系之一。鲁菜取材广泛，选料精细，讲究丰满实惠，烹调方法全面，精于制汤，善以葱调味。鲁菜在烹制海鲜上有独到之处，尤其对海珍和小海味的烹制，堪称一绝。代表有：葱烧海参、烩	¿Cuáles son las cuatro principales cocinas chinas? <i>Lu Cai</i> También conocida como la cocina de Shandong, es una de las cuatro principales cocinas de China. Lu Cai toma una amplia gama de materiales, una excelente selección de materiales, presta atención a la plenitud y los beneficios, a los métodos de cocción completos, a la refinada sopa, al buen condimento con cebollas. La cocina Lu tiene su propio sabor único en la cocina de mariscos, especialmente	¿Cuáles son los cuatro menú de China? <i>Remolacha</i> China es uno de los cuatro menú de China. La remolacha se basa en una amplia gama de materiales selectivos, la buena cocina, la buena cocina, la precisión de la sopa y el sabor a la cebolla. La comida de los mariscos para cocinar, especialmente para la cocina de Jen y la pequeña sabor de mar, es una buena cocinera. Los representantes son los siguientes: alambre de cebolla, huevos de risotto,	¿Cuáles son los cuatro estilos culinarios chinos? <i>Lu cai</i> También conocida como la cocina de Shandong, es uno de los cuatro principales estilos culinarios de China. Esta utiliza una amplia variedad de ingredientes seleccionados meticulosamente, presta especial atención a sus beneficios y a todo el proceso de cocción. Los expertos destacan la sopa y el condimento con cebolla. La cocina de Shandong destaca por su forma de cocinar el marisco, especialmente la forma de cocinar los frutos del mar, por lo que es justo calificarla como

乌鱼蛋、蟹黄鱼翅、德州扒鸡、奶汤核桃肉等。	para la cocina de Haizhen y Xiaohaiwei. Los representantes incluyen: cebolla, pepino de mar , calamares, calamares , faisán de Texas, faisán de Texas , sopa de leche , carne de nuez , etc.	aletas de cangrejo, carteristas de Texas, sopa de nueces de nueces, etc.	única. Entre lo más destacado encontramos: pepino de mar asado con cebolleta, estofado de huevos de pez negro, aleta de tiburón con cangrejo, pollo estofado Dezhou y sopa de nueces y carne.
Texto 4			
世界文化遗产——长城 约公元前 220 年，一统天下的秦始皇，将修建于早些时候的一些断续的防御工事连接成一个完整的防御系统，用以抵抗来自北方的侵略。在明代（公元 1368～1644 年），又继续加以修筑，使长城成为世界上最长的军事设施。它在文化艺术上的价值，足以与其在历史和战略上的重要性相媲美。 概况	Patrimonio Cultural de la Humanidad - Gran Muralla Alrededor del 220 a. C., Qin Shihuang, quien dominó el mundo, conectó algunas de las fortificaciones intermitentes construidas anteriormente en un completo sistema de defensa para resistir la agresión del norte. En la dinastía Ming (1368-1644 dC), se continuó construyendo para hacer de la Gran Muralla la instalación militar más larga del mundo. Su valor en la cultura y el arte es comparable a su importancia histórica y estratégica. Visión general La Gran Muralla es un antiguo	Patrimonio Mundial - la gran ciudad En los 220 años anteriores a la Yu, el emperador Qin, el emperador Qin, construirá una serie de fortificaciones intermitentes que se conectarán a un sistema de defensa completo para resistir a la agresión del Norte.La construcción de la gran ciudad se ha convertido en la más larga de las instalaciones militares del mundo en la generación mañana (1368 a 1644).Su valor en el arte y el arte culturales es comparable a su importancia histórica y estratégica. Generalidades La gran ciudad es una de las obras más	Patrimonio Cultural de la Humanidad – La Gran Muralla Alrededor del 220 a. C., Qin Shi, emperador de la dinastía Qin, conectó algunos tramos de la muralla construidos anteriormente en un completo sistema de defensa para resistir la agresión del norte. La construcción prosiguió durante la dinastía Ming (1368-1644 d. C.) para hacer de la Gran Muralla la construcción militar más larga del mundo. Su valor en la cultura y el arte es comparable a su importancia histórica y estratégica. Información general La Gran Muralla es una antigua construcción de defensa con la mejor ingeniería y el tiempo

<p>长城是中国也是世界上修建时间最长、工程量最大的一项古代防御工程。自公元前七八世纪开始，延续不断修筑了 2000 多年，分布于中国北部和中部的广大土地上，总计 长度达 50000 多千米，被称之为“上下两千多年，纵横十万余里”。如此浩大的工程不仅 在中国就是在世界上，也是绝无仅有的，因而在几百年前就与罗马斗兽场、比萨斜塔等列为中古世界七大奇迹之一。</p>	<p>proyecto de defensa con el tiempo de construcción más largo y la mayor cantidad de ingeniería en China y del mundo entero. Desde los siglos VII y VIII aC, se ha construido de forma continua durante más de 2.000 años. Se distribuye en vastas tierras en el norte y centro de China, con una longitud total de más de 50.000 kilómetros. Se denomina "dos mil años arriba y abajo, y tiene más de 100.000 millas". ". Un proyecto tan vasto no solo está en China sino también en el mundo, y es único. Por lo tanto, fue catalogado como una de las Siete Maravillas de la Edad Media con el Coliseo y la Torre Inclinada de Pisa hace cientos de años.</p>	<p>antiguas de defensa de China, la mayor de las cuales se construyó en el mundo. A partir del siglo VII antes del dólar de los Estados Unidos, la construcción de más de 2.000 años, que se remonta a más de 50.000 kilómetros de tierra en el norte y el Centro de China, se ha mantenido en un total de más de 50.000 kilómetros, lo que se denomina "más de 2.000 años y más de más de 100.000 kilómetros". Este gran proyecto, no sólo en China, sino también en el mundo, es uno de los siete milagros del mundo central, que hace cientos de años, se enmarcó en el campo de los arcos romanos, la torre de pizza, etc.</p>	<p>de construcción más largo de China y del mundo entero. Se estuvo construyendo de forma continua durante más de 2000 años desde los siglos VII y VIII a. C. La muralla se extiende por las tierras del norte y centro de China con un total de más de 50000 quilómetros, motivo por el que existe una expresión china que la describe como «larga en espacio y tiempo». Por lo tanto, este no es solo un gran proyecto en China, sino en el mundo entero; es algo único. Por esta razón, hace cientos de años se la catalogó junto con el Coliseo y la Torre de Pisa como una de las Siete Maravillas de la Edad Media.</p>
Texto 5			
<p>中医的历史</p> <p>在 3000 多年前的殷商甲骨文中，中国已经有关于医疗</p>	<p>Historia de la medicina china</p> <p>En los huesos del oráculo del Yin Shang, hace más de 3.000 años, China ha registrado atención médica y de</p>	<p>Historia de la medicina china</p> <p>En China, hace más de 3.000 años, se ha documentado la atención de la salud y más de una docena de</p>	<p>Historia de la medicina tradicional china</p> <p>En los huesos oraculares de la dinastía Shang, conocida entonces como Yinshang, de hace más de 3000 años ya hay constancia de</p>

<p>卫生以及十多种疾病的记载。周代已经使用望、闻、问、切等诊病方法和药物、针灸、手术等治疗方法。秦汉时期，形成了《黄帝内经》这样具有系统理论的著作。此书是现存最早的一部中医理论性经典著作。张仲景所著的《伤寒杂病论》，专门论述了多种杂病的辨证诊断、治疗原则，为后世的临床医学奠定了发展的基础。汉代外科学已具有较高水平。据《三国志》记载，名医华佗已开始使用全身麻醉剂“麻沸散”进行各种外科手术。</p>	<p>salud y más de 10 enfermedades. La dinastía Zhou ha utilizado métodos de diagnóstico, tratamiento, acupuntura y cirugía. En las dinastías Qin y Han, se formó el libro "Sistema del Emperador Amarillo". Este libro es el clásico teórico más antiguo de la medicina china. El "Tratado sobre las enfermedades febriles" de Zhang Zhongjing analiza específicamente los principios del diagnóstico y cura de diferenciación del síndrome y el tratamiento de varias enfermedades diversas, lo que sentó las bases para el desarrollo de la medicina clínica posterior.</p>	<p>enfermedades. En el transcurso de las semanas se han utilizado métodos de tratamiento, tratamiento y tratamiento, así como tratamiento médico, acupuntura, cirugía, etc. La época de qinhan fue una obra de teorías sistemáticas como el menú del rey Huang. El libro es el primero de los libros clásicos teóricos de la medicina china. La teoría del diagnóstico y tratamiento de múltiples enfermedades, que se centra en el diagnóstico y el tratamiento de múltiples enfermedades, ha sentado las bases para el desarrollo de la medicina clínica en las generaciones venideras.</p>	<p>tratamientos médicos y de más de diez enfermedades diferentes. Durante la dinastía Zhou ya se utilizaban los cuatro métodos de diagnóstico (observación, auscultación y olfacción, interrogatorio y toma del pulso) y tratamientos, así como la acupuntura, la moxibustión y la cirugía como métodos de cura. En las dinastías Qin y Han se escribió el libro <i>Canon de medicina interna del emperador Amarillo</i>, el clásico teórico más antiguo que recoge los principios de la medicina china. El libro <i>Tratado sobre el daño por frío y enfermedades misceláneas</i> de Zhang Zhong Jing analiza específicamente el diagnóstico de varias enfermedades misceláneas y su cura, lo que sentó las bases para el desarrollo de la medicina clínica posterior. La cirugía de la dinastía Han ya tenía un nivel bastante alto. Según los «registros de los Tres Reinos», el famoso médico Hua Tuo ya había empezado a utilizar el anestésico <i>Mafeisan</i> en operaciones quirúrgicas.</p>
--	---	---	--

