

Percepción de los estudiantes de traducción sobre la traducción automática neuronal

Translation students' perceptions about neural machine translation

Elke Cases Berbel

Universidad Complutense de Madrid
España

ONOMÁZEIN 67 (marzo de 2025): 265-282

DOI: 10.7764/onomazein.67.12

ISSN: 0718-5758



Elke Cases Berbel: Departamento de Estudios Románicos, Franceses, Italianos y Traducción, Universidad Complutense de Madrid, España. | E-mail: ecases@ucm.es

Fecha de recepción: octubre de 2020

Fecha de aceptación: marzo de 2021

Resumen

El lanzamiento de la herramienta de traducción automática neuronal (TAN) en línea de DeepL en 2017 ha revolucionado el mundo de la traducción. Tanto es así que muchos estudiantes del grado de Traducción e Interpretación han comenzado a utilizarla de forma habitual, sin tener en cuenta las debilidades que este traductor en línea puede traer consigo. De forma unánime los aprendientes defienden que esta herramienta produce textos de calidad. Para ver si DeepL realmente traduce con parámetros de calidad superiores a las de un traductor humano, la docente ha realizado un experimento con estudiantes de la asignatura de traducción inversa español-alemán, del último curso del grado de Traducción e Interpretación de la Universidad Complutense de Madrid. Para ello, se ha dividido a los participantes en dos grupos: uno tendrá como objetivo la posesición de la traducción de DeepL, y el otro, la traducción sin ayuda tecnológica. En este artículo se presenta tanto dicha actividad como sus resultados y las conclusiones que se pueden extraer de este ejercicio didáctico.

Palabras clave: traducción; traducción automática; didáctica de la traducción; traducción inversa; DeepL.

Abstract

The launch of the online neural machine translation (NMT) tool from DeepL in 2017 has revolutionized the translation world. As a result, many translation students have started to use it indiscriminately, without taking into account the weaknesses that this online translator brings with it. The learners unanimously stated that this tool produces quality texts. To see if DeepL really translates with higher quality parameters than a human translator, the lecturer conducted an experiment with students of the back translation subject Spanish-German, taught in the last year of the Bachelor of Translation and Interpretation at the Universidad Complutense de Madrid. For this purpose, the participants have been divided into two groups, with the aim of post-editing the translation of DeepL, as well as the translation of the same text but without technological assistance. This article presents both, the performed activity as well as its results and the conclusions that can be drawn from this didactic exercise.

Keywords: translation; automatic translation; translation didactics; reverses translation; DeepL.

1. Introducción

Este artículo tiene el propósito de confrontar la percepción de los estudiantes sobre la herramienta del traductor automático neuronal en línea DeepL, así como examinar si, aun siendo una buena herramienta de ayuda a la traducción, requiere de la posesición y revisión de un traductor humano. Para ello se ha llevado a cabo un breve resumen de la historia de la traducción, que servirá para poner en perspectiva los principios y el desarrollo de la traducción automática (TA), lo que nos permitirá mostrar un enfoque más global de la evolución de esta herramienta. Después se presenta el experimento realizado junto con los resultados obtenidos y las conclusiones extraídas de estos.

La traducción automática es una herramienta que se persigue desde que comenzó el desarrollo de las tecnologías, ya que la traducción es una necesidad real para poder comprendernos mejor. La globalización sin la traducción sería inexistente. Así pues, tras decenios intentado crear una máquina que traduzca por sí sola, finalmente en 1966 el Comité asesor de tratamiento automático del lenguaje (ALPAC, por sus siglas en inglés), dirigido por John R. Pierce, concluyó en un informe que la traducción automática (TA) jamás tendría la calidad que le imprime un traductor humano (ALPAC, 1966: 124 y ss.).

No fue hasta 1991 cuando la idea de Warren Weaver de la traducción automática estadística (TAE), formulada en 1949, se retomó y desarrolló por el centro de investigación Thomas J. Watson, adscrito a la empresa multinacional estadounidense de tecnología y consultoría IBM (Brown y otros, 1993). Desde ese momento y hasta 2016, esta forma de TA ha sido la más utilizada en el mundo de la traducción.

La TAE funciona básicamente mediante la alineación de diferentes corpus ya traducidos por un traductor humano. No se rige por códigos lingüísticos, sino que opera por n-gramas: pequeños grupos de palabras (suelen ser entre cuatro y ocho), que aparecen tanto en la lengua de partida como en la de llegada (Werthmann y Witt, 2014: 98). Después de buscar todas las posibilidades, la TAE propone la que estadísticamente más se repite. Este algoritmo estadístico aplicado al aprendizaje produce un modelo de idioma. Cuanto más aparecen ciertas palabras o frases juntas, más las propondrá (Eberle, 2008).

