



TRABAJO DE FIN DE GRADO

FACULTAD DE CIENCIAS HUMANAS Y SOCIALES

GRADO EN TRADUCCIÓN E INTERPRETACIÓN

**ESTUDIO COMPARATIVO DE TRES
TRADUCTORES AUTOMÁTICOS EN LÍNEA:**

DEEPL, YANDEX Y APERTIUM

Autora: Mónica Adán Soriano

Directora: Profesora M^a Luisa Romana García

Madrid, junio 2019

Resumen: Este trabajo tiene la finalidad de comparar traductores automáticos en línea para así determinar cuál es el traductor más avanzado para un texto técnico. Para ello, primero, habrá un análisis de la evolución histórica de la traducción automática y sus usos, así como los programas desarrollados para ello. Se explicarán además los diferentes sistemas de traducción automática que existen y cómo funcionan. En la parte experimental, se escogerá un texto de carácter técnico en español y de oraciones que supongan un reto para un traductor y se procesará en los tres traductores automáticos escogidos según su modalidad para realizar una traducción al inglés. Una vez obtenidas las respuestas, se analizarán los errores cometidos por los traductores automáticos, concluyendo así con los errores más comunes y el mejor traductor automático en línea así como los usos que se le pueden dar.

Palabras clave: traducción automática, sistemas basados en estadística, Yandex, sistemas neuronales, DeepL, sistemas basados en reglas, Apertium, errores de traducción.

Abstract: The aim of this dissertation is to do a comparative research analysis on three automatic translation programs available on the internet to determine which is the most advanced for a technical text. First, we will do an analysis on the historic evolution of automatic translation and the several uses given to it, as well as the diverse programs developed for it and how they work. For the empirical analysis, a technical text in Spanish will be processed through three different translation programs using different systems to translate this text into English. Once the results have been obtained and gathered, they will be analyzed to state the most common mistakes for each system and the best out of the three, as well as the uses that could be given to these programs.

Key words: automatic translation, statistical-based systems, Yandex, neuronal systems, DeepL, rule-based systems, Apertium, translation mistakes.

ÍNDICE

1.	Introducción.....	3
1.1.	Objeto.....	3
1.2.	Objetivo	3
2.	Metodología.....	5
3.	Estado de la cuestión	6
3.1.	Historia de la traducción automática	6
3.2.	Aplicaciones de la traducción automática	12
4.	Hipótesis y pregunta de investigación	15
5.	Marco teórico.....	16
5.1.	Tipos de Traducción Automática.....	16
5.2.	Parámetros de evaluación de la Traducción Automática.....	20
6.	Análisis	22
6.1.	Traductores automáticos	22
6.2.	Texto elegido	25
6.3.	Análisis del texto	25
6.4.	Revisión post-traducción	30
7.	Conclusión.....	32
8.	Bibliografía.....	35

1. INTRODUCCIÓN

1.1. OBJETO

La traducción es una de las profesiones más antiguas en este mundo. Desde hace miles de años la traducción ha ayudado a la evolución de la humanidad al colaborar con la comunicación entre diferentes culturas y sociedades. Poco a poco y gracias a la difusión del conocimiento y los grandes avances tecnológicos y del conocimiento humano, el trabajo de la traducción es cada vez más asequible y nos aseguramos un trabajo más fiel. Hoy en día somos incapaces de imaginarnos una traducción realizada sin recursos tecnológicos o incluso sin recursos de internet. El internet resulta un pozo sin fondo de conocimientos, que, a pesar de no ser fiable, si se tiene un buen criterio al usarlo, puede ser un gran aliado a la hora de realizar una traducción. Todos los diccionarios y glosarios imaginables se pueden encontrar en internet, y siempre estarán actualizados, a diferencia del tradicional diccionario de papel, que pasados los años se acaba quedando antiguo, ya que no incluye las actualizaciones o las nuevas incorporaciones a la lengua. Y es por este desarrollo del campo de la tecnología aplicada a la lingüística que surgieron los traductores automáticos, aplicaciones en línea que los más optimistas consideran que pueden llegar a sustituir ese oficio, que existe desde hace milenios.

El objeto de este Trabajo de Fin de Grado es hacer un estudio comparativo del trabajo realizado por tres traductores automáticos que se puedan usar de manera gratuita en internet. Cada vez es mayor el número de personas que recurren a este tipo de herramientas con todo tipo de finalidades, ya sea entender un texto o publicar un artículo en otro idioma. En la mayoría de los casos estos traductores en línea se complementan con una revisión por una tercera persona que conozca el idioma, aunque no sea un traductor profesional. Sin embargo, como estudiante de traducción e interpretación en mi último curso, considero que, a pesar de los avances existentes, aún es necesaria la intervención de un traductor profesional al hacer uso de estas aplicaciones.

1.2. OBJETIVO

Con este trabajo se pretende, en primer lugar, exponer una pequeña introducción sobre la evolución de la traducción hasta llegar a la traducción automática en línea y el desarrollo que ha supuesto. Tras esto, se analizará el funcionamiento de tres de estos traductores que dispongan un sistema basado en reglas lingüísticas, basado en estadística

y, por último, un sistema neuronal, para determinar así cuál de estos es el que menos errores comete. También se verá cuáles son los errores más recurrentes de cada sistema y se analizará el tipo de error.

Y, así, el objetivo es demostrar que incluso los traductores automáticos con un programa más avanzado aún no están preparados para sustituir el trabajo de un traductor profesional, a pesar de los grandes progresos que se han conseguido en este campo. Y que, aunque su uso puede ser de gran ayuda para los profesionales, no se deberían usar con estos fines por personas que no tengan un conocimiento amplio sobre la lingüística de un idioma. Su uso para una mejor comprensión de un texto está completamente justificado y puede ofrecer soluciones muy fiables, pero también puede llevar al engaño en ciertas circunstancias en las que la composición de una frase sea más complicada.

2. METODOLOGÍA

Para la realización de este Trabajo de Fin de Grado se han usado sobre todo fuentes secundarias de estudios previamente realizados por diferentes investigadores y docentes, tanto lingüísticas como historiográficas. Esta información obtenida de fuentes secundarias se ha recopilado, resumido y explicado de manera congruente con los objetivos de este análisis. Se ha desarrollado así un estudio comparativo partiendo de la hipótesis explicada en el apartado anterior para evaluar los programas de TA Traductor de Google, DeepL y Apertium. Se han elegido estos tres programas por su accesibilidad gratuita en internet y porque cada uno dispone de un sistema diferente.

En primer lugar se presenta un breve resumen sobre la evolución de los sistemas de TA que se han ido desarrollando, y se explican sus principales aplicaciones. En segundo lugar se exponen la hipótesis y las preguntas de investigación del presente trabajo, y a continuación se aclara el marco teórico con los diferentes sistemas y se elige un sistema de evaluación para el posterior resultado de los TA. A continuación, se realiza el análisis de un texto con cierto grado técnico procesado por los tres sistemas mencionados y se evalúan los errores obtenidos en los resultados. Por último, se exponen las conclusiones sacadas del conjunto del estudio.

3. ESTADO DE LA CUESTIÓN

3.1. HISTORIA DE LA TRADUCCIÓN AUTOMÁTICA

Hasta principios del siglo XX solo se había especulado con la posibilidad de crear diccionarios numéricos para una posible traducción automatizada a base de códigos. Fue en el año 1933 cuando, por separado, dos científicos desarrollaron ideas similares. El primero, el francés George Astrouni, diseñó un método para almacenar diccionarios y así encontrar la palabra equivalente en otro idioma, lo que hoy en día se conocería como un diccionario bilingüe mecanizado (Hutchins W. J., 2001). Y, por otro lado, el ruso Petr Smirnov-Troyanskii; su idea pareció tener más relevancia. Consistía en tres fases de las que solo una estaba automatizada: la primera fase consistía en un análisis del TO por un lingüista que lo pasaba a las formas y funciones sintácticas más simples, para que así una máquina lo transformara en estas formas y funciones sintácticas de la LM, convirtiéndolo así en el TM, que necesitaba entonces la tercera fase, en la que un lingüista de LM editara el texto para introducir las estructuras normalizadas de la LM.

Sin embargo, no existen pruebas claras de que estas ideas se llevaran a la práctica. Los primeros que materializaron una intención de traducción a través del ordenador fueron Andrew D. Booth, un criptógrafo de Inglaterra, y Warren Weaver, de la Fundación Rockefeller en el año 1947 y más específicamente aún en el año 1949, en el que Weaver escribió un memorándum sobre las ideas y desarrollos establecidos en este tema. En el memorándum, Weaver especifica que lo redacta con la esperanza de que alguien con mejores técnicas, conocimientos e imaginación lo utilizara como un estímulo, con comentarios y sugerencias acerca de la posibilidad de solucionar el problema mundial de la traducción con el uso de ordenadores de gran capacidad, flexibilidad y velocidad (Weaver, W., 1949). La mayor preocupación o problema que existía en la época era la comunicación entre países. Al finalizar la Segunda Guerra Mundial, se desarrolla a grandes pasos la comunicación internacional tanto para organizaciones internacionales como fue la Comunidad Europea como para fines comerciales y científicos que tenían el objetivo de compartir información y desarrollo.

Aunque ya habían comenzado algunas investigaciones al respecto, fue el memorándum de W. Weaver el que espoleó en los años 1950 la creación de grupos de

investigación en Europa y Estados Unidos con suficiente financiación¹ en el ámbito de la traducción automática con ordenador. Los primeros sistemas se basaban mucho en la criptografía desarrollada durante la Segunda Guerra Mundial, como explica W. Weaver en su memorandum, es decir, pasar un lenguaje a un código para después tener que descifrar ese código; se pretendía que ese código expresara las ideas más básicas comunes a todos los idiomas (Weaver, W., 1949).

Se produjo una gran decepción debido a las dudas y los pocos éxitos que habían surgido de la gran inversión inicial. Esta decepción se puede achacar a que el objetivo de estas investigaciones era, en su base, diferente a los actuales; en esa época se estaba buscando una traducción totalmente automática de gran calidad (*fully automatic high-quality translation, FAHQT*) que no se diferenciara de las producidas por los traductores profesionales. El personaje más notorio que criticó la investigación dedicada a esta finalidad fue el filósofo Bar-Hillel, en un informe del año 1959 sobre la evolución de la mecanización de la traducción en EE. UU. y Gran Bretaña. Como conclusión de su informe, Bar-Hillel propone que la investigación de la TA se destine a buscar, ya sea una ayuda al traductor profesional o una traducción completamente automática, pero de mala calidad, ya que consideró imposible llegar a alcanzar la FAHQT.

A causa de las dudas que había suscitado el informe de Bar-Hillel, la Academia Nacional de Ciencias de EE. UU. estableció en el año 1964 un Comité asesor del procesamiento automático del lenguaje (Automatic Language Processing Advisory Committee, ALPAC) y dos años después publicaron el informe acerca de la situación de la traducción automática en cuanto a costes, calidad y prospecciones de futuro (ALPAC, 1966). Este informe ejerció una gran influencia en la investigación de la TA, ya que concluyó que no había falta de traductores, que con los traductores profesionales de la época eran suficientes y que no había ninguna perspectiva en un futuro cercano de que la TA ofreciera una traducción útil de textos. Para el informe ALPAC se realizaron investigaciones empíricas que comparaban las traducciones realizadas por humanos y las traducciones de TA y pos-editadas para su normalización, y concluyeron que se necesitaba más tiempo, más dinero y más recursos, y que la calidad no era tan buena. Es

¹ En EE. UU. se estima que en el año 1958 cerca de un millón y medio de dólares se invirtieron en el desarrollo de la traducción automática con 100 personas dedicadas a ello, la mayoría licenciados en humanidades y ciencias. (Hillel, 1959)

por esto que el informe ALPAC concluyó con recomendaciones para invertir los recursos destinados a la TA en recursos para la ayuda de traductores, como podían ser diccionarios automáticos o la lingüística computacional (ALPAC, 1966).

