

Universidad Femenina del Sagrado Corazón
Facultad de Traducción, Interpretación y Ciencias de la
Comunicación
Escuela Profesional de Traducción e Interpretación

ANÁLISIS DE CALIDAD DE TRADUCCIÓN DE DOS TRADUCTORES
AUTOMÁTICOS UTILIZADOS EN UN ESTUDIO DE TRADUCCIÓN
EN LIMA, 2023

Tesis presentada por:

ROXANA JAZMIN BARRANTES ROJAS

Para obtener el Título Profesional de Licenciada en Traducción e
Interpretación

Línea de Investigación: Transformación digital en la traducción e Interpretación

Asesor

Dr. Jesús Fernando Cornejo Sánchez
<https://orcid.org/0000-0003-3468-8854>

Lima – Perú

2023

Los miembros del jurado han aprobado el estilo y el contenido de la tesis sustentada por:

ROXANA JAZMIN BARRANTES ROJAS

JESÚS FERNANDO CORNEJO SÁNCHEZ

Asesor: Nombre(s) y Apellidos

ROSSANA MARÍA SORIANO VERGARA

Presidente de Jurado: Nombre(s) y Apellidos

DANIEL SANTOS SILVA NIEVES

Miembro de Jurado 1: Nombre(s) y Apellidos

VIVIAN KATIANA MORALES SÁNCHEZ

Miembro de Jurado 2: Nombre(s) y Apellidos

JESÚS FERNANDO CORNEJO SÁNCHEZ

Miembro de Jurado 2: Nombre(s) y Apellidos



Mg. Rossana Soriano Vergara
Decana de la Facultad de Traducción,
Interpretación y Ciencias de la
Comunicación

Declaratoria de Originalidad del Asesor

Facultad o Escuela de Posgrado:	Traducción, Interpretación y Ciencias de la Comunicación
Escuela Profesional o Programa Académico:	Traducción e Interpretación
Dependencia a la que pertenece el docente asesor:	Departamento Académico de Idiomas y Ciencias de la Comunicación
Docente asesor que verifica la originalidad:	Dr. Jesús Fernando Cornejo Sánchez
ORCID:	0000-0003-3468-8854
Título del documento:	Análisis de la calidad de la Traducción de dos traductores automáticos utilizados en un estudio de Traducción en Lima, 2023
Autora(s) del documento:	Roxana Jazmin Barrantes Rojas
Mecanismo utilizado para detección de originalidad:	Software Turnitin
Nº de trabajo en Turnitin: (10 dígitos)	2408538382
Porcentaje de <u>similitud</u> detectado:	17 %
Fuentes originales de las similitudes detectadas:	Fuentes de internet = 17 % Publicaciones = 1 % Trabajos del estudiante = 5 %

(imagen del % similitud y fuentes detectadas)

tesis

INFORME DE ORIGINALIDAD



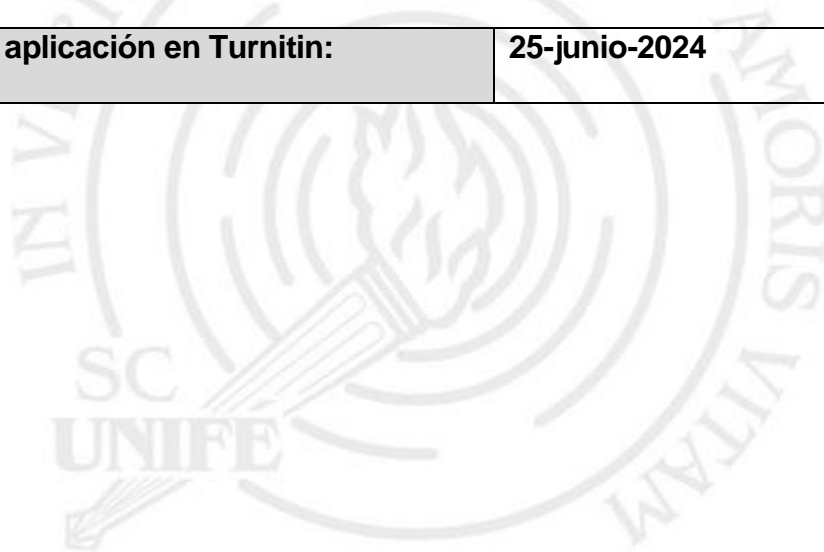
FUENTES PRIMARIAS

1	hdl.handle.net Fuente de Internet	4%
2	Submitted to Universidad Femenina del Sagrado Corazón Trabajo del estudiante	3%
3	repositorio.ucv.edu.pe Fuente de Internet	2%
	unc.aws.openrepository.com	2%

El docente asesor declara ha revisado el informe de similitud y expresa que el porcentaje señalado cumple con las “Normas Internas de Investigación e Innovación” establecidas por la Universidad Femenina del Sagrado Corazón.

Fecha de aplicación en Turnitin:

25-junio-2024



RESUMEN

Esta investigación se centró en evaluar la calidad de la traducción automática mediante el análisis de dos traductores automáticos, utilizando la métrica de evaluación SAE J2450, previamente establecida. Se optó por un enfoque cuantitativo, de nivel exploratorio y se llevó a cabo mediante un diseño de estudio transeccional exploratorio en Lima, Perú. La muestra consistió en dos motores de traducción automática, DeepL y Yandex, a los cuales se les sometió un fragmento de un documento técnico del campo de la automoción. Luego de evaluar ambos traductores automáticos, se determinó que la calidad promedio de sus traducciones es medianamente baja, con un puntaje de 74.50 %. Esto se atribuyó en gran medida a la presencia significativa de errores terminológicos y sintácticos. Asimismo, el resultado de calidad de DeepL fue de un 72.39 %, mientras que Yandex alcanzó un 76.62 %. Ambos casos revelaron una calidad de traducción medianamente baja, con errores de terminología como los más frecuentes. Por último, para mejorar la calidad en las traducciones se recomienda la intervención de un traductor profesional que pueda corregir y optimizar los posibles errores identificados durante el proceso de traducción automática.

Palabras clave: traducción automática, calidad, error de traducción, SAE J2450, traductor profesional

ABSTRACT

This study focused on evaluating the quality of machine translation by analyzing two machine translators using the previously established SAE J2450 evaluation metric. The study chose a quantitative, exploratory approach and was performed using an exploratory cross-sectional study design in Lima, Peru. The sample included two machine translation programs, DeepL and Yandex, and a fragment of a technical document from the automotive field was submitted on these platforms. After the evaluation of both machine translators, the average quality of their translations was found to be moderately low, with a score of 74.50 %. This was largely due to the significant presence of terminological and syntactic errors. Likewise, the quality score for DeepL was 72.39 %, while Yandex reached 76.62 %. In both cases, the translation quality was moderately low, with terminology errors being the most frequent. Finally, to improve the quality in translations, the intervention of a professional translator who can correct and enhance possible errors identified during the machine translation process is recommended.

Keywords: machine translation, quality, translation error, SAE J2450, professional translator

Agradezco a Dios por siempre estar presente y por ser mi fuente de confianza en el camino hacia mis metas.

A mis padres, por su apoyo incondicional, por hacer posible que llegue hasta aquí. Su orgullo en mí es mi mayor logro.

A mí misma, por persistir en mis sueños y nunca dejar de creer en mí. Reconozco mi singularidad y valoro mi determinación.

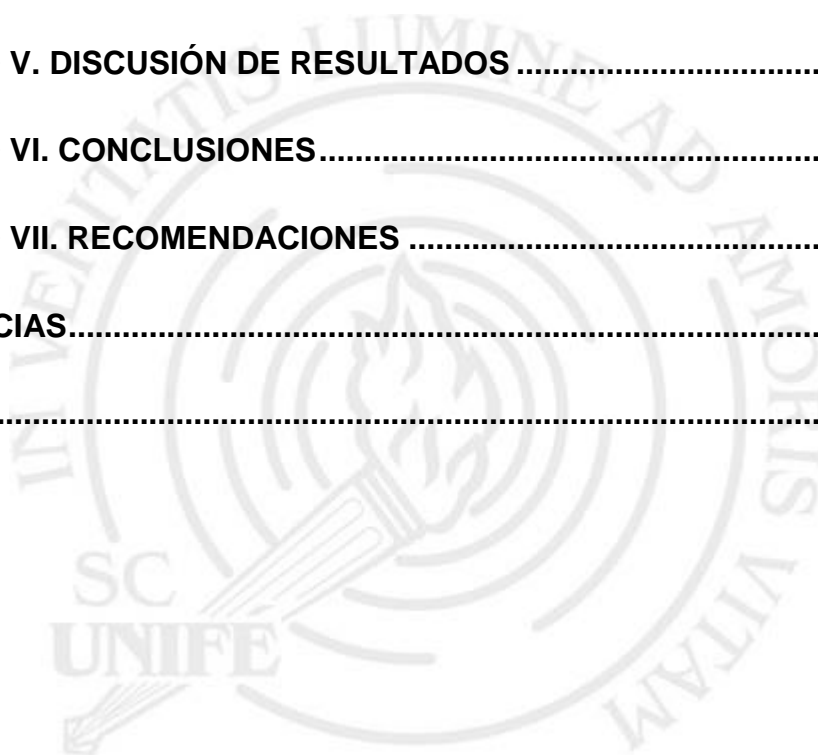
A todas las personas que formaron parte de mi camino, por escucharme y brindarme el apoyo necesario para alcanzar mis objetivos.

Y a mi asesor, Dr. Jesús Cornejo, por su guía en la elaboración de esta investigación y por simplificar lo que parecía difícil.

ÍNDICE

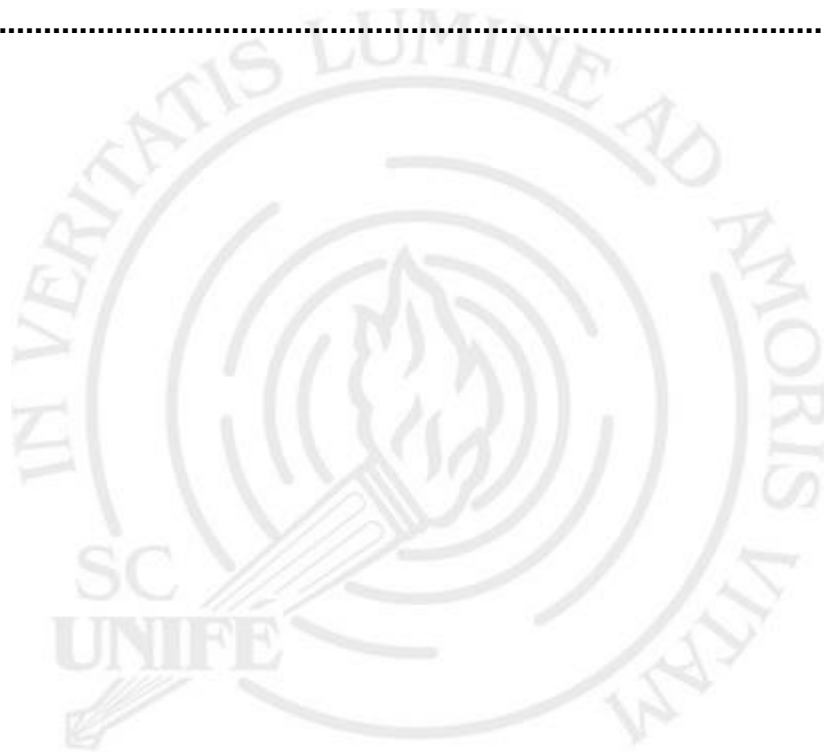
RESUMEN	5
LISTA DE TABLAS	9
LISTA DE FIGURAS.....	10
CAPÍTULO I. PROBLEMA DE INVESTIGACIÓN	13
1.1. Planteamiento del problema:.....	13
1.2. Justificación	16
1.3. Delimitación y limitación de la investigación	17
1.3.1. Delimitación	17
1.3.2. Limitación	18
1.4. Objetivos de la investigación	18
1.4.1. Objetivo general	18
1.4.2. Objetivos específicos.....	18
CAPÍTULO II. MARCO TEÓRICO	19
2.1. Antecedentes:.....	19
2.2. Bases teóricas	26
2.2.1. La traducción.....	26
2.2.2. La traducción automática	28
2.2.3. Tipos de traducción automática.....	30
2.2.4. Calidad de traducción	32
2.2.5. Calidad de la traducción neuronal y la traducción híbrida	35
2.2.6. Evaluación de calidad de traducción	36
2.2.7. La métrica SAE J2450	38

CAPÍTULO III. MÉTODO	44
3.1. Nivel, tipo y diseño de la investigación	44
3.2. Operacionalización de variables.....	46
3.3. Muestra	50
3.4. Técnica e instrumento de recolección de datos.....	51
3.5. Técnicas de procesamiento y análisis de datos	55
3.6. Método de análisis.....	56
CAPÍTULO IV. RESULTADOS.....	57
CAPÍTULO V. DISCUSIÓN DE RESULTADOS	63
CAPÍTULO VI. CONCLUSIONES.....	68
CAPÍTULO VII. RECOMENDACIONES	70
REFERENCIAS.....	72
ANEXOS	79



LISTA DE TABLAS

TABLA 1	46
TABLA 2	53
TABLA 3	54
TABLA 4	57
TABLA 5	59
TABLA 6	61



LISTA DE FIGURAS

FIGURA 1.....	58
FIGURA 2.....	60
FIGURA 3.....	62



INTRODUCCIÓN

En los últimos tiempos, la traducción automática ha ganado una fuerte presencia en el ámbito de la traducción profesional. A pesar de su prominencia, esta no ha sido acompañada de una debida evaluación de calidad para probar su capacidad para ofrecer un nivel de calidad consistente a los profesionales del sector. Esta falta de conocimiento se refleja en la escasa comprensión sobre la contribución o perjuicio que los sistemas de traducción automática podrían aportar a las traducciones profesionales, especialmente en el campo de la traducción especializada.

