

---

# Precisión en la traducción de metadatos: desafíos de la traducción automática en artículos científicos

---

*Accuracy in metadata translation: challenges of machine translation in scientific articles*

---

**Javier TARANGO, Haydee MENCHACA, Fidel GONZÁLEZ-QUIÑONES,  
Efraín Alfredo BARRAGAN-PEREA**

Universidad Autónoma de Chihuahua, Avenida Universidad s/n, Chihuahua, México 31174.  
jtarango@uach.mx, gmenchaca@uach.mx, fgonzalez@uach.mx, ebarragan@uach.mx

## Resumen

La traducción de documentos científicos en general y de los metadatos en particular (título del artículo, resumen y palabras clave) es un tema de importancia en la formación de investigadores cuya lengua materna es el español, publican en español y requieren del uso parcial del idioma inglés en estos elementos clave. El objetivo de esta investigación es analizar los niveles de efectividad de la traducción humana en comparación con la traducción automática, recurriendo al desarrollo de un estudio cuantitativo, descriptivo, transeccional y no experimental, evaluando los niveles de calidad en la traducción de 410 resúmenes de artículos científicos provenientes de 27 revistas del área de humanidades publicados en el año 2022 en el Sistema de Clasificación de Revistas Mexicanas de Ciencia y Tecnología. La evaluación de los datos se basa en una rúbrica compuesta por 15 criterios de tipo sintáctico-semántico, léxico-semántico, morfosintáctico, ortotipográfico y pragmático-cultural, evaluando la traducción humana (original) y de tres traductores automáticos de libre acceso (Google, Microsoft y Systran). El estudio analiza fortalezas y debilidades de la traducción humana y automática, con lo cual se identifican elementos de efectividad de ambas formas de traducción aplicadas a textos breves con fines de favorecer a investigadores con limitantes en el bilingüismo español-inglés.

**Palabras clave:** Traducción humana. Traducción automática. Metadatos. Resúmenes. Artículos científicos. Revistas científicas mexicanas en humanidades. Comunicación científica.

## 1. Introducción

En los procesos de localización de temáticas y documentos específicos, en cualquier base de datos o sistema de información (manual y electrónico), los metadatos se vuelven indispensables (para el caso: títulos, resumen y palabras clave) como elementos fundamentales en el sentido de que son condensados del documento, ya que establecen el tema a partir del cual se aportan las novedades, cobrando importancia tanto para el artículo científico como para el objeto de estudio. Este tema de

## Abstract

The translation of scientific documents in general and metadata in particular (article title, abstract, and keywords) is an important issue in the training of researchers whose mother tongue is Spanish, who publish in Spanish and who require partial use of the English language in these key elements. The objective of this research is to analyze the effectiveness of human translation compared to machine translation using a quantitative, descriptive, cross-sectional, and non-experimental study. This study evaluates the translation quality of 410 abstracts of scientific articles from 27 humanities journals published in 2022 in the Mexican Journal Classification System for Science and Technology. The data evaluation is based on a rubric composed of 15 criteria of a syntactic-semantic, lexical-semantic, morphosyntactic, orthotypographical, and pragmatic-cultural nature, assessing human (original) translation and that of three freely available machine translation tools (Google, Microsoft, and Systran). The study analyzes the strengths and weaknesses of human and machine translation, thereby identifying elements of effectiveness of both forms of translation applied to short texts in order to support researchers with limitations in Spanish-English bilingualism.

**Keywords:** Human translation. Machine translation. Metadata. Abstracts. Scientific articles. Mexican scientific journals in the humanities. Scientific communication.

estudio resulta valioso, tanto para investigadores de la traducción, como para profesionales de la organización y representación del conocimiento, especialmente ante la abundancia de información y la complejidad en el acceso preciso a ella.

En general, los metadatos ofrecen un futuro prometedor ya que constituyen un área fundamental de la investigación académica y la práctica profesional, alcanzando en su beneficio a diversas personas e instituciones (Mayernik, 2020). En específico, las palabras clave guardan una función

principal, ya que permiten caracterizar documentos para mejorar la localización, tanto del documento completo como de fragmentos dentro del mismo, con lo cual son consideradas como uno de los principales lenguajes de indización (Lardera y Hjørland, 2021).

Históricamente debe resaltarse que, al finalizar la segunda Guerra Mundial el idioma inglés empieza a cobrar relevancia como lengua franca o lengua vehicular de la ciencia, impactando descomunalmente las disciplinas científicas, las cuales lo emplean casi de forma exclusiva para la divulgación y comunicación del conocimiento generados a través de diversos medios (Vicente Cubero, 2019). Debe tenerse en cuenta que la gran mayoría de la comunidad científica internacional no posee dicho idioma como lengua materna, no obstante, esto no ha sido un impedimento para que la proporción de artículos escritos en inglés y firmados por autores no anglófonos aumente en los últimos años (López Navarro et al., 2017).

Existen perspectivas que justifican la adopción de un lenguaje científico universal, lo cual ofrecería un efecto positivo en la comunidad científica facilitando el acceso a la literatura científica, además de permitir la comunicación con otros científicos en cualquier parte del mundo (Ulitkin et al., 2021), con lo cual se generan vínculos y redes de investigación, tanto de los autores como en beneficio de las propias revistas (Beltrán-Santoyo et al., 2021). Sin embargo, algunos otros autores consideran que el inglés como lengua vehicular tiene varias desventajas, ya que se convierte en un desafío para los autores académicos no nativos, quienes deben someter sus textos a procesos de edición y revisión, lo cual implica un costo elevado que no necesariamente se puede asumir (Vardaro et al., 2019; Acosta Plata, 2021).

Los estudios sobre el uso del idioma inglés para fines académicos por parte de la comunidad científica no anglófona señalan que el uso eficaz del lenguaje es de los retos más trascendentales que enfrenta cualquier investigador sin importar su lugar de origen o lengua materna. Así, los comentarios de los revisores acerca del uso de un idioma distinto aparecen recurrentemente en todas las evaluaciones, sin exceptuar a las de los autores cuya lengua materna es el inglés. A pesar de estos señalamientos, otros estudios exponen la presión a la que han sido sometidos los miembros de la comunidad investigadora no anglófona para publicar en inglés en lugar de realizarlo en su lengua materna. Aunque tampoco se puede negar que gran parte de los investigadores eligen en qué idioma publicar, lo cual es influenciado por su deseo de extender su audiencia más allá de las fronteras nacionales (López Navarro et al., 2017).

Respecto a las barreras lingüísticas, éstas pueden ser un obstáculo para acceder a la cuantiosa información generada; y es que es tan considerable la cantidad de información que se produce, que es improbable satisfacer la demanda de traducciones confiando de forma exclusiva en traductores humanos profesionales (Pineda González y Tarango, 2018). Consecuentemente, el uso de la traducción automática puede resultar un reto de la lingüística computacional, ya que, al imitar a la capacidad humana, podría llegarse a la generación de estructuras de procesamiento del lenguaje tan similares que pongan en duda la capacidad humana en sí (Fernández Torné y Matamala, 2021). Existen diversos sistemas de traducción automática basados en diferentes enfoques de traducción y algoritmos, los cuales tienen ventajas y desventajas, pero una característica que comparten todos es que por lo general son usados como dispositivos totalmente automáticos, es decir, traducen un texto fuente a un idioma destino sin la intervención de humanos (Wang et al., 2021; Rivera-Trigueros, 2022).