Durante estos años, la investigación de la traducción automática se había centrado en la traducción de materiales técnicos y científicos de ruso al inglés para EE. UU. a causa de la Guerra Fría, en la que ambos países aspiraban al mayor desarrollo tecnológico y científico posible, por lo que necesitaban saber qué nuevos descubrimientos realizaba el país oponente. Por otro lado, para Canadá tenían más relevancia las traducciones entre francés e inglés (sobre todo EN>FR) debido a la cultura bilingüe del país, que hacía necesaria la traducción de textos de todo tipo: políticos, científicos, literarios, comerciales..., que superaba la oferta del mercado. Y, por último, en Europa crecía la demanda de traducciones a pasos agigantados por la necesidad de traducir todos los textos científicos, técnicos, administrativos y legales para aquellos países que formaban parte de la Comunidad Europea (Hutchins W. J., 2001, págs. 6-7).

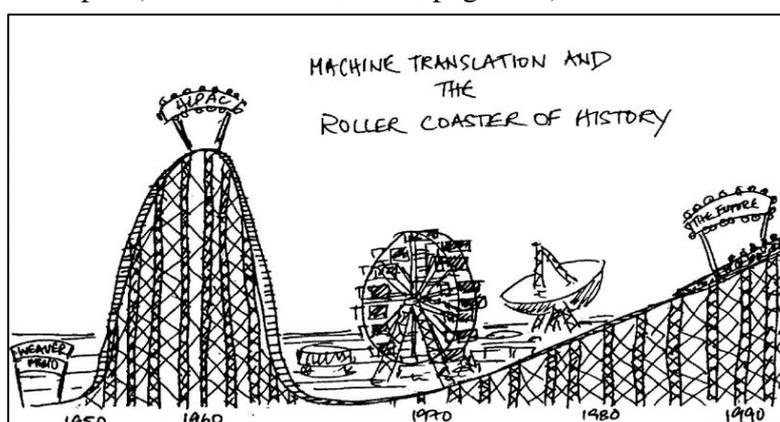


Ilustración 1: Viñeta sobre la historia de la TA. (Arnold, Balkan, Meijer, Humphreys, & Sadler, 1994)

Debido a la conclusión del informe del ALPAC, la inversión en el campo de la TA disminuyó considerablemente en EE. UU.; la ilustración simboliza esta reducción del interés en EE. UU. por la TA tras el informe del ALPAC. Sin embargo, en Europa y Canadá no tuvo tanta repercusión y se continuó con la investigación.

En EE. UU., Peter Toma desarrolló en los años 1960 el sistema Systran (*system translation*) en la Universidad de Georgetown que más adelante, en 1970, se utilizó en las fuerzas aéreas de EE. UU. y en el proyecto Apollo-Soyuz de la NASA (Hutchins & Somers, 1992, pág. 176)

En Canadá destacó entre muchos otros por su repercusión el grupo de investigación de traducción automática de la Universidad de Montreal (*Traduction Automatique à l'Université de Montréal, TAUM*) que desarrolló el programa Météo para traducir boletines informativos sobre el tiempo meteorológico. Este programa se instaló en el año 1976 y sigue funcionando hoy en día, con numerosas actualizaciones. Météo está únicamente destinado al par de idiomas EN-FR, se trata de un sistema de traducción directa que primero realiza la transferencia léxica y después la sintáctica. Su éxito se centra en el hecho que es un vocabulario de un subcampo muy limitado, la meteorología, y por lo tanto los dobles significados son muy escasos y la estructura de las frases es siempre similar. El éxito de Météo actuó a su vez como publicidad para la TA y su desarrollo, demostrando que en una aplicación precisa tenía resultados satisfactorios y de bajo coste económico (Hutchins & Somers, 1992, págs. 207-220).

En Europa, en esta época destacó el grupo de estudios de la traducción automática (*Groupe d'Etudes de la Traduction Automatique, GETA*) de la Universidad de Grenoble, en Francia. Este grupo inició su trabajo con el desarrollo de un sistema interlingua, pero tras años sin éxito, en 1971 diseñaron un sistema de transferencia denominado Ariane, centrado en las traducciones del alemán, japonés y ruso al francés, (Hutchins & Somers, 1992, págs. 221-238), en el que más tarde se basó el sistema Eurotra de la entonces Comunidad Europea, que antes había intentado adaptar el programa Systran comprado a los EE. UU. a las grandes necesidades de traducción inicialmente de inglés y francés, que más tarde se amplió a otros idiomas de la CE. Eurotra fue el proyecto de TA más ambicioso de la época, basado en la interlingua hasta 1981, año en que la CE decidió dejar de financiar el proyecto. Sin embargo, las investigaciones y avances de este periodo sentaron las bases para la actual traducción asistida por ordenador (Aranda, 2012, págs. 36, 37).

En los años 1980 se desarrollaron numerosos programas que se basaban en un sistema que no se había visto antes, las memorias de traducción; en esta década se generaron programas como Transit (1984) y Trados (1984), al que más tarde se integró SDLX (1998) creando SDL Trados. Estas herramientas se basan en un almacenamiento de segmentos de frases en diccionarios bilingües que se pueden reutilizar más adelante cuando vuelvan a surgir segmentos similares. Además de este nuevo sistema de memorias de traducción, también cabe destacar el surgimiento de Internet.

Ya en el año 1988 se creó el primer programa de traducción en internet. El pionero de este servicio fue el centro de Systran en París, que ofrecía estos servicios a los suscriptores de Minitel²: los usuarios podían enviar un texto y recibir una traducción automática a un coste aproximado de 0,01€ la palabra, y tardaba un minuto en traducir 22 líneas (Gaspari & Hutchins, 2007). En septiembre de 1992 CompuServe, una antigua compañía de comunicación estadounidense, ofreció a sus miembros la posibilidad de traducción online en diferentes fórums; Mary Flanagan, una lingüista computacional, se encargó de evaluar los resultados de este servicio. El programa traducía a una velocidad de 3.000 palabras por minuto y ofrecía una traducción de baja calidad, pero el propio CompuServe también ofrecía los servicios de una posesición de un traductor profesional. Según Flanagan, CompuServe aprovechó la oportunidad de mercado en un entorno en el que una rápida visión del contenido principal era más importante que una alta calidad de texto (Gaspari & Hutchins, 2007). También se observó que únicamente el 15 % de los usuarios de este servicio requerían la posesición, mientras que al restante 85 % les bastaba con la traducción de baja calidad realizada por el programa (Flanagan, 1996).

Pocos años después cambió radicalmente el panorama de la traducción online: el motor de búsqueda Altavista³ y la empresa Systran Software Inc. presentaron el programa de traducción online gratuito Babelfish,⁴ que integraba diez idiomas. Un informe realizado años después de su lanzamiento por los lingüistas Hutchins y Gaspari, mostró que los principales usuarios de este programa estaban más interesados en las combinaciones EN>ES, EN>FR, DE>EN, ES>EN y EN>DE (Gaspari & Hutchins, 2007). Este informe también demostró que más de la mitad de los textos que se introducían en el sistema eran de menos de 5 palabras, y que solo un 25 % era de más de 20 (Yang & Lange, 2003). Desde entonces han surgido numerosos programas de traducción en línea gratuitos. Los lingüistas Hutchins y Gaspari además realizaron una evaluación para este mismo informe mencionado anteriormente sobre los datos que les ofrecieron tres programas de traducción online: Yahoo! Babel Fish, FreeTranslation y Systran, en la cual llegaron a la conclusión de que es más común la traducción de textos

² Servicio accesible a través del teléfono predecesor del World Wide Web (www), donde los franceses podían realizar transacciones similares a las que se hacen hoy en día con internet: comprar billetes, chatear o buscar productos.

³ Creado en 1998, en el año 2003 la empresa Overture Service Inc. compró este motor, siendo a su vez comprada por Yahoo! en el año 2004.

⁴ Yahoo! Babelfish tras la compra de Yahoo! en el 2004, que pasó a ser el traductor de Bing en 2012, con Microsoft como propietario.

sencillos y más bien cortos, de una palabra o una frase, lo que puede indicar que los usuarios son personas que conocen el idioma al que se realiza la traducción, pero necesitan una ayuda menor para la plena comprensión. Los idiomas más usados eran inglés, español y francés. El hecho de que no haya una gran cantidad de textos completos y extensos que se traduzcan en un TA online demuestra que los usuarios conocen las limitaciones de estos programas y no se fían de la calidad de su traducción al completo (Gaspari & Hutchins, 2007).

El lingüista Hutchins es el encargado de realizar un compendio de programas de traducción que existen incluyendo sistemas comerciales de TA online y herramientas de ayuda a la traducción para la Asociación Europea para la Traducción Automática (*European Association for Machine Translation, EAMT*) solo accesible para los miembros de esta. Este tipo de programa de traducción ha evolucionado a pasos agigantados. No es pública la información acerca de los algoritmos que utilizan, pero, por ejemplo, el traductor de Google puede utilizar todos los recursos de su motor de búsqueda para recopilar toda la información bilingüe de internet; además también usa recursos semánticos y terminológicos que sean accesibles en la web. Y, finalmente, una de las mayores desventajas que suponen estos programas es la imposibilidad de confidencialidad de los textos traducidos, es decir, cualquier texto que estos programas procesen será almacenado en sus sistemas.

Sin embargo, la posesición de traducciones realizadas automáticamente en la actualidad recibe una gran importancia. Se están organizando conferencias y fórums acerca de este asunto para discutir su importancia económica y científica. Por un lado, la posesición de la TA puede agilizar el proceso de traducción permitiendo que un traductor profesional sea capaz de traducir una cantidad mayor de texto en un tiempo más reducido. Y, por el otro, puede mejorar los propios sistemas de TA siendo capaces de reutilizar las ediciones que se han realizado para evolucionar y ofrecer traducciones más correctas.

La TA también está desarrollándose en nuevos campos con otras aplicaciones, como es la interpretación automática, es decir, desarrollar la TA en el reconocimiento automático del habla (*automatic speech recognition, ASR*) y la TA. En el año 2014 una empresa coreana, CSLi, compró Systran y ahora está desarrollando un sistema de análisis de la voz y traducción que usan los aparatos de Samsung como móviles o *tablets*. En los últimos años han surgido numerosas aplicaciones para traducir la voz, sobre todo para teléfonos móviles. Estos programas o aplicaciones pretenden posibilitar la comunicación

de una persona en su idioma nativo con un interlocutor que hable en otra lengua mediante la TA del habla en directo (Poibeau, 2017, págs. 237-245). Diversas empresas de comunicación ven este nuevo desarrollo por la TA del habla como una gran oportunidad, y muchas de ellas están invirtiendo en ello como, por ejemplo, Skype, WeChat, Facebook o Google.