En el primer capítulo de este trabajo se expone la problemática actual que motiva esta investigación. En este contexto, se presenta la justificación del estudio, junto con el objetivo general y los objetivos específicos que orientan la investigación.

En el segundo capítulo se presentan los estudios y trabajos que sirvieron de antecedentes para la investigación. Además, se menciona algunas bases teóricas que hablan de la traducción automática y la calidad que estos pueden proporcionar. Posteriormente se presenta más conceptos que ayudan a la comprensión de este estudio.

El tercer capítulo se enfoca en la metodología del estudio, se detalla el método, el enfoque, el nivel y el tipo, así como también el instrumento de análisis utilizado para la recolección de datos, el proceso de análisis y procesamiento de la información.

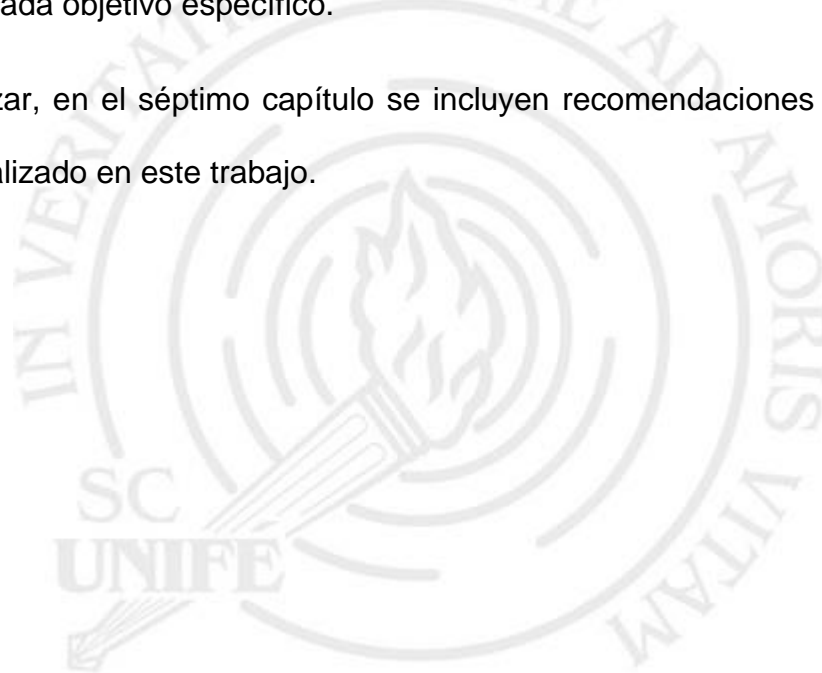
En el cuarto capítulo se detallan los resultados, partiendo del objetivo general hasta los objetivos específicos. Estos resultados derivan del análisis de las fichas correspondientes. En términos generales se encontró que ambos traductores

alcanzaron un nivel promedio de calidad del 74.50 %. Para los objetivos específicos, se observó que DeepL tuvo un puntaje del 72.39 % y Yandex un puntaje del 76.62 %.

En el quinto capítulo, se aborda la discusión de los resultados estableciendo su correlación con las bases teóricas de la investigación. Asimismo, se lleva a cabo una comparación a nivel metodológico entre los resultados obtenidos en este estudio y los de investigaciones previas consideradas como antecedentes.

En el sexto capítulo del estudio, se presenta las conclusiones tanto a nivel general como por cada objetivo específico.

Para finalizar, en el séptimo capítulo se incluyen recomendaciones derivadas del análisis realizado en este trabajo.



CAPÍTULO I. PROBLEMA DE INVESTIGACIÓN

En el siguiente capítulo se abordan cuatro puntos importantes; el planteamiento del problema, la justificación de la investigación, la delimitación y limitación; y por último se presentan los objetivos, tanto el general como los específicos.

1.1. Planteamiento del problema:

Hoy en día, en diversos ámbitos, se tienen en cuenta varios aspectos para cumplir ciertas necesidades del usuario o cliente. Según lo establecido en la norma ISO 9000 (2005) de la Organización Internacional de Normalización, se debe mencionar el aspecto de la calidad como toda característica inherente de todo producto o servicio que permite cumplir con aquellas necesidades que se requieren. Dado esto, si se quiere hablar del cumplimiento de necesidades, es importante resaltar el grado de calidad de aquel producto o servicio brindado. Además, para lograr la calidad se deben cumplir los requisitos demandados por el cliente, en los cuales se debe priorizar la eficacia de la realización del objetivo (Valls et al., 2017).

Si bien la traducción más que un producto es un servicio el cual se brinda a aquellos que la necesiten, desde hace algunos años se ha hecho uso de diversas herramientas tecnológicas como los diccionarios en línea, las memorias de traducción o incluso traductores automáticos y son de estos últimos que se tiene un mayor interés. A nivel internacional, desde hace décadas se han implementado diversos métodos para poner en evaluación los traductores automáticos y de esta manera comprobar su calidad. Sin embargo, todavía no se está cien por ciento de

acuerdo sobre la definición y los parámetros en relación con la calidad de traducción, la cual puede ser mejor para la industria de la traducción (Koby et al., 2014).

La traducción automática tiene como base principal el poder reformular automáticamente las palabras de un texto en una lengua origen por palabras de otra lengua, la lengua meta, con el fin que se mantenga el sentido general de la oración (Koskenniemi et al., 2012). Para Alawneh y Sembok (2011) es importante incluir a la traducción automática en la lingüística computacional, ya que este tipo de traducción utiliza softwares para poder traducir un texto de una lengua a otra.

En el Perú, no se cuenta con la información adecuada con respecto al tema de la calidad de la traducción automática, considerando la implementación de sistemas de evaluación de traducción. Este problema lleva al traductor a la incertidumbre de no saber con claridad si el uso de traductores automáticos sería correcto o no durante el proceso de la traducción. En el código de ética del Colegio de Traductores del Perú (CTP) se indica que la calidad debe ser el “principal atributo del servicio ofrecido” (2017, p.3). Esta corta definición solo abarca el tema de la calidad desde un punto de vista general y deja a una libre interpretación lo que un traductor profesional debe brindar como servicio y producto al cliente. Por otro lado, si bien las tecnologías de traducción todavía siguen siendo consideradas como un avance en progreso, es importante resaltar que la traducción automática tiene una gran trayectoria de investigación. Desde 1966 cuando se presentó el informe del *Automatic Language Processing Advisory Committee* (ALPAC) ya se veía la idea de plantear un método de evaluación para los traductores automáticos. En la

actualidad, España es uno de los países que más tiene presente el avance de la traducción automática, tomado en cuenta los diversos motores de traducción automática y por ende el interés de la calidad de estos motores. Todo esto es un área que siempre está en cuestión e investigación. Asimismo, en el Perú, el tema de la calidad parece estar desapercibido o simplemente se entiende que está inmersa en el trabajo del traductor. Esto ocasiona que muchas veces el traductor solo realiza sus proyectos sin enfocarse en el proceso de calidad.

Hasta el día de hoy, se cuenta con diversos trabajos de investigación que tienen en cuenta el planteamiento de métodos de evaluación. En el año 1993, el Grupo Asesor de Expertos en Estándares de Ingeniería del Lenguaje (EAGLES, por sus siglas en inglés) inició un proyecto tomando en cuenta las normas ISO 9126, el propósito era lograr desarrollar un marco general como base para iniciar el rubro de la evaluación de las tecnologías de la lengua. Tiempo después gracias al proyecto anterior se iniciaron nuevos proyectos como el *International Standards for Language Engineering (ISLE)* en el año 2000, proyecto en colaboración entre Europa y Estados Unidos. Dentro de este proyecto es crucial señalar que tuvo tres grupos de trabajo y en el caso del Grupo de Trabajo de Evaluación (EGW, por sus siglas en inglés) se dedicó a la evaluación de la traducción automática (Calzolari et al., 2003).

Hay investigaciones como es el caso del Marco de Referencia para la Evaluación Automática (FEMTI, por sus siglas en inglés), el cual fue resultado del proyecto de ISLE y tiene como objetivo desarrollar un conjunto de métodos de manera estructurada para evaluar sistemas de traducción automática (Calzolari et al., 2003; Quah, 2006). En el 2013 con un enfoque más cercano a la posesión, el

centro de investigaciones TAUS planteó unas directrices para poder evaluar la calidad de traducciones automáticas.

En el presente proyecto, lo que se busca indagar es acerca de la calidad de las traducciones obtenidas de traductores automáticos. Este tema es muy poco conocido y se cuenta con pocos trabajos de investigación a nivel nacional. Debido a esto, con este estudio se espera contribuir en el conocimiento de la calidad en la traducción, además de poner en conocimiento a los profesionales del campo de la traducción de los nuevos avances en los traductores automáticos.

Entonces, conociendo todo lo anterior, este proyecto de investigación busca responder la siguiente interrogante como problema general, ¿Cuál es la calidad de dos traductores automáticos utilizados en un estudio de traducción en Lima, 2023?

1.2. Justificación

A continuación, se presenta la justificación de aspectos teórico, práctico y metodológico de la presente investigación:

La justificación teórica se fundamenta en la importancia de proponer nuevos enfoques con respecto a la calidad de la traducción obtenida por sistemas de traducción automática. Asimismo, esta investigación se apoya en los aportes de Koby et al. (2014) quienes mencionan que una traducción de alta calidad implica transmitir completamente el mensaje presente en el texto, incluyendo aspectos como la denotación, la connotación, el matiz y el estilo. El objetivo es reformular el texto de manera que se pueda leer como si hubiera sido creado alguien que habla el idioma nativo. Además, cuando se trata de traducción automática, aquí se exige

un nivel suficiente de posesión para garantizar unos niveles mínimos de precisión y fluidez, logrando así una buena calidad en la traducción.

De la misma manera, la presente investigación cuenta con una justificación práctica, ya que con este trabajo se espera beneficiar a futuros traductores, profesionales relacionados al rubro de la traducción y público interesado en el tema. También, se espera aportar nuevos conocimientos en el campo de la evaluación de calidad con un enfoque a los traductores automáticos, dado que no se cuenta con muchos estudios sobre el tema.

Por último, la contribución metodológica se justifica mediante el uso del método científico. Este estudio adoptó un enfoque cuantitativo con un nivel exploratorio y se diseñó como una investigación transeccional exploratoria. Con el fin de analizar la calidad de traductores automáticos, se aplicó una prueba de evaluación, la cual respalda la credibilidad y la validez de los hallazgos en los estudios (Creswell y Creswell, 2018). El instrumento utilizado fue la métrica de evaluación SAE J2450, que permite evaluar la calidad de la traducción independientemente del idioma de partida, el idioma destino o si la traducción fue una traducción humana o automática.

1.3. Delimitación y limitación de la investigación

1.3.1. Delimitación

La investigación actual, enfocada en evaluar el nivel de calidad de traducción proporcionado por dos motores de traducción automática, se realizó durante el año 2023, en Lima, Perú, como escenario principal. En términos de su delimitación

teórica, este estudio se fundamenta en diversos conceptos relacionados con la calidad de la traducción, centrándose específicamente en la evaluación de los traductores automáticos neuronales y híbridos.

1.3.2. Limitación

En cuanto a las limitaciones identificadas, es importante mencionar el uso de una métrica de evaluación de traducción poco conocida, así como la falta de una bibliografía extensa que aborde la evaluación de calidad de los diversos tipos de traductores automáticos. Además, se debe considerar el impacto de las actualizaciones periódicas de las plataformas de traducción automática, las cuales pueden variar los resultados de la traducción a lo largo del tiempo.

1.4. Objetivos de la investigación

1.4.1. Objetivo general

Evaluar la calidad de dos traductores automáticos utilizados en un estudio de traducción en Lima, 2023.

1.4.2. Objetivos específicos

Analizar la calidad de la traducción automática neuronal del traductor DeepL utilizado en un estudio de traducción en Lima, 2023.

Analizar la calidad de la traducción automática híbrida estadística-neuronal del traductor Yandex utilizado en un estudio de traducción en Lima, 2023.

CAPÍTULO II. MARCO TEÓRICO

En este capítulo, se exponen los antecedentes que respaldan la investigación, así también se desarrollan temas fundamentales que forman parte de las bases teóricas, tales como la traducción automática, la calidad y las métricas de evaluación. Posteriormente, se incluyen algunos términos clave para facilitar la comprensión del presente trabajo.

2.1. Antecedentes

En las siguientes líneas, se explican los antecedentes organizados cronológicamente hasta los más actuales:

Adán (2019) en su investigación, tuvo como objetivo comparar tres traductores automáticos (DeepL, Yandex y Apertium) de diferentes tipos con el propósito de encontrar el óptimo para el uso en textos técnicos. El estudio empleó un enfoque cualitativo, con un nivel descriptivo y se decidió usar una ficha de análisis como instrumento de investigación. Este trabajo contó con un corpus de dos artículos científicos que contenían un lenguaje técnico, los cuales se tradujeron mediante los tres traductores automáticos. Para el análisis de los errores, se utilizó la clasificación propuesta por Waddington, quien propone tres categorías, como la deontología, la expresión y el significado. Luego de obtener las traducciones, se realizó el análisis de errores y la suma de la cantidad total de errores de cada tipo de traductor. Los resultados mostraron que DeepL tuvo un total de 22 errores, Yandex tuvo 39 errores y Apertium tuvo 75 errores. Tanto DeepL como Yandex tuvieron una mayor presencia de errores de terminología e idiomática. En el caso

de Apertium, se obtuvo más errores de sinsentidos y sintaxis. A modo de conclusión, se mencionó que en los tres TA existe una mayor cantidad de errores léxicos y semánticos. En cuanto a la mejor opción de TA, se decidió por DeepL dado que se obtuvo menos errores.