Texto 6			
<p>三亚后海：探寻隐世渔村的浪漫情调</p> <p>后海村，又称藤海渔村，藏于山海之间，就连许多三亚本地人都未曾踏足此地。位于该村东北侧的蜈支洲岛，每年接待着上百万名来自世界各地的游客，与渔村的静谧安逸形成鲜明对比。后海村虽默默无闻，但其绝美风光曾一度通过电视画面走进千家万户——中央电视台2000年拍摄千禧年第一缕阳光的地方便是三亚后海。</p>	<p>Houhai en Sanya: explorando el ambiente romántico del pueblo de pescadores escondido</p> <p>El pueblo de Houhai, también conocido como el pueblo de pescadores de Tenghai, está escondido entre las montañas y el mar, y muchas personas locales en Sanya nunca han pisado este lugar. Ubicada en el lado noreste del pueblo, la isla de Wuzhizhou recibe a millones de visitantes de todo el mundo, en marcado contraste con la tranquilidad del pueblo de pescadores. Aunque se desconoce el pueblo de Houhai, su hermoso paisaje una vez ingresó a las casas a través de las pantallas de televisión. La filmación del primer rayo de sol del milenio por parte de CCTV en el 2000 es conveniente para Sanya Houhai.</p>	<p>El romanticismo en la búsqueda de una aldea de pesca invisible.</p> <p>La aldea de atrás, conocida también como la aldea de pescadores del Mar de Kato, se esconde entre los mares de montaña, y ni siquiera muchos de los tres subnacionales de la población local han pisado el lugar. La isla del CICAA, situada al nordeste de la aldea, recibe cada año a millones de turistas de todo el mundo, lo que contrasta con el silencio de las aldeas de pesca. En la televisión central, en el año 2000, en los primeros días de la luz del Milenio, fue el mar de atrás.</p>	<p>Houhai, Sanya: explorando el ambiente romántico del pueblo de pescadores escondido</p> <p>El pueblo de Houhai, también conocido como el pueblo de pescadores de Tenghai, está escondido entre las montañas y el mar, tanto que incluso hay muchas personas de Sanya que nunca antes han pisado este lugar. Situada en el lado noreste del pueblo, la isla de Wuzhizhou recibe a millones de visitantes de todo el mundo, en marcado contraste con la tranquilidad del pueblo de pescadores. Aunque el pueblo de Houhai no es muy conocido, su hermoso paisaje una vez entró a las casas a través de las pantallas de televisión. La retransmisión del primer rayo de sol del milenio por parte de CCTV en el año 2000 fue una gran oportunidad para el pueblo de Houhai.</p>

5. CUESTIONARIO

Para la realización de este trabajo, pensamos que sería interesante conocer cuál es la opinión de otros estudiantes de traducción con chino como lengua C. Esto nos ayudará a comprobar si en la práctica los servicios de TA que ofrecen Google y Baidu resultan de utilidad o no a los estudiantes como herramienta para tareas de documentación. En este capítulo presentaremos una descripción detallada de ese cuestionario, el análisis de los resultados y las conclusiones.

5.1. Descripción del cuestionario

Para informar a los estudiantes del propósito de este cuestionario, primero hemos realizado una breve presentación en la que se explica cuál es el objetivo principal: conocer la opinión de otros estudiantes sobre la TA del chino al español. Para garantizar que las opiniones sean reales, es decir, de estudiantes que llevan el tiempo suficiente estudiando chino como para tener ya una opinión sólida sobre el tema, hemos enviado el cuestionario a alumnos de traducción de tercero y cuarto curso del Grado de Traducción e Interpretación. En total hemos obtenido un total de 16 respuestas, 7 son alumnos de tercero y 9 son de cuarto. Todos están matriculados en el grado y no hay estudiantes de intercambio, a pesar de que hay estudiantes no nativos que estudian la combinación lingüística chino-español. Prácticamente la mitad de los encuestados hace más de tres años que estudia chino. El cuestionario consta de un total de 17 preguntas, las cuales se describen a continuación:

1. La primera pregunta es para confirmar **cuántos años hace que estudian chino**; deben escoger una opción de entre las siguientes: menos de 1 año, entre 2 y 3 años y más de 3 años. Es obligatoria.
2. **¿Alguna vez has tenido que buscar bibliografía original en chino?** Los estudiantes deben responder sí o no. Es obligatoria.
3. **¿Cómo has solucionado el problema de la barrera lingüística al leer los textos?** Aquí ofrecemos cuatro opciones: diccionarios, enciclopedias, traductores automáticos y otros. En esta última opción los alumnos pueden proponer la opción que deseen. Además, en esta pregunta se pueden seleccionar varias opciones, por lo que en los resultados aparecerán varias combinaciones posibles. Es obligatoria.
4. **¿Consideras que para un estudiante de chino como lengua C es fácil acceder y comprender bibliografía en chino?** Los estudiantes deben responder sí o no. Es obligatoria.
5. **¿Crees que hay suficientes recursos en línea para obtener y/o entender bibliografía en lengua china?** Los estudiantes deben responder sí o no. Es obligatoria.

6. **¿Alguna vez has utilizado algún traductor automático de internet?** Los estudiantes deben responder sí o no. Es obligatoria.
7. **¿Cuál has utilizado o utilizas con más frecuencia?** Aquí ofrecemos varias opciones: Google Traductor, Baidu Translate, Microsoft Bing y otros. En esta última opción los alumnos pueden proponer la opción que deseen. Además, en esta pregunta se pueden seleccionar varias opciones, por lo que en los resultados aparecerán varias combinaciones posibles. Es obligatoria.
8. **¿Crees que la traducción automática resulta útil para tareas de documentación? Por ejemplo, imagina que estás traduciendo un texto sobre medicina tradicional china y necesitas buscar información sobre el tema en páginas web escritas originalmente en chino, ¿crees que la traducción automática te ayudaría a comprender la información?** Los estudiantes deben responder sí o no. Es obligatoria.
9. **¿En alguna ocasión has utilizado la traducción automática con otros fines, por ejemplo, como diccionario o para traducir textos?** Los estudiantes deben responder sí o no. Es obligatoria.
10. **En caso afirmativo, di qué motor de traducción automática has utilizado.**
Esta es una respuesta libre, es decir, los estudiantes pueden poner la opción que más se ajuste a su realidad. En caso negativo no es necesario que respondan. Es obligatoria.
11. **¿Cómo valorarías tu experiencia con la traducción automática? Puntúa de 1 (muy mala experiencia) a 10 (muy buena experiencia).** En esta pregunta los estudiantes deben valorar su experiencia personal con la TA con una puntuación de entre 1 y 10 puntos, siendo 1 el valor más bajo y 10 el valor más alto. Es obligatoria.
12. **¿Crees que la traducción automática del chino al español podría resultar útil en contextos especializados, como medicina, economía, etc.?** Los estudiantes deben responder sí o no. Es obligatoria.
13. **¿De qué crees que trata este texto: «La gran ciudad es una de las obras más antiguas de defensa de China, la mayor de las cuales se construyó en el mundo. A partir del siglo VII antes del dólal de los Estados Unidos, la construcción de más de 2.000 años, que se remonta a más de 50.000 kilómetros de tierra en el norte y el Centro de China, se ha mantenido en un total de más de 50.000 kilómetros, lo que se denomina "más de 2.000 años y más de más de 100.000 kilómetros".»?** Esta es una pregunta de respuesta libre, es decir, los estudiantes deben responder de forma breve cuál creen que es el tema del texto. No hay opciones correctas, simplemente deseamos comprobar si son capaces de identificar de qué se está hablando. Es obligatoria.

14. **Puntúa de 1 (muy bajo) a 10 (muy alto) el grado de comprensión del siguiente texto: «Evaluación del Comité del Patrimonio Mundial. Alrededor del 220 a. C., Qin Shihuang, quien dominó el mundo, conectó algunas de las fortificaciones intermitentes construidas anteriormente en un completo sistema de defensa para resistir la agresión del norte. En la dinastía Ming (1368-1644 dC), se continuó construyendo para hacer de la Gran Muralla la instalación militar más larga del mundo. Su valor en la cultura y el arte es comparable a su importancia histórica y estratégica.».** En esta pregunta deben evaluar la calidad del texto según la cantidad de información que son capaces de comprender. Para ello, deben puntuarlo con entre 1 y 10 puntos, siendo 1 el valor más bajo y 10 el valor más alto. Es obligatoria.
15. **En general, ¿cuál crees que es el principal fallo en las traducciones automáticas del chino? Puedes seleccionar varias opciones.** Esta es una pregunta en la que les proponemos diferentes opciones, de las cuales pueden seleccionar más de una: contenido, estructura del texto, léxico, gramática y otros. En esta última opción los alumnos pueden proponer la opción que deseen. Es obligatoria.
16. **¿Crees que la traducción automática es una posible solución para los estudiantes de chino como lengua C a la hora de realizar tareas de documentación sobre un tema? Si no es así, ¿cuál crees que sería una posible solución?** Esta es una pregunta de respuesta libre en la que los estudiantes pueden expresar cuál es su opinión personal sobre la pregunta que se les está planteando. Es obligatoria.
17. **Observaciones.** Este apartado es opcional. Los estudiantes pueden realizar los comentarios que consideren oportunos si así lo desean.

5.2. Análisis de datos

La muestra obtenida con este cuestionario es de un total de 16 alumnos de traducción de chino de tercero y cuarto curso. Como vemos en los gráficos, más de la mitad de los alumnos (56,3%) ya lleva más de tres años estudiando chino y prácticamente todos han tenido que buscar bibliografía original en chino alguna vez (gráficos 16 y 17).

