PITFALLS IN MACHINE TRANSLATION: A STUDY ON WEB-BASED NEWS ARTICLES

TRAMPAS EN LA TRADUCCIÓN AUTOMÁTICA: UN ESTUDIO BASADO EN TEXTOS PERIODÍSTICOS EN LÍNEA

Sara Cecilia Moreno-McMahon
cecymcmahon@msn.com
Yasmine Beale-Rivaya
Professor of Spanish Linguistics, Texas State University
NEH Distinguished Professor in the Humanities
yb10@txstate.edu

Recibido: 7 de septiembre 2021. Aceptado: 1 de diciembre 2021

Abstract: A quality translation involves a process that goes beyond matching words, phrases, and sentences between two languages. Machine Translation (MT) systems use applications to translate text from one language to another without the intervention of a human translator (TH), resulting in poor translations. Through this research, we identified the most frequent errors in translations produced by automatic translation systems from English to Spanish. We analyze MT's limitations regarding grammatical, semantic, and lexical elements in a comparative and statistical analysis. Google Translate™ (GT) and Bing Translator™ (BT) were used to translate twenty news stories from English to Spanish. Both services ensure that, through neural machine translation (TAN), systems can "think" for themselves without human intervention. TAN

systems have achieved impressive results using state-of-the-art sentence processing techniques. Although these systems appear to be very promising, they still produce many errors in their translations.

Keywords/Palabras claves: Automatic translation/Traducción automática, errors in translation/errores de traducción, post-edition/posedición, journalism in Spanish/periodismo en español, Spanish in the United States/español de Los Estados Unidos.

Resumen: Una traducción de calidad conlleva un proceso que va más allá de emparejar palabras, frases y oraciones entre dos idiomas. Los sistemas de traducción automática (TA) utilizan aplicaciones para traducir un texto de un lenguaje a otro sin la intervención de un traductor humano (TH) y, como resultado producen traducciones poco satisfactorias. Mediante esta investigación, identificamos los errores más frecuentes en las traducciones que producen los sistemas de TA del inglés al español. Analizamos las limitaciones de la TA en lo que respecta a los elementos gramaticales, semánticos y léxicos en un análisis comparativo y estadístico. Google Translate (GT) y Bing Translator™ (BT) sirvieron para traducir veinte noticias periodísticas del inglés al español. Ambos servicios aseguran que, a través de la traducción automática neuronal (TAN), los sistemas pueden "pensar" por si mismos sin la intervención humana. Los sistemas de TAN han logrado impresionantes resultados mediante técnicas de procesamiento de oraciones con la última tecnología de vanguardia. Aunque estos sistemas aparentan ser muy prometedores, aun producen un gran número de errores en sus traducciones.

1. Introducción

La traducción automática (TA) se ha convertido en una herramienta indispensable entre las agencias de traducción y las grandes empresas internacionales. Para facilitar el trabajo del traductor humano (TH) y, para aliviar la demanda

que requiere el mercado internacional, se han implementado procesos para la internacionalización de la comunicación digital. Existe una plétora de servicios de traducciones que han mejorado su tecnología para proporcionar los beneficios de productividad que las empresas requieren. Las compañías que operan las traducciones automáticas hacen posible la traducción en poco tiempo y a bajo costo. Según el sitio de traductores Proz.com, el costo promedio de traducciones por traductores humanos (TH) es de diez centavos por palabra o veinticinco dólares por hora para las traducciones del inglés al español. El servicio de traducción de Google es gratuito para los usuarios con un límite de hasta cinco mil palabras. Para las traducciones en grandes cantidades de texto, Google ofrece el servicio de traducciones con mensualidades de 20 dólares, lo cual facilita la traducción de hasta un millón de palabras por mes[1]. El bajo costo de las traducciones automáticas pone en desventaja al TH, dado que la mayoría de los usuarios desconocen si los resultados de las traducciones son de calidad. El propósito de esta investigación es identificar los errores más frecuentes en las traducciones que producen los sistemas de TA del inglés al español. Analizamos las limitaciones de la TA en lo que respecta a los elementos gramaticales, semánticos y léxicos en un análisis comparativo entre la TA y el TH.

Pese a los beneficios de la TA, el proceso de traducción puede ser muy complejo, si se toma en cuenta que muchas palabras tienen múltiples significados, las oraciones pueden tener varias interpretaciones posibles y, ciertas estructuras gramaticales en un idioma pueden no existir o traducirse claramente a otro. Además, existen factores extralingüísticos a considerar en una traducción exitosa, como un amplio conocimiento de la cultura y el contexto. El proceso para una traducción de especialidad requiere ciertos aspectos como el estilo, la terminología y, en muchas ocasiones guías de estilo y una investigación minuciosa. Los textos de especialidad, como el periodismo, exigen ciertos criterios en la redacción. La función principal de los géneros periodísticos es de infor-

mar y la información puede transmitirse en diferentes estilos o con diferente intención. El objetivo siempre tiene que ser el mismo: fácil de entender para los lectores, especialmente en un contexto multilingüe como lo es los Estados Unidos (Guerrero 126).[2]

Las características que distinguen el español estadounidense de otros países monolingües son las variedades morfosintácticas y fonológicas de los distintitos países que forman la población de hispanohablantes en los Estados Unidos (Jiménez 45). Lo que estas variedades del español tienen en común es una indiscutible influencia del inglés, que se manifiesta según el tiempo de residencia del hablante en los Estados Unidos. A nivel morfológico, la prefijación y sufijación nominal se presenta en las combinaciones léxicas derivadas del inglés adaptadas al español, como *lunch box* / *lonchera*, o *truck driver* / *troquero*. Las formas infinitivas del inglés al sistema verbal español resultan en conjugaciones como, *to leak* / *liquear*; *to print* / *printear; to quit* / *cuitear* (Pacchiarotti 165).

Los calcos semánticos dan un nuevo significado a palabras españolas en virtud de su semejanza con el inglés, como, *application* / aplicación; *registration* / registración. También se forman dobletes léxicos en la lengua receptora, dado que adquieren el significado de la lengua donante y se sobreponen con términos que tienen el mismo significado en español como, atender / asistir; aplicación / solicitud (Pacchiarotti,165). Los hispanohablantes en Estados Unidos también usan ciertas frases inglesas, que eliminan la forma equivalente del español al incluir calcos sintácticos como "no son tus negocios", del inglés *it's none of your business* (Torres 412). John Lipsky explica que este tipo de calcos no viola ninguna regla sintáctica o de selección léxica del español, sino que se incorpora en el repertorio de modismos regionales (239).