En cuanto a las perspectivas de los expertos en materia de traducción, hay opiniones encontradas sobre el uso de la traducción automática. Algunos afirman que puede ser de gran utilidad a la hora de traducir textos técnicos debido a que, gracias a su precisión y características semánticas, sintácticas y estilísticas, son idóneos para obtener una traducción automática aceptable (Fernández Torné y Matamala, 2021). Otros tantos consideran que la completa automatización con traducción automática se encuentra muy lejos de la realidad actual ya que los textos producidos requieren de revisión (post-edición) por traductores humanos para obtener un resultado apropiado (Ferrer Año, 2019). Muchos otros están de acuerdo en que los principales obstáculos de la traducción automática no son computacionales sino lingüísticos, de ambigüedad léxica, complejidad sintáctica, diferencias de vocabulario entre idiomas, de elíptica y de construcciones no gramaticales. Son considerados problemas que implican la extracción del significado de frases y textos desde el análisis de signos escritos y la producción de oraciones y textos con un significado equivalente (Viver Sorolla, 2018).

Tomando como referencia las perspectivas anteriores, este estudio tiene como finalidad analizar el tipo de errores que se incurre con más frecuencia la traducción generada por los traductores automáticos. Todo ello en comparación con la traducción original de resúmenes como unidades aisladas de artículos de revistas científicas. El propósito es probar si la traducción automática ha llegado a un nivel que rivalizaría con la traducción humana, o que, por lo contrario, está todavía

lejana de las capacidades humanas de traducción. Por tanto, se examinan a detalle las condiciones sintáctico-semánticas, léxico-semánticas, morfosintácticas, ortotipográficas y pragmático-culturales de traducción humana y automática que observan los metadatos (específicamente en el resumen al convertirlo a abstract) en artículos científicos publicados durante el año 2022 incluidos en revistas del área de humanidades del Sistema de Clasificación de Revistas Mexicanas de Ciencia (México).

## 2. Implicaciones en la calidad de la traducción automática ante la ausencia de bilingüismo

La función del traductor es entendida como un fenómeno lingüístico-cultural, basado en dos visiones fundamentales al caso: la teoría funcionalista de la traducción y el análisis del discurso. La correcta aplicación de los principios teóricos antes mencionados, han demostrado tener especial relevancia por sus contribuciones a la traducción en general y específicamente en la estructura del discurso científico. Las barreras geográficas y lingüísticas han constituido un reto a la transmisión del conocimiento a lo largo de la historia. La principal estrategia para solventar dichas barreras comunicativas es el uso de una lengua común, también conocida como lengua vehicular, lengua franca o lingua franca (Acosta Plata, 2021) y en ámbitos académicos, como lengua vehicular de la ciencia.

El crecimiento sostenido del mercado de la traducción afecta a todos los países del mundo y ha llevado a algunos a investigar e implementar programas de traducción automática como una posible solución al aumento de la demanda. En la medida en que la comunidad científica comience a reconocer esta barrera lingüística, que se valore la producción científica en otros idiomas y que se mejore la accesibilidad a la producción escrita en inglés mediante su traducción, será posible que los usuarios de este conocimiento puedan beneficiarse. Aunque existen recomendaciones acerca de comenzar a evaluar de forma equitativa las publicaciones realizadas en idiomas diferentes al inglés, no parece haber directrices claras sobre cómo poner estas medidas en práctica (Acosta Plata, 2021).

En la literatura científica existe cierta discrepancia al querer identificar un concepto preciso sobre lo que se entiende por traducción de calidad, pero sí consiguen concordar en condiciones para identificar una serie de rasgos generales que identifiquen su definición (Pérez Rodríguez y Huertas Abril, 2016; Martínez Mateo, 2022). Para definir tal concepto y adaptarlo a las circunstancias de interés,

será necesario enfatizar en la definición de su objetivo, necesidades del cliente y la adecuación de textos según sus circunstancias. De esta forma es posible identificar los fines de la traducción: (1) comprensión del contexto, un caso de uso en el que la traducción debe entenderse, pero no es necesaria una calidad imoluta; y (2) difusión, donde la calidad es un factor imprescindible del proceso de traducción.

De acuerdo con Ragni y Nunes Vieira (2011), Moor-kens (2018) y Chatzikoumi (2020) los criterios básicos pueden, por lo tanto, resumirse como: la fluidez en el idioma de destino, que incluye gramaticalidad y naturalidad; adecuación como en la equivalencia semántica y pragmática entre el texto fuente y el texto meta; y cumplimiento con las posibles especificaciones del solicitante. Durante los últimos años se han desarrollado multitud de técnicas de evaluación, que se han utilizado para todo tipo de sistemas de traducción automática basados en reglas (usos gramaticales más adecuados), estadísticos (frecuencias de formas de traducción), híbridos (combinación de patrones gramaticales y frecuencias) y neuronales (donde se consideran esquemas de pensamiento). En la literatura internacional, las técnicas de evaluación se clasifican como métricas automáticas o humanas (o manuales).

Para mantener los estándares de calidad en la traducción es necesario evitar errores que la lleven a una condición de carencia de rigor, evitando con ello, los que se determinan durante el proceso de la labor de traducción y los que aparecen después de finalizar este proceso (Cossío Montalvo y Rodríguez, 2021; Roble, 2022). Una determinada expresión puede considerarse “correcta” en el sentido de que es funcional o “incorrecta”, según el contexto. En cambio, para llevar a cabo el análisis del texto meta, se propone su propia clasificación de errores de traducción automática, después de analizar algunas de las más importantes clasificaciones que se identifican en la literatura científica.

La clasificación propuesta para evaluar las condiciones de la traducción debe ajustarse de manera más adecuada al texto a analizar y se deben considerar los siguientes elementos tipo (Hidalgo, 2009; Gramigna, 2013; Cossío Montalvo y Rodríguez, 2021): léxico-semánticos, morfosintácticos, ortotipográficos, pragmático-culturales y retóricos.

Tales características son necesarias en la toma de decisiones del traductor al desarrollar su trabajo: identificar el texto meta, su intención y características de los interesados en el trabajo. Este criterio se rige por patrones lingüísticos y pragmáticos, identificados y clasificados por Maldonado y Liébana González (2021) como errores que debe

evitar el traductor y por lo tanto poseer conocimientos suficientes para que no sucedan, tema que es estudiado posteriormente como una condición de control de los procesos de calidad de la traducción. Los principales aspectos a considerar se resumen en los siguientes errores: (1) léxicos (mantener la consistencia del contenido); (2) sintácticos (aunque el orden de las expresiones es diferente al original, mantiene aspectos concordantes); (3) léxico-sintácticos (aceptación inadecuada de oraciones mal construidas); (4) léxico-semánticos (referido a las alteraciones en el sentido de las oraciones); (5) sintáctico-semánticos (alteración del sentido de las oraciones, aunque la traducción sea correcta); y (6) léxico-sintáctico-semánticos (son errores acumulados en un mismo contexto: mala elección de acepciones de términos, sintaxis incorrecta y problemas de sentido en el entendimiento del sentido del texto).