Sin embargo, en 2016 irrumpió la traducción automática neuronal (TAN) de la mano de Google en el mundo de la traducción y con ella una nueva forma de entender la TA (Wu y otros, 2016). A diferencia de la TAE, la TAN siempre tiene en cuenta la frase entera, es decir que, de un mapeo lineal¹ (como es el caso de la TAE), pasamos a uno no lineal, basado en

1 Las palabras y las frases son representadas de forma numérica mediante vectores (Bengio y otros, 2003).

la representación vectorial de la lengua de partida y de llegada que se establece en varias etapas intermedias y con varios grados de abstracción (Schmalz, 2019: 198 y ss.). Consecuencia de esto es que los textos producidos tienen estructuras más naturales.

En agosto de 2017 la empresa alemana *Linguee* decide lanzar al mercado una herramienta TAE llamada *deep learning* o *DeepL*. La diferencia con el TAN de Google es que no produce redes neuronales recurrentes², sino redes neuronales convolucionales³ (Merkert, 2017). Esta red se solía usar para el reconocimiento de imágenes. La gran ventaja de las redes neuronales convolucionales es que todas las palabras se pueden traducir de forma paralela, además de incluir bibliotecas optimizadas (Merkert, 2017). Esto significa que aprende de sus errores e incluye la solución en sus algoritmos⁴. Por otro lado, *DeepL* ha sido alimentada de todas las bases de datos alineadas de las que *Linguee* (de donde nace esta empresa) ya disponía.

Hoy en día ya existen estudios suficientes que muestran que la calidad de este TAN es aceptable, aunque aún comete bastantes errores. Sus desventajas más significativas son que (1) entrenar estos modelos tan complejos requiere una gran cantidad de datos y parámetros, lo que se traduce en un elevado coste, y que (2) precisa determinar la arquitectura o topología de la red adecuada a cada tarea (Costa-Jussà, 2018). Esto se traduce en que *DeepL* requiere de la revisión de un traductor humano (Aguilar Canal, 2019; Grundvall, 2019).

Con la llegada de esta herramienta, cuya calidad es muy superior a los TA anteriores, nos encontramos, no obstante, con un problema en el grado de traducción e interpretación. Muchos estudiantes confían ciegamente en ella, sin plantearse qué errores puede cometer. Para investigar hasta qué punto los aprendientes (así como los traductores profesionales) pueden hacer uso de este programa en su día a día y con la intención de concienciar a estos de la necesidad de dominar la técnica de la traducción, así como para darles una visión más real de esta herramienta, hemos llevado a cabo un experimento de traducción con este programa.

2. Ejercicio con DeepL en el aula

El ejercicio de *DeepL* propuesto se realizó en clase de traducción inversa ES-DE (español-alemán) de la Universidad Complutense de Madrid, donde participaron 28 estudiantes del último

2 Redes en las que las neuronas se realimentan con sus propias salidas de forma directa o indirecta (Jordan, 1990).

3 Un tipo especial de arquitectura de red neuronal, especialmente eficaz para el reconocimiento de signos naturales como imágenes, imágenes volumétricas, vídeo, habla, música y texto (Le Cun, 2019: 385).

4 Sitio de instrucciones que se deben ejecutar. En la mayoría de los casos, estas instrucciones se ejecutan por un ordenador e incluyen operaciones matemáticas, pruebas, repeticiones, etc. (Le Cun, 2019: 381).

curso. Los textos seleccionados para este experimento son literarios, ya que consideramos que este tipo de texto especializado es el más ligado a la cultura, por lo que también es el más difícil de traducir por DeepL. Nos parece relevante que la traducción del texto propuesto posea cuantiosos errores, para que así los estudiantes puedan reconocerlos y subsanarlos.

2.1. Sujetos

El grupo estaba compuesto por un 74 % de mujeres y un 26 % de hombres de entre 21 y 28 años. Se ha efectuado en un solo curso, ya que DeepL, como hemos explicado en la introducción, aprende de sus errores y evoluciona, por lo que no tiene sentido alargarlo en el tiempo. Los errores que comete este año se pueden ver subsanados en un breve periodo de tiempo y se producirán otros, por lo que los estudiantes de diferentes cursos no trabajarían con las mismas traducciones, lo que haría esta comparación inútil.

Todos los participantes habían cursado al menos cuatro años de alemán, por lo que presentaban un nivel de este idioma de como mínimo un B1 y B2 según el Marco Común Europeo de Referencia para las Lenguas. Además, un 37 % ya había residido en un país germanoparlante (estancias Erasmus, veranos, etc.). El grupo también contaba con una estudiante Erasmus nativa alemana.

2.2. Presentación del ejercicio

La encuesta inicial, cerrada, preguntaba por el uso de DeepL y la respuesta fue unánime: todos lo utilizaban en su día a día y se declaraban buenos "poseedores". Además, también coincidieron en afirmar que el resultado de su traducción con DeepL tendría más calidad que la realizada sin esta herramienta.

Tras esto, se dividió a la clase en dos grupos y se les entregaron dos párrafos del libro *La Colmena*, del premio Nobel de Literatura Camilo José Cela (Cela, 1951). Elegimos una traducción literaria, ya que este tipo de texto tiene una relación muy estrecha con la cultura de la lengua de partida, por lo que es la más propensa a sufrir errores por parte de DeepL (Cases, 2019).