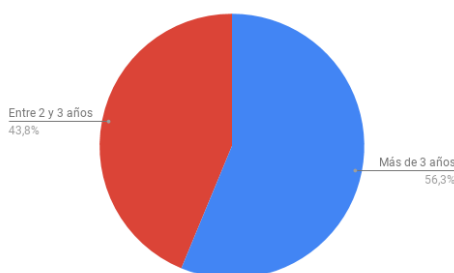


Gráfico 16. Representación de los resultados de la pregunta 1 del cuestionario

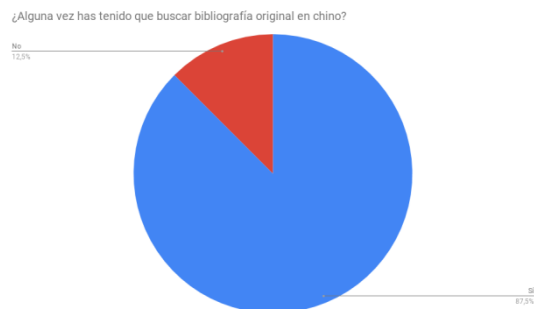


Gráfico 17. Representación de los resultados de la pregunta 2 del cuestionario

A la hora de comprender bibliografía en chino existen varias posibilidades, de las cuales ofrecimos las siguientes: diccionarios, enciclopedias, traductores automáticos o señalar otros métodos. En el siguiente gráfico (gráfico 18) vemos que la opción más frecuente es la de utilizar únicamente traductores automáticos (50%). La segunda opción más escogida por los estudiantes han sido los diccionarios (25%) y la tercera, el uso de diccionarios y traductores automáticos (12,6%). Al tratarse de una pregunta de opción múltiple han aparecido nuevas opciones y combinaciones, ambas tienen el mismo porcentaje (6,3%) y son: textos paralelos, diccionarios y enciclopedias; diccionarios, traductores automáticos y vídeos.

Vemos que la TA es la opción más utilizada porque es la más rápida de todas, aunque no siempre es la más fiable. Hoy en día, si encontramos una página web en una lengua extranjera Google nos ofrece la opción de que con un solo clic se traduzca toda la página. Esta opción nos puede ayudar a tener una idea general de qué está hablando el texto, pero no de los detalles, especialmente si la combinación lingüística con la que trabajamos es tan opuesta como la de chino-español. Por esta razón el uso de los diccionarios sigue siendo importante, pero los estudiantes cada vez lo utilizan menos. Su uso está quedando restringido a profesionales del ámbito de la lengua, ya que, aunque son una fuente fiable, ralentizan mucho nuestro trabajo al tener que buscar el significado palabra por palabra. Vemos que los estudiantes también optan por consultar textos paralelos y vídeos, aunque estas opciones son menos frecuentes.



Gráfico 18. Representación de los resultados de la pregunta 3 del cuestionario

El siguiente gráfico (gráfico 19) justifica el alto uso de la TA representado en el gráfico 18. Las dificultades de los estudiantes de traducción a la hora de comprender textos en chino hacen que opten por la traducción automática. Es el recurso más rápido para traducir grandes cantidades de texto del chino al español, pero la calidad es tan deficiente que debe combinarse con otras herramientas, como los diccionarios o textos paralelos. No obstante, la mayoría de los estudiantes (62,5%, gráfico 20) cree que todas estas herramientas no son suficientes para resolver el problema de la comprensión de textos en chino. A día de hoy, tanto la comprensión de los textos como la búsqueda de información adicional a menudo son tareas difíciles que llevan mucho tiempo.

¿Consideras que para un estudiante de chino como lengua C es fácil acceder y comprender bibliografía en chino?

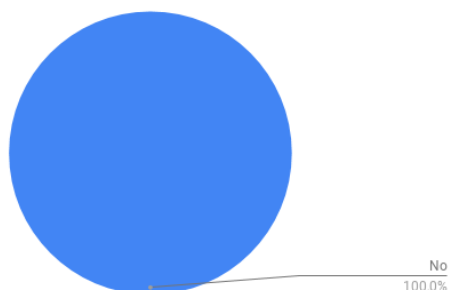


Gráfico 19. Representación de los resultados de la pregunta 4 del cuestionario

¿Crees que hay suficientes recursos en línea para obtener y/o entender bibliografía en lengua china?

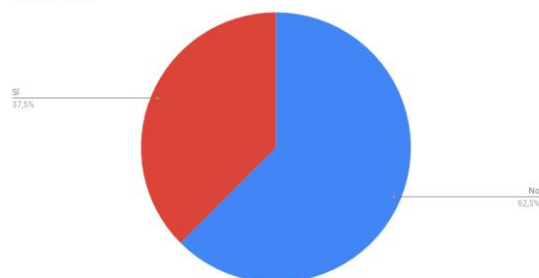


Gráfico 20. Representación de los resultados de la pregunta 5 del cuestionario

Las respuestas a las preguntas anteriores confirman que todos los alumnos encuestados han utilizado la TA en alguna ocasión (gráfico 21), bien para buscar bibliografía en chino o con otros fines. Para hacer uso de la TA, bien podemos hacerlo a través de programas o través de traductores automáticos de internet. Todos los estudiantes encuestados han utilizado o utilizan los traductores automáticos de internet. Entre los más habituales encontramos (gráfico 22), en primer lugar, Google Traductor con un 61,1% y, en segundo lugar, Baidu Translate con un 22,2%. Otro traductor automático señalado por los alumnos es el de Youdao (16,7%), que también utilizan en combinación con el de Google y el de Baidu. Intuimos que estos datos reflejan la realidad de que Google es mucho más conocido debido a su alcance mundial y, por tanto, es un motor mucho más entrenado. Además, sus resultados suelen ser más inteligibles que los de otros motores.

¿Alguna vez has utilizado algún traductor automático de internet?

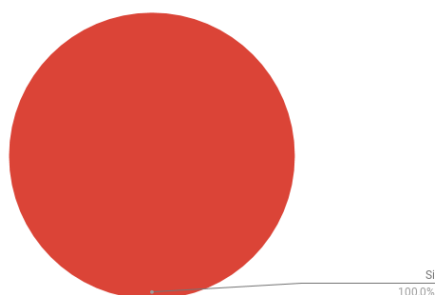


Gráfico 21. Representación de los resultados de la pregunta 6 del cuestionario

¿Cuál has utilizado o utilizas con más frecuencia?

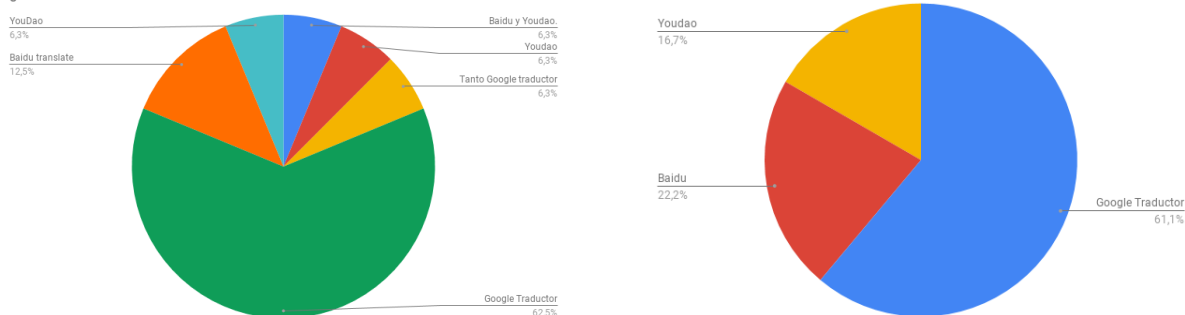


Gráfico 22. Representación de los resultados de la pregunta 7 del cuestionario

Un 68,8% de los encuestados está de acuerdo en que la TA puede resultar útil a la hora de realizar tareas de documentación (gráfico 23). Tal vez no nos ofrezca todos los detalles del texto que estamos leyendo, pero sí nos puede ayudar a tener una idea general de lo que se está hablando, lo que nos da un punto del que partir en nuestra

búsqueda de documentación. Puede ofrecernos otro tipo de información útil, como recursos u otras fuentes bibliográficas donde buscar.

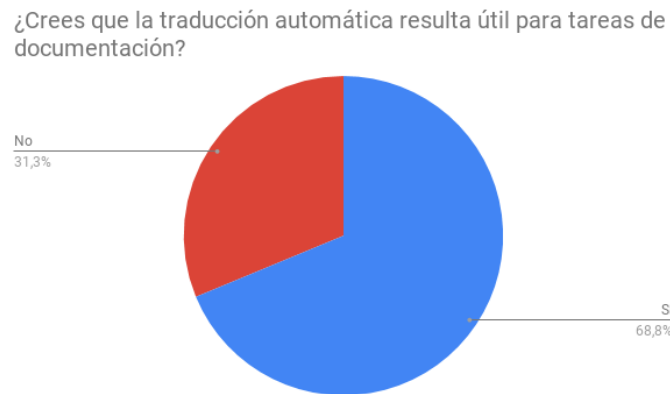


Gráfico 23. Representación de los resultados de la pregunta 8 del cuestionario

No obstante, la traducción automática también puede servir como herramienta para otros fines, por ejemplo como diccionario o simplemente como traductor automático. El 66,7% de los estudiantes encuestados ya ha probado otros usos diferentes de la TA (gráfico 24). Si viajas a un país del que no conoces el idioma, la TA puede llegar a ser una herramienta bastante útil para comunicarte, aunque siempre debemos ir con cuidado. La TA cumplirá su función comunicativa siempre y cuando intentemos traducir frases sencillas, pero si complicamos mucho su trabajo podría causarnos más problemas que ventajas.