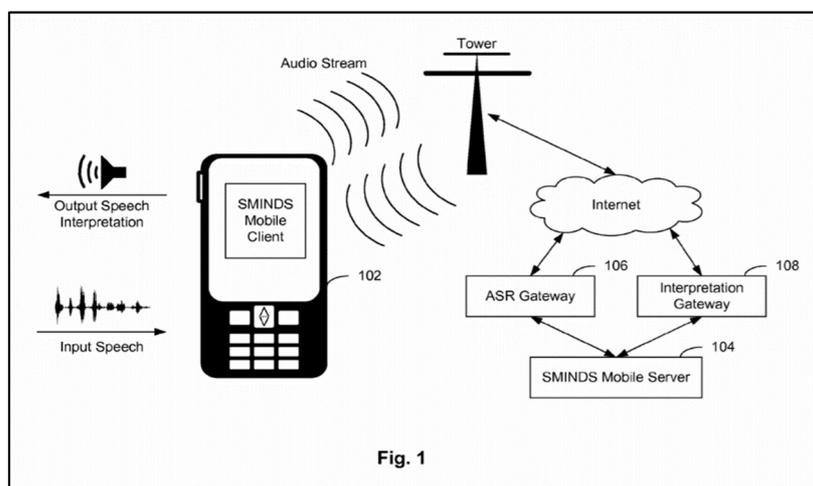


Figura 1: Patente de un sistema de interpretación automática.
(Estados Unidos Patente nº US 9.430,467 B2, 2016)

En la ilustración 2 podemos ver una patente de Google de un servicio telefónico denominado SMINDS, que muestra cómo usando un aparato de comunicación como puede ser el teléfono móvil es posible reconocer, enviar y recibir información de voz a través de un servidor web de interpretación. En la patente se especifica que con este invento se pretende ofrecer una interpretación que resulta de un proceso de reconocimiento de voz (ASR), una generación de una traducción a través de un sistema de interlingua y devolver la traducción al usuario en el formato que se especifique, ya sea hablado o escrito (Estados Unidos Patente nº US 9.430,467 B2, 2016, págs. 1-2).

3.2. APLICACIONES DE LA TRADUCCIÓN AUTOMÁTICA

Con la gran variedad de programas de traducción automática que existen en la actualidad, cualquiera que disponga de un ordenador y acceso a internet puede utilizarlos. La traducción automática es empleada por personas de diferentes edades, profesiones, intereses; empresas e incluso instituciones internacionales.

Según el departamento de Lenguajes y Sistemas Informáticos de la Universidad de Alicante, la traducción informática tiene tres usos principales en los que se pueden reunir todas las aplicaciones de la TA. En primer lugar, se encuentra la asimilación, en este caso, la traducción que se obtiene a través de TA sirve para que el usuario pueda

tener una idea general de la temática del texto, la calidad del TM no es relevante, ya que el usuario no pretende obtener la información exacta del texto. En segundo lugar, tenemos la comunicación. Como su propio nombre indica, este uso de la TA tiene la finalidad de facilitar la comunicación entre las personas. Este tipo de TA se suele usar en chats electrónicos como Facebook o el correo electrónico. En este caso, dependiendo de la conversación puede o no necesitarse una buena TA. Y, por último, tenemos el uso de la TA para la divulgación, es decir, hacer público un texto para el interés general. En esta situación es muy importante la calidad del texto, ya que no es de uso individual sino general.

En la actualidad se han desarrollado programas para turistas en el que, ya sea hablado o escrito, el usuario del programa inserta un texto en el programa y este lo traduce al idioma del país que esté visitando, para así poder comunicarse con los locales. Un ejemplo de este programa es el desarrollado por la empresa japonesa Logbar⁵, denominado Ili, que se puede usar sin internet. En estas situaciones, la TA puede ser de gran utilidad ya que las necesidades de traducción no son muy específicas y el vocabulario suele ser sencillo y sin grandes tecnicismos, además de que se puede acotar el vocabulario para traducir, evitando así dobles significados. (Martín Juarrós, 2017)

Sin embargo, la traducción para asuntos de contratos, negociaciones, leyes o asuntos de mayor importancia puede llegar a generar grandes problemas. En estos casos, sería éticamente correcta su utilización por parte de un traductor profesional para consultar alguna duda o posible redacción de una traducción o para una posesición. Este último se está comenzado a usar con más asiduidad para agilizar el proceso de traducción y así invertir menos tiempo y dinero (Martín Juarrós, 2017). No obstante, es necesario ser conscientes de que, si la TA no es de buena calidad y contiene muchos errores, es posible que el traductor tarde más tiempo en realizar la posesición que en realizar todo el trabajo desde cero, y aun así no tendría la misma calidad.

Una empresa que esté realmente interesada en comunicar correctamente sus documentos utilizará siempre un traductor profesional; sin embargo, hoy en día, una gran cantidad de empresas recurren a la TA en internet, con textos que contienen errores, sin darle mayor importancia. Aquí es donde entra la integridad de cada empresa y su interés

⁵ <https://logbar.jp/en/index.html>

por ofrecer un servicio profesional y adecuado a sus clientes con un texto escrito de manera correcta.

The problem is that the demand for translation in the modern world far outstrips any possible supply. Part of the problem is that there are too few human translators, and that there is a limit on how far their productivity can be increased without automation. In short, it seems as though automation of translation is a social and political necessity for modern societies which do not wish to impose a common language on their members (Arnold, Balkan, Meijer, Humphreys, & Sadler, 1994)

4. HIPÓTESIS Y PREGUNTA DE INVESTIGACIÓN

Este Trabajo de Fin de Grado parte de la hipótesis de que los traductores automáticos en línea pueden servir de gran ayuda a un traductor profesional, a la vez que no son fiables y, por el momento, no pueden sustituir el oficio de un traductor profesional. Ya que en este trabajo se van a estudiar tres tipos diferentes de traductores automáticos en línea para poder así observar la evolución de los diferentes sistemas, estas preguntas se aplicarán a la traducción que resulte la más fiel y correcta, es decir, se aplicarán al sistema de traducción automática más avanzado. Para desarrollar esta hipótesis se formulan las siguientes preguntas de investigación:

- ¿Cuáles son los errores más recurrentes en los diferentes traductores automáticos?
- ¿Ofrece algún traductor automático un trabajo final sin errores?
- ¿Cómo sería el trabajo de una revisión realizada después de un traductor automático?
- ¿Cuál es el futuro de los traductores automáticos en línea?

5. MARCO TEÓRICO

5.1. TIPOS DE TRADUCCIÓN AUTOMÁTICA

A lo largo del desarrollo de los diferentes programas de traducción automática han surgido diferentes métodos de traducción. A continuación, voy a ofrecer un pequeño resumen con una breve explicación de los sistemas que han existido hasta el día de hoy.

Para empezar, es necesario diferenciar si es un sistema multilingüe o bilingüe. Los bilingües, como su propio nombre indica, son aquellos que solo realizan traducciones entre dos idiomas; dentro de estos están los unidireccionales, que solo van de la lengua A a la lengua B, y los bidireccionales, que también realizan traducciones a la inversa (B→A). Existen también sistemas multilingües que disponen de más de dos idiomas para generar combinaciones lingüísticas, pero el hecho de que tenga por ejemplo cinco lenguas no significa que pueda traducir en todas las combinaciones; puede ser que traduzca de un solo idioma a otros cuatro, haciendo así únicamente cuatro combinaciones lingüísticas. Además, para ser un sistema plenamente multilingüe, es necesario que el sistema no cambie de análisis de LO cuando cambie la LM. Así, si la LO es inglés, el análisis del texto inglés deberá ser siempre igual independientemente de que la LM sea francés, chino o español, y esto puede inducir a muchos errores si el campo léxico no está muy limitado. Un ejemplo de sistema multilingüe puro es Ariane, de GETA (Hutchins & Somers, 1992, págs. 69-71).

Por otro lado, los sistemas de TA se pueden dividir en los basados en reglas y los sistemas en corpus lingüísticos, o los híbridos, que se basan en ambas modalidades.

Sistemas basados en reglas

Dentro de los sistemas basados en reglas están los sistemas de traducción directa y los de traducción indirecta, que son sistemas de transferencia y sistemas interlingua.

Sistemas de traducción directa

Primero se explicará el sistema de traducción directa. Los programas desarrollados con este sistema se conocen como programas de TA de primera generación. Este sistema produce una traducción directa de la LO a la LM sin ninguna interpretación o representación intermedia. Estos sistemas se basan sobre todo en diccionarios bilingües, es decir, en una traducción palabra a palabra, recolocando después las palabras en la LM

para que la estructura de la frase sea similar a aquellas de la LM, pero no existe un análisis sintáctico. Este sistema fue el utilizado por los programas que surgieron en los años 50 y 60.

Sistemas de traducción indirecta de transferencia

En segundo lugar, existen los sistemas de traducción indirecta de transferencia. Son más complejos, ya que integran un grado de análisis sintáctico que ayuda a evitar una traducción palabra por palabra, pudiendo así dar un resultado más idiomático siempre que la información del análisis sea correcta.

Sistemas de interlingua

Y, por último el sistema más ambicioso, denominado «de interlingua», que supone una representación abstracta del contenido que se va a traducir; desde esta representación se genera la traducción en LM. La representación intermedia debe incluir toda la información para poder generar un texto en LM sin necesidad de mirar el original. En un principio, la idea original era crear una interlingua universal que pudiera servir de intermediaria entre cualquier idioma; se intentó usar el esperanto o el volapük pero hoy en día ya no es así y se suele usar el inglés como lengua ‘pivot’ (Poibeau, 2017, págs. 24- 32).

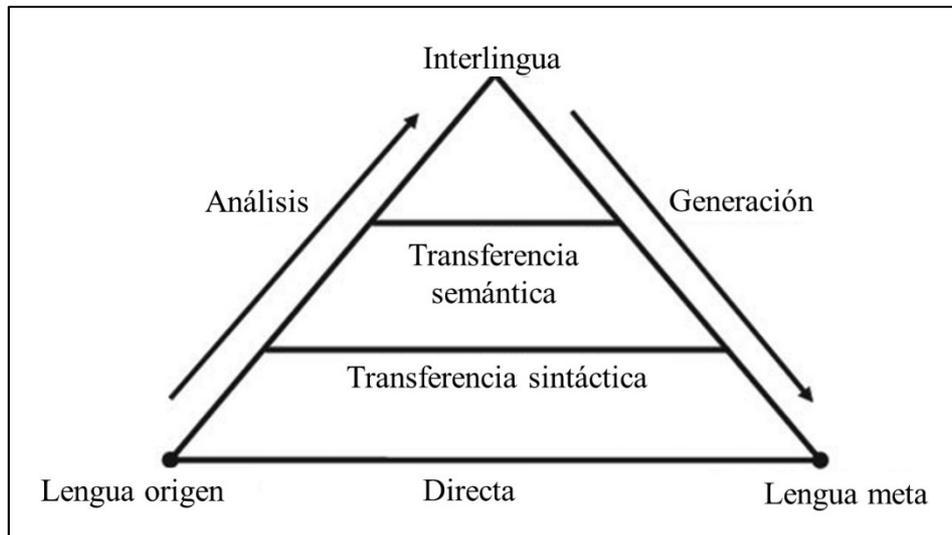


Figura 2: Triángulo de Vauquois (Trujillo, 1999).

Sistemas basados en corpus lingüísticos

Con el transcurso del tiempo, el incremento de la cantidad de información y documentación que existía sobre la traducción revolucionó el uso de la TA. Es relevante explicar la noción de compendios lingüísticos paralelos o *bitextos*, que consisten en un corpus de textos traducidos en dos o más idiomas que se alinean. Este tipo de recurso es muy común para traductores profesionales, y en la actualidad es el enfoque predominante de la TA. Todos los textos previamente traducidos, ya sea mediante TA o por un traductor profesional, quedan almacenados en la red y se pueden usar como documentación para futuras traducciones. De este sistema de almacenaje de textos ha surgido la TA basada en ejemplos y la TA basada en estadísticas.

