

# Traducción automática en la era digital: una evaluación de calidad de ChatGPT frente a DeepL y Reverso en artículos científicos médicos

María Victoria Espel Espel

Trabajo Final de Máster  
Tutor: Dr. Carles Prado Fonts  
Máster de Traducción y Tecnologías  
Universitat Oberta de Catalunya  
Junio de 2024



Universitat Oberta  
de Catalunya

Índice	
Resumen .....	3
Resum .....	4
Abstract.....	5
1. Introducción.....	6
2. Justificación de la propuesta.....	7
3. Objetivos e hipótesis.....	8
4. Metodología.....	9
5. Marco teórico.....	14
5.1. Traducción automática.....	14
5.1.1. Sistemas de traducción automática.....	14
5.2. Inteligencia artificial.....	18
5.2.1. ChatGPT .....	19
5.3. Textos especializados .....	21
5.3.1. Características del lenguaje médico publicado.....	22
5.3.2. Traducción automática en textos especializados .....	23
6. Estudio de caso .....	24
6.1. Resultados.....	24
6.1.1. Análisis global .....	25
6.1.2. Análisis por categoría .....	26
6.2. Impacto del <i>prompt</i> en ChatGPT .....	38
6.2.1. Análisis global .....	38
7. Conclusiones.....	45
8. Bibliografía.....	48
9. Anexos.....	53

# Resumen

En la era digital, la traducción automática se ha convertido en una herramienta con mayor presencia en diversos campos de la traducción. Este estudio se enfoca en evaluar la calidad de las traducciones de artículos científicos médicos realizadas con ChatGPT, comparándolas con las de DeepL y Reverso. El objetivo principal es determinar si ChatGPT, con su inteligencia artificial avanzada y recursos de búsqueda, puede producir traducciones de mayor calidad que DeepL y Reverso, herramientas diseñadas específicamente para la traducción automática.

La metodología se dividió en tres fases, la documentación de los sistemas de traducción basados en redes neuronales y estadística y la función de ChatGPT como herramienta para así crear dos *prompts*, uno sencillo similar al proceso de DeepL y Reverso y otro más completo con características de la tipología textual elegida. En la segunda fase, se seleccionaron los fragmentos de dos artículos científicos médicos para ser traducidos por los tres motores. Finalmente, se realizó un análisis manual de las traducciones siguiendo las métricas de TAUS DQF para medir la eficacia de cada herramienta.

Los resultados revelan que ChatGPT no supera a DeepL en calidad de traducción y Reverso es el peor de los tres. Sin embargo, se destaca la utilidad de ChatGPT en la redacción asistida, mientras que DeepL demostró ser más eficiente en la traducción de terminología. En conclusión, aunque ChatGPT es una herramienta eficiente para la generación de textos y provisión de información, requiere supervisión del contenido que genera. Este estudio es relevante por el creciente uso de ChatGPT en diversos ámbitos profesionales, destacando la necesidad de supervisión de las áreas de mejora en la traducción automática.

**Palabras clave:** traducción automática, textos especializados, artículo científico médico, inteligencia artificial, motores de traducción automática, sistemas de traducción de redes neuronales

# Resum

En l'era digital, la traducció automàtica s'ha convertit en una eina amb major presència en diversos camps de la traducció. Aquest estudi s'enfoca a avaluar la qualitat de les traduccions d'articles científics mèdics realitzades amb ChatGPT, comparant-les amb les de DeepL i Reverso. L'objectiu principal és determinar si ChatGPT, amb la seva intel·ligència artificial avançada i recursos de cerca, pot produir traduccions de major qualitat que DeepL i Reverso, eines dissenyades específicament per a la traducció automàtica.

La metodologia es va dividir en tres fases, la documentació dels sistemes de traducció basats en xarxes neuronals i estadística i la funció de ChatGPT com a eina per a així crear dos prompts, un senzill, similar al procés de DeepL i Reverso i un altre més complet amb característiques de la tipologia textual escollida. En la segona fase, es van seleccionar els fragments de dos articles científics mèdics per a ser traduïts pels tres motors. Finalment, es va fer una anàlisi manual de les traduccions seguint les mètriques de TAUS DQF per a mesurar l'eficàcia de cada eina.

Els resultats revelen que ChatGPT no supera a DeepL en qualitat de traducció i Reverso és el pitjor de tots tres. No obstant això, es destaca la utilitat de ChatGPT en la redacció assistida, mentre que DeepL va demostrar ser més eficient en la traducció de terminologia. En conclusió, encara que ChatGPT és una eina eficient per a la generació de textos i provisió d'informació, requereix supervisió del contingut que genera. Aquest estudi és rellevant pel creixent ús de ChatGPT en diversos àmbits professionals, destacant la necessitat de supervisió de les àrees de millora en la traducció automàtica.

**Paraules clau:** traducció automàtica, textos especialitzats, article científic mèdic, intel·ligència artificial, motors de traducció automàtica, sistemes de traducció de xarxes neuronals.

# Abstract

In the digital era, machine translation has become an increasingly prevalent tool across various fields of translation. This study evaluates the quality of medical scientific article translations produced by ChatGPT, comparing them with those of DeepL and Reverso. The primary objective is to determine whether ChatGPT, with its advanced artificial intelligence and search capabilities, can produce higher-quality translations than DeepL and Reverso, tools specifically designed for machine translation.

The methodology was divided into three phases: documentation of neural network and statistical-based translation systems, examination of ChatGPT's function as a tool to create two prompts (one simple, similar to the process of DeepL and Reverso, and another more complete with characteristics of the chosen text typology). In the second phase, excerpts from two medical scientific articles were selected to be translated by the three engines. Finally, a manual analysis of the translations was conducted using TAUS DQF metrics to measure the effectiveness of each tool.

The results reveal that ChatGPT does not surpass DeepL in translation quality, with Reverso being the worst among the three. However, ChatGPT's utility in assisted writing is highlighted, while DeepL proved to be more efficient in translating terminology. In conclusion, although ChatGPT is an effective tool for text generation and information provision, it requires content supervision. This study is relevant due to the growing use of ChatGPT in various professional fields, emphasizing the need for oversight in areas for improvement in machine translation.

**Keywords:** machine translation, specialized texts, medical scientific article, artificial intelligence, machine translation engines, neural network translation systems

# 1. Introducción

En un mundo cada vez más globalizado y digitalizado, la traducción automática ha emergido como una herramienta fundamental para facilitar la comunicación entre idiomas. Recientemente, con la llegada de ChatGPT y la buena acogida que ha tenido por parte de los usuarios de internet, se ha convertido en una herramienta nueva de traducción. En este contexto, el presente estudio se adentra en la evaluación comparativa de tres motores de traducción: ChatGPT, DeepL y Reverso.

La traducción automática ha evolucionado considerablemente en los últimos años, desempeñando un papel crucial en la vida profesional de los traductores y en la difusión rápida de conocimientos científicos. En particular, la creciente prominencia de ChatGPT, una herramienta basada en inteligencia artificial plantea preguntas intrigantes sobre su eficacia y relevancia, tanto en la traducción no especializada, como en la traducción de textos especializados, donde la precisión y la terminología especializada son imperativas.

El objetivo principal de este estudio es determinar si ChatGPT, a pesar de no estar específicamente diseñado para la traducción, al contrario que DeepL y Reverso, supera en eficiencia a estos motores de traducción automática en artículos científicos médicos con los pares de lenguas inglés y español. Al abordar esta pregunta, buscamos ofrecer una comprensión más profunda de las capacidades y limitaciones de estos sistemas en un contexto especializado.

Este estudio contribuye al cuerpo existente de conocimiento al proporcionar una evaluación sistemática y comparativa de la competencia de los traductores automáticos en el campo de la traducción de textos científicos médicos. Al destacar las fortalezas y debilidades de cada sistema evaluado, esperamos ofrecer información valiosa para los profesionales de traducción y de materias afines.

Para dar contexto a este trabajo comenzamos presentando los objetivos e hipótesis que presentamos y seguido hicimos un breve repaso por los diferentes sistemas de traducción automática, presentamos los motores de traducción que utilizamos y a la herramienta ChatGPT y su función como herramienta de traducción. Como último apartado de nuestro marco teórico, presentamos las características del lenguaje especializado médico y hablamos de la traducción automática en este campo.

Presentamos la metodología del trabajo en la cual se expondrán los textos de trabajo y el método de evaluación manual que empleamos. Y por último, exponemos los

resultados y reflexiones sobre las ventajas y limitaciones de las herramientas de traducción evaluadas.

## 2. Justificación de la propuesta

La presente investigación se centró en probar la efectividad de diferentes motores de traducción automática, ChatGPT, DeepL y Reverso, en la traducción de artículos científicos del ámbito médico. La importancia de esta investigación radica en determinar qué motor de traducción ofrece resultados óptimos en la traducción de textos especializados, en particular en el área médica. Esto permitirá que los usuarios puedan acceder a artículos científicos que no estén disponibles en su lengua materna para poder satisfacer sus necesidades de información, ya sean a nivel personal, académico o profesional.

La evaluación comparativa entre ChatGPT, DeepL y Reverso es crucial para comprender qué motor de traducción es más eficaz en este contexto. DeepL y Reverso son motores de traducción reconocidos y ampliamente utilizados, diseñados específicamente para traducir textos de manera precisa y coherente. En contraste, ChatGPT se basa en inteligencia artificial y su capacidad para traducir textos ha sido objeto de estudio y debate en el mundo de la traducción. Un artículo en el blog de Pangea Global (Dimitriadou, 2023), habla sobre la eficacia de ChatGPT en comparación con otros métodos de traducción automática. El artículo destaca que ChatGPT mejora la traducción, pero presenta desafíos en términos de precisión y contexto. De manera similar, el análisis de CCJK (Bashir, 2023) señala que el impacto de ChatGPT en la industria de la traducción sigue siendo un tema de discusión, especialmente en términos de costos y recursos de tiempo y concluyen que, a pesar de reducir los costos y acelerar el proceso de traducción, la calidad de las traducciones automáticas aún puede ser inconsistente (Bashir, 2023). Además, un artículo de Larroyed (2023) proporciona una visión general de las innovaciones en la tecnología de traducción de ChatGPT y el debate sobre su efectividad, en cual se evalúa el rendimiento de ChatGPT en la traducción de patentes, mostrando que ChatGPT todavía tiene limitaciones en la terminología técnica y la precisión legal (Larroyed, 2023).

Finalmente, Northwood (2023) destaca el debate sobre si ChatGPT ha avanzado en la traducción de expresiones idiomáticas y elementos sutiles del lenguaje, pero también menciona que su uso en situaciones críticas aún es objeto de debate debido a posibles errores y malinterpretaciones (Northwood, 2023).

Los resultados de la presente investigación beneficiarán a aquellos usuarios que necesitan traducir artículos científicos para informarse cuando estos no están disponibles en su lengua materna. Asimismo, también puede servir de ayuda a aquellos investigadores que necesiten traducir sus propios artículos a una lengua diferente a la suya, ya que podrán elegir la herramienta según su eficiencia y la necesidad que necesiten cubrir.

Los resultados obtenidos pueden ayudar a desarrollar o apoyar teorías relacionadas con la traducción automática y su eficacia en la traducción de textos especializados. Este trabajo proporciona información adicional sobre la efectividad de la traducción automática en textos científicos, así como una mayor comprensión del funcionamiento de ChatGPT en el ámbito de la traducción, un campo aún en fase de estudio debido a su novedosa incorporación.

En resumen, la presente investigación tiene como objetivo abordar la necesidad de identificar el motor de traducción más eficaz para artículos científicos médicos, con el fin de facilitar el acceso a la información científica y contribuir al avance del conocimiento en el campo de la traducción automática especializada.

### 3. Objetivos e hipótesis

El objetivo de este trabajo es evaluar y comparar la eficiencia de los motores de traducción automática ChatGPT, DeepL y Reverso en la traducción de artículos científicos en el ámbito médico con el fin de determinar cuál de ellos ofrece los mejores resultados. ¿Es capaz ChatGPT de producir traducciones con mayor calidad al contar con inteligencia artificial y otros recursos de búsqueda? ¿O de lo contrario, DeepL y Reverso pueden producir traducciones de calidad al ser herramientas creadas específicamente para la traducción?

En el estudio comparativo de Lettere Dragosavljevich (2023) sobre Google Translate y ChatGPT, se expone que ChatGPT, en su versión 3.5, da mejores resultados que Google Translate gracias a que el primero tiene acceso a recursos actuales de internet,



lo que le permite referenciar varias versiones del texto y encontrar traducciones adecuadas y por qué es un modelo lingüístico grande, por lo que aplica pragmática y sensibilidad a la traducción mejorando su calidad. Este resultado también está reafirmado por el estudio de Jiao (2023) en el que también se compara ChatGPT con Google Translate, y lo hace en sus versiones 3.5 y 4, siendo esta última la más eficiente entre ambas versiones, por lo que la considera una herramienta comercial competitiva con Google Translate. Sin embargo, en investigaciones recientes hay falta de consenso sobre cuál de estos dos motores de traducción son mejores, por un lado, los hallazgos de Işım y Balcioglu (2023) resaltan la efectividad de las herramientas de traducción automática como Google Translate en el ámbito de las traducciones sencillas. Aun así, destaca sus limitaciones en la precisión y la coherencia de sus traducciones, por lo que recomienda que estas se supervisen debido a que hay errores pueden alterar el significado del texto. Por otro lado, Dogukan et.al (2023), destaca que la inteligencia artificial es más eficiente en el ámbito de la traducción e incluso se atreve a vaticinar que esta profesión puede ser más susceptible a ser sustituida por la inteligencia artificial al considerarla como una actividad repetitiva al ser una tarea de predicción. Basada en la literatura previa existente, este trabajo pretende comprobar la hipótesis de que ChatGPT puede llegar a producir traducciones de mayor calidad que los motores de traducción de redes neuronales.

## 4. Metodología

El presente trabajo, se desarrolló en diferentes fases. En la fase preliminar dimos forma al marco teórico del trabajo en el que se hizo un repaso por los sistemas de traducción automática y sobre ChatGPT como herramienta de traducción para poder saber cómo estos sistemas pueden influir en la traducción de ciertos términos y si son diferentes a los propuestos por ChatGPT al emplear otro tipo de métodos de traducción. También se hace un repaso de las características de los textos especializados, haciendo hincapié en sus características del lenguaje, concretamente el lenguaje médico publicado y la aplicación de la traducción automática en el campo de los textos médicos.