Otra propuesta de evaluación de la calidad de la traducción (sea automática o humana), es la presentada por Aguilar Canal (2019) y Liebling et al. (2022), quienes clasifican los elementos a considerar, tales como: evaluación cuantitativa, calidad de la traducción, dimensión macrotextual, comparación entre el texto original y el traducido, terminología adecuada, así como adecuación pragmática.

Además de los anteriores modelos de medición de la calidad de la traducción, existen dos visiones fundamentales al respecto, una basada en la necesidad de transmitir el verdadero sentido del texto original y la otra, la consideración de utilizar la traducción automática con la tecnología idónea ante la ausencia de conocimiento de otro idioma. Para que una traducción sea considerada de buena calidad, debe transmitir el contenido del texto original desde el punto de vista connotativo como denotativo. Ante esto, muchas veces estas traducciones automáticas terminan siendo aceptables por sus usuarios, quienes en su mayoría no poseen conocimiento de traducción profesional. El grado de aceptación de una traducción se puede valorar según la opinión o la impresión del receptor de la lengua meta (Cossío Montalvo y Rodríguez, 2021).

Si bien, se ha profundizado en los aspectos de calidad en la traducción automática, es importante identificar los elementos a evaluar para definir la calidad de la traducción humana, estas son las siguientes: fidelidad e inteligibilidad, adecuación y fluidez, clasificación de oraciones completas y tiempo de post-edición. En oposición, la traducción automática no es expresiva y menos objetiva, por tanto, se mide: organización de la información lingüística, información morfológica, información sintáctica, información de semántica de oraciones, léxico y morfología (Vardaro et al., 2019).

La conceptualización y evolución de la traducción automática se caracteriza como aquel proceso mediante el cual se utilizan herramientas de software de computadora con lo cual sea posible pasar de forma adecuada del lenguaje natural al estandarizado (Rueda Álvarez y Álvarez Méndez, 2022). Cabe mencionar que la traducción automática es empleada por usuarios en todo el mundo, ya sea porque no dominan un idioma extranjero o precisan de su inmediatez. De la misma forma, se ha visto un creciente uso de esta herramienta en traductores en formación y profesionales. En los últimos años, como solución ante la barrera idiomática, se han implementado estos sistemas de traducción automática en conversaciones en línea (chats) con el fin de establecer contacto con cualquier parte del mundo (Cossío Montalvo y Rodríguez, 2021).

La traducción automática investiga los enfoques para traducir texto de un lenguaje natural a otro. Es un subcampo de la lingüística computacional que extrae ideas de la lingüística, la informática, la teoría de la información, la inteligencia artificial y la estadística. Durante mucho tiempo tuvo mala reputación porque se percibía como de baja calidad. Especialmente, la sociedad ha sido testigo de un gran progreso en la calidad de la traducción automática, lo que la hizo interesante también para su uso en la industria de la traducción. Su calidad es todavía inferior a la traducción humana, pero eso no significa que no tenga buenos usos prácticos (Sepesy Maučec y Donaj, 2020).

Los métodos de traducción automática se podrían emplear de diferentes maneras, por ejemplo, comprender el texto ligeramente, sin importar la redacción, sino que basta la transmisión del mensaje del texto origen y posibilitar la comunicación entre personas de diferentes lenguas. No obstante, en la actualidad hay información publicada día a día destinada a campos específicos, lo cual conlleva a la necesidad de utilizar traductores automáticos por la rapidez y la demanda económica que involucra (Cieza Delgado, 2020).

La traducción automática es posible que suceda bajo diversas dimensiones y propósitos, tales como la traducción automática en sí y la traducción automática estadística. La idea principal de los sistemas de traducción automática basada en ejemplos es recopilar textos y sus traducciones en los que poder basarse para generar sus propias traducciones. El sistema segmenta las oraciones y busca coincidencias en otras traducciones (Mohamed et al., 2021; Rodríguez del Rosario, 2021). Cuando se han extraído las partes útiles, se recombinan y se realizan los cambios necesarios para generar la nueva traducción. Los sistemas de traducción automática estadística necesitan grandes corpus para realizar las traducciones.

Estos sistemas alinean textos y utilizan algoritmos de análisis para calcular cuál es la probabilidad de que un término se traduzca por otro término sin perder el sentido. El sistema *Candide* fue el primer sistema de traducción automática estadística (López Pereira, 2019; Llopis Prats, 2020).

La traducción automática ha demostrado ser de gran utilidad en el campo de investigación; desde años atrás, ha sido utilizada para entender de una manera más rápida y como un apoyo o herramienta en una determinada población, siendo que estos avances demandaron de tiempo y esfuerzo más allá de lo comúnmente requerido (Cieza Delgado, 2020). No obstante, la principal dificultad que enfrentan los sistemas de traducción automática es la ambigüedad léxica (por homografía), por polisemia (que se da cuando una palabra tiene dos o más significados) y estructural (desde un punto de vista gramatical). Varios términos son homógrafos si se escriben igual, pero tienen significados diferentes, mientras que son polisémicos si tienen varios significados relacionados entre sí. El tipo más común es la ambigüedad léxica categorial (Llopis Prats, 2020).

Es posible encontrar varios tipos de interferencia: léxica, sintáctica, cultural y de convenciones genéricas relacionadas con la macroestructura del texto. Las más comunes son la sintáctica y la léxica, especialmente por la constante aparición de neologismos. Por otro lado, los falsos amigos también son característicos de los textos científicos. Son palabras que se escriben de manera muy parecida, pero que tienen significados distintos en dos lenguas (Llopis Prats, 2020).

Al tratarse del uso de máquinas, el traductor automático difícilmente identifica ambigüedades del lenguaje, en cambio, un traductor humano se basa en cinco habilidades cognitivas: origen del texto original, ausencia de comprensión del destinatario final, equivalencias lingüísticas entre el texto original y el traducido, e identificación de aspectos socioculturales (Fernández Torné y Matala, 2021).

Asimismo, la tecnología asociada con la traducción ha dado paso a la creación de herramientas útiles, como la traducción asistida por ordenador y la traducción automática. La traducción automática es una traducción producida con tecnología avanzada, sin intervención de traductores humanos. Por tanto, la traducción automática o denominada *machine translation* en inglés emplea un sistema o software de computadora para realizar la traducción (Cossío Montalvo y Rodríguez, 2021). Además, profesionales y académicos del ámbito de la traducción y de la lingüística computacional consideran el surgimiento de nuevos sistemas de traducción automática, afirmando que

su calidad es tal que suele no diferenciarse de la traducción humana o bien, se comportan al nivel de cualquier traductor promedio. Por ejemplo, en el caso de las noticias, algunos académicos han llegado a considerar que la traducción automática supera a la traducción humana, no obstante, estos sistemas de traducción automática, aunque arrojan productos de calidad, no han seguido la aplicación de medios de evaluación adecuada de la traducción (Briva-Iglesias, 2021).

