

TRADUCCIÓN AUTOMÁTICA, LINGÜÍSTICA TEXTUAL Y ESTILÍSTICA CONTRASTIVA

Alejandro Rojas Walker

COLECCIÓN CILAMPA



TRADUCCIÓN AUTOMÁTICA,
LINGÜÍSTICA TEXTUAL
Y ESTILÍSTICA CONTRASTIVA

COLECCIÓN CILAMPA

TRADUCCIÓN AUTOMÁTICA,
LINGÜÍSTICA TEXTUAL
Y ESTILÍSTICA CONTRASTIVA

Alejandro Rojas Walker



EDICIONES

ESCUELA DE LITERATURA
Y CIENCIAS DEL LENGUAJE

Alejandro Rojas Walker,
Traducción automática, lingüística textual y estilística contrastiva. / Alejandro Rojas Walker. – 1.ª ed. –
Heredia: Ediciones Escuela de Literatura y Ciencias del Lenguaje, Universidad Nacional, 2022.
56 p.: 21x 14 cm.

ISBN 978-9930-9768-5-2

1. TRADUCCIÓN 2. ESTUDIOS DE TRADUCCIÓN E INTERPRETACIÓN



COLECCIÓN CILAMPA

Traducción automática, lingüística textual y estilística contrastiva
© Alejandro Rojas Walker

Primera edición, 2022
ISBN: 978-9930-9768-5-2

© Escuela de Literatura y Ciencias del Lenguaje, Universidad Nacional, Costa Rica
Teléfono: (506) 2562-4051
Portal electrónico: <http://www.literatura.una.ac.cr/>
Heredia, Costa Rica

Corrección filológica: *Sherry E. Gapper, Carlos Francisco Monge y Francisco Vargas Gómez*
Diseño preliminar: *Comisión Editorial de la ELCL*, con la colaboración de *Marta Lucía Gómez*
Diseño definitivo: *Daniela Hernández*
Diagramación: *Francisco Vargas Gómez*
Control de calidad: *Sherry E. Gapper, Carlos Francisco Monge y Francisco Vargas Gómez*

Prohibida la reproducción total o parcial. Todos los derechos reservados.
Hecho el depósito de ley.

Impreso bajo demanda en la Escuela de Literatura y Ciencias del Lenguaje,
Universidad Nacional (Heredia, Costa Rica).

ÍNDICE

Prólogo.....	11
INTRODUCCIÓN	13
1. ANTECEDENTES LOCALES.....	17
Traducción de estilo en textos científico-médicos	17
Análisis estilístico contrastivo y traducción automática	19
2. LA TRADUCCIÓN AUTOMÁTICA	23
Breve historia.....	23
Definición	25
Tipos.....	26
La traducción automática basada en reglas	26
La traducción automática estadística.....	26
La traducción automática neuronal.....	27
Ventajas, desventajas y retos de la traducción automática	28
Posedición y traducción automática	31
3. LA LINGÜÍSTICA TEXTUAL Y LOS GÉNEROS TEXTUALES	33
4. EL ESTILO Y LA ESTILÍSTICA CONTRASTIVA	37
Indicadores de estilo	38
Estilística contrastiva	39
5. LA TAN, LA LINGÜÍSTICA TEXTUAL Y LA ESTILÍSTICA CONTRASTIVA: APLICACIÓN Y UTILIDAD PRÁCTICA	41
Índole del experimento	41
Resultados del experimento	51
BIBLIOGRAFÍA.....	53

PRÓLOGO

Los Estudios de Traducción e Interpretación (ETI) en la Escuela de Literatura y Ciencias del Lenguaje de la Universidad Nacional (Costa Rica) cumplen en 2022 treinta años de desarrollo ininterrumpido. A lo largo del periodo, tanto el grupo de académicos que acogió esta disciplina como su principal área de interés, como los estudiantes de traducción de la institución, han atendido una gran diversidad de asuntos relacionados con el fenómeno en cuestión¹.

Desde luego, en muchos casos el trabajo se ha centrado en aspectos que podrían considerarse «tradicionales» (la traducción de elementos culturales en textos literarios, o los tocantes a la traducción de terminología especializada en textos técnicos o científicos). Sin embargo, en no pocas ocasiones se han tratado de igual manera algunos de los más novedosos aspectos disciplinares. Son notables las iniciativas orientadas a analizar diferentes sistemas o herramientas automatizadas de traducción, principalmente en sus aplicaciones prácticas. Por lo general, tales proyectos se ejecutan en el marco de los trabajos finales de graduación que desarrollan los estudiantes de la Maestría Profesional en Traducción Inglés-Español (UNA).

Como se expone más adelante en estas mismas páginas, tanto la traducción asistida por computadora como la automática han sido objeto de estudio en la UNA, al menos desde 2003, cuando se presentó un proyecto de graduación que buscaba valorar la calidad de una traducción producida por medio de un sistema automático en contraposición a la elaborada por un traductor humano. Con aquel visionario trabajo como ejemplo, el más reciente de estos proyectos dedicados a la traducción por computadora se centró en la denominada traducción automática neuronal, variante que en la actualidad está a la vanguardia de los sistemas automatizados de traducción. Ha sido precisamente esta última iniciativa la que dio origen al presente volumen.

Fiel a su cometido («la edición [...] de obras de referencia y demostrada solvencia académica [y] la creación de una interesante oferta de materiales didácticos»), la Colección Cilampa ofrece al lector interesado este tomo: un texto introductorio a la traducción automática, tema que se conjuga además

1 Al respecto, véase Gapper, Sherry E. (2018). Veinte años de estudios aplicados sobre traducción. En Sherry E. Gapper (ed.). *Memoria del V Congreso Internacional de Lingüística aplicada*. Heredia: Ediciones Escuela de Literatura y Ciencias del Lenguaje (pp. 221-230).

con nociones básicas de lingüística textual y estilística. Con tal objetivo y tras revisar una serie de antecedentes pertinentes, se exploran en la segunda sección de este cuadernillo asuntos cardinales sobre esta modalidad de traducción (su historia, definición, tipos, ventajas, desventajas y relación con la posesición). A partir de ellos, en los apartados tercero y cuarto, se desarrollan ideas básicas sobre lingüística del texto, géneros textuales, estilo y estilística contrastiva. Mediante la quinta y última parte, se ejemplifican las aplicaciones y utilidad prácticas que la traducción automática y los procesos de posesición pueden suponer para los investigadores y profesionales de la traducción.

FRANCISCO VARGAS GÓMEZ

INTRODUCCIÓN

A pesar de los notables avances acaecidos en el campo de la traducción automática durante los dos últimos decenios, en Costa Rica persiste un ambiente de desconfianza y rechazo a esta forma de traducción. A manera de ejemplo, se puede citar una publicación cuyo título resulta ilustrativo por sí solo de la situación recién referida: «Why Machine Translation is Not Good For Business in Costa Rica» (López, 2015). El artículo gira en torno a un experimento realizado por la empresa de traducción Verba Link e implicó la comparación de las traducciones producidas por un traductor humano y por un conocido sistema de traducción automática, con el objetivo de determinar cuál de las dos traducciones obtenía un mejor resultado. Ambas traducciones fueron cuidadosamente evaluadas por una especialista en control de calidad que actuó como jurado durante el experimento.

Evaluadas ambas traducciones, la especialista identificó las fortalezas y debilidades de cada una. Se concluyó que la traducción humana contaba con una precisión excelente e interpretaba el texto original a la perfección al tiempo que no se identificaron errores en ella. Por otro lado, sobre la traducción automática se señaló que la aplicación informática había realizado un buen trabajo al proporcionar lo esencial del texto. No obstante, también se dictaminó que su utilidad se limitaba a servir como un medio para realizar una lectura rápida del texto original o como instrumento en casos en que el lector necesite obtener una idea general del texto en lengua original. Además, se detectó que las oraciones que produjo el traductor automático eran disjuntas y que la gramática, legibilidad y fluidez de la traducción debían corregirse para que el texto tuviera sentido. Finalmente, se concluyó que el traductor automático resultaba una herramienta funcional en ciertos casos, cuyas limitaciones no la hacen realmente útil para entornos o actividades de tipo profesional, pues una traducción de calidad debe ser precisa y, en general, no debe parecer una traducción. Por tales razones, quienes participaron en el experimento se pronunciaron a favor de la traducción humana como el principal recurso en lo que compete a la traducción profesional.

Situaciones como la descrita acentúan los prejuicios sobre la traducción automática extendidos entre los potenciales usuarios y el público en general. Así, cuando se habla de traducción automática, se piensa en aquellos «primeros» traductores automáticos basados en reglas gramaticales que traducían palabra por palabra, dando como resultado, en el mejor de los casos, una traducción

poco natural desde el punto de vista del hablante promedio de la lengua meta. También se suele pensar que las deficiencias que tales sistemas presentaban en el pasado —en términos de la calidad del producto final con respecto a la calidad de las traducciones que podía ofrecer un traductor humano— persisten en las actuales herramientas de traducción automática.

Pese a que actualmente los traductores automáticos generan traducciones similares cada vez más a las producidas por traductores humanos, lo cierto es que aún enfrentan grandes desafíos. Uno de ellos es la incapacidad de los sistemas informáticos para distinguir el estilo del texto de partida del estilo de los textos paralelos que se producen en la cultura meta, como lo haría normalmente un traductor humano, para luego tratar de reproducirlo en el texto de salida. Esta característica ocasiona que las computadoras produzcan traducciones en apariencia adecuadas, pero que, para el ojo entrenado, pueden distanciarse del estilo deseado para el tipo y género textual que se esté traduciendo.

El panorama expuesto ha hecho que se pasen por alto los adelantos en el campo de la lingüística computacional que han tomado lugar a lo largo de los últimos lustros y que le ha permitido a la traducción automática mejorar más allá de lo esperado. Por tanto, incluso con el fácil acceso a la información que caracteriza a la sociedad contemporánea, muchos traductores desconocen la traducción automática y las ventajas que puede ofrecerles en su actividad profesional.

El grado de desinformación que parece existir en el país alrededor de la traducción automática, aunado a la falta de investigaciones sobre su efectividad, a la propagación de mitos y al rechazo prematuro de muchos traductores hacia esta tecnología, justifican una reflexión como primer acercamiento al verdadero potencial que podría tener, sobre todo, en el ámbito profesional. Tal es el objetivo de estas páginas, a lo largo de las cuales se prestará especial atención a sus aplicaciones con respecto a los procesos de posesición y a asuntos de orden estilístico con respecto a un género textual dado. En sus aspiraciones más altas, se pretende que este volumen contribuya a que tanto los traductores como las empresas en el país comprendan que la traducción automática no debe verse como una enemiga ni como un recurso inútil o muy limitado, sino como una aliada que puede contribuir en gran medida a realizar el trabajo de forma más eficiente.

Con ello en mente, en las diferentes secciones que componen este cuadernillo se plantearán asuntos introductorios con respecto al tema. Tras una primera sección, que repasa diferentes fuentes bibliográficas a manera de antecedentes, se plantean diversos asuntos de particular importancia sobre los sistemas de traducción automática actuales, entre los que destacan los diferentes tipos, sus ventajas y sus desventajas. Las secciones siguientes tratan asuntos propios de la lingüística textual y de la estilística contrastiva. En el

último apartado se ofrece un ejemplo de aplicación de todas las nociones antes referidas, mediante el cual se pretende ilustrar las bondades de los actuales sistemas de traducción automática.

1. ANTECEDENTES LOCALES

Desde sus inicios en la década de 1930 la traducción automática ha sido tema de interés para investigadores y estudiosos en diversas partes del mundo. En Costa Rica ha motivado solamente algunas pocas iniciativas relacionadas con dicho objeto de estudio. Lo mismo podría decirse de temas como la posesición y la estilística contrastiva, al menos en lo que respecta a sus aplicaciones a la traducción. Por suerte, la situación es diferente si hablamos de la traducción de textos especializados de corte científico-médico, género al que se hará alusión directa en estas páginas. Así, son varios los trabajos relacionados con la traducción de resúmenes o *abstracts* de textos especializados de diversos campos —como la psicología, la medicina y las biociencias—, aunque lo cierto es que ninguno se enfoca realmente en aspectos estilísticos relativos a estos géneros textuales ni tampoco en las aplicaciones de los sistemas automatizados a su traducción.

TRADUCCIÓN DE ESTILO EN TEXTOS CIENTÍFICO-MÉDICOS

Adriana Castro Benítez (2012), en *Oportunidades de mejora para la traducción de resúmenes de artículos de revistas científico-médicas costarricenses: el caso de la “Revista Médica de la Universidad de Costa Rica”*, identifica ciertos obstáculos en el proceso de elaboración y traducción del resumen de revistas científico-médicas costarricenses. Castro centra su análisis en la traducción de la macro y microestructura, con el fin de identificar maneras que permitan cumplir con los estándares de calidad establecidos. Al concluir el estudio, la autora resalta la importancia de establecer la responsabilidad del traductor durante la elaboración y traducción de resúmenes, para que la calidad concuerde con lo que al respecto exigen las normas establecidas (p. ix).