Después de adentrarnos en la materia de los textos especializados y la traducción automática, pasamos a la fase de selección de los textos de trabajo. Se seleccionaron dos artículos científicos médicos publicados en revistas especializadas. El primero de ellos, titulado *The role of cardiologists in stroke prevention and treatment: position paper of the*

*European Society of Cardiology Council on Stroke* (2017), fue extraído de la revista *European Heart Journal* de la European Society of Cardiology. El segundo artículo, *Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report* (2019), fue seleccionado de la revista *The New England Journal of Medicine*.

Para la preparación de los textos, que se utilizaron tanto para su traducción, como para su posterior evaluación, se extrajeron dos fragmentos de cada artículo, cada uno con una extensión de 500 palabras aproximadamente cada uno. Respecto al artículo relacionado con el ictus, se seleccionaron las siguientes secciones: *Carotid stenting* (p.1571), *Pulmonary veins isolation* (p. 1572), *Left atrial appendage occlusion or removal* (p.1572 – 1573), y las conclusiones (p. 1573) del estudio. Por otro lado, en el caso del artículo centrado en enfermedades pulmonares asociadas al uso del cigarrillo electrónico, únicamente se tradujo la sección *Select Clinical Vignette with Radiographs* (p. 910 – 911). Se decidió trabajar con dos fragmentos que no superaran las 1000 palabras en total, debido a que se sometieron a una evaluación manual.

La siguiente fase consiste en la preparación del caso práctico, y el primer paso para ello es traducir los textos con las diferentes herramientas propuestas: ChatGPT, DeepL y Reverso.

Para el caso de ChatGPT, se hicieron dos tipos de traducción: una empleando comandos simples para traducir del inglés al español con el objetivo de equiparar su funcionalidad con la ofrecida por DeepL y Reverso, por lo que solo se le indicó los pares de lenguas y se pegó el texto a traducir, por lo que el *prompt* empleado es el siguiente: «traduce del español al inglés el siguiente texto:».

También se creó otro *prompt* más específico, en el que se incluyeron detalles como el contexto sobre el texto, la tipología textual, lugar de publicación, pares de idiomas, y se le asignó el rol de traductor con el fin de determinar si al proporcionar dicha información adicional esta herramienta se posiciona como la más eficiente entre las tres para la traducción de artículos científicos médicos, o si bien, aunque no destaque como la primera opción, pueda mejorar sus propios resultados. El *prompt* que se utiliza para la segunda traducción es el siguiente: «imagina que eres un traductor especializado en textos médicos. Traduce del inglés al español el siguiente artículo científico que se publicará en una prestigiosa revista científica. El vocabulario debe ser especializado y el registro debe ser formal:».

El proceso de traducción implica el ingreso de los textos en los sistemas de traducción automática. Para DeepL y Reverso, se adjuntaron los textos en un documento

en formato de texto en el sistema y se seleccionan los idiomas de origen y destino, en este caso inglés y español. Para ChatGPT, se utilizan los comandos específicos de acuerdo con el protocolo establecido y el texto se copió directamente en el cuadro de chat después de escribir el *prompt*, debido a que hasta el momento en el que se realizó la investigación, la versión gratuita de ChatGPT no permite adjuntar archivos.

Para poder comparar las traducciones de ambos textos, es necesario tener una traducción humana de referencia, que en este caso la realicé yo misma utilizando como referencia las recomendaciones del *Libro Rojo*, el *Diccionario de dudas y dificultades de traducción del inglés médico* de Fernando A. Navarro (2023e) para que la evaluación manual tenga una coherencia y una referencia a la hora de catalogar los posibles errores o posibles preferencias terminológicas.

Como última fase, tenemos la evaluación de la calidad de las traducciones. Para ello se utilizó la métrica TAUS DQF para la revisión manual, la cual utiliza parámetros de tipología de errores. Siguiendo su tipología de errores, se pudieron evaluar las traducciones con un modelo estándar y dinámico de calidad (TAUS, 2024).

Las traducciones se insertaron fragmento a fragmento dentro del Excel que contiene las métricas de evaluación. Con la información que fuimos introduciendo, se calculó el nivel de calidad de cada texto, para ello se rellenó la plantilla con el número o nombre de referencia de cada fragmento, seguido del texto de origen, a continuación, se añadía la traducción hecha por nuestras herramientas de traducción automática y luego la traducción humana para tenerla de referencia y comparar los posibles errores. Una vez detectados, se categorizaron según su tipología y gravedad.

Para clasificar los tipos de errores contamos con 4 categorías diferentes, propuestas por TAUS DQF, que cuentan con sus propias subcategorías, las cuales vamos a explicar a continuación para poder entender qué se está evaluando de estos textos. Debemos aclarar que hay categorías y subcategorías que propone TAUS DQF que se han eliminado debido a que no aplicaban a nuestro contexto, como pueden ser las categorías de error de código informático.

La primera categoría de análisis es la **fidelidad**, la cual es la falta de precisión del texto de origen (TO), teniendo en cuenta las diferencias autorizadas previamente que puedan existir por parte del cliente que quiere que se traduzca el texto. Este tipo de error, por ejemplo, pueden ser errores de traducción como los falsos amigos. Además, tenemos las siguientes subcategorías:

- **Incorporación:** el texto meta (TM) añade texto que no forma parte del TO.

- **Omisión:** falta texto en el TM que aparece en el TO.
- **Error de traducción:** el TM no representa fielmente el contenido del TO.
- **Sobretraducción:** el TM es más específico que el TO.
- **Subtraducción:** el TM es menos específico que el TO.
- **Coincidencia de memoria de traducción inexacta:** Se proporciona una traducción como coincidencia exacta de un sistema de memoria de traducción (MT), pero en realidad es incorrecta.

En segundo lugar, tenemos la **fluidez**, la cual incluye errores relacionados con la forma o el contenido de un texto, independientemente de si es una traducción o no. Cuenta con las siguientes subcategorías:

- **Puntuación:** error de puntuación según la configuración del receptor del texto o según el estilo.
- **Ortografía:** errores relacionados con la ortografía de las palabras.
- **Gramática:** errores gramaticales o de la sintaxis del texto que no estén relacionadas con la ortografía ni con la ortotipografía.
- **Registro gramatical:** la TM tiene un registro gramatical incorrecto, como el uso de pronombres o formas verbales informales cuando se requieren sus homólogos formales.

La tercera categoría es la **terminología**, la cual abarca errores terminológicos, es decir, aquellos términos del campo del texto que se traduce incorrectamente o se traducen por otro término que no se espera dentro de ese contexto. La subcategoría de esta clasificación es:

**Inconsistencia en el uso de terminología:** La terminología se utiliza de forma incoherente/ inconsistente en el texto.

Por último, tenemos el **estilo**, los cuales son errores que estén relacionados con la falta de fluidez del texto por la forma de la redacción. Sus subcategorías son:

- **Falta de estilo:** al TM le falta estilo. A pesar de que se puede llegar a entender su contenido, está redactado de una manera que no denota fluidez.
- **Estilo inconsistente:** inconsistencia del estilo a lo largo del TM.
- **Idiomático:** el contenido es gramatical, pero no idiomático.

Una vez identificada la tipología de los errores, los clasificamos según su nivel de gravedad. TAUS DQF nos presenta 5 valores de gravedad que vamos a presentar de mayor a menor:

- **Muy grave:** errores que puedan tener implicaciones sanitarias, de seguridad, legales o financieras, que infrinjan las directrices geopolíticas de uso, que dañen la reputación de la empresa, que provoquen el bloqueo de la aplicación o que modifiquen o tergiversen negativamente la funcionalidad de un producto o servicio, o que puedan considerarse ofensivos.
- **Grave:** errores que pueden confundir o inducir al usuario a equivocarse o dificultar el uso adecuado del producto/servicio debido a un cambio significativo en el significado o porque los errores aparecen en una parte visible o importante del contenido.
- **Menor:** errores que no conducen a la pérdida de significado y que no confundirían ni engañarían al usuario, pero que son notorios, por lo que disminuye la calidad estilística, la fluidez o la claridad, o hacen que el contenido sea menos atractivo.
- **Neutro:** esta categoría se utiliza para registrar información adicional, problemas o cambios por hacer que no cuentan como errores, por ejemplo, reflejan la elección o el estilo preferido de un revisor, son errores repetidos o cambios de instrucciones/glosarios aún no aplicados, un cambio por hacer del que el traductor no es consciente.
- **Prestigio:** se utiliza para elogiar un logro excepcional.

Por defecto, estos errores reciben una puntuación individual que contribuye a la puntuación global del texto, determinando así si la traducción de los artículos es aceptable. Los errores muy graves tienen una ponderación de 10 puntos, los graves suman 5 puntos, los menores cuentan con 1 punto y los errores neutrales no afectan la puntuación. Esta escala se establece de esta manera porque cada texto tiene permitido un número máximo de errores puntuados para cumplir con el baremo del TAUS DQF. La métrica preestablecida por el sistema permite un máximo de 50 puntos de errores por cada 1000 palabras. En otras palabras, para que los motores de traducción aprueben su traducción, no pueden superar los 25 puntos cada traducción.

## 5. Marco teórico

A continuación, haremos un breve repaso por los diferentes sistemas de traducción automática de estadística y redes neuronales, presentaremos las herramientas seleccionadas basadas en estos sistemas, DeepL y Reverso. Además, contextualizaremos qué es ChatGPT y su funcionamiento como herramienta de traducción. Asimismo, discutiremos las características de los textos especializados, su lenguaje, y la aplicación de la traducción automática en los textos médicos.

### 5.1. Traducción automática

La traducción automática es la aplicación de tecnologías informáticas a textos de una lengua A y que se traduzcan a una lengua B sin necesidad de que intervenga un humano (Bermer, 2003, como se citó en Sánchez y Rico, 2020). Puede parecer un proceso sencillo, pero va más allá de sustituir palabras por otras, es un proceso complejo en el que el sistema aplica sobre el texto complejos conocimientos sobre lingüística como son la morfología y la sintaxis (Sánchez y Rico, 2020). Estos programas de traducción automática han podido salir adelante gracias a otros campos de estudios como la lexicografía, la lingüística computacional, las ciencias de la computación y la ingeniería lingüística (Quah, 2006, como se citó en Sánchez y Rico, 2020).

#### 5.1.1. Sistemas de traducción automática

A continuación, haremos un repaso de los sistemas de traducción automática basados en modelos estadísticos y redes neuronales, los cuales son utilizados por los motores de traducción seleccionados para este estudio, DeepL y Reverso. Comprender el funcionamiento de estos motores de traducción permite entender las discrepancias que puedan existir entre ChatGPT y los motores neuronales, así como evaluar su consistencia. Esto nos proporciona una base para contrastar estos conocimientos con el análisis de las traducciones.

### **5.1.1.1. *Sistemas de traducción automática basada en estadística***

Los sistemas de traducción automática basada en estadística nacen a la mitad de los años 80 y se sustentan en la probabilidad de que una oración en la lengua de origen le corresponda una traducción en la lengua meta (Sánchez y Rico, 2020).

Para que este sistema funcione, necesita un corpus monolingüe de la lengua meta, lo más grande posible, y otro paralelo con traducciones entre la lengua de origen y la lengua madre (Parra, 2022). Se parte de la idea de que no existe una única traducción posible y que una misma oración puede tener diferentes traducciones (Sánchez y Rico, 2020). Lo que caracteriza este tipo de sistemas es que asocian cada una de las posibles traducciones a una probabilidad según su verosimilitud (Forcada et al., 2016, como se citó en Sánchez y Rico, 2020).

Estos sistemas cuentan con tres componentes (Parra, 2022):

- 1. El modelo del lenguaje:** calcula la probabilidad en la que una frase en la lengua de destino sea correcta, se encarga de la fluidez de la traducción y se entrena con sus corpus monolingües en la lengua meta que deben ser lo más grande posible.
- 2. El modelo de traducción:** el que establece la correspondencia entre la lengua original y la lengua meta y se entrena utilizando un corpus alineado a nivel oracional. En la frase de entrenamiento, el sistema estima la probabilidad de una traducción a partir de las traducciones que aparecen en el corpus de entrenamiento.
- 3. El codificador:** es el que busca dentro de todas las traducciones posibles la más probable en cada caso. Una vez se dé un modelo de lenguaje y un modelo de traducción, es capaz de crear todas las traducciones posibles y propone la más probable.

La investigación en esta área siguió avanzando y logró alcanzar resultados aceptables e irrumpieron en la vida del traductor profesional hace unos años. Estas mejoras fueron en todos los componentes, desde la forma de alinear subordinadas, hasta la incorporación de la lingüística como parte del entrenamiento o más técnicas del procesamiento del lenguaje natural y aprendizaje automático (Parra, 2022).

### 5.1.1.2. *Sistemas de traducción automática basados en redes neuronales*

Este sistema son redes neuronales artificiales que representan de forma numérica las palabras y las frases a través de vectores y empleando técnicas de aprendizaje profundo (Sánchez y Rico, 2020). Estas redes neuronales intentan emular el funcionamiento de las neuronas del cerebro, basado en las interconexiones que se producen entre ellas y que aprenden a través de entrenamiento. Las redes neuronales están interconectadas, por lo que reciben estímulos que hacen que se activen a los que están conectadas (Sánchez y Rico, 2020).

Las redes neuronales artificiales están agrupadas en distintos niveles de capas llamados modelo de perceptrón multicapa (Sánchez y Rico, 2020):

La capa de entrada está formada por neuronas que reciben información, la cual se suministra y se envía a la siguiente capa. Esta siguiente capa es la que está oculta y formada por neuronas que reciben el estímulo de la capa de entrada y sobre las cuales aprende, por lo que en esta se produce el entrenamiento. La última capa se llama capa de salida, la cual recibe información de las neuronas de la capa oculta y dan un resultado final con el cual se valorará si el entrenamiento y el aprendizaje han sido adecuados (Sánchez y Rico, 2020). Por lo tanto, las neuronas de cada capa envían estímulos a las neuronas inmediatamente a la que le siguen y que están fuera de la capa (Sánchez y Rico, 2020).

La neurona recibe un estímulo de las neuronas que están en la capa anterior y esta misma genera una activación al crear una salida. Este proceso de activación depende de la activación de otras neuronas y del peso de las conexiones establecidas entre ellas (negativo o positivo). Estas aprenden según se van entrenando y asignando el peso adecuado de las conexiones a un resultado acorde a los resultados que se esperan (Sánchez y Rico, 2020).

La traducción neuronal, las redes neuronales, se entrenan con corpus paralelos de cientos de miles de oraciones.