Es importante tener en consideración dos perspectivas sobre el tema de estudio: la primera, es que los traductores, las empresas de traducción y los servicios de traducción de los gobiernos y las organizaciones internacionales deben ser responsables de la calidad de su producto (Rivera-Trigueros, 2022); la segunda perspectiva radica en que, impulsada por la globalización y los desarrollos de rápido crecimiento de la era digital, ha proliferado la comunicación entre idiomas y culturas a través de la traducción (Asscher y Glikson, 2023).

### 3. Metodología

La delimitación del estudio comprende lo siguiente: (1) el universo de estudio fue el Sistema de Clasificación de Revistas Mexicanas de Ciencia y Tecnología (México), del cual sólo se consideraron las revistas pertenecientes al área de humanidades (27 publicaciones) y se utilizaron los 410 artículos publicados durante el año 2022; y (2) de cada artículo seleccionado sólo se utilizó la parte específica de los metadatos referente al resumen, elemento que al ser traducido se convierte en abstract.

El encuadre metodológico de la investigación se basa en los siguientes criterios: por su enfoque paradigmático se trata de una investigación cuantitativa; por su naturaleza se trata de una investigación no experimental; por su finalidad se trata de una investigación descriptiva; y por su temporalidad es un diseño transeccional.

Para la realización de la presente investigación, se desarrollaron las siguientes actividades metodológicas.

Tomando como referencia la literatura científica se derivó una rúbrica para evaluar las condiciones de la traducción de los abstract de los documentos participantes, la cual se compone de 15 criterios específicos de evaluación para identificación de errores de traducción, agrupados así: sintáctico-semántico (un criterio); léxico-semántico (cinco criterios); morfosintácticos (cuatro criterios); ortotipográficos (dos criterios); y pragmático-culturales (tres criterios).

En cada criterio a evaluar los errores de traducción, se identificaron dos niveles de gravedad:

gravedad baja (con un descuento de -10 puntos de 100) y gravedad alta (con un descuento de -20 puntos de 100), dando al final un puntaje total en contra y como resultado, un puntaje global de evaluación de cada producto (Tabla I). Este criterio, de inicio fue una asignación arbitraria, considerando para ello la experiencia profesional de los investigadores en procesos de evaluación de la calidad de la traducción, tomando en consideración las implicaciones de cada rasgo de traducción y la calidad de los mismos.

<i>Errores de Traducción</i>	<i>Gravedad</i>	
	<i>Baja</i>	<i>Alta</i>
<i>Tipo sintáctico-semántico:</i>		
1. El orden sintáctico altera el sentido	-10	-20
<i>Tipo léxico-semántico:</i>	-10	-20
2. Errores de sentido con respecto al termino	-10	-20
3. Término se mantiene en lengua de origen	-10	-20
4. Se adicionan palabras	-10	-20
5. Se omiten palabras	-10	-20
6. Selección léxica con respecto a los términos polisémicos	-10	-20
<i>Tipo morfosintáctico:</i>	-10	-20
7. Inadecuación en la categoría gramatical	-10	-20
8. Distribución de los componentes de la oración	-10	-20
9. Inadecuación en la concordancia (género, número, persona)	-10	-20
10. Inadecuación en las construcciones -ing del inglés (gerund y present participle)	-10	-20
<i>Tipo ortotipográfico:</i>	-10	-20
11. Uso de mayúsculas y minúsculas	-10	-20
12. Signos de puntuación	-10	-20
<i>Tipo pragmático-cultural:</i>	-10	-20
13. Inadecuación con respecto al registro	-10	-20
14. Inadecuación con respecto a frases idiomáticas, interjecciones y locuciones	-10	-20
15. Inadecuación con respecto a los nombres propios, ciudades, países, festividades	-10	-20
Puntaje total en contra	x /100	

*Tabla I. Rúbrica de evaluación de errores de traducción*

Para ello, se hizo una adaptación a una escala preexistente (UNESCO, 2023) y se adaptó de la siguiente manera en cinco rangos de calificación final por traducción: 0-29 (inaceptable, es una traducción no válida); 30-59 (insuficiente, no alcanza los niveles exigidos); 60-79 (suficiente, se puede utilizar realizando modificaciones de importancia); 80-99 (satisfactorio, se puede utilizar realizando algunas correcciones de poca importancia); y 100 (destacado, se puede utilizar tal cual). Es importante indicar que los resúmenes originales fueron sometidos a metodologías de evaluación para identificar que en su formato original no se utilizaron medios de traducción automática, de lo contrario se descartaron en su participación en el proceso de evaluación. Esta escala ofrece una serie

de perspectivas cualitativas que definen la calidad de la traducción, a lo que fue necesario asignar una calificación cuantitativa para mejor percepción de la condición observada y así optimizar el procesamiento estadístico de los datos.

Una vez integrada la rúbrica, se procedió a solicitar la validación de cinco expertos en traducción. La rúbrica inicial proponía una escala de -1 (gravedad baja) y -2 (gravedad alta). Sin embargo, dos de los evaluadores sugirieron modificarla a -10 y -20 respectivamente por considerar que arrojaba calificaciones objetivas de calidad con 80 puntos como calificación mínima aprobatoria. Respecto a los criterios específicos de evaluación, de inicio eran 13, no obstante, dada la importancia de la dimensión léxico-semántica, se sugirió agregar dos más. Ambas recomendaciones fueron aceptadas y atendidas.

Ante la totalidad de datos recolectados, se procedió a generar una base electrónica con los resúmenes publicados, constituida de los siguientes elementos: (1) nombre de la revista, institución editora de la revista, tipo de institución, resumen original en español, abstract (traducido de forma humana por los autores y aprobado por la propia revista) y su posterior traducción usando tres sistemas de traducción automática de acceso libre (Google, Microsoft y Systran).

Una vez agrupadas las traducciones (una humana y tres automáticas), se procedió a la aplicación de la rúbrica antes descrita, así como su respectiva evaluación, identificando para ello, criterios específicos de error presentes en el texto y asignando una calificación final en una escala de 0 a 100 puntos. Para el análisis estadístico de los datos se utilizó el software estadístico SPSS versión 29.

La reproductividad del estudio se fundamenta en los siguientes análisis estadísticos desarrollados para probar la robustez de la investigación: análisis de contingencias por institución generadora, tipo de institución y pos revista; rangos de error en la distribución de promedios; obtención de medias por sistema de traducción, por criterio específicos y por niveles de gravedad; y análisis de diagramas de cajas. No obstante, deben tenerse en cuenta dos condiciones: (1) la aplicación en otros contextos disciplinares o geográficos, los cuales están determinados específicamente por un nivel de bilingüismo bajo de los investigadores; y (2) los cambios tecnológicos recientes con la inclusión de la inteligencia artificial, los cuales han superado ampliamente cualquier expectativa previa.