Sistemas basados en ejemplos

La TA basada en ejemplos fue desarrollada por el japonés Makoto Nagao que consideró que, en vez de desarrollar nuevos diccionarios y métodos de análisis y transferencia, sería una mejor idea usar directamente fragmentos de traducciones de corpus bilingües (Poibeau, 2017, pág. 109). El sistema basado en ejemplos funciona en tres fases: primero, el programa intenta buscar en el corpus lingüístico fragmentos de la frase que se pretende traducir en la LO y los almacena, después el sistema busca equivalencias en los textos alineados traducidos y, por último, el programa intenta combinar los fragmentos de la traducción para conseguir una frase adecuada en la LM. Este tipo de sistema se implementó sobre todo en programas de traducción para lenguas asiáticas por su similitud en la estructura, que no existe en el resto de las lenguas y, por otro lado, también fue muy utilizado en subcampos específicos, donde el lenguaje es más limitado y donde se suele usar la misma terminología y expresiones.

Este sistema basado en ejemplos se ha unido en muchas ocasiones con el sistema basado en estadística, demostrando ofrecer resultados muy satisfactorios. A continuación se explicarán los sistemas basados en estadística.

Sistemas basados en estadística

El objetivo de este enfoque es acumular un corpus lingüístico muy amplio para extraer automáticamente lexicones bilingües; así, a cada palabra se le ofrecen diferentes posibles traducciones y se les da una puntuación en función de sus probabilidades de ser correcta. Encontrar la traducción más adecuada implica que el programa deba pasar por

una multitud de opciones, ya que cada palabra tiene múltiples equivalencias, o puede que no sea posible traducirse a la LM. Sin embargo, estos sistemas, a pesar de su gran complejidad, no han sido capaces de prever los problemas encontrados; el corpus bilingüe existente no es suficiente para textos más específicos o técnicos. Una de las soluciones que se ha propuesto es la integración de conocimientos lingüísticos en los programas para que puedan representar mejor la relación entre la sintáctica y la semántica. El éxito de este sistema ha hecho que se integre en otros sistemas de transferencia o de interlingua obteniendo mejores resultados.

Los programas con sistemas basados en reglas siguen en funcionamiento a pesar de que se hayan desarrollado sistemas más modernos y con mejores resultados, ya que existen lenguas que no disponen de un gran corpus lingüístico en el que basarse para realizar traducciones, como por ejemplo el vasco.

Sistemas basados en aprendizaje neuronal

El último sistema que se ha desarrollado es la TA neuronal. Este sistema consiste en crear un sistema en el que haya que introducir una cantidad mínima de información. Es el propio sistema que utiliza un ‘codificador’, que analiza todos los textos introducidos en el sistema como parte del ‘entrenamiento’; a partir de ahí el ‘decodificador’ es capaz de generar una traducción de cualquier frase. Este sistema es capaz de analizar varios tipos de información a la vez, como la sintáctica, la semántica o el contexto comunicativo, generando así una traducción mucho más fiable. Al ser capaz de analizar frases en su contexto, los programas que se basen en TA neuronal tienen más posibilidades de generar traducciones que se separen más del original pero que aun así sean correctas; pueden separarse más del tradicional ‘palabra por palabra’ y acercarse a una lengua meta más natural. Este sistema recibe críticas por su falta de base teórica: es decir, cuando el sistema se encuentra con una frase o palabra que desconoce no se basa en teoría aprendida para solucionarlo, sino que depende de su buen funcionamiento, esto es el aprendizaje previo que ha realizado el programa con los textos introducidos. En el año 2016 se instaló este sistema por primera vez en un TA en internet, y ahora, la mayoría de los programas han comenzado a usar este sistema (Systran, Facebook, etc.): ha supuesto una revolución para la TA, y se espera una rápida progresión en un futuro cercano.

5.2. PARÁMETROS DE EVALUACIÓN DE LA TRADUCCIÓN AUTOMÁTICA

Se puede afirmar que la principal conclusión a la que han llegado la mayoría de los traductólogos en torno a la cuestión de la equivalencia es que el efecto deseable en la lengua meta es el de facilitar la comprensión por parte del lector sin renunciar a la fidelidad con respecto al texto origen (Aranda, 2012, pág. 148).

Para conseguir el objetivo de este estudio, es necesario establecer unos parámetros para evaluar la traducción generada por los diferentes traductores automáticos que van a formar parte del estudio. Estas traducciones generadas se evaluarán como si fuera una traducción realizada por cualquier alumno que curse el grado de Traducción e Interpretación en la Universidad. Para ello se van a utilizar los parámetros establecidos por la profesora María Luisa Romana⁶ de la Universidad Pontificia de Comillas que se verán en la tabla a continuación. Este sistema de clasificación de errores está muy generalizado en todo España, según el estudio que realizó Christopher Waddington en el 2017 (Waddington, 2017, pág. 19). Ya que la finalidad de este trabajo es únicamente diferenciar qué tipo de error realiza cada TA y la frecuencia relativa de los errores, durante el análisis no se puntuará el trabajo de los TA, sino que solo se comentará qué tipo de error comete y qué TA comete más errores y con qué frecuencia.

⁶ M^a Luisa Romana García, material didáctico de la asignatura Teoría de la Traducción, Universidad Pontificia Comillas.

SIGNIFICADO	Ideas desarticuladas		SS	
	Discrepancia semántica (TO)	Contrasentido	CS	
		Falso sentido	FS	
		No mismo sentido	NMS	
EXPRESIÓN	Estructuras	Desarticulación sintáctica (anacoluto)	Por segmentación	SINT
			Por concordancia	
			Por indeterminación	
			Por mala subordinación	
			Por régimen pronominal	
			Por régimen preposicional	
			Por modificación incompleta	
			Por errata	
	Estilística	Disgrafías	Ortografía	ORT
			Ortotipografía	OTP
			Formato	FOR
		Terminología		LEX
		Idiomática	Registro inadecuado	
Calco			Calco específico	
	Calco asentado			
DEONTOLOGÍA	Inadaptación	Omisión	OM	
		Adición	AD	
	Documentación		DOC	

6. ANÁLISIS

Para realizar esta investigación sobre el funcionamiento de la traducción automática en línea, primero se explicarán los tres traductores automáticos que han sido seleccionados, accesibles gratuitamente en internet para cualquier persona. Y después se explicará el texto que se procesará por los diferentes programas para su análisis.

6.1. TRADUCTORES AUTOMÁTICOS

Apertium

En primer lugar tenemos Apertium⁷, un programa libre de traducción automática para desarrollar sistemas basados en reglas de traducción automática. Actualmente dispone de 49 pares de idiomas que se consideran fiables por su propia página web. Destacan diferentes idiomas y dialectos provenientes de Europa, y en particular, todos los dialectos de España. Este programa fue creado por el grupo de investigación Transducens de la Universidad de Alicante, en conjunto con varios proyectos de colaboración con otras universidades españolas (Grupo Transducens, 2007, pág. 1). El hecho de que sea un sistema de *software* libre, es decir, sin restricciones en cuanto a código y a datos, facilita que se pueda adaptar a nuevos pares de lenguas o incluirse en nuevas aplicaciones. Apertium dispone de la licencia GNU General Public License⁸ para que así se pueda

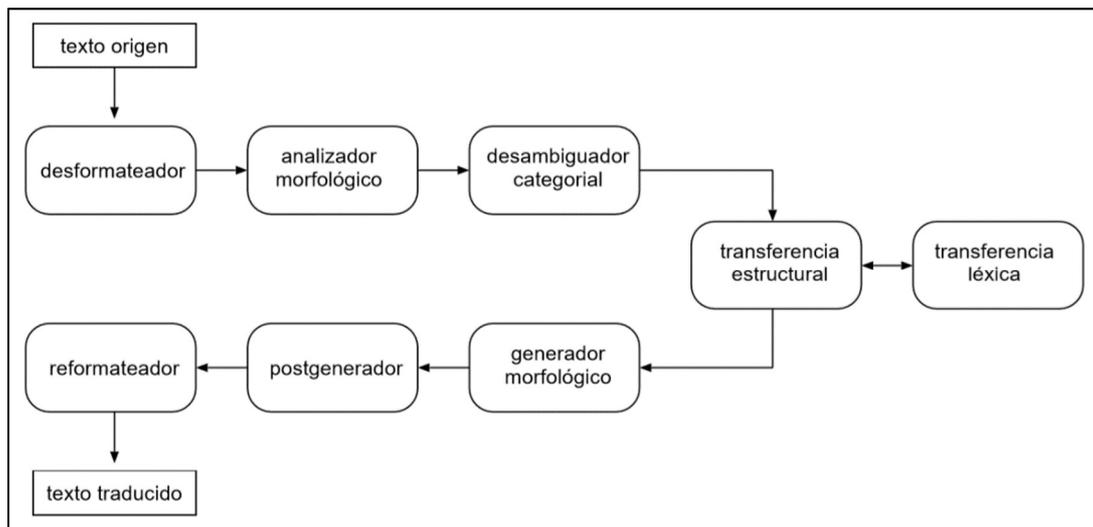


Figura 3: Módulos de un sistema de traducción automática realizado en con el *software* libre de Apertium. (Grupo Transducens, 2007, pág. 9)

⁷ <https://www.apertium.org/index.spa.html?dir=eng-spa#translation>

⁸ GNU es un sistema operativo de *software* libre que respeta la libertad de los usuarios para ejecutar, estudiar, distribuir y redistribuir copias del programa a gusto del consumidor.

comenzar a desarrollar nuevos traductores automáticos sin que sea necesario comenzar de cero. Así, “cualquier persona que tenga las habilidades informáticas y lingüísticas necesarias puede adaptar o ampliar Apertium para producir nuevos o mejores sistemas de traducción automática, incluso para nuevos pares lingüísticos” (Grupo Transducens, 2007, pág. 12).

Apertium consiste en un sistema de traducción basado en reglas, es un sistema de traducción indirecta de transferencia, que como se ha explicado anteriormente integran un grado de análisis sintáctico que ayuda a evitar una traducción de palabra por palabra, pudiendo así dar un resultado más idiomático que los sistemas de traducción directa. En la figura 3 podemos ver las diferentes fases por las que pasa texto como un análisis morfológico y una transferencia estructural y léxica basada en las reglas lingüísticas de las que dispone el programa.

A pesar de ser un programa de *software* libre y disponer una capacidad de desarrollo ilimitada que lo puede convertir en un traductor automático muy fiable, es necesario un elaborado trabajo previo por parte de una persona. Por lo tanto, podemos suponer que será el traductor menos eficiente de los tres, aunque se podría trabajar con el sistema para hacerlo eficiente en determinados campos que el traductor requiera.

Yandex.Translate

En la actualidad no es fácil encontrar sistemas de traducción que usen el sistema basado en estadística, ya que la mayoría han comenzado a implementar el sistema neuronal para un mejor resultado. Incluso el Traductor de Google ya usa un traductor automático neuronal excepto en su formato de traducción de imagen, en la que puede traducir el texto que se encuentre en una imagen, que continúa usando traducción automática por estadística (Google Translate, 2019).