Forcada (como se cita en Sánchez y Rico, 2020), dice que se requieren de tres corpus para trabajar en la TA neuronal:

1. **Corpus de entrenamiento:** son corpus de mayor envergadura entre 100 mil y 1 millón de oraciones alineadas. Su función es enseñar y entrenar a redes neuronales automáticas.



2. **Corpus de desarrollo:** estos corpus están formados entre mil y 3 mil frases alineadas que se utilizan para detener el entrenamiento en un momento concreto y evitar sobreentrenar el algoritmo con datos para los que ya se conoce el resultado.
3. **Corpus de test:** están formados entre mil y 3 mil frases alineadas que indican el rendimiento del sistema de traducción automática neuronal. Este corpus no se utiliza en el corpus de entrenamiento.

La traducción neuronal utiliza de algoritmo la retropropagación para establecer el peso de las conexiones entre las redes neuronales artificiales, en donde se propagan los datos de entrada con el fin de obtener una predicción, se calcula el error respecto a la salida.

### 5.1.1.3. *DeepL y Reverso*

DeepL es un sistema de traducción automática basada en redes neuronales artificiales, las cuales se entrenan con millones de textos traducidos y que además cuentan con un equipo de investigación propio que han introducido innovaciones en la metodología de las redes neuronales en cuatro ámbitos (DeepL, 2021).

- **Arquitectura neuronal:** muchos sistemas de traducción públicos utilizan modificaciones directas en la arquitectura *transformer* (DeepL, 2021), la cual consiste en redes neuronales que aprenden contexto y significado mediante las relaciones con datos secuenciales. Esta arquitectura aplica una técnica matemática de atención (NVIDIA, 2022). DeepL tiene arquitectura en modelos de atención, pero se diferencia en la topología, lo que logra una mejora en la calidad de traducción (DeepL, 2021).
- **Datos de entrenamiento:** se obtienen a través de la adquisición selectiva de datos de entrenamiento específicos de la calidad para lograr una mejor traducción. Para ello se desarrollaron rastreadores webs únicos que encuentran traducciones automáticamente en internet y evalúan la calidad.
- **Metodología de entrenamiento:** se entrena a las redes neuronales con el método de aprendizaje supervisado, el cual consiste en mostrar a las redes neuronales los mismos ejemplos una y otra vez para que comparen continuamente sus propias traducciones con los datos de entrenamiento. Si hay

alguna diferencia en los pesos de las neuronas, se ajustan de la manera correspondiente. Además, se han integrado técnicas de otros ámbitos del aprendizaje automático para entrenar redes neuronales cuyo resultado son innovaciones significativas.

- **Tamaño de las redes neuronales:** se entrenan con millones de parámetros. Son tan grandes que solo pueden entrenarse si están distribuidas en clústeres de ordenadores de gran tamaño destinados a ese fin. El equipo de investigación ha logrado que las redes neuronales pequeñas, sean capaces de producir traducciones de calidad casi igual que las grandes gracias a los parámetros que usan.

Reverso es un motor de traducción asistida basado en redes neuronales (Reverso, 2024b). La herramienta de traducción automática incluye en memorias de traducción de sistemas reconocidos como Across, Trados, Dejavú, así como en motores de traducción automática como Reverso, Babelfish y Systran (Reverso, 2024c).

Además de contar con un sistema de traducción automática similar a DeepL, Reverso ofrece diccionarios bilingües en varios idiomas con ejemplos de uso. Asimismo, dispone de una herramienta denominada Context, la cual permite encontrar la traducción de una palabra o frase en diferentes contextos (Reverso, 2024a).

Reverso también proporciona otras herramientas como un corrector ortográfico y gramatical, un diccionario de sinónimos y analogías, y un diccionario colaborativo donde los usuarios de la plataforma pueden proponer traducciones y votar a favor o en contra, entre los miembros que estén suscritos.

## 5.2. Inteligencia artificial

John McCarthy (2007) define la inteligencia artificial como una disciplina científica y técnica dedicada a la creación de sistemas y programas informáticos capaces de realizar tareas inteligentes. Estas tareas incluyen la capacidad para alcanzar objetivos en el mundo de manera similar a como lo hacen los seres humanos y otros organismos. Aunque la inteligencia artificial está relacionada con la comprensión de la inteligencia humana, no se limita a imitar únicamente procesos biológicos observables. La inteligencia artificial se basa en la capacidad de las máquinas para ejecutar procedimientos computacionales que pueden ser considerados inteligentes, aunque aún

no existe una definición sólida de inteligencia que no dependa de relacionarla con la inteligencia humana.

### 5.2.1. ChatGPT

ChatGPT es un sistema de chat basado en inteligencia artificial (IA), creado por OpenAI, una empresa dedicada a la investigación y desarrollo de IA con la misión de beneficiar a toda la humanidad (OpenAI, 2024). Este sistema permite mantener conversaciones en las que se pueden hacer preguntas de manera convencional, proporcionando respuestas naturales y fluidas gracias al modelo de lenguaje GPT-3.5 en su versión gratuita y GPT-4 en su versión de pago (Fernández, 2024). Estos modelos cuentan con más de 176 millones de parámetros y han sido entrenados con grandes cantidades de texto para realizar diversas tareas relacionadas con el lenguaje, desde la generación de texto hasta la traducción (Fernández, 2024). El aprendizaje de ChatGPT se logra mediante la técnica de refuerzo a partir de retroalimentación humana (por sus siglas en inglés RLHF), que implica la simulación de conversaciones humanas y la adaptación de sus respuestas en función de la precisión reflejada en el diálogo humano natural (Arimetrics, 2024).

Algunas de las formas en las que se puede utilizar ChatGPT son las siguientes:

- **Generador de texto:** esta herramienta puede producir textos coherentes y bien escritos en una amplia gama de estilos, como música, poemas, artículos, historias y respuestas a preguntas. Además, puede generar contenido en diversos temas e idiomas, incluyendo resúmenes de noticias, descripciones de productos y narrativas. En el ámbito educativo, la producción de textos como ensayos, monografías, trabajos finales, tesinas y tesis son elementos relevantes y una herramienta básica en el proceso de evaluación (Diego Olite et al., 2023).
- **Aumento de la precisión en los sistemas de búsqueda:** el modelo puede incrementar la precisión de los sistemas de búsqueda de información, ya que es capaz de entender el contexto y la intención de las consultas de los usuarios (Ortiz, 2023).
- **Resolución de problemas:** el modelo es posible analizar situaciones y ofrecer soluciones o respuestas a las preguntas formuladas (Diego Olite et al., 2023). Por ejemplo, detectar problemas de código HTML.

- **Generar contenido para diferentes ámbitos:** por ejemplo, se puede utilizar en el ámbito de la atención al cliente para crear una línea de chatbot, crear descripciones de productos en el ámbito del comercio electrónico o generar contenido para redes sociales (Ortiz, 2023).
- **Traducir textos a diferentes idiomas,** que será el uso que le daremos a esta herramienta en el presente trabajo y poner a prueba su eficiencia.

### 5.2.2. *ChatGPT como herramienta de traducción*

La traducción automática ha experimentado avances significativos en los últimos años gracias al desarrollo de modelos del lenguaje avanzados basados en inteligencia artificial (Dongxing y Zuying, 2022). ChatGPT ha demostrado ser capaz de traducir textos en varios idiomas, sin embargo, para comprender cómo esta herramienta lleva a cabo esta tarea, es crucial examinar el proceso de creación de *prompts* y estrategias utilizadas para influir en la calidad y precisión de estas traducciones. Los *prompts* son instrucciones o pautas que se le indican a ChatGPT para orientar la generación de texto en una dirección específica (Barquilla, 2024).

Según exponen en Generación Aprende (2023), la manera más efectiva de que ChatGPT genere textos de calidad es asignándole un rol o una orden específica, además de la actividad que va a llevar a cabo. Por ejemplo, un comando sencillo puede ser «traduce esta frase al inglés», y seguidamente se pega la frase en cuestión en el cuadro de diálogo (Generación Aprende, 2023). Ese comando es un comando sencillo, pero se puede perfilar más para que sea más específico, como puede ser: «traduce este email como si fueras angloparlante estadounidense. Utiliza palabras propias del inglés de negocios, reemplazando frases o palabras del texto original por sinónimos o frases que signifiquen lo mismo» (Generación Aprende, 2023). Como podemos observar, aquí ya se añade el tipo de texto que se va a traducir (email), la variante del lenguaje (inglés de Estados Unidos) y el campo de especialidad del texto (negocios).

De una manera similar, Sui He (2024) presenta en su estudio sobre la creación de *prompts* para la traducción varias propuestas metodológicas. Inicialmente, presenta cinco propuestas diferentes, entre ellos un *prompt* sencillo en el que da el comando de traducción a ChatGPT indicándole los pares de lenguas y el uso profesional de la traducción. Posteriormente, otro al que añade elementos específicos que considerarían traductores profesionales, tales como la función del texto, el destinatario, el momento y

lugar de recepción y el medio en el cual se transmite el texto. Finalmente, incorpora al *prompt* sencillo, roles específicos como el de traductor y el de autor. Al evaluar los diferentes comandos, concluye que los *prompts* más efectivos son aquellos que asignan roles claros a ChatGPT, lo que resulta en una traducción y redacción del texto más eficiente y el *prompt* sencillo, al ser más directo, destaca sobre el segundo *prompt*.

### 5.3. Textos especializados

La traducción de textos especializados representa un desafío único debido a la complejidad y especificidad de los contenidos. Estos textos se distinguen por la precisión terminológica y el rigor conceptual necesarios para garantizar la correcta transmisión de información entre especialistas y otros destinatarios, ya sean aprendices o el público general. En este apartado, se exploran las características distintivas de los textos especializados, centrándonos en los elementos lingüísticos y estructurales que los conforman.

El texto especializado se distingue de otros textos por los elementos que intervienen en su proceso de producción y recepción del mismo (Domènech, 2006). Para que un texto sea considerado especializado, su emisor debe ser especialista en su campo, mientras que su destinatario puede ser otro especialista del mismo campo o materias afines, aprendices o bien el público general (Domènech, 2006).

En este tipo de textos, hay un conjunto de recursos lingüísticos específicos del ámbito comunicativo, entre ellos están la temática, el interlocutor, la situación comunicativa o la intención, los cuales se utilizan para garantizar la comprensión entre las personas del ámbito de especialidad (Guantiva, et al., 2008).

La estructura de los textos especializados se ajusta a la intención comunicativa, ya que es necesario que crear la estrategia para comunicarse con el receptor, debido a que este tiene una actitud expectante respecto al texto que va a leer, y el léxico el cual, según su uso específico, podrá ser más o menos opaco dependiendo de la especialización y de la variación expresiva (Guantiva, et al., 2008).

Es fundamental tener claros el registro, el género y la tipología textual a la hora de traducir y asegurar que se respetan, ya que forman parte de la adecuación del texto. El registro es el que responde a las formas lingüísticas que condicionan las funciones comunicativas, es decir, las características que logran que un texto sea más o menos

especializado y que orientan en la producción de un texto y en la expectativa del lector. Estos elementos son el campo (especialidad), el tenor (relación entre interlocutores) y el modo (canal) (Guantiva, et al., 2008).

El género es el tipo de texto, normalmente determinado por un factor histórico cultural, cuyas características pueden cambiar según la lengua meta, y la tipología textual que está ligada a la teoría para la clasificación científica de los textos (Guantiva et al., 2008). En nuestro caso, estaríamos hablando de un artículo científico, cuyas características trataremos más a continuación.

### 5.3.1. Características del lenguaje médico publicado

En lo que atañe a este trabajo, lo que más nos interesa, debido a que es lo que vamos a analizar, son las características del lenguaje médico a nivel académico.

El lenguaje de especialidad tiene la función primordial de transmitir conocimientos lingüísticos que se presentan en ámbitos de especialización concretos (Streiger, 1993, como se citó en: Ariza, 2009). El lenguaje de la medicina, como en otras materias, ha creado su propia forma de comunicarse y se puede clasificar según dos criterios:

- Eje horizontal que tiene que ver con la especialidad del texto.
- Eje vertical que tiene que ver con la especialización.

Como mencionamos anteriormente, los textos especializados pueden crearse para comunicarse entre otros expertos, por lo que la comunicación entre médicos se caracteriza por su precisión y economía en circunstancias específicas (Schifko, 2001, como se citó en Ariza, 2009).

Dentro de este lenguaje escrito de la medicina, una de las características que tiene son las abreviaturas, las cuales usan en su día a día (Ariza, 2009). Estas abreviaturas pueden ser, por ejemplo, ACVA < Accidente cardiovascular agudo o TAC < Tomografía Axial Computarizada. Estas siglas y acrónimos se utilizan como recurso lingüístico para simplificar y agilizar los textos escritos (Ariza, 2009). Pero, a diferencia de lo que es el lenguaje oral, el escrito no presenta ni truncamientos ni elipsis, al contrario, tenemos un lenguaje pomposo y rebuscado que, en ocasiones, suele ser excesivo de palabras para expresar una sola idea (Ariza, 2009).

En ocasiones, también es frecuente que, en el lenguaje médico publicado, se utilicen metáforas. Sobre todo, en especialidades como la neumología que utiliza símiles bélicos.

### 5.3.2. Traducción automática en textos especializados

La traducción automática ha avanzado notablemente en los últimos años gracias a las innovaciones en inteligencia artificial y al aprendizaje profundo. Sin embargo, la traducción en textos especializados, como los médicos, presenta desafíos que requieren enfoques y técnicas que conoce un traductor para asegurar su precisión y su coherencia. En este apartado analizaremos los avances y los desafíos de la traducción automática aplicada en textos especializados.

Los avances en la traducción automática se deben al desarrollo de modelos neuronales y modelos *Transformers*. Estos modelos han logrado mejorar significativamente la capacidad de los motores de traducción para comprender y traducir textos complejos especializados, así como el manejo del contexto y de la terminología específica, mejorando la calidad de las traducciones de campos especializados (Soto et al., 2019) (Baniata et al., 2021).

Un estudio publicado en la revista *Neural machine translation of clinical texts between long distance languages* (Soto et al., 2019), destaca cómo las redes neuronales recurrentes pueden optimizarse para traducir textos clínicos con alta precisión. Esta investigación subraya la importancia de ajustar los parámetros y las configuraciones para mejorar el rendimiento de la traducción automática de los textos médicos (Soto et al., 2019). Estos resultados los consigue al crear su propio motor de traducción automática.