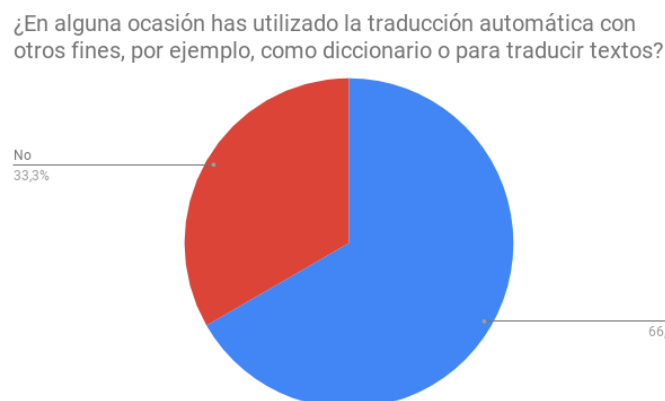


Gráfico 24. Representación de los resultados de la pregunta 9 del cuestionario

Estos son los motores que han escogido o escogen cuando necesitan utilizar la TA para otros fines además de las tareas de documentación (gráfico 25). Una vez más, Google Traductor es la opción más frecuente (44,4%). No obstante, nos ha llamado la atención que un 27,8% de los encuestados ha respondido que no utiliza la TA con otros fines además de para documentarse. Esto quiere decir que ellos se encargan de realizar toda la traducción y, en caso de duda, consultan diccionarios u otras fuentes.

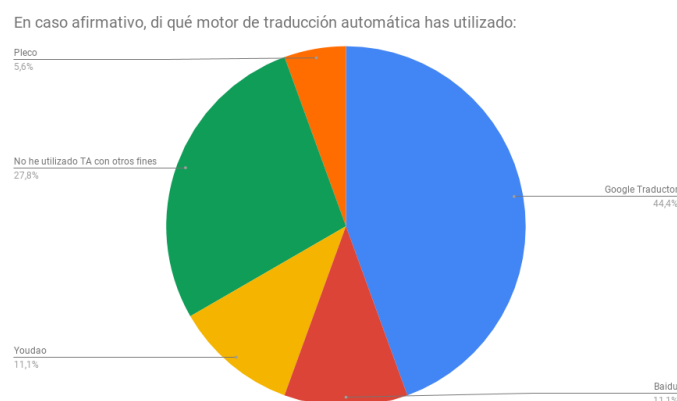


Gráfico 25. Representación de los resultados de la pregunta 10 del cuestionario

El hecho de que un alto porcentaje de los encuestados utilice la TA solo para documentarse podría deberse al hecho de que conocen las deficiencias de la TA en la combinación chino-español. Los resultados de baja calidad que ofrece muchas veces impiden que la TA pueda utilizarse con otras finalidades, sobre todo para traducir. En el siguiente gráfico (gráfico 26) podemos observar que casi la mitad de los encuestados califica su experiencia con la TA con un 5, es decir, han obtenido unos resultados muy mejorables. En general, vemos que las puntuaciones no distan mucho entre sí, ya que todos los estudiantes puntúan su experiencia con la TA con entre 5 y 7 puntos. En este caso no hemos especificado a qué experiencia nos referíamos, si a la experiencia con la combinación inglés-español o chino-español. Tal vez por eso hay alumnos que la puntúan con un 7, ya que los resultados de las traducciones hacia o del inglés siempre son de más calidad.

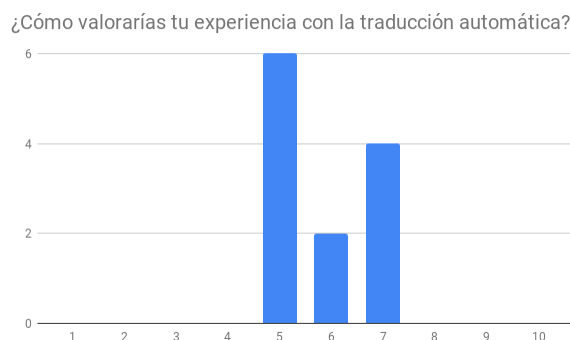


Gráfico 26. Representación de los resultados de la pregunta 11 del cuestionario

En cuanto al grado de especialidad de los textos, vemos que también hay una opinión prácticamente unánime. El 81,3% (gráfico 27) considera que la TA no es útil en contextos más especializados como medicina, economía, etc. Tal y como mostramos a lo largo del trabajo, el grado de especialización es uno de los grandes problemas de la TA en general y, en concreto, de la TAN. La TA todavía necesita más investigación y desarrollo para poder ofrecer un resultado de calidad en traducciones de ámbitos específicos.



Gráfico 27. Representación de los resultados de la pregunta 12 del cuestionario

Las preguntas 13 y 14 están pensadas de forma que podamos, por una parte, comprobar si son capaces de comprender un texto traducido con TA y, por otra, saber cómo valoran los resultados de la TA. En la primera (preg. 13) proponemos un fragmento traducido con TA de los que hemos utilizado para este trabajo y preguntamos a los estudiantes si son capaces de identificar de qué se está hablando. Como podemos ver (gráfico 28), la gran mayoría (66,7%) ha podido reconocer de qué monumento se habla: la Gran Muralla. Las otras opciones dadas por los estudiantes no especificaban la Gran Muralla, pero sí han intuido que se hablaba de un edificio histórico importante. Es un fragmento traducido por Google Traductor, lo que significa que la calidad de la TA es más alta y, por lo tanto, es más fácil de entender.

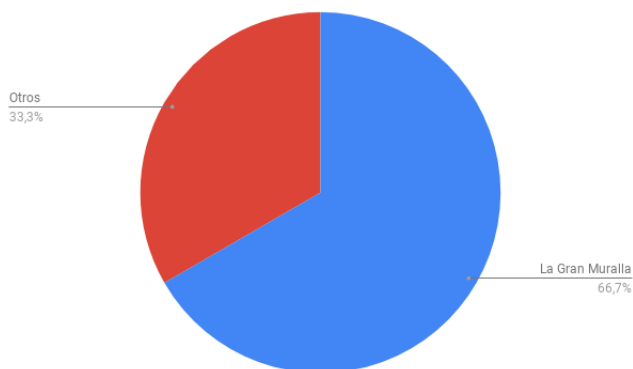


Gráfico 28. Representación de los resultados de la pregunta 13 del cuestionario

En la pregunta 14 (gráfico 29) hemos pedido a los estudiantes que puntuaran un fragmento de un texto, también utilizado en este trabajo, del 1 al 10, siendo 1 el valor más bajo y 10 el valor más alto. En esta ocasión el texto también trataba sobre la Gran Muralla. Podemos observar que la puntuación más frecuente ha sido un 7 y, teniendo en cuenta la baja calidad de la TA del chino-español, es una puntuación relativamente alta. Esto quiere decir que las ideas principales del texto pueden comprenderse sin mucha dificultad. Es importante aclarar también que se trata de un fragmento traducido por Google Traductor, lo que hace que la calidad sea mejor.

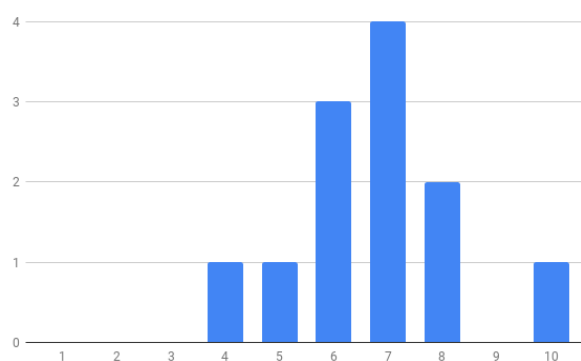


Gráfico 29. Representación de los resultados de la pregunta 14 del cuestionario

Hemos querido conocer la opinión de los estudiantes sobre cuáles creen que son los principales fallos de la TA del chino al español. Tenían cinco opciones, cuatro dadas por nosotros y una respuesta libre en la que podían añadir otras opciones. Los resultados son los siguientes (gráfico 30): léxico y estructura del texto 28% respectivamente, gramática 24% y contenido 20%. Esto confirma lo que hemos ido viendo a largo del trabajo: traducir el léxico e identificar y trasladar las estructuras del chino al español es otro de los principales problemas de la TA CH-ES.



Gráfico 30. Representación de los resultados de la pregunta 15 del cuestionario

Para finalizar nuestro cuestionario, hemos querido saber qué opinan ellos de la TA como herramienta de documentación y que, en caso de que crean que no es útil, cuál

creen que es la mejor solución para comprender bibliografía en chino. Estas son algunas de las respuestas obtenidas:

- Creo que es útil para entender de qué habla el texto. Sin embargo, para documentación no. Siempre es mejor documentarse con libros o páginas web especializadas en temas de cultura china.
- Considero que puede ser útil como filtro para decidir si un texto es de interés o no, según lo que se busque. Puede dar una idea aproximada de los puntos principales que trata. Si se quiere comprender el texto con profundidad no creo que sea una herramienta útil.
- Diccionarios más intuitivos y con puntos de gramática.
- Si no es muy específico, sí.
- Creo que es un buen sitio por el que empezar para captar el concepto del texto.
- Sí, aunque debe utilizarse como apoyo y no como único método para resolver los posibles problemas.
- La traducción automática puede ayudar para dar una pequeña introducción y para hacerse una noción sobre un tema.
- Buscar esa información en inglés o alemán.
- Sí, pero solo para tener una idea de las cosas que estás buscando.

Estas respuestas podríamos resumirlas en que la TA es una buena herramienta para hacerse una idea general sobre qué trata un texto, es decir, tener un punto desde el que empezar a investigar. Por otro lado, la TA resulta más útil en contextos generales, ya que en temáticas muy especializadas más que ayudar podría causarnos más problemas y dar lugar a confusión. Otro punto importante a resaltar es que la TA no debe ser nuestra única opción para poder documentarnos, sino que siempre debe ser una herramienta más que tendremos que complementar con otras (gráfico 31).