Sistema de traducción	Mínimo	Máximo	Media	Desviación estándar
Original	40	100	84.83	13.382
Microsoft	40	100	83.90	12.361
Systran	40	100	83.62	14.051
Google	47	99	80.90	10.960

Tabla II. Comparación de resultados globales por sistema de traducción evaluado

#### 4. Análisis de resultados

De acuerdo a la calificación general obtenida por cada sistema de traducción evaluado, es posible observar que tanto la traducción humana (original) como la traducción automática obtienen un puntaje superior a los 80 puntos para lograr alcanzar la media, que, según la escala de evaluación utilizada en esta investigación adaptada de las propuestas de UNESCO (2023) para asignación de calificaciones cuantitativas y cualitativas sobre el desempeño académico, se agrupan dentro del rango de satisfactorio. De acuerdo a los resultados de las medias estadísticas, la mejor

calificación global se asocia con la traducción humana (original) y la más baja corresponde al traductor automático de Google (Tabla II).

De acuerdo a los cinco rangos probables de calificación final por traductor, la Tabla III permite observar una mayor concentración en los cuatro casos dentro del rango de Satisfactorio seguido del Suficiente. Aunque existen diferencias en la distribución general por escalas de calificación agrupada, se observan solo algunas diferencias notorias a favor de la traducción humana en comparación con la automática. Así también, no se representan ventajas entre un sistema automático en comparación con los demás.

Calificación final agrupada	Traducción humana (original) %	Google %	Microsoft %	Systran %
30-59 Insuficiente	3.7	6.6	5.1	7.6
60-79 Suficiente	19.8	36.1	15.1	14.1
80-99 Satisfactorio	51.7	57.3	68.3	61.5
100 Destacado	24.9	0.0	11.5	16.8

Tabla III. Calificación final agrupada por traductor

Tipo de institución y escala	Traducción original %	Google %	Microsoft %	Systran %	
Universidades públicas federales	30-59 Insuficiente	6.7	6.7	3.3	5.8
	60-79 Suficiente	18.3	30.0	15.8	17.5
	80-99 Satisfactorio	44.2	63.3	60.0	50.8
	100 Destacado	30.8	0.0	20.8	25.8
Centros públicos de investigación	30-59 Insuficiente	1.9	0.0	9.3	11.1
	60-79 Suficiente	16.7	35.2	13.0	18.5
	80-99 Satisfactorio	38.9	64.8	72.2	63.0
	100 Destacado	42.6	0.0	5.6	7.4
Universidades públicas estatales	30-59 Insuficiente	2.0	9.5	4.0	6.0
	60-79 Suficiente	22.1	43.7	15.1	12.1
	80-99 Satisfactorio	57.3	46.7	72.4	68.3
	100 Destacado	18.6	0.0	8.5	13.6
Universidades privadas	30-59 Insuficiente	12.5	0.0	4.0	12.5
	60-79 Suficiente	12.5	12.5	15.1	0.0
	80-99 Satisfactorio	25.0	87.5	72.4	62.5
	100 Destacado	50.0	0.0	8.5	25.0
Universidades públicas estatales con apoyo solidario	30-59 Insuficiente	3.4	0.0	10.3	17.2
	60-79 Suficiente	17.2	17.2	13.8	10.3
	80-99 Satisfactorio	75.9	82.8	72.4	55.2
	100 Destacado	3.4	0.0	3.4	17.2

Tabla IV. Análisis de contingencia entre calificación agrupada y tipo de institución

En relación con la evaluación de medias en los sistemas de traducción por criterios específicos, se presenta un resumen de las medias estadísticas de cada uno de los 15 criterios específicos que se evalúan de los cuatro sistemas de traducción (Tabla V). Para estos casos de análisis por criterios específicos de evaluación, debe indicarse que, a

menor media estadística los resultados son más favorables ya que no se presentaron evidencias de errores de gravedad baja o alta. Esto representa que sólo se mencionarán aquellas incidencias con mayor presencia, consideradas debilidades dentro de cada sistema de traducción.

Criterios de evaluación	Original		Google		Microsoft		Systran	
	Media	Desv.	Media	Desv.	Media	Desv.	Media	Desv.
<i>Tipo sintáctico-semántico</i>								
El orden sintáctico altera el sentido	.68	3.429	.54	3.236	.39	2.770	.59	3.375
<i>Tipo léxico-semántico</i>								
Errores de sentido con respecto al término	1.85	5.504	2.54	6.168	2.29	6.025	3.22	7.088
Término se mantiene en lengua de origen	2.76	6.291	1.76	4.879	1.83	5.263	1.07	3.933
Se adicionan palabras	.63	2.981	.59	2.644	.34	2.398	.24	1.834
Se omiten palabras	.71	3.315	.24	2.084	.32	2.244	.44	2.579
Original Selección léxica con respecto a los términos polisémicos	.59	3.150	.73	3.490	.80	3.645	1.46	5.024
<i>Tipo morfosintáctico</i>								
Inadecuación en la categoría gramatical	.66	2.850	.39	2.177	.39	2.061	.49	2.570
Distribución de los componentes de la oración	.20	1.702	.49	2.664	.22	1.769	.29	2.192
Inadecuación en la concordancia (género, número, persona)	.37	2.236	.17	1.474	.07	.853	.27	2.139
Inadecuación en las construcciones -ing del inglés (gerund y present participle)	.07	1.103	.12	1.303	.05	.988	.05	.988
<i>Tipo ortotipográfico</i>								
Uso de mayúsculas y minúsculas	5.49	5.031	8.32	3.874	8.17	4.233	7.10	4.856
Signos de puntuación	.32	1.889	.46	2.218	.46	2.105	.32	1.889
<i>Tipo pragmático-cultural</i>								
Inadecuación con respecto al registro	.32	1.889	.24	1.834	.27	1.896	.27	1.896
Inadecuación con respecto a frases idiomáticas, interjecciones y locuciones	.07	1.103	.10	1.207	.05	.698	.05	.698
Inadecuación con respecto a los nombres propios, ciudades, países, festividades	.46	2.622	.51	2.794	.49	2.842	.51	2.964

Tabla V. Evaluación de medias de la traducción humana (original) y automática por criterios específicos

Criterio de evaluación	Original		Google		Microsoft		Systran	
	Nº	%	Nº	%	Nº	%	Nº	%
<i>Tipo sintáctico-semántico</i>								
El orden sintáctico altera el sentido	11	2.7	39	9.5	8	2.0	12	2.9
<i>Tipo léxico-semántico</i>								
Errores de sentido con respecto al término	31	7.6	39	9.5	38	9.3	58	14.1
Término se mantiene en lengua de origen	40	9.8	19	4.6	26	6.3	12	2.9
Se adicionan palabras	6	1.5	3	.7	5	1.2	2	.5
Se omiten palabras	9	2.2	4	1.0	4	1.0	5	1.2
Original Selección léxica con respecto a los términos polisémicos	9	2.2	11	2.7	12	2.9	26	6.3
<i>Tipo morfosintáctico</i>								
Inadecuación en la categoría gramatical	4	1.0	2	.5	1	.2	4	1.0
Distribución de los componentes de la oración	2	.5	5	1.2	2	.5	4	1.0
Inadecuación en la concordancia (género, número, persona)	3	.7	1	.2	0	0	4	1.0
Inadecuación en las construcciones -ing del inglés (gerund y present participle)	1	.2	1	.2	1	.2	1	.2
<i>Tipo ortotipográfico</i>								
Uso de mayúsculas y minúsculas	1	.2	2	.5	6	1.5	6	1.5
Signos de puntuación	1	.2	1	.2	0	0	1	.2
<i>Tipo pragmático-cultural</i>								
Inadecuación con respecto al registro	1	.2	2	.5	2	.5	2	.5
Inadecuación con respecto a frases idiomáticas, interjecciones y locuciones	1	.2	1	.2	0	0	0	0
Inadecuación con respecto a los nombres propios, ciudades, países, festividades	5	1.2	6	1.5	7	1.7	5	1.2