A pesar de los avances, la traducción automática aplicada en textos especializados sigue presentando varios desafíos, como la traducción correcta de la terminología específica del campo del que trata el texto y las jergas propias de dicho campo. Los textos médicos, como hemos visto anteriormente, contienen términos técnicos y abreviaciones que requieren de una comprensión profunda del contexto para que su traducción sea adecuada.

Según el estudio realizado por Baniata (2021) sobre la traducción de dialectos árabes mediante modelos *Transformers*, los cuales emplean subunidades léxicas pueden

ofrecer una solución al problema que hay con la terminología que desconoce el motor de traducción y poder mejorar la calidad de la traducción. Este enfoque permite a los modelos transformar y ajustar mejor los términos técnicos y específicos dentro de un contexto adecuado, lo que resulta en traducciones más precisas y útiles para el ámbito médico.

A pesar de las capacidades avanzadas de los modelos de traducción automática, la intervención humana sigue siendo crucial para asegurar la precisión y la adecuación de las traducciones en contextos especializados. Un motor de traducción automática puede ofrecer una primera versión que luego debe ser revisada y corregida por un experto para poder garantizar la coherencia y la precisión terminológica que requiere el texto según su especialidad. Como afirma Soto (2019) en su estudio, la traducción automática puede ser útil y ha mejorado con el paso del tiempo, pero es necesaria la intervención humana para poseer los resultados obtenidos.

## 6. Estudio de caso

Esta sección del caso de estudio se divide en dos partes. La primera parte es el análisis de las traducciones de los motores de traducción automática y las traducciones de ChatGPT en las cuales se emplea el *prompt* sencillo. La segunda parte del estudio se basa en comparar las traducciones de ChatGPT con el *prompt* sencillo y el *prompt* completo con el fin de analizar si este cambio en el *prompt* consigue presentar mejoras en la traducción, superando la puntuación de los motores de traducción automática o si mejora su propia traducción respecto a la propuesta con el *prompt* sencillo.

### 6.1. Resultados

En las siguientes secciones, presentaremos el análisis detallado de las traducciones que hemos analizado con TAUS DQF. En primer lugar, presentaremos la comparación general con las cifras obtenidas por la métrica para cada uno de los motores de traducción, DeepL y Reverso y las traducciones que realiza ChatGPT con el *prompt* sencillo, ya que presentan variaciones. Luego, haremos un análisis más detallado, utilizando ejemplos de



las traducciones, centrándonos en las de las cuatro categorías, fidelidad, fluidez, terminología y estilo, con el fin de respaldar los resultados obtenidos.

El análisis se presentará en forma de tabla en la cual se incluye el texto original (TO), la traducción humana (TH) y las traducciones de los diferentes motores de traducción (DeepL, Reverso y ChatGPT). En el caso de ChatGPT siempre se indicará en la tabla si la traducción es del *prompt* simple o completo.

Los errores aparecerán en negrita y en ocasiones podemos encontrar omisiones del texto en los fragmentos, las cuales se indicarán de la siguiente manera: [...].

### 6.1.1. Análisis global

Texto	Fidelidad	Fluidez	Terminología	Estilo	Total
<b>DeepL_Ictus</b>	3	0	6	4	18
<b>DeepL_Pulmón</b>	3	3	3	3	27
<b>Rev_Ictus</b>	6	0	9	4	43
<b>Rev_Pulmón</b>	3	1	6	2	31
<b>GPT_Ictus_S</b>	5	0	6	3	30
<b>GPT_Pulmón_S</b>	3	1	6	2	23

Tabla 1. Resultados traducción motores de redes neuronales y prompt simple de ChatGPT

En esta tabla se presentan los resultados obtenidos según la puntuación de cada error dependiendo de su gravedad. En un análisis inicial, observamos que el motor de traducción que mejores resultados en términos de puntuación es DeepL, a pesar de que no consigue aprobar la traducción del texto sobre enfermedades pulmonares por un margen muy estrecho, quedando ligeramente por debajo del límite permitido de 25 puntos. En segunda posición se encuentra ChatGPT, que también no logra aprobar uno de los textos, en este caso el que trata sobre el ictus. En última posición, está Reverso, el cual no consigue aprobar ninguna de las traducciones, presentando en uno de los casos un número de errores que prácticamente duplica el máximo permitido.

La categoría que presenta una mayor cantidad de errores en ambos textos, y en las traducciones de todos los motores evaluados, es la de terminología. Esto significa que los motores de traducción no son eficientes a la hora de traducir terminología médica

especializada, lo cual es un problema, ya que, como se discutió en la sección dedicada a los textos especializados, es precisamente la terminología específica del campo la que distingue a un texto especializado de uno no especializado. Por consiguiente, es fundamental que dicha terminología se respete rigurosamente en la traducción.

Al analizar los datos de DeepL, resulta interesante observar que parece haber más errores en el texto aprobado que en el texto no aprobado. Esto se explica por la gravedad de los errores: mientras que en el texto aprobado predominan errores de categoría leve, en el texto no aprobado son más frecuentes los errores de categoría grave, por lo que se suman más puntos a pesar de ser menos errores.

## 6.1.2. Análisis por categoría

En esta sección vamos a analizar los errores más destacados, en cada una de las categorías, que han tenido los motores de traducción evaluados para poder comprender mejor cuáles han sido sus errores y sus aciertos.

### 6.1.2.1. *Fidelidad*

En la categoría de fidelidad analizamos aquellos errores que carecen de adecuación respecto al texto de origen. Es de las categorías que más errores presenta y contabilizando los errores de ambos textos, podemos encontrar errores de todas las subcategorías: incorporación, omisión, error de traducción, sobretraducción y subtraducción.

Un error común que se repite en todas las traducciones propuestas es la traducción de la palabra *procedure*, la cual se traduce por «procedimiento». Esta propuesta de traducción no es incorrecta, pero tal como nos indica el Libro Rojo de Fernando A. Navarro (2023g), debemos tener en cuenta que según el contexto del texto la traducción puede ser diferente y referir a: intervención quirúrgica, técnica quirúrgica, operación, maniobra, técnica, método, conducta, protocolo, reglamento, norma, modo o manera (Navarro, 2023g). Justamente es el caso de este fragmento, en la misma entrada del diccionario da como ejemplo *ablative procedures* que se debe traducir por «técnica ablativa»:

<b>TO</b>	<i>Ablation procedures do not obviate the need of anticoagulation in high-risk patients [...].</i>
<b>TH</b>	Las <b>técnicas ablativas</b> no evitan la necesidad de anticoagulantes en los pacientes de alto riesgo [...].
<b>DeepL</b>	Los <b>procedimientos de ablación</b> no obvian la necesidad de anticoagulación en los pacientes de alto riesgo [...].
<b>Reverso</b>	Los <b>procedimientos de ablación</b> no obvian la necesidad de anticoagulación en pacientes de alto riesgo [...].
<b>ChatGPT simple</b>	Los <b>procedimientos de ablación</b> no eliminan la necesidad de anticoagulación en pacientes de alto riesgo [...].

Tabla 2. fragmento 9, *The role of cardiologists in stroke prevention and treatment*

Lo mismo ocurre con la traducción de la palabra *bleeding* que en español puede ser tanto «sangrado» como «hemorragia». Al estar hablando del ictus debemos traducir por «hemorragia» y no por «sangrado», ya que en español este último se utiliza para hacer «referencia a la salida normal de sangre en procesos fisiológicos como la menstruación o el puerperio» (Navarro, 2023a, definición 1).

Este error lo comete tanto Reverso y como ChatGPT:

<b>TO</b>	<i>It might be a future alternative for patients at high risk of stroke with simultaneous high risk for <b>bleeding</b>.</i>
<b>TH</b>	Puede ser una futura alternativa para pacientes con alto riesgo de sufrir un ictus con riesgo simultáneo de <b>hemorragia</b> .
<b>DeepL</b>	Podría ser una alternativa futura para pacientes con alto riesgo de ictus y alto riesgo simultáneo de <b>hemorragia</b> .
<b>Reverso</b>	Podría ser una alternativa futura para pacientes con alto riesgo de accidente cerebrovascular con alto riesgo simultáneo de <b>sangrado</b> .
<b>ChatGPT simple</b>	Podría ser una alternativa futura para pacientes con alto riesgo de accidente cerebrovascular con riesgo simultáneo de <b>sangrado</b> .

Tabla 3. Fragmento 9: *The role of cardiologists in stroke prevention and treatment*

Durante el análisis de las traducciones, se ha detectado que en algunos motores de traducción se han producido incorporaciones al texto meta que no aparecen en el texto original. Este problema se detectó en dos ocasiones en la traducción de DeepL:

<b>TO</b>	<i>A radiograph of the chest showed opacities in both lungs that were suggestive of infection or acute lung injury (Fig. 2A).</i>
<b>TH</b>	Una radiografía de tórax mostró opacidades en ambos pulmones que indicativo de infección o lesión pulmonar aguda (Fig. 2A).
<b>DeepL</b>	<b>La radiografía de tórax mostró oclusión torácica.</b> Una radiografía de tórax mostró opacidades en ambos pulmones sugestivas de infección o lesión pulmonar aguda (Fig. 2A).
<b>Reverso</b>	Una radiografía de tórax mostró opacidades en ambos pulmones que sugerían infección o lesión pulmonar aguda (fig. 2A).
<b>ChatGPT simple</b>	Una radiografía de tórax mostró opacidades en ambos pulmones sugestivas de infección o lesión pulmonar aguda (Fig. 2A).

Tabla 4. Fragmento 6: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

<b>TO</b>	<i>Cytologic testing of bronchoalveolar-lavage specimens showed a neutrophil predominance (78%) with no eosinophils (0%) and a moderate number of lipid-laden macrophages on oil red O staining.</i>
<b>TH</b>	Las pruebas citológicas de las muestras del lavado broncoalveolar mostraron un predominio de neutrófilo (78%) sin presencia de eosinófilo (0%) y un número moderado de macrófagos cargados de lípidos en la tinción con aceite rojo O.
<b>DeepL</b>	Las pruebas citológicas de las muestras de lavado broncoalveolar mostraron un predominio de neutrófilos (78%) con ausencia de eosinófilos (0%) y un número moderado de macrófagos cargados de lípidos en la tinción con aceite rojo O. <b>Los estudios infecciosos, incluida la analítica sanguínea, mostraron que la neumonitis era una enfermedad crónica.</b>
<b>Reverso</b>	Las pruebas citológicas de muestras de lavado broncoalveolar mostraron predominio de neutrófilos (78%) sin eosinófilos (0%) y un número moderado de macrófagos cargados de lípidos en la tinción de aceite de O rojo.
<b>ChatGPT simple</b>	La prueba citológica de muestras de lavado broncoalveolar mostró un predominio de neutrófilos (78%) sin eosinófilos (0%) y un número moderado de macrófagos con contenido lipídico en la tinción con rojo aceite O.

Tabla 5. Fragmento 13: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

Como podemos observar, la información que se añade a ambos fragmentos no está relacionada con la información que ofrece el texto original, al contrario, añade

información nueva que puede confundir al lector al presentar un diagnóstico incorrecto, por lo que falta a la veracidad del texto.

Por el lado contrario, Reverso omite un fragmento del texto original, afectando de igual manera en la veracidad del texto:

<b>TO</b>	<i>Carotid stenting and carotid endarterectomy and similarly effective for patients.</i>
<b>TH</b>	Tanto la colocación de endoprótesis en la arteria carótida, como la endarterectomía carotídea son efectivos para los pacientes.
<b>DeepL</b>	La colocación de stents carotídeos y la endarterectomía carotídea son igualmente eficaces para los pacientes.
<b>Reverso</b>	
<b>ChatGPT simple</b>	La colocación de stents carotídeos y la endarterectomía carotídea son igualmente efectivas para los pacientes.

Tabla 6. Fragmento 5: *The role of cardiologists in stroke prevention and treatment*

Otro error que detectamos fue la subtraducción de algunos términos, es decir, no especificar un término que sí es específico en el texto original. En esta ocasión vemos este error en la traducción del término *hospital ED*, el cual hace referencia al servicio de urgencias del hospital. Reverso, en vez de especificar que el paciente acudió al servicio de urgencias de un hospital, simplemente especifica que acudió al hospital:

<b>TO</b>	<i>A 17-year-old male patient with no clinically significant medical history presented to a hospital ED [...].</i>
<b>TH</b>	Un paciente de 17 años, sin antecedentes de interés, acude al servicio de urgencias [...].
<b>DeepL</b>	Un paciente varón de 17 años sin antecedentes médicos clínicamente significativos se presentó en el <b>SU de un hospital</b> [...].
<b>Reverso</b>	Un paciente masculino de 17 años sin historia clínica significativa presentado a un <b>hospital</b> [...].
<b>ChatGPT simple</b>	Un paciente varón de 17 años sin antecedentes médicos clínicamente significativos se presentó en un servicio de emergencia hospitalario [...].

Tabla 7. Fragmento 3: *Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report*

Este tipo de error no es grave, pero sí es una falta a la fidelidad del texto original, ya que al no especificar que el paciente va a urgencias, hay una falta de contexto y puede

entenderse que quizás ya tenía una cita programada para consultar la causa de sus síntomas y recibir un tratamiento.

Otro problema que se ve reflejado en Reverso es que hay términos que no es capaz de traducir, este ejemplo lo vemos en el título del informe con el término *e-Cigarette*:

<b>TO</b>	<i>Pulmonary Illness Related to <b>E-Cigarette</b> Use in Illinois and Wisconsin — Final Report</i>
<b>TH</b>	Enfermedades pulmonares relacionadas con el uso del <b>cigarrillo electrónico</b> en Illinois y Wisconsin — informe final
<b>DeepL</b>	Enfermedades pulmonares relacionadas con el uso de <b>cigarrillos electrónicos</b> en Illinois y Wisconsin - Informe final
<b>Reverso</b>	Enfermedad pulmonar relacionada con el uso de la <b>e-Cigarette</b> en Illinois y Wisconsin - Informe final
<b>ChatGPT simple</b>	Enfermedad Pulmonar Relacionada con el Uso de <b>Cigarrillos Electrónicos</b> en Illinois y Wisconsin — Informe Final

Tabla 8. Fragmento 1: *Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report*

Al presentarse este tipo de errores con términos actuales que todavía se están introduciendo en nuestro vocabulario, podemos sospechar que la base de datos del motor de traducción no está actualizada, por lo que no puede proporcionar una traducción.

### 6.1.2.2. *Fluidez*

Los problemas de fluidez que hemos encontrado en las traducciones pertenecen principalmente al artículo científico sobre enfermedades pulmonares, lo que lo convierte en el texto con más problemas de este tipo.