En los Estados Unidos no existe una lengua oficial dado que, la Constitución garantiza el derecho a sus ciudadanos y residentes a usar cualquier lengua. A pesar del gran crecimiento demográfico y económico, el español aún no se ha

impuesto en el país por razones de prestigio social, científico y cultural como en el caso de la lengua inglesa que goza del apoyo y el prestigio en el medio lingüístico (Marcos Marín, 112). Es por esta razón, que se encuentran errores frecuentes en los textos redactados en español debido a la falta de conciencia de la población y de una norma que garantice una comunicación entre sus usuarios (113).

Las noticias disponibles que se redactan originalmente en español, tales como Telemundo o Univisión, son ejemplos de las pocas fuentes informativas disponibles para los hispanohablantes. A pesar de la creciente influencia de los hispanohablantes en EE. UU., la cobertura de noticias de la nación tiende a servir a personas que se comunican principalmente en inglés.[3] La población hispana de EE. UU. alcanzó un récord de 60,6 millones en 2019, un aumento de 930.000 en comparación al año anterior[4], según las estimaciones de población de la Oficina del Censo de EE. UU. recientemente publicadas (Bustamante sin pág.). Según el reporte anual de 2019 del Instituto Cervantes la población en EE. UU. es de 328,2 millones de personas. El 17,8 % se identifica como hispano y, de este grupo el 71,6 % habla español en sus hogares. Las proyecciones del informe 2019 del Instituto Cervantes indican que, en el 2060, EE. UU. será el segundo país hispanohablante del mundo, después de México, y casi uno de cada tres estadounidenses será hispano. Cerca de 22 millones de alumnos estudian español como segundo idioma según datos del año 2019 (Fernández Vítores, David, y García Montero 10).

Pese a la demanda, los medios de comunicación no cuentan con fuentes de noticias suficientes para los hablantes del idioma español en EE. UU. Los resultados de una investigación realizada por Jorge Ignacio Covarrubias muestran que la mayoría de los medios de noticias tanto impresos como digitales son distribuidos gratuitamente y contienen un gran número de anuncios de propaganda para poder subsistir (7). Covarrubias también encontró que más del 60 % de los periodistas consultados respondieron que los periódicos en es-

pañol contribuyen a la conservación del idioma, y una gran mayoría de los encuestados prefieren las noticias en español. De los 93 participantes que respondieron específicamente a la pregunta: "Para los hispanos, ¿conviene más ofrecerles medios en español o bilingües?", el 61.29 % prefiere las noticias en español, el 30 % bilingües y el 8 % cualquiera de las dos (7-41).

La introducción de la tecnología digital en la comunicación ha fomentado el incremento de neologismos cibernéticos en inglés adaptados al español. El español de Estados Unidos gira en torno a la diglosia que se ajusta a la compleja diversidad del español estadounidense. El español informal se encuentra con mayor frecuencia en las redes sociales. Por otro lado, el español formal sirve para la comunicación en contextos gubernamentales, instituciones y artículos periodísticos.

Según Covarrubias, "Los medios digitales representan una proporción ligeramente mayor al de los medios únicamente impresos". Lo que significa que, con el auge de la información digital y el fácil acceso a los dispositivos electrónicos, los medios de comunicación dependen de la TA en línea para hacer llegar las noticias al mundo (6). Gracias a la ayuda de los servicios gratuitos de TA, las noticias se pueden traducir a una infinidad de idiomas en un instante. Sin embargo, estas traducciones son un ejemplo de los retos que aún existen dentro de la practicidad y velocidad de esta tecnología.

2. Estado actual de la TA:

La TA fue una de las primeras aplicaciones creadas en las computadoras para descifrar códigos durante la Segunda Guerra Mundial. El Memorando Weaver, creado por el matemático Warren Weaver, marcó la pauta para la investigación en la TA (Bowker y Buitrago 37-38). Desde el inicio de las traducciones automáticas en los años cincuenta, los expertos en inteligencia artificial aseguraban que la traducción se podría automatizar para que estas máquinas pudieran traducir tex-

tos técnicos. Por consiguiente, podrían reducir el costo de las traducciones principalmente para las empresas, organismos internacionales y servicios de inteligencia militar. Los resultados no fueron obviamente muy satisfactorios dado que los sistemas no estaban lo suficientemente desarrollados como para reproducir los procesos similares al de un humano. Una traducción de calidad implica la capacidad de reproducir procesos cognitivos muy complejos y difíciles de crear, sin importar los avances y mejoras de la tecnología (Oliver 28).

La TA estuvo dominada por los sistemas basados en conocimientos lingüísticos en sus inicios. En la década de los noventa se implantaron las memorias de traducción y los sistemas estadísticos de traducción que consistían en extraer el conocimiento de corpus paralelos. Los investigadores centraron su atención en el desarrollo de herramientas de traducción asistidas por computadora, como son los sistemas de memoria de traducción. Las herramientas de Traducción Asistida por Computadora (TAC) no tenían como objetivo realizar la tarea de traducción. Su función era ayudar a los traductores profesionales en tareas automatizadas, como la búsqueda de diccionarios o comparar sus traducciones con los textos traducidos previamente. Poco a poco, la traducción automática y los sistemas de memoria se fueron integrando a los procesos de traducción y, por ende, permitieron a los traductores profesionales hacer uso de ambos tipos de tecnología (Bowker y Buitrago 38).

Hoy en día, el modelo neuronal es la tecnología de vanguardia de la traducción automática. Un modelo neuronal es un tipo de modelo estadístico formado por un conjunto de unidades de proceso simple, densamente conectadas entre sí (Casacuberta y Peris 66). Estos sistemas representan de forma numérica las palabras y las frases. Con estos sistemas se utilizan las técnicas del aprendizaje profundo, cuya función consiste en imitar el cerebro humano para realizar funciones cognitivas, como la resolución de problemas y el aprendizaje. Un humano adquiere conocimiento a través de la experiencia y, de esa misma manera las neuronas artificiales lo adquieren (Chen 6).

Algunas compañías se han sumado a la lista de servidores

de la TA como Google, Amazon y Bing, en una contienda por conseguir la tecnología más avanzada. La página de servicios de traducciones Amazon Translate™ explica los beneficios de la nueva tecnología de TA a base de redes neuronales (RN). Según Amazon, las RN son una forma de automatización para ofrecer traducciones más naturales y precisas que los algoritmos estadísticos tradicionales y, de traducción basados en reglas. Además, presenta la TA como una alternativa más eficaz que la TH. Amazon Translate™ asegura que los sistemas de TA son la mejor opción para empresas con miras a la internacionalización de páginas web: "Para los equipos de traductores humanos resulta muy difícil ponerse al día con un contenido dinámico. Con Amazon Translate™, se pueden traducir fácilmente enormes volúmenes de contenido generado por el usuario al instante."