Luis Fernando González Salas (2015), en *La calidad comunicativa de la traducción al inglés de los resúmenes de artículos biomédicos costarricenses: ¿el traductor científico o el científico traductor?*, evalúa las particularidades de la traducción al inglés de los resúmenes de revistas biomédicas realizada por traductores costarricenses con diferente nivel de conocimiento científico o traductológico. Concluye que las diferencias en la calidad lingüística y comunicativa de las traducciones de los resúmenes de revistas biomédicas parecen

guardar relación con el perfil técnico-académico de los traductores y plantean la necesidad de establecer estrategias de mejora para garantizar que las traducciones cumplan con las exigencias establecidas en el ámbito internacional (p. ix).

En cuanto a la constitución de las sociedades modernas, Castoriadis sostiene que las mismas están repletas de significaciones menores, menos relevantes, que cambian constantemente para que las significaciones mayores y centrales no cambien nunca, incluso cuando estas ya hayan perdido el sentido con el que emergieron (Castoriadis, 1998). En consecuencia, existe en las sociedades modernas una suerte de repetición. Esta constante negación de la creación conlleva a un ocultamiento de lo imaginario y hace que la historia y las sociedades mismas se perciban como gobernadas por una instancia superior a ellas, y que lo imaginario se vea como falso o como lo opuesto a la «realidad».

Ana Luisa Monge Naranjo (2016) se centra en la traducción, del español al inglés, de resúmenes de investigaciones del campo de la psicología y elabora una propuesta preliminar de mejora de la función comunicativa de este tipo de resúmenes con base en el análisis de dos rasgos estilísticos: la proporción de verbos y la extensión de las oraciones. Concluye su trabajo, titulado *Aspectos estilísticos en la traducción de resúmenes de investigaciones de psicología: la proporción de verbos y la extensión de las oraciones*, con una propuesta para los traductores de resúmenes del español al inglés y subraya la necesidad de replicar el estudio con muestras más amplias para generar datos estadísticamente significativos (p. viii).

Adriana Bolaños Carpio (2014) en *Consideraciones bioéticas en la traducción del protocolo de estudio clínico Variability of Parkinson's Disease Biomarker Analytes*, analiza las consideraciones que los traductores deben tener en cuenta para resguardar los principios bioéticos de quienes participan en estudios clínicos. Al concluir su investigación, Bolaños determina que los problemas de traducción encontrados afectan en diferente grado los derechos de los participantes en un estudio clínico y destaca que existen convenciones específicas para este género textual que los traductores deben conocer con el fin de generar traducciones adecuadas (p. vi).

Las iniciativas centradas en la estilística y su relación con la traducción de textos especializados no son recurrentes en el ámbito local y, cuando se han dado, se han limitado a abordar solo algunos de los tantos rasgos estilísticos que se pueden considerar a la hora de traducir un texto especializado: además de analizar la proporción de verbos y la extensión de las oraciones —como en el caso de Monge—, también merecen atención, por ejemplo, el uso de las oraciones según su complejidad o modo, las categorías gramaticales (como los sustantivos, adjetivos y adverbios), los primeros elementos de las oraciones, el

grado de densidad léxica, el nivel promedio de legibilidad y la incidencia de voces pasivas, entre otros aspectos.

ANÁLISIS ESTILÍSTICO CONTRASTIVO Y TRADUCCIÓN AUTOMÁTICA

La adecuación estilística asociada a la traducción automática no es una novedad. Ha sido una respuesta a la necesidad de mejorar la calidad de las traducciones automáticas desde sus inicios. En el contexto canadiense, por ejemplo, existe una cierta tradición investigativa sobre el tema. Para 1994, Chrysanne DiMarco y Keith Mah ya habían desarrollado y presentado —por medio de un artículo titulado «A Model of Comparative Stylistics for Machine Translation»— un modelo formal de estilística comparada que podía utilizarse como componente de una teoría general computacional del estilo. Tal modelo puede implementarse en un sistema de traducción automática, permitiéndole tomar decisiones más informadas sobre las elecciones de traducción y sus potenciales efectos estilísticos (p. 21).

En España, Antonia Montes Fernández (2007), en «La textología contrastiva – método de análisis para la traducción de textos publicitarios», explica la forma en que se puede aplicar un enfoque basado en la textología contrastiva al estudio de los anuncios publicitarios impresos, con el propósito de poner de relieve las convenciones textuales y las prácticas discursivas que rigen dicho género textual. Montes muestra, por medio de un análisis cualitativo, los recursos prototípicos de las prácticas discursivas en alemán y en español, basada en un corpus significativo de anuncios publicitarios, y examina algunas consecuencias que derivan de aplicar el enfoque contrastivo a la traducción del género textual en cuestión (p. 150). Si bien la investigación de Montes no aborda de forma directa el tema de la estilística ni el de la traducción automática, sí propone un modelo basado en la comparación de elementos convencionales propios de un género textual en lenguas diferentes cuya relación con las nociones de estilo aplicadas a la traducción resulta evidente.

Más recientemente, también en España, en su tesis doctoral *Lenguaje controlado aplicado a la traducción automática de prospectos farmacéuticos* (2015), Ángel Luis Seoane Vicente identifica las dificultades que el sistema Google Translator Toolkit encuentra al traducir prospectos farmacéuticos del inglés al español peninsular. Una vez identificadas las dificultades y definidas las reglas del lenguaje controlado, Seoane traduce, en idénticas condiciones, textos controlados a fin de contrastar los resultados con los obtenidos en la traducción inicial no controlada. El investigador concluye que el uso de lenguajes controlados es positivo si se cumplen los requisitos operativos de partida planteados en su estudio: «de un lado, un sistema de traducción auto-

mático estadístico muy potente; del otro, un género textual donde abundan las estructuras sintácticas y frases prefijadas» (Resumen). Cabe destacar que, en su trabajo, Seoane alude a un sistema de traducción automática estadística, y no a un traductor automático neuronal, considerados en la actualidad como la vanguardia de la traducción automática.

En «Aplicaciones del análisis estilístico contrastivo-funcional para la traducción de géneros digitales», Alicia Bolaños Medina (2000) señala que el análisis estilístico contrastivo textual y funcional proporciona información sobre las convenciones de un género dado en las dos lenguas de trabajo (el español y el inglés, en su caso), útil para los traductores al redactar un texto que funcione plenamente en la lengua meta (p. 39). Partiendo de las teorías de la estilística funcional rusa y apoyándose en los conceptos de estilo y géneros funcionales, Bolaños Medina elabora una metodología de análisis que aúna parámetros pragmáticos, comunicativos y lingüísticos. Su objetivo era aplicar el análisis estilístico contrastivo a un corpus de páginas de sitios web con el fin de establecer conclusiones sobre la especificidad de este género textual electrónico y evaluar la utilidad del método analítico para el traductor profesional (p. 39). El estudio de Bolaños Medina llama la atención principalmente porque toma las teorías de la estilística contrastiva funcional y las integra en el análisis de un corpus para descubrir las diferencias y similitudes de los géneros textuales electrónicos redactados en inglés y español.

En Jamaica, Mariana González Boluda (2010) publica «Estudio comparativo de traductores automáticos: Systran, Reverso y Google», artículo en el cual evalúa tres programas de traducción automática disponibles en Internet y presenta, a través de un análisis de traducciones de oraciones y frases breves del inglés al español, las principales dificultades que enfrenta cada uno de estos sistemas de traducción automática. González analiza los problemas y determina cuál de los tres traductores automáticos ofrece mejores resultados (p. 187). A propósito de este estudio, debe decirse que, aunque se examinó el traductor de Google (que también se utiliza como ejemplo más adelante en este volumen), aquella valoración se realizó hace más de diez años, cuando aún era un sistema de tipo estadístico; por tanto, en aquel momento dicho sistema tenía una arquitectura distinta a la actual (de tipo neuronal).

De vuelta al ámbito nacional, Carlos Mora Campos (2003) realizó un estudio comparativo: *Un encuentro con la traducción automática. Análisis comparativo de un texto traducido bajo diferentes modalidades de traducción*. Presenta los resultados de un experimento que implicó traducir un mismo texto utilizando diferentes modalidades de traducción. Al inicio, establece algunas hipótesis que toman en consideración el uso de dos programas de traducción, la traducción tradicional y el conocimiento del traductor. Para poner a prueba sus hipótesis, Mora Campos proporciona un mismo texto a diferentes traductores, quienes emplean distintas modalidades de traducción, y luego analiza el resultado desde la perspectiva sintáctica y léxica para valorar la calidad de

cada una de las traducciones. Por último, analiza el tiempo que tomó realizar la traducción y su costo, para finalmente plantear conclusiones acerca del uso de los programas informáticos de traducción como herramientas de trabajo por parte del traductor (p. iv). Si bien el experimento de Mora Campos analiza la traducción automática en términos de calidad y rentabilidad, no entra en los terrenos de la posesición y, más importante aún, el experimento se realizó hace casi veinte años, cuando aún no habían tomado lugar algunos de los más importantes avances en el campo de la traducción automática.

Como nota final a este recuento bibliográfico, debe decirse que, aunque en otras latitudes sí se han realizado estudios sobre posesición y traducción automática, en el país no parece haberse desarrollado ninguna investigación de este tipo; es esta pues una de las razones que motivan la publicación de estas páginas. En la misma línea, hay que destacar que, si bien varias de las fuentes recién reseñadas se centran en asuntos de tipo estilístico y contrastivo, ninguna se propuso plantear formas para aprovechar las posibilidades que ofrecen los traductores automáticos actuales con respecto a tales asuntos. Ello, precisamente, constituye uno de los propósitos de las secciones que se desarrollan a continuación.

2. LA TRADUCCIÓN AUTOMÁTICA

BREVE HISTORIA

La traducción automática (TA) también conocida como *traducción por computadora, automatizada o instantánea*, no es el producto de iniciativas recientes, ni tampoco de desarrollos tecnológicos nuevos; lo cierto es que lleva entre nosotros más de setenta años. Sus orígenes se remontan a finales de los años cuarenta, a 1946 para ser precisos, cuando se presentó al mundo el ENIAC: el primer ordenador electrónico de propósito general de la historia y, con él, surgieron las primeras ideas de traducir una lengua natural por medio de una computadora. Ahora bien, al hablar de los orígenes de la traducción automática, es obligatorio mencionar a Warren Weaver, un informático estadounidense que trabajó como matemático durante la Segunda Guerra Mundial y que, en 1947, postuló la posibilidad de utilizar computadoras digitales para traducir documentos entre lenguas naturales (véase Gamal, 2017, p. 4).

Un ejemplo que ilustra mejor esta noción, que entrelaza y tensa lo material con lo simbólico, es el del uso de metáforas para describir el interior de una casa y cómo la casa misma se transforma a través de las experiencias que esas metáforas puedan sugerir. Así, el espacio objetivo de una casa —sus esquinas, pasillos, habitaciones, sótano, medidas— es mucho menos importante que los atributos poéticos que le puedan ser asignados y que generalmente son cualidades con un valor imaginario y figurativo que es posible nombrar y sentir. De esta forma, la casa puede estar encantada, puede ser acogedora, mágica o descrita como una prisión. El espacio, entonces, adquiere un sentido emocional e incluso racional a través de un proceso poético en el que los elementos vacantes o anónimos son transformados en lo que en realidad le da significado al espacio mismo (Said, 1978).

En 1949, Weaver propuso por medio de un memorándum varios métodos —basados en sus conocimientos de criptografía, estadística, teoría de la información, lógica y universales lingüísticos— que posteriormente incentivarían las investigaciones sobre la traducción automática en Estados Unidos (véase Hutchins, 2006, p. 376). En 1954, IBM, en colaboración con la Universidad de Georgetown, realizó la primera demostración relacionada con la confiabilidad de la traducción automática. Durante el experimento, fue posible traducir 250 palabras del ruso al inglés de 49 oraciones seleccionadas de forma cuidadosa. El éxito mediático de aquella demostración fue tal que el Ministerio de Defen-

sa de Estados Unidos y otros grupos invirtieron grandes sumas de dinero en proyectos para traducir del ruso, francés y alemán al inglés; a partir de entonces, también se establecieron las bases de reconocidos sistemas de traducción automática, como SYSTRAN y METAL (Gamal, 2017, p. 5).

Después de aquella demostración de 1954 y durante el decenio que va de 1956 a 1966, se acentuó el optimismo alrededor de la traducción automática. Fue una época en la cual, a pesar de que no se alcanzaron las metas deseadas, se produjeron grandes descubrimientos teóricos basados en investigaciones en el campo de la lingüística y la informática. Así, en 1964 se constituyó el ALPAC (Automatic Language Processing Advisory Committee), una entidad creada para evaluar la situación de la traducción automática, cuyo informe, publicado dos años más tarde, pondría fin a este periodo de grandes expectativas y daría paso a una época de desilusiones en Estados Unidos (Gamal, 2017, p. 5).