En las tres traducciones proporcionadas por los motores de traducción se detecta una inconsistencia en la relación sustantivo-adjetivo del término *pediatric intensive care unit*. Esta inconsistencia se debe a que los motores de traducción no identifican adecuadamente cuál es la referencia del adjetivo *pediatric*, lo que resulta en una falta de concordancia en el género. Tanto en inglés como en español, *pediatric* hace referencia a «unidad», diferenciando así entre la unidad de cuidados intensivos para adultos y la unidad de cuidados intensivos pediátricos.

<b>TO</b>	<i>Within hours, he was transferred to the <b>pediatric intensive care unit</b> owing to respiratory deterioration.</i>
<b>TH</b>	En cuestión de horas, se le trasladó a la <b>unidad de cuidados intensivos pediátrica</b> a causa de un deterioro respiratorio.
<b>DeepL</b>	A las pocas horas, fue trasladado a la <b>unidad de cuidados intensivos pediátricos</b> debido al deterioro respiratorio.
<b>Reverso</b>	En pocas horas fue trasladado a la <b>unidad de cuidados intensivos pediátricos</b> debido al deterioro respiratorio.
<b>ChatGPT simple</b>	En cuestión de horas, fue trasladado a la <b>unidad de cuidados intensivos pediátricos</b> debido al deterioro respiratorio.

Tabla 9. Fragmento 9: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

En la traducción sobre enfermedades pulmonares de DeepL también detectamos dos errores gramaticales de género. El primer error es traducir *the patient* por «la paciente» (tabla 10), a pesar de que al inicio del texto se especifica que el paciente es un varón: *A 17-year old male patient*.

El segundo error se produce entre la concordancia entre sujeto y verbo, ambos están separados por un listado de pruebas, por lo que los motores de traducción no detectan la concordancia y traducen las palabras de manera independiente, haciendo una traducción literal:

<b>TO</b>	<i><b>Infectious workup</b>, including blood cultures, testing for the human immunodeficiency virus (HIV), testing for urinary histoplasma and blastomyces antigens, polymerase-chain-reaction (PCR) panel for nasopharyngeal respiratory virus, PCR panel for enteric pathogens, and bronchoalveolar-lavage studies including bacterial and fungal cultures and pneumocystis stain, <b>was negative</b>.</i>
<b>TH</b>	<b>Los análisis infecciosos</b> , incluidos los hemocultivos, las pruebas para la inmunodeficiencia humana (VIH), pruebas para antígenos de histoplasma y blastomicosis en la orina, panel de reacción en cadena de la polimerasa (PCR) para virus respiratorios nasofaríngeos, panel PCR para agentes entéricos y estudios de lavado broncoalveolar que incluyen los cultivos bacterianos y fúngicos y la tención para neumocistosis, <b>dieron resultado negativo</b> .

<b>DeepL</b>	<b>Las pruebas infecciosas</b> , incluidos los hemocultivos, las pruebas del virus de la inmunodeficiencia humana (VIH), las pruebas de antígenos urinarios de histoplasma y blastomyces, el panel de reacción en cadena de la polimerasa (PCR) para el virus respiratorio nasofaríngeo, el panel de PCR para patógenos entéricos y los estudios de lavado broncoalveolar, incluidos los cultivos bacterianos y fúngicos y la tinción para pneumocystis, <b> fueron negativos</b> .
<b>Reverso</b>	<b>Análisis infeccioso</b> , incluyendo hemocultivos, pruebas para el virus de inmunodeficiencia humana (VIH), pruebas para antígenos de histoplasma urinario y blastomycosis, panel de reacción en cadena de polimerasa (PCR) para el virus respiratorio nasofaríngeo, panel de PCR para patógenos entéricos, y los estudios de lavado broncoalveolar incluyendo cultivos bacterianos y fúngicos y tinción de pneumocystis, <b> fue negativo</b> .
<b>ChatGPT simple</b>	<b>El estudio de infecciones</b> , incluidos cultivos sanguíneos, pruebas para el virus de la inmunodeficiencia humana (VIH), pruebas para antígenos de histoplasma y blastomycosis en orina, panel de <u>PCR</u> para virus respiratorios nasofaríngeos, panel de PCR para patógenos entéricos, y estudios de lavado broncoalveolar incluyendo cultivos bacterianos y fúngicos y tinción para pneumocystis, <b> fue negativo</b> .

Tabla 101. Fragmento 14: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

Como punto a favor respecto a los otros dos motores, DeepL fue capaz de adaptar *infections workup* al plural al reconocer el listado de pruebas, dándole fluidez al texto en español, concordando con la propuesta en la traducción humana. En este fragmento también hay un error gramatical por parte de Reverso al omitir el artículo al traducir «análisis infeccioso».

Estos problemas no son graves y con el tiempo y un buen entrenamiento de los motores de traducción pueden corregirse.

### 6.1.2.3. Terminología

Se consideran errores terminológicos aquellos términos relacionados con el campo especialización del texto y su contexto que estén mal traducidos. En esta categoría hay dos clasificaciones, la primera es marcar el error simplemente como error terminológico y la segunda es clasificarlo como error de inconsistencia a lo largo del texto. El motor de



traducción que tiene menos errores de este tipo es DeepL, con un total de 9 errores entre los dos textos frente a Reverso y ChatGPT que tienen 15 y 11 respectivamente. Al haber tantos errores sobre terminología, solo mencionaremos los casos más destacables.

Uno de los errores más graves que detectamos se comete en los tres motores de traducción y es traducir *co-morbidities* por «comorbilidades». Al hacer una búsqueda en el diccionario Libro rojo, nos indica que *cormobidity*, en singular, hace referencia «al número de personas afectadas por una enfermedad determinada, en una población durante un periodo de tiempo» (Navarro, 2023b, definición 1). En esa misma entrada, en la acepción 2, se especifica que el uso en plural de este término se refiere a aquellas enfermedades que aparecen con frecuencia y a la misma vez que otras, por lo que cuando aparece este término en plural en inglés, debe traducirse por «enfermedad concomitante» (Navarro, 2023b, definición 2).

<b>TO</b>	<i>Additionally, many patients with AF are elderly with other <b>co-morbidities</b> and there is limited information on the risk/benefit of these procedures in this population.</i>
<b>TH</b>	Además, muchos pacientes con FA son ancianos y presentan <b>enfermedades concomitantes</b> y la información sobre el riesgo/beneficio de estas técnicas en esta población es limitada.
<b>DeepL</b>	Además, muchos pacientes con FA son ancianos con otras <b>comorbilidades</b> y existe información limitada sobre el riesgo/beneficio de estos procedimientos en esta población.
<b>Reverso</b>	Además, muchos pacientes con FA son ancianos con otras <b>comorbilidades</b> y hay información limitada sobre el riesgo/beneficio de estos procedimientos en esta población.
<b>ChatGPT simple</b>	Además, muchos pacientes con fibrilación auricular son ancianos con otras <b>comorbilidades</b> y hay información limitada sobre el riesgo/beneficio de estos procedimientos en esta población.

Tabla 112. Fragmento 10: *The role of cardiologists in stroke prevention and treatment*

Otro error grave que, a simple vista, puede pasar desapercibido o como un error menor al pensar que no afecta en el significado del texto, es la traducción errónea del término *C-reactive protein level*. Al consultar el diccionario de Fernando A. Navarro (2023d), lo primero que desaconseja es traducir el término por «proteína C reactiva», ya que da pie a interpretar que estamos hablando de un tipo de proteína C cuando en realidad

estamos hablando de «una proteína anómala que reacciona por precipitación con el polisacárido C», por lo tanto, hay que emplear el uso de guion para hacer referencia al polisacárido C (Navarro, 2023d): proteína C-reactiva.

<b>TO</b>	<i>His laboratory results showed leukocytosis (white-cell count, 18,000 cells per cubic millimeter) with a neutrophil predominance (94%) and no eosinophils (0%), as well as elevated inflammatory markers with a <b>C-reactive protein</b> level of 32 mg per deciliter and an erythrocyte sedimentation rate of 68 mm per hour.</i>
<b>TH</b>	Los resultados de laboratorio mostraron leucocitosis (recuento de leucocitos 18.000 células por milímetro cúbico) con predominio de neutrófilos (94%) y ausencia de eosinófilo (0%), así como marcadores inflamatorios elevados con un nivel de <b>proteína C-reactiva</b> de 32 mg por decilitro y una velocidad de sedimentación globular de 68 mm por hora.
<b>DeepL</b>	Los resultados de laboratorio mostraron leucocitosis (recuento de leucocitos, 18.000 células por milímetro cúbico) con predominio de neutrófilos (94%) y ausencia de eosinófilos (0%), así como marcadores inflamatorios elevados con un nivel de <b>proteína C reactiva</b> de 32 mg por decilitro y una velocidad de eritrosedimentación de 68 mm por hora.
<b>Reverso</b>	Sus resultados de laboratorio mostraron leucocitosis (recuento de glóbulos blancos, 18.000 células por milímetro cúbico) con predominio de neutrófilos (94%) y ausencia de eosinófilos (0%), así como elevados marcadores inflamatorios con un <b>Cnivel de proteína reactiva</b> de 32 mg por decilitro y una tasa de sedimentación de eritrocitos de 68 mm por hora.
<b>Chat GPT simple</b>	Sus resultados de laboratorio mostraron leucocitosis (conteo de células blancas, 18,000 células por milímetro cúbico) con predominio de neutrófilos (94%) y sin eosinófilos (0%), así como marcadores inflamatorios elevados con un nivel de <b>proteína C reactiva</b> de 32 mg por decilitro y una velocidad de sedimentación de eritrocitos de 68 mm por hora.

Tabla 123. Fragmento 7: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

Además, como podemos observar en la traducción de Reverso, se invirtió el orden del término, por lo que ni siquiera podría valorarse como una posible opción correcta. Con estos dos ejemplos evidenciamos la importancia de que las traducciones de textos especializados exigen un gran conocimiento en la materia, por lo que deben realizarla un traductor especializado.

Otro error que compromete la comprensión de la información del texto es la traducción literal de *standard care* por «estándar de atención» por parte de Reverso y ChatGPT:

<b>TO</b>	<i>Other ongoing studies in AF patients with contraindications for long term anticoagulation compare LAA occlusion with <b>standard of care</b>.</i>
<b>TH</b>	Otras investigaciones en curso comparan el cierre de la OAI con el <b>tratamiento estándar</b> en pacientes con FA y con contraindicaciones para la anticoagulación a largo plazo.
<b>DeepL</b>	Otros estudios en curso en pacientes con FA con contraindicaciones para la anticoagulación a largo plazo comparan la oclusión de la OI con el <b>tratamiento estándar</b> .
<b>Reverso</b>	Otros estudios en curso en pacientes con FA con contraindicaciones para la anticoagulación a largo plazo comparan la oclusión de LAA con el <b>estándar de atención</b> .
<b>ChatGPT simple</b>	Otros estudios en curso en pacientes con fibrilación auricular con contraindicaciones para la anticoagulación a largo plazo comparan el cierre de la LAA con el <b>estándar de atención</b> .

Tabla 134. Fragmento 23: *The role of cardiologists in stroke prevention and treatment*

Otros errores terminológicos comunes que hemos detectado es la falta de consistencia en la elección de los términos y mantenerla a lo largo del texto. Por ejemplo, en la traducción de Reverso vemos como traduce *strokes* de dos maneras diferentes, en el título del artículo lo traduce por «ictus» (tabla 15), pero en el cuerpo del artículo lo traduce por «ACV», las siglas de accidente cardiovascular, en los casos como *primary stroke prevention* (tabla 16) y *acute phase of stroke* (tabla 16).

El mismo tipo de error ocurre con DeepL y su traducción para el artículo sobre el ictus. En el texto original aparece la sigla LAA (*left atrial appendage occlusión*), y lo traduce por su equivalente en español OAI (orejuela auricular izquierda) (tabla 17), pero más adelante en el artículo, cuando aparece solo la sigla, es decir sin especificar su significado entre paréntesis, traduce LAA por OI (orejuela izquierda) (tabla 18). Esto puede provocar confusión en el lector, principalmente en aquel que no tenga conocimientos sobre este campo al poder interpretar que está hablando de otra cosa.

Otra inconsistencia que detectamos en el mismo texto, en DeepL, es la traducción de *stent*. Cuando en el texto original se habla de *carotid stening*, es decir, la práctica de

colocar estens en la arteria carótida, traduce *stening* por «endoprótesis» (tabla 19), que es la forma recomendada de traducción al español de *stent* (Navarro, 2023h, definición 1). Sin embargo, cuando aparece la palabra *stent* la traduce por la forma adaptada en español «estent» (tabla 3).

#### 6.1.2.4. *Estilo*

Respecto a estilo, hemos detectado que tanto DeepL como Reverso presentan una redacción más literal, siguiendo una estructura más parecida al texto original, que ChatGPT.

En uno de los fragmentos del artículo sobre las enfermedades pulmonares se narra cronológicamente el estado y el tratamiento que recibe el paciente, para la apropiada adecuación de la narración, el original en inglés utiliza expresiones como *on day two*, *on day six*, que no se logran traducir de manera apropiada en español, sino que presentan una traducción literal. Tanto DeepL, como Reverso, traducen estas expresiones como «el día 2» y el «día 4», que, para este contexto, no son la traducción más adecuada, ya que puede dar a entender que hace referencia al día 2 o al día 6 del mes en el que pasa la acción. En este tipo de contexto es preferible utilizar expresiones idiomáticas del español más apropiadas como «el segundo día» o «el cuarto día», tal y como utiliza ChatGPT en su traducción, lo que consigue que la redacción del artículo en la lengua meta sea más adecuada y tenga fluidez.

<b>TO</b>	<i>He began receiving high-dose intravenous glucocorticoids <b>on day 2</b> of the hospitalization, and antibiotics were discontinued <b>on day 4</b> because an infectious cause was deemed to be unlikely.</i>
<b>TH</b>	<b>El segundo día</b> de hospitalización, comenzó a recibir una dosis alta de glucocorticoides intravenosa y <b>al cuarto día</b> se retiraron los antibióticos al descartarse la causa infecciosa.
<b>DeepL</b>	Comenzó a recibir altas dosis de glucocorticoides intravenosos <b>el día 2</b> de la hospitalización, y se suspendieron los antibióticos <b>el día 4</b> porque se consideró improbable una causa infecciosa.
<b>Reverso</b>	Comenzó a recibir dosis altas de glucocorticoides intravenosos <b>el día 2</b> de la hospitalización, y los antibióticos se suspendieron <b>el día 4</b> porque se consideró improbable una causa infecciosa.