El director de ingeniería de GT, Macduff Hughes, afirmó que, en el 2016 se produjo un gran cambio en el que Google incorporó también la traducción automática a base de RN. A diferencia del antiguo método que funcionaba traduciendo sintagmas y palabra por palabra, ahora traduce la oración completa. La simplicidad de este sistema se encuentra en que no es necesario considerar la sintaxis de la LM. Con el nuevo método, Google puede utilizar los millones de traducciones existentes en la web y así seleccionar la traducción más acertada. La calidad de la traducción depende de la cantidad de traducciones con que la RN esté entrenada y condicionada para hacerlo. Al introducir un nuevo idioma al sistema de RN, es fundamental que existan en la web los datos fiables para entrenar al sistema (Rubio sin pág.).

3. Método del estudio

Los métodos para la evaluación automática consisten en sistemas métricos que pueden comparar los resultados de la TA con la traducción profesional de un humano. Las evaluaciones de paridad entre humano-máquina, por lo general

se miden exclusivamente a nivel de oración, debido a que la mayoría de los sistemas de TA no consideran el contexto más allá de los límites de las oraciones. (Laubli, 656). Al seleccionar los criterios de evaluación y a través de una comparación lingüística pudimos determinar si los errores y ambigüedades alteran el mensaje original, o si las traducciones resultan coherentes. Las aplicaciones de GT y BT en los dispositivos móvil facilitaron el proceso de traducción de los textos, con un promedio de 1500 palabras por cada artículo de noticias, suficientes para una evaluación comprensiva.

Tanto los textos periodísticos originales en inglés como su correspondiente traducción al español en GT y BT se introdujeron a una hoja de cálculo para formar un corpus de análisis paralelo. Pudimos identificar y clasificar en los textos traducidos los errores según el tipo y, de esta manera, se puede examinar si existe una diferencia entre las traducciones de Bing y Google.

4. Evaluación cuantitativa

Los resultados de esta investigación muestran que los errores identificados en GT y BT son muy similares. Estos errores, obtenidos de veinte artículos periodísticos, se clasificaron en siete categorías. De acuerdo con esta clasificación, se analizaron las diferencias entre GT y BT. Se realizó la prueba de CHI-Cuadrada de homogeneidad con la hipótesis de investigación (o alternativa) para determinar si existen diferencias significativas entre los errores de las dos plataformas de TA. Como vemos en la tabla 1 y en la figura 1, la frecuencia de los errores muestra que existen pequeñas diferencias entre GT y BT con respecto a los errores por categoría. En las tablas A y B podemos observar que las diferencias observadas en la frecuencia de errores en ambas plataformas no son significativas. Se puede interpretar que ambas plataformas presentan la misma frecuencia de errores en las 20 historias analizadas.

Tabla 1. Frecuencia de errores identificados en GT y BT por categoría.

	Google Translate	Bing Translator
Ambigüedad	40	35
Calco léxico literal	315	324
Calcos estructurales	117	117
Calco semántico	52	53
Falsos amigos	34	44
Frase idiomática	23	22
Traducción incompleta	18	22
Total	**599**	**617**

Figura 1. Comparación de la frecuencia de errores de Google y Bing.

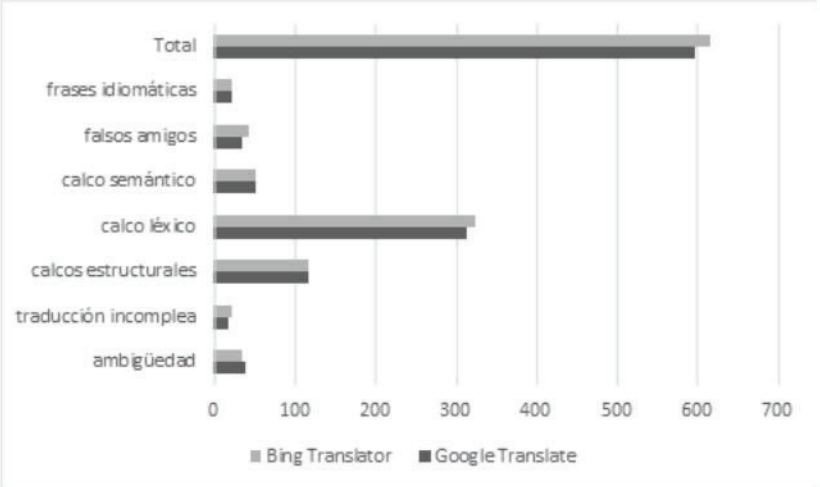


Tabla A. Análisis de homogeneidad con Chi cuadrado (X^2).

Frecuencia observada			
	Google	**Bing**	
Ambigüedad	40	35	75
Traducción incompleta	18	22	40
Calcos estructurales	117	117	234
Calco léxico	315	324	639
Calco semántico	52	53	105
Falsos amigos	34	44	78
Frases idiomáticas	23	22	45
	599	617	1216
Frecuencia esperada			
	Google	**Bing**	
Ambigüedad	36.89154569	38.10845431	75
Traducción incompleta	19.67549103	20.32450897	40
Calcos estructurales	115.1016225	118.8983775	234
Calco léxico	314.3159693	324.6840307	639
Calco semántico	51.64816396	53.35183604	105
Falsos amigos	38.36720751	39.63279249	78
Frases idiomáticas	22.16694079	22.83305921	45
	576	595	1171

Tabla B. Prueba de homogeneidad.

	Google		**Bing**	
	Observada	Esperada	Observada	Esperada
Ambigüedad	40	36.89154569	35	38.10845431
Traducción incompleta	18	19.67549103	22	20.32450897
Calcos estructurales	117	115.1016225	117	118.8983775
Calco léxico	315	314.3159693	324	324.6840307
Calco semántico	52	51.64816396	53	53.35183604
Falsos amigos	34	38.36720751	44	39.63279249
Frases idiomáticas	23	22.16694079	22	22.83305921
Resultado: 1.841368				

En la tabla 2, se observan los resultados en un esquema más detallado con los datos obtenidos de las traducciones de GT. No existen diferencias significativas entre las dos plataformas, y por lo tanto se analizaron solo los datos de Google para presentar estadísticas descriptivas de los errores en 20 artículos de noticias.

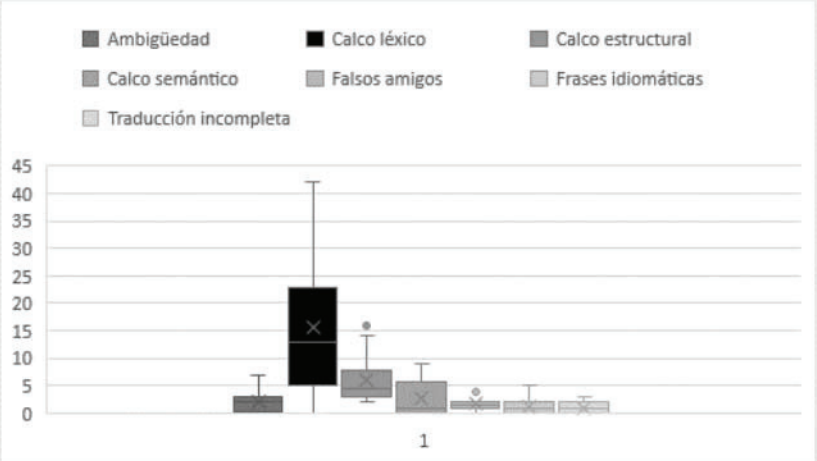
Tabla 2. Frecuencia de errores en GT por categoría y por artículo.