Tabla VI. Evaluación de niveles de gravedad alta en traducción original y automática por criterios específicos

Dado que, en los cuatro sistemas de traducción se observa alta coincidencia de medias estadísticas en criterios con mayor incidencia, estos se resumen de la siguiente forma: (1) traducción original (errores de sentido con respecto al término=

1.85, término se mantiene en lengua de origen= 2.76 y uso de mayúsculas y minúsculas= 5.49); (2) Google (errores de sentido con respecto al término= 2.54, término se mantiene en lengua de origen= 1.76 y uso de mayúsculas y minúsculas=

8.32); (3) Microsoft Systran (errores de sentido con respecto al término= 2.29, término se mantiene en lengua de origen= 1.83 y uso de mayúsculas y minúsculas= 8.17); y (4) Systran (errores de sentido con respecto al término= 3.26, término se mantiene en lengua de origen= 1.46 y uso de mayúsculas y minúsculas= 7.10).

Respecto a la evaluación de los niveles de gravedad en los sistemas de traducción por criterios específicos, sólo se consideraron las frecuencias de artículos que observaron condiciones de gravedad alta, esto según los resultados de cada uno de los sistemas de traducción por criterio específico evaluado (Tabla VI). De forma específica se presentan a continuación las situaciones más críticas en situaciones de gravedad alta sobre la base de 410 artículos analizados:

En la traducción original se distinguen los siguientes criterios: (1) errores de sentido con respecto al término (7.6 %); y (2) término se mantiene en lengua de origen (frecuencia: 9.8 %).

Google observa resultados de mayor gravedad en los siguientes criterios de evaluación: (1) Errores de sentido con respecto al término (9.5 %).

Microsoft observa gravedad alta en: (1) errores de sentido con respecto al término (9.3 %); y (2) término se mantiene en lengua de origen (6.3 %).

Systran observa condiciones de alta gravedad en: (1) errores de sentido con respecto al término (14.1 %); y (2) Selección léxica con respecto a los términos polisémicos (6.3 %).

El análisis de contingencia entre calificaciones agrupadas por tipo de traducción y tipo de institución editora de las revistas participantes en el estudio, permitió comparar los resultados de la escala de evaluación de la traducción (niveles de insuficiente, suficiente, satisfactorio y destacado) y las formas de traducción incluidas en la investigación (original, Google, Microsoft y Systran), todo ello en concordancia con las variables del estudio: tipo de institución y revistas.

La importancia de analizar las traducciones según el tipo de institución (1) de donde proviene la edición de las revistas científicas radica en que cada tipo puede mostrar un diferente perfil en cuanto a la importancia que se da o promueve a la comunicación científica. Aun cuando la Tabla IV merece un análisis detallado, sobresalen los siguientes hallazgos: (1) se observan mejores resultados en la traducción original producida de forma humana en los cinco tipos de instituciones en comparación con los tres traductores automáticos; (2) en el nivel más alto de la escala de evaluación (Destacado), la traducción humana supera notablemente a la traducción automática; (3) tanto en la traducción automática como en la

humana, el rango de mayor frecuencia tiende a agruparse en el nivel de Satisfactorio; y (4) los rangos de Insuficiente tienden a mostrar presencia generalizada en todos los tipos de instituciones, no obstante, sus porcentajes en general son bajos.

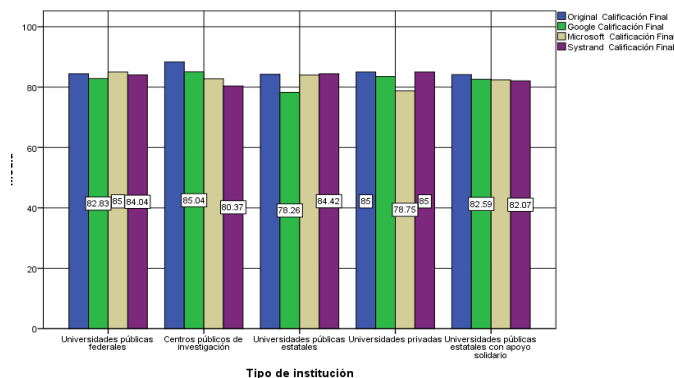


Figura 1. Medias de promedios generales por forma de traducción y tipo de institución

Otra visión del análisis de resultados globales, se presenta como una concentración de datos que evalúan en su conjunto el comportamiento de las formas de traducción estudiadas y el tipo de institución editora de las revistas participantes, en esta ocasión basado en medias de promedios generales. Los resultados expresados en la Figura 1, muestran uniformidad en los niveles de calidad de las distintas alternativas estudiadas y aunque se observen diferencias, tales no son representativas de una superioridad marcada en los promedios generales. No obstante, se puede reconocer que la traducción humana (original), en todos los casos observó ligeras ventajas competitivas en relación con la traducción automática.

Respecto a las medias estadísticas de calificaciones finales por revista evaluada, estas resultaron de la sumatoria de promedios de las evaluaciones de cada uno de los resúmenes. Los resultados generales de las 27 revistas evaluadas son los siguientes: (1) en la traducción original, siete obtuvieron un puntaje de 69-79 (nivel de Suficiente), en tanto, 20 se ubican en un rango de 80-95 (nivel de Satisfactorio); (2) el traductor automático Google arroja que una revista observó una media de 56 (nivel de Insuficiencia), nueve revistas obtuvieron un rango de 65-77 (nivel de Suficiente) y 17 un rango de 80-95 (nivel de Satisfactorio); (3) el traductor automático Microsoft observa que seis revistas ofrecen una media general de evaluación en el nivel de Suficiencia dentro del rango de 73-79 y 21 revistas obtuvieron medias que van de 80-93 puntos bajo el cali-

ficativo de Satisfactorio; y (4) en el traductor automático Systran, seis revistas ofrecen una media general de evaluación en el nivel de Suficiencia dentro del rango de 74-79 y 21 revistas obtuvieron medias que van de 80-90 puntos bajo el calificativo de Satisfactorio.

## 5. Conclusiones

Esta investigación surgió de forma coyuntural ante la amplia divulgación que se ha presentado en torno al uso de la Inteligencia Artificial en diversos contextos, especialmente en el académico y en la comunicación científica. Aunque no se aborda a la Inteligencia Artificial como tema central, si se relaciona con el uso de la tecnología para la solución de prácticas científicas ordinarias que lleven a investigadores no bilingües a solucionar problemas de traducción, al menos en elementos aislados de un documento científico a través de la traducción de metadatos, tales como el título del documento, el resumen y las palabras clave, evitando la dependencia de traductores profesionales, a quienes incluso, no se está en la posibilidad de solventar económicamente tales servicios.