La empresa Yandex es una empresa de origen ruso fundada en 2000 como empresa tecnológica de motor de búsqueda con diferentes servicios para el consumidor como búsqueda de noticias, correo electrónico y demás. Y poco a poco, según se ganaba un espacio en el mercado, desarrolló otros servicios, como mapas, anuncios, redes sociales, mensajería instantánea y finalmente, en el año 2011, la empresa lanzó su propio

sistema de traducción Yandex.Translate⁹, basado en traducción automática estadística. Este programa comenzó traduciendo únicamente de inglés y ucraniano a ruso y viceversa, y más tarde se han ido ampliando hasta llegar en la actualidad a 95 idiomas o dialectos, entre los que se incluyen el vasco o el gallego.

Ya en el año 2017, la empresa rusa comenzó a barajar la utilización de un sistema neuronal para su traductor, pero por el momento no ha sido implementado. Aunque en España, como en otras partes de Europa y del mundo, el motor de búsqueda más común sea Google, que representa del 80 al 90 por ciento de las búsquedas en internet en algunos países como Inglaterra y Francia, el motor de búsqueda nacional en Rusia suponía el 64 por ciento de las búsquedas en el año 2010, mientras que Google únicamente disponía del 22 por ciento (Agarwal & Round, 2011, págs. 122-125). Esto puede deberse a que en un principio Google no encontró fácil adaptarse al idioma ruso, lo que suponía un problema de para la comprensión.

DeepL

Por último, tenemos el sistema basado en la traducción neuronal automática. Este supone el sistema más avanzado hasta el momento, que ofrece los mejores resultados, que se acercan mucho a un resultado natural y correcto cuando no se exige un texto muy enrevesado o técnico. Existen ya muchos programas que implementan este sistema como el Traductor de Google o el Traductor de Microsoft. Pero para este trabajo se ha elegido DeepL¹⁰.

Este sistema fue creado por el equipo de Linguee de origen alemán en el año 2017. DeepL directamente comenzó siendo un sistema de conexiones neuronales, no como el resto de los traductores automáticos que actualizaron sus tecnologías. Únicamente cuenta con 9 idiomas, lo que supone 72 combinaciones lingüísticas (alemán, español, francés, inglés, italiano, neerlandés, polaco, portugués y ruso) siendo el portugués y el ruso los añadidos más recientemente (DeepL, 2019).

Este programa se considera el mayor competidor contra el Traductor de Google, que durante muchos años fue referente de la traducción automática. De momento, DeepL solo ofrece la traducción de frases o textos en estos 9 idiomas, mientras que Google ofrece

⁹ <https://translate.yandex.com/>

¹⁰ <https://www.deepl.com/translator>

la traducción de voz, texto e imágenes en numerosos idiomas (Pastor, 2019). Así que, en cuanto a variedad, Google continúa siendo el mayor exponente; sin embargo, en cuanto a calidad, diferentes blogs de profesionales califican las traducciones de DeepL de superiores a las de Google o cualquier otro programa de traducción automática (AT Language Solutions, 2019).

6.2. TEXTO ELEGIDO

El texto elegido para procesar por los diferentes programas es un texto dirigido al público general que tenga cierto conocimiento o interés por el tema, que a la vez incorpora algunos elementos técnicos y un vocabulario más específico. Se trata de un texto científico expositivo que presenta de forma objetiva la situación y los efectos de los microplásticos en el océano (Anexo 1). El texto ha sido extraído de la introducción de un artículo de una revista científica de la editorial Springer, denominada *Current Environmental Health Reports*, y el artículo se titula «Microplastics in Seafood and the Implications for Human Health», escrito en el año 2018; trata sobre los recientes descubrimientos de los efectos de los microplásticos que se encuentran en el océano sobre los alimentos marinos que los humanos consumimos y cómo esto afecta a nuestra salud.

6.3. ANÁLISIS DEL TEXTO

Una vez se han procesado este mismo texto por los tres diferentes programas de traducción que usan sistemas diferentes, se han analizado los errores y se han dividido en las categorías de errores expuestas anteriormente, para así poder comparar el funcionamiento de los sistemas neuronales, estadísticos y basados en reglas. El texto elegido procesado por los tres programas analizados, además de una traducción realizada por la autora de este TFG que se ha considerado *tertium comparationis* para la consecuencia de la finalidad de este estudio, se encontrarán en anexo al final del trabajo, añadiendo en comentarios los errores que se han encontrado en cada apartado, así como la categoría de cada error. Dentro del capítulo de análisis se podrá encontrar un apartado del texto con los respectivos comentarios para poder así exponer brevemente una conclusión de los errores encontrados en cada texto y las diferencias con los otros. En la tabla que se encuentra a continuación se indica el número de errores que ha cometido cada TA, divididos en apartados según la categoría, así como el total de errores.

Errores DeepL		Errores Yandex		Errores Apertium	
Sin sentido (SS)	1	Sin sentido (SS)	3	Sin sentido (SS)	26
Falso sentido (FS)	1	Falso sentido (FS)	2	Falso sentido (FS)	—
No mismo sentido (NMS)	1	No mismo sentido (NMS)	—	No mismo sentido (NMS)	1
Anacoluto (SINT)	1	Anacoluto (SINT)	5	Anacoluto (SINT)	24
Ortografía (ORT)	—	Ortografía (ORT)	—	Ortografía (ORT)	1
Ortotipografía (OTP)	1	Ortotipografía (OTP)	—	Ortotipografía (OTP)	1
Formato (FOR)	1	Formato (FOR)	5	Formato (FOR)	3
Terminología (LEX)	8	Terminología (LEX)	12	Terminología (LEX)	10
Idiomaticidad (ID)	7	Idiomaticidad (ID)	7	Idiomaticidad (ID)	7
Omisión (OM)	—	Omisión (OM)	3	Omisión (OM)	—
Adición (AD)	1	Adición (AD)	2	Adición (AD)	2
Total DeepL	22	Total Yandex	39	Total Apertium	75

Después de haber analizado los errores encontrados en los textos traducidos por los diferentes TA, se han sumado y se ha expuesto el total de errores de cada TA en la tabla anterior. Hay una clara diferencia entre los tres TA y los sistemas que estos usan. Sería interesante ver cómo se comportan con este mismo texto, otros TA que usen los mismos sistemas de traducción y ver las diferencias entre sistemas similares. Pero, por razones de espacio y de tiempo, en este trabajo solo se analizan estos tres TA.

En primer lugar, DeepL con un sistema de traducción basado en redes neuronales, en segundo lugar, Yandex con un sistema basado en estadística y, por último, Apertium con un sistema basado en reglas. El sistema de conocimiento neuronal es la última tecnología que ha surgido en el mundo de la traducción automática y por lo tanto era de espera que obtuviera los mejores resultados con únicamente 22 errores totales. El sistema basado en estadística es también de los mejor considerados en el mundo de la traducción automática, aunque en la actualidad está siendo sustituido en todos los motores por el sistema neuronal. Yandex ya se ha planteado cambiar a neuronal, pero por el momento continua con estadística y ha cometido 39 errores, casi llegando a duplicar la cantidad de errores cometidos por DeepL. Y, además tenemos el sistema basado en reglas que fue uno de los primeros que se idearon para desarrollar la TA, pero ya ha quedado más que demostrado que este sistema no es el más idóneo. La única razón por la que un programa como Apertium existe es porque constituye un programa de *software* libre que otros individuos o grupos pueden comprar y modificar para adaptar a una traducción más específica que siga ciertas reglas. Así se puede mejorar sus resultados haciéndolo más

exacto para ciertas traducciones que se quieran obtener. Sin embargo, el programa Apertium sin modificar ha demostrado ser el que peor resultados ha obtenido sin duda, cometiendo 75 errores, casi el doble que el programa de Yandex, es decir, casi cuadruplica los 22 errores de DeepL.

A continuación pasaré a comentar los resultados que más han llamado mi atención y que considero más relevantes para este estudio, dando ejemplos de ciertos errores. Sin embargo, en caso de querer observar los errores señalados al completo en el propio texto, estarán adjuntos en anexo como se ha mencionado anteriormente.

El primer texto para analizar será el de DeepL ya que es el que menos errores contiene, únicamente 22. De estos 22 errores, la mayoría han sido errores de terminología (8) y de Idiomaticidad (7); dentro de este último se engloban errores como calcos y el uso de un registro inadecuado. Ambos tipos de errores, terminología e Idiomaticidad, hacen referencia a la estilística de la expresión del texto meta. Sin embargo, los fallos de significado, deontología y estructura no han sido tan numerosos, aunque sí ha habido. Un error que destacar es, por ejemplo, la traducción de dos palabras diferentes en el texto origen *trash* y *litter* por la misma en el texto meta *basura*.

Texto origen EN	Traducción de DeepL ES
Discarded plastic materials enter the marine environment as trash, industrial discharge, or litter through inland waterways, wastewater outflows, and transport by winds or tides.	Los materiales plásticos desechados entran (LEX) al ambiente marino (LEX) como basura , descargas industriales o basura (LEX) a través de las vías fluviales interiores, desagües de aguas residuales y transporte (SS) por vientos o mareas.

Las palabras *trash* y *litter* sí que podrían traducirse por ‘basura’ en ciertos contextos, y no sería un error, pero en este caso es necesario diferenciar, ya que claramente está haciendo referencia a elementos diferentes. En este caso, *trash* se podría traducir por ‘residuo’ y *litter* por ‘basura’, para que así, además, no se caiga en una repetición. En este párrafo también podemos encontrar otros errores como el calco de la palabra ‘entrar’ como traducción de *enter*. Esto puede calificarse como calco que lleva a un error de registro, ya que sería más correcto decir que los plásticos ‘llegan’ o ‘terminan’ o ‘acaban’ en el océano, ya que el verbo ‘entrar’ puede implicar una acción por parte del

plástico que no es real. Asimismo, se ha traducido *marine environment* por ‘ambiente marino’ cuando la traducción correcta sería ‘entorno’. Esto supone un error de terminología. Y, por último, en la última frase hay un sinsentido en la frase que resulta de un calco de la estructura sintáctica del original; sería más correcto decir ‘transportado por vientos’ para que encajara mejor en la frase.

El resto de los errores encontrados en el texto son similares, de términos que no son los adecuados para el contexto o de preposiciones erróneas. Sin embargo, el sentido general del texto no se pierde, y es posible seguir una lectura continuada y conocer lo que pretende transmitir, por lo tanto, esta traducción es fiel a la finalidad del texto a pesar de no proporcionar una traducción 100 % correcta.

El segundo texto analizado es el del motor de búsqueda ruso Yandex. De los 39 errores cometidos por este programa, destacan por encima del resto los errores de terminología con 12 errores, seguidos de 7 errores de Idiomaticidad. Los errores cometidos por este programa son más serios y sí que dificultan la lectura del texto. En algunos casos se han insertado mayúsculas donde no debía, no se han modificado la numeración para que encajara con la española, se han dejado palabras sin traducir o se ha inventado palabras. Como ejemplo, vamos a ver el mismo párrafo que se ha usado como ejemplo con DeepL.

Texto origen EN	Traducción de Yandex ES
Discarded plastic materials enter the marine environment as trash, industrial discharge, or litter through inland waterways, wastewater outflows, and transport by winds or tides.	Los materiales plásticos desechados entran (LEX) en el ambiente (LEX) marino como basura , literalmente figuran (SS), o basura (LEX) a través de las vías actualizadables (SINT) interiores, las corrientes de aguas residuales, y el transporte (SS) por los vientos o la libertad (LEX).