El texto A se extrajo del primer capítulo, primer párrafo, y constaba de 168 palabras:

Doña Rosa va y viene por entre las mesas del Café, tropezando a los clientes con su tremendo trasero. Doña Rosa dice con frecuencia "leñe" y "nos ha merengao". Para doña Rosa, el mundo es su Café, y alrededor de su Café, todo lo demás. Hay quien dice que a doña Rosa le brillan los ojillos cuando viene la primavera y las muchachas empiezan a andar de manga corta. Yo creo que todo eso son habladurías: doña Rosa no hubiera soltado jamás un buen amadeo de plata por nada de este mundo. Ni con primavera ni sin ella. A doña Rosa lo que le gusta es arrastrar sus arrobos, sin más ni más, por entre las mesas. Fuma tabaco de noventa, cuando está a solas, y bebe ojén, buenas copas de ojén, desde que se levanta hasta que se acuesta. Después tose

y sonrío. Cuando está de buenas, se sienta en la cocina, en una banqueta baja, y lee novelas y folletines, cuanto más sangrientos, mejor: todo alimenta.

El texto B se extrajo del primer capítulo, séptimo párrafo, y constaba de 167 palabras:

A don Jaime Arce, que tiene un gran aire a pesar de todo, no hacen más que protestarle letras. En el Café, parece que no, todo se sabe. Don Jaime pidió un crédito a un Banco, se lo dieron y firmó unas letras. Después vino lo que vino. Se metió en un negocio donde lo engañaron, se quedó sin un real, le presentaron las letras al cobro y dijo que no podía pagarlas. Don Jaime Arce es, lo más seguro, un hombre honrado y de mala suerte, de mala pata en esto del dinero. Muy trabajador no es, ésa es la verdad, pero tampoco tuvo nada de suerte. Otros tan vagos o más que él, con un par de golpes afortunados, se hicieron con unos miles de duros, pagaron las letras y andan ahora por ahí fumando buen tabaco y todo el día en taxi. A don Jaime Arce no le pasó esto, le pasó todo lo contrario. Ahora anda buscando un destino, pero no lo encuentra.

Elegimos dos textos no muy extensos, para que así los aprendientes pudieran sentirse cómodos con él y no tuvieran la presión añadida del tiempo, pues se les daba solo 30 minutos para cada uno. Por otro lado, la intención era que ambos textos tuvieran una longitud y un grado de dificultad parecidos.

Tras la entrega de ambos textos, se pidió al primer grupo que tradujera con DeepL y pose-ditara. El segundo grupo tradujo mediante ordenadores, con los correctores ortográficos y gramaticales desconectados, sin la TAN ni buscadores, y únicamente con la ayuda de diccionarios electrónicos monolingües y bilingües. A los 30 minutos se dio por finalizada la actividad y se intercambiaron los grupos. Esta vez el primer grupo traducía con la única ayuda de diccionarios electrónicos, a la vez que el segundo debía pose-ditar con DeepL.

2.3. Extracción de errores producidos por DeepL

En paralelo, nosotros llevamos la misma actividad y extrajimos todos los errores cometidos por DeepL. Para ello, nos basamos en el criterio etiológico basado en la taxonomía Fondecyt 1140651 (Ferreira, 2014) y nos centramos en su categoría interlingual, que está dividida en (1) transferencia, (2) traducción literal, (3) cambio de código, (4) interferencia de la L2, (5) calco y (6) falso cognado. Además, esta subdivide las transferencias (grupo 1) en (1.1) tr. de estructuras gramaticales, (1.2) tr. de estructuras sintácticas, (1.3) tr. léxicas, (1.4) tr. morfológicas, (1.5) tr. gráficas-ortográficas y (1.6) tr. discursivas.

Para conseguir una mejor visualización, como muestra la tabla 1, decidimos dividir los errores asimismo en colores. Así pues, marcamos los errores de transferencia en amarillo, las traducciones literales en verde, los cambios de código en turquesa, la interferencia de la lengua de llegada en lila, el calco en marrón y las palabras cognadas en cerceta. Con esta distinción, el aprendiente podrá reconocer mejor los errores recurrentes, tanto de DeepL como los suyos propios.