	Ambigüedad	Calco léxico	Calco estructural	Calco semántico	Falsos amigos	Frases idiomáticas	Traducción incompleta	Total
1	1	5	4	2	1	0	2	15
2	1	20	2	1	2	0	1	27
3	7	22	7	1	1	5	2	45
4	2	12	2	0	1	0	1	18
5	2	11	5	9	2	1	1	31
6	1	13	5	7	2	2	3	33
7	3	5	3	0	0	1	0	12
8	0	2	3	0	2	0	0	7
9	2	0	3	5	0	0	2	12
10	2	21	11	6	4	0	0	44
11	6	35	16	1	2	2	2	64
12	4	42	14	6	1	4	0	71
13	3	32	8	5	4	0	1	53
14	0	5	7	0	0	0	1	13
15	2	3	4	1	4	0	0	14
16	0	6	9	0	4	2	0	21
17	0	13	6	0	2	1	0	22
18	3	23	2	0	1	1	0	30
19	1	22	2	6	1	1	0	33
20	0	23	4	2	0	2	3	34
	40	**315**	**117**	**52**	**34**	**22**	**18**	**599**

El análisis de una sola plataforma de traducción nos permite observar los resultados de cada categoría con más de-

talle. La figura 2 muestra la distribución de errores representados en cajas de acuerdo con el número de errores. La categoría de calco léxico cuenta con el mayor número de errores, mientras que la categoría de traducción incompleta está representada con el menor número de errores. Esto también se puede observar en los totales por categorías en la tabla 2. Independientemente de las categorías, todas las observaciones del número de errores son sesgadas y por lo tanto asimétricas. Es decir, la mayoría de las historias tienden a tener un número bajo de errores mientras que uno o dos artículos periodísticos acumularon un mayor número de errores. Por ejemplo, en la historia 12 de la tabla 2 en la categoría del calco léxico se encontraron 42 errores.

Figura 2. Distribución de valores en las categorías de errores.



En la figura 2, la barra horizontal en medio de la caja representa la mediana y la x la media. Se observa en forma cuantitativa y consistente con los diagramas de las distribuciones que la categoría de calco léxico tiene un promedio mayor de número de errores por historia (una mediana de 13 errores por historia). Esta categoría tiene en promedio la dispersión más grande del número de errores (desviación estándar de

11.8 errores). Las demás categorías tienen promedios (medianas) que van de 1 a 5 errores por historia con poca dispersión en comparación con otras categorías. La distribución en una escala horizontal no es simétrica y la mayoría de los errores se encuentran en las categorías que tienden a tener un valor relativamente menor (categorías en azul). La categoría de calcos léxicos (en negro) pertenece a la minoría en el diagrama horizontal, pero con un valor relativamente mayor.

La tabla 3 indica que el promedio de errores por historia en esta categoría es de cerca de 16 errores y, la mayoría de las historias tienen menos de 22 errores. La tabla 3 presenta las distribuciones y los resultados cuantitativos del análisis con las estadísticas de media central y de dispersión del número de errores por categoría. Hemos observado que los valores altos alteran la media de una manera significativa sin dar una medida adecuada del centro de la distribución.

Tabla 3. Comparación de la media, mediana y desviación estándar.

Categoría	Media	Mediana	Desviación estándar
Ambigüedad	2.00	2	1.9
Calco léxico	15.75	13	11.8
Calcos estructurales	5.85	5	4.03
Calco semántico	2.60	1	2.95
Falsos amigos	1.70	2	1.38
Frase idiomática	1.10	1	1.41
Traducción incompleta	0.95	1	1.05

Aunque la categoría de calcos léxicos sobresale en la frecuencia de errores, en una distribución estadística podemos identificar las diferencias de valores entre todas las categorías.

5. Evaluación y análisis

Para interpretar el significado de un texto, un traductor

analiza el mensaje de la lengua fuente (LF) y todos sus componentes, incluyendo un conocimiento amplio de la gramática, sintaxis y semántica de la lengua meta. El número de errores en las traducciones al español y los resultados mostraron que tanto GT como BT presentaron serios errores semánticos y sintácticos que afectaron la calidad y fidelidad de las traducciones. Los resultados muestran que los sistemas de TA actuales tienen muchas limitaciones al traducir oraciones complejas y, por consiguiente, tienden a transferir la estructura y el léxico del inglés al español, con resultados muy literales.

Al evaluar la calidad de los textos traducidos, encontramos que el principal problema se encuentra en los calcos léxicos y los falsos amigos por el grado de semejanza de las palabras entre el inglés y el español en los errores del vocabulario. La categoría de calcos estructurales muestra las formas inglesas que se adaptan al español incluyendo la omisión del artículo, la estructura de la oración, el tiempo verbal, errores en el uso de preposiciones o incoherencias que cambian el significado de las frases.

La tabla 4 representa los datos organizados según la categoría y el número de errores. A su vez, estos se dividieron en dos partes: la primera representa el grupo de los errores relacionados con la semántica, que lo forman las irregularidades en el uso de los cognados falsos, la ambigüedad y lo calcos léxicos. El segundo grupo de la tabla 4 enumera el uso incorrecto u omisión de artículos y preposiciones. Se identificaron algunos errores producidos por la adaptación a la normativa en inglés, tales como el abuso del pretérito pluscuamperfecto, concordancia de sujeto y verbo y, en algunos casos, la concordancia del tiempo verbal y de género.

Tabla 4. Frecuencia de errores y categorías.

Errores semánticos	**Google Translate**	**Bing Translator**
Ambigüedad	40	35

Calco léxico literal	315	324
Calco semántico	52	53
falsos amigos	34	44
Traducción incompleta	18	22
Calcos estructurales		
construcción verbal	1	2
concordancia de género	10	7
frase idiomática	23	24
falta del artículo	14	13
preposiciones	7	7
tiempo verbal	10	12
estructura oracional	52	52

La evaluación de los 20 artículos seleccionados demuestra que tanto Google como Bing produjeron traducciones muy similares: 599 errores en GT y 617 en BT en total. El mayor número de errores pertenece a la categoría de calco léxico literal: GT con 315 errores y BT con 324. Ambas aplicaciones fallaron en la selección apropiada del significado del léxico. Por esta razón, los resultados sugieren que los sistemas de TA no son capaces de traducir algunas de las palabras que presentan polisemia. Debido a que los sistemas de TA operan a nivel oracional, sigue siendo un reto emparejar la semántica del léxico con el contexto.