Según Hutchins (2006), el «informe ALPAC» concluía que la traducción automática era más lenta, de menor calidad y el doble de costosa en términos económicos que la traducción humana y, por lo tanto, no era recomendable continuar invirtiendo en su desarrollo (p. 377). Este mensaje caló en las empresas y organismos que financiaban los proyectos de investigación en el campo, lo que ocasionó que entre 1966 y 1976 no se realizaran estudios en Estados Unidos sobre el tema. Afortunadamente, el informe no tuvo el mismo impacto en espacios como Canadá o Europa, pues en estas regiones se continuó desarrollando la traducción automática. Tal fue el caso que en 1976 se presentó el sistema MÉTÉO en Canadá y, tiempo después, el sistema EUROTRA en Europa (Gamal, 2017, pp. 5-6).

Entre 1970 y 1990 se emprendieron numerosos proyectos a raíz del éxito de MÉTÉO. Al contrario de lo sucedido con las investigaciones llevadas a cabo durante la década de 1970 —que se centraban en una interlingua—, las iniciativas durante los ochenta adoptaron un enfoque basado en la transferencia en tres etapas, orientado a la sintaxis y fundamentado en las reglas gramaticales y de formalización del léxico que estaban influenciadas por las teorías lingüísticas de la época (Hutchins, 2006, p. 378). También durante la década de 1980 empezó el desarrollo de las herramientas de traducción asistidas por computadora, las cuales representaron un gran avance en el campo. Sin embargo, la traducción automática aún se enfrentaba a los mismos desafíos al intentar traducir un texto de un idioma a otro sin que se perdiera el significado (Gamal, 2017, p. 6). Posteriormente y desde mediados de la década siguiente, con la aparición del internet y la llegada de la globalización, los avances en el desarrollo de la traducción automática han sido sustanciales, en parte debido a la necesidad de las empresas de localizar productos y servicios, por lo que cada día se llevan a cabo más investigaciones que apuntan al mejoramiento de la tecnología disponible (Hutchins, 2006, p. 382).

DEFINICIÓN

El concepto de traducción automática (*machine translation* o MT, en inglés) la han definido por una gran cantidad de investigadores y empresas interesados en el estudio de este campo de la lingüística computacional. En las fuentes sobre el tema es posible encontrar una variedad de definiciones del término, entre las que destaca, por la variedad de ámbitos lingüísticos a los que hace referencia, la de Berner (2003), quien lo define de la siguiente manera:

Machine translation (MT) is the use of computer software to translate text or speech from one natural language into another. Like translation done by humans, MT does not simply involve substituting words in one language for another, but the application of complex linguistic knowledge: morphology, syntax, semantics, and understanding of concepts such as ambiguity. (p. 6)

Llamativa es la propuesta del diccionario en línea *Cambridge Dictionary*, mediante la cual se define el término como «the process of changing text from one language into another language using a computer» (s. f.). Si bien esta última definición no engloba la complejidad de la traducción automática a la que sí hace alusión la de Berner, lo cierto es que coinciden en que ambas la describen como un proceso que involucra el uso de computadoras o software especializados para el traslado de un texto de una lengua natural a otra *sin la intervención humana*.

Por último, cabe destacar la diferencia entre la traducción automática y las herramientas de traducción asistida por computadora —herramientas TAC, por sus siglas en español o *CAT tools*, como se les denomina en inglés— tales como las memorias de traducción. De acuerdo con Retana (2008), una memoria de traducción «[c]onsiste en una base de datos que almacena traducciones anteriores en segmentos originales con su correspondiente construcción equivalente», cuya función principal es «[...] asociar la oración que se está traduciendo con una exacta o muy semejante que se encuentra almacenada en la base de datos del sistema» (pp. 34-36). Como resulta evidente, una herramienta TAC no es lo mismo que la traducción automática, puesto que las herramientas TAC requieren dos componentes (un *software* y un traductor humano), uno de los cuales asiste al otro; en otras palabras, el *software* asiste al traductor o el traductor asiste al *software*. Diferente resulta el caso de la traducción automática, ya que estos sistemas ejecutan todo el proceso de manera autónoma; en otras palabras, la computadora lleva a cabo la traducción por sí sola sin asistencia del humano.

TIPOS

En los últimos años se han desarrollado diversas arquitecturas de traducción automática que, por sus características, ofrecen resultados distintos. Según Tertoolen (2012), los sistemas de traducción automática se clasifican según dos enfoques principales: el enfoque basado en reglas o normas y el basado en corpus (p. 142). Los enfoques basados en reglas se subdividen a su vez en tres sistemas vistos como tradicionales: los sistemas de traducción directa, los sistemas interlinguales y los sistemas de transferencia. Los basados en corpus, por otra parte —también conocidos como *enfoques empíricos*—, se subdividen en principio en dos tipos: la traducción automática basada en ejemplos (TABE) y la traducción automática estadística (TAE) (Tertoolen, 2012, pp. 142-143). A este último hay que agregar un tercer tipo de traducción automática de más reciente desarrollo: la traducción automática neuronal (TAN).

De acuerdo con Molina (2019), pese a que existen muchos tipos de traducción automática, actualmente los sistemas de traducción de este tipo más comunes son la traducción automática basada en reglas (TABR), la traducción automática estadística y la traducción automática neuronal (p. 6). A continuación, se expondrán brevemente las características principales y el funcionamiento de cada una de estas tres formas de traducción automática.

LA TRADUCCIÓN AUTOMÁTICA BASADA EN REGLAS

Tripathi y Sarkhel (2010) explican que la TABR (*Rule-Based Machine Translation* en inglés) construye reglas lingüísticas basadas en la información morfológica, sintáctica y semántica de la lengua de origen y la de llegada. Para ello, el sistema utiliza millones de diccionarios bilingües, aunque también pueden ser diccionarios monolingües y diccionarios especializados, con el objetivo de traducir textos con terminología específica para cada par de lenguas (p. 389). Estos sistemas primero analizan el texto a traducir y luego realizan una representación simbólica interna haciendo uso de recursos lingüísticos y estructuras gramaticales para recrear el texto, reemplazando las palabras del texto original con los equivalentes más apropiados en la lengua meta.

LA TRADUCCIÓN AUTOMÁTICA ESTADÍSTICA

Tripathi y Sarkhel señalan que la TAE (*Statistical Machine Translation* en inglés) fue introducida en 1949 por Warren Weaver. Estos sistemas utilizan métodos estadísticos para generar una versión traducida utilizando corpus monolingües y bilingües (2010, p. 390). Para la compañía pionera en sistemas de traducción automática SYSTRAN (2020), a diferencia de los sistemas TABR, la TAE ofrece un proceso de traducción más rápido; no obstante, su desempeño depende en gran medida de los corpus multilingües existentes, ya

que se necesita un mínimo de dos millones de palabras para un campo específico y aún más para una lengua en términos generales (párr. 8). Pese a que esta tecnología hace posible alcanzar un umbral de calidad aceptable, para la mayoría de las empresas resulta muy difícil hacerse con corpus multilingües tan grandes como para crear los modelos de traducción necesarios para que el sistema opere con resultados satisfactorios. Otra desventaja de la TAE es que requiere una gran capacidad por parte de los sistemas informáticos para poder ejecutar los modelos de traducción que permiten obtener un rendimiento aceptable (SYSTRAN, 2020, párr. 8).

LA TRADUCCIÓN AUTOMÁTICA NEURONAL

Según Forcada (2017), la TAN (*Neural Machine Translation*, en inglés) es una nueva variante de la traducción automática basada en corpus, entrenada con enormes corpus compuestos de pares de segmentos de la lengua de origen con sus respectivas traducciones; en otras palabras, se entrena con gigantescas memorias de traducción que contienen cientos de miles o incluso millones de unidades de traducción (p. 292). Coccozza (2018) agrega que los sistemas TAE y TAN comparten características en común. Por ejemplo, ambos sistemas se basan en datos y sus núcleos estructurales genéricos carecen de conocimiento lingüístico (p. 23). Por otra parte, la diferencia más notable entre estas dos tecnologías es que la TAN utiliza un procedimiento computacional distinto: el uso de redes neuronales, las cuales imitan el funcionamiento de las neuronas del cerebro humano, cuya salida o activación depende del estímulo que reciben de otras «neuronas» y de la fortaleza de las conexiones a través de las cuales transitan estos estímulos (Forcada, 2017, p. 292).

Según Coccozza, uno de los principales rasgos distintivos de la TAN es su capacidad de «aprender» gracias a las redes neuronales. Es particular en estos sistemas su capacidad de traducir oraciones completas y no solo frases o fragmentos, como lo hace la TAE basada en frases; por lo tanto, las redes neuronales pueden almacenar significado a nivel oracional (2018, pp. 23-26). Parra (2018) explica que una desventaja de los sistemas TAN es que, en el caso de frases muy extensas, suelen ofrecer peores resultados que los sistemas TAE (p. 26). De acuerdo con Forcada, estos sistemas son difíciles de entrenar, incluso más que los TAE, ya que necesitan enormes corpus paralelos que, por lo general, no están disponibles para traductores independientes o compañías de traducción pequeñas; ello sin mencionar que esta tecnología también requiere un *hardware* exclusivo y que los tiempos de entrenamiento, en ocasiones, pueden ser en muy extensos (2017, p. 302).

Pese a las desventajas, Parra señala que el potencial de la TAN es tal que ya se están entrenando motores que, además de textos, incorporan imágenes o incluso archivos multimedia con resultados muy prometedores (2018, p. 25). Por el momento, solo las grandes corporaciones como Amazon, Booking.com,

eBay, Google o Microsoft, pueden invertir en esta costosa tecnología, pero se espera que en un futuro no muy lejano este paradigma termine por desplazar a la TAE —que ha liderado el campo de la traducción automática por muchos años— y que cada vez más traductores y agencias de traducción logren incorporar esta herramienta en sus labores diarias (Parra, 2018, p. 26).

VENTAJAS, DESVENTAJAS Y RETOS DE LA TRADUCCIÓN AUTOMÁTICA

En general, la traducción automática supone numerosos beneficios, sobre todo desde el punto de vista profesional y tanto para el cliente como para el traductor, según explica Shift Languages (2020), una agencia de traducción argentina que ofrece servicios de traducción y posesición. Shift Languages subraya algunas desventajas; por ejemplo, debe tenerse en cuenta que no todos los géneros textuales se prestan para ser traducidos por medio de un traductor automático. De una traducción que requiera el sello oficial de un traductor certificado, no se admite la opción automática. Si el texto por traducir no está impreso o no se ha laborado en un archivo electrónico editable, será casi imposible pretraducirlo; será necesario digitalizarlo y convertirlo a un formato editable antes de iniciar el proceso, lo cual aumentaría el costo del encargo de traducción. Tampoco en el caso de textos de *marketing* y publicidad o literarios parece adecuado utilizar un sistema automatizado de traducción, ya que la traducción de tales géneros textuales requiere pensamiento creativo y, en múltiples ocasiones, transcreación por parte del traductor, por lo que lo aconsejable sería optar por un traductor humano. Además, Shift Languages también subraya asuntos relacionados con la confidencialidad como un potencial obstáculo, ya que el cliente debe estar enterado de que cuando solicita una traducción automática casi siempre la información se comparte con terceros, es decir, con los proveedores (Google o Microsoft, por ejemplo) de la interfaz de programación de aplicaciones (API, por sus siglas en inglés) que se utiliza para pretraducir los textos mediante conexión de internet. Por lo tanto, si la privacidad es un factor importante para el cliente, esta podría llegar a representar un inconveniente si se piensa utilizar una herramienta de traducción automática (párr. 2-6).

No menos problemática resulta la falta de capacidad de los sistemas de traducción automática para identificar contextos, situaciones comunicativas, públicos meta y especificaciones del encargo de traducción, entre otros factores extratextuales que los traductores humanos deben considerar a la hora de realizar su labor profesional. Ante la incapacidad de los sistemas para evaluar factores externos que caracterizan el contexto de producción de los textos y condicionan la forma final que deben adoptar, surgen diferentes problemas

que se trasladan a la traducción, siendo uno de ellos los de orden estilístico. Según Chrysanne DiMarco (1994), en «Stylistic Choice in Machine Translation»,

[w]e have all had the experience of reading a translated text, and feeling that it was not quite right, but not being able to articulate exactly why this was so. The text may be divided into the correct number of paragraph units, the semantic content may all be there, but something in the way the material is presented strikes a false note: while the meaning of the source text may be intact, the style of the target text is faulty. (p. 32)

La inquietud de DiMarco sobre el estilo resulta pertinente no solo en lo relativo a las traducciones efectuadas por seres humanos, sino que también y de manera particular cuando se trata de aquellas elaboradas por medio de sistemas de traducción automática. La situación parece agravarse más si se considera que, al tiempo que buena parte de la investigación en el campo se centra en el rendimiento de los sistemas informáticos de traducción, no parece que se pueda decir lo mismo de la relación estilo-traducción automática.