<b>ChatGPT simple</b>	Comenzó a recibir glucocorticoides intravenosos en dosis altas en <b>el segundo día</b> de hospitalización y los antibióticos se suspendieron <b>el cuarto día</b> porque se consideraba improbable una causa infecciosa.
-----------------------	---

Tabla 20. Fragmento 15: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin

En ocasiones también hay traducciones redundantes, que a pesar de que aparezcan en el original, podrán evitarse en el texto meta empleando sinónimos que permitan transmitir lo mismo sin repetir las mismas palabras. Esto se ve en el siguiente fragmento:

<b>TO</b>	<i>The aim of cardiologists should be to help other specialists (neurologists, stroke physicians, radiologists, neurosurgeons, etc.) to improve the outcomes of stroke patients and to decrease the risk of stroke among those, who are at risk for stroke.</i>
<b>TH</b>	El objetivo de los cardiólogos debe ser ayudar a otros especialistas (neurólogos, médicos especializados en el ictus, radiólogos, neurocirujanos, etc.) <b>a mejorar los resultados de los pacientes con ictus y reducir las probabilidades de sufrirlo entre los pacientes de riesgo.</b>
<b>DeepL</b>	El objetivo de los cardiólogos debe ser ayudar a otros especialistas (neurólogos, médicos especialistas en ictus, radiólogos, neurocirujanos, etc.) <b>a mejorar los resultados de los pacientes con ictus y a disminuir el riesgo de ictus entre los que tienen riesgo de padecerlo.</b>
<b>Reverso</b>	El objetivo de los cardiólogos debe ser ayudar a otros especialistas (neurólogos, médicos de accidente cerebrovascular, radiólogos, neurocirujanos, etc.) para mejorar los resultados de los pacientes de accidente cerebrovascular <b>y para disminuir el riesgo de accidente cerebrovascular entre aquellos, que están en riesgo de accidente cerebrovascular.</b>
<b>ChatGPT simple</b>	El objetivo de los cardiólogos debe ser ayudar a otros especialistas (neurólogos, médicos de accidentes cerebrovasculares, radiólogos, neurocirujanos, etc.) <b>a mejorar los resultados de los pacientes con accidente cerebrovascular y disminuir el riesgo de accidente cerebrovascular entre aquellos que están en riesgo de sufrirlo.</b>

Tabla 2114. Fragmento 16: The role of cardiologists in stroke prevention and treatment

Como podemos ver, tanto en el original, como en las traducciones de los motores de traducción al traducir palabra por palabra del original, suenan todas igual de redundantes al repetir ictus o accidente cerebrovascular dos veces.

## 6.2. Impacto del *prompt* en ChatGPT

Después del análisis de los motores de traducción, vamos a presentar los resultados del análisis que se realizó comparando las traducciones de ChatGPT con el *prompt* simple y con el *prompt* completo y remarcar las mejoras respecto a un *prompt* u otro y comprobar si al añadir más información al *prompt* se presentan mejoras en la traducción, o de lo contrario empeora.

### 6.2.1. Análisis global

Texto	Fidelidad	Fluidez	Terminología	Estilo	Total
GPT_Ictus_S	5	0	6	3	29
GPT_Ictus_C	5	0	6	4	35
GPT_Pulmón_S	3	1	6	3	22
GPT_Pulmón_C	2	1	6	2	22

Tabla 22. Resultados de las traducciones propuestas por ChatGPT con los prompts simple y completo

En un primer análisis de cifras generales, se observa que, según los resultados de las traducciones, el artículo sobre el ictus traducido con el *prompt* más completo presenta una calificación insuficiente en cuanto a la calidad de la traducción con un total de 35 puntos, 10 puntos por encima de los 25 puntos permitidos. Este resultado es 5 puntos superior en comparación con la traducción en la que utilizamos el *prompt* simple. Además, con el nuevo *prompt*, se evidencia una ligera disminución en la calificación general de la traducción, a pesar de que solo se logra una reducción de un punto en las categorías de precisión y estilo.

Por otro lado, en la traducción del artículo sobre enfermedades pulmonares presenta una puntuación similar respecto a la traducción con el *prompt* simple con una diferencia de un punto. Esta traducción mantiene los mismos puntos en todas las categorías de análisis, a excepción de la de estilo, que empeora, subiendo un punto más, sumando un total de 3.

Con este primer resultado tenemos la impresión de que al utilizar el *prompt* más completo, no se logra mejorar, de manera notable, la traducción respecto a los errores previos registrados.

En el siguiente análisis por categorías veremos si esto es ciertamente así.

## 6.2.2. Análisis por categoría

Tal y como hicimos en el apartado anterior, se analizarán exhaustivamente las distintas categorías para identificar los errores que se presentan en las traducciones. En esta sección se presentará un análisis comparativo entre las dos versiones traducidas con el *prompt* más básico y el *prompt* más completo para evaluar si el *prompt* más completo es capaz de corregir los problemas existentes, si añade nuevos errores o, si de lo contrario, mejora otros aspectos del texto.

En esta ocasión, para ilustrar mejor los hallazgos, destacaremos los errores del fragmento en negrita y subrayaremos las mejoras para que se puedan distinguir las referencias.

### 6.2.2.1. Fidelidad

En esta primera categoría, que es la segunda con más cantidad de problemas, el *prompt* más completo no consigue mejorar el resultado en ninguna de las dos traducciones. Aun así, se analizarán los fragmentos en los que se detectan mejoras, a pesar de que no se corrigen los errores previos, y aquellos en los que puedan aparecer errores nuevos.

En el siguiente fragmento, vemos una mejora en la redacción del texto al añadir la especificación de que la sigla utilizada está en inglés y pertenece al término que acabamos de mencionar, técnica que aplican muchos traductores en los textos en los que aparecen siglas en inglés que pueden no tener una equivalencia en español. Esta mejora implica que se le da un contexto al lector sobre el significado de los términos:

<b>TO</b>	<i>Open carotid endarterectomy (CEA) and stenting have been extensively studied for primary stroke prevention (outside the acute phase of stroke) and after transient ischaemic <b>attacks</b> or minor strokes (modified Rankin scale 0-2).</i>
<b>TH</b>	La endarterectomía carotídea abierta (CEA) y la colocación de endoprótesis se ha estudiado ampliamente para la prevención primaria del ictus (fuera de la fase grave del ictus) y después de <b>accidentes</b> isquémicos transitorios o ictus leves (escala de clasificación modificada 0-2).

<b>ChatGPT simple</b>	La endarterectomía carotídea abierta (CEA) y la colocación de stents carotídeos han sido ampliamente estudiadas para la prevención primaria del accidente cerebrovascular (fuera de la fase aguda del accidente cerebrovascular) y después de <b>ataques</b> isquémicos transitorios o accidentes cerebrovasculares leves (escala de Rankin modificada 0-2).
<b>ChatGPT completo</b>	La endarterectomía carotídea abierta ( <u>CEA</u> , por sus siglas en inglés) y la colocación de stents carotídeos han sido ampliamente estudiadas para la prevención primaria del accidente cerebrovascular (fuera de la fase aguda del accidente cerebrovascular) y después de <b>ataques</b> isquémicos transitorios o accidentes cerebrovasculares leves (escala de Rankin modificada 0-2).

Tabla 23. Fragmento 3: The role of cardiologists in stroke prevention and treatment

En el siguiente fragmento, se añade un error que no existía previamente. En este caso, la traducción con el *prompt* completo traduce *fluids* como «fluidos» y no como «líquidos» como ya se había traducido previamente con el *prompt* simple. La dificultad que presenta *fluid* tal y como nos explica Libro Rojo (2023f) es que «en español, ‘fluido’ es cualquier sustancia en estado líquido o gaseoso; en el lenguaje médico inglés, en cambio, el término *fluid* se utiliza casi siempre de forma impropia en el sentido más restringido de **líquido**», por lo que en este contexto el término apropiado es «fluidos».

<b>TO</b>	<i>He sought care with his <u>primary care provider</u> and at multiple EDs for his gastrointestinal symptoms during the week before his admission and was treated with intravenous <b>fluids</b> and given metronidazole, levofloxacin, and an antiemetic agent.</i>
<b>TH</b>	Acudió a su <u>médico de cabecera</u> , y a varios servicios de urgencias, por síntomas gastrointestinales durante la semana anterior a su ingreso y le trataron con <b>líquidos</b> por vía intravenosa y le suministraron metronidazol, levofloxacino y antiemético.
<b>ChatGPT simple</b>	Buscó atención con su <u>proveedor de atención primaria</u> y en varios servicios de emergencia por sus síntomas gastrointestinales durante la semana antes de su ingreso y fue tratado con <b>líquidos</b> intravenosos y recibió metronidazol, levofloxacino y un agente antiemético.
<b>ChatGPT completo</b>	Buscó atención con su <u>médico de atención primaria</u> y en múltiples salas de emergencias por sus síntomas gastrointestinales durante la semana



	anterior a su ingreso y recibió tratamiento con <b>fluidos</b> intravenosos, metronidazol, levofloxacino y un agente antiemético.
--	---

Tabla 2415. Fragmento 4: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

A pesar de que en este fragmento hay un error que no había anteriormente, vemos como mejora la traducción de *primary care provider* proponiendo como traducción «médico de atención primaria» siendo un equivalente más adecuado al ser una expresión idiomática del español y se despegas de la literalidad del texto original que propuso el *prompt* simple: «proveedor de atención primaria».

En el siguiente fragmento del artículo sobre el ictus, vemos una incorporación adicional en la traducción, fenómeno que hasta el momento no había ocurrido en ChatGPT. En este fragmento el autor se refiere a los ancianos como «población» para evitar la repetición en el texto al hablar de la información limitada que reciben. Como veremos a continuación, ChatGPT añade que esta población es de riesgo, cuando en el original no se proporciona esta información:

<b>TO</b>	<i>Additionally, many patients with AF are elderly with other <b>co-morbidities</b> and there is limited information on the risk/benefit of these procedures in this population.</i>
<b>TH</b>	Además, muchos pacientes con FA son ancianos que presentan <b>enfermedades concomitantes</b> y la información sobre el riesgo/beneficio de estas técnicas es limitada en esta población.
<b>ChatGPT simple</b>	Además, muchos pacientes con fibrilación auricular son ancianos con otras <b>comorbilidades</b> y hay información limitada sobre el riesgo/beneficio de estos procedimientos en esta población.
<b>ChatGPT completo</b>	Además, muchos pacientes con FA son ancianos con otras <b>comorbilidades</b> y hay información limitada sobre el riesgo/beneficio de estos procedimientos en esta población de <u>riesgo cada vez mayor.</u>

Tabla 25. Fragmento 10: The role of cardiologists in stroke prevention and treatment

En este mismo fragmento podemos observar que este cambio de *prompt* no consigue solucionar la traducción de *co-morbidities*.

### 6.2.2.2. *Fluidez*

En esta categoría, solo se identificaron errores de fluidez en la traducción del artículo sobre enfermedades pulmonares. En particular, se observó una falta de concordancia de género en el adjetivo «pediátricos», que ya apareció en las traducciones que analizamos anteriormente y que, como podemos ver, persiste. Este error resalta la importancia de una hacer una revisión detallada de las traducciones que se hacen utilizando herramientas de traducción automática para asegurar la precisión y adecuación gramatical en la traducción de textos médicos especializados.

### 6.2.2.3. *Terminología*

La categoría de terminología sigue siendo la que presenta más errores. En esta ocasión, como veremos durante el análisis, observaremos como el *prompt* más completo es capaz de solucionar uno de los problemas terminológicos existentes en la traducción con el *prompt* básico, pero también empeora la traducción con un error nuevo.

El error nuevo que se añade en la traducción del artículo sobre el ictus es en la traducción de *carotid stening* siendo uno de los errores más graves que se detectan entre todos los motores de traducción que hemos analizado al traducir este término por «estenosis carotídea»:

<b>TO</b>	<i>Carotid stenting</i>
<b>TH</b>	Colocación de endoprótesis en la arteria carótida
<b>ChatGPT simple</b>	Stent caritoideo
<b>ChatGPT completo</b>	Estenosis carotídea

Tabla 166. Fragmento 2: *The role of cardiologists in stroke prevention and treatment*

En términos generales, la estenosis carotídea ocurre cuando el diámetro de la luz carotídea disminuye, lo cual puede provocar un ictus o un infarto cerebral (Quirón Salud, 2024). Por otro lado, la colocación de endoprótesis en la carótida es el método con el que se consigue abrir la arteria carótida y evitar que se obstruya utilizando una endoprótesis, o el también conocido esten (MedlinePlus, 2023). Este error es el que ha hecho que la traducción quede suspendida al marcarse como error «muy grave» porque compromete la precisión y la coherencia del texto, afectando negativamente a la reputación del autor

y la revista en la que se pueda publicar el artículo científico. Este error aumenta 10 puntos más a la puntuación total.

Tal y como hemos visto en el análisis de las categorías anteriores, en esta categoría también detectamos como en la traducción se mantienen errores que aparecían en la traducción anterior, pero que se mejora la redacción del texto. Uno de esos casos lo encontramos en el siguiente fragmento:

<b>TO</b>	<i>His <b>condition</b> worsened, and new respiratory symptoms developed, so he presented again to the ED, where he was found to have hypoxemia [...].</i>
<b>TH</b>	Su <b>estado</b> empeoró y surgieron nuevos síntomas respiratorios, por lo que acudió de nuevo a <u>urgencias donde le detectaron hipoxemia</u> [...].
<b>ChatGPT simple</b>	Su <b>condición</b> empeoró y desarrolló nuevos síntomas respiratorios, por lo que volvió a presentarse en el <u>servicio de emergencia, donde se encontró con hipoxemia</u> [...].
<b>ChatGPT completo</b>	Su <b>condición</b> empeoró y desarrolló nuevos síntomas respiratorios, por lo que volvió a presentarse <u>en la sala de emergencias, donde se encontró que tenía hipoxemia</u> [...].

Tabla 177. Fragmento 5: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

Aunque no logra corregir correctamente el término *condition*, el cual debe traducirse por «estado» o «situación», ya que en este contexto hace referencia al estado de la enfermedad del paciente (Navarro, 2023c, definición 2), se mejoran otros aspectos. Concretamente, se mejora la traducción de la sigla *ED* como «sala de emergencias» en vez de «servicio de emergencia» a portando más naturalidad al lenguaje de la traducción. Así mismo, mejora la traducción de *he was found to have hypoxemia*, la cual tradujo la primera vez como «se encontró con hipoxemia», que como podemos observar, ChatGPT obvió el verbo *have* logrando una expresión no idiomática. En esta nueva versión, propone «se encontró que tenía hipoxemia», que, a pesar de no ser una traducción perfecta, es una expresión un poco más acorde a como se puede expresar en español el significado del texto original.