En la Tabla 5, presentamos un cuadro comparativo de las traducciones de la TA (GT y BT) y la de un traductor humano (TH) y la frecuencia de errores.

Tabla 5. Ejemplo de error según categoría.

Calco léxico	LF	he couldn't **afford** to go to the doctor without insurance
	TA	no podía **permitirse** ir al médico sin seguro
	TH	no podía **darse el lujo** de ir al médico sin seguro
Calco semántico	LF	Canada is a top desired destination for **would-be emigrant** Americans.
	TA	Canadá es un destino ideal para **los aspirantes a estadounidenses emigrantes.**
	TH	Canadá es uno de los destinos principales para **los estadounidenses que desean emigrar a otros países**.
Falsos amigos	LF	Nominated in 1993 by former President Bill Clinton, the liberal **justice** was the legal force behind many of the successes of the women's movement
	TA	Nominada en 1993 por el ex presidente Bill Clinton, **la justicia** liberal fue la fuerza legal detrás de muchos de los éxitos del movimiento de mujeres
	TH	Nominada en 1993 por el expresidente Bill Clinton, **la juez** liberal fue la principal impulsora de los derechos de las mujeres.
Frases idiomáticas	LF	we're not out of the woods yet
	TA	aun no estamos fuera del bosque
	TH	aun no estamos fuera de peligro

Calcos

Se define como calco a la traducción literal de una palabra o sintagma de la LF incorporada a la LM, con o sin adaptación ortográfica. Estas se dividen en dos categorías: el calco léxico y el semántico. Gómez Capuz (2009) propone varias subcategorías del calco léxico, como el literal y el aproximado. Para el propósito de este estudio, emplearemos el calco

léxico literal, que se refiere a la transferencia de una palabra o compuesto nominal de la LF a la LM con un significado literal sin sentido en la lengua receptora e imita la estructura morfológica del modelo de la lengua original (8). Una gran mayoría de los errores en los textos analizados corresponden a las traducciones del calco léxico literal, que son una imitación exacta de palabras o sintagmas de la LF. La tabla contiene algunos ejemplos del calco léxico literal en la traducción.

Tabla 6. Calco léxico literal.

LF	TA	TH
sit-down conversation	conversación sentada	conversación en persona / extendida
he is doing very well	está haciendo muy bien	se siente bien / se siente mejor
walked back that request	caminó hacia atrás esa solicitud	canceló la solicitud
heard back from a staff member	escucharon de un miembro del personal	un empleado respondió
he spent his entire life savings	pasó toda su vida de ahorros	gastó todos sus ahorros

El calco léxico se limita a un contexto oracional. En la tabla 7, podemos observar el calco léxico literal en un contexto más amplio. La frase *school-based*, traducida como *basados en las escuelas* constituye un problema de traducción dado que es una frase inexistente en español. En el ejemplo A3 (tabla 7), el TH puede asociar inmediatamente la frase de la LF con las frases o palabras existentes en la lengua española, en este caso *escolares*. Los calcos ingleses no siempre se traducen de forma adecuada y afectan a la fidelidad que se busca en una traducción.

Tabla 7. Calco léxico literal en contexto.

A1	LF	...as **school-based** nutrition programs serve fewer students, districts are losing money.
A2	TA	...como los programas de nutrición **basados en las escuelas** atienden a menos estudiantes, los distritos están perdiendo dinero.
A3	TH	...como los programas **escolares** de nutrición atienden a menos estudiantes, los distritos están perdiendo dinero.

El siguiente ejemplo muestra los calcos en relación con el sujeto y la construcción verbal y, cómo al infiltrarse la estructura inglesa en la española interfiere con la semántica.

Tabla 8. Calco léxico en construcción verbal.

B1	LF	the president is doing very well.
B2	TA	el presidente estaba haciendo muy bien.
B3	TH	el presidente se siente mejor.

El factor del contexto en el proceso de traducción es crucial. Ambos ejemplos presentan una traducción aislada del contexto. Si incluimos el contexto en el ejemplo B1 de la tabla 8: *The physician tried to tamp down panic, declaring during a news conference Saturday the president was "doing very well,"* la frase *doing very well* puede traducirse de diferentes maneras, según el estilo y contexto. La LF indica que el presidente se siente mejor tras una condición médica. El verbo *hacer* en la LM es irrelevante en este contexto, por lo tanto, el TH en B1 puede explorar las opciones existentes en la LM para su elección: "está mejorando", "se siente mejor", "se está recuperando", entre otras. El ejemplo B2, es un prototipo del calco sintáctico o anglicismo. El ejemplo C2 de la tabla 9 se podría interpretar de diferentes maneras si se considera que, aunque la forma verbal "ha estado" del pretérito perfecto en tercera persona singular no concuerda con el sujeto. Una manera que el TH podría corregir la concordancia verbal sería: "Toda mi vida he estado." O cambiar la estructura completa-

mente sin perjudicar el mensaje original: "He pasado toda mi vida con él." También puede ser "Toda mi vida a girado a su alrededor". La correspondencia que el TH encuentra en *junto a él* concuerda con el significado del texto original.

Tabla 9. Calco léxico literal.

C1	LF	My whole life has been around him.
C2	TA	Toda mi vida ha estado a su alrededor
C3	TH	Toda mi vida la he pasado junto a él. /toda mi vida ha girado en torno a él.

Detectamos las limitaciones de GT en un caso de polisemia de la palabra *left*. Se observa la falta de concordancia verbal entre el léxico y la oración.

Tabla 10. Calco léxico literal y polisemia.

D1	LF	*Devon Austin left with her husband for Costa Rica.*
D2	TA	Devon Austin **izquierda**, con su esposo a Costa Rica.
D3	TH	Devon Austin **se fue** con su esposo a Costa Rica.

El ejemplo D2 (tabla 10) es un caso característico del problema de traducción que con frecuencia se presentó en la categoría de calco léxico literal. La selección errónea de una palabra distorsiona el sentido de la oración y en otros casos, la elección de una palabra parece tener sentido, pero la semántica de la oración cambia completamente según el contexto.

Tabla 11. Calco léxico literal en oraciones.

E1	LF	*It's **overwhelming, she** said. "There are so many people like me **looking".***
E2	TA	**Es abrumador**", dijo. "Hay tanta gente como yo **mirando**".
E3	TH	**Es de preocuparse**, dijo. "Hay tanta gente como yo **buscando trabajo**".

La TA generó una traducción en un concepto aislado don-

de aparentemente el resultado es correcto, (ver E2, tabla 11). E2 parece indicar que algo es tan abrumador y que obliga a alguien a mirar algo. La traducción por un TH (E3) es más específica y, aunque también aislada del contexto no aparenta concordancia con de la oración del inglés. Si a estas oraciones agregamos el contexto original del idioma fuente, el resultado de las traducciones cambia drásticamente.