En cuanto a las ventajas que ofrece a los profesionales en traducción, Shift Languages destaca el hecho de que la traducción automática es más rápida y barata que la traducción tradicional, lo cual redundaría en un beneficio no solo para la empresa y el traductor, sino también para el cliente, cuyas demandas de inmediatez y bajos costos son constantes. Asimismo, la agencia argentina indica que, si bien la traducción automática puede representar una amenaza para muchos traductores, lo cierto es que también les ofrece múltiples beneficios que pueden dar pie a mayores oportunidades de trabajo, entre otras razones porque cada vez más agencias de traducción solicitan servicios de posesición. Otra ventaja es que les permite a los traductores ofrecer el servicio de traducción automática cuando lo crean oportuno. Por ejemplo, la traducción automática permite al traductor tomar encargos en circunstancias en que un cliente necesita traducir un texto muy extenso, pero tiene un presupuesto muy bajo, o cuando el cliente necesita traducir un texto no tan extenso, pero tiene muy poco tiempo.

Quizás la ventaja más importante para el traductor profesional es el notable aumento de la productividad que puede suponerle la utilización de un traductor automático. De acuerdo con Shift Languages, se estima que un traductor traduce en promedio 2000 palabras al día (en una jornada de ocho horas), lo que equivaldría a unas 300 palabras por hora, y que posedita cerca de 4800 palabras al día o unas 600 palabras por hora. Incluso la desventaja que supone que los honorarios por concepto de servicios de posesición sean entre un 30 % y un 50 % menores que la tarifa por servicios de traducción puede contrarrestarse con el aumento en la productividad antes referido. En consecuencia, si se realiza de manera adecuada, los servicios de posesición podrían llegar a ser más rentables que la misma traducción estándar (Shift Languages, párr. 7-10).

En cuanto a la dimensión técnica, los retos a los que todavía se enfrenta la traducción automática son varios. Al respecto, León (2015) explica que traducir es una tarea muy complicada para automatizar y, en particular, la subtarea de derivar el significado de un texto puede convertirse en una misión colosal, incluso cuando los pares de lenguas seleccionados no sean tan distantes (p. 415). Por ejemplo, para un sistema de traducción automática que solo analiza la sintaxis del texto fuente palabra por palabra, busca un equivalente en el diccionario para cada palabra, las une y da como resultado un texto de salida, surgen tres problemas inmediatos: primero, que algunas palabras pueden no tener un equivalente en la lengua meta; segundo, que algunas palabras pueden tener más de un significado y, por lo tanto, múltiples propuestas de traducción y, tercero, que el orden gramaticalmente correcto de las palabras puede diferir de una lengua a la otra (León, 2015, p. 415).

Aunque dista mucho de lo ideal, una posible solución a los problemas recién planteados sería reemplazar la palabra por un conjunto de palabras que se aproximen tanto como resulte posible al significado del texto original. En consecuencia, sería necesario modificar el sistema de traducción automática para que ejecute un análisis a nivel oracional, lo que conduciría al orden correcto de las palabras y su desambiguación. Sin embargo, tal procedimiento podría generar nuevos problemas. Algunas combinaciones de palabras específicas, por ejemplo, podrían tener más de un significado (las frases idiomáticas, por mencionar alguna), ciertas combinaciones de palabras apenas definidas podrían introducir significados que difieren de lo que se espera (como sería el caso de las colocaciones), el significado específico de determinadas palabras no siempre está determinado por la función morfológica que cumplen dentro de la oración, y algunas oraciones podrían tener más de un mapeo de palabras que se aplica a las funciones morfológicas (León, 2015, pp. 415-416). Es evidente que la traducción automática aún le quedan por resolver diferentes problemas de orden morfológico, sintáctico y semántico, entre otros.

Por su parte, Tertoolen (2012) identifica al menos dos tipos de problemas: los lingüísticos y los tecnológicos. Entre los problemas lingüísticos, señala la arbitrariedad de la lengua natural, ya que no existe una relación entre el signo y el significado; la productividad de una lengua (las reglas gramaticales permiten la creación de una infinita cantidad de construcciones sintácticas); y la transmisión cultural, puesto que cada lengua es producto de la evolución histórica que se transmite entre generaciones (p. 146). En lo que respecta a los problemas tecnológicos, Tertoolen menciona la falta de reglas claras (las computadoras necesitan reglas precisas para funcionar de manera correcta y, debido a la arbitrariedad de las lenguas naturales, no se cuenta con ellas), la incapacidad de realizar razonamientos de sentido común y, por último, la dificultad de aprender cosas nuevas (2012, p. 146).

POSEDICIÓN Y TRADUCCIÓN AUTOMÁTICA

Al tratar la traducción automática, es ineludible referirse a la posesición, dado que por lo general las traducciones que generan este tipo de sistemas informáticos deben ser poseídas. Según Sánchez-Gijón (2016), poseer implica efectuar «la edición de segmentos obtenidos mediante traducción automática (TA)» (p. 152). Dicha actividad ha experimentado un gran auge durante los últimos años y ha venido a ejemplificar la manera en que la tecnología puede revolucionar no solo el entorno de trabajo, sino la propia naturaleza de la labor que ejecuta un profesional determinado, el traductor en este caso (Sánchez-Gijón, 2016, p. 152). En otras palabras, la industria de la traducción parece demandar cada vez con mayor insistencia un profesional que debe ser capaz no solo de traducir de forma correcta, sino también de poseer de manera competente. Con el paso del tiempo la posesición se va convirtiendo en una competencia más que el traductor debe dominar si desea mantenerse competitivo, ya que la demanda de este servicio en el mercado aumenta día con día. Así, hay que aceptar que la traducción automática representa una oportunidad para ampliar el rango de servicios que un profesional de la traducción puede llegar a ofrecer: sin importar cuán desarrollados estén los sistemas de traducción automática ni cuán especializados sean en la actualidad, al menos por ahora siempre será necesario un poseedor que mejore la calidad de las traducciones realizadas por un *software* de traducción automática.

3. LA LINGÜÍSTICA TEXTUAL Y LOS GÉNEROS TEXTUALES

La aproximación a la traducción desde la dimensión textual supone la reflexión sobre determinados conceptos que resultan fundamentales a dicho enfoque. Es necesario entonces comprender, al menos en primera instancia, lo que debe entenderse cuando se habla de texto, lingüística del texto, tipos de texto y géneros textuales. Según Alexopoulou (2011), a finales de los años sesenta surgen las teorías de la lingüística del texto y las tipologías textuales que abordan el texto como una unidad específica de estudio, quedaban atrás así el estructuralismo de Ferdinand de Saussure y el generativismo de Noam Chomsky (s. p.). Alexopoulou también explica que la lingüística textual conceptualiza el texto como un producto de la actuación lingüística y de la interacción social; por lo tanto, el concepto de texto se convierte en una noción crucial para esta rama de la lingüística. Añade que la preocupación por explicar la naturaleza de los textos llevó a la búsqueda de un sistema que ordenara los distintos tipos de texto con el fin de definirlos, clasificarlos y construir un sistema tipológico con el objetivo de desentrañar la naturaleza de las estructuras textuales (2011, s. p.). Es de vital importancia comprender qué es un texto. Bernárdez (1982) nos dice que es

[la u]nidad lingüística comunicativa fundamental, producto de la actividad verbal humana, que posee carácter social. Se caracteriza por su cierre semántico y comunicativo y por su coherencia, debida a la intención comunicativa del hablante de crear un texto íntegro, y a su estructuración mediante dos conjuntos de reglas: las de nivel textual y las del sistema de la lengua. (p. 85)

Esta definición muestra que el texto tiene como mínimo tres dimensiones. En primer lugar, la dimensión comunicativa; es decir, se basa en la capacidad de los seres humanos de relacionarse consigo mismos y con los demás mediante el uso del lenguaje. En segundo lugar, la dimensión pragmática, ya que un texto ha de tomarse como lenguaje en uso. Al producir textos los hablantes utilizan el lenguaje con una intención en mente y dentro de situaciones concretas. En tercer lugar, la dimensión estructural, que hace alusión al empleo por parte del hablante de mecanismos y reglas lingüísticos a los que necesariamente debe acudir al construir textos con tal de que resulten coherentes y cohesionados (Alexopoulou, 2011, s. p.).

Beaugrande y Dressler (1997) proponen un modelo centrado en el usuario, esto es, en el emisor y el receptor, para diferenciar a los textos comunicativos de los no comunicativos. Para ellos, un texto es un acontecimiento comuni-

cativo que debe cumplir siete normas de textualidad: cohesión, coherencia, intencionalidad, aceptabilidad, situacionalidad, intertextualidad e informatividad; si un texto no satisface alguna de tales condiciones, no comunica, y si un texto no comunica, no puede ser analizado como si fuera un texto genuino (p. 35). Este razonamiento resulta de particular interés cuando se habla de traducción automática: ya que interesa que los sistemas de traducción automática produzcan textos genuinos que puedan ser utilizados en situaciones reales, es primordial que resulte claro qué es y qué no es un texto genuino.

Otros autores que aportaron a la clasificación de los textos fueron Teun van Dijk y Mijail Bajtín. Van Dijk (1980) sostiene que las regularidades del texto se manifiestan en la macroestructura y en la superestructura (en Alexopoulou, 2011, s. p.). La microestructura organiza los elementos del texto a nivel oracional a través de la coherencia semántica. Sin embargo, son el esquema estructural y el contenido global los que permiten clasificar el texto dentro de una tipología y un género textual en concreto (Alexopoulou, 2011, s. p.). Esta idea remite a la teoría de los géneros discursivos de Bajtín, uno de los aportes más significativos al estudio de la temática en el siglo xx (Alexopoulou, 2011, s. p.). Para Bajtín (1999), la diversidad de los géneros textuales o discursivos es inmensa porque a diario se produce una amplia gama de textos que varían según la situación comunicativa y que, a la vez, dan origen a otros nuevos a medida que se amplía y se vuelve más complejo el repertorio de géneros discursivos (p. 248). Por lo tanto, se puede observar que la producción de textos está muy ligada a las prácticas sociales y a la cultura misma, porque cada acto comunicativo sucede en un contexto social determinado.

En cuanto a las tipologías textuales, son numerosas las propuestas que han surgido a lo largo de los años. Klaus Brinker (1988) sostiene que el texto tiene un carácter «accional» y que cada texto tiene su respectiva función de la cual se derivan las distintas tipologías con sus correspondientes géneros. Brinker propone cinco funciones: informativa, de contacto, apelativa, obligativa y declarativa. Estas funciones textuales dan origen a tipos de texto, que se clasifican a partir de las mismas cinco denominaciones principales (en Alexopoulou, 2011, s. p.). Por otro lado, Werlich (1979) propone cinco tipos de texto muy relacionados con la estructura cognitiva humana: los narrativos, los descriptivos, los expositivos, los argumentativos y los instructivos (en Sánchez, 2006, pp. 84-85).

Con respecto a los géneros textuales, estos se clasifican desde dos perspectivas: la primera se enfoca en el tipo de enunciados, mientras que la segunda lo hace en la función. Bajtín (1999) propone una clasificación que varía según el tipo de enunciados y señala que se debe prestar atención a la diferencia entre los géneros textuales primarios o simples y los géneros textuales secundarios o complejos. Los géneros primarios son los constituidos por enunciados de la comunicación inmediata y se relacionan con contextos cotidianos, informales y familiares, por ejemplo, las cartas, los diálogos del diario vivir, los

mandatos, las interjecciones y las onomatopeyas, entre muchos otros. Por otra parte, los géneros secundarios absorben y reúnen diversos géneros discursivos simples, dando origen a otros géneros más complejos. En este segundo tipo de géneros discursivos, el uso de la lengua va más allá de la comunicación inmediata. Algunos ejemplos de géneros discursivos secundarios o complejos son las novelas, las investigaciones científicas, los textos periodísticos, los textos legales, las enciclopedias y las bitácoras (p. 250). De acuerdo con la propuesta de Bajtín, una determinada función (científica, técnica, periodística, oficial, cotidiana) crea determinados géneros que se logran identificar por el tipo de enunciados.

Cassany (2008) describe los géneros discursivos como dinámicos, situados y organizados en forma y contenido; además de que construyen y reproducen estructuras sociales y delimitan comunidades discursivas (pp. 12-13). Dentro de una comunidad discursiva el género textual ejerce tres funciones diferentes: la cognitiva, la interpersonal y la social y gremial. La función cognitiva «construye y formaliza el conocimiento de una disciplina», la función interpersonal «construye la identidad (la imagen o *face*) del autor, miembro de la comunidad» y la función social y gremial «contribuye a establecer el estatus de los miembros de la comunidad discursiva en la sociedad» (2008, p. 13). Por lo tanto, la noción de género textual abarca cuestiones tanto formales como de contenido. En lo que concierne a la traducción, resulta de suma importancia reconocer los géneros textuales y sus características convencionales en cada par de lenguas con las que se trabaje, ya que de ello dependerá que un texto resulte adecuado según las prácticas discursivas de la cultura y lengua de llegada.