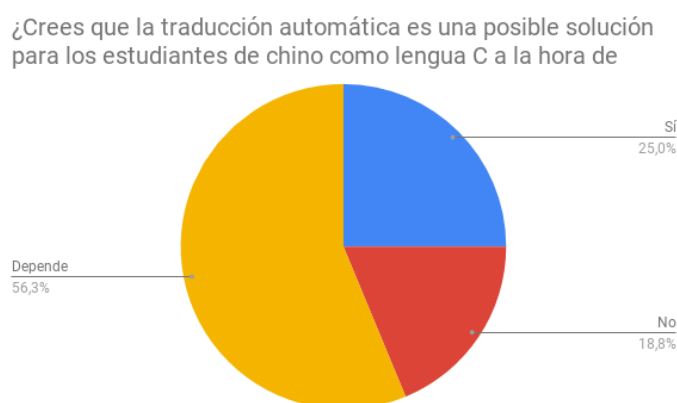


Gráfico 31. Representación de los resultados de la pregunta 16 del cuestionario

En la última pregunta hemos añadido un apartado de observaciones por si los alumnos querían añadir alguna otra opinión o comentario sobre el tema en cuestión. Dos personas han respondido lo siguiente:

- La traducción automática ayuda a aquellos que todavía no son expertos en el idioma y que, por consiguiente, no pueden entender el contenido del texto con una primera lectura. No es una solución al 100% fiable, pero sirve como ayuda para después elaborar la traducción de un texto. Muchas veces leemos las oraciones y, aunque entendamos el significado de las palabras, no comprendemos el sentido de la oración. Gracias al traductor automático podemos descifrar este sentido (aunque a veces resulte inútil) para luego reformular una oración en la lengua meta adecuada.
- Siempre utilizo traducción automática chino-inglés, inglés-chino. Con español el nivel de traducción es muy bajo.

5.3. Conclusiones

Tras el análisis de los resultados obtenidos con este cuestionario, hemos llegado a diferentes conclusiones. La principal es que la TA no resulta útil si se utiliza como única herramienta en tareas de documentación de bibliografía en chino. Nos puede ayudar a formarnos una idea general acerca de qué habla el texto, pero no a comprenderlo en profundidad. Por eso, es recomendable seguir profundizando en nuestra investigación con otras herramientas, como diccionarios y/o enciclopedias.

Hemos podido comprobar que el traductor automático más utilizado, con indiferencia del uso que se le dé, es Google Traductor. No obstante, sí hemos podido observar que los alumnos de cuarto curso son los que han propuesto otras opciones además de Google Traductor, como por ejemplo Youdao, lo que es comprensible, ya que, a más años de estudio del chino, más recursos se conocen.

Todos los estudiantes han probado en alguna ocasión la TA y, por la puntuación que han dado, su experiencia es muy mejorable. Todos coinciden en que en textos con temáticas para cualquier tipo de público podría llegar a funcionar. Por el contrario, tienen claro que en textos especializados no podrían hacer uso de la TA de ninguna manera, ya que esto podría llegar a entorpecer más el resultado de la traducción.

Por último, todos están de acuerdo también en que actualmente existen muy pocos recursos para poder acceder y comprender bibliografía original en chino. Este es el motivo principal que les lleva a utilizar la TA en muchas ocasiones, aunque esta no resuelva todos sus problemas. La clave está en saber escoger la combinación adecuada de herramientas que faciliten nuestro trabajo y nos ahorren tiempo, ya que como traductores tanto el proceso de documentación como el tiempo son aspectos muy importantes.

6. CONCLUSIONES

En este último capítulo del cuerpo del trabajo presentaremos las conclusiones a las que hemos llegado y si hemos podido cumplir con todos los objetivos que nos planteamos al inicio de la investigación. Empezaremos respondiendo a los objetivos y preguntas de investigación de este trabajo:

- 1. Conocer qué es la traducción automática.** Hemos conocido en profundidad en qué consiste exactamente la TA de la mano de profesionales como Hutchins y Somers, quienes nos han ofrecido definiciones detalladas sobre la TA. La TA es mucho más que traducir palabras, también juegan un papel muy importante las matemáticas, ya que todos los sistemas incluyen alguna fórmula o algoritmo en su funcionamiento. El breve repaso por la historia de la TA nos ha ayudado a entender toda su evolución e importancia. Esta evolución está marcada por la globalización y los rápidos avances de la era tecnológica que estamos viviendo. Si bien hubo un periodo de estancamiento tras el informe ALPAC, hoy es un sector en crecimiento en el que todavía queda mucho por descubrir y mejorar.
- 2. Ver el funcionamiento de la TA.** Hemos conocido cuáles son los tipos de TA más utilizados y su funcionamiento, aunque nos hemos centrado más en el sistema de TA en auge: los sistemas neuronales. Sin embargo, no es suficiente con conocerlos, también es importante saber identificar cómo es un buen sistema de TA. Para ello, hemos investigado qué son las normas de calidad de TA de la mano de MQM (2015). Estas normas nos han servido para entender mejor qué tipos de error comete la TA y por qué. Los factores más importantes que deben tenerse en cuenta para obtener una buena traducción son la adecuación, la fluidez y la terminología.
- 3. Comprobar la eficacia de la TA.** A pesar de que cada vez se utiliza más en los procesos de traducción, la TA no solo posee grandes ventajas, también tiene sus propias dificultades. No deja de ser una traducción llevada a cabo por una máquina, por lo que toda la parte subjetiva que puede tener un texto la omite, especialmente en textos con una gran carga emocional. Por esta razón, en traducción literaria no se suele utilizar TA. Por el contrario, esta suele funcionar mucho mejor en textos técnicos, es decir, de carácter objetivo. La principal gran ventaja de la TA es que es mucho más rápida y rentable que la traducción humana. Sin embargo, la desventaja es que cuanto más grande es la distancia entre dos lenguas, peores son los resultados que ofrece. Esto es lo que ocurre en el caso del chino-español. La lejanía y las diferencias entre ambas culturas hace que exista muy poco material traducido y, por ende, pocos documentos que los motores de traducción puedan utilizar para entrenarse.
- 4. Realizar una tarea de posedición.** Como hemos podido ver, cuando la TA falla las consecuencias pueden llegar a ser desastrosas en algunas ocasiones. Por eso es imprescindible realizar una posedición de la TA, especialmente en textos que cuenten

con muchos detalles importantes (como por ejemplo un manual de instrucciones). La posesición también se rige por una serie de normas de calidad imprescindibles para que el resultado final sea perfecto. Actualmente, las más conocidas y respaldadas por profesionales son las que propone TAUS (2016), las cuales siguen muchas empresas de traducción. Estas se dividen en dos tipos de normas, según el resultado que deseemos obtener: un texto con la calidad suficiente para que puedan entenderse las ideas básicas o bien un texto con una calidad igual o superior a la que realizaría un traductor humano. En nuestro caso, la calidad de los resultados obtenidos era muy baja en general, por lo que hemos tenido que realizar una posesición de calidad igual o superior a la de un traductor humano. No obstante, en algunas ocasiones el grado de ininteligibilidad era tan alto que hemos tenido que reescribir el texto de nuevo.

5. Conocer la situación de la traducción automática chino-español. Para poder realizar este trabajo hemos tenido que buscar información sobre cuál es la situación actual de la TA tanto en España como en China. Esta información ha sido importante para entender el funcionamiento y los resultados obtenidos con Google Traductor y Baidu Translate. Así, hemos podido saber que ambos traductores trabajan ya con TAN. No obstante, uno es de alcance mundial (Google) y el otro se limita al país asiático (Baidu). En España la TA existe y, a pesar de que su uso está en expansión, todavía no está suficientemente implantada en las empresas españolas de servicios lingüísticos. En China, la TAN también es un mercado en expansión, aunque necesita mucho más input y experiencia que Google.

6. Comparar los traductores automáticos de Google y Baidu. Tras el análisis de todos los textos y la clasificación de los errores encontrados, podemos afirmar con seguridad que Google Traductor es el que más ha facilitado nuestro proceso de traducción. Con las versiones de Google hemos podido tener una idea general del texto desde la primera lectura y, una vez identificados los errores, el porcentaje de texto que hemos podido aprovechar para la posesición ha sido mucho mayor. Por el contrario, Baidu ha sido el que más partes ininteligibles tenía, por lo que prácticamente hemos tenido que reescribir casi todos los textos. Como hemos mencionado anteriormente, esto se debe a que, primero, la cantidad de input del que dispone Google es mucho más alta que la de Baidu, ya que Google es utilizado a nivel mundial mientras que el uso de Baidu se limita a China. Esta falta de entrenamiento se nota y mucho en los resultados que ofrece. Entre los problemas más frecuentes, tanto de Google Traductor como de Baidu Translate, encontramos el de mala traducción, el de ininteligibilidad y el de terminología. Estos errores explican también los resultados incoherentes o sin sentido que obtenemos muchas veces o que vemos en la tele, en publicidad, etc. No obstante, vemos que ambos respetan la ortografía y la cantidad de elementos del texto, es decir, cometen pocas omisiones, adiciones y no dejan partes sin traducir.

- 7. Proponer una versión poseditada.** Siguiendo la norma de calidad de TAUS (2016) para obtener un texto con una calidad igual o superior a la de un traductor humano hemos realizado nuestra propuesta de posedición. Para ello, hemos tenido que llevar a cabo una fase previa de documentación sobre todos los temas tratados para saber exactamente de qué hablaban los textos y encontrar soluciones a los problemas de traducción que nos han surgido.
- 8. Saber cuál es la opinión de los estudiantes.** La opinión de los estudiantes podemos resumirla en que la TA no es un método para documentarse, sino una herramienta más que nos puede ser de utilidad en la fase de documentación previa de un proceso de traducción. Si no todos, la gran mayoría de los estudiantes hemos utilizado o utilizamos la TA para poder comprender grandes cantidades de texto que, si intentáramos comprender solo con la ayuda de un diccionario, nos llevaría demasiado tiempo.