En la investigación de Aguilar (2019) se tuvo como objetivo analizar la calidad del traductor automático DeepL en textos especializados con la combinación de idiomas inglés-español. La investigación tuvo un enfoque mixto, de tipo básico y nivel descriptivo. Se realizó un estudio de caso y también se utilizó la técnica de análisis de contenido. La unidad de análisis utilizada fue la calidad de traducción de un traductor automático, la muestra consistió en 5 tipos de documentos de diferentes áreas temáticas como aeronáutica, arquitectura, agricultura, química y minería. La evaluación fue analizada haciendo uso de una ficha de análisis para la recolección de datos y se basó en la propuesta de evaluación y clasificación del trabajo de Martínez (2014). Dicha propuesta señala una serie de errores al igual que una puntuación para cada texto traducido. Luego de un análisis micro y macrotectual, se concluyó que la calidad del traductor DeepL es aceptable, pero no estable. En el proceso de evaluación, se obtuvo calificaciones diferentes, buenas y malas, por parte de los evaluadores (traductores y revisores). Por último, se mencionó que es necesaria la presencia de un traductor humano para la revisión del texto traducido, y que la calidad del traductor depende de la carga terminológica de los documentos al igual que la temática de estos.

Esteban (2020) tuvo como objetivo en su investigación comprobar la calidad de los motores de traducción automática neuronal (Google Translate y Microsoft

Translator) en textos del ámbito jurídico con la combinación de idiomas inglés-catalán para su posterior posesión. El trabajo tuvo un enfoque mixto, ya que se realizaron diversas pruebas para catalogar y analizar la calidad brindada por los motores de traducción. La evaluación tuvo una muestra de diferentes fragmentos de textos (corpus), además de ser un análisis comparativo, se utilizaron dos principales métodos de evaluación, una evaluación automática mediante métricas específicas y una evaluación humana realizada por traductores profesionales. Se contrastó ambos resultados para una mejor valoración. Luego del trabajo de investigación y evaluación, se concluyó que, en relación con la comparación de los dos motores de traducción, la diferencia es relativamente baja. Con ambos motores, se generaron traducciones cohesionadas y coherentes, se pudo entender la intención del texto y por lo general se respetó el tono formal que estaba presente en el texto origen. De igual manera, en la elección de terminología de ambos motores de traducción no hubo un mayor error, se tradujo en la mayoría de los casos de manera correcta. Asimismo, se presentaron algunos errores en las estructuras gramaticales, en el estilo de la traducción y por último en ambos motores se presentaron problemas en el uso de mayúsculas, siendo Microsoft Translator quien tuvo una mayor cantidad de errores de este tipo.

Llopis (2020) tuvo como objetivo analizar la traducción automática de un artículo científico y a su vez identificar los distintos tipos de errores. La investigación tuvo un enfoque cualitativo y como parte de su metodología se utilizó un corpus de 1956 palabras de un artículo científico. Se escogió DeepL como traductor automático y también el análisis de los errores se basó en la métrica MQM

(*Multidimensional Quality Metrics*). Después del análisis de cada tipo de error, se dio como resultado que la traducción obtenida presentó errores en cada tipo, pero también tuvo algunos aciertos frente a problemas de sentido. A modo de conclusión, se detalló que la traducción ofrecida por DeepL solo permitió entender las ideas principales del texto, ya que los errores fueron bastante notorios ante la evaluación del autor. El error de terminología estuvo más presente en la traducción y, por lo contrario, el error con relación a la convención local no afectó mucho a la traducción.

Bermúdez (2020) tuvo como objetivo analizar la traducción automática de sitios web médicos en la combinación de idiomas español- inglés y también español-francés además de brindar una propuesta de posesición. La investigación tuvo un enfoque cuantitativo, fue un estudio de caso y tuvo una metodología inductiva y descriptiva. La evaluación fue realizada solo por un traductor y para todo el proceso se seleccionó el motor de traducción automática Google Translate y el texto elegido fue una colección de diferentes secciones de una página web. Para evaluar la traducción primero se tomó en cuenta los aspectos de la fluidez y la adecuación, para esto se utilizó un baremo con el cual se dio un puntaje de uno (1) a cinco (5). Después de este primer paso, se decidió categorizar los errores en base a la métrica MQM (*Multidimensional Quality Metrics*). En los resultados de la investigación se mencionó que, en base al análisis de adecuación y fluidez, se pudo considerar con una calidad aceptable a la traducción de Google Translate. Con respecto a los errores, en ambas combinaciones de idiomas, el tipo de error que tuvo un mayor porcentaje fue el *Mistranslation* con un 30 % (inglés-español) y un 39 % (francés-español) y los errores de menos porcentaje fueron de *Addition* y *Typography*. Como

conclusión, se mencionó que los traductores automáticos son solo una de las muchas herramientas para el traductor y que no pueden ser vistos como método único para todo el proceso traductor.

Maldonado y Liébana (2021) presentaron como objetivo en su artículo científico señalar algunas limitaciones lingüísticas que tienen diferentes traductores automáticos. Como parte de la metodología, se evaluó desde un enfoque práctico, el comportamiento de tres traductores automáticos (DeepL, Google Translate y Bing Microsoft Translator) frente a la traducción de unidades léxicas aisladas. Además, se identificó los problemas más comunes en los traductores automáticos (errores léxicos, sintácticos, semánticos, ortográficos, etc.). En los resultados, se señaló que los TA presentaron carencias al traducir unidades léxicas, dado que muchos de los TA solo muestran las mejores opciones en base de sus corpus de textos paralelos. En las conclusiones, se detalló que los avances tecnológicos de informática benefician mucho a los traductores, herramienta para buscar equivalentes de unidades léxicas.

En su estudio de investigación, Bolaños y Ramos (2021) se propusieron analizar los errores de traducción automática de una página web de noticias (inglés-español). La investigación tuvo un enfoque cualitativo, descriptivo y con un diseño de estudio de caso. La evaluación de errores se basó en la clasificación dada por Allen. El corpus del estudio fue la página web "TMZ", se extrajo una muestra de 25 noticias, las cuales se analizaron haciendo uso de una ficha de análisis. Después del proceso de evaluación, se obtuvo un total de 148 errores por parte de la traducción automática. De este total, los errores de tipo gramatical y/o sintáctico

fueron los más frecuentes (60 errores), luego los errores de tipo terminológico (51 errores), los errores de transferencia (29 errores), los errores mecánicos (4 errores), los errores en base al tiempo verbal (3 errores) y, por último, los errores de puntuación (1 error). A modo de conclusión, la investigación destacó que los errores más comunes en los textos de noticias son los relacionados con el sentido, la terminología y la concordancia.

Otro estudio relevante para esta investigación es el de Rodríguez (2021), quien se propuso detallar como se desarrollaron dos sistemas de traducción automática, uno basado en las estadísticas y otro en redes neuronales, utilizando la herramienta MTUOC, un proyecto de investigación y transferencia tecnológica llevado a cabo en la *Universitat Oberta de Catalunya*. La metodología empleada fue cualitativa, se siguió el diseño y creación de un corpus paralelo inglés-español como primer paso. Se realizó un análisis comparativo entre los resultados de los textos meta generados por los dos sistemas creados y dos traductores, Google Translate y Yandex. La evaluación mediante el sistema BLEU indicó que, en todos los casos, se logró una traducción general satisfactoria debido al uso de diferentes corpus en el entrenamiento. En el caso de la traducción automática estadística, el sistema MTUOC-Translator obtuvo un puntaje de 48.50 y Yandex un puntaje de 69.79, en ambos casos la traducción se considera buena y fluida. Sin embargo, para el caso de Yandex, los errores principales se enfocan en la terminología especializada. Con respecto a los resultados de la traducción automática neuronal, el sistema MTUOC-Translator obtuvo un 57.82 y Google Translate un 69.39. En el caso de Google Translate, se presentaron errores de elección en los tiempos verbales y algunos

errores de terminología especializada, En conclusión, entre los dos motores de traducción elaborados, el sistema de traducción automática neuronal alcanzó mejores resultados de traducción.

En el estudio de Rubio (2022), el objetivo principal analizar la traducción de una serie audiovisual mediante la combinación de la traducción automática y la posesición. Para este trabajo, se tuvo un enfoque cualitativo, teniendo como corpus el primer episodio de una serie audiovisual. La evaluación de calidad se basó en el análisis de la subtitulación en español y posteriormente se clasificó los errores según el marco *Multidimensional Quality Metrics* (MQM), que incluye la precisión, fluidez, convención local y estilo. Después de todo el proceso de análisis y evaluación se obtuvo como resultados que el traductor automático junto con la posesición fue insuficiente y generó un texto de baja calidad según los estándares aceptados. Se observó una gran incidencia de errores menores, errores mayores, errores nulos y errores críticos, aunque estos últimos fueron menos frecuentes, su impacto fue significativo. En conclusión, el estudio menciona que, si bien la traducción pudo resultar aceptable para el consumidor, no alcanzó los estándares mínimos requeridos de calidad.

Lai (2023) presentó como objetivo analizar las características de los errores de la traducción automática neuronal de un artículo periodístico en los idiomas español-chino. El estudio adoptó un enfoque cualitativo y utilizó la lista de errores del TAUS DQF (precisión, fluidez, terminología y *verity*) para la clasificación de errores. El corpus de investigación estuvo compuesto por un reportaje de un diario español con 1299 palabras. Para llevar a cabo el análisis, se tuvo como recursos

tecnológicos a 2 traductores automáticos (Google Translate y DeepL) y la opción de traducción que brinda ChatGPT-3.5. Adicionalmente, como parte de todo este proceso, se desarrolló un estudio exploratorio sin posesición y, posteriormente, el análisis de los errores con posesición. Los resultados mostraron que tanto la TA neuronal como la inteligencia artificial para la traducción no siempre logran transferir con precisión el contenido del texto origen al texto meta. También se evidenció una confusión en la equivalencia léxica, una gran cantidad de errores gramaticales en chino y la inconsistencia en la terminología seleccionada. A modo de conclusión, el estudio indica que es preferible seguir las convenciones chinas de la denominación en textos españoles para mejorar la traducción de las TA neuronales y de ChatGPT-3.5.

2.2. Bases teóricas

A continuación, se expone las bases teóricas de la presente investigación. Para iniciar con el desarrollo de los temas, se considera fundamental abordar algunas definiciones de la traducción por parte de diferentes autores.

2.2.1. La traducción

La traducción es un área que ha sido estudiada desde distintos puntos de vista, cada autor expresa sus ideas y conceptos según la base de su formación y sus conocimientos. De manera general se entiende que la traducción es un traspase de palabras de una lengua a otra, pero a lo largo de los años esta vaga definición ha ido tomando diversos cambios muy importantes a señalar.

Como primer concepto se presenta los planteamientos de Warwal (2015) quien señala que se puede considerar traducción como la comprensión del significado de un texto en una lengua origen y posteriormente realizar una producción de un texto equivalente en otra lengua o también llamada lengua meta. Este proceso llamado traducción permite que se comunique el mismo mensaje de un idioma a otro manteniendo la misma idea y el mismo significado. El producto de una traducción suele denominarse un texto meta.

Un concepto más complejo con respecto a la traducción es el que se presenta en la investigación dada por Koby et al. (2014) quienes mencionan que para poder definir la traducción es importante diferenciar dos puntos de vista distintos (amplio y concreto). La definición amplia de la traducción se refiere a esta como una creación de contenido de destino que corresponde a un contenido de origen según ciertas especificaciones. Esta definición al ser de cierta manera general permite decir que se puede llamar “traducción” a la traducción humana, la traducción automática sin editar, la traducción automática poseditada, lo que incluye a la traducción automática completa y ligeramente poseditada, la localización, la traducción completa (el texto destino incluye una frase por cada frase del texto origen), la traducción resumida (solo se tiene en cuenta ciertas frases del texto origen), un *gisting* (una traducción rápida y superficial realizada por un humano o por una máquina), la traducción técnica- científica, la traducción jurídica, la traducción literaria, entre otros. Esta definición permite decir que la traducción es multimodal y abarca diferentes aspectos.

En contraposición, la definición concreta de la traducción distingue entre la traducción propiamente dicha (*per se*) y otras actividades, que se nombran como un “*translation-plus*”. La traducción propiamente dicha define a la traducción como la transferencia de un texto escrito de origen a un texto escrito de destino de una extensión aproximadamente equivalente. La traducción debe transmitir todo el significado del texto origen y solo se realizará los ajustes necesarios para su adecuación cultural, sin añadir, omitir, considerar o adaptar nada más. A diferencia, en las actividades consideradas como “*translation-plus*” no se suele transmitir todo el significado del texto origen. Se puede considerar el resumen (la condensación del material después de la traducción) y la transcreación (considerada como la modificación del mensaje original de forma creativa después de la traducción).

Después de mencionar estos dos grandes aportes a la definición de la traducción, se puede concluir que la idea que se tiene de traducción va cambiando y que se encuentra en una evolución constante. Se puede mencionar que ahora la traducción también tiene relación con las modificaciones que un traductor hace al texto origen, como es el caso del resumen y la transcreación.

2.2.2. La traducción automática

En muchos estudios de la traducción, se puede destacar la relación que tiene esta disciplina con la lingüística. Uno de los campos dentro de esta ciencia es la lingüística computacional que se dedica a estudiar la interacción entre la informática y la lingüística. El mayor objetivo que se plantea es el poder recrear a través de las computadoras la capacidad del lenguaje humano. Es en este

contexto donde se puede situar la traducción automática, un tema de gran interés para la presente investigación, por lo que resulta crucial presentar algunas definiciones dadas por autores.

El concepto de traducción automática (TA) se refiere a los sistemas informáticos que se encargan de producir traducciones con o sin ayuda de una persona. En este sentido, se excluye las herramientas informáticas de traducción como los diccionarios en línea, bancos de datos terminológicos, entre otros que sirven de ayuda a los traductores en el proceso de traducción. A pesar de que es cierto que existe una traducción humana asistida por una máquina (MAHT, por sus siglas en inglés) y una traducción automática asistida por una persona (HAMT, por sus siglas en inglés) y que ambas pueden considerarse como una traducción asistida por ordenador (CAT, por sus siglas en inglés), la base principal de la TA es la idea de una automatización del proceso de la traducción (Hutchins, 1995).