4. EL ESTILO

Y LA ESTILÍSTICA CONTRASTIVA

Huang (2015) define el estilo como «[...] the distinctive way of using language which maintains some consistency in all one's writings» (p. 19). Al respecto, añade que Leech y Short se refieren al término como «the way in which language is used in a given context, by a given person, for a given purpose [...]» (1981, p. 10; en Huang, 2015, p. 17). Estos autores ponen de relieve la singularidad del estilo, que resulta de las escogencias que el hablante o escritor realiza de forma intencionada (Huang, 2015, p. 17). A estas definiciones, que remiten sobre todo al uso que cada individuo hace del lenguaje según sus intenciones particulares, Huang añade la de Crystal, quien define el concepto como «any situationally distinctive use of language, and of the choices made by individuals and social groups in their use of language» (1999, p. 323, en Huang, 2015, p. 18). Evidentemente en esta segunda definición la perspectiva desde la cual se acota el término en cuestión es más amplia: ya no solo se trata del uso distintivo que una persona haga de la lengua, sino que se ha elevado a la esfera social y situacional. A ello hay que añadir, además, una característica adicional: esas formas distintivas varían según los usos y los usuarios de la lengua. Por tal razón, Huang también acude a Cheng, quien define estilo como «the different varieties typically used by the language users of one specific language variety (for instance, standard language, dialect, social dialect, etc.) on different occasions» (1989, pp. 1-2, en Huang, 2015, p. 18).

Si el estilo depende de la regularidad con que un grupo de hablantes utiliza la lengua de determinadas formas en contextos específicos, es razonable pensar que tales regularidades pueden ser identificadas y descritas con un grado de precisión más que aceptable. Esta idea concuerda con lo propuesto por Enkvist, quien describe el estilo de un texto desde una perspectiva cuantitativa como «the aggregate of contextual probabilities of its linguistic items» (1964, p. 28, en Huang, 2015, p. 17). El estilo guarda una relación con la frecuencia de uso de determinados elementos lingüísticos que necesariamente deben existir dentro de un contexto específico; por lo tanto, el estilo de un texto dado solo se puede identificar a través de un análisis exhaustivo de las frecuencias de uso de los elementos lingüísticos que lo representan en todos sus niveles (Huang, 2015, p. 17). Enkvist también postula que, para identificar el estilo de un pasaje, las frecuencias de uso de sus diferentes elementos lingüísticos en distintos niveles deben compararse con las características correspondientes en otro texto o corpus de textos que sea considerado como la norma y que, por ende, se tiene por representativo de las características estilísticas del género textual

al que pertenece el pasaje bajo estudio (1964, p. 29, citado en Huang, 2015, p. 17). Esto quiere decir que el estilo puede ser considerado como una especie de desviación o apego a ciertas convenciones textuales.

En lo atinente a la traducción y según DiMarco (1994), el estilo influye en todos los niveles: léxico, sintáctico, semántico y textual. La elección particular de las palabras, el ordenamiento específico de las oraciones o la selección de aquellos elementos lingüísticos a transferir en la traducción, todo transmite un efecto estilístico particular. Las escogencias lingüísticas que el traductor hace al optar por varias maneras de comprender el texto fuente en el idioma de destino tienen una repercusión de tipo pragmático y, por lo tanto, afectan el significado que se transmite al lector (p. 32). Por esta razón, es fundamental prestar atención a la constitución del estilo a la hora de traducir sin importar la modalidad, ya sea que se trate de traducción entendida en el sentido tradicional (la realizada por humanos) o automática.

INDICADORES DE ESTILO

Leech y Short (1981) elaboran una lista de aspectos lingüísticos y estilísticos que se pueden estudiar en un texto o conjunto de textos con el fin de caracterizar estilos particulares. Tales aspectos se clasifican en (1) figuras retóricas o literarias, (2) categorías léxicas, (3) categorías gramaticales y (4) cohesión y contexto (pp. 75-80, en Huang, 2015, pp. 17-18), tal y como se muestra en la tabla 1.

Tabla 1. Lista de aspectos lingüísticos y estilísticos relacionados con el estilo según Leech y Short (1981)

A Checklist of Linguistic and Stylistic Categories	
a.	Lexical categories: (1) general (lexical complexity, degree of formality, descriptive or evaluative, general or specific, etc.); (2) nouns; (3) adjectives; (4) verbs; (5) adverbs [A2–A4 are about the frequency and semantic features]
b.	Grammatical categories: (1) sentence types (statements, questions, commands, exclamations, etc.); (2) sentence complexity (not only the length of sentence but also whether coordination, subordination, and parataxis are used); (3) clause types (the ratio of different types of clauses); (4) clause structure (the distribution of clause elements); (5) noun phrases; (6) verb phrases; (7) other phrase types [B5–B7 are about the ratio and distribution of types of phrases]; (8) word classes (the ratio between lexical words and functional words and their distributions, respectively); (9) general (the use of general types or grammatical constructions for special effect)
c.	Figures of speech: (1) grammatical and lexical schemes; (2) phonological schemes; (3) tropes
d.	Cohesion and context: (1) cohesion (the use of various cohesive devices, such as connectives, cross-reference by pronouns, etc.); (2) context (involves some characteristic use of narrative mode, for instance, the relationship between addresser and addressee revealed by the use of personal pronouns)

Fuente: Tomado de Huang (2015, pp. 17-18).

Para Huang, la caracterización de un estilo dado se logra al identificar las regularidades o patrones de uso de cada uno de los indicadores en la tabla 1 por medio de un análisis de las distribuciones estadísticas de los elementos lingüísticos específicos (2015, pp. 17-18). En lo que respecta a la traducción, la caracterización y posterior contraste del estilo de determinado género textual en las lenguas de partida y llegada permitiría evidenciar las adecuaciones estilísticas necesarias por implementar al pasar de una lengua a la otra siempre que el fin sea elaborar un texto meta funcional en el contexto de llegada. La comparación ha de centrarse en tres elementos clave: primero, en la regularidad de los patrones lingüísticos específicos; segundo, en la frecuencia de elementos lingüísticos específicos; y tercero, en una referencia que se toma como norma para la comparación (Huang, 2015, p. 18). De describir mejor la índole de este proceso contrastivo nos ocuparemos en el siguiente apartado.

ESTILÍSTICA CONTRASTIVA

De acuerdo con Bolaños Medina (2000), la estilística contrastiva estudia, desde una perspectiva funcional, las variaciones de uso entre textos redactados, en lenguas y culturas diferentes, que pertenecen a un mismo estilo funcional. Su objetivo fundamental es identificar las convenciones textuales propias de cada género textual en el seno de un sistema lingüístico-cultural dado, para luego hacer hincapié en sus diferencias y semejanzas con las propias de otro u otros sistemas (p. 42). Señala que, para contrastar este tipo de convenciones, es necesario establecer dos conceptos básicos: el de estilo y el de géneros funcionales. La autora define el estilo en términos funcionales como «aquel sistema de medios lingüísticos interrelacionados que cumplen un propósito comunicativo definido y se identifica con un ámbito profesional o del conocimiento humano determinado» (p. 42). Esta definición de estilo añade una dimensión adicional al concepto tal y como se había entendido en páginas anteriores: el de campo, idea por medio de la que se relacionan usos concretos del lenguaje, no solo con intenciones comunicativas determinadas, sino con actividades específicas, lo cual hace aun más evidente el carácter funcional de los géneros discursivos. Describe la idea de género funcional como «un acto comunicativo identificable caracterizado por un conjunto de propósitos comunicativos que los miembros de la comunidad profesional o académica en la que normalmente se produce reconocen y comprenden» (p. 42). Además, señala que los géneros funcionales presentan una estructura y una serie de convenciones muy marcadas (p. 42).

En cuanto al análisis propiamente dicho, Bolaños Medina describe varios modelos; señala que el modelo de análisis estilístico contrastivo de Galperin (1981) consta de cinco fases bien definidas que resume de la siguiente manera:

1. Fase taxonómica. El objetivo de esta fase es definir a qué estilo funcional pertenece el texto que se va a analizar. En esta etapa se analizan factores extralingüísticos.
2. Fase de comprensión de contenido. Implica buscar un acercamiento global al tema textual.
3. Fase semántica. En ella se estudian los significados de palabras por separado, aunque sin olvidar el contexto.
4. Fase estilística. Comprende el análisis del estilo mediante la detección de los procedimientos estilísticos usados por el autor, así como la organización de las frases que forman las diferentes SPU (*supra-phrasal units*), es decir, los diferentes marcos que componen el enunciado.
5. Fase funcional. Dedicada a profundizar en la información conceptual contenida en todo el texto. Para ello, se consideran los datos obtenidos como un todo y se sintetizan los diferentes procedimientos. Toda la información recolectada adquiere sentido al relacionarse con los objetivos comunicativos determinados por una situación dada.

(Bolaños Medina, 2000, p. 42)

Al referirse al enfoque de Shveitser (1997), Bolaños Medina (2000) comenta que este aboga por un acercamiento multidimensional, que refleje el carácter polifacético de los objetos descritos y que combine la comparación de los parámetros funcionales y de sus manifestaciones lingüísticas con la comparación de estructuras textuales y estrategias de diferentes variedades funcionales de un idioma (p. 42). También señala que Shveitser se interesa no solo por comparar los textos originales, sino también por analizar traducciones, tanto cualitativa como cuantitativamente (Bolaños Medina, p. 42).

El tercer diseño de estudio que Bolaños Medina destaca es el modelo de análisis estilístico aplicado a la traducción por Bravo (1999), de gran relevancia para la didáctica de la traducción porque se integra de manera perfecta en el proceso traductor como una fase más y tiene en consideración la definición del encargo de traducción (Bolaños Medina, p. 43). Partiendo de parámetros funcionales y extralingüísticos en el modelo se plantea el análisis de los niveles gráfico, grafológico, gramatical y léxico, basándose en los postulados de Crystal. Luego, con la información que se obtiene, se elabora un cuadro descriptivo general de los textos contrastados en el que se evidencien las semejanzas y diferencias lingüístico-culturales tipotextuales que se hayan identificado entre las dos lenguas y que servirán para establecer conclusiones sobre la estrategia más adecuada para su traducción (Bolaños Medina, p. 43). Como ya se apuntaba, este modelo parece ser de particular utilidad para el campo de la traducción, ya que permite precisar las convenciones de un género textual en los sistemas lingüísticos con que se trabaja, lo cual permite plantear las adaptaciones necesarias en cuanto a rasgos estilísticos concretos.

5. LA TAN, LA LINGÜÍSTICA TEXTUAL Y LA ESTILÍSTICA CONTRASTIVA: APLICACIÓN Y UTILIDAD PRÁCTICA

Para ilustrar la aplicación de las nociones expuestas en los apartados anteriores, parece conveniente incluir un ejemplo práctico. Así, el experimento que se describe de seguido muestra la forma en que se incorporan a un proceso de traducción que implica también una necesaria etapa de posesición.

ÍNDOLE DEL EXPERIMENTO

En términos generales, mediante el experimento se pretendió analizar el estilo de un texto de salida o traducción generada por el sistema de traducción automática neuronal de Google en contraposición con los patrones de estilo observados en un conjunto de textos pertenecientes al mismo género textual escritos originalmente en lengua meta. El texto original es un protocolo de ensayo clínico que fue traducido del español al inglés por medio del recién nombrado sistema TAN. En términos específicos, del protocolo referido solamente se tradujo y comparó la sección denominada «Justificación». El uso del sistema TAN de Google resulta particularmente adecuado para este tipo de tareas ya que, además de que se trata de una herramienta de acceso abierto, también posee una arquitectura informática que se encuentra entre las más avanzadas de los sistemas TAN de carácter general que existen en la actualidad. Por lo demás, resulta necesario aclarar que durante el experimento solamente se estudiaron los rasgos de estilo de las traducciones generadas por el motor de Google del español al inglés. En cuanto al procedimiento seguido, el experimento se desarrolló en seis etapas, las cuales se detallan a continuación.

ETAPA 1. LA RECOPIACIÓN DE LOS TEXTOS

De previo, se establecieron tres parámetros de carácter general para seleccionar tanto el texto original como el conjunto de textos en lengua meta: (1) que la extensión de los textos oscilara entre las 100 y 200 palabras, (2) que todos los textos trataran sobre investigaciones relacionadas con la COVID-19 y el SARS-CoV-2, de manera que la terminología y el contenido fuesen lo más homogéneos posible y (3) que todos los casos refirieran a investigaciones intervencionistas, en otras palabras, estudios clínicos en los que un grupo de participantes ha recibido algún tipo de intervención o tratamiento con el fin de

evaluar los resultados biomédicos o relacionados con la salud (ClinicalTrials.gov, «Glossary of Common Site Terms», 2021, disponible en: <https://clinicaltrials.gov/ct2/about-studies/glossary>). Una vez fijados los parámetros recién referidos, se seleccionaron un ensayo clínico en español o texto original y un corpus de textos paralelos en inglés.