TABLA 1

Criterio etiológico seguido (Ferreira, 2014)

TIPOS DE ERRORES	
TRANSFERENCIA	1 - ESTRUCTURA GRAMATICAL
	2 - ESTRUCTURA SINTÁCTICA
	3 - LÉXICA
	4 - MORFOLÓGICA
	5 - GRÁFICA ORTOGRÁFICA
	6 - DISCURSIVA
TRADUCCIÓN LITERAL	
CAMBIO DE CÓDIGO	
INTERFERENCIA L2	
CALCO	
FALSO COGNADO	

Una vez elegidos los criterios de corrección y pasar ambos textos por DeepL, los errores encontrados fueron los siguientes:

2.3.1. Texto A

Doña (3) Rosa kommt und geht zwischen den Tischen des Cafés und stolpert mit ihrem gewaltigen Arsch über die Kunden. Doña (3) Rosa sagt oft "leñe" (3,5) und "nos ha merengao" (3, 5). Für Doña (3) Rosa ist die Welt ihr Café (2), und um ihr Café herum alles andere. Es gibt Leute, die sagen, dass Doña (3) Rosas Augen (3) leuchten, wenn der Frühling kommt und die Mädchen anfangen, kurzärmlig (3) zu gehen. Ich denke, das ist alles Gerede: Doña (3) Rosa hätte nie ein gutes Stück⁵ Silber (3) für irgendetwas auf dieser Welt losgelassen (3). Weder mit Feder noch ohne Feder (3). Doña (3) Rosa zieht gerne ihre Arroben (3), einfach so, zwischen die Tische. Sie raucht Tabak aus den 90er⁶ (3) Jahren, wenn sie allein ist, und trinkt Ojén, gute (3) Gläser Ojén, vom Moment, in dem sie aufsteht, bis sie ins Bett geht. Dann hustet er (2) und lächelt. Wenn es ihr gut geht, sitzt (1) sie in der Küche, auf einem (1) niedrigen Hocker und liest Romane und Folklore (3), je blutiger, desto besser: Alles ernährt sich.

5 Amadeo de plata: moneda de plata de cinco pesetas con el busto del rey Amadeo (RAE y ASALE, 2019).

6 Cajetilla de tabaco de picadura que valía 90 céntimos.

Con los colores, vemos que casi todos los errores cometidos por DeepL son de transferencia. Además, aquellas palabras que no comprende, como “leñe”, o expresiones como “nos ha merengao”, no las traduce y las transcribe en la lengua de partida. Creemos que esto es un error, ya que ningún alemán va a entenderlas, por lo que obstruye la comprensión de la traducción. DeepL debería buscar expresiones similares en el idioma de llegada, como por ejemplo “Menschenskinder” o “das war ja klar!”. Por otro lado, evita adaptar los signos de puntuación, como las comillas, a la lengua de llegada. Además, todas aquellas palabras que tienen una interrelación con hechos culturales, como el amadeo de plata, las traduce de forma literal, obviando su significado real. También cabe resaltar que, como el español es un idioma *pro-drop*, que tiende a omitir los pronombres personales, DeepL opta por incluir el masculino, ya que sabe que en alemán no se debe omitir. Esto es un error habitual que comete en la combinación lingüística ES-DE (español-alemán).

Así pues, y sin entrar en detalle en los errores cometidos por DeepL, ya que eso supondría otro artículo y nos alejaría de nuestro objetivo primario, al realizar la comparación de los estándares de calidad de un texto traducido con y sin DeepL, tras un recuento de todos los errores cometidos, la tabla 2 nos muestra los siguientes resultados:

TABLA 2

Errores de DeepL en el texto A

1 – TR. ESTRUCTURA GRAMATICAL	2
2 – TR. ESTRUCTURA SINTÁCTICA	2
3 – TR. LÉXICA	16
4 – TR. MORFOLÓGICA	0
5 – TR. GRÁFICA ORTOGRÁFICA	2
6 – TR. DISCURSIVA	0
TRADUCCIÓN LITERAL	2
CAMBIO DE CÓDIGO	1
INTERFERENCIA L2	0
CALCO	0
FALSO COGNADO	0
TOTAL	25

Fuente: elaboración propia.

Vemos que los errores más frecuentes de DeepL fueron de transferencia, ya que suman 22 de los 25 errores totales. Además, cabe resaltar que las incorrecciones léxicas, con 16 ejemplos, han sido de lejos los más habituales. Por otro lado, la calidad de la traducción ha sido claramente insuficiente, ya que en 168 palabras se ha equivocado en 25 ocasiones, una cifra que invalida esta traducción⁷.

2.3.2. Texto B

El texto B es el séptimo párrafo del primer capítulo:

An **Don** (3) Jaime Arce, der trotz allem eine großartige **Luft** (3) hat, **tun** (1) sie nichts anderes als **Protestbriefe** (3) an ihn (**schreiben**). Im Café scheint es, **dass nicht**, alles bekannt ist. Don Jaime bat um einen Kredit von einer Bank, sie gaben ihm und er unterschrieb einige **Briefe** (3). Dann kam, was kam. Er **kam** (3) in ein Geschäft, wo sie ihn betrogen haben, er blieb ohne einen **echten** (3), sie überreichten ihm die **Rechnungen** (3) und er sagte, er könne sie nicht bezahlen. **Don** (3) Jaime Arce ist, **ganz sicher** (6), ein ehrlicher Mann mit Pech, mit **einem schlechten Fuß** (3) in **diesem** (3) Geldgeschäft. Er ist kein harter Arbeiter, das ist die Wahrheit, aber er hatte auch kein Glück. Andere **als faul** (1) oder mehr als er, mit ein paar glücklichen **Schlägen** (3), bekamen **ein paar tausend hart**, bezahlten die Rechnungen und **gehen jetzt herum** (6) **und** (6) rauchen guten Tabak und (**fahren**) **den ganzen Tag im Taxi** (1). Das geschah nicht mit **Don** (3) Jaime Arce, das Gegenteil **geschah** (6) mit ihm. Jetzt sucht er nach einem **Ziel** (3), aber er kann es nicht finden.