Un concepto más actualizado sobre la traducción automática es el que brinda Rico (2017) quien manifiesta que la traducción automática es un proceso y no un producto. Se debe ver a la TA como un instrumento con el que el usuario pueda interactuar para realizar su labor. Este proceso debe ser cíclico, ya que todas las etapas al momento de trabajar en un proyecto de traducción deben estar interconectadas. En este aspecto, el rol del traductor tiene un enfoque más amplio y debe estar presente en cada etapa de la traducción, a modo de administrador que utiliza un sistema de traducción automática.

Por lo expuesto anteriormente se puede concluir que el avance que ha tenido la traducción automática es realmente interesante. A lo largo de los años, desde el inicio y aparición de la traducción automática los progresos en su investigación han ido incrementando y todo aporte es recibido por parte de los profesionales de la traducción para poder mejorar el servicio que se brinda. Se debe dejar en claro que la traducción automática no puede ser vista como el producto final, sino que es solo una etapa más dentro del proceso completo de un proyecto de traducción.

2.2.3. Tipos de traducción automática

A lo largo de las investigaciones sobre la traducción automática se han ido desarrollando sistemas que emplean la informática y la tecnología computacional para realizar traducciones. Desde las primeras etapas, la traducción automática ha requerido de grandes cambios y transformaciones, dando a lugar posteriormente a distintos tipos de sistemas o también llamados traductores automáticos, los cuales son importantes mencionar para esta investigación.

El primer tipo de traducción automática es la traducción basada en reglas (RBMT, por sus siglas en inglés). Durante el proceso de este sistema, se codifica toda la información recogida de las fuentes lingüísticas (diccionarios bilingües, glosarios terminológicos, etc.). Luego, junto con la aplicación de una serie de reglas específicas de ambas lenguas, se logra obtener la traducción de textos de una lengua origen a una lengua meta (Córdoba et al., 2015). Un claro ejemplo de este tipo de traducción es el sistema SYSTRAN, que estuvo integrado en el traductor ofrecido por Google hasta el año 2007.

El siguiente tipo de sistema es el traductor automático basado en la estadística (SMT, por sus siglas en inglés), el cual se desprende del uso de reglas específicas para poder adquirir un conocimiento automático de la traducción a partir del uso de datos masivos, conocido como corpus. Estos constan de millones de textos ya traducidos, disponibles en internet, lo que permite al traductor asociar palabras, frases o incluso un grupo más extenso de palabras con sus equivalentes en textos paralelos y a su vez calcular la probabilidad de una traducción correcta entre ambas lenguas. El traductor en base a la estadística escoge la opción más frecuente en su corpus para proponer el equivalente en la traducción. Todo esto genera que para obtener mejores resultados en la traducción se debe contar con un corpus con una gran cantidad de traducciones fiables (Maldonado y Liébana, 2021).

Otro tipo de traducción es el sistema basado en redes neuronales (NMT, por sus siglas en inglés) el cual es considerado como uno de los más recientes en el campo de la traducción automática. Maldonado y Liébana (2021) señalan que este sistema al igual que la traducción automática estadística requiere de un corpus con textos paralelos. Las redes neuronales artificiales que se encuentran en estos sistemas tienen la intención de pretender imitar el funcionamiento de las redes neuronales propias del cerebro humano. Para comprender esto es importante mencionar que las neuronas del cerebro reciben y procesan la información para realizar conexiones entre sí y asociar la información nueva con informaciones previas. Del mismo modo, las redes neuronales artificiales utilizan estas técnicas de aprendizaje de manera automática para poder “aprender” a

traducir a partir del corpus de textos paralelos, los cuales incluyen información de tipo lingüística y no lingüística. Mediante todo el complejo proceso, el traductor a base de redes neuronales observa el tipo de palabra (verbo, sustantivo, adjetivo, etc.) y aprende a buscar un equivalente dentro de su corpus para proponer un equivalente adecuado.

El último tipo de traducción automática importante de mencionar es la traducción automática híbrida. Este enfoque, según Anugu y Ramesh (2020) nace al unir dos o más sistemas de traducción automática, lo que se busca es extraer los beneficios de ambos TA y a su vez intentar abordar las deficiencias de uno mediante el uso del otro. Esta unión se refiere a que dentro de su motor traductor se combina dos diferentes TA, puede ser una TA basada en reglas junto con una TA basada en la estadística o lo que se maneja actualmente una TA basada en la estadística con una TA que usa redes neuronales. Este tipo de traducción suele usarse en un determinado software que analiza ambas opciones de traducción, ambos TA brindan su propia traducción completa y el objetivo es escoger la mejor opción.

En suma, se puede decir que el proceso de la traducción automática ha tenido un gran y complejo avance a lo largo de los años. Esto permite que la traducción mejore cada vez más y se tenga una mejor opinión de lo que es el uso de los traductores automáticos.

2.2.4. Calidad de traducción

Dentro de todo proceso de traducción, ya sea mediante una traducción humana o automática, el tema de la calidad está inmerso y se espera que el

resultado final cuente con una calidad óptima. Sin embargo, no se llega a un total acuerdo con respecto a una definición propiamente dicha de la calidad de la traducción. Se entiende que el significado puede variar de acuerdo con los propósitos de la traducción. Para esta investigación se considera pertinente plantear algunas definiciones por parte de autores que se enfoquen en el campo de la traducción automática dado que es el tema del cual se tiene un mayor interés.

Para Koby et al. (2014) la calidad de la traducción puede ser descrita desde dos diferentes puntos de vista. Se puede mencionar una definición de calidad desde una perspectiva amplia y también desde una perspectiva más concreta. Dentro de esta perspectiva concreta se encuentra una clara definición de lo que es la calidad con respecto al uso de traductores automáticos. Una traducción de alta calidad se consigue cuando el mensaje de la lengua origen se transfiere completamente a la lengua meta y se incluye aspectos como la denotación, connotación y el estilo. El texto en lengua meta debe mantener el uso correcto de la gramática y el vocabulario para que se consiga un texto culturalmente apropiado. Cuando se trate de una traducción automática, se requiere de un suficiente nivel de posesión para garantizar la precisión y la fluidez del texto.

Desde otra perspectiva, Hutchins (1995) señala que la calidad en la traducción obtenida mediante los sistemas de traducción automática sigue en proceso. Los TA todavía pueden mejorarse. Para que se mejore la calidad de los traductores automáticos se requiere desarrollar métodos más sofisticados. El

sistema puede mejorarse a través de la selección de ciertos tipos de textos con una gramática y vocabulario determinado sobre un campo temático concreto. Además, el texto en lengua origen podría estar escrito en un registro controlado del lenguaje, lo que se refiere a restringir parte del vocabulario para evitar problemas de homonimia y polisemia en estructuras complejas. Otra opción para mejorar el rendimiento de los TA es la realización de una preedición del texto origen con el fin de adaptarlo antes de su pase por el TA. Una última opción es dirigir aquellos problemas que dificulten el desarrollo óptimo del traductor a una persona (traductor) para que este se encargue de resolverlos durante el proceso de traducción. Todo lo anteriormente mencionado permitiría mejorar la calidad de un sistema de traducción automática.

Para Comparin (2017) La calidad del resultado de los diferentes traductores automáticos es inconsistente y depende de factores, como el paradigma del sistema de traducción automática que se utiliza. Además, debido a la diversidad de motores de traducción automática, la calidad varía considerablemente. Por lo tanto, se hace necesario realizar evaluaciones de calidad no solo para definir posibles mejoras, sino también poder obtener características únicas de cada TA, como los errores más comunes y el rendimiento en base a la calidad.

Aunque la definición de calidad de traducción puede diferir según los autores o incluso según el propósito que se la dará a la traducción, en el ámbito de la traducción automática aún se requiere la supervisión de los profesionales de la traducción. Es decir que, a pesar del considerable avance de estos

sistemas, la revisión humana sigue siendo crucial para garantizar un trabajo de alta calidad.

2.2.5. Calidad de la traducción neuronal y la traducción híbrida

La calidad de las traducciones automáticas puede variar significativamente según el tipo de sistema utilizado. En este caso, es relevante señalar la opinión de algunos autores respecto a la calidad de la traducción neuronal y la traducción híbrida, ya que son los sistemas de traducción presentes en el trabajo.

De acuerdo con Liu y Zhu (2023), la traducción automática neural (NMT) ha alcanzado grandes resultados y ha proporcionado muchos beneficios a la industria de la traducción y compañías de tecnología. Este tipo de sistema se emplea para mejorar la precisión y fiabilidad de la traducción. Sin embargo, su principal desafío radica en la calidad insuficiente de las traducciones, lo que no cumple con las expectativas de los usuarios. Este problema se manifiesta en errores léxicos, inexactitudes en los tiempos verbales, fallos sintácticos y confusión de la estructura, entre otros aspectos. Esta perspectiva crítica es importante al momento de evaluar la efectividad y utilidad de los sistemas de traducción neuronal en el contexto de la investigación.

En cuanto a la traducción híbrida para Anugu y Ramesh (2020), el empleo de diversas técnicas en un conjunto de dos sistemas de traducción, conocido como traducción híbrida, mejora la calidad de traducción al intentar superar las limitaciones de uno de los sistemas mediante el uso del otro. Sin embargo, debido al amplio campo del vocabulario, la naturaleza del léxico y las irregularidades

lingüísticas, la calidad de traducción no consigue alcanzar el nivel de precisión y fluidez proporcionado por una traducción realizada por un profesional.

Con lo visto anteriormente, se entiende que tanto la traducción automática neuronal como la traducción híbrida ofrecen avances significativos en la industria de la traducción. No obstante, ambas enfrentan desafíos en cuanto a la calidad de las traducciones, lo cual limita su eficacia para obtener resultados de alta calidad y se crea la necesidad de contar con la orientación y revisión de un profesional para garantizar mejores resultados.

2.2.6. Evaluación de calidad de traducción

A continuación, se presenta otro tema muy importante con relación a la calidad de las traducciones. Si bien, no existe una definición exacta de lo que es la calidad, diversos autores mencionan que para poder medir la calidad de un texto traducido es imprescindible someter dicha traducción por un proceso de evaluación de calidad. En el caso de la traducción automática, el tema de la calidad es aún más importante. Para esto es importante mencionar en qué consiste la evaluación de calidad.

Para Esteban (2020) es importante saber medir y evaluar la calidad de las traducciones. En el caso del uso de diferentes sistemas de traducción automática se debe evaluar la calidad para comparar sistemas, realizar mejoras y entrenar al traductor para mejorar los resultados brindados. Este proceso de evaluación de calidad es complejo dado que implica poner a prueba factores lingüísticos y extralingüísticos.

En el proceso de la evaluación de calidad de traductores automáticos, es importante distinguir entre la evaluación humana y la evaluación automática, en la cual se evalúa la calidad de traducción en base a algoritmos informáticos. A continuación, se presentará una breve definición de ambos tipos de evaluación con el fin de complementar la documentación para el presente estudio.

Moorkens et al. (2018) señalan que, para definir la evaluación humana en relación con la calidad de los traductores automáticos, es importante mencionar que cualquier proceso de evaluación suele variar considerablemente en función de quien realiza dicha evaluación y la finalidad del texto. La evaluación humana es subjetiva, se requiere de más tiempo y en ocasiones puede implicar un gasto adicional a la traducción. Para este proceso se tiene en cuenta factores como la precisión, la fluidez, la legibilidad, la comprensión y la aceptabilidad, los cuales van a determinar la evaluación de calidad. Generalmente, los revisores a cargo de la evaluación separan el texto a traducir por segmentos y se guían mediante escalas de Likert.

En contraste, el caso de la evaluación automática de la calidad de los traductores automáticos es mucho más compleja. Esteban (2020) menciona que, en el proceso de evaluación automática, son los algoritmos y otros procesos los que se encargan de calcular puntuaciones numéricas y se suelen basar en diferentes aspectos según la métrica. Estas métricas deben estar relacionadas con la opinión o el trabajo de una persona traductora. La relación entre la métrica y la opinión de la persona logra en sí el concepto de calidad, la métrica sola no puede definir la calidad. Actualmente, ya existen diversas métricas que evalúan

de manera automática, entre las más conocidas resaltan las métricas BLEU (*Bilingual Evaluation Understudy*) y *Dynamic Quality Framework* de TAUS.

A modo de conclusión, ambas maneras de evaluación ayudan a mejorar el nivel de calidad en las traducciones. En los dos tipos de evaluación, se puede observar ventajas y también desventajas. Sin embargo, el objetivo principal en ambos casos es ayudar en la labor traductora y proporcionar mejores resultados en las traducciones finales.

2.2.7. La métrica SAE J2450

Uno de los métodos de evaluación de calidad basado en métricas es el uso de la evaluación con las métricas SAE J2450, la cual resulta imprescindible para la presente investigación y se definirá en detalle a continuación.

La Sociedad de Ingenieros de Automoción (SAE por sus siglas en inglés) es una organización que se encarga del estudio de nuevas tecnologías en el campo de la automoción. En 1997, la SAE convocó un comité de estudio en conjunto con representantes de General Motors, Ford y Chrysler, y también con representantes de firmas traductoras para desarrollar y establecer un sistema de evaluación de traducciones con el fin de mantener un estándar y mejorar la calidad de las traducciones de los diferentes servicios brindados por las compañías automotrices. Este comité de estudio tuvo el nombre oficial de J2450 (SAE, 2001).

Esta métrica se basa en errores que el evaluador según su experiencia traductora logra detectar. Esta métrica puede aplicarse independientemente del

idioma de origen o del método de traducción (es decir, traducción humana, traducción asistida por ordenador o traducción automática).

La métrica SAE J2450 consiste en cuatro etapas. La primera etapa comprende las siete categorías principales de error, en la segunda etapa se encuentran las dos subcategorías secundarias para clasificar los errores detectados (grave y leve). La tercera etapa consta de dos meta-reglas que el evaluador debe seguir: cuando un error es ambiguo, siempre se debe elegir la categoría que según el evaluador es más importante; en caso de alguna duda, el evaluador debe elegir siempre la opción “grave” en lugar de “leve”. Una última etapa consiste en la valoración numérica que se le otorga a cada error detectado. Mientras menor sea el resultado del total de las valoraciones numéricas mayor será la calidad de la traducción.