El texto en español o texto original. El texto original seleccionado es un protocolo de ensayo clínico tomado del Registro Español de Ensayos Clínicos (REEG), base de datos gratuita a la que se puede acceder desde la página electrónica de la Agencia Española de Medicamentos y Productos Sanitarios. Como ya se había mencionado, de dicho texto se utilizó únicamente la sección denominada «Justificación». A continuación, se presenta el contenido del texto en cuestión.

Cuadro 1. Texto original

Justificación:

El objetivo de este ensayo clínico es investigar si la administración de la vacuna para prevenir la COVID-19 es segura para mujeres embarazadas y sus bebés en gestación, así como las cantidades de anticuerpos que producen. También se estudiará si la vacuna en investigación puede prevenir la COVID-19 en las mujeres vacunadas durante todo el embarazo.

Las mujeres embarazadas se dividirán en dos grupos: un grupo que recibirá la vacuna en investigación y otro grupo que recibirá el placebo. Ambos mediante la inyección en el músculo de la parte superior del brazo. Las mujeres embarazadas asignadas a recibir placebo podrán recibir la vacuna en investigación a partir del mes de alumbramiento.

Durante el estudio se realizarán una serie de visitas al centro del ensayo clínico: visitas de inyección, así como visitas de seguimiento.

La duración de este estudio es de alrededor de 7-10 meses, dependiendo de si la vacuna se le administra a la participante durante el embarazo o poco después del nacimiento de su bebé. El bebé participará en el ensayo clínico durante un máximo de 6 meses.

Fuente: Tomado del Registro Español de Ensayos Clínicos (REEG).

Los textos paralelos. Para contar con un «modelo» de estilo con el que contrastar el estilo del texto de salida (la traducción realizada por medio del sistema TAN de Google) se reunió un conjunto de treinta textos paralelos en inglés (treinta protocolos de ensayo clínico) extraídos de la página electrónica ClinicalTrials.gov (disponible en: <https://clinicaltrials.gov/>), base de datos administrada por la U.S. National Library of Medicine. De cada uno de tales textos, se procesó la sección «Study Description», por representar la sección análoga a la sección «Justificación» del texto original. Los treinta textos de

este corpus monolingüe cumplían, como es de esperar, con los criterios pre-establecidos: su extensión se encontraba en un rango de entre 100 y 200 palabras, todos trataban sobre investigaciones relacionadas con la COVID-19 y el SARS-CoV-2 (para garantizar en la medida de lo posible la homogeneidad en cuanto a la terminología y el contenido) y, por último, todos los casos referían a investigaciones intervencionistas. A continuación, se presenta un ejemplo de los textos recolectados:

Cuadro 2. Ejemplo de texto paralelo

Brief Summary:

This retrospective cohort study aims to evaluate the effect of proton pump inhibitor (PPI) or histamine-2 receptor antagonist (H2RA) on the clinical outcomes and positivity rates of Coronavirus Disease-19 (COVID-19).

Detailed Description:

Previous researches on the association between the use of PPI and H2RA and the treatment and prevention of COVID-19 have reported inconsistent findings. This study aimed to evaluate the effect of H2 receptor antagonists and proton pump inhibitors on the clinical outcomes and positivity rates of COVID-19 using the Korean National Health Insurance (NHI) data. This is a retrospective cohort study using the NHI COVID-19 dataset, which includes health insurance data of individuals who underwent COVID-19 PCR test in South Korea. This study aims to evaluate the effect of the use of PPI and H2RA on the positivity rates and clinical outcomes of COVID-19.

Fuente: Tomado de ClinicalTrials.gov.

ETAPA 2. LA TRADUCCIÓN AL INGLÉS

Recopilados los textos, se procedió a traducir el texto original. Para ello, como ya se ha indicado en repetidas ocasiones, se utilizó el traductor automático neuronal de acceso abierto de Google, denominado Google Translate (<https://translate.google.co.cr/>). Mediante el cuadro 3 se muestra la traducción o texto de salida tal y como la produjo Google Translate.

ETAPA 3. EL ANÁLISIS DE LOS TEXTOS PARALELOS Y DE LA TRADUCCIÓN

Al analizar el estilo tanto del conjunto de textos paralelos en inglés como de la traducción realizada por medio de Google Translate, se tomó en consideración la frecuencia de uso de seis categorías de rasgos estilísticos, las cuales se detallan a continuación.

Uso de las oraciones. Las oraciones se analizaron según su cantidad, extensión, modalidad, complejidad y distribución. El análisis con respecto a la cantidad refirió al número de oraciones que en promedio se contabilizan en una cantidad determinada de palabras. En cuanto a la extensión, las oraciones se clasificaron según la cantidad de palabras que las conformaban: las oraciones cortas iban de 1 a 20 palabras, las medianas de 21 a 45 palabras, las largas de 46 a 70 palabras y las extralargas de 71 palabras en adelante (véase el documento *Analyzing Your Own Style*, disponible en: <https://cmsw.mit.edu/writing-and-communication-center/analyzing-style/>). El análisis relativo a la modalidad de las oraciones implicó clasificarlas según fueran afirmativas, negativas, interrogativas, imperativas o exclamativas. En lo que corresponde a su complejidad, las oraciones fueron clasificadas, a partir de la tipología de Crystal y Davy (2016), en simples, compuestas, complejas o mixtas (pp. 46-49). Finalmente, la caracterización a partir de la distribución significó identificar la frecuencia con que las oraciones, ya fuesen cortas, medianas, largas o extralargas, se ubicaban al inicio, en el medio o al final de los párrafos

Cuadro 3. Texto de salida o traducción original producida por Google Translate

Justification:

The objective of this clinical trial is to investigate whether the administration of the vaccine to prevent COVID-19 is safe for pregnant women and their unborn babies, as well as the amounts of antibodies they produce. It will also study whether the investigational vaccine can prevent COVID-19 in vaccinated women throughout pregnancy.

Pregnant women will be divided into two groups: a group that will receive the investigational vaccine and another group that will receive the placebo. Both by injecting into the muscle of the upper arm. Pregnant women assigned to receive placebo will be able to receive the investigational vaccine from the month of delivery.

During the study there will be a series of visits to the clinical trial center: injection visits as well as follow-up visits.

The duration of this study is around 7-10 months, depending on whether the vaccine is given to the participant during pregnancy or shortly after the birth of her baby. The baby will participate in the clinical trial for up to 6 months.

Fuente: Elaboración propia por medio del traductor Google Translate.

Categorías léxicas. Los textos fueron analizados para identificar la frecuencia de uso de sustantivos, adjetivos, verbos, adverbios, preposiciones, pronombres y verbos auxiliares según los valores porcentuales que cada categoría representaba en el conjunto de los textos.

Primeros elementos de las oraciones. El estudio de esta categoría implicó reconocer la recurrencia de aquellos elementos que se ubicaban al inicio de las oraciones. Los primeros elementos se clasificaron según fueran el sujeto de la oración, predicado de la oración, cláusulas subordinadas, frases, palabras o fórmulas de transición (conectores), palabras interrogativas, verbos u otros elementos varios (compárese con el documento *Analyzing Your Own Style*).

Grado de densidad léxica. Para caracterizar este aspecto, se acudió al analizador sintáctico automático *Analyze My Writing* (disponible e: <https://www.analyzemywriting.com>). Desde esta herramienta informática, se define el grado de densidad léxica como «the number of *lexical words* (or content words) divided by the total number of words [...]» («Lexical Density», párr. 1), entendidas las *lexical words* o categorías léxicas como las palabras que proporcionan información acerca del tema de un texto dado: sustantivos, adjetivos, verbos y adverbios. Así, «[n]ouns tell us the subject, adjectives tell us more about the subject, verbs tell us what they do, and adverbs tell us how they do it» (*Analyze My Writing*, «Lexical Density», 2022, párr. 2-4). Los artículos, las preposiciones, las conjunciones y los verbos auxiliares, entre otros, son denominadas *function words*, cuya naturaleza es más bien gramatical; se consideraron, entonces, como palabras que proporcionan relativamente poca información sobre el tema tratado en un texto (*Analyze My Writing*, «Lexical Density», 2022, párr. 2-4). Con ello en mente, se determina que la densidad léxica «is simply the percentage of words in written (or spoken) language which give us information about what is being communicated. With regard to writing, *lexical density is simply a measure of how informative a text is*» (*Analyze My Writing*, «Lexical Density», 2022, párr. 2-4).

Nivel promedio de legibilidad. Según se establece en la página electrónica *Analyze My Writing*, el índice de legibilidad es «an estimation of how difficult a text is to read. The estimation is made by measuring a text's complexity» («What is a Readability Index?», 2022, párr. 1). Existen múltiples índices que se pueden utilizar para medir la legibilidad de los textos. No obstante, todos acuden a la complejidad a la hora de realizar sus estimaciones. Al respecto, en la página de *Analyze My Writing* se señala lo siguiente:

Every index does this a little bit differently and emphasizes particular aspects of text complexity. Some emphasize syllable counts while others look only at word and sentence lengths. This is why there are more than a few readability indices available to us. However, the core idea of each is the same; readability is essentially a measure of text complexity. («What is a Readability Index?», 2022, párr. 1-2)

Así, entre las maneras en que se puede determinar la complejidad de un texto dado se encuentran las mediciones relativas a la extensión de las palabras y oraciones o el conteo de sílabas. Una vez que se cuenta información sobre aspectos como los recién referidos, se comparan la complejidad del texto

con la facilidad con la que los lectores lo comprenden; a partir del resultado de dicha comparación se elabora una fórmula por medio de la cual predecir cuán difícil de leer resultará un texto dado según su grado de complejidad (Analyze My Writing, «What is a Readability Index?», 2022, párr. 1). Aspecto importante por tomar en consideración sobre los índices es que sus predicciones se establecen con respecto a contextos específicos. Por ejemplo, los índices de legibilidad que puedan establecerse por medio de Analyze My Writing para un determinado texto se corresponderán con el nivel requerido dentro del sistema educativo formal de los Estados Unidos para que el lector comprenda dicho texto sin dificultades; «[t]hus, *the lower the index, the easier the text is to read*, and conversely, *the higher the index, the more difficult the text is to read*» (Analyze My Writing, «Interpreting the Numbers», párr. 1). Por medio de la tabla 2 se puede visualizar de mejor manera la relación que establece el índice utilizado en esta página electrónica con el nivel educativo dentro de los Estados Unidos.

Tabla 2. Relación índice de legibilidad/años de escolaridad en EE. UU. según Analyze My Writing

Readability Index	Level	Examples
3 and below	Emergent and early readers	Picture and early reader books
3-5	Children	Chapter books
5-8	Young adult	Advertising copy, young adult literature, some news articles
8-12	General adult	Novels, news articles, blog posts, political speeches
12-16	Undergraduate	College textbooks
16 and above	Graduate, post-graduate, professional	Scholarly journal and technical articles

Fuente: Tomado de Analyze My Writing, «Interpreting the Numbers» (2022, párr. 7).

Incidencia de voces pasivas. Por medio de esta, la última categoría, se hace alusión a la recurrencia de la voz pasiva en un texto, conjunto de textos o género textual dado. Se establece a partir de la cantidad de voces pasivas por cada cierta cantidad de oraciones.

Para analizar los textos atinentes al experimento se acudió al analizador sintáctico automático o *parser* Analyze My Writing. Esta herramienta permitió examinar tanto el corpus de textos paralelos en inglés como la traducción generada por medio del traductor automático neuronal de Google. Su uso permitió analizar de forma automatizada la mayoría de los rasgos estudiados, aunque en algunos casos el análisis tuvo que realizarse de forma manual —el promedio de oraciones por cada cierta cantidad de palabras, por ejemplo, así como la frecuencia de uso con respecto a la extensión, la complejidad y el modo de las oraciones—. El análisis practicado a cada uno de los textos que conformaron el corpus de textos paralelos se ilustra por medio del texto paralelo 2 (Cuadro 4) y de los resultados de dicho análisis.

Uso de las oraciones:

1. Cantidad promedio de oraciones por cada 100 palabras: 5,5 oraciones
2. Extensión de las oraciones: 50 % oraciones cortas (entre 1 y 20 palabras) y 50 % oraciones medianas (entre 21 y 45 palabras)
3. Distribución de las oraciones: oraciones cortas al inicio y en medio del texto; oraciones medianas en medio y al final del texto
4. Modo de las oraciones: 100 % oraciones afirmativas
5. Complejidad de las oraciones: 50 % simples, 33 % complejas y 17 % mixtas
6. Primeros elementos de las oraciones: 83 % sujetos y 17 % frases

Cuadro 4. Texto paralelo 2

Brief Summary:

The purpose of this study is to investigate if Quercetin Phytosome is beneficial for the treatment of COVID-19.

Detailed Description:

This is a randomized, open-labelled and controlled study aimed to investigate the adjuvant benefits of Quercetin Phytosome in community-based subjects with confirmed SARS-CoV-2 infection (by RT-PCR). The study has two arms. In one arm the subjects will receive standard COVID-19 care as per the hospital/physician guidelines, whereas in the other arm the subjects will receive standard COVID-19 care + Quercetin Phytosome. The treatment will continue for 30 days. It is proposed that Quercetin Phytosome will contribute to boosting the natural immunity of the subjects and will help in preventing the COVID-19 disease progression i.e. preventing the need of hospitalization.