En esta traducción vemos que se han repetido los mismos parámetros que en el texto A. De nuevo, la mayoría de los errores, a excepción de dos traducciones literales, se han centrado en errores de transferencia. De estos, la mayoría han vuelto a ser problemas léxicos. Además, ha cometido dos omisiones de verbos. Esto arroja los siguientes resultados:

TABLA 3

Errores de DeepL en el texto B

1 – TR. ESTRUCTURA GRAMATICAL	3
2 – TR. ESTRUCTURA SINTÁCTICA	0
3 – TR. LÉXICA	13
4 – TR. MORFOLÓGICA	0
5 – TR. GRÁFICA ORTOGRÁFICA	0

7 Aquí cabe resaltar que en la traducción literaria se debe tener en cuenta no solo la imagen que quiere transmitir el texto original, sino también la belleza de este. Debido al espacio limitado del artículo, en este experimento no vamos a entrar a valorar esto.

6 – TR. DISCURSIVA	4
TRADUCCIÓN LITERAL	2
CAMBIO DE CÓDIGO	0
INTERFERENCIA L2	0
CALCO	0
FALSO COGNADO	0
TOTAL	22

Fuente: elaboración propia.

En este segundo texto, DeepL ha cometido 22 errores en 167 palabras, un porcentaje muy parecido a los cometidos en el primer texto. La calidad de la traducción, de nuevo, es insuficiente. Al igual que en el texto A, también ha llevado a cabo faltas de léxico, lo que demuestra que una de las debilidades de DeepL es la elección de las equivalencias terminológicas.

3. Resultados y discusión

Tras la realización de la actividad detallada en el punto 2 y realizadas ya las correcciones de las traducciones entregadas, obtuvimos los siguientes resultados:

3.1. Resultados del texto A

Este texto resultó quizás más fácil que el texto B, ya que los estudiantes tardaron menos en su traducción. La mayoría de grupos que poseditaban tardaron menos de 15 minutos en finalizar esta actividad, mientras que los grupos que traducían gastaron los 30 minutos. Sin embargo, no hubo ninguno que tuviera que entregar sin haber terminado la traducción.

3.1.1. Resultados de la posesión del texto A

La tabla 4 presenta la media de los errores de los siete grupos que realizaron la posesión:

TABLA 4

Media de los errores de los grupos que poseditaron el texto A

1 – TR. ESTRUCTURA GRAMATICAL	3
2 – TR. ESTRUCTURA SINTÁCTICA	0,8
3 – TR. LÉXICA	12,2
4 – TR. MORFOLÓGICA	0

5 – TR. GRÁFICA ORTOGRÁFICA	2
6 – TR. DISCURSIVA	0,4
TRADUCCIÓN LITERAL	0,8
CAMBIO DE CÓDIGO	0,6
INTERFERENCIA L2	0
CALCO	0
FALSO COGNADO	0
TOTAL	19,8

Fuente: elaboración propia.

La traducción sigue incluyendo errores léxicos, lo que nos enseña que los aprendientes han tomado por supuesto que la elección del programa de TAN era la acertada. De hecho, de media, los estudiantes han descubierto, o en su caso cometido, únicamente 5,2 errores menos que en el texto original de DeepL sin posesición. Eso nos enseña que, aunque los discentes tenían el convencimiento de que dominaban la posesición de DeepL, los datos revelan lo contrario. Cabe resaltar que ningún grupo fue consciente de los errores de puntuación de DeepL, por lo que los eliminó.

3.1.2. Resultados de la traducción del texto A sin ayuda tecnológica

La tabla 5 expone los errores que de media han cometido aquellas parejas que no disponían de TAN y que consumieron los 30 minutos de los que disponían:

TABLA 5

Resultados medios del grupo sin ayuda tecnológica del texto A

1 – TR. ESTRUCTURA GRAMATICAL	17
2 – TR. ESTRUCTURA SINTÁCTICA	1,5
3 – TR. LÉXICA	4,25
4 – TR. MORFOLÓGICA	0
5 – TR. GRÁFICA ORTOGRÁFICA	2,75
6 – TR. DISCURSIVA	0
TRADUCCIÓN LITERAL	0
CAMBIO DE CÓDIGO	0
INTERFERENCIA L2	0

CALCO	0
FALSO COGNADO	0
TOTAL	25,5

Fuente: elaboración propia.