La métrica SAE J2450 comprende siete categorías que se basan en la detección de diferentes tipos de errores. El primer tipo es el error terminológico (*Wrong Term* - WT), este error mencionado en la métrica puede asociarse a la definición dentro la clasificación de errores dado por Vilar et al. (2006), ellos mencionan que, en el caso de los traductores automáticos, estos no siempre pueden encontrar el equivalente correcto para un término y al momento de decidir un posible equivalente, este resulta erróneo.

El segundo tipo es el error sintáctico (*Syntactic Error* - SE), en este caso siguiendo la definición de Vilar et al. (2006), el error surge al encontrar un desorden en las palabras dentro de una oración, además también puede

considerar dentro de esta categoría cuando ocurre un reordenamiento a nivel de las oraciones.

El tercer tipo se refiere a un error de omisión (*Omission* - OM), este tipo de error se puede asociar a la definición de Vilar et al. (2006), quienes mencionan que este error ocurre cuando se omiten y faltan palabras principales en el texto, tales como los verbos o sustantivos.

El cuarto tipo es el error morfológico o de concordancia (*Word Structure or Agreement Error* - SA), este tipo de error puede asociarse con la definición de Temnikova (2010), quien menciona que en este error se suele escoger la palabra correcta, pero se coloca de manera incorrecta en la traducción, como es el caso de género y número.

El quinto tipo es el error ortográfico (*Misspelling* - SP), este error se asocia a la definición de Vilar et al. (2006), quienes señalan que cuando se ingresan ciertos términos desconocidos al traductor automático, los resultados en muchas ocasiones terminan teniendo errores en la ortografía que pueden modificar el sentido de la palabra.

El sexto tipo es el error de puntuación (*Punctuation Error* - PE), para Temnikova (2010), este error sucede cuando se omite el uso de algún signo de puntuación o cuando se escoge un signo de puntuación incorrecto.

El séptimo y último tipo es el error de otro tipo (*Miscellaneous Error* - ME), en este tipo de error se incluyen todo tipo de error lingüísticos relacionados con la lengua meta que no son claramente atribuibles a una de las categorías

anteriores. Según la clasificación de Temnikova (2010) se da el caso del error en la traducción de expresiones idiomáticas, debido a que los traductores automáticos suelen traducir estas frases de manera literal.

En suma, esta métrica resulta de gran ayuda en la evaluación de calidad de traducciones y para el presente trabajo el caso de traducciones obtenidas por traductores automáticos, en la cual la presencia de un traductor-revisor a modo de evaluador será crucial.



2.3 Definición de términos

- Traducción técnica: traducción de un tema en específico que requiere conocimientos de terminología especializada o información técnica (Hurtado, 2001).
- Traducción automática: proceso computarizado de convertir un texto de un idioma a otro, teniendo al traductor humano como administrador y no operario del proceso (Rico, 2017).
- Errores de traducción: desviaciones o inexactitudes que se presentan al momento de pasar un texto de un idioma a otro, esto afecta a la fidelidad, coherencia o comprensión del mensaje original. (Hurtado, 2001).
- Métricas de evaluación: estándares o medidas que son utilizadas para cuantificar el rendimiento o la eficacia de un sistema o proceso, permitiendo así evaluar su calidad en base a resultados cuantificables (House, 2014).
- Calidad: conformidad de estándares establecidos para un producto o servicio, de acuerdo con la evaluación de su efectividad, precisión y el cumplimiento de expectativas o requisitos por parte del cliente (Valls et al., 2017).
- Revisión: proceso de análisis de un proceso o producto con el objetivo de identificar errores, mejoras o asegurar la calidad y precisión (Mossop et al., 2019).

2.4 Hipótesis

Hernández et al. (2018) destacan que las hipótesis indican las afirmaciones provisionales que buscan probarse en una investigación y dar una posible respuesta a las preguntas de investigación. No obstante, también reconocen que no todos los trabajos requieren la formulación de hipótesis, ya que esto depende del enfoque de la investigación, aspecto esencial a considerar. En concordancia con lo anterior, esta investigación, al tener un enfoque exploratorio, no incluye hipótesis, dado que busca comprender y explorar una situación o problema sin preconcepciones específicas.



CAPÍTULO III. MÉTODO

En esta sección, se describe la metodología de la investigación, incluyendo el tipo, nivel y diseño seleccionados, junto con la justificación de esta elección. También, se detallan las variables, dimensiones e indicadores, así como la muestra y la técnica e instrumento utilizados para la recolección de datos. Finalmente, se describe el procesamiento y análisis de datos, así como el método utilizado.

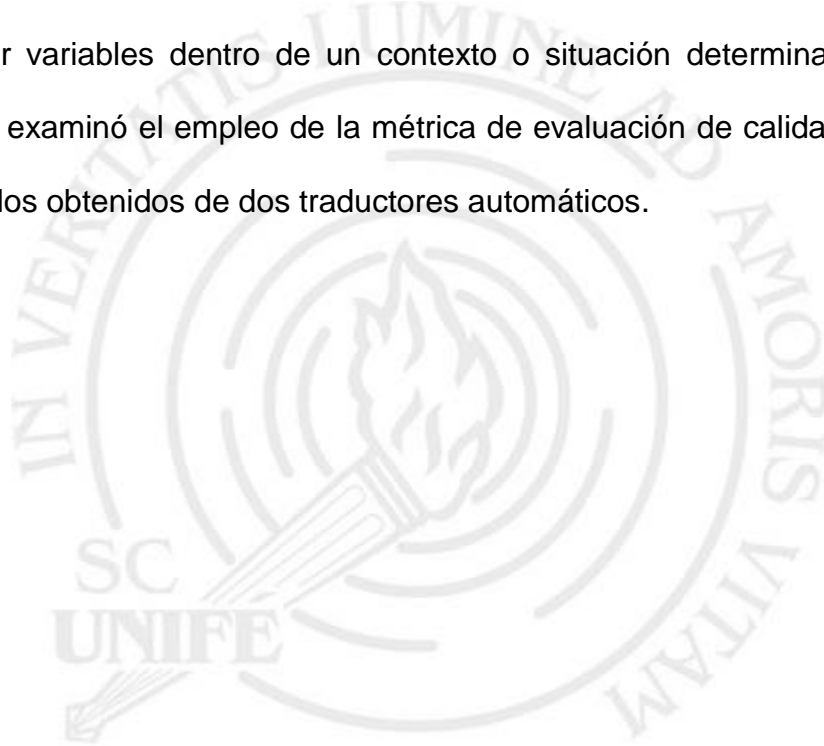
3.1. Nivel, tipo y diseño de la investigación

El presente estudio optó por un enfoque cuantitativo dado que se tuvo que describir y analizar la calidad de traducción de dos traductores automáticos. De acuerdo con Hernández et al. (2018) durante el proceso de investigaciones cuantitativas se utiliza la recolección de datos para medir los resultados de manera objetiva y se debe evitar cualquier fenómeno subjetivo.

Para Hernández et al. (2018) al hablar de una investigación de nivel exploratorio, el propósito es explorar aspectos poco abordados con respecto a un tema específico. En este nivel, la prioridad es indagar y recopilar información nueva y poco conocida para enriquecer el conocimiento en el ámbito de la investigación. Siguiendo este concepto, el presente trabajo se clasificó como exploratorio, puesto que se buscó examinar y analizar el uso de una métrica de evaluación de calidad poco reconocida en el campo de la traducción, además de contar con la participación de un grupo de estudio.

Además, esta investigación fue de tipo básica, conforme a lo indicado por Hernández et al. (2018), el propósito de este tipo de investigaciones es ampliar la comprensión sobre un tema en particular. En este contexto, el presente trabajo procuró aportar nuevos conocimientos para contribuir al campo de la evaluación de la calidad, focalizándose especialmente en el empleo de métricas de evaluación.

Finalmente, el diseño de la presente investigación fue transeccional exploratorio. Como señalan Hernández et al. (2018), este tipo de diseño implica realizar una exploración inicial de datos en un momento específico con el propósito de describir variables dentro de un contexto o situación determinados. En este estudio, se examinó el empleo de la métrica de evaluación de calidad utilizada en los resultados obtenidos de dos traductores automáticos.



3.2. Operacionalización de variables

Tabla 1

Operacionalización de variables

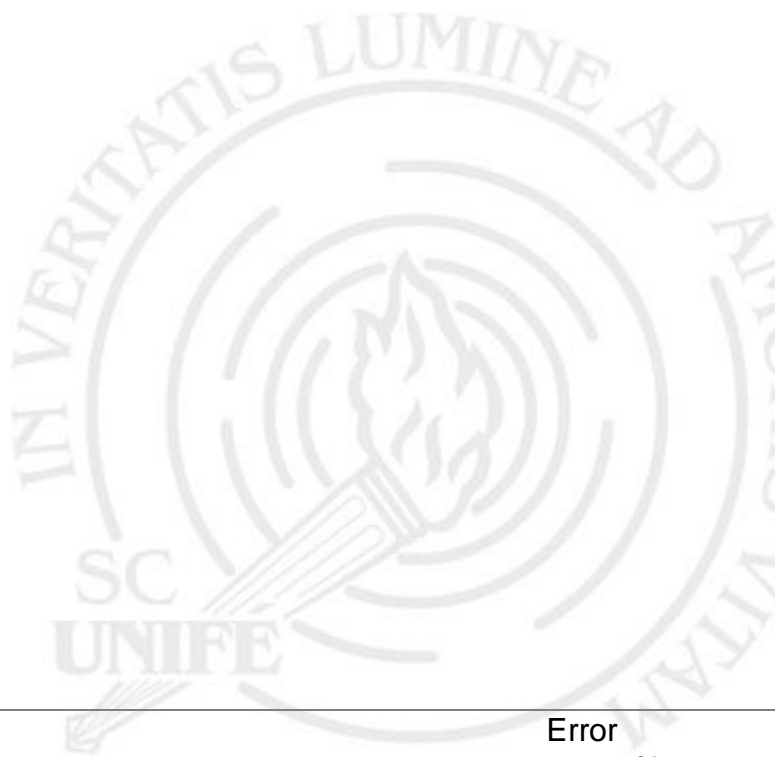
Variable	Definición conceptual	Definición operacional	Dimensiones	Indicadores
Calidad de traducción	Una traducción de alta calidad se logra cuando el mensaje de la lengua origen se transfiere completamente a la lengua meta y se incluye aspectos como la denotación, connotación y el estilo. El texto en lengua meta debe mantener el uso correcto de la gramática y el vocabulario para que se consiga un texto culturalmente apropiado. Cuando se trate de una	Para medir la calidad de la traducción, se utilizó la métrica de evaluación SAE J2450, la cual se enfoca en la traducción de textos del área automotriz.	Error terminológico	Término que no respeta el glosario de términos del cliente Término que no es usado en las traducciones estándar de la lengua origen en el campo temático Término incoherente con otras traducciones del término de la lengua origen en el mismo documento o tipo de documento, a menos que el contexto de término de la lengua origen justifique el uso de un término de la lengua de destino diferente (por ambigüedad) Término que denota un concepto totalmente diferente en la lengua meta

traducción automática, se requiere de un suficiente nivel de posesición para garantizar la precisión y la fluidez del texto. (Koby et al., 2014).

Error sintáctico El término en lengua origen tiene asignado una categoría gramatical incorrecta en la oración del texto meta

El texto de destino contiene una estructura de frase incorrecta, por ejemplo, una oración subordinada relativa cuando se necesita una frase verbal. Las palabras de la lengua de llegada son correctas, pero están en un orden lineal incorrecto según las reglas sintácticas de

		la lengua meta
	Error de omisión	Se omite parte del texto o gráficos que pertenecen del texto en lengua origen
	Error morfológico o de concordancia	Un error de estructura de la palabra se produce cuando una palabra o un término que, por lo demás, es correcto, se expresa en una forma morfológica incorrecta (caso, género, número, tiempo, prefijo, sufijo, infijo, etc.). Un error de concordancia se produce cuando dos o más palabras del texto meta no coinciden en ninguna forma de inflexión
	Error ortográfico	No respeta la ortografía según el glosario del cliente No respeta las normas ortográficas de la lengua destino No respeta el sistema de escritura, el texto



está escrito de manera incorrecta e inadecuada

Error de puntuación

El texto meta contiene errores según las reglas de puntuación de la lengua meta

Error de otro tipo

Errores lingüísticos relacionados con la lengua meta que no son claramente atribuibles a una de las categorías anteriores



3.3. Muestra

Como unidad de análisis se tuvo la evaluación de la calidad, para esta investigación se hizo uso de dos traductores automáticos. El traductor automático DeepL, el cual es un sistema alemán desarrollada por DeepL SE en el año 2017 y desde sus inicios contó con el uso de redes neuronales en su sistema. Asimismo, se utilizó el traductor Yandex Translate, sistema que pertenece a la compañía Yandex. En 2011, se desarrolló su motor de traducción que originalmente se basa en la estadística, pero en el 2017 incluyó el uso de una traducción en base a redes neuronales y actualmente manejan un sistema híbrido que combina ambas traducciones automáticas para optar por la mejor opción de traducción. Ambos traductores fueron aplicados en un texto semi técnico. Asimismo, como unidad muestral se tuvo en cuenta un extracto de un documento técnico del área de la automoción. Se utilizó una guía explicativa del uso de un nuevo servicio para autos y se seleccionó solo algunos segmentos de las primeras páginas. Este tipo de texto técnico tiene como características mantener un lenguaje preciso y claro, trata de evitar ambigüedades y se enfoca en transmitir información de manera directa y comprensible. Además, cuenta con una terminología específica, se suele utilizar términos técnicos, siglas, acrónimos y detalles específicos.

3.4. Técnica e instrumento de recolección de datos

El presente trabajo tuvo como técnica de investigación la observación, dado que se planteó observar los errores de traducción en relación con la calidad a través del uso de una métrica de evaluación realizada por parte de traductores-revisores de un estudio de traducción. Para Hernández, et al. (2018), la observación consiste en el registro sistemático y confiable de algunas situaciones a través de un conjunto de categorías y subcategorías.