Fuente: Tomado de ClinicalTrials.gov.

Categorías léxicas:

1. Sustantivos: 36,44 %
2. Adjetivos: 5,08 %
3. Verbos: 14,41 %
4. Preposiciones: 15,25 %
5. Pronombres: 1,69 %
6. Verbos auxiliares: 4,24 %

Grado de densidad léxica: 55,93 %

Nivel promedio de legibilidad: 14,45 (según índice de Analyze My Writing)

Incidencia de voces pasivas: 17 %

Para continuar, se ofrecen los resultados del análisis practicado a la traducción original elaborada por medio de Google Translate (véase § Cuadro 1):

Uso de las oraciones:

1. Cantidad promedio de oraciones por cada 100 palabras: 5,1 oraciones
2. Extensión de las oraciones: 50 % oraciones cortas y 50 % oraciones medianas
3. Distribución de las oraciones: oraciones cortas al inicio, en medio y al final del texto; oraciones medianas al inicio, en medio y al final
4. Modo de las oraciones: 100 % oraciones afirmativas
5. Complejidad de las oraciones: 50 % simples y 50 % complejas
6. Primeros elementos de las oraciones: 75 % sujetos, 12,5 % frases y 12,5 % pronombres

Categorías léxicas:

1. Sustantivos: 28,99 %
2. Adjetivos: 7,69 %
3. Verbos: 8,88 %
4. Adverbios: 2,37 %
5. Preposiciones: 18,93 %
6. Pronombres: 1,18 %
7. Verbos auxiliares: 4,14 %

Grado de densidad léxica: 47,93 %

Nivel promedio de legibilidad: 13,07 (según índice de Analyze My Writing)

Incidencia de voces pasivas: 25 %

ETAPA 4. LA COMPARACIÓN DE RASGOS ESTILÍSTICOS

Procesado el conjunto de textos paralelos en inglés, fue posible contraponer los patrones identificados en el corpus con los patrones observados en la traducción original, con respecto a cada uno de los rasgos de estilo analizados. El resultado de tal contraposición se presenta por medio de la tabla 3.

Tabla 3. Comparación de patrones de estilo textos paralelos/traducción original

Rasgo analizado	Muestras de textos paralelos (171 oraciones en total)	Traducción original (8 oraciones en total)
Promedio de oraciones por cada 100 palabras	aprox. 4 por cada 100 palabras	aprox. 5 por cada 100 palabras
Frecuencias de uso con respecto a la extensión de las oraciones	35 % cortas	50 % cortas
	54 % medianas	50 % medianas
	8 % largas	--
	1 % extralargas	--
Frecuencias de uso con respecto a la complejidad de las oraciones	63 % simples	50 % simples
	33 % complejas	50 % complejas
	1 % compuestas	--
	3 % mixtas	--
Frecuencias de uso con respecto al modo	98 % afirmativas	100 % afirmativas
	2 % negativas	--
Patrones de distribución de las oraciones según la extensión	cortas: en cualquier posición	cortas: en cualquier posición
	medianas: en cualquier posición	medianas: en cualquier posición
	largas: en cualquier posición	--
	extralargas: en el medio y al final	--
Frecuencias de uso de las categorías léxicas	40 % sustantivos	29 % sustantivos
	8 % adjetivos	8 % adjetivos
	10 % verbos	8 % verbos
	2 % adverbios	2 % adverbios
	14 % preposiciones	19 % preposiciones
	0,7 % pronombres	1 % pronombres
Frecuencias de uso de los primeros elementos de las oraciones	4 % verbos auxiliares	4 % verbos auxiliares
	85 % sujetos	75 % sujetos
	6 % frases	12 % frases
	4 % palabras/fórmulas transición	13 % otros (pronombres)
	3 % verbos	--
Incidencia de voz activa y pasiva	2 % cláusulas subordinadas	--
	57 % voces activas	75 % voces activas
Nivel de legibilidad	43 % voces pasivas	25 % voces pasivas
	18	13
Grado de densidad léxica	59 %	47,9 %

Fuente: Elaboración propia a partir del análisis y de la traducción original.

ETAPA 5. LA COMPARACIÓN DE RASGOS ESTILÍSTICOS

El contraste entre los patrones de estilo identificados tanto en el corpus de textos paralelos como en la traducción original permitió detectar las diferencias entre ambos. Las diferencias a su vez permitieron establecer un conjunto de cambios por implementar —durante el proceso de posesición— en la traducción original con el fin de que su estilo se adecuara al de los textos paralelos. La lista de los principales cambios por implementar es la siguiente:

- reducir el número promedio de oraciones por cada 100 palabras en 1 oración;
- reducir la cantidad de oraciones cortas un 15 % y aumentar la cantidad de oraciones medianas y largas;
- aumentar la cantidad de oraciones simples y disminuir la cantidad de oraciones complejas;
- aumentar la cantidad de sustantivos aproximadamente en un 10 %;
- aumentar la cantidad de sujetos como primer elemento de las oraciones aproximadamente en un 10 % y reducir la cantidad de frases y pronombres que aparecen en primera posición;

- aumentar la incidencia de voces pasivas aproximadamente un 20 %;
- aumentar el grado de densidad léxica aproximadamente en un 10 %; y
- aumentar el nivel promedio de legibilidad en aproximadamente 5 puntos según el índice utilizado por Analyze My Writing.

Cuadro 5. Traducción poseditada

Rationale:

This study is aimed to investigating if vaccine administration to prevent COVID-19 is safe for pregnant women and their unborn babies, observing the number of antibodies they produce, and examining if the vaccine under investigation prevents COVID-19 in vaccinated women throughout pregnancy.

Pregnant women participating in the clinical trial will be distributed into two groups. A group will receive the investigational vaccine while another group the placebo, both administered by intramuscular injection in the upper arm. Pregnant women assigned the placebo will receive the investigational vaccine from the month of delivery on.

During the period of the study, a series of visits to the clinical trial center will be made— injection visits as well as follow-up visits.

The study will last from 7 to 10 months, depending on whether the vaccine is administered to the participants during pregnancy or shortly after the birth of her baby. The baby will participate in the clinical trial for a maximum period of 6 months.

Fuente: Elaboración propia a partir de la traducción original elaborada con Google Translate.

El resultado de implementar las adecuaciones recién señaladas —es decir, de realizar el proceso de posesición que garantizaría la adecuación estilística de la traducción original a los patrones de estilo detectados en el corpus de textos paralelos— se presenta en el cuadro 5.

ETAPA 6. EL ANÁLISIS DE LA TRADUCCIÓN POSEDITADA

Una vez implementadas las adecuaciones antes descritas y efectuada la traducción poseditada, se procedió a un nuevo análisis comparativo, esta vez para contraponer los patrones de estilo presentes en el corpus de textos paralelos con los que pueden identificarse en la traducción poseditada. Su resultado se muestra mediante la tabla 4. Con el fin de observar el resultado de los ajustes realizados, se incluyen de nuevo los resultados del análisis practicado a la traducción original (tercera columna de izquierda a derecha, tabla 4).

Tabla 4. Comparación de patrones de estilo textos paralelos/traducción original/traducción poseidita

Rasgo analizado	Muestra de textos paralelos (171 oraciones en total)	Traducción original (8 oraciones en total)	Traducción poseidita (7 oraciones en total)
Promedio de oraciones por cada 100 palabras	aprox. 4 por cada 100 palabras	aprox. 5 por cada 100 palabras	aprox. 4 por cada 100 pal.
Frecuencias de uso con respecto a la extensión de las oraciones	35 % cortas	50 % cortas	43 % cortas
	54 % medianas	50 % medianas	57 % medianas
	8 % largas	--	--
	1 % extralargas	--	--
Frecuencias de uso con respecto a la complejidad de las oraciones	63 % simples	50 % simples	71 % simples
	33 % complejas	50 % complejas	29 % complejas
	1 % compuestas	--	0 % compuestas
	3 % mixtas	--	0 % mixtas
Frecuencias de uso con respecto a la modalidad	98 % afirmativas	100 % afirmativas	100 % afirmativas
	2 % negativas	--	--
Patrones de distribución de las oraciones según la extensión	cortas: cualquier posición	cortas: cualquier posición	cortas: en medio y final
	medianas: cualquier posición	medianas: cualquier posición	medianas: al inicio y en medio
	largas: cualquier posición	--	--
	extralargas: en medio y final	--	--
Frecuencias de uso de las categorías léxicas	40 % de sustantivos	29 % de sustantivos	34 % de sustantivos
	8 % de adjetivos	8 % de adjetivos	8 % de adjetivos
	10 % de verbos	8 % de verbos	11 % de verbos
	2 % de adverbios	2 % de adverbios	1,25 % de adverbios
	14 % de preposiciones	19 % de preposiciones	18 % de preposiciones
	0,7 % de pronombres	1 % de pronombres	0,6 % de pronombres
	4 % de verbos auxiliares	4 % de verbos auxiliares	3 % de verbos aux.
Frecuencias de uso de los primeros elementos de las oraciones	85 % sujetos	75 % sujetos	86 % sujetos
	6 % frases	12 % frases	14 % frases
	4 % palabras/fórm. transición	13 % otros (pronombres)	--
	3 % verbos	--	--
	2 % cláusulas subordinadas	--	--
Incidencia de voz activa y pasiva	57 % voces activas	75 % voces activas	57 % voces activas
	43 % voces pasivas	25 % voces pasivas	43 % voces pasivas
Nivel de legibilidad	18	13	15,5
Grado densidad léxica	59 %	48 %	54,3 %

Fuente: Elaboración propia a partir del análisis del corpus, de la traducción original y de la traducción poseidita.

RESULTADOS DEL EXPERIMENTO

En términos generales, la traducción automática neuronal o TAN es una herramienta eficaz para traducir el género textual estudiado (protocolos de ensayos clínicos) del español al inglés, de aceptable calidad y con ahorro considerable en tiempo y esfuerzo, en comparación a una traducción humana. No obstante, para obtener tales resultados y beneficios, es fundamental que se utilice en combinación con un sistema o modelo de posesición que forme parte del proceso de traducción. Tal modelo debe permitir la identificación de adecuaciones de estilo precisas por implementar en las traducciones de salida que produzca el sistema TAN, esto con el fin de que, tras el proceso de posesición, tales traducciones gocen del grado de aceptación requerido por el usuario meta (siempre que esto sea lo que se necesite). A su vez, las adecuaciones y el modelo mismo han de elaborarse a partir del contraste entre los patrones de estilo que siguen las traducciones de salida que produce el sistema

TAN utilizado y los patrones identificables en un conjunto de textos paralelos originalmente escritos en lengua meta y representativos del género textual que se esté traduciendo. Aunque a primera vista pueda parecer que el proceso de posesición resulta muy complicado por la gran cantidad de parámetros que deben considerarse, puede resultar sencillo gracias al excelente rendimiento que han llegado a tener los sistemas TAN (el traductor de Google, en este caso, no cometió omisiones de contenido, cambios en el sentido o errores de puntuación u ortografía, a menos, claro, que estuvieran presentes en el texto original). Finalmente, los resultados obtenidos a partir del experimento llevado a cabo podrían variar de un género textual a otro; incluso, de una sección dentro del mismo protocolo de ensayo clínico a otra. Esto es, que las características del experimento no permiten extrapolar sus hallazgos más allá de los casos estudiados.

BIBLIOGRAFÍA CITADA

- Alexopoulou, A. (2011). Tipología textual y comprensión lectora en E/LE. *Revista Nebrija de Lingüística aplicada a la enseñanza de Lenguas*, 9, s. p.
- Analyze My Writing. (2022). <https://www.analyzemywriting.com/>
- Analyze My Writing. (2022). Interpreting the Numbers. https://www.analyzemywriting.com/readability_indices.html
- Analyze My Writing. (2022). Lexical Density. https://www.analyzemywriting.com/lexical_density.html
- Analyze My Writing. (2022). What is a Readability Index? https://www.analyzemywriting.com/readability_indices.html
- Bajtín, M. (1999). *Estética de la creación verbal* (10.^a ed.). México D. F.: Siglo XXI.
- Beaugrande, R. A. de, & Dressler, W. U. (1997). *Introducción a la lingüística del texto*. Barcelona: Ariel Lingüística.
- Bernárdez, E. (1982). *Introducción a la lingüística del texto*. Madrid: Espasa-Calpe.
- Berner, S. (2003). «Lost in translation»: Cross-lingual communication and virtual academic communities. En *5th Annual Conference on World Wide Web Applications* (pp. 10-12). <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.130.2973>
- Bolaños Carpio, A. (2014). *Consideraciones bioéticas en la traducción del protocolo de estudio clínico: Variability of Parkinson's Disease Biomarker Analytes*. [Trabajo final de graduación, Universidad Nacional, Costa Rica]. UNA Translation Research Access Manager. <https://www.mogap.net/pmt/AdrianaBolanos.pdf>
- Bolaños Medina, A. (2000). Aplicaciones del análisis estilístico contrastivo-funcional para la traducción de géneros digitales. *Vector Plus*, 16, 39-47. https://acceda.cris.ulpgc.es/bitstream/10553/7505/5/0231633_00016_0003.pdf
- Cabré, M. T. (2002). Textos especializados y unidades de conocimiento: metodología y tipologización (I). En J. García Palacios & M. T. Fuentes Morán (Eds.). *Texto, terminología y traducción* (pp. 15-36). Salamanca: Almar.
- Cabré, M. T., & i Castellví, M. T. C. (1993). *La terminología: teoría, metodología, aplicaciones*. Barcelona: Antártida/Empúries.
- Cassany, D. (2008). Metodología para trabajar con géneros discursivos. En P. Salaburu Etxeberria & I. Ugarteberu Gastañares (Eds.). *Espezialitate hizkerak eta terminologia III: espezialitate hizkeren didaktika eta komunikazioa* (pp. 9-24). País Vasco: Argitalpen Zerbitzua.