Con este ejemplo queda bastante claro que no resulta sencillo leer el texto y comprender lo que pretende transmitir. En este único párrafo ya se pueden encontrar 6 errores, entre ellos los que ya fueron cometidos por DeepL como el uso del verbo ‘entrar’, el uso del término ‘ambiente marino’ o la repetición de la palabra ‘basura’. Además, hay otros errores como la traducción de *industrial discharge* por ‘literalmente figuran’ lo que no tiene sentido alguno o el uso de la palabra ‘actualizadables’ como traducción de

inland waterways que no solo no tiene sentido, sino que además es una palabra que no existe en español por lo tanto es una errata. Y, por último, la traducción de *tides* por ‘libertad’, que también carece de sentido. Es posible, que la palabra *marea* se haya podido usar como sinónimo de *libertad* en algún texto más literario, y de ahí que el programa de estadística las haya emparejado como una traducción. En el resto del texto los errores no son tan notorios como en este mismo párrafo, pero sí que dificultan la lectura, aunque no interfieran en el significado general.

En estos dos TA, los errores más recurrentes han sido de terminología y de Idiomática, como hemos mencionado anteriormente. Este tipo de errores son los que un traductor humano no profesional también tendría más posibilidades de realizar, como no elegir la palabra correcta o caer en un calco.

Por último tenemos el texto procesado por Apertium, que ha realizado 75 errores, colocándolo de manera destacada en el peor TA de los tres. Al contrario que en los otros dos TA, los errores realizados por Apertium se encuentran sobre todo en las categorías de sin sentido (SS), ideas que no se encuentran conectadas unas con otras y hacen que la frase no tenga sentido y en anacolutos (SINT), es decir, desarticulaciones sintácticas que llevan a la inconsecuencia de la construcción del discurso. Estos dos tipos de errores dificultan en mayor medida la lectura y comprensión del texto. Para poder hacer una comparativa de los tres, a continuación, se expondrán los errores cometidos en el mismo párrafo que con los TA anteriores.

Texto origen EN	Traducción de Apertium ES
Discarded plastic materials enter the marine environment as trash, industrial discharge, or litter through inland waterways, wastewater outflows, and transport by winds or tides.	Discarded (SS) Los materiales plásticos introducen (LEX) el entorno marino cuando (LEX) basura, caudal industrial, o ensuciar (LEX) a través de inland waterways (SS), wastewater outflows (SS), y transporte (SS) por vientos o mareas

Esta traducción se diferencia más de las dos anteriores. Destacan sobre todo las faltas de traducción del programa, ya que ha dejado tres elementos sin traducir, supuestamente porque no disponía de ninguna regla para traducir elementos con un grado de tecnicidad más alto como son *discarded*, *inland waterways* o *wastewater outflows*. Sin

embargo, a pesar de haber dejado palabras sin traducir, Apertium no ha inventado ninguna palabra que no existiera en español, como es en el caso de Yandex. Por otro lado, ha tenido un error de terminología la usar el verbo ‘introducir’ para traducir *enter* cambiando así el sentido de la frase, al igual traduciendo *as* por ‘cuando’ dejando la frase sin sentido y sin transmitir lo mismo que el texto origen. Además, ha tenido el mismo error que los otros dos TA con la palabra ‘transporte’. Sin embargo, es el único de los tres TA que ha traducido correctamente la palabra *environment* por ‘entorno’ en vez de ‘ambiente’.

Además de en este párrafo, el texto completo no es legible ni comprensible, ha generado todo tipo de errores, destacando las palabras que se han dejado sin traducir y los errores de concordancia, de falta de artículos y de preposiciones de la elección de términos inadecuados. También cabe destacar que el único error de ortografía ha sido cometido por Apertium, al añadir un acento a un ‘cuando’ que no debería llevarlo. La traducción en general realizada por Apertium carece de sentido y no permite conseguir llegar a entender de qué trata el texto.

6.4. REVISIÓN POST-TRADUCCIÓN

También resulta interesante y relevante para los objetivos de este trabajo analizar cómo resultaría el trabajo de posesición de un texto que ha sido procesado por estos programas para su traducción. Como se ha mencionado anteriormente, cada vez es más común recurrir a estas herramientas como sustituto de un traductor profesional. Para ciertos textos puede ser una buena opción siempre y cuando haya una posesición y revisión por parte de un profesional. La revisión es un trabajo de los traductores que ya se realizaba anteriormente. Sin embargo, hoy en día es común ver a traductores que ofrecen sus servicios de posesición o revisión para una traducción realizada por un TA que resulta más barato para las empresas que contratan los servicios y requiere menos tiempo por parte del traductor, o al menos así debería ser.

Si comenzamos por la traducción realizada por el programa DeepL, la posesición resultaría muy sencilla de realizar. Los errores no impiden la comprensión de ninguna de las frases del texto y los errores consisten principalmente en calcos y el uso de un registro inadecuado fácilmente identificables por un traductor profesional. Este texto no supone ningún reto muy complicado para un traductor profesional y no es necesaria una documentación muy exhaustiva, por lo que el uso de DeepL con una posterior edición por

parte de un traductor profesional sería un buen método de ahorrar tiempo y dinero en esta traducción.

Por otro lado, con las traducciones realizadas por Yandex.Translate y Apertium no ocurriría lo mismo. Los tipos de errores realizados por estos TA son mucho más complejos y dificultan la comprensión y lectura del resultado. El hecho de tener que leer este texto para editar la traducción supondría que el trabajo del traductor que realice la edición posterior se alargara más intentando reelaborar las frases y los términos que ha ofrecido el TA. En el caso de Yandex, muchas de las palabras que aparecen sin sentido, harían que el traductor tuviera que volver al texto origen y realizar el trabajo de documentación. Con la traducción realizada por Apertium, sería imposible que ni siquiera un traductor profesional pudiera modificar ese texto para obtener una traducción bien hecha y fiel al original, en este caso supondría un reto mucho mayor para el traductor que una simple traducción. Asimismo, resulta más fácil para un traductor formular una frase en la lengua meta desde el texto origen que disponer de una frase mal formulada en la lengua meta y tener que reformularla para que resulte más natural. En general, este proceso podría alargarse más que si fuese el mismo traductor profesional el que realizase la traducción desde un principio como afirma la Asociación Española de Traductores, Correctores e Intérpretes “[...] muchos textos procedentes de la traducción automática son tan macarrónicos que empezar desde cero lleva menos tiempo (y dinero) que «arreglarlos»” (Equipo ASETRAD, 2012).

De esto se concluye que, en el caso de un TA con aprendizaje neuronal, este podría resultar un recurso muy útil para un traductor profesional o para una empresa que busque una primera traducción para su posterior revisión y edición. Esto en caso de que se trate de un texto sencillo sin lenguaje demasiado técnico ni literario que pudiera causar que el TA cometiera errores difíciles de resolver para el traductor profesional. En el periodo de posesión del traductor profesional se podría llevar a cabo una traducción más oblicua, es decir, separarse de la traducción literal a la que tienden los TA y generar una traducción con una transformación y combinación de elementos que busquen una mayor naturalidad (Vinay & Darbelnet, 1995, págs. 30-41).

Estos traductores también pueden servir de ayuda para las personas que no busquen un servicio profesional y únicamente quieran comprender la idea general de un texto, y no quieran usar esa traducción para ningún fin más allá del personal.

7. CONCLUSIÓN

El interés de este trabajo surge del avance de las nuevas tecnologías y de la supuesta amenaza que esto significa para el trabajo de los traductores profesionales. Tras cinco años estudiando la carrera de traducción e interpretación, no ha habido una sola asignatura en la que no se haya valorado la importancia de las herramientas de traducción automática, ya sea positiva o negativamente. Debido a esto he considerado muy interesante realizar este Trabajo de Fin de Grado en el que poder valorar cómo han evolucionado los diferentes TA y cómo funcionan los diversos sistemas que existen, además de investigar la posibilidad de su aplicación con una posterior edición profesional.

Los programas de traducción automática en internet comenzaron a aparecer sobre la década de 1980 y desde entonces no han hecho más que desarrollarse y avanzar tecnológicamente. Existe un gran interés en estos programas gracias a la globalización y el mundo interconectado en el que vivimos. Todas las empresas e instituciones que busquen una expansión y reconocimiento internacional necesitan exponer sus servicios a diferentes países en sus idiomas nativos. Por lo tanto, cada año que pasa es más necesaria la labor de un traductor, lo que supone un gran gasto para las empresas y el papel de los TA es muy importante y puede traer un gran número de ventajas.

Sí, es cierto que poseer las traducciones de un motor de TA de mala calidad puede implicar toda una retraducción y una inversión de tiempo mayor, pero si la calidad del motor es buena, el aumento de la productividad puede ser significativo. (Collantes, Mallo, Parra, Serrano, & Quiñones, 2018)

Y aquí entra en escena otra amenaza para los traductores: si el trabajo lo hace una TA y ellos solo han de revisar el texto, el beneficio que obtengan de ese trabajo no será tan alto como el de una traducción; sin embargo, el tiempo invertido también es mucho menor lo que permite trabajar en más proyectos en menos tiempo.

Siempre y cuando se cuente de partida con un motor bien entrenado que ofrezca unos resultados de calidad, el aumento de la productividad que ofrece la TA permite también al traductor ofrecer plazos de entrega más ajustados, lo que en la mayoría de los casos supone una ventaja competitiva. (Collantes, Mallo, Parra, Serrano, & Quiñones, 2018)

A lo largo de este Trabajo de Fin de Grado hemos podido ver como los resultados obtenidos por los tres TA estudiados diferían muchos los unos de los otros. El sistema más novedoso y avanzado que obtiene mejores resultados es el sistema neuronal que en la actualidad ya usan diferentes programas de traducción en línea, como el Traductor de Google y Reverso. De los 22 errores cometidos por DeepL, ninguno podía llegar a entorpecer la comprensión del texto o el posterior trabajo de edición por parte de un traductor profesional, sino todo lo contrario, facilitar y ofrecer buenas soluciones de traducción para así hacer el trabajo del traductor más eficiente y eficaz.

Los sistemas de Traducción asistida por ordenador también han generado un gran avance y ayuda para los traductores profesionales y ahora uno de estos sistemas de TAO, SDL Trados ha incorporado en su sistema la interfaz de DeepL por lo que ambos programas podrán ayudar conjuntamente a generar una traducción de calidad (Presentamos la API de DeepL, 2018). Estos programas van a continuar desarrollándose debido a su gran interés laboral y económico, ya que suponen un gasto mucho menor de personal y de dinero. Sin embargo, los otros dos sistemas analizados en este estudio han quedado obsoletos, a no ser que se genere un corpus lingüístico específico y propio para una traducción; pero esto requeriría un trabajo previo de creación de ese corpus que, por lo tanto, supondría una mayor inversión de tiempo y dinero, que si fuera a largo plazo podría llegar a ser rentable.

Por contraste, existen ciertos tipos de textos que no se ven tan ‘amenazados’ por la TA; entre estos podemos encontrar los textos literarios, los poemas, los textos humorísticos, aquellos que presenten un gran número de figuras literarias difíciles de captar para el programa, etc. En este caso, es más difícil que el papel del traductor se vea sustituido por un programa informático ya que ocurriría lo mencionado anteriormente sobre el incremento de tiempo en corregir una traducción mal formulada (Barrusto, 2014).

Otros problemas que hemos podido ver en el análisis realizado para este TFG han sido las palabras que pueden tener varias traducciones diferentes y que, si el TA no es capaz de identificar el contexto (que en muchas ocasiones no lo es), no podrá elegir la adecuada. Esto lo hemos podido ver sobre todo con la traducción de *environment* por ‘medio ambiente’ en vez de ‘entorno’. Pero este tipo de modificación no supone ningún reto para un traductor profesional. Otro asunto por resolver es el hecho de que los TA realizan unas traducciones muy pegadas al original, es decir, la estructura sintáctica no varía y existen ocasiones que requieren una reformulación de la frase al completo para

mostrar naturalidad en la LM y, por el momento, los TA no han demostrado una gran habilidad en este campo. Aunque si continúan desarrollándose y aumentando su corpus lingüístico, es muy posible que en unos años sí que dispongan de esta capacidad, así como de la capacidad para identificar el contexto y poder así elegir la traducción más adecuada de cada palabra.