Esta tabla revela que los errores léxicos se han reducido a 4,25, casi 12 errores menos. Esto deja patente que los aprendientes, cuando no tienen ejemplos de DeepL, saben discriminar los términos correctos y los aplican con mayor certeza que la TAN. Además, cometieron menos errores léxicos que los que realizaron los grupos de posesición, ya que ellos tuvieron que informarse y meditar sobre qué término es el más correcto.

Sin embargo, las debilidades de este grupo se manifiestan en los errores gramaticales, donde han introducido 17. Esto se debe a que la lengua alemana es su lengua C, y no están acostumbrados a escribir en este idioma. De hecho, esta asignatura es la única inversa ES-DE que cursan en todo el grado.

3.1.3. Comparación de los errores cometidos en el texto A

TABLA 6

Comparativa de los errores de los dos grupos y de DeepL en el texto A

ERRORES	DEEPL	POSEDICIÓN	SIN SOPORTE TECNOLÓGICO
1 – TR. ESTRUCTURA GRAMATICAL	2	3	17
2 – TR. ESTRUCTURA SINTÁCTICA	2	0,8	1,5
3 – TR. LÉXICA	16	12,2	4,25
4 – TR. MORFOLÓGICA	0	0	0
5 – TR. GRÁFICA ORTOGRÁFICA	2	2	2,75
6 – TR. DISCURSIVA	0	0,4	0
TRADUCCIÓN LITERAL	2	0,8	0
CAMBIO DE CÓDIGO	1	0,6	0
INTERFERENCIA L2	0	0	0
CALCO	0	0	0
FALSO COGNADO	0	0	0
TOTAL	25	19,8	25,5

Fuente: elaboración propia.

3.2. Resultados del texto B

A continuación presentamos las traducciones realizadas por los grupos en relación al texto B. Se volvió a repetir el tiempo consumido en el texto A de forma casi idéntica. Los grupos dedicados a la posesición consumieron entre 15 y 20 minutos, y los grupos sin ayuda tecnológica hicieron uso de los 30 minutos. Uno de estos grupos, además, no fue capaz de acabar la traducción. Debido a esto, la traducción queda invalidada y no se incluye en los resultados.

3.2.1. Resultados de la posesición del texto B

La tabla 7 recoge los resultados del grupo dedicado a la posesición del texto B:

TABLA 7

Media de los errores de los grupos de posesición del texto B

1 – TR. ESTRUCTURA GRAMATICAL	7,5
2 – TR. ESTRUCTURA SINTÁCTICA	0
3 – TR. LÉXICA	8,5
4 – TR. MORFOLÓGICA	0
5 – TR. GRÁFICA ORTOGRÁFICA	0
6 – TR. DISCURSIVA	3
TRADUCCIÓN LITERAL	1,25
CAMBIO DE CÓDIGO	0
INTERFERENCIA L2	0
CALCO	0
FALSO COGNADO	0
TOTAL	20,25

Fuente: elaboración propia.

Vemos que la localización de los errores léxicos no ha sido satisfactoria, ya que todavía presentan 8,5 errores en el texto. Por otro lado, los errores en la estructura gramatical han aumentado, por lo que, al corregir DeepL, han incluido errores que antes no existían. Además, solo han localizado de media un error discursivo y solo una minoría ha descubierto la traducción literal de expresiones.

3.2.2. Resultados de la traducción del texto B sin ayuda tecnológica

Los resultados de los seis grupos que entregaron esta actividad a tiempo se ven recogidos en la tabla 8:

TABLA 8

Resultados medios del grupo sin ayuda tecnológica del texto B

1 – TR. ESTRUCTURA GRAMATICAL	10,3
2 – TR. ESTRUCTURA SINTÁCTICA	1
3 – TR. LÉXICA	6
4 – TR. MORFOLÓGICA	0
5 – TR. GRÁFICA ORTOGRÁFICA	0,5
6 – TR. DISCURSIVA	0,3
TRADUCCIÓN LITERAL	0
CAMBIO DE CÓDIGO	0
INTERFERENCIA L2	0
CALCO	0
FALSO COGNADO	0
TOTAL	18

Fuente: elaboración propia.

Esta tabla nos muestra que los errores más cometidos por los participantes en esta actividad han sido de tipo gramatical, seguidos de términos inapropiados. Sin embargo, no han cometido ni una traducción literal y únicamente dos han cometido un error discursivo. Vemos que todos los errores cometidos por este grupo han sido de transferencia, lo que demuestra que los participantes deben ahondar en el estudio de la lengua alemana, su gramática y su ortografía.

3.2.3. Comparación de los errores cometidos en el texto B

La tabla 9 presenta una comparativa de los errores cometidos por DeepL y por los dos grupos de estudiantes, divididos en 7 grupos:

TABLA 9

Comparativa de los errores de los dos grupos y de DeepL en el texto B

ERRORES	DEEPL	POSEDICIÓN	SIN SOPORTE TECNOLÓGICO
1 – TR. ESTRUCTURA GRAMATICAL	3	7,5	10,3
2 – TR. ESTRUCTURA SINTÁCTICA	0	0	1
3 – TR. LÉXICA	13	8,5	6
4 – TR. MORFOLÓGICA	0	0	0
5 – TR. GRÁFICA ORTOGRÁFICA	0	0	0,5
6 – TR. DISCURSIVA	4	3	0,3
TRADUCCIÓN LITERAL	2	1,25	0
CAMBIO DE CÓDIGO	0	0	0
INTERFERENCIA L2	0	0	0
CALCO	0	0	0
FALSO COGNADO	0	0	0
TOTAL	22	20,25	18

Fuente: elaboración propia.