Asimismo, el instrumento que se utilizó para la recolección de datos fue la métrica SAE J2450. Esta métrica fue desarrollada por la *Society of Automotive Engineers (SAE)*, su objetivo principal se basa en la revisión de traducciones de textos de automoción (SAE, 2001). La métrica no se diseñó originalmente para traducciones donde el estilo, registro y tono puedan desempeñar un papel crucial, como en traducciones de marketing o publicidad. Sin embargo, su aplicación, aunque con ajustes, ha empezado a extenderse a otros sectores industriales recientemente, como la biología, incluyendo productos farmacéuticos, dispositivos médicos, entre otros (Martínez, 2014).

De igual manera, la métrica SAE J2450 pretende medir objetivamente la calidad de traducciones independientemente de la lengua origen, la lengua meta y el método de traducción (traducción humana, traducción asistida o traducción automática) (SAE, 2001). Para el uso de esta métrica en el presente trabajo se contó con la validación de 3 profesionales especialistas en el área temática.

La métrica se basa en una clasificación de errores y una subclasificación de estos. Para la primera clasificación, se cuenta con siete categorías: error

terminológico, error sintáctico, omisión, error morfológico o de concordancia, error de ortografía, error de puntuación y error de otro tipo. El error terminológico se refiere cuando el término no respeta el glosario del cliente, cuando se elige un término que no es usado en las traducciones estándar la lengua origen en el campo temático, cuando el termino no tiene relación con otras traducciones en el mismo documento, a menos que el contexto de término de la lengua origen justifique el uso de un término distinto o cuando se emplea un término que denota un concepto totalmente diferente en la lengua meta. El error sintáctico se refiere cuando el término en lengua origen tiene asignado una categoría gramatical incorrecta en la oración del texto meta, cuando el texto de destino contiene una estructura de frase incorrecta, por ejemplo, una oración subordinada relativa cuando se necesita una frase verbal y por último cuando las palabras de la lengua de llegada son correctas, pero están en un orden lineal incorrecto según las reglas sintácticas de la lengua meta. El error de omisión es cuando se omite partes del texto o gráficos que pertenecen al texto origen. El error morfológico o de concordancia se produce cuando una palabra o un término que, por lo demás, es correcto, se expresa en una forma morfológica incorrecta (caso, género, número, tiempo, prefijo, sufijo, infijo, etc.). Un error de concordancia se produce cuando dos o más palabras del texto meta no coinciden en ninguna forma de inflexión. El error ortográfico se da cuando no se respeta la ortografía según el glosario del cliente, no se respeta las normas ortográficas de la lengua destino o no se respeta el sistema de escritura, el texto está escrito de manera incorrecta e inadecuada. El error de puntuación se da cuando el texto meta contiene errores según las reglas de puntuación de la lengua meta. En la última categoría, los errores de otro tipo consisten en cualquier error

lingüístico relacionado con la lengua meta que no son claramente atribuibles a una de las categorías anteriores. Los errores de cada categoría pueden subclasificarse como graves o leves, con una puntuación numérica asociada a cada tipo de error y a su respectivo nivel de gravedad como se observa en la tabla siguiente (SAE, 2001):

Tabla 2

Tipos de error, clasificación y puntuación

Tipo de error: (abreviación en inglés)	Subclasificación: (abreviación en inglés)	Puntuación: Grave / Leve
a. Error terminológico (WT)	Grave (s)	5/2
b. Error sintáctico (SE)	Leve (m)	4/2
c. Omisión (OM)		4/2
d. Error morfológico o de concordancia (SA)		4/2
e. Error ortográfico (SP)		3/1
f. Error de puntuación (PE)		2/1
g. Error de otro tipo (ME)		3/1

Fuente: SAE (2001)

Para la recolección de resultados se tiene en cuenta como primer paso la suma del resultado de la multiplicación de la cantidad de errores detectados por la puntuación obtenida de cada uno de estos (grave o leve), este paso se repite para cada tipo de error (7 tipos de error). El resultado tiene el nombre de “Puntuación ponderada por categoría”. Para conocer la puntuación total del documento traducido se debe sumar cada puntuación ponderada y posteriormente dividir el resultado entre el total de palabras del texto origen.

Tabla 3*Análisis de resultado*

Tipo de error	n.º de errores x Puntaje (grave)		n.º de errores x Puntaje (leve)		Puntuación ponderada por categoría
Error terminológico (ET)	<u> x 5 </u>	+	<u> x 2 </u>	=	
Error sintáctico (ES)	<u> x 4 </u>	+	<u> x 2 </u>	=	
Omisión (OM)	<u> x 4 </u>	+	<u> x 2 </u>	=	
Error morfológico o de concordancia (EM)	<u> x 4 </u>	+	<u> x 2 </u>	=	
Error ortográfico (EO)	<u> x 3 </u>	+	<u> x 1 </u>	=	
Error de puntuación (EP)	<u> x 2 </u>	+	<u> x 1 </u>	=	
Error de otro tipo (EOT)	<u> x 3 </u>	+	<u> x 1 </u>	=	

Nota: Esta puntuación es resultado según la métrica SAE J2450

Puntuación total del documento: (suma de las puntuaciones ponderadas ÷ el número de palabras en el texto origen).

$\underline{ET} + \underline{ES} + \underline{OM} + \underline{EM} + \underline{EO} + \underline{EP} + \underline{EOT} = (\text{Suma de las puntuaciones ponderadas})$

$\underline{\text{Suma de las puntuaciones ponderadas}} \div \underline{\text{el número de palabras en el texto}}$

$\underline{\text{origen}} = \underline{\text{Puntuación total del documento}}$

Fuente: SAE (2001)

3.5. Técnicas de procesamiento y análisis de datos

Para esta investigación, se estructuró el proceso en tres fases. En la primera etapa, se procedió a la selección del texto que sería utilizado en los traductores automáticos, centrándose específicamente en el ámbito técnico de la automoción.

Para la segunda fase, se elaboró una ficha de identificación y conteo de errores, la cual fue entregada al grupo de estudio. Este grupo estaba compuesto por dos traductores anónimos con más de un año de experiencia en revisión que validaron los errores encontrados y con profesionales del área de corrección de textos y corrección de estilo. La tarea se centró en identificar los errores presentes en las traducciones generadas por dos traductores automáticos DeepL y Yandex Translate. Además, se les solicitó asignar puntajes conforme a los parámetros establecidos. La ficha también sirvió para que el revisor contara los errores encontrados en ambas opciones de traducción. Cabe destacar que se estableció un tiempo específico para la entrega de la prueba de revisión.

Posteriormente, se recolectaron las fichas de ambos traductores para luego enviar estas fichas con un traductor profesional en el área de corrección de textos y corrección de estilo para que diera una última revisión. En la fase final, la tercera etapa, se procedió al análisis de los resultados. Este análisis incluyó la elaboración de un informe descriptivo en Excel, seguido por la creación de gráficos que resumieron y visualizaron de manera efectiva los hallazgos obtenidos durante la revisión.

3.6. Método de análisis

Para esta investigación, se utilizó un análisis descriptivo para resumir, organizar y presentar los datos de manera informativa y comprensible. Se hizo uso de gráficos y tablas para el proceso de los resultados y de acuerdo con Diz (2016) al hacer uso de la estadística descriptiva, se pone en práctica un conjunto de métodos que permiten describir datos de manera ordenada.



CAPÍTULO IV. RESULTADOS

En el cuarto capítulo se desarrollan los resultados obtenidos tras finalizar el proceso de análisis, los cuales están vinculados con el objetivo general y los objetivos específicos de la investigación.

Los resultados recopilados a través de las fichas de análisis y las pruebas se muestran a continuación mediante tablas y gráficos. Estos describen el análisis de las traducciones automáticas según el objetivo general de la investigación el cual es evaluar la calidad de dos traductores automáticos utilizados en un estudio de traducción en Lima, 2023.

Tabla 4

Análisis de resultado del traductor DeepL y del traductor Yandex

Tipo de error	Puntuación ponderada por categoría			
	DeepL		Yandex	
	R1	R2	R1	R2
Error terminológico (ET)	20	24	19	23
Error sintáctico (ES)	8	8	8	8
Omisión (OM)	2	0	0	0
Error morfológico o de concordancia (EM)	0	0	0	4
Error ortográfico (EO)	0	0	0	0
Error de puntuación (EP)	3	1	3	1
Error de otro tipo (EOT)	11	8	5	1

Puntuación y porcentaje promedios de calidad del traductor DeepL:
0.27597/ 72.39 %

Puntuación y porcentaje promedios de calidad del traductor Yandex:
0.23376/ 76.62 %

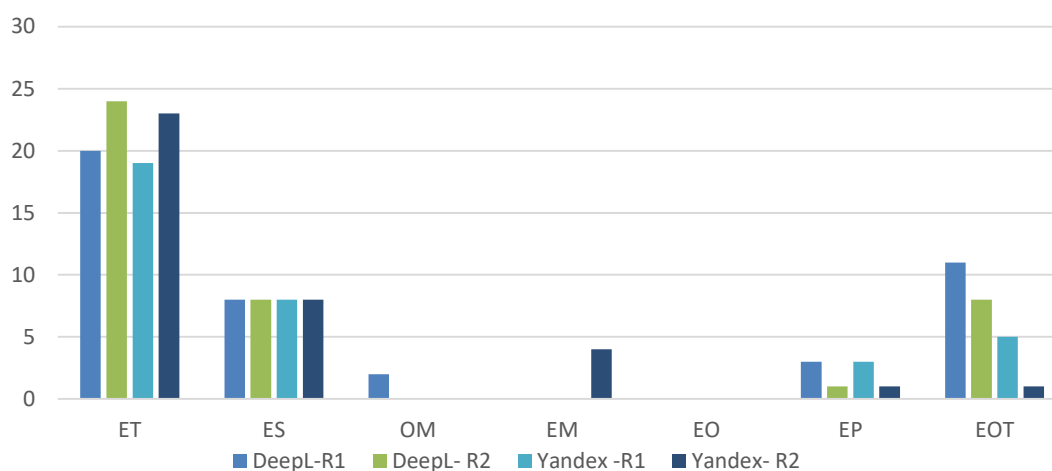
Porcentaje promedio de calidad de ambos traductores: 74.50 %

Nota: Esta puntuación es resultado según la métrica SAE J2450

R1: Revisor 1 / R2: Revisor

Figura 1

Resultados del traductor DeepL y del traductor Yandex



Descripción:

En todas las pruebas realizadas comparando ambos traductores automáticos, se observó que el error más significativo, encontrado por los revisores, fue el error terminológico (ET), se obtuvo un puntaje entre 19 y 23. El siguiente error que estuvo presente fue el error sintáctico (ES) con un puntaje promedio de 8 puntos. Los errores de otro tipo (EOT) tuvieron puntajes variados, dese 1 hasta 11. En cuanto al error de puntuación (EP) estuvo presente en todas las pruebas con un puntaje mínimo de 1 y máximo de 3. El error morfológico o de concordancia (EM) estuvo presente en una de las pruebas con un puntaje de 4. Asimismo, el error de omisión (OM) registró un puntaje de 2 al estar presente en una prueba. Es importante señalar que no se encontraron problemas relacionados con el error ortográfico (EO) en ninguna de las pruebas. En términos de la calidad, en ambas traducciones automáticas se disminuye en gran medida debido a los altos puntajes en errores terminológicos y sintácticos, los cuales tienen un impacto considerable en la evaluación global.

A continuación, en la siguiente tabla se muestra los resultados del primer objetivo específico, el cual es analizar la calidad de la traducción automática neuronal del traductor DeepL utilizado en un estudio de traducción en Lima, 2023.

Tabla 5

Análisis de resultado de la traducción neuronal de DeepL

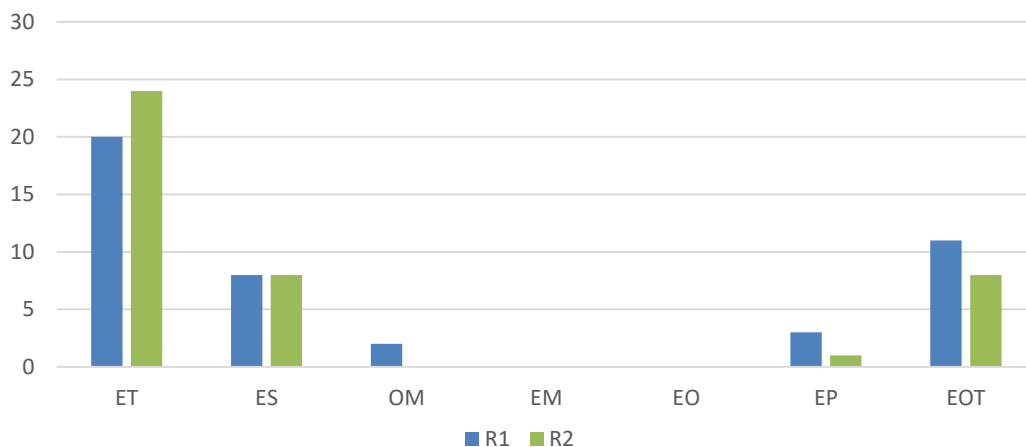
Tipo de error	N.º de errores graves		N.º de errores leves		Puntuación ponderada por categoría	
	R1	R2	R1	R2	R1	R2
Error terminológico (ET)	4	4	0	2	20	24
Error sintáctico (ES)	2	2	0	0	8	8
Omisión (OM)	0	0	1	0	2	0
Error morfológico o de concordancia (EM)	0	0	0	0	0	0
Error ortográfico (EO)	0	0	0	0	0	0
Error de puntuación (EP)	1	0	1	1	3	1
Error de otro tipo (EOT)	0	0	11	8	11	8
Puntuación total del documento y porcentaje de calidad R1: <u>0.28571/ 71.42 %</u>						
Puntuación total del documento y porcentaje de calidad R2: <u>0.26623/ 73.37 %</u>						
Porcentaje promedio de calidad: 72.39 %						

Nota: Esta puntuación es resultado según la métrica SAE J2450

R1: Revisor 1 / R2: Revisor 2

Figura 2

Resultados de la traducción neuronal de DeepL



Descripción:

La figura detalla los resultados en ambas pruebas sobre el traductor DeepL, donde se observó que el error terminológico (ET) tuvo un peso mayor, alcanzando un puntaje entre 20 y 24 según ambos revisores, quienes identificaron 8 errores graves y 2 errores leves. Por otro lado, el error de otro tipo (EOT) recibió un puntaje mínimo de 8 y máximo de 11, con 19 errores leves. En cuanto al error sintáctico (ES) y de acuerdo con ambos revisores se registró un puntaje de 8 con 4 errores graves. El error de puntuación (EP) obtuvo un puntaje mínimo de 1 y máximo de 3 con solo un error grave y 2 errores leves. El error de omisión (OM) solo estuvo presente en una de las pruebas con 1 error considerado leve. Es importante señalar que no se encontraron problemas relacionados con el error morfológico o de concordancia (EM) y el error ortográfico (EO). En términos de calidad de traducción, esta disminuye en gran medida debido al alto puntaje de errores terminológicos y sintácticos, así como a la cantidad de errores graves. No obstante, la ausencia de los errores EM y EO contribuyó positivamente a la calificación final de la calidad.