La respuesta al principal interrogante de este trabajo (¿Los estudiantes de chino no nativos pueden utilizar la TA para acceder y comprender bibliografía escrita originalmente en chino?) es que pueden utilizarla como una herramienta más en combinación con otras, pero nunca utilizar exclusivamente la traducción automática, y esto se debe sobre todo a la falta de entrenamiento de los motores con textos traducidos en la combinación chino-español. Hasta hace algunos años, China apenas tenía contacto con otros países del mundo, sobre todo con los más lejanos, como España. La apertura del país hizo que empezara a conocerse tanto la lengua como la cultura chinas, pero otros países ya le llevaban ventaja. Cuando empezaron a llegar las primeras traducciones del chino al español sobre temas culturales, en España ya había muchas traducciones de otros países. Estas traducciones ya formaban parte del input de los motores de TA que traducen al español, como Google Traductor. Es cierto que la cultura china cada vez es más conocida y ya hay una importante comunidad china en nuestro país, pero todavía falta mucho para que la TA sea realmente útil. Lo mismo ocurre en China. Cuando empezó a añadir input sobre otros idiomas, países como Estados Unidos o Europa ya estaban muy avanzados.

No obstante, el traductor que más nos puede ayudar para realizar tareas de documentación es Google Traductor. Su porcentaje de error es mucho más bajo que el de Baidu translate y, a pesar de que no podamos comprender el texto en su totalidad, sí se puede llegar a tener una idea general sobre el texto y así decidir si puede sernos de utilidad o no. Con Baidu Translate hemos aprendido que no resulta rentable utilizarlo, al menos no en grandes cantidades de texto. Como hemos podido comprobar, en ocasiones no ha ofrecido buenas traducciones ni siquiera para los títulos. Por lo tanto, todavía no podemos calificarla como una herramienta útil ni para tareas de documentación ni mucho menos para traducir.

Este trabajo nos ha servido para conocer más en profundidad el ámbito de la TA y ser conscientes de la importancia que tiene hoy día. Hemos podido conocer dos aspectos principales: por un lado, entender mejor el funcionamiento de la TA en general y, por otro, conocer cuál es la situación de la TA en la combinación chino-español, ya que el chino es nuestra lengua C y será una de las lenguas con las que trabajemos como traductores. Antes de realizar este proyecto, creíamos que la TA era algo mucho más sencillo, algo que había creado alguien desconocido para ahorrarse el tener que aprender idiomas y que por eso a veces los resultados eran tan desastrosos. No obstante, hemos podido comprobar que no es así.

La TA nació hace ya muchos años con el fin de agilizar las comunicaciones entre los países del mundo mediante el uso de un solo idioma. No obstante, este ha sido siempre un objetivo demasiado ambicioso y difícil de llevar a cabo, ya que nadie está dispuesto a renunciar a su lengua y cultura, es decir, a la propia identidad. Por este motivo se optó por la traducción a través de ordenadores, ya que así se obtendrían resultados con más rapidez. Los avances en las investigaciones sobre TA poco a poco fueron creando los sistemas que aún todavía se siguen utilizando (TAE, TABR) y que sirvieron de base para los nuevos (TAN).

Sin embargo, crear un sistema capaz de traducir todos los idiomas del mundo también es un proyecto muy ambicioso que, aunque no es imposible, sí requiere de mucho tiempo. El tiempo es la clave para que poco a poco la TA se perfeccione y logre alcanzar los mismos niveles de calidad que podría ofrecer un traductor humano. En el presente, únicamente puede utilizarse como apoyo o como una herramienta más en un proceso de traducción, ya que la calidad final todavía no es lo suficientemente buena como para utilizar solo TA. Por ello, es muy importante también la fase de posedición. Durante esta fase se revisa y/o modifica todo aquello que la TA no haya podido resolver correctamente. Esta es la práctica más habitual hoy en día en muchas empresas de todo el mundo, ya que es la forma más rápida y eficaz de traducir y lograr un resultado de calidad. Estas son las palabras clave que describen a la perfección la TA: rapidez, eficacia y rentabilidad.

También es importante tener en cuenta que no todos los traductores automáticos son iguales, por lo que no ofrecen los mismos resultados ni la misma calidad. Por ello, es imprescindible conocer las normas de calidad de la TA para poder identificar y clasificar los errores que nos encontremos y así ver qué traductor automático nos ayudará a ser más eficaces o, por el contrario, ralentizará nuestro trabajo. Estas normas son las que han ayudado a comprobar qué traductor automático puede resultarnos más útil y a poseditar para lograr un buen resultado final. Asimismo, también nos han

servido para aprender nuevos tipos de error propios de la TA y a repasar los que ya conocíamos.

El gran descubrimiento de este trabajo ha sido la TAN. Hasta ahora habíamos escuchado muy poco sobre este nuevo sistema de TA, pero tras la investigación que hemos llevado a cabo hemos podido conocerla más en profundidad. Lo que más nos ha llamado la atención es que es un sistema que se rige con algoritmos, es decir, con las matemáticas, y jamás pensamos que la lengua podría llegar a estar tan relacionada con ellas. Conocer sus ventajas y desventajas también nos ha ayudado a realizar la comparación de Google y Baidu, ya que ambos motores trabajan ya con este sistema. En este punto la TA ha dado un gran paso, ya que ahora, además de TAN, también hablamos de inteligencia artificial.

La rápida implementación de la TAN en empresas y aplicaciones de todo el mundo ha abierto un controvertido debate sobre la invasión de la inteligencia artificial cada vez más en nuestro día a día. Aquí es donde surge una de las preguntas que muchos traductores humanos se plantean últimamente: ¿la TA sustituirá a la figura del traductor humano? Tal y como muchos expertos afirman, no pensamos que sea así. Es cierto que la TA ha reducido el tiempo y coste de las traducciones, pero si los traductores automáticos funcionan es gracias a los traductores humanos. Todos los sistemas los han diseñado los humanos y así seguirá siendo, las máquinas no son capaces de crear e inventar por sí solas, siempre hay un humano detrás. Lo mismo ocurre con la TA: siempre será necesaria la figura humana para crear y actualizar el input que utiliza la TA y para realizar la posedición.

Con respecto a la traducción del chino al español, creemos que todavía queda mucho camino por recorrer, tanto al traductor de Google como al de Baidu. Asimismo, también pensamos que esta evolución a mejor de la TA CH-ES no tardará en llegar, ya que ahora la tecnología avanza a un ritmo de vértigo, y más aún cuando se trata de importar un producto (en este caso el idioma y la cultura chinas) a un país. En los últimos años, el descubrimiento de China ha suscitado un gran interés en España y viceversa, los chinos también se sienten muy atraídos por nuestro idioma y nuestra cultura. Esto provoca que la demanda de contenido tanto en España como en China sea cada vez más alta. Además, China es un país que también cuenta con un gran desarrollo tecnológico, lo que creemos que facilitará esta mejora de sus servicios de traducción. De cualquier manera, en ambos países la investigación continuará, pues tarde o temprano la globalización les obligará a adaptarse a las nuevas necesidades de la población.

Creo que este es un campo muy amplio y en expansión, pero del que todavía existen muy pocas investigaciones. En el futuro nos gustaría seguir con este proyecto para poder contribuir, en la medida de lo posible, en la mejora de la calidad de la traducción

automática del chino al español. Una de las posibles vías de investigación sería la creación de un nuevo motor de traducción solo para esta combinación lingüística que tuviera en cuenta todos los elementos del chino que son diferentes a los del español, por ejemplo la puntuación o el orden de las palabras, y que no tuviera tantos problemas a la hora de realizar una traducción con terminología especializada, algo que Google y Baidu todavía no pueden resolver con eficacia, en especial Baidu. Otra opción sería investigar más en profundidad el funcionamiento de la TAN y cómo podríamos adaptarla a las necesidades de la traducción chino-español.

7. BIBLIOGRAFÍA

FUENTES CONSULTADAS Y CITADAS

- Baidu Baike (2018). 百度翻译 (Baidu Translate). Recuperado de:
<https://baike.baidu.com/item/%E7%99%BE%E5%BA%A6%E7%BF%BB%E8%AF%91> [Consulta: 8 diciembre 2018].
- Cabré, M. T. (1992). *La terminologia. La teoria, els mètodes, les aplicacions*. Barcelona: Empúries.
- Carrozo, M. (2017). «Promocionando la Traducción Automática Neuronal: “Es muy fácil superar a Google”». *Unbabel Blog*. Recuperado de:
<https://unbabel.com/blog/es/promocionando-la-traduccion-automatica-neuronal-es-muy-facil-superar-google/> [Consulta: 17 noviembre 2018].
- Casacuberta, F.; Peris, A. (2017). «Traducción automática neuronal». *Revista Tradumàtica. Número 15*, pp. 66–74. Recuperado de:
<https://doi.org/10.5565/rev/tradumatica.203> [Consulta: 6 diciembre 2018]
- Centelles, J. (2013). «Machine Translation for Chinese-Spanish: Experimenting with online Statistical and Rule-Based Paradigms», Tesis doctoral. Recuperado de:
<https://upcommons.upc.edu/bitstream/handle/2099.1/19031/PFC-jordiCentelles.pdf;sequence=4> [Consulta: 6 diciembre 2018]
- Córdoba, I.; Rico, C. *et al* (2015). «Estudio de viabilidad para la implantación de la traducción automática en la empresa VITAE». Recuperado de:
https://www.academia.edu/13019044/Estudio_de_viabilidad_para_la_implantaci%C3%B3n_de_la_traducci%C3%B3n_autom%C3%A1tica?auto=download [Consulta: 6 diciembre 2018]
- Costa-Jussà, M. R.; Aldón, D.; Fonollosa, J. A. (2017). «Chinese-Spanish Neural Machine Translation Enhanced with Character and Word Bitmap Fonts», *UPCommons*. Recuperado de:
<https://upcommons.upc.edu/bitstream/handle/2117/104271/bitmap-mt.pdf> [Consulta: 6 febrero 2019]
- Dixit (2013). «La traducción automática». *Dixit Traducción e Interpretación*. Recuperado de: <http://www.dixit.es/la-traduccion-automatica/> [Consulta: 27 octubre 2018]
- ETransFair (2016-2019). «How to achieve innovative, inclusive and fit-for-market specialized translator training? –A transferable model for training institutions», *ETransFair*. Recuperado de: <https://etransfair.eu/about/etransfair> [Consulta: 16 marzo 2019]