Tabla 12. Calco léxico literal en contexto.

F1	LF	*Even companies that pay a third of what she once earned won't call her back.* "It's overwhelming," she said. "There are so many people like me **looking**".
F2	TA	Incluso las empresas que pagan un tercio de lo que una vez ganó no la devolverán la llamada. "**Es abrumador**", " dijo. "Hay tanta gente como yo **mirando**."
F3	TH	Incluso las empresas que pagan un tercio de lo que una vez ganó, no le devolverán la llamada. "**Es de preocuparse**", dijo. "Hay tanta gente como yo **buscando trabajo**."

Se observa que un concepto se puede traducir de varias formas, dado que la interpretación de la LF podría dar resultados bastante amplios. Por ejemplo, la palabra *looking* en inglés tiene múltiples significados en español, incluyendo *mirar* y *buscar*. En casos de polisemia, el traductor puede hacer las distinciones semánticas siempre y cuando el resultado de la traducción concuerde con el contexto. El problema de la pérdida léxica en la TA podría ser al mismo tiempo un síntoma o una causa de un problema más grave de los sistemas actuales. Cuando se presenta una oración ambigua en una traducción, el contexto es imprescindible para elegir la traducción que maximice la probabilidad en la claridad en la oración (Vanmassenhove, 1).

Calcos semánticos

A diferencia del calco léxico, el calco semántico modifica el significado de una palabra ya existente en la lengua receptora

y tan solo se limita a alterar los significados de esa palabra. (Gómez Capuz 11). El calco semántico puede aportar un significado parcial o total a la lengua receptora y pueden ser parónimos o no parónimos. Según Gómez Capuz, los calcos semánticos parónimos presentan una similitud formal con palabras de otra lengua que facilita la transferencia semántica, como en el caso de *carpet* y carpeta. Este tipo de calco también se les conoce como "falsos amigos", que veremos en más detalle en la siguiente sección. Los calcos semánticos no parónimos tienen una similitud conceptual entre palabras de dos lenguas, pero no formal ni etimológica, por ejemplo, "perro caliente" y *hot dog*. (14). En la siguiente tabla, se listan ejemplos en donde se detectan los calcos semánticos parónimos.

Tabla 13. Calco semántico (parónimo).

LF	LM	TH
...igniting *protests and violence here*	**. . . e n c e n d i e n d o** protestas y violencia	**...provocando** protestas y violencia
Monica helped ***handle*** *the piles of paperwork.*	Mónica ayudó a **manejar** los montones de papeleo.	Mónica ayudó con el papeleo acumulado.
...were ***struggling*** *with flood damages.*	...estaban **luchando** contra los daños por las inundaciones.	...estaban **sufriendo** por los daños de las inundaciones.
A ***parent*** *with two children*	Un **padre** con dos hijos	Un **padre o una madre** con dos hijos.
swing *states*	los estados **oscilantes**	Estados **indecisos**
Advocates ***target*** *Latinos*	Los defensores **apuntan** a los latinos	Activistas **intentan captar a** los latinos

Falsos amigos

A pesar de las diferencias genealógicas del inglés, una lengua germánica, y el español, una lengua romance, ambos idiomas aún comparten ciertos rasgos de un parentesco (Santos 17). El traductor tiene que poder reconocer las diferencias

semánticas en las similitudes ortográficas entre los cognados falsos y verdaderos. Se consideran cognados verdaderos cuando dos palabras en dos idiomas tienen la misma ortografía, etimología y significado como en los siguientes pares en inglés y español: *patio* y patio *art* y arte, *angel* y ángel. Los falsos amigos (FA) son pares de palabras que tienen una ortografía y pronunciación similar, pero con un significado muy diferente. Estas similitudes pueden conducir a asociaciones falsas, malentendidos, o distorsión del contexto (Vesiburgs, 628). No todos los FA son iguales, dado que se pueden clasificar de diversas maneras.

Estas clasificaciones incluyen a los FA semánticos cuyas palabras son ortográfica o fonológicamente similares en varias lenguas dadas y están relacionadas en su etimología, ya sea porque ambas proceden de una lengua en común como el latín, en *library* y librería. Los FA semánticos a su vez se clasifican en totales y parciales (Vesiburgs, 628). Los FA totales son aquellas palabras cuyos significados son completamente distintos, como en *salad* y salado. Los FA parciales son aquellos que en dos o más lenguas comparten algunos significados mientras difieren en otros (Chamizo 75). Un ejemplo de FA parcial se encuentra en la tabla 5 en la palabra *justice,* y su cognado en español *justicia.* Ambos idiomas comparten los significados de cualidad de ser justo, imparcialidad; pero difieren en el contexto de *justice* como un miembro de la Suprema Corte (*Supreme Court).* Las opciones que concuerdan con el contexto original, como en este ejemplo, la traducción que empareja con la LF es *juez.* La mayoría de los errores identificados en las traducciones dentro de esta categoría, pertenecen al tipo de cognados falsos parciales. Un ejemplo de cognado falso parcial se encontró en la palabra en inglés *move* y su cognado en español mover. Ambas palabras comparten el significado de hacer que un cuerpo deje el lugar y pase a ocupar otro, menear o agitar una cosa o parte del cuerpo. El siguiente ejemplo muestra la disparidad entre la semántica de ambos cognados y cómo el contexto es determinante en la elección de una palabra u otra.

Tabla 14. Falsos amigos.

G1	LF	*Migrant farmworkers, many coronavirus positive,* ***move*** *north from Florida to other states.*
G2	TA	Trabajadores agrícolas migrantes, muchos coronavirus positivos, se **mueven** hacia el norte de Florida a otros estados.
G3	TH	Los trabajadores agrícolas migrantes, muchos de ellos portadores del coronavirus, **se trasladan/mudan** de Florida a otros estados del norte.

Un cognado falso de la palabra en inglés *move* en español puede significar trasladar, mover, mudar, cambiar entre otras (G2 tabla 14). *Mover* no concuerda con la oración de la LF. Aunque *mover* comparte algunas características ortográficas y semánticas con *move*, en el contexto de trasladarse de una región a otra, *mover* pierde su lugar por una palabra que concuerde mejor con el contexto.