Para analizar el siguiente objetivo específico, se recopilaron las siguientes tablas y gráficos, teniendo en cuenta que el objetivo es analizar la calidad de la traducción automática híbrida estadística-neuronal del traductor Yandex, utilizado en un estudio de traducción en Lima, 2023.

Tabla 6

Análisis de resultado de la traducción híbrida neuronal-estadística de Yandex

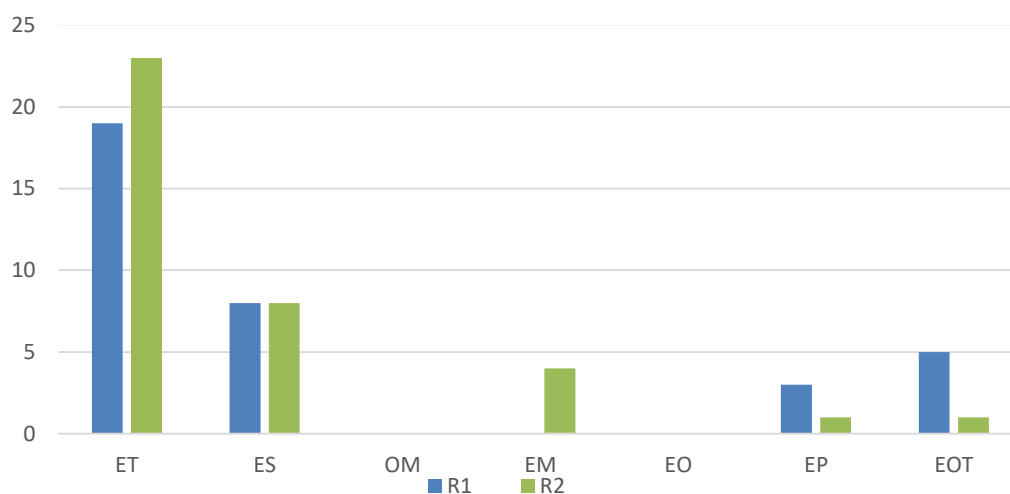
Tipo de error	N.º de errores graves		N.º de errores leves		Puntuación ponderada por categoría	
	R1	R2	R1	R2	R1	R2
Error terminológico (ET)	3	3	2	4	19	23
Error sintáctico (ES)	2	2	0	0	8	8
Omisión (OM)	0	0	0	0	0	0
Error morfológico o de concordancia (EM)	0	0	0	2	0	4
Error ortográfico (EO)	0	0	0	0	0	0
Error de puntuación (EP)	1	0	1	1	3	1
Error de otro tipo (EOT)	0	0	5	1	5	1
Puntuación total del documento y porcentaje de calidad R1: <u>0.22727 / 77.27 %</u>						
Puntuación total del documento y porcentaje de calidad R2: <u>0.24025/ 75.97 %</u>						
Porcentaje promedio de calidad: 76.62 %						

Nota: Esta puntuación es resultado según la métrica SAE J2450

R1: Revisor 1 / R2: Revisor 2

Figura 3

Resultado de la traducción híbrida neuronal – estadística de Yandex



Descripción:

La figura detalla los resultados en ambas pruebas sobre el traductor Yandex, donde se observó que el error terminológico (ET) tuvo mayor peso, alcanzando un puntaje entre 19 y 23 según ambos revisores, quienes encontraron 6 errores graves y 6 errores leves. El error sintáctico (ES) obtuvo un puntaje promedio de 8, se encontró 2 errores graves en ambas pruebas. El error de otro tipo (EOT) recibió un puntaje mínimo de 1 y máximo de 5, los revisores encontraron en total 6 errores leves. El error de puntuación (EP) obtuvo un puntaje mínimo de 1 y máximo de 3 con solo un error grave y 2 errores leves. El error morfológico o de concordancia (SA) solo estuvo presente en una de las pruebas con 2 errores leves. Es importante señalar que no se encontraron problemas relacionados con el error de omisión (OM), y de ortografía (EO). En términos de la calidad de traducción, esta disminuye en gran medida debido al alto puntaje de errores terminológicos y sintácticos, así como a la gran cantidad de errores considerados grave. No obstante, la ausencia de los errores OM y EO contribuyó positivamente a la calificación final de la calidad.

CAPÍTULO V. DISCUSIÓN DE RESULTADOS

En esta sección del quinto capítulo, se discuten los resultados de la presente investigación teniendo en cuenta una comparación entre los hallazgos de los antecedentes previamente mencionados y los resultados de esta investigación. Asimismo, se muestra a detalle las coincidencias y diferencias entre ambos trabajos.

Con respecto al objetivo general de esta investigación, que consistió en evaluar la calidad de dos traductores automáticos utilizados en un estudio de traducción en Lima, 2023, se obtuvo un promedio de 74.50 % como puntaje de calidad de traducción. Para Comparin (2017), la calidad de los resultados de diferentes traductores automáticos es inconsistente y depende de ciertos factores. Esto hace que se requiera de una evaluación de calidad no solo para identificar posibles áreas de mejora, sino también distinguir las características únicas de cada TA, como los errores que podrían ser más frecuentes y su rendimiento en términos de calidad. En relación con los resultados, las traducciones de ambos TA obtuvieron diferentes puntajes y también se presentó diferentes tipos de errores. En cuanto a los errores, tanto DeepL como Yandex mostraron la mayor cantidad de errores de tipo terminológico y sintáctico, lo que redujo su puntaje final de calidad en ambos casos. Finalmente, el traductor automático DeepL obtuvo un puntaje de 0.27597 que representa el 72.39 % como porcentaje de calidad mientras que el traductor automático Yandex obtuvo un puntaje de 0.23376, equivalente al 76.62 %. Estos puntajes se consideran de calidad media-baja ya que se necesita promedios del 90 % a 100 % para alcanzar una calidad óptima.

Los resultados de este trabajo coinciden parcialmente con la investigación de Adán (2019) quién concluyó que la traducción de Yandex al igual que DeepL mostraron una mayor presencia de errores en terminología e idiomática. En el caso de DeepL, de un total de 22 errores, 8 fueron de terminología y 7 de idiomática. Por otro lado, Yandex presentó un total de 39 errores, de los cuales 12 fueron de terminología y 7 de idiomática. Respecto a Apertium, se observó un total de 75 de errores con una mayor incidencia de errores de sinsentidos y sintaxis, con 26 y 24 errores respectivamente. En conjunto, los tres traductores automáticos tuvieron errores a nivel léxico y semántico.

Asimismo, es importante mencionar que ambos trabajos no comparten una metodología similar. En el caso de la investigación de Adán (2019) tuvo un enfoque cualitativo y se utilizó una ficha de análisis con un corpus de diferentes textos, mientras que, para esta investigación se tuvo un enfoque cuantitativo y se utilizó una prueba con una métrica para recolectar los datos necesarios. Para finalizar con el objetivo general, este tiene coincidencia en relación con los resultados con el trabajo de Adán (2019) ya que en ambos se encontró una mayor presencia de errores de terminología. No obstante, ambas investigaciones tienen diferentes enfoques de estudio, cualitativo y cuantitativo.

Para el primer objetivo específico de esta investigación, que consistió en analizar la calidad de la traducción automática neuronal del traductor DeepL utilizado en un estudio de traducción en Lima, 2023, se obtuvo un puntaje de calidad de traducción del 72.39 %. Este porcentaje sugiere que la calidad obtenida por este traductor se considera de calidad media-baja. En cuanto a los resultados, la

traducción generada por DeepL mostró una alta incidencia de errores terminológicos y de otro tipo, reduciendo de esta manera su puntaje final de calidad. Según Liu & Zhu (2023), el principal desafío de los sistemas de traducción automática neuronal radica en su insuficiente calidad de traducción, que no logra satisfacer las necesidades de los usuarios. Esta baja calidad se manifiesta en errores léxicos, uso incorrecto de tiempos verbales, fallos sintácticos y confusión en la estructura, entre otros aspectos.

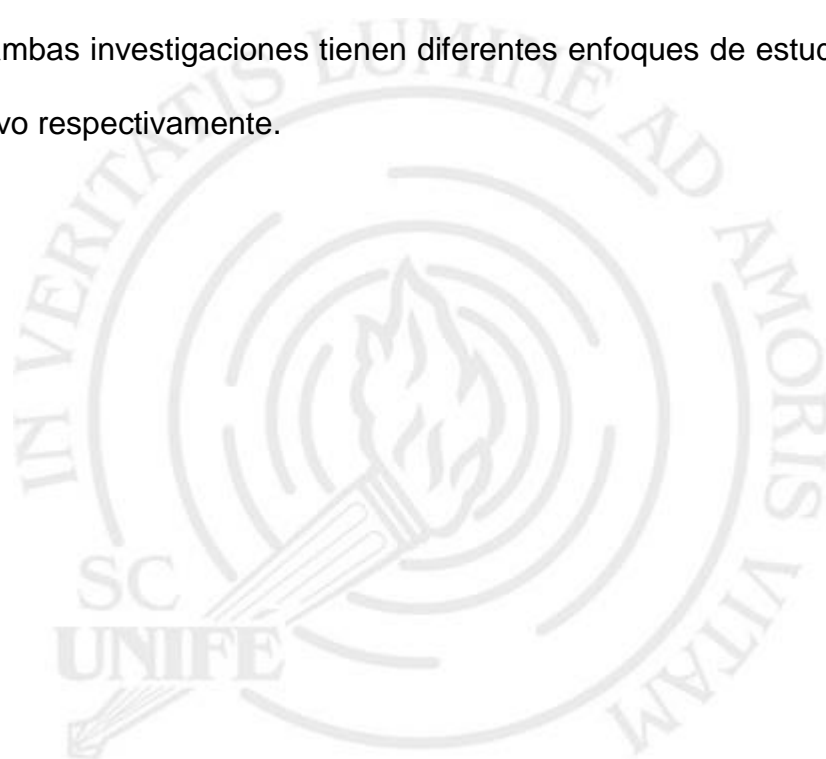
Los resultados de este trabajo coinciden parcialmente con la investigación de Llopis (2020), quién concluyó que la traducción de DeepL solo permitió entender las ideas principales del texto traducido, pero se encontraron errores en las cinco categorías establecidas (precisión, fluidez, terminología, convención local y estilo). El error de terminología fue el más frecuente en la traducción, mientras que, los errores relacionados con la convención local tuvieron un impacto menor en la traducción.

Asimismo, es importante mencionar que ambos trabajos no comparten una metodología similar, ya que el trabajo de Llopis (2020) se enfocó en un análisis cualitativo y utilizó un corpus con una cantidad de palabras específica, mientras que el presente trabajo tuvo un enfoque cuantitativo y se basó en pruebas para recolectar los datos necesarios. Para finalizar con el primer objetivo específico, este tiene coincidencia en relación con los resultados con el trabajo de Llopis (2020) ya que en ambos se encontró una mayor presencia de errores de terminología. No obstante, ambas investigaciones tienen diferentes enfoques de estudio, cualitativo y cuantitativo.

Para el segundo objetivo específico de esta investigación, que consistió en analizar la calidad de la traducción automática híbrida estadística-neuronal del traductor Yandex utilizado en un estudio de traducción en Lima, 2023, se obtuvo un puntaje de calidad de traducción del 76.62 %. Este porcentaje indica que la calidad obtenida por este traductor se considera de calidad media-baja. Los resultados proporcionados por Yandex mostraron un nivel alto de errores terminológicos, sintácticos y de otro tipo, lo cual redujo su puntaje final de calidad. Para Anugu y Ramesh (2020), el empleo de diversas técnicas en un conjunto de dos sistemas de traducción (traducción híbrida) mejora la calidad de traducción al intentar superar las limitaciones de uno de los sistemas mediante el uso del otro. No obstante, debido al amplio del vocabulario, la naturaleza del léxico y las irregularidades lingüísticas, la calidad de traducción no logra alcanzar el nivel de la traducción con presencia de un profesional.

Los resultados de este trabajo coinciden con la investigación de Rodríguez (2021), quien señaló que la traducción de Yandex, a pesar de haber tenido una traducción general satisfactoria alcanzando un puntaje de 69.79 según la evaluación BLEU, presentó errores principalmente en la parte terminológica. De igual manera, al realizar el análisis de Google Translate, el cual logró un puntaje de 69.39, se encontraron errores de elección de tiempos verbales y errores de terminología. No obstante, con los motores elaborados de traducción MTUOC-Translator se observó un cambio positivo en el puntaje de la traducción automática neuronal, consiguiendo así mejores resultados de traducción.