Vemos una inversión en el tipo de error entre DeepL y las traducciones realizadas sin ayuda tecnológica: DeepL cometió solo 3 errores gramaticales por 10,25 de los estudiantes, que, sin embargo, incluyeron menos de la mitad de faltas léxicas, 6 errores por los 13 del programa TAN. Si incluimos los grupos que posecieron, vemos que los que menos errores léxicos cometen son los que no tienen ningún soporte tecnológico.

Esto pone de manifiesto la necesidad de que los estudiantes no se fíen ciegamente de programas de TA, sino que confíen en sus conocimientos lingüísticos y de lo que los alemanes llaman *Sprachgefühl*, que no es más que la sensibilidad e intuición que los humanos tenemos para con el lenguaje, tanto materno como adquirido.

En lo que los estudiantes sí se ven superados por la máquina es en sus conocimientos gramaticales, ya que los grupos sin soporte tecnológico han realizado 10,3 errores, 7,3 más que la máquina. Esto indica que estos aún tienen un largo camino que recorrer en su estudio de la lengua alemana.

4. Conclusiones

No podemos obviar que los grupos con DeepL han necesitado mucho menos tiempo que aquellos con soporte tecnológico, lo que indica que este avance tecnológico hace más efectivos a los traductores. Los grupos que poseían tardaron casi la mitad que aquellos que traducían sin ayuda.

Si bien es cierto que una de las debilidades de DeepL es la transferencia del léxico, es también una ventaja del sistema, pues es más fácil aportarle léxico que estructuras gramaticales.

Sin embargo, vemos que en ambos grupos se repiten los mismos parámetros: en los grupos del texto A y B, cuando los aprendientes ejercieron de poseedores, no prestaron la atención necesaria a los errores léxicos, y en ambos, cuando los aprendientes ejercieron de traductores sin soporte tecnológico, cometieron más errores gramaticales y menos léxicos. Esto indica que, a la hora de poseer, los aprendientes tienen un enfoque erróneo. Procuran cambiar estructuras, en vez de fijarse en la terminología aplicada.

Por otro lado, los estudiantes sin soporte tecnológico cometieron la misma cantidad de errores o incluso menos, como es el caso del texto B, que la máquina. Esto enseña a los aprendientes a que ellos pueden, por sí solos, imprimir una calidad a sus traducciones que el programa TAN aún no puede (al menos en textos literarios y periodísticos).

Si nos fijamos en el enfoque didáctico, todos los aprendientes por unanimidad han encontrado que esta actividad les ha ayudado mucho a entender mejor los errores que deben buscar a la hora de corregir una traducción realizada por una herramienta TAN y que se deben fiar más de su instinto de traductor. Los 28 participantes han admitido por unanimidad que este ejercicio les ha “abierto los ojos” y que de ahora en adelante su uso del programa DeepL no será tan indiscriminado, sino con la atención puesta más en el léxico utilizado por esta herramienta. Igualmente, todos los aprendientes han encontrado esta actividad muy útil y de gran ayuda para ejercer su profesión de forma más consciente.

Como docentes, y tal como defienden Burbat y Varela Salinas (2018), tenemos la obligación de romper viejos tabús relacionados con el traductor automático neuronal y enseñar un uso adecuado y consciente de este.

Para concluir no nos queda sino apoyar parcialmente la declaración de John R. Pierce, incluyendo, eso sí, la tesis de Wu y su equipo: La máquina no podrá superar a corto y medio plazo la competencia lingüística y traductora humana, aunque sí le podrá ayudar a ser más efectivo. La labor de revisar lo producido por un TAN se deberá realizar siempre por un humano, si queremos imprimir calidad a cualquier traducción. Siempre habrá expresiones, recursos lingüísticos y juegos de palabras que solo un humano podrá descifrar y reformular en la lengua de llegada.

5. Bibliografía

AGUILAR CANAL, Enrique, 2019: *Evaluación de la calidad del traductor automático DeepL del inglés al español en textos especializados*, Repositorio de la Universidad César Vallejo [<http://repositorio.ucv.edu.pe/handle/20.500.12692/39871>, fecha de consulta: 26 de julio de 2020].

ALPAC, 1966: *Languages and machines: computers in translation and linguistics*, Originary: National Academy of Sciences, National Research Council. Washington, D.C.: National Academy of Sciences [<http://www.hutchinsweb.me.uk/ALPAC-1996.pdf>, fecha de consulta: 15 de octubre de 2020].