Calcos estructurales

Uno de los errores más comunes en el proceso de traducción es la confusión sintáctica en las preposiciones, artículos y la estructura de la oración. Cuando se sigue el modelo de la LF, se crean distorsiones en el significado y la coherencia en la LM. Los calcos identificados en la TA afectan principalmente la estructura sintáctica, preposiciones u omisiones de artículos. En el análisis de los textos se identificaron 117 errores estructurales en GT y BT. Los errores gramaticales que se muestran en la tabla 15 son ejemplos de la transferencia de las estructuras inglesa a la española. La omisión o adición de artículos no concuerda con las reglas de uso españolas. Los calcos estructurales pueden producirse por diferentes mecanismos, como el orden de las palabras que dan como resultado traducciones ambiguas. Aunque estos tipos de errores se consideren menos graves, se han incluido para demostrar las transferencias morfosintácticas de la LF a la LM. Sin embargo, la elección errónea de una preposición puede crear confusión en la semántica, como en el siguiente ejemplo.

Tabla 15. Calco estructural,

H1	LF	*Vasquez needed his father to recover from the coronavirus.*
H2	TA	Vásquez necesitaba **a** su padre para recuperarse del coronavirus
H3	TH	Vásquez necesitaba **que** su padre se recuperara del coronavirus.

El calco del inglés de la proposición *to* al *a* en ocasiones pasa desapercibido. En este caso no se puede ignorar por la ambigüedad que resulta en la traducción. Para resolver este dilema, podríamos preguntarnos si Vásquez es el enfermo que necesita a su padre para poderse recuperar del coronavirus o, si el padre de Vásquez es el enfermo que necesita recuperarse. Aún sin el contexto se puede deducir que la recuperación de Vásquez depende de una intervención médica y no necesariamente de su padre. Veamos el contexto de la LF: "*Vasquez, 40, who has lived with muscular dystrophy since birth, depends on his father to be his arms and legs*". Según el contexto podemos concluir que Vásquez necesitaba que su padre se recuperara por ser el encargado de su cuidado.

Tabla 16. Tipos de calco estructural.

LF	LM	Sugerencia de TH
The eyes are said to be windows to the soul	Se dice que los ojos son ventanas del alma.	Se dice que los ojos son **las** ventanas del alma.
demand limits on voting by mail.	exigen límites **a** votar por correo.	exigen límites **para** votar por correo.
And does it matter if it's for the greater good of the economy?	¿**Y** importa si es por el bien de la economía?	¿E importa si es por el bien de la economía? / y, ¿importa si es por el bien de la economía?
Hernandez's aunt was admitted to the hospital.	La tía de Hernández **fue ingresada** en el hospital.	Internaron a la tía de Hernández en el hospital.

Frases idiomáticas

Tradicionalmente las frases idiomáticas se consideran uno de los factores más idiosincrásicos de una comunidad de hablantes y por lo tanto su traducción puede ser compleja (Molina 429). Según Molina Plaza, en la búsqueda de la correspondencia de las unidades fraseológicas de una lengua a otra no se pueden establecer correspondencias automáticas dado que su complejidad requiere identificar la complejidad de los aspectos léxicos, semánticos, pragmáticos, textuales y discursivos de la expresión idiomática en el contexto de la LM. Todo depende de la dificultad de la frase y de la competencia del traductor (429). Los resultados de este estudio sugieren que las traducciones de GT y BT aún no cuentan con la suficiente competencia para traducir frases idiomáticas del inglés al español. Existen factores culturales y diferencias lingüísticas que influyen en la interpretación y traducción de modismos en las traducciones entre dos lenguas que solo el traductor que conoce ampliamente dichos factores podría descifrar. De los 20 artículos evaluados se identificaron 23 frases idiomáticas, de las cuales BT falló en cada una de ellas, mientras GT solo pudo captar una de ellas.

Table 17. Frases idiomáticas.

LF	TA	TH
business-as-usual	negocios como de costumbre	como si nada hubiera ocurrido
There's no rewinding the clock	No hay forma de rebobinar el reloj	No hay forma de regresar el tiempo
at the drop of a dime	a la caída de un centavo	en un abrir y cerrar de ojos
Justin is cut from the same cloth	Justin esta cortado de una tela similar	Justin está cortado por la misma tijera/ por el mismo patrón

Conclusión

Pese a ciertas ventajas de la TA, es fundamental subrayar que el proceso de posedición beneficia a cualquier traducción realizada automáticamente. Los resultados del análisis sugieren que la TA aún presenta serios errores en la falta de comprensión del contexto más allá de la oración. Las traducciones que producen los motores de TA pueden ser útiles si se incluye el proceso de posedición, ya que los resultados que ofrecen estos sistemas no logran la misma fidelidad que las traducciones revisadas por un TH. La combinación de ambos métodos se ha convertido en el modelo a seguir tanto por traductores profesionales, como por empresas proveedoras de servicios traducción.

La mayoría de los errores se encontraron en las categorías de los calcos léxicos literales, estructurales y semánticos. Se detectó que la elección del léxico fue el principal problema cuando la TA no considera el contexto. Tanto GT como BT fallaron en múltiples ocasiones al no poder solventar los problemas de sinonimia. Las traducciones relacionadas con los textos periodísticos no captaron la esencia y la brevedad que requieren los encabezados. Pudimos identificar que los patrones de selección léxica fueron aleatorios en su mayoría.

Tabla 18. Traducciones de encabezados periodísticos.

LF	Advocates Target Latinos for ACA Enrollment
TA	Los defensores **apuntan** a los latinos para la inscripción de ACA
TH	Activistas se enfocan en los latinos para inscripciones de ACA

El ejemplo de la tabla 18 muestra que el TH toma en cuenta que esta oración es un encabezado para una historia periodística, omite el artículo y selecciona el vocabulario que mejor se adapta al contexto.

Los sistemas de TA carecen de la habilidad de interpretar no solo el contexto, sino formular oraciones con la estructura gramatical española. La función de la traducción automática

neuronal (NMT por sus siglas en inglés) promete que estos sistemas superan a los sistemas predecesores que traducían frase por frase. Los resultados de esta investigación muestran que las traducciones de TA aún son insuficientes para la adecuación y fluidez del texto. Cuando la TA se aplica a las traducciones de textos de especialidad, como los textos periódicos, puede a ser un factor determinante en la calidad de las traducciones.

Representantes de Google afirman que el sistema de redes neuronales (RN) ha sido entrenado con una gran cantidad de traducciones y condicionado para mejorar la calidad. El control de calidad en Google Translate™ se basa en la aportación de usuarios quienes alertan a Google a través de mensajes o en las redes sociales. El sistema no corrige cada traducción incorrecta porque prefieren que el modelo haga su trabajo de "aprendizaje" y en ocasiones tendrán que corregir traducciones que sean ofensivas o engañosas (Rubio sin pág.).

La evaluación de la calidad de la TA ha sido objeto de controversias en la investigación de la industria de servicios de traducciones. Las evaluaciones de traducciones por traductores profesionales son más confiables a pesar de los altos costos que estas requieren. Sin embargo, persiste una tendencia en confiar en estudiantes y evaluadores novatos que a veces muestran un nivel limitado de competencia en los idiomas involucrados. Esto sucede con algunas compañías que ofrecen servicios de TA, ya que los traductores y evaluadores no expertos carecen del conocimiento de traducción y, por lo tanto, pueden no notar sutiles diferencias que hacen que una traducción sea más adecuada que otra. Cuando se enfrentan a una traducción que es difícil de editar, tienden a aceptar la traducción de TA en lugar de intentar mejorarla (Laubli 657).