Los resultados obtenidos en esta investigación, aunque descriptivos, analizaron una situación compleja para la comunicación de la ciencia, donde, además de responder al objetivo planteado, surge una amplia variedad de hallazgos de los cuales se podrían obtener diversas inferencias sobre el papel de los actores involucrados en esta clase de procesos. Tales actores son: los propios editores de las revistas al aceptar metadatos traducidos como válidos, los investigadores (autores de los artículos) en la aceptación de sus limitaciones y en la búsqueda de soluciones precisas respecto a la cuestión de traducción, así como las instituciones generadoras de conocimiento quienes buscan posicionar a sus revistas a través de la aplicación de diversos criterios de rigor científico.

Concretamente, es posible llegar a dos conclusiones generales: (1) la traducción humana sigue ofreciendo los mejores parámetros de calidad en la estructura de metadatos en un idioma diferente al original; (2) Las alternativas a sistemas de traducción de libre acceso, aunque con diversos detalles no representativos, solucionan el problema de la traducción de metadatos, siempre y cuando, estén escritos de forma adecuada en su idioma original (español); y (3) pareciera que la situación ideal fuera la traducción mixta entre lo humano y lo automático, especialmente cuando existe una ausencia de traductor profesional y se logre la adquisición de una cultura lingüística genérica del idioma al que se traduce (ortografía,

léxico, estilo de redacción, etc.), esto por parte del autor o de los autores de los textos científicos.

Para responder al objetivo planteado, se afirma que la traducción automática, en términos generales y de acuerdo con los sistemas de traducción participantes (Google, Microsoft y Systran), ofrece soluciones similares a la traducción humana hecha por autores de manuscritos científicos para ser aceptados en su publicación en la revista correspondiente. Tal afirmación, sin embargo, no es generalizable a todos los conceptos, especialmente si se considera que sólo se tocaron dos elementos distintivos: circunscribirse a revistas mexicanas (en otros países pueden suceder cuestiones diferentes por su cercanía o no al idioma inglés) y del área de humanidades (la condición en otras disciplinas científicas puede ofrecer retos diferentes).

Al identificar las diferencias sintáctico-semánticas, léxico-semánticas, morfosintácticas, ortotipográficas y pragmático-culturales entre la traducción humana y automática de metadatos de artículos científicos (resúmenes) de las revistas en cuestión, puede manifestarse, que cada resumen evaluado generalmente no presenta errores de traducción, siendo que, en las condiciones más adversas, únicamente sobresalen por su frecuencia tres criterios con resultados fallidos en la traducción: (1) errores de sentido con respecto al término; (2) término se mantiene en lengua de origen; y (3) uso de mayúsculas y minúsculas.

Respecto a la factibilidad de comprobar la efectividad de los traductores automáticos participantes en el estudio (Google, Microsoft y Systran), en su condición de ofrecer soluciones efectivas en traducción de metadatos de artículos científicos (resúmenes) del área de humanidades como alternativa para investigadores que demandan recurrir a ellos, puede afirmarse que es factible que el investigador recurra a su uso. Aunque se observaron diferencias entre los tres sistemas, su funcionalidad es similar y las diferencias no son representativas para decidirse por alguno de los tres traductores automáticos.

## Notas

- (1) El sistema educativo mexicano de nivel superior es ampliamente clasificatorio en gran número de subsistemas educativos independientes. Para el caso, sólo participan cinco categorías, las cuales coinciden en la generación de publicaciones científicas propias. Cada subsistema observa diversos elementos socio-históricos que los caracteriza. Las universidades públicas federales son las de mayor dimensión poblacional y presupuestal, incluso son las más antiguas; las universidades públicas estatales, en general son de tamaño medio y se ubican en estados fuera de la capital federal y se derivan de las anteriores; los centros públicos de investigación son modelos institucionales alternos tendientes a promover la investigación

por encima de la docencia; las universidades públicas estatales con apoyo solidario, son instancias de reciente creación con financiamiento estatal y federal, tal es el caso de las universidades indígenas y multiculturales (además de otras), mismas que se ubican en regiones que demandan satisfacer necesidades locales de estudiantes con bajas posibilidades de movilidad; las universidades privadas, sin importar su condición o características distintivas, son aquellas instituciones que no recibe financiamiento del gobierno.

## Declaración de autoría

*Javier Tarango*: Conceptualization (lead), Writing – original draft (lead), Writing – review & editing (equal).

*Haydee Menchaca*: Investigation (lead), Data curation (lead).

*Fidel González-Quiñones*: Methodology (lead), Formal analysis (lead), Validation (lead), Writing – review & editing (equal).

*Efraín Alfredo Barragán-Perea*: Writing – review & editing (equal).