No se puede negar el hecho de que los traductores automáticos hayan llegado al ámbito profesional y la necesidad de su integración. Es necesario que los traductores profesionales, así como aquellos estudiantes que se estén preparando para serlo, tengan un amplio conocimiento de cómo manejar y hacer un mejor y correcto uso de estas herramientas. La posibilidad de que los traductores automáticos sustituyan plenamente la labor de un profesional, aunque sea en un texto sencillo, se encuentra en un futuro muy lejano y, por supuesto, queda descartado que se pueda poner en manos de un TA un texto técnico, especializado o con interés literario. Es necesario considerar estas herramientas como un gran aliado del traductor profesional y conocer las ventajas que puede traer a esta profesión y ayudar a su desarrollo adecuado. Al igual que en las últimas décadas, en los próximos años este campo de estudio va a recibir una gran atención con grandes inversiones y estudios sobre él, por lo tanto se desarrollará con gran velocidad y es imprescindible que los traductores formen una parte fundamental de este proyecto.

8. BIBLIOGRAFÍA

- Agarwal, M., & Round, D. K. (2011). The Emergence of Global Search Engines: Trends in History and Competition. *Competition Policy International*, 115-134.
- ALPAC. (1966). *Language and Machines: Computers in Translation and Linguistics*. Washington D.C.: National Research Council.
- Aranda, M. (2012). *La influencia del uso de herramientas de traducción asistida por ordenador sobre la calidad y la creatividad*. Madrid: Universidad Pontificia de Comillas.
- Arenas, A. G. (diciembre de 2010). Exploring Machine Translation on the Web. *Revista tradumàtica*(8). Obtenido de Revista tradumàtica: <http://www.fti.uab.cat/tradumatica/revista>
- Arnold, D., Balkan, L., Meijer, S., Humphreys, R., & Sadler, L. (1994). *Machine Translation: an Introductory Guide*. Londres: NCC Blackwell.
- AT Language Solutions. (12 de febrero de 2019). *Google Translate vs DeepL*. Obtenido de AT Language Solutions: <https://www.at-languagesolutions.com/atblog/google-translate-vs-deepl/>
- Barrusto, H. (2014). *Evolución y Aplicaciones de la Traducción Automática*. Madrid: Universidad Pontificia Comillas.
- Collantes, C., Mallo, J., Parra, C., Serrano, R., & Quiñones, H. (noviembre de 2018). Pásate al lado oscuro: ventajas de la traducción automática para el traductor profesional. *La Linterna del Traductor*. Obtenido de La Linterna del Traductor: <http://www.lalinternadeltraductor.org/n17/ventajas-traduccion-automatica.html>
- Dureja, M., & Gautam, S. (noviembre de 2015). Speech-to-Speech Translation: a Review. *International Journal of Computer Applications*, 129(13), 28-30.
- Ehsani, F., Master, D., & Zuber, E. D. (2016). *Estados Unidos Patente n° US 9.430,467 B2*.
- Equipo ASETRAD. (2012). *Todo lo que siempre quiso saber sobre la traducción*. Obtenido de ASETRAD: https://asetrad.org/images/asetrad/pdf/como_encargar_trad/GIR_espaol.pdf
- Flanagan, M. (1996). Two Years Online: Experiences, Challenges and Trends. *Expanding MT Horizons: Proceedings of the Second Conference of the Association for Machine translation in the Americas* (págs. 192-197). Montreal: The Association.

- Forcada, M. L., Ginestí-Rosell, M., Nordfalk, J., O'Regan, J., Ortiz-Rojas, S., Pérez-Ortiz, J. A., Tyers, F. M. (2011). Apertium: a free/open-source platform for rule-based machine translation. *Machine Translation*, 127-144.
- Gaspari, F., & Hutchins, J. (2007). Online and Free! Ten Years of Online Machine Translation: Origins, Development, Current Use and Future Prospects. *Proceedings of the Machine Translation Summit XI*, (págs. 199-206).
- Gómez Grajales, C. A. (23 de junio de 2015). *The Statistics behind Google Translate*. Obtenido de Statistics Views: <https://www.statisticsviews.com/details/feature/8065581/The-statistics-behind-Google-Translate.html>
- Grupo Transducens. (2007). *Apertium, una plataforma de código abierto para el desarrollo de sistemas de traducción automática*. Alicante: Servicio de Publicaciones de la Universidad de Cádiz.
- Hillel, Y. B. (1959). *Report on the State of Machine Translation in the United States and Great Britain*. Jerusalem: Hebrew University.
- Hutchins, J. (2010). *Compendium of Translation Software: Directory of commercial machine translation systems and computer-aided translation support tools*. European Association for Machine Translation. Ginebra: European Association for Machine Translation. Obtenido de <http://eamt.org/compendium/index.php>
- Hutchins, W. J. (2001). Machine Translation over Fifty Years. *Histoire, Epistemologie, Langage*, XXII(1), 7-31.
- Hutchins, W., & Somers, H. (1992). *An Introduction to Machine Translation*. San Diego: Academic Press Limited.
- Lotz, S., & van Rensburg, A. (2016). Omission and other sins: Tracking the quality of online machine translation output over four years. *Stellenbosch Papers in Linguistics*, 46, 77-97.
- Martín Juarrós, V. (2017). *La formación de poseedores mediante el uso de las TIC: estado de la cuestión y una propuesta de estrategia de formación*. Universitat de Vic-Universitat Central de Catalunya y Universidad San Jorge.
- Pastor, J. (14 de marzo de 2019). *DeepL vs Google Translate: ¿quién gana la batalla de los traductores online?* Obtenido de Xataka: <https://www.xataka.com/servicios/deepl-vs-google-translate-quien-gana-batalla-traductores-online>
- Poibeau, T. (2017). *Machine Translation*. Londres - Cambridge: The MIT Press.

- Presentamos la API de DeepL*. (20 de marzo de 2018). Obtenido de Blog de DeepL:
<https://www.deepl.com/es/blog/20180305.html>
- Smith, M., Love, D. C., Rochman, C. M., & Neff, R. A. (16 de agosto de 2018). Microplastics in Seafood and the Implications for Human Health. *Current Environmental Health Reports*, 5, 375-386. Obtenido de <https://link.springer.com/article/10.1007/s40572-018-0206-z>
- Trujillo, A. (1999). *Translation Engines: Techniques for Machine Translation*. Nueva York: Springer.
- Vinay, J.-P., & Darbelnet, J. (1995). *Comparative Stylistics of French and English: a Methodology for Translation*. Amsterdam/Filadelfia: John Benjamins Publishing Company.
- Waddington, C. (2017). Should translations be assessed holistically or through error analysis? *HERMES - Journal of Language and Communication in Business*, 14(26), 15-37. doi:<https://doi.org/10.7146/hjlcb.v14i26.25637>
- Weaver, W. (1949). *Translation*. Nueva York: The Rockefeller Foundation.
- Weiss, R. J., Chorowski, J., Jaitly, N., Wu, Y., & Chen, Z. (12 de junio de 2017). *Sequence-to-Sequence Models Can Directly Translate Foreign Speech*. Obtenido de Arxiv: <https://arxiv.org/pdf/1703.08581.pdf>
- Yandex.com. (2018). *History*. Obtenido de Yandex: <https://yandex.com/company/history/2018>
- Yang, J., & Lange, E. (2003). Going live on the Internet. En H. Somers, *Computers and translation: a translator's guide* (págs. 191-210). Amsterdam: John Benjamins.

9. ANEXOS

TEXTO ORIGINAL	TRADUCCIÓN PERSONAL
<p>Since the 1960s, plastic production has increased by approximately 8.7% annually, evolving into a \$600 billion global industry. Approximately eight million metric tons of plastics enter the oceans annually, and conservative estimates suggest 5.25 trillion plastic particles currently circulate in ocean surface waters. While some plastics enter oceans from maritime operations, 80% is suspected to originate from land-based sources.</p>	<p>Desde los años 60, la producción de plástico ha aumentado un 8,7 % cada año y se ha convertido en una industria mundial valorada en 600 mil millones de dólares. Cerca de 8 millones de toneladas de plásticos acaban en los océanos cada año, y ciertas estimaciones prudentes sugieren que, en la actualidad, 5,25 billones de partículas plásticas circulan por las aguas superficiales de los océanos. Aunque algunos plásticos llegan al mar mediante actividades marinas, se sospecha que el 80 % provienen de la tierra.</p>
<p>Discarded plastic materials enter the marine environment as trash, industrial discharge, or litter through inland waterways, wastewater outflows, and transport by winds or tides. Waste generation and waste leakage are inextricably linked and proportionally associated with economic development,</p>	<p>El material plástico desechado termina en el medio marino como residuo, vertido industrial o basura a través de canales, salidas de aguas residuales y transportado por vientos o mareas. La generación de residuos y su fuga son inherentes y están asociados proporcionalmente al desarrollo económico, la infraestructura</p>

<p>local infrastructure, and legislation.</p> <p>Today, uncollected waste accounts for 75% of these land-based discharges, while the remaining 25% comes from within the waste management system.</p>	<p>local y la legislación. En la actualidad, la basura no recolectada supone un 75 % de estos vertidos terrestres, mientras que el 25 % restante proviene del sistema de gestión de residuos.</p>
<p>When plastics enter the ocean, the rate of degradation and persistence of plastics varies by polymer, shape, density, and the purpose of the plastic itself. These characteristics also govern where in the water column plastics may be found. For example, more buoyant plastics are more likely to be carried by ocean currents and wind across the environment.</p>	<p>El grado de degradación y persistencia de los plásticos que llegan al océano depende del polímero, la forma, la densidad y la finalidad del propio plástico. Estas características también influyen en la profundidad oceánica a la que se encuentra el plástico. Por ejemplo, los plásticos más flotantes tienen una mayor probabilidad de que las corrientes oceánicas y los vientos los arrastren por el entorno.</p>
<p>Additionally, when plastics are exposed to natural forces like sunlight and wave action, plastics will degrade into microplastics—defined as plastic particles under 5 mm in size. This definition commonly includes plastic pieces in the nano-scale, < 1 μm in size.</p>	<p>Además, cuando los plásticos se exponen a fuerzas naturales, como la luz solar o la acción de las olas, se degradarán en microplásticos. Estos se definen como partículas de plástico de menos de 5 milímetros que incluyen fragmentos de plástico de la nanoescala, de menos de 1 micrómetro.</p>