- Fu, A. (1999). «The Research and Development of Machine Translation in China». *MT Summit VII*. Recuperado de: <http://mt-archive.info/MTS-1999-Fu.pdf> [Consulta: 6 febrero 2019]
- German Research Center for Artificial Intelligence (DFKI); QTLaunchPad (2015). «Multidimensional Quality Metrics (MQM) Definition», *QTLaunchPad*. Recuperado de: <http://www.qt21.eu/mqm-definition/definition-2015-12-30.html> [Consulta: 9 febrero 2019]
- Gilbert, W. K.; Chang, H. (1963). «Machine Translation of Chinese». *Scientific American*, 208 (6), pp. 124–137. Recuperado de: <https://pdfs.semanticscholar.org/ddad/955d1ec9fc0bbfd495d063c5bdbc7e1ba6aa.pdf> [Consulta: 8 diciembre 2018]
- Hurtado, A. (1995). «La didáctica de la traducción. Evolución y estado actual». *Perspectivas de la traducción*. Eds. P. Fernández Nistal y J. M. Bravo Gozalo. Valladolid: Universidad de Valladolid, pp. 49-74.
- Hutchins, W. J.; Somers, H. L. (1992). *An Introduction to machine translation*. London [etc.]: Academic Press.
- Justo, D. (2018). «Esto es lo que pasa cuando el traductor de Google no funciona como debería». *Cadena SER*. Recuperado de: http://cadenaser.com/ser/2018/04/12/gente/1523514631_113076.html [Consulta: 27 octubre 2018]
- Koehn, P., & Knowles, R. (2017). «Six Challenges for Neural Machine Translation». *Proceedings of the First Workshop on Neural Machine Translation*, pp. 28–39. Recuperado de: <http://www.statmt.org/wmt17/> [Consulta: 27 abril 2019]
- Liang, Q., Wu, F. y He, G. (2003). «An Overview of Translation in China: Practice and Theory». *Translation Journal*, 7 (2). Recuperado de: <http://translationjournal.net/journal/24china.htm> [Consulta: 6 febrero 2019]
- Parra, C. (2018). «Historia de la traducción automática». *La linterna del traductor*, 16. Recuperado de: http://www.lalinternadeltraductor.org/pdf/lalinterna_n16.pdf [Consulta: 6 diciembre 2018].
- Massardo, I., van der Meer, J., O'Brien, *et al* (2016). «MT Post-Editing Guidelines», *TAUS*, 16–18. Recuperado de: <https://github.com/ygraham/direct-assessment/blob/master/post-editing-guidelines> [Consulta: 4 mayo 2019]
- Temple, J. (2018). «Baidu muestra en directo la eficacia de su nuevo traductor simultáneo». Traducción de Mariana Díaz. *MIT Technology Review*. Recuperado de: <https://www.technologyreview.es/s/10104/baidu-muestra-en-directo-la-eficacia-de-su-nuevo-traductor-simultaneo> [Consulta: 6 diciembre 2018]
- Tilde (2017). «Comparative Evaluation: Statistical vs. Neural Machine Translation», *Tilde*. Recuperado de: <https://www.tilde.com/about/news/316> [Consulta: 27 abril 2019]

- Torres-Hostench, O.; Presas, M. *et al* (2016). «El uso de traducción automática y posesición en las empresas de servicios lingüísticos españolas: Informe de investigación ProjeTA 2015», *Dipòsit Digital de Documents de la UAB*. Recuperado de: <https://ddd.uab.cat/record/148361> [Consulta: 3 noviembre 2018]
- Tradel Barcelona (2017). «Errores de los traductores automáticos». *Tradel Barcelona traductores jurados y técnicos*. Recuperado de: <https://blog.tradel-barcelona.com/2017/01/31/errores-de-los-traductores-automaticos/> [Consulta: 27 octubre 2018]
- Vilarnau, J. (2006). «Àmbits d'ús de la traducció automática». *Revista Tradumàtica: Tecnologies de La Traducció. Número 4*. Recuperado de: <http://www.fti.uab.cat/tradumatica/revista> [Consulta: 6 diciembre 2018]
- Viver, 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. Recuperado de: <http://uvadoc.uva.es/bitstream/10324/33981/1/TFG-O-1431.pdf> [Consulta: 9 febrero 2019]
- WIKIPEDIA (2018). *Traductor de Google*. Recuperado de: https://es.wikipedia.org/wiki/Traductor_de_Google [Consulta: 8 diciembre 2018].
- Wu, Y., Schuster, M., *et al* (2016). «Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation». Recuperado de: <https://arxiv.org/pdf/1609.08144v2.pdf> [Consulta: 6 diciembre 2018]
- Zhongjun, He. (2015). «Baidu Translate: Research and Products», *61 Proceedings Of the ACL 2015 Fourth Workshop on Hybrid Approaches to Translation (HyTra)*, pp. 61–62. Recuperado de: <http://www.aclweb.org/anthology/W15-4110> [Consulta: 10 febrero 2019]

FUENTES CONSULTADAS Y NO CITADAS

- Baidu Baike (2018). 百度机器翻译 (Baidu Traducción Automática). Recuperado de: <https://baike.baidu.com/item/%E7%99%BE%E5%BA%A6%E6%9C%BA%E5%99%A8%E7%BF%BB%E8%AF%91> [Consulta: 8 diciembre 2018]
- Ding, Y., Liu, Y., Luan, H., & Sun, M. (2017). «Visualizing and Understanding Neural Machine Translation». Recuperado de: http://nlp.csai.tsinghua.edu.cn/~ly/papers/acl2017_dyz.pdf [Consulta: 27 abril 2019]
- Herranz, M. (2014). «Poseditores: ¿Por qué la traducción automática en crudo necesita posesición?», *Pangeanic*. Recuperado de: <https://pangeanic.es/conocimiento/poseditores-por-que-la-traducccion-automatica-en-crudo-necesita-posedicion/#> [Consulta: 6 diciembre 2018]

- Herranz, M. (2018). «Traducción automática: el informe ALPAC». *Pangeanic*. Recuperado de: <https://pangeanic.es/conocimiento/traduccion-automatica-informe-alpac/#> [Consulta: 6 diciembre 2018]
- Oliver A., Vázquez M. *et al* (2017). «Machine Translation post-editing at TransPerfect – The ‘Human’ Side of the Process». *Revista Tradumàtica: Tecnologies de la Traducció*, 15, 1-11. Recuperado de: <https://doi.org/10.5565/rev/tradumatica.201> [Consulta: 29 septiembre 2018]
- Rao, D. (2018). «The Real Problems with Neural Machine Translation», *Deliprao*. Recuperado de: <http://deliprao.com/archives/301> [Consulta: 27 abril 2019]
- Rico, C.; Martín, A. *et al* (1997). «New Trends in Machine Translation», *Meta*, 42(4), pp. 605–615. Recuperado de: <https://doi.org/10.7202/003822ar> [Consulta: 3 noviembre 2018]
- SDL Trados (2018). «¿Qué es la traducción automática?», *SDL Trados*. Recuperado de: <https://www.sdltrados.com/es/solutions/machine-translation.html> [Consulta: 6 diciembre 2018]
- SYSTRAN (2018). «¿Qué es la traducción automática? Traducción automática basada en reglas vs. traducción automática estadística», *SYSTRAN*. Recuperado de: <http://www.systran.es/systran/tecnologia-de-traduccion/que-es-la-traduccion-automatica/> [Consulta: 6 diciembre 2018]
- Traducción Automática (2011). «Tipos de traducción automática», *Traducción automática*. Recuperado de: <http://traduccionautomaticaevenin.blogspot.com/2011/01/tipos-de-traduccion-automatica.html> [Consulta: 6 diciembre 2018]
- Xinyu, Z. (2017). «Traducción de los nombres de los platos chinos y su relación con herramientas de traducción automática», *Dipòsit Digital de Documents de la UAB*. Recuperado de: https://ddd.uab.cat/pub/trerecpro/2017/hdl_2072_307046/xinyu_zhang_tfm.pdf [Consulta: 8 diciembre 2018]

DICCIONARIOS

- Harper Collins (1979). Collins Dictionary (13.a ed.). Consultado en: <https://www.collinsdictionary.com/>
- Naver (2019). LINE Dictionary: Chinese-English (versión 1.4.6) [Aplicación móvil]. Descargado de: https://play.google.com/store/apps/details?id=com.naver.android.globaldictcnen&hl=en_US
- Pleco Inc (2019). Pleco Chinese Dictionary (Versión 3.2.64) [Aplicación móvil]. Descargado de: <https://play.google.com/store/apps/details?id=com.pleco.chinesesystem>

Real Academia Española (2001). *Diccionario de la lengua española* (22.a ed.).

Consultado en: <https://dle.rae.es/>

TEXTOS ANALIZADOS

China Baike (2007). 中医的历史 (*Zhōngyī de lìshǐ*) [Historia de la medicina tradicional china]. Recuperado de:

<http://www.chinabaike.com/article/292/304/2007/2007021945441.html> [Consulta: 28 enero 2019]

Hainan Daily (2018). 三亚后海：探寻隐世渔村的浪漫情调 (*Sānyà hòuhǎi: tàn xún yǐn shì yú cūn de làngmàn qíngdiào*) [Houhai, Sanya: explorando el ambiente romántico del pueblo de pescadores escondido]. Recuperado de:

<http://www.hi.chinanews.com/travel/content/2018-09-19/21696.html> [Consulta: 28 enero 2019]

Oficina Nacional de Turismo (2005). 中国的四大菜系分别是哪些? (*Zhōngguó de sì dà cài xìfēn bié shì nǎ xiē?*) [¿Cuáles son los cuatro estilos culinarios chinos?]