Existe un mito sobre la traducción automática, al igual que ocurre cuando se plantea la posibilidad de automatizar tareas humanas. Por mucho tiempo, se ha creído que las maquinas reemplazarían a las personas y, por lo tanto, los traductores humanos se convertirían en una especie en ex-

tinción. La TA aún carece de la inteligibilidad, la fidelidad y la precisión requeridas en una traducción de calidad. La comprensión de un texto requiere de un conocimiento muy completo de la lengua de origen, de los mecanismos que procesan la información lingüística y del conocimiento cultural en la lengua fuente y meta (Oliver 31).

El ámbito de la traducción requiere que los traductores profesionales tengan un nivel avanzado no solo en los idiomas, sino un alto conocimiento en la materia de especialidad o la habilidad de investigación. Los últimos avances en la tecnología de la traducción automática prometen que la calidad ha superado sus expectativas y aseguran que la calidad ha llegado a un nivel casi como el de las traducciones realizadas por un humano. La TA, en una era de globalización, se ha convertido en la norma y en una necesidad en la industria de la traducción, pero al mismo tiempo reconocemos que la traducción humana y la evaluación posterior a la TA son indispensables para una eficacia completa.

Firefox about:blank

Works Cited/Obras Citadas

"Average Rates Charged for Translations." *www.Proz.Com*, 2020, https://search.proz.com/employers/rates.

Bustamante, Luis Noe, et al. "U.S. Hispanic Population Surpassed 60 Million in 2019, but Growth Has Slowed." *Pew Research Center*, 7 de julio de 2020.

Bowker, Lynne y Buitrago Ciro, Jairo. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Emerald Group Publishing, 2019.

Chamizo Domínguez, Pedro. "Variaciones representacionales y falsos amigos." *Contrastes*, no. 10, 2005, pp. 73 -103.

Casacuberta, Francisco y Peris, Álvaro. "Traducción automática neuronal". *Revista Tradumàtica*, núm. 15, diciembre de 2017, pp. 66 -72.

Chen, Zhi. "Un estudio empírico del alcance de la traducción automática del español al chino. Caso de estudio de GNMT sobre las expresiones metafóricas y metonímicas". *Círculo de lingüística aplicada a la comunicación*, vol. 83, julio de 2020, pp. 1- 24.

Covarrubias, Jorge Ignacio. "El periodismo en español en los Estados Unidos". Observatorio, marzo de 2016, pp. 1- 86.

Fernández Vítores, David, y García Montero, Luis. El español: una lengua viva, informe 2019. Anuario, Instituto Cervantes, 2019, pp. 1-96.

Guerrero, María José Hernández. "Técnicas específicas de la traducción periodística". *Quaderns*, núm. 13, 2006, pp. 125–39.

Gómez Capuz, Juan. "El tratamiento del préstamo lingüístico y el calco en los textos de bachillerato y en las obras divulgativas". *Tonos: Revista electrónica de estudios filológicos*, núm. 17, julio de 2009, pp. 1-24.

Jiménez, Rosario Outes. "El español escrito en el ámbito educativo de EE.UU.: la acreditación docente en California". *Glosas*, vol. 9, núm. 3, septiembre de 2017, pp. 44–57.

Laubli, Samuel, et al. "A Set of Recommendations for Assessing Human Machine Parityin Lan uage Translation." Journal of Artificial Intelligence Research, vol. 67, mar. 2020, pp. 653–72.

56 de 149 27/06/2022, 17:25

Lipski, John M. "La lengua española en los Estados Unidos: Avanza a la vez que retrocede". *Revista española de lingüística*, vol. 33, núm. 2, 2003, pp. 231–60.

Marcos Marín Francisco A. "RIUSS y el papel de La traducción en el español de los EUA". *II International Colloquiumon Languages, Cultures, Identity, in School and Society*, 2016, pp. 110–18.

Molina Plaza, Silvia. "La traducción de las unidades fraseológicas inglés- español:" *Las palabras del traductor*, 2004, pp. 427–43.

Moreno-Fernández, Francisco. "Español estadounidense: perfiles lingüísticos y sociales". *Glosas*, vol. 9, núm. 2, marzo de 2017, pp. 10–23.

Oliver González, Antoni, et al. *Traducción y tecnologías*. Editorial UOC, 2011.

Rubio, Isabel. "El ingeniero que te permite hablar más de 100 idiomas" | *Tecnología* | *EL PAÍS*, diciembre 2019, https://elpais.com/tecnologia/2019/11/25/actualidad/1574676898_992149.html.

Santos, Pascual, and Aquilino Sánchez. "El inglés y el español desde una perspectiva cuantitativa y distributiva: equivalencias y contrastes." *Estudios ingleses de la Universidad Complutense*, vol. 19, 2011, pp. 15–44.

Vanmassenhove, Eva, et al. "Lost in Translation: Loss and Decay of Linguistic Richness in Machine Translation.", junio de 2019.

Notas

1 Ver “Average Rates Charged for Translations” para más información sobre el costo de traducción.

2 La Academia Norteamericana de la Lengua Española (ANLE) se encarga de determinar las normas del español estadounidense formal, recogiendo como propios aquellos términos que han alcanzado un nivel de uso “culto” (“Estadounidismo no es espanglish”). Algunos de los objetivos principales de la ANLE son: “preservar el uso de la lengua y cuidar el uso de la variante hispano-estadounidense, ofrecer referencia a las instituciones públicas y privadas en el uso de las normas del español de EE. UU. y fortalecer un bilingüismo auténtico que enriquezca la cultura de Estados Unidos” (Academia Norteamericana de la Lengua Española).

3 Ver: Moreno-Fernández, Francisco (2017). Moreno-Fernández propone una realidad del español de Estados Unidos en torno a la diglosia (11). La variedad estadounidense se adquiere de forma oral y en los ambientes familiares donde la presencia del inglés es más evidente con la inclusión de préstamos, calcos, alternancia de códigos y anglicismos tales como la troca o camión; el torque o pavo. La subvariedad culta o general se adquiere a través de la familia, en el ámbito de la educación, contextos comunitarios más generales, en manifestaciones escritas y con menor influencia del inglés. Esta se utiliza en los medios de comunicación social, que coincide con las manifestaciones cultas del español de otros países hispanos (12-13).

4 Ver: Bustamante (2020) con respecto al crecimiento de la población hispana desde 1970 al 2019, según estimaciones de población de la Oficina del Censo de EE. UU.