## Referencias

- Acosta Plata, Marta (2021). El lenguaje controlado y la traducción automática como vehículos de transmisión del conocimiento científico [Tesis de Máster en Traducción Especializada, Universidad de Vic – Universidad Central de Cataluña].
- Aguilar Canal, Enrique Obed (2019). Evaluación de la calidad del traductor automático DeepL del inglés al español en textos especializados [Tesis, Licenciatura en Traducción e Interpretación, Universidad César Vallejo].
- Asscher, Omri; Glikson, Ella (2023). Human evaluations of machine translation in an ethically charged situation. // *New Media & Society*. 25:5, 1087-1107. <https://doi.org/10.1177/14614448211018833>
- Beltrán-Santoyo, Giovanni; Ruíz-Huerta, Estela Aurora; Gómez-Bernal, Juan Miguel (2021). La importancia e influencia del idioma inglés dentro del campo científico. // *Revista Lengua y Cultura*, 35:5, 46-5. <https://doi.org/10.29057/lc.v3i5.7499>
- Briva-Iglesias, Vicent (2021). Traducción humana vs. traducción automática: análisis contrastivo e implicaciones para la aplicación de la traducción automática en traducción jurídica. // *Mutatis Mutandis*. 14:2, 571-600. <https://doi.org/10.17533/udea.mut.v14n2a14>
- Cieza Delgado; Yessika Yanina (2020). Análisis de errores sintácticos de la traducción automática en textos médicos [Tesis de Licenciatura en Traducción e Interpretación, Universidad César Vallejo].
- Cossío Montalvo, Rebeca Tatiana; Rodríguez, Rosario (2021). Errores de traducción de los chats comerciales del inglés al español en un traductor automático, Lima 2021 [Tesis, Licenciatura en Traducción e Interpretación, Universidad César Vallejo].
- Chatzikoumi, Eirini (2020). How to evaluate machine translation: A review of automated and human metrics. // *Natural Language Engineering*. 26:2, 137-161. <https://doi.org/10.1017/S1351324919000469>
- Fernández Torné, Anna; Matamala, Anna (2021). Human evaluation of three machine translation systems: From quality to attitudes by professional translators. // *Vigo International Journal of Applied Linguistics*. (8, 97-121. <https://doi.org/10.35869/vial.v0i18.3366>
- Ferrer Añó, Núria (2019). Análisis de la utilidad de herramientas de traducción automática en traducción especializada: La traducción inglés-español de sentencias del Tribunal Europeo de Derechos Humanos [Tesis de Máster en Tecnologías de la Información y la Comunicación en la Enseñanza y Tratamiento de Lenguas, UNED].
- Gramigna, Anita (2013). Estética y relación con el pensamiento científico: El papel del lenguaje y el modelo de la investigación contemporánea. // *Thémata: Revista de Filosofía*. (47): 121-137.
- Hidalgo, Sandra Lorena (2009). Lenguaje, ciencia y metafórica de la realidad: Aproximaciones a la estética del lenguaje científico en Nietzsche. // *Visitas Al Patio*. 3, 113–129. <https://doi.org/10.32997/2027-0585-vol.0-num.3-2009-2265>
- Lardera, Marco; Hjørland, Birger (2021). Keyword. // *Knowledge Organization*. 48:6, 430-456. <https://doi.org/10.5771/0943-7444-2021-6-430>
- Liebling, Daniel; Heller, Katherine; Robertson, Samantha; Deng, Wesley (2022). Opportunities for human-centered evaluation of machine translation systems. Findings of the Association for Computational Linguistics: NAACL, 229–241. <https://doi.org/10.18653/v1/2022.findings-naacl.17>
- López Navarro, Irene; Moreno Fernández, Ana I.; Rey Rocha, Jesús (2017). Dificultades de los investigadores españoles para publicar en revistas internacionales: Métricas, editores y multilingüismo. // *Teknokultura*. 14:1, 13-33. <https://doi.org/10.5209/tekn.54142>
- López Pereira, Ariana (2019). Traducción automática neuronal y traducción automática estadística: Percepción y productividad. *Revista Tradumàtica*. // *Tecnologies de la Traducció*. (17): 1-19. <https://doi.org/10.5565/rev/tradumatica.235>
- Llopis Prats, Marta (2020). La traducción automática aplicada a textos científicos: análisis de la traducción de un artículo científico generada por DeepL [Trabajo de Final de grado, Universitat Jaume I]. <https://bit.ly/3NXXU3n>
- Maldonado, María Concepción; Liébana González, María (2021). Los motores de traducción automática y su uso como herramienta lexicográfica en la traducción de unidades léxicas aisladas. // *Círculo de Lingüística Aplicada a la Comunicación*. 88: 189-211. <https://dx.doi.org/10.5209/clac.77002>
- Martínez Mateo, Roberto (2022). Evaluar la calidad de las traducciones profesionales: Propuesta de un modelo mixto. Editorial Comares.
- Mayernik, Matthew S. (2020). Metadata. // *Knowledge Organization* 47:8, 696-713. <https://doi.org/10.5771/0943-7444-2020-8-696>
- Mohamed, Shereen A.; Elsayed, Ashraf A.; Hassan, Y. F.; Abdou, Mohamed A. (2021). Neural machine translation: Past, present, and future. // *Neural Comput & Applic*. 33: 15919–15931. <https://doi.org/10.1007/s00521-021-06268-0>
- Moorkens, Joss (2018). What to expect from Neural Machine Translation: A practical in-class translation evaluation exercise. // *The Interpreter and Translator Trainer*. 12:4, 375-387. <https://doi.org/10.1080/1750399X.2018.1501639>
- Pérez Rodríguez, Vanessa; Huertas Abril, Cristina (2016). La importancia de la calidad en el ámbito de la traducción: estudio comparativo entre traductores autónomos, traductores en plantilla y clientes de servicios de traducción. // *Futhark*. 11, 85-105. <http://doi.org/10.12795/futhark.2016.i11.06>
- Pineda González, Laura Lorena; Tarango, Javier (2018). Barreras lingüísticas entre profesores universitarios. // *Cuadernos Fronterizos*. 43, 37-40. <https://doi.org/10.20983/cuadfront.2018.43.15>
- Ragni, Valentina; Nunes Vieira, Lorena (2021). What has changed with neural machine translation? A critical review of human factors. // *Perspectives*. 30: 137-158. <https://doi.org/10.1080/0907676x.2021.1889005>

- Rivera-Trigueros, Irene (2022). Machine translation systems and quality assessment: a systematic review. // *Lang Resources & Evaluation*. 56: 593–619. <https://doi.org/10.1007/s10579-021-09537-5>
- Roble, Brandon (2022). ¿Cómo se evalúa la calidad de la traducción automática? // *ABC Translations*. <https://bit.ly/3pEnlas>
- Rodríguez del Rosario, Carlota (2021). Creación de motores de traducción automática (estadística y neuronal) inglés-español especializados en el campo de la aviación con la herramienta Mtuoc [Tesis de Máster de Traducción y Tecnologías, Universitat Oberta de Catalunya].
- Rueda Álvarez, Juana del Carmen; Álvarez Méndez, Domingo David (2022). El uso de traductores automáticos para la traducción técnica italiano-español. // *Ciencia Latina: Revista Científica Multidisciplinar*. 6:4, 4642-4657. [https://doi.org/10.37811/cl\\_rcm.v6i4.2963](https://doi.org/10.37811/cl_rcm.v6i4.2963)
- Sepesy Mauček, Mirjam; Donaj, Gregor (2020). Machine Translation and the Evaluation of Its Quality, En Ali Sadollah y T. Tilendra Shishir Sinha, *Recent Trends in Computational Intelligence*. 143-158. IntechOpen. <https://doi.org/10.5772/intechopen.89063>
- Ulitkin, Ilya; Filippova, Irina; Ivanova, Natalia; Poroykov, Alexey (2021). Automatic evaluation of the quality of machine translation of a scientific text: The results of a five-year-long experiment. // *E3S Web Conf*. 284: 1-12. <https://doi.org/10.1051/e3sconf/202128408001>
- Unesco (2023). Sistema de Información de Tendencias Educativas en América Latina: Observatorio regional de políticas de educación. <https://siteal.iiep.unesco.org/>
- Vardaro, Jennifer; Schaeffer, Moritz; Hansen-Schirra, Silvia (2019). Translation Quality and Error Recognition in Professional Neural Machine Translation Post-Editing. // *Informatics*. 6:3, 41. <https://doi.org/10.3390/informatics6030041>
- Vicente Cubero, Santiago (2019). El inglés como lengua vehicular para los experimentos científicos en el aula de primaria [Tesis de grado en Educación Primaria, Universidad de Valladolid].
- Viver Sorolla, Paula (2018). La evaluación de las herramientas de traducción automática (TA) desde la perspectiva del traductor: Google Translate, Bing, Babylon y Systran [Tesis de pregrado, Universidad de Valladolid]. <https://bit.ly/42Qprbk>
- Wang, Haifeng; Wu, Hua; He, Zhongjun, He; Huang, Liang; Ward Church, Kenneth Ward (2021). Progress in Machine Translation. // *Engineering*. (18): 143-153. <https://doi.org/10.1016/j.eng.2021.03.023>

---

Enviado: 2025-07-28. Segunda versión: 2025-12-23.  
Aceptado: 2025-12-23.

---