Asimismo, es importante mencionar que ambos trabajos no comparten una metodología similar. En el caso de la investigación de Rodríguez (2021) tuvo un enfoque cualitativo y se utilizó un análisis comparativo con un corpus de diferentes textos, mientras que, para esta investigación se tuvo un enfoque cuantitativo y se utilizó una prueba con una métrica para recolectar los datos necesarios. Para finalizar con el segundo objetivo específico, este tiene coincidencia en relación con los resultados con el trabajo de Rodríguez (2021) ya que en ambos se encontró una presencia relevante de errores de terminología para el traductor Yandex. No obstante, ambas investigaciones tienen diferentes enfoques de estudio, cualitativo y cuantitativo respectivamente.



CAPÍTULO VI. CONCLUSIONES

En esta investigación, tras la evaluación de los datos recopilados y la elaboración de las discusiones correspondientes, se han alcanzado las siguientes conclusiones, las cuales se alinean con el objetivo general y los objetivos específicos planteados.

1. Respecto al objetivo general, el cual fue evaluar la calidad de dos traductores automáticos utilizados en un estudio de traducción en Lima, 2023, se concluyó que, en promedio, la calidad de traducción de ambos traductores automáticos fue del 74.50 %. Este porcentaje de calidad se clasifica como medio-bajo, dado que para que un texto logre una alta calidad los promedios varían del 90 % a 100 %. En este caso, ambos traductores presentaron un gran número de errores terminológicos graves y también errores sintácticos, lo cual redujo el promedio final de la calidad.

2. Respecto al primer objetivo específico, el cual fue analizar la calidad de la traducción automática neuronal del traductor DeepL, utilizado en un estudio de traducción en Lima, 2023, se concluyó que, en promedio, la calidad de traducción del traductor automático fue del 72.39 %. Este porcentaje de calidad se clasifica como medio-bajo, dado que para que un texto logre una alta calidad los promedios varían del 90 % a 100 %. Ambas pruebas presentaron errores terminológicos, lo cual redujo en gran medida el porcentaje final de calidad.

3. Respecto al segundo objetivo específico, el cual fue analizar la calidad de la traducción automática híbrida estadística-neuronal del traductor Yandex,

utilizado en un estudio de traducción en Lima, 2023, se concluyó que, en promedio, la calidad de traducción del traductor automático fue del 76.62 %. Este porcentaje de calidad se clasifica como medio-bajo, dado que para que un texto logre una alta calidad los promedios varían del 90 % a 100 %. En ambas pruebas, los revisores encontraron un gran número de errores terminológicos y errores sintácticos, lo que causó una reducción en el porcentaje final de calidad.



CAPÍTULO VII. RECOMENDACIONES

En este apartado, después de detallar el análisis de los resultados y presentar las conclusiones correspondientes, se brindan algunas recomendaciones con la finalidad de promover y mejorar el análisis de la calidad de traducción.

Para futuras investigaciones:

1. Incentivar más estudios con un enfoque cuantitativo teniendo en cuenta el uso de una muestra más amplia para así lograr aumentar los resultados y los conocimientos sobre la calidad y los errores de la traducción automática. De este modo, la base de datos respecto a este tema será más amplia.

Para las Escuelas de Traducción e Interpretación:

2. Incluir en el Currículo la preedición con relación a la traducción automática, lo cual es un tema asociado a este trabajo de investigación y es muy importante ampliar las investigaciones al respecto. Este tema ayudaría a mejorar la calidad de las traducciones y más aún en el caso de los traductores automáticos.

Para los traductores:

3. Promover el uso de métricas de evaluación de calidad en los diversos ámbitos de la traducción, así como en diferentes áreas temáticas. Es importante tener un método de evaluación para comprobar la calidad que se tiene en los proyectos de traducción. Se invita a todo profesional relacionado

al rubro de la traducción implementar el uso de estas métricas para obtener un mejor producto de traducción.



REFERENCIAS

- Adán, M. (2019). *Estudio comparativo de tres traductores automático en línea: DeepL, Yandex y Apertium*. [Trabajo de fin de grado]. Archivo digital.
<https://repositorio.comillas.edu/xmlui/handle/11531/31624>
- Anugu, A., & Ramesh, G. (2020). A Survey on Hybrid Machine Translation. *E3S Web of Conferences*, 184, 01061.
<https://doi.org/10.1051/e3sconf/202018401061>
- Aguilar, E. (2019). *Evaluación de la calidad del traductor automático DeepL del inglés al español en textos especializados, Lima, 2019* [Tesis de licenciatura, Universidad Cesar Vallejo]. Archivo digital.
<https://hdl.handle.net/20.500.12692/39871>
- Alawneh, M. F., & Sembok, T. M. (2011). Rule-Based and Example-Based Machine Translation from English to Arabic. *2011 Sixth International Conference on Bio-Inspired Computing: Theories and Applications*.
<https://doi.org/10.1109/bic-ta.2011.76>
- Bermúdez, C. (2020). *Estudio y análisis de la traducción automática de sitios web médicos: Propuesta de posesición de un estudio de caso* [Trabajo de fin de grado]. Archivo digital.
<http://hdl.handle.net/10481/62583>

- Bolaño, R., Ramos, D. (2021). *Errores de traducción automática del inglés al español de una página web de noticias, Lima, 2021* [Tesis de licenciatura, Universidad Cesar Vallejo]. Archivo digital
<https://hdl.handle.net/20.500.12692/77923>
- Calzolari, N., McNaught, J., Palmer, M., Zampolli, A.(eds.). 2003. *ISLE Final Report. ISLE Deliverable D14.2, Pisa, 24 pp.*
http://www.ilc.cnr.it/EAGLES96/isle/ISLE_D14.2.zip
- Colegio de Traductores del Perú. (2017). Código de Ética.
<https://www.colegiotraductores.org.pe/wp-content/uploads/2018/04/Código-de-Ética-revisado-AGE-2017.pdf>
- Córdoba, I., Rico, C., Ortiz, M., Doquin, A., Arevalillo, J., Arcedillo, M., y Cabero, J. (2015). *Estudio de viabilidad para la implantación de la traducción automática en la empresa VITAE*. Universidad europea ABACUS.
<http://hdl.handle.net/11268/4362>
- Comparin, L. (2017). *Quality in machine translation and human post-editing : error annotation and specifications*. [Trabajo de maestría]. Archivo digital.
https://repositorio.ul.pt/bitstream/10451/27969/1/ulfl233806_tm.pdf
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches*. Fifth edition. Los Angeles, SAGE.
https://spada.uns.ac.id/pluginfile.php/510378/mod_resource/content/1/creswell.pdf

Diz, E. (2016). *Estadística básica, introducción a la estadística con R*. Bogotá, Colombia: Ediciones de la U. Obtenido de <https://bit.ly/3kYQPz0>

Esteban, S. (2020) *Evaluación de la calidad de la traducción de motores de traducción automática neuronal en textos del ámbito jurídico Comparación del rendimiento de Google Translate y Microsoft Translator en la combinación lingüística EN > CA* [Tesis de maestría, Universitat Autònoma de Barcelona]. Archivo digital. <https://ddd.uab.cat/record/249935>

Hernández-Sampieri, R. & Mendoza, C (2018). Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta. *Revista Universitaria Digital De Ciencias Sociales (RUDICS)*, 10(18), 92–95.

<https://doi.org/10.22201/fesc.20072236e.2019.10.18.6>

House, J. (2014). Translation Quality Assessment: Past and present. In *Palgrave Macmillan UK eBooks* (pp. 241–264).

https://doi.org/10.1057/9781137025487_13

Hurtado, A. (2001). *Traducción y Traductología. Introducción a la traductología*. Madrid, España, Ediciones Cátedra.

Hutchins, J. (1995) Machine translation: a brief history. En E.F.K. Koerner & R.E. Asher (eds.) *Concise history of the language sciences: from the Sumerians to the cognitivists* (pp. 431-445). Oxford: Pergamon Press.
<https://books.google.com/books?id=VCqLBQAAQBAJ&printsec=frontcover>

Koby, G., Fields, P., Hague, D., Lommel, A., & Melby, A. (2014) Defining Translation Quality. *Tradumàtica: Tecnologies De La Traducció*, 12, 413.
<https://doi.org/10.5565/rev/tradumatica.76>

Koskenniemi, K., Linden, K., Carlson, L., Vainio, M., Arppe, A., Lennes, M., Westerlund, H., Hyvarinen, M., Bartis, I., Nuolijarvi, P., and Piehl, A. (2012). *Suomen kieli digitaalisella aikakaudella – The Finnish Language in the Digital Age*. META-NET White Paper Series (Series Editors)
<http://www.meta-net.eu/whitepapers/overview>

Lai, H. (2023). *Análisis de las características de los errores de traducción automática neuronal y su posesión: Un estudio de las traducciones (del español al chino) de un artículo periodístico* [Trabajo de fin de Máster, Universitat Autònoma de Barcelona]. Archivo digital.
https://ddd.uab.cat/pub/trerecpro/2023/289027/HaohongLai_TFM2023.pdf

Liu, S., & Zhu, W. (2023). An analysis of the evaluation of the translation quality of neural machine translation application systems. *Applied Artificial Intelligence*, 37(1).

<https://doi.org/10.1080/08839514.2023.2214460>

Llopis, M. (2020). *La traducción automática aplicada a textos científicos: análisis de la traducción de un artículo científico generada por DeepL* [Trabajo de fin de grado, Universitat Jaume]. Archivo digital

<http://hdl.handle.net/10234/189524>

Maldonado González, M. C., & Liébana González, M. (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, 189–212. <https://doi.org/10.5209/clac.77002>

Martínez, R. (2014). A deeper look into metrics for translation quality assessment (TQA): a case study. *Miscelánea: A Journal of English and American Studies*, 49: 73-93.

<http://www.miscelaneajournal.net/index.php/misc/article/view/170/126>

Moorkens, J., Castilho, S., Gaspari, F., & Doherty, S. (Eds.). (2018). Translation Quality Assessment. *Machine Translation: Technologies and Applications*.

<https://doi.org/10.1007/978-3-319-91241-7>

Mossop, B., Hong, J. H., & Teixeira, C. R. (2019). Revising and editing for translators. In *Routledge eBooks*.

<https://doi.org/10.4324/9781315158990>

Norma Internacional ISO 9000. (2005). Sistemas de gestión de la calidad.

Fundamentos y vocabulario.

<https://www.iso.org/obp/ui/es/#iso:std:iso:9000:ed-3:v1:es>

Rico, C. (2017). La formación de traductores en Traducción Automática. *Tradumàtica. Tecnologies de la Traducció*, 15, 75-96.

<https://doi.org/10.5565/rev/tradumatica.200>

Rodríguez del Rosario, C (2021). *Creación de motores de traducción automática (estadística y neuronal) inglés-español especializados en el campo de la aviación con la herramienta MTUOC* [Trabajo de Fin de Máster, Universitat Oberta de Catalunya, Facultad de Artes y Humanidades. Barcelona, España]. Archivo digital

<https://openaccess.uoc.edu/bitstream/10609/133768/7/crodriguezdelrTFM0721memoria.pdf>

Rubio, N. (2022). *La calidad de los procesos de traducción automática y posesión en un producto audiovisual. El caso de la serie El juego del calamar* [Trabajo de fin de Máster, Universitat de Vic - Universitat Central de Catalunya]. Archivo digital.

http://repositori.uvic.cat/bitstream/handle/10854/7407/trealu_a2022_rubio_noemi_calidad_procesos.pdf?sequence=1&isAllowed=y

SAE J2450 (2001). *SAE J2450 Translation Quality Metric Task Force, Quality Metric for Language Translation of Service Information*

<https://www.sae.org/standardsdev/j2450p1.htm>

- Temnikova, I. (2010). Cognitive Evaluation Approach for a Controlled Language Post-Editing Experiment. *Proceedings Of The Seventh International Conference On Language Resources And Evaluation (LREC'10)*, 17-23
http://www.lrec-conf.org/proceedings/lrec2010/pdf/437_Paper.pdf
- Valls, W., Román, V., Chica, C., Salgado, G. (2017). *La calidad de servicio: vía segura Para Alcanzar la competitividad (1° Edición)*. Editorial Mar Abierto.
https://issuu.com/marabiertouleam/docs/la_calidad_del_servicio_wtest/33
- Vilar, D., Xu, J., D'Haro, L. F., & Ney, H. (2006). Error analysis of statistical machine translation output. *Proceedings Of The Fifth International Conference On Language Resources And Evaluation (LREC'06)*, 697-702. <https://www.dfki.de/~davi01/papers/vilar:lrec2006:errorAnalysis.pdf>
- Warwal, S. (2015). Translation process and problem of translation in world classics. *An International Multidisciplinary Research e-Journal*, 1(1), 125-133.
https://www.researchgate.net/publication/316655699_TRANSLATION_PROCESS_AND_PROBLEM_OF_TRANSLATION_IN_WORLD_CLASSICS

ANEXOS

Anexo 1: Matriz de consistencia

Problema General	Objetivos	Hipótesis	Categoría	Subcategoría	Metodología
¿Cuál es la calidad de dos traductores automáticos utilizados en un estudio de traducción en Lima, 2023?	<p>Objetivo general:</p> <p>Evaluar la calidad de dos traductores automáticos utilizados en un estudio de traducción en Lima, 2023.</p> <p>Objetivos específicos:</p> <p>Analizar la calidad de la traducción automática neuronal del traductor DeepL utilizado en un estudio de traducción en Lima, 2023.</p> <p>Analizar la calidad de la traducción automática híbrida</p>	No aplica	La calidad	Error terminológico	<p>Enfoque de investigación: Cuantitativo</p> <p>Nivel de investigación: Exploratorio</p> <p>Tipo de investigación: Básica</p> <p>Diseño de investigación: Transeccional exploratorio</p>
				Error sintáctico	
				Error de omisión	
				Error morfosintáctico o de concordancia	
				Error ortográfico	
				Error de puntuación	

	estadística-neuronal del traductor Yandex utilizado en un estudio de traducción en Lima, 2023.				
--	--	--	--	--	--