<p>The extent of plastic degradation depends on factors including polymer type, age, and environmental conditions like weathering, temperature, irradiation, and pH. Over time, plastic particles contaminate the marine ecosystem and the food chain, including foodstuffs intended for human consumption. <i>In vivo</i> studies have demonstrated that nanoplastics can translocate to all organs. Evidence is evolving regarding relationships between micro- and nanoplastic exposure, toxicology, and human health.</p>	<p>El nivel de degradación del plástico depende de variantes que incluyen el tipo de polímero, la antigüedad y las condiciones ambientales como el clima, la temperatura, la irradiación y el pH. Con el tiempo, las partículas de plástico contaminan el ecosistema marino y la cadena alimenticia, incluso los alimentos destinados al consumo humano. Los estudios <i>in vivo</i> han demostrado que los nanoplasticos pueden llegar a todos los órganos. Cada vez existen más pruebas en cuanto a la relación entre micro y nanoplasticos; la toxicidad y la salud humana.</p>
<p>Nutritional authorities advise Americans to double their seafood consumption; however, awareness or concerns about microplastics in seafood could lead consumers to reduce their consumption. Research to understand and reduce human health risks is critical in order to simultaneously protect consumers and support their nutritional health.</p>	<p>Las autoridades nutricionales recomiendan a los estadounidenses que dupliquen el consumo de alimentos marinos, sin embargo, la conciencia sobre los microplásticos en los alimentos marinos podría llevar a los consumidores a reducir su ingesta. Es imprescindible llevar a cabo investigaciones para entender y reducir los riesgos de la salud</p>

	humana y así proteger a los consumidores y ayudar a su salud nutricional.
--	---

TEXTO ORIGINAL	TRADUCCIÓN DEEPL
<p>Since the 1960s, plastic production has increased by approximately 8.7% annually, evolving into a \$600 billion global industry. Approximately eight million metric tons of plastics enter the oceans annually, and conservative estimates suggest 5.25 trillion plastic particles currently circulate in ocean surface waters. While some plastics enter oceans from maritime operations, 80% is suspected to originate from land-based sources.</p>	<p>Desde la década de 1960, la producción de plástico ha aumentado aproximadamente un 8,7% anual, convirtiéndose en una industria global de 600.000 millones de dólares. Aproximadamente ocho millones de toneladas métricas de plásticos entran a los océanos anualmente, y estimaciones conservadoras sugieren que actualmente circulan 5,25 billones de partículas plásticas en las aguas superficiales del océano. Mientras que algunos plásticos entran en los océanos desde operaciones marítimas, se sospecha que el 80% proviene de fuentes terrestres.</p>
<p>Discarded plastic materials enter the marine environment as trash, industrial discharge, or litter through inland waterways, wastewater outflows, and transport by winds or tides. Waste generation and waste leakage are inextricably linked and proportionally associated with economic development, local</p>	<p>Los materiales plásticos desechados entran al ambiente marino como basura, descargas industriales o basura a través de las vías fluviales interiores, desagües de aguas residuales y transporte por vientos o mareas. La generación de residuos y las fugas de residuos están inextricablemente vinculadas y proporcionalmente asociadas con el desarrollo económico, la infraestructura</p>

Comentado [MA1]: Calco ID

Comentado [MA2]: Calco asentado ID

Comentado [MA3]: AD adición innecesaria, metric tonnes se traduce por toneladas simplemente

Comentado [MA4]: Termino no adecuado, calco del TO ID

Comentado [MA5]: Reformulación correcta

Comentado [MA6]: Formato, exige una reformulación para su correcta expresión en español FOR

Comentado [MA7]: Terminología inadecuada: entorno marino LEX

Comentado [MA8]: Se traduce trash y litter por basura en ambos casos. Repetición y mala terminología LEX

Comentado [MA9]: Calco del original que genera un sin sentido en la lengua meta SS el sin sentido ya se encuentra en el texto original

Comentado [MA10]: Copia del original repitiendo la palabra residuos, estructura no muy común en el español ID

<p>infrastructure, and legislation. Today, uncollected waste accounts for 75% of these land-based discharges, while the remaining 25% comes from within the waste management system.</p>	<p>local y la legislación. Hoy en día, los residuos no recogidos representan el 75% de estos vertidos en tierra, mientras que el 25% restante procede del sistema de gestión de residuos.</p>
<p>When plastics enter the ocean, the rate of degradation and persistence of plastics varies by polymer, shape, density, and the purpose of the plastic itself. These characteristics also govern where in the water column plastics may be found. For example, more buoyant plastics are more likely to be carried by ocean currents and wind across the environment.</p>	<p>Cuando los plásticos entran al océano, la tasa de degradación y persistencia de los plásticos varía según el polímero, la forma, la densidad y el propósito del plástico en sí. Estas características también determinan en qué parte de la columna de agua se encuentran los plásticos. Por ejemplo, es más probable que los plásticos más flotantes sean transportados por las corrientes oceánicas y el viento a través del medio ambiente.</p>
<p>Additionally, when plastics are exposed to natural forces like sunlight and wave action, plastics will degrade into microplastics—defined as plastic particles under 5 mm in size. This definition commonly includes plastic pieces in the nano-scale, < 1 µm in size.</p>	<p>Además, cuando los plásticos están expuestos a fuerzas naturales como la luz solar y la acción de las olas, los plásticos se degradan en microplásticos, definidos como partículas plásticas de menos de 5 mm de tamaño. Esta definición incluye comúnmente las piezas de plástico en la escala nanométrica, < 1 µm en tamaño.</p>

Comentado [MA11]: Muy INformal, registro inadecuado ID

Comentado [MA12]: Terminología inadecuada: grado LEX

Comentado [MA13]: Mala terminología LEX

Comentado [MA14]: Discrepancia semántica, no mismo sentido NMS

Comentado [MA15]: Terminología inadecuada LEX

Comentado [MA16]: Error de terminología, no se usa esta palabra, fragmentos LEX

Comentado [MA17]: Anacoluto por régimen preposicional SINT

<p>The extent of plastic degradation depends on factors including polymer type, age, and environmental conditions like weathering, temperature, irradiation, and pH. Over time, plastic particles contaminate the marine ecosystem and the food chain, including foodstuffs intended for human consumption. In vivo studies have demonstrated that nanoplastics can translocate to all organs. Evidence is evolving regarding relationships between micro- and nanoplastic exposure, toxicology, and human health.</p>	<p>El grado de degradación del plástico depende de factores como el tipo de polímero, la edad y las condiciones ambientales como la intemperie, la temperatura, la irradiación y el pH. Con el tiempo, las partículas de plástico contaminan el ecosistema marino y la cadena alimentaria, incluidos los alimentos destinados al consumo humano. Estudios in vivo han demostrado que los nanoplasticos pueden translocarse a todos los órganos. La evidencia está evolucionando en cuanto a las relaciones entre la exposición a micro y nanoplasticos, la toxicología y la salud humana.</p>
<p>Nutritional authorities advise Americans to double their seafood consumption; however, awareness or concerns about microplastics in seafood could lead consumers to reduce their consumption. Research to understand and reduce human health risks is critical in order to simultaneously protect consumers and support their nutritional health.</p>	<p>Las autoridades nutricionales aconsejan a los estadounidenses que dupliquen su consumo de mariscos; sin embargo, la conciencia o las preocupaciones sobre los microplásticos en los mariscos podrían llevar a los consumidores a reducir su consumo. La investigación para comprender y reducir los riesgos para la salud humana es fundamental para proteger al mismo tiempo</p>

Comentado [MA18]: Falso sentido: erosión FS

Comentado [MA19]: Error ortotipográfico, faltan cursivas OTP

Comentado [MA20]: Término no adecuado, esta palabra no existe en español LEX

Comentado [MA21]: Error de término, calco ID

Comentado [MA22]: Esta palabra no engloba toda la comida proveniente del mar, no es la adecuada LEX

Comentado [MA23]: LEX

	a los consumidores y apoyar su salud nutricional.
--	--

ERRORES	
SIN SENTIDO (SS)	1
FALSO SENTIDO (FS)	1
NO MISMO SENTIDO (NMS)	1
ANACOLUTO (SINT)	1
ORTOGRAFÍA (ORT)	—
ORTOTIPOGRAFÍA (OTP)	1
FORMATO (FOR)	1
TERMINOLOGÍA (LEX)	8
IDIOMATICIDAD (ID)	7
OMISIÓN (OM)	—
ADICIÓN (AD)	1
TOTAL	22

TEXTO ORIGINAL	TRADUCCIÓN YANDEX
<p>Since the 1960s, plastic production has increased by approximately 8.7% annually, evolving into a \$600 billion global industry. Approximately eight million metric tons of plastics enter the oceans annually, and conservative estimates suggest 5.25 trillion plastic particles currently circulate in ocean surface waters. While some plastics enter oceans from maritime operations, 80% is suspected to originate from land-based sources.</p>	<p>Desde la década de 1960, la producción de plástico ha aumentado aproximadamente un 8,7% anual, evolucionando en una industria global de \$600 mil millones. Aproximadamente ocho millones de toneladas métricas de plásticos entran en los océanos anualmente, y estimaciones conservadoras sugieren 5.25 partículas de plástico de conejo actualmente circulan en las aguas superficiales del océano. Mientras que algunos plásticos entran en los océanos a partir de operaciones marítimas, se sospecha que el 80% proceden de Fuentes terrestres.</p>
<p>Discarded plastic materials enter the marine environment as trash, industrial discharge, or litter through inland waterways, wastewater outflows, and transport by winds or tides. Waste generation and waste leakage are inextricably linked and proportionally associated with economic development,</p>	<p>Los materiales plásticos desechados entran en el ambiente marino como basura, literalmente figuran, o basura a través de las vías actualizadables interiores, las corrientes de aguas residuales, y el transporte por los vientos o la libertad. La generación y las fugas de residuos están inextricablemente Roger y</p>

Comentado [MA24]: Calco en la estructura ID

Comentado [MA25]: Calco ID

Comentado [MA26]: Anacoluto SINT

Comentado [MA27]: Estilo incorrecto FOR

Comentado [MA28]: Adición, no es necesario. AD

Comentado [MA29]: Termino no adecuado ID

Comentado [MA30]: Falta preposición SINT

Comentado [MA31]: Falta de ortotipografía OTP

Comentado [MA32]: Omisión OM

Comentado [MA33]: Error sin sentido SS

Comentado [MA34]: T. inadecuado LEX

Comentado [MA35]: Error de formato FOR

Comentado [MA36]: Terminología inadecuada LEX

Comentado [MA37]: Error sin sentido SS

Comentado [MA38]: Repetición del término LEX

Comentado [MA39]: Palabra inventada, error SINT

Comentado [MA40]: Registro inadecuado ID

Comentado [MA41]: Error en la palabra elegida LEX

Comentado [MA42]: Errata, sin sentido SS
Además de omisión OM

<p>local infrastructure, and legislation.</p> <p>Today, uncollected waste accounts for 75% of these land-based discharges, while the remaining 25% comes from within the waste management system.</p>	<p>proporcionalmente asociadas con el desarrollo económico, la infraestructura local y la legislación. Hoy en día, los residuos no recogidos representan el 75% de estos vertidos en tierra, mientras que el 25% restante proviene del sistema de gestión de residuos.</p>
<p>When plastics enter the ocean, the rate of degradation and persistence of plastics varies by polymer, shape, density, and the purpose of the plastic itself. These characteristics also govern where in the water column plastics may be found. For example, more buoyant plastics are more likely to be carried by ocean currents and wind across the environment.</p>	<p>Cuando los plásticos entran en el océano, la velocidad de degradación y la persistencia de los plásticos varía según el polímero, la forma, la densidad y el propósito del propio plástico. Estas características también rigen donde en la columna de agua se pueden encontrar plásticos. Por ejemplo, es más probable que los plásticos más boyantes sean transportados por las corrientes oceánicas y el viento a través del medio ambiente.</p>
<p>Additionally, when plastics are exposed to natural forces like sunlight and wave action, plastics will degrade into microplastics—defined as plastic particles under 5 mm in size. This definition commonly includes plastic pieces in the nano-scale, < 1 µm in size.</p>	<p>Además, cuando los plásticos están expuestos a las fuerzas naturales como la luz solar y la acción de las olas, los plásticos se degradan en microplásticos, definidos como