- Castro Benítez, A. (2012). *Oportunidades de mejora para la traducción de resúmenes de artículos de revistas científico-médicas costarricenses: El caso de la Revista Médica de la Universidad de Costa Rica*. [Trabajo final de graduación, Universidad Nacional, Costa Rica]. UNA Translation Research Access Manager. <https://www.mogap.net/pmt/AdrianaCastro.pdf>
- Cheng, Y. (1989). *Yingyu Yutixue* [English stylistics]. Shanghai: Shanghai Foreign Language Education Press.
- Cocozza Garro, G. (2018). *Corpus-based machine translation: a study case for the e-government of Costa Rica*. [Trabajo final de maestría, Universidad Autónoma de Barcelona, España]. https://ddd.uab.cat/pub/trerecpro/2018/tfg_96164/TFM_Corpus-Based_Machine_Translation_A_Study_Case_for_the_e-Government_of_Costa_Rica_GCocozza.pdf
- Crystal, D., & Davy, D. (2016). *Investigating English style*. Londres: Routledge. [edición electrónica]
- DiMarco, C. (del 5 al 8 de octubre de 1994). Stylistic choice in machine translation. En Association for Machine Translation in the Americas. *Proceedings of the First Conference of the Association for Machine Translation in the Americas* (pp. 32-39). <https://aclanthology.org/1994.amta-1.5.pdf>
- DiMarco, C., & Mah, K. (1994). *A model of comparative stylistics for machine translation*. *Machine translation*, 9(1), 21-59. <https://www.jstor.org/stable/40007981>
- Enkvist, N. Erik. (1964). On defining style: An essay in applied linguistics. En Spencer John (Ed.). *Linguistics and style* (3-56). Londres: Oxford University Press.
- Forcada, M. L. (2017). Making sense of neural machine translation. *Translation Spaces*, 6(2), 291-309. DOI 10.1075/ts.6.2.06for
- Gamal Othman, M. (2017). *Resumen de la Traducción Automática*. [Trabajo final de maestría, Universidad Nacional de Educación a Distancia, España].
- González, A. O. (1 de abril de 2021). Cómo crear nuestro propio sistema de traducción automática neuronal. *The Conversation*. <https://theconversation.com/como-crear-nuestro-propio-sistema-de-traducion-automatca-neuronal-155769>
- González Boluda, M. (2010). Estudio comparativo de traductores automáticos en línea: Systran, Reverso y Google. *Núcleo*, 22(27), 187-216. http://saber.ucv.ve/ojs/index.php/rev_n/article/view/2565
- González Salas, L. F. (2015). *La calidad comunicativa de la traducción al inglés de los resúmenes de artículos biomédicos costarricenses: ¿el traductor científico o el científico traductor?* [Trabajo final de graduación, Universidad Nacional, Costa Rica]. UNA Translation Research Access Manager. <https://www.mogap.net/pmt/LuisGonzalez.pdf>
- Huang, Libo. (2015). *Style in translation: A corpus-based perspective*. Berlín/Shanghai: Springer/Shanghai Jiao Tong University Press.

- Hutchins, J. (2006). Machine translation: History. En Keith Brown (Ed.). *Encyclopedia of Language & Linguistics* (2.^a ed.) (pp. 375-383). Oxford: Elsevier Pergamon.
- Invenmar. (s. f.). Longitud de las oraciones y los párrafos. *Invenmar*. <http://www.invenmar.org.co/redcosteral/invenmar/docs/660longoracpar.htm>
- Leech, G., & Short, M. (1981). *Style in fiction: A linguistic introduction to English fictional prose*. Londres: Longman.
- León Calderón, S. (2015). Building a controlled natural language framework for real-time machine translation. *Revista Lenguas Modernas*, 23, 413-422. DOI 10.15517/RLM.V0I23.22362
- López, J. (2015). Why machine translation is not good for business in Costa Rica. *The Costa Rica Star*. <https://news.co.cr/why-machine-translation-is-not-good-for-business-in-costa-rica/38639/>
- Machine translation. (s. f.). *Cambridge Dictionary*. <https://dictionary.cambridge.org/es/diccionario/ingles/machine-translation>
- Marsh, M. (1999). Algunas consideraciones sobre la traducción médica. *Centro Virtual Cervantes*. <https://cvc.cervantes.es/lengua/aproximaciones/marsh.htm>
- Molina Baños, A. (2019). *Creación de un motor de traducción automática estadístico (EN>ES) para textos del ámbito farmacéutico. Comparación con otros motores de traducción automática neuronal existentes*. [Trabajo de fin de máster, Universitat Autònoma de Barcelona]. <https://ddd.uab.cat/record/210864>
- Monge Naranjo, A. L. (2016). *Aspectos estilísticos en la traducción de resúmenes de investigaciones de psicología: la proporción de verbos y la extensión de las oraciones*. [Trabajo final de graduación, Universidad Nacional, Costa Rica]. Repositorio Académico Institucional (RAI) de la Universidad Nacional. <https://repositorio.una.ac.cr/bitstream/handle/11056/14299/Monge%20Naranjo%2c%20Ana%20Luisa.pdf?sequence=4&isAllowed=y>
- Montes Fernández, A. (2007). La Textología Contrastiva – método de análisis para la traducción de textos publicitarios. *Lebende Sprachen*, 4, 150-155. https://rua.ua.es/dspace/bitstream/10045/45549/1/2007_Montes_Lebende-Sprachen.pdf
- Mora Campos, C. (2003). *Un encuentro con la traducción automática. Análisis comparativo de un texto traducido bajo diferentes modalidades de traducción*. [Trabajo final de graduación, Universidad Nacional, Costa Rica]. UNA Translation Research Access Manager. <https://www.mogap.net/pmt/carlosmora.pdf>
- Mossop, B. (2014). *Revising and editing for translators*. Nueva York: Routledge.
- Mugierza, P. (2019). *Manual de traducción inglés-español de protocolos de ensayos clínicos* (2.^a ed.). Barcelona: Fundación Dr. Antonio Esteve.
- Parra, C. (2018). ¿Cómo ha evolucionado la traducción automática en los últimos años? *La Linterna del Traductor*, 16, 20-27. http://www.lalinternadeltraductor.org/pdf/lalinterna_n16.pdf

- Registro Español de Ensayos Clínicos (REEC). (s. f.). <https://reec.aemps.es/reec/public/web.html>
- Retana Mora, C. (2008). *Las memorias de traducción en el proceso y el producto de la traducción especializada en Costa Rica*. [Trabajo final de graduación, Universidad Nacional, Costa Rica]. Repositorio Académico Institucional (RAI) de la Universidad Nacional. <https://repositorio.una.ac.cr/bitstream/handle/11056/17795/Las%20memorias%20de%20traducción%20en%20el%20proceso%20y%20el%20producto%20de%20la%20traducción%20especializada%20en%20Costa%20Rica.pdf?sequence=1&isAllowed=y>
- Sáenz-Campos, D. (2005). Protocolos para investigación clínica: aspectos prácticos para su preparación. *fármacos*, 18, 1-2. <https://www.binasss.sa.cr/revistas/farmacos/v18n1-2/art6.pdf>
- Sánchez, A. C. (2006). Textos, tipos de texto y textos especializados. *Revista de filología*, 24, 77-90.
- Sánchez-Gijón, P. (2016). La posesición: hacia una definición competencial del perfil y una descripción multidimensional del fenómeno. *Sendeban*, 27, 151-162. DOI <https://doi.org/10.30827/sendeban.v27i0.4016>
- Seoane Vicente, Á. L. (2015). *Lenguaje controlado aplicado a la traducción automática de prospectos farmacéuticos*. [Tesis doctoral, Universidad de Alicante, España]. Repositorio Institucional de la Universidad de Alicante. https://rua.ua.es/dspace/bitstream/10045/53587/1/tesis_angel_luis%20seoane.pdf
- Shift Languages. (2020). Detrás del clic III: ventajas y oportunidades de la traducción automática. *Shift Languages*. <https://shiftlanguages.com/ar/detras-del-clic-iii-ventajas-y-oportunidades-de-la-traducion-automatica/>
- SYSTRAN. (2020). ¿Qué es la traducción automática? Traducción automática basada en reglas vs. traducción automática estadística. *Systran beyond language*. <https://www.systransoft.com/es/systran/tecnologia/que-es-la-traducion-automatica/>
- Tertoolen, R. C. (2012). Desarrollos en la traducción automática: esperando aún una traducción de alta calidad. *Caracteres: estudios culturales y críticos de la esfera digital*, 1(2), 141-147.
- Tripathi, S. & Sarkhel, J. K. (2010). Approaches to machine translation. *Annals of Library and Information Studies*, 57, 388-393. <http://nopr.niscair.res.in/bitstream/123456789/11057/4/ALIS%2057%284%29%20388-393.pdf>
- U.S. National Library of Medicine. (s. f.). Interventional study (clinical trial). *ClinicalTrials.gov*. <https://clinicaltrials.gov/ct2/about-studies/glossary>
- U.S. National Library of Medicine. (s. f.). Learn About clinical trials. *ClinicalTrials.gov*. <https://clinicaltrials.gov/ct2/about-studies/learn#WhatIs>
- Weaver, W. (4 de marzo de 1947). Carta a Norbert Wiener. Cambridge, Massachusetts. <https://aclanthology.org/www.mt-archive.info/Weaver-1947-original.pdf>

Weaver, W. (15 de julio de 1949). «Translation». Carlsbad, Nuevo México. <https://aclanthology.org/1952.earlymt-1.1.pdf>

Cilampa (segunda serie) es la colección de cuadernillos temáticos y de divulgación científica de la Escuela de Literatura y Ciencias del Lenguaje; tiene como propósitos mayores, primero, la edición, digital e impresa, de obras de referencia y demostrada solvencia académica; segundo, la creación de una interesante oferta de materiales didácticos. Se trata de publicaciones breves, no superiores a ochenta páginas, destinadas a dar cuenta del desarrollo de las distintas áreas de conocimiento inscritas en el seno de la Escuela de Literatura y Ciencias del Lenguaje; publicaciones cuidadosas en el desarrollo del contenido técnico, concebidas mediante líneas editoriales específicas, escritas en un estilo llano y explicativo, de naturaleza monográfica o miscelánea, diseñadas con esmero y editadas en arreglo con los modernos preceptos editoriales.

COMISIÓN EDITORIAL DE LA ESCUELA DE LITERATURA Y CIENCIAS DEL LENGUAJE
UNIVERSIDAD NACIONAL (COSTA RICA)

D. Gabriel Baltodano
D.^a Sherry E. Gapper
D.^a Mayra Loaiza
D. Carlos Francisco Monge
D. Jimmy Ramírez Acosta
D. Andrew Smith
D. Francisco Vargas Gómez

Impreso en Heredia

Fiel a su cometido («la edición [...] de obras de referencia y demostrada solvencia académica [y] la creación de una interesante oferta de materiales didácticos»), la Colección Cilampa ofrece al lector interesado este tomo, un texto introductorio a la traducción automática, que se conjuga con nociones básicas de lingüística textual y estilística. Tras revisar una serie de antecedentes, se exploran asuntos cardinales sobre esta modalidad de traducción (su historia, definición, tipos, ventajas, desventajas y relación con la posesición), a partir de cuales reflexiona sobre lingüística del texto, géneros textuales, estilo y estilística contrastiva, para finalizar con un ejemplo de las aplicaciones y utilidad prácticas que la traducción automática y los procesos de posesición pueden suponer para los investigadores y profesionales de la traducción.

ALEJANDRO ROJAS WALKER

Graduado de la Maestría Profesional en Traducción (Inglés-Español) de la Universidad Nacional (UNA). Cuenta con estudios universitarios en literatura inglesa (Universidad de Costa Rica) e ingeniería informática (UNED). Colabora con el Programa de Traducción e Interpretación (ProgTI) de la Escuela de Literatura y Ciencias del Lenguaje (UNA). Actualmente labora para la empresa Lionbridge.