BENGIO, Yoshua, Réjean DUCHARME, Pascal VINCENT y Christian JANVIN, 2003: "A Neural Probabilistic Language Model", *Journal of Machine Learning Research* 3 [<http://www.jmlr.org/papers/volume3/bengio03a/bengio03a.pdf>, fecha de consulta: 15 de octubre de 2020].

BROWN, Peter, Stephen DELLA PIETRA, Vincent DELLA PIETRA y Robert MERCER, 1993: "Las matemáticas de la traducción automática estadística: estimación de parámetros", *Lingüística Computacional*, 263-311.

BURBAT, Ruth, y Marza José VARELA SALINAS, 2018: *First International Conference on Research in Multilingualism: Innovation and New Challenges*, 14-16 de marzo, Universidad de Oviedo [Disponible en https://riuma.uma.es/xmlui/bitstream/handle/10630/16310/Google%20Translate%20and%20DeepL_%20Breaking%20Taboos%20in%20Translator%20Training_presentation.pdf?sequence=3&isAllowed=y, fecha de consulta: 26 de julio de 2020].

CASES, Elke, 2019: "Die Zukunft der Übersetzung mit der neuronalen maschinellen Übersetzung", *Wer schafft, der schafft's. X.FAGE-Kongress*, 7 al 9 de octubre, Logroño: FAGE.

CELA, Camilo José, 1951: *La colmena*, Madrid: Emecé Editores.

COSTA-JUSSÀ, Marta, 2018: "From Future to Paradigm: Deep Learning in Machine Translation", *Journal of Artificial Intelligence Research* 61, 947-974.

EBERLE, Kurt, 2008: Integration von regel- und statistikbasierten Methoden in der Maschinellen Übersetzung. *Journal for Language Technology and Computational Linguistics*, 37-70.

FERREIRA, Anita, 2014: Proyecto de investigación FONDECYT - CONICYT no.1140651: "El feedback correctivo escrito directo e indirecto en la adquisición y aprendizaje del español como lengua extranjera". Universidad de Chile 2014-2017.

GRUNDVALL, Jessica, 2019: *Was heißt „gut genug“ bei maschinell übersetzten Texten? Eine Studie der Qualität an zwei von DeepL übersetzten Textbeispielen*, Universität Vaasa [<https://>

osuva.uwasa.fi/bitstream/handle/10024/9663/osuva_8805.pdf?sequence=1&isAllowed=y, fecha de consulta: 15 de octubre de 2020].

JORDAN, Michael, 1990: "Attractor Dynamics and Parallelism in a Connectionist Sequential", *Artificial neural networks. IEEE Press*, 112-127.

LE CUN, Yann, 2019: *Quand la machine apprend. La révolution des neurones artificiels et de l'apprentissage profon*, París: Odile Jacob.

MERKERT, Johannes, 2017: *Maschinelle Übersetzer: DeepL macht Google Translate Konkurrenz*, 29 de agosto [<https://feedlywatch.blogspot.com/2017/08/maschinelle-ubersetzer-deepL-macht.html>, fecha de consulta: 15 de octubre de 2020].

REAL ACADEMIA ESPAÑOLA [RAE] y ASOCIACIÓN DE ACADEMIAS DE LA LENGUA ESPAÑOLA [ASALE], 2019: *Diccionario de la Lengua Española [DLE]*.

SCHMALZ, Antonia, 2019: *Künstliche Intelligenz, iit-Themenband* [https://link.springer.com/content/pdf/10.1007%2F978-3-662-58042-4_12.pdf, fecha de consulta: 15 de octubre de 2020].

WEAVER, Warren, 1949: *Traducción automática de las lenguas*, Cambridge: MIT Press.

WERTHMANN, Antonina, y Andreas WITT, 2014: "Maschinelle Übersetzung - Gegenwart und Perspektiven" en Gerhard STICKEL (ed.): *Translation and Interpretation in Europe. Contributions to the Annual Conference 2013 of EFNIL in Vilnius*, Fránkfort del Meno/Berlín/Berna/Bruselas/Nueva York/Oxford/Viena: Lang, 79-103.

WU, Yonghui, Mike SCHUSTER, Zhifeng CHEN, Quoc V. LE, Mohammad NOROUZI, Wolfgang MACHEREY, Maxim KRİKUN, Yuan CAO, Qin GAO, Klaus MACHEREY, Jeff KLINGNER, Apurva SHAH, Melvin JOHNSON, Xiaobing Liu, Łukasz KAISER, Stephan GOUWS, Yoshikiyo KATO, Taku KUDO, Hideto KAZAWA, Keith STEVENS, George KURIAN, Nishant PATIL, Wei WANG, Cliff YOUNG, Jason SMITH, Jason RIESA, Alex RUDNICK, Oriol VINYALS, Greg CORRADO, Macduff HUGHES y Jeffrey DEAN, 2016: "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation", 26 de septiembre, Cornell University [<http://arxiv.org/abs/1609.08144>, fecha de consulta: 15 de octubre de 2020].