Recuperado de: http://www.gov.cn/test/2006-03/28/content_238209.htm [Consulta: 28 enero 2019]

Oficina Nacional del Patrimonio Cultural (2006). 世界文化遗产——长城 (*Shìjiè wénhuà yí chǎn——Chángchéng*) [Patrimonio Cultural de la Humanidad – La Gran Muralla China]. Recuperado de: http://www.gov.cn/test/2006-03/28/content_238209.htm [Consulta: 28 enero 2019]

Public Holidays (2019). 2019与2020年春节 (*2019 yǔ 2020 nián chūnjié*) [La Fiesta de la Primavera de 2019 y 2020]. Recuperado de: <https://publicholidays.cn/zh/spring-festival/> [Consulta: 28 enero 2019]

Yanlin, Y. (杨雁琳) (2017). *Zhōngguó shūfǎ zào yì zuì shēn de shí dà jiā, bù kě bù zhī* (中国书法造诣最深的十大家, 不可不知) [Los diez calígrafos más importantes de China que debes conocer]. Recuperado de: <https://www.cqcb.com/reading/2017-07-15/404375.html> [Consulta: 28 enero 2019]

FIGURAS

German Research Center for Artificial Intelligence (DFKI); QTLaunchPad (2015).

«Multidimensional Quality Metrics (MQM) Definition», *QTLaunchPad*. Recuperado de: <http://www.qt21.eu/mqm-definition/definition-2015-12-30.html> [Consulta: 9 febrero 2019]

Herramientas para traductores (2012). «Traducción automática: Google Translator Toolkit». Recuperado de:

<https://herramientasparatraductores.wordpress.com/2012/09/13/traduccion-automatica-google-translator-toolkit/> [Fecha de consulta: 6 diciembre 2018]

- Parra, C. (2018). «Historia de la traducción automática». *La linterna del traductor*, 16.
Recuperado de: http://www.lalinternadeltraductor.org/pdf/lalinterna_n16.pdf
[Consulta: 6 diciembre 2018]
- Useless paperclip: Blog de informática y traducción (2017). «Traducción automática basada en reglas». Recuperado de: <https://uselesspaperclip2.wordpress.com/>
[Consulta: 17 noviembre 2018]
- Wibawa, A., A. Nafalski *et al* (2013). «Potential of an educational application of a language statistical translation system». Recuperado de:
<file:///C:/Users/USER/Downloads/11-Wibawa-A.pdf> [Consulta: 8 mayo 2019]

8. ANEXOS

I. Tabla con las puntuaciones para obtener el grado de calidad de los títulos

	GOOGLE		
TEMA	COMPRENSIÓN	EXPRESIÓN	PRAGMÁTICO
Año Nuevo	4	4	3
Caligrafía	4	3	4
Cocina	4	4	4
Muralla china	4	4	4
Medicina	4	4	3
Turismo	4	4	4

	BAIDU		
TEMA	COMPRENSIÓN	EXPRESIÓN	PRAGMÁTICO
Año Nuevo	4	4	1
Caligrafía	3	3	1
Cocina	3	3	2
Muralla china	2	4	1
Medicina	4	4	3
Turismo	2	3	1

II. Tabla con las puntuaciones para obtener el grado de calidad de los textos

	GOOGLE		
TEMA	COMPRENSIÓN	EXPRESIÓN	PRAGMÁTICO
Año Nuevo	3	3	3
Caligrafía	2	2	1
Cocina	1	1	1
Muralla china	4	4	4
Medicina	2	2	3
Turismo	4	3	4

	BAIDU		
TEMA	COMPRENSIÓN	EXPRESIÓN	PRAGMÁTICO
Año Nuevo	2	2	2
Caligrafía	1	1	1
Cocina	1	1	1
Muralla china	3	2	1
Medicina	1	1	1
Turismo	1	1	1

III. Tabla con la suma total de errores de Google

TEMA	ADICIÓN	MALA TRADUCCIÓN	OMISIÓN	NO TRADUCIDO	GRAMÁTICA	REGISTRO GRAMATICAL	INCONSISTENCIA	ORTOGRAFÍA	TIPOGRAFÍA	ININTELIGIBILIDAD	TERMINOLOGÍA
Año Nuevo	1	6	0	0	5	0	1	0	0	0	3
Caligrafía	0	1	1	0	1	0	0	0	2	13	2
Cocina	0	2	2	2	1	0	0	0	5	8	6
Muralla china	0	2	2	0	3	0	1	0	5	1	4
Medicina	0	4	4	0	5	0	0	0	4	2	2
Turismo	0	0	0	0	1	0	0	0	0	0	5
TOTAL	1	15	9	2	16	0	2	0	16	24	22

IV. Tabla con la suma total de errores de Baidu

TEMA	ADICIÓN	MALA TRADUCCIÓN	OMISIÓN	NO TRADUCIDO	GRAMÁTICA	REGISTRO GRAMATICAL	INCONSISTENCIA	ORTOGRAFÍA	TIPOGRAFÍA	ININTELIGIBILIDAD	TERMINOLOGÍA
Año Nuevo	0	9	0	0	2	0	3	0	6	0	3
Caligrafía	0	1	3	0	0	0	1	0	3	19	1
Cocina	0	2	0	0	1	0	1	0	3	10	5
Muralla china	4	5	3	0	2	0	0	0	6	0	3
Medicina	0	7	4	0	3	0	2	0	6	3	2
Turismo	0	5	1	0	0	0	1	0	3	1	2
TOTAL	4	29	11	0	8	0	8	0	27	33	16

V. Cuestionario

Mi nombre es María Páez, alumna de 4º de Traducción e Interpretación de la UAB. Estoy realizando mi TFG sobre el uso que los estudiantes de chino como lengua C pueden hacer de los traductores automáticos de Google y Baidu para poder acceder y comprender bibliografía original en chino. Para ello, necesito que respondas el siguiente cuestionario, que no te llevará más de cinco minutos. No hay respuestas verdaderas o falsas, responde lo que más se ajuste a tu realidad. Los datos son totalmente anónimos y confidenciales y solo se usarán para fines académicos.

1. ¿Cuántos años hace que estudias chino?
 - Menos de 1 año
 - Entre 2 y 3 años
 - Más de 3 años
2. ¿Alguna vez has tenido que buscar bibliografía original en chino?
 - SI
 - NO
3. ¿Cómo has solucionado el problema de la barrera lingüística al leer los textos?
 - Diccionarios
 - Enciclopedias
 - Traductores automáticos
 - Otros
4. ¿Consideras que para un estudiante de chino como lengua C es fácil acceder y comprender bibliografía en chino?
 - SI
 - NO
5. ¿Crees que hay suficientes recursos online para obtener y/o entender bibliografía en lengua china?
 - SI
 - NO
6. ¿Alguna vez has utilizado algún traductor automático de internet?
 - SI
 - NO
7. ¿Cuál has utilizado o utilizas con más frecuencia?
 - Google Traductor
 - Baidu Translate
 - Microsoft Bing
 - Otros
8. ¿Crees que la traducción automática resulta útil para tareas de documentación?
Por ejemplo, imagina que estás traduciendo un texto sobre medicina tradicional

china y necesitas buscar información sobre el tema en páginas web escritas originalmente en chino, ¿crees que la traducción automática te ayudaría a comprender la información?

- SI
- NO

9. ¿En alguna ocasión has utilizado la traducción automática con otros fines, por ejemplo, como diccionario o para traducir textos?

- SI
- NO

10. En caso afirmativo, di qué motor de traducción automática has utilizado:

- Respuesta libre

11. ¿Cómo valorarías tu experiencia con la traducción automática? Puntúa de 1 (muy mala experiencia) a 10 (muy buena experiencia).

- DE 1 A 10

12. ¿Crees que la traducción automática del chino al español podría resultar útil en contextos más especializados, como medicina, economía, etc.?

- SI
- NO

13. ¿De qué crees que trata este texto: «*La gran ciudad es una de las obras más antiguas de defensa de China, la mayor de las cuales se construyó en el mundo. A partir del siglo VII antes del dólal de los Estados Unidos, la construcción de más de 2.000 años, que se remonta a más de 50.000 kilómetros de tierra en el norte y el Centro de China, se ha mantenido en un total de más de 50.000 kilómetros, lo que se denomina "más de 2.000 años y más de más de 100.000 kilómetros".*»?

- Respuesta libre

14. Puntúa de 1 a 10 el grado de comprensión del siguiente texto: «*Evaluación del Comité del Patrimonio Mundial. Alrededor del 220 a. C., Qin Shihuang, quien dominó el mundo, conectó algunas de las fortificaciones intermitentes construidas anteriormente en un completo sistema de defensa para resistir la agresión del norte. En la dinastía Ming (1368-1644 dC), se continuó construyendo para hacer de la Gran Muralla la instalación militar más larga del mundo. Su valor en la cultura y el arte es comparable a su importancia histórica y estratégica.*»

- Puntuar de 1 a 10

15. En general, ¿cuál crees que es el principal fallo en las traducciones automáticas del chino? Puedes seleccionar varias opciones
- CONTENIDO
 - ESTRUCTURA DEL TEXTO
 - LÉXICO
 - GRAMÁTICA
 - OTROS
16. ¿Crees que la traducción automática es una posible solución para los estudiantes de chino como lengua C a la hora de realizar tareas de documentación sobre un tema? Si no es así, ¿cuál crees que sería una posible solución?
- Respuesta libre
17. Observaciones:

¡Muchas gracias por tu colaboración!