En este fragmento, vemos que no se traduce correctamente el término *positive end-expiratory pressure* que «son aquellas técnicas que generan una presión positiva en las vías aéreas que facilita la eliminación de secreciones y previene el colapso de las vías aéreas» (Pereira, 2018). Ambas versiones presentan la traducción «presión positiva al final de la respiración» siendo esta propuesta una traducción literal del término en inglés.

<b>TO</b>	<i>He was intubated and mechanically ventilated, receiving a high Fio2 and <b>positive end-expiratory pressure</b>, and met the criteria for <u>moderate ARDS</u>.</i>
<b>TH</b>	Se le intubó y recibía ventilación mecánica con Fio2 alto y <b>presión telespiratoria positiva</b> y cumplía los criterios de <u>SDRA moderado</u> .
<b>ChatGPT simple</b>	Fue intubado y ventilado mecánicamente, recibiendo una alta Fio2 y <b>presión positiva al final de la espiración</b> , y cumplió con los criterios para síndrome de <u>distrés respiratorio agudo moderado</u> .
<b>ChatGPT completo</b>	Fue intubado y ventilado mecánicamente, recibiendo un alto Fio2 y <b>presión positiva al final de la espiración</b> , y cumplió con los criterios para <u>SDRA moderado</u> .

Tabla 188. Fragmento 10: *Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report*

Como podemos observar, a pesar de no poder solventar el error de traducción de *positive end-expiratory pressure*, sí es capaz de corregir la traducción «distrés respiratorio agudo moderado» de la traducción del *prompt* simple que hace referencia a *moderate ARDS*, que ya no solo no respeta el uso de las siglas que tienen su equivalencia reconocida en español, sino que es el nombre de la enfermedad que propone no existe. Con el uso del *prompt* completo es capaz de traducirlo correctamente.

El único error terminológico que logra solucionar es la traducción de *occluders* por «oclusores», el cual se tradujo por «ocludedores» (tabla 29) con el *prompt* simple creando una palabra que no existe.

Otros errores que tampoco logra corregir son las traducciones que vimos anteriormente de los términos *C-reactive protein* (tabla 30), y «estándar de atención» (tabla 31).

#### 6.2.2.4. *Estilo*

En cuanto a los errores estilísticos más relevantes, se observa que, en su mayoría, no se han solucionado con el nuevo *prompt*. Por ejemplo, en el texto sobre enfermedades pulmonares, el término *discharged home* se traduce de manera literal como «dado de alta a casa» (tabla 32) cuando simplemente podría traducirse como «dado de alta», ya que se sobreentiende que, al salir del hospital, el paciente irá a su domicilio.

En un análisis previo, se destacó que ChatGPT mejoraba la traducción de expresiones como *on day 2* y *on day 6*, traduciéndolas como «el segundo día» y «el sexto día» respectivamente. Estas traducciones eran más precisas y naturales en comparación con otros motores de traducción que utilizaban términos como «el día 2» lo que podía generar confusión. No obstante, en esta ocasión, ChatGPT da un paso atrás en precisión al traducir las expresiones por «el día 2» (tabla 32) y «el día 6» (tabla 33). Estas traducciones son menos naturales y crean un problema idiomático que antes no existía, afectando negativamente la fluidez y claridad del texto.

Esta persistencia de errores estilísticos resalta la necesidad de una revisión y corrección detallada de las traducciones automáticas, especialmente en textos médicos donde la precisión y claridad son cruciales. La capacidad de una traducción automática de entender y reproducir el estilo natural del lenguaje es esencial para asegurar la calidad y coherencia del texto traducido.

## 7. Conclusiones

En la comparación entre ChatGPT frente a DeepL y Reverso, hemos comprobado que DeepL destaca en la eficiencia de traducción por encima de ChatGPT. Sumando las puntuaciones, DeepL destaca en las categorías de terminología con 9 puntos respecto a los 12 de ChatGPT y en la categoría de precisión con 6 puntos, mientras que ChatGPT tiene 8. Por otro lado, vemos que ChatGPT va por delante en fluidez habiendo hecho solo 1 error y en estilo con 5 puntos frente a los 7 puntos de DeepL. Esta ventaja se debe a la función de generación de textos de ChatGPT cuya función para redactar con lenguaje natural le permite separarse más de las estructuras originales del texto en inglés, siendo más eficiente en este sentido. Por lo que podemos concluir que el potencial de ChatGPT está más enfocado en la redacción del texto, mientras que con DeepL podemos asegurarnos una mejor terminología y precisión del texto.

A pesar de que el nivel de calidad de los textos presentados es aceptable, es necesario hacer una postedición de las traducciones ofrecidas por todos los motores de traducción. Como hemos observado en el análisis, existen errores de traducción terminológicos que no se pueden dejar pasar por alto, especialmente en el campo donde estos errores pueden acarrear problemas de salud y desinformación.

Además de los errores lingüísticos, también debemos prestar atención a las incorporaciones de información nueva al texto y a las omisiones de contenido por parte del motor de traducción, ya que pueden comprometer la veracidad del texto original y desinformar al lector.

Respecto al cambio de *prompt*, hemos comprobado que agregar características propias del texto no mejora los problemas de carácter terminológico ni de fidelidad, sin embargo, sí que consigue mejorar la fluidez y el estilo del texto.

A pesar de que no mejora aspectos importantes como los términos especializados, sí que consigue mejorar la redacción de la traducción al especificar el tipo de texto que va a traducir y al asignarle un rol. Por lo tanto, concluimos que, a pesar de que ChatGPT actualmente no consigue estar a la altura de DeepL, es útil para mejorar la fluidez y el estilo de la traducción.

En resumen, ambas herramientas requieren una revisión posterior a la traducción para corregir los errores presentes. Según nuestra preferencia, podemos elegir entre DeepL y ChatGPT: si buscamos una herramienta que traduzca mejor la terminología especializada y nos permita enfocar la postedición en la redacción, debemos optar por DeepL. Si preferimos una herramienta que presente un texto fluido con expresiones idiomáticas del idioma de destino, podemos optar por ChatGPT y luego realizar una corrección más exhaustiva dedicada a la terminología.

Si extrapolamos estos resultados al campo más amplio de la traducción automática y la inteligencia artificial, podemos observar tendencias relevantes. La capacidad que tienen los modelos como ChatGPT sugiere que los motores de traducción automática podrían beneficiarse significativamente al integrar IA en sus sistemas, lo que puede ayudar a comprender y reproducir el contexto y el estilo del lenguaje humano con mayor precisión. Esto no solo mejoraría la calidad de las traducciones automáticas, sino que también este método podría aplicarse en otras áreas en las que es crucial la naturalidad y la fluidez del texto.

Por otro lado, la fortaleza de DeepL en la precisión y la terminología especializada, indica un futuro en el que la combinación de diferentes herramientas de inteligencia artificial podría ofrecer soluciones de mayor calidad y adaptadas a las necesidades específicas.

El continuo desarrollo y perfeccionamiento de estas tecnologías promete mejoras en la calidad de las traducciones y abrir fronteras en la interacción humano – máquina, lo

que optimizaría a largo plazo la forma en la que abordamos la comprensión y producción del lenguaje.

## 8. Bibliografía

- Alonso Martín, J.A. et al. (2012) *Tecnologías del lenguaje*. Editorial UOC. 2012.
- Arimetrics (2024). *Qué es ChatGPT*. Arimetrics. <https://www.arimetrics.com/glosario-digital/chatgpt>
- Ariza, M.A. (2009). La comunicación especializada de la medicina en español, francés e inglés: niveles profesional y académico. *Lebende Sprachen*, (54)3, 127-130. <https://doi.org/10.1515/les.2009.030>
- Baniata, L.H., Isaac. K. E., et al. (2021). A Transformer-Based Neural Machine Translation Model for Arabic Dialects That Utilizes Subword Units. *Sensors* (21)19, 6509. <https://doi.org/10.3390/s21196509>
- Barquilla, Y. (2024). *Qué es un prompt en Inteligencia Artificial*. BeeDigital. <https://www.beedigital.es/tendencias-digitales/que-es-un-prompt-en-inteligencia-artificial/>
- Generación Aprende. [Generación Aprende] (21 de septiembre de 2023). *Cómo TRADUCIR con CHATGPT: Prompts Infalibles* [vídeo]. Youtube. <https://www.youtube.com/watch?v=1FiG62ca9GY>
- Bashir, S. (2023). *ChatGPT and the Future of Translation*. CCJK. <https://www.ccjk.com/chatgpt-and-the-future-of-translation/>
- DeepL. (2021). *¿Cómo funciona DeepL?*. DeepL. <https://www.deepl.com/es/blog/how-deepl-works>
- Diego Olite, F. M.; Morales Suarez, I. R. y Vidal Ledo, M. J. (2023). Chat GPT: origen, evolución, retos e impactos en la educación. *REMS*. (37)2. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-21412023000200016](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412023000200016)
- Dimitriadou, A. (2023). *ChatGPT's translation strengths & limitations*. Pangea. <https://www.pangea.global/blog/chatgpts-translation-strengths-limitations/>
- Dogukan, Y. et al. (2023) AI-Driven Labor Substitution Evidence from Google Translate and ChatGPT. *INSEAD* working paper. <https://dx.doi.org/10.2139/ssrn.4400516>



Domènech, O. (2006) Capitulo I. Textos especializados y niveles de especialización. En: *Textos especialitzats i variació vertical: la diversitat terminològica com a factor discriminant del nivell d'especialització d'un text*. 25 – 74. Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada.

Dongxing L., Zuying L. (2022). An Improved Transformer-Based Neural Machine Translation Strategy: Interacting-Head Attention. *Computational Intelligence and Neuroscience*. (2022). ID 2998242, 21. <https://doi.org/10.1155/2022/2998242>

Fernández, Y. (2024). *ChatGPT: qué es, cómo usarlo y qué puedes hacer con este chat de inteligencia artificial GPT*. Xataka. <https://www.xataka.com/basics/chatgpt-que-como-usarlo-que-puedes-hacer-este-chat-inteligencia-artificial>

Guantiva Acosta, R.; Cabré Castellví, M. T.; Castellà Lidon, J. M. (2018). Clasificación de textos especializados a partir de su terminología. *Íkala, revista de lenguaje y cultura*. (13)19, 15-39. Universidad de Antioquia, Medellín, Colombia.

He, S. (2024). *Prompting ChatGPT for Translation: A Comparative Analysis of Translation Brief and Persona Prompts* [Tesis doctoral, School of Culture and Communication]. Swansea University. Arxiv. <https://arxiv.org/pdf/2403.00127>

Işim, Ç. y Selim Balcioglu, Y. (2023). *ChatGPT Performance of Translate* [, Faculty of Business Administration, Business Administration Department] Universidad técnica de Gezbe. [https://www.researchgate.net/profile/Yavuz-Balcioglu/publication/369370220\\_CHATGPT\\_PERFORMANCE\\_OF\\_TRANSLATE/links/641815cc66f8522c38bc16ad/CHATGPT-PERFORMANCE-OF-TRANSLATE.pdf](https://www.researchgate.net/profile/Yavuz-Balcioglu/publication/369370220_CHATGPT_PERFORMANCE_OF_TRANSLATE/links/641815cc66f8522c38bc16ad/CHATGPT-PERFORMANCE-OF-TRANSLATE.pdf)

Jiao, W. et al. (2023). Is ChatGPT a Good Translator? Yes with GPT-4 the engine. *Tencent AI lab*. [https://www.researchgate.net/publication/367359399\\_Is\\_ChatGPT\\_A\\_Good\\_Translator\\_A\\_Preliminary\\_Study](https://www.researchgate.net/publication/367359399_Is_ChatGPT_A_Good_Translator_A_Preliminary_Study)

Larroyed, A. (2023). Redefining Patent Translation: The Influence of ChatGPT and the Urgency to Align Patent Language Regimes in Europe with Progress in Translation Technology. *GRUR International*. 72(11), 1009 — 1017. <https://doi.org/10.1093/grurint/ikad099>

Layden, J. E., Ghinai, I., et al. (2019). Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report. *The New England Journal of Medicine*, (382)10, 903 — 916. DOI: 10.1056/NEJMoa1911614

Lettere Dragosovljevič, M. G. (2023). *Estudio comparativo del traductor de idiomas Google Translate y ChatGPT vs a utilización de algoritmos de Machine Translation de Python* [Trabajo previo a la obtención del título de ingeniero de sistemas y computación]. Facultad de Ingeniería, Escuela de Sistemas. Pontificia Universidad del Ecuador. <https://repositorio.puce.edu.ec/server/api/core/bitstreams/174e3851-3eaa-436e-8e50-b961b790e888/content>

McCarthy, J. (2007). *What is artificial intelligence?* Computer Science Department Stanford University. <https://www-formal.stanford.edu/jmc/whatisai.pdf>

Merritt, R. (2022). *¿Qué es un modelo transformer?* NVIDIA. <https://la.blogs.nvidia.com/blog/que-es-un-modelo-transformer/>

Navarro, F. A. (2023). Bleeding. En: *Diccionario de dudas y dificultades de traducción del inglés médico*, de: <https://www.cosnautas.com/es/libro/6990-bleeding>

Navarro, F. A. (2023). Comorbidity. En: *Diccionario de dudas y dificultades de traducción del inglés médico*, de: <https://www.cosnautas.com/es/libro/12751-comorbidity>

Navarro, F. A. (2023). Condition. En: *Diccionario de dudas y dificultades de traducción del inglés médico*, de: <https://www.cosnautas.com/es/libro/13088-condition>

Navarro, F. A. (2023). C-reactive protein. En: *Diccionario de dudas y dificultades de traducción del inglés médico*, de: <https://www.cosnautas.com/es/libro/8809-c-reactive-protein>

Navarro, F. A. (2023). *Diccionario de dudas y dificultades de traducción del inglés médico* (4.ª edición). Edición en línea (versión 4.04). Madrid: Cosnautas.

Navarro, F. A. (2023). Fluid. En: *Diccionario de dudas y dificultades de traducción del inglés médico*, de: <https://www.cosnautas.com/es/libro/22988-fluid>

Navarro, F. A. (2023). Procedure. En: *Diccionario de dudas y dificultades de traducción del inglés médico*, de: <https://www.cosnautas.com/es/libro/46923-procedure>

Navarro, F. A. (2023). Stent. En: *Diccionario de dudas y dificultades de traducción del inglés médico*, de: <https://www.cosnautas.com/es/libro/55313-stent>

Northwood, A. (2023). *The Great ChatGPT Debate: Boon or Bane for Humanity?* Medium. <https://medium.com/@alexnorthwood/the-great-chatgpt-debate-boon-or-bane-for-humanity-e8c5c9610f45>

Ortiz, P. (2023) *Chat GPT: qué es, para qué sirve y su aplicación en la economía [explicado por Chat GPT]*. EDEM. <https://edem.eu/chat-gpt-que-es-para-que-sirve-y-su-aplicacion-en-la-economia-explicado-por-chat-gpt/>

Parra Escartín, C. (2022). *¿Cómo ha evolucionado la traducción automática en los últimos años?* La Linterna del Traductor. <https://lalinternadeltraductor.org/n16/traduccion-automatica.html>

Pereira Brun, M., Caña Pino, A., y Apolo Arenas, M. D. (2018). Técnicas de presión espiratoria positiva (PEP) en la fibrosis quística. Revisión bibliográfica. *Cuestiones de fisioterapia: revista universitaria de información e investigación en Fisioterapia*. (47)2. 131 —141. <https://dialnet.unirioja.es/servlet/articulo?codigo=6739441#:~:text=Las%20t%C3%A9cnicas%20de%20presi%C3%B3n%20espiratoria,colapso%20de%20las%20v%C3%ADas%20a%C3%A9reas>

Reverso. (2024a). *Context*. Reverso. <https://context.reverso.net/traduccion/about>

Reverso. (2024b). *Nuestros planes*. Reverso. <https://documents.reverso.net/Pricing.aspx?lang=es#:~:text=%C2%BFC%C3%B3mo%20funciona%20la%20traducci%C3%B3n%20de,empresas%20y%20empresas%20de%20traducci%C3%B3n>

Reverso. (2024c). *Translate text*. Reverso. <https://www.reverso.net/text-translation>

Sánchez Ramos, M. M. y Rico Pérez, C. (2020a). La automatización de la traducción automática: definición y evaluación de la traducción automática. *Traducción automática:*

*conceptos clave, procesos de evaluación y técnicas de posesición.* 1 – 9. Editorial Comales.

Sánchez Ramos, M. M. y Rico Pérez, C. (2020b). Sistemas de traducción automática. *Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición.* 11 – 30. Editorial Comales.

Soto, X.; Olats Pérez-De-Vinaspre, G.L; Ornoz, M. (2019). Neural machine translation of clinical texts between long distance. *Journal of the American Medical Informatics Association*, 26(12), 1478–1487.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7647170/pdf/ocz110.pdf>

Taus. (2024). *Start tracking error with DQF-MQM.* Taus. <https://info.taus.net/dqf-mqf-error-typology-template-download>

Widimsky, P., Wolfram, D., et al. (2018). The role of cardiologists in stroke prevention and treatment: position paper of the European Society of Cardiology Council on Stroke. *European Heart Journal*, (39)17, 1567–1573. <https://doi.org/10.1093/eurheartj/ehx478>

## 9. Anexos

<b>TO</b>	<i>On day 2 of hospitalization, <b>the patient</b> underwent bronchoscopy, which showed normal appearing bronchi.</i>
<b>TH</b>	El segundo día de hospitalización, <b>el paciente</b> fue sometido a una broncoscopia la cual mostró bronquios de aspecto normal.
<b>DeepL</b>	El segundo día de hospitalización, <b>la paciente</b> fue sometida a una broncoscopia, que mostró bronquios de aspecto normal.
<b>Reverso</b>	El día 2 de la hospitalización, <b>el paciente</b> fue sometido a broncoscopia, que mostró bronquios de apariencia normal.
<b>ChatGPT simple</b>	En el segundo día de hospitalización, el paciente se sometió a broncoscopia, que mostró bronquios de apariencia normal.

Tabla 10. fragmento 12: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

<b>TO</b>	<i>The role of cardiologists in <b>stroke</b> prevention and treatment: position paper of the European Society of Cardiology Council on Stroke</i>
<b>TH</b>	El papel de los cardiólogos en la prevención y tratamiento del <b>ictus</b> : documento de posición del Consejo del Ictus de la Sociedad Europea de Cardiología
<b>DeepL</b>	El papel de los cardiólogos en la prevención y <b>el</b> tratamiento del <b>ictus</b> : documento de posición del Consejo de Ictus de la Sociedad Europea de Cardiología
<b>Reverso</b>	El papel de los cardiólogos en la prevención y el tratamiento del <b>ictus</b> : documento de posición del Consejo de la Sociedad Europea de Cardiología sobre el ictus
<b>ChatGPT simple</b>	El papel de los cardiólogos en la prevención y tratamiento del <b>accidente cerebrovascular</b> : documento de posición del Consejo de Accidente Cerebrovascular de la Sociedad Europea de Cardiología

Tabla 15. Fragmento 1: The role of cardiologists in stroke prevention and treatment

<b>TO</b>	<i>Open carotid endarterectomy (CEA) and stenting have been extensively studied for <b>primary stroke prevention</b> (outside the <b>acute phase of stroke</b>) and after transient ischaemic attacks or minor strokes (modified Rankin scale 0-2).</i>
<b>TH</b>	La endarterectomía carotídea abierta (CEA) y la colocación de endoprótesis se ha estudiado ampliamente para la <b>prevención primaria del ictus</b> (fuera de la

	<b>fase grave del ictus</b> ) y después de accidentes isquémicos transitorios o ictus leves (escala de clasificación modificada 0-2).
<b>DeepL</b>	La endarterectomía carotídea abierta (EAC) y la colocación de estents se han estudiado ampliamente para la <b>prevención primaria del ictus</b> (fuera de la <b>fase aguda del ictus</b> ) y tras ataques isquémicos transitorios o ictus leves (escala Rankin modificada 0-2).
<b>Reverso</b>	La endarterectomía abierta de la carótida (ACE) y el stent se han estudiado extensamente para la <b>prevención primaria del ACV</b> (fuera de la <b>fase aguda del ACV</b> ) y después de ataques isquémicos transitorios o accidentes cerebrovasculares menores (escala de Rankin modificada 0-2).
<b>ChatGPT simple</b>	La endarterectomía carotídea abierta (CEA) y la colocación de <u>stents carotídeos</u> han sido ampliamente estudiadas para la <b>prevención primaria del accidente cerebrovascular</b> (fuera de la <b>fase aguda del accidente cerebrovascular</b> ) y después de ataques isquémicos transitorios o accidentes cerebrovasculares leves (escala de Rankin modificada 0-2).

Tabla 16. Fragmento 3: The role of cardiologists in stroke prevention and treatment

<b>TO</b>	<i>Left atrial appendage (LAA) occlusion is emerging as a possible alternative to anticoagulation for patients carrying very high bleeding risk, especially those that have already suffered intracranial haemorrhage because of anticoagulation.</i>
<b>TH</b>	El cierre de la orejuela auricular izquierda ( <b>OAI</b> ) es una posible alternativa novedosa a la anticoagulación para pacientes con alto riesgo de hemorragia, especialmente aquellos que ya han sufrido hemorragias cerebrales por anticoagulación.
<b>DeepL</b>	La oclusión de la orejuela auricular izquierda ( <b>OAI</b> ) se perfila como una posible alternativa a la anticoagulación para los pacientes con un riesgo hemorrágico muy elevado, especialmente los que ya han sufrido una hemorragia intracraneal debido a la anticoagulación.
<b>Reverso</b>	La oclusión del apéndice auricular izquierdo ( <b>LAA</b> ) está emergiendo como una posible alternativa a la anticoagulación para pacientes con riesgo hemorrágico muy alto, especialmente aquellos que ya han sufrido hemorragia intracraneal debido a la anticoagulación.

<b>ChatGPT simple</b>	El cierre de la orejuela izquierda ( <b>LAA</b> ) está emergiendo como una posible alternativa a la anticoagulación para pacientes con un riesgo de sangrado muy alto, especialmente aquellos que ya han sufrido hemorragias intracraneales debido a la anticoagulación.
-----------------------	--

Tabla 17. Fragmento 17: The role of cardiologists in stroke prevention and treatment

<b>TO</b>	<i>In two studies patients with device implanted into the <b>LAA</b> had fewer haemorrhagic strokes, but more ischaemic strokes compared with warfarin.</i>
<b>TH</b>	En dos estudios, los pacientes con un dispositivo implantado en la <b>OAI</b> han sufrido menos ictus hemorrágicos, pero más ictus isquémicos en comparación con la warfarina.
<b>DeepL</b>	En dos estudios, los pacientes con el dispositivo implantado en la <b>OI</b> sufrieron menos ictus hemorrágicos, pero más isquémicos, en comparación con la warfarina.
<b>Reverso</b>	En dos estudios, los pacientes con dispositivo implantado en la <b>LAA</b> tuvieron menos accidentes cerebrovasculares hemorrágicos, pero más accidentes cerebrovasculares isquémicos en comparación con la warfarina.
<b>ChatGPT simple</b>	En dos estudios, los pacientes con dispositivo implantado en la <b>LAA</b> tuvieron menos accidentes cerebrovasculares hemorrágicos, pero más accidentes cerebrovasculares isquémicos en comparación con la warfarina.

Tabla 18. Fragmento 20: The role of cardiologists in stroke prevention and treatment

<b>TO</b>	<b><i>Carotid stenting</i></b>
<b>TH</b>	Colocación de <b>endoprótesis</b> en la arteria carótida
<b>DeepL</b>	Colocación de <b>endoprótesis</b> carotídea
<b>Reverso</b>	<b>Stent carotídeo</b>
<b>ChatGPT simple</b>	<b>Stent caritoideo</b>

Tabla 19. Fragmento 2: The role of cardiologists in stroke prevention and treatment

<b>TO</b>	<i>One study, comparing <b>LAA occluders</b> versus novel oral anticoagulant apixaban is ongoing.</i>
<b>TH</b>	Hay una investigación en curso que está comparando los <b>cierres</b> de <b>OAI</b> con el nuevo anticoagulante oral apixaban.

<b>ChatGPT simple</b>	Un estudio en curso compara los <b>ocludedores</b> de LAA con el anticoagulante oral novedoso apixaban.
<b>ChatGPT completo</b>	Un estudio que compara <b>oclusores</b> de la LAA con el anticoagulante oral novedoso apixaban está en curso.

Tabla 29. Fragmento 22: The role of cardiologists in stroke prevention and treatment

<b>TO</b>	<i>His laboratory results showed leukocytosis (white-cell count, 18,000 cells per cubic millimeter) with a neutrophil predominance (94%) and no eosinophils (0%), as well as elevated inflammatory markers with a <b>C-reactive protein</b> level of 32 mg per deciliter and an erythrocyte sedimentation rate of 68 mm per hour.</i>
<b>TH</b>	Los resultados de laboratorio mostraron leucocitosis (recuento de leucocitosis, 18.000 células por milímetro cúbico) con predominio de neutrófilos (94%) y ausencia de eosinófilo (0%), así como marcadores inflamatorios elevados con un nivel de <b>proteína C-reactiva</b> de 32 mg por decilitro y una velocidad de sedimentación globular de 68 mm por hora.
<b>ChatGPT simple</b>	Sus resultados de laboratorio mostraron leucocitosis (conteo de células blancas, 18,000 células por milímetro cúbico) con predominio de neutrófilos (94%) y sin eosinófilos (0%), así como marcadores inflamatorios elevados con un nivel de <b>proteína C reactiva</b> de 32 mg por decilitro y una velocidad de sedimentación de eritrocitos de 68 mm por hora.
<b>ChatGPT completo</b>	Sus resultados de laboratorio mostraron leucocitosis (recuento de glóbulos blancos, 18,000 células por milímetro cúbico) con predominio de neutrófilos (94%) y ausencia de eosinófilos (0%), así como marcadores inflamatorios elevados con un nivel de <b>proteína C reactiva</b> de 32 mg por decilitro y una velocidad de sedimentación de eritrocitos de 68 mm por hora.

Tabla 3019. Fragmento 7: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

<b>TO</b>	<i>Other ongoing studies in AF patients with contraindications for long term anticoagulation compare LAA occlusion with <b>standard of care</b>.</i>
<b>TH</b>	Otras investigaciones en curso comparan el cierre de la OAI con el <b>tratamiento estándar</b> en pacientes con FA y con contraindicaciones para la anticoagulación a largo plazo.



<b>ChatGPT simple</b>	Otros estudios en curso en pacientes con fibrilación auricular con contraindicaciones para la anticoagulación a largo plazo comparan el cierre de la LAA con el <b>estándar de atención</b> .
<b>ChatGPT completo</b>	Otros estudios en curso en pacientes con FA con contraindicaciones para la anticoagulación a largo plazo comparan la oclusión de la LAA con el <b>estándar de atención</b> .

Tabla 31. Fragmento 23: The role of cardiologists in stroke prevention and treatment

<b>TO</b>	<i>On day 2 of hospitalization, the patient underwent bronchoscopy, which showed normal appearing bronchi.</i>
<b>TH</b>	<b>Durante el segundo día</b> de hospitalización, el paciente se sometió a broncoscopia, que mostró bronquios de apariencia normal.
<b>ChatGPT simple</b>	<b>En el segundo día</b> de hospitalización, el paciente se sometió a broncoscopia, que mostró bronquios de apariencia normal.
<b>ChatGPT completo</b>	<b>El día 2 de hospitalización</b> , el paciente se sometió a broncoscopia, que mostró bronquios de aspecto normal.

Tabla 32. Fragmento 12: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report

<b>TO</b>	<i>The patient was <b>discharged home</b> on hospital <b>day 6</b> with instructions to continue an oral glucocorticoid–tapering regimen for 6 weeks.</i>
<b>TH</b>	<b>Al sexto día</b> , el paciente fue <b>dado de alta</b> con indicaciones de ir disminuyendo progresivamente el tratamiento oral de glucocorticoides durante 6 semanas.
<b>ChatGPT simple</b>	El paciente fue <b>dado de alta a su hogar el sexto día</b> de hospitalización con instrucciones de continuar un régimen de reducción gradual de glucocorticoides orales durante 6 semanas.
<b>ChatGPT completo</b>	El paciente fue <b>dado de alta</b> a casa el <b>día 6</b> de hospitalización con instrucciones de continuar un régimen de reducción gradual de glucocorticoides orales durante 6 semanas.

Tabla 33. Fragmento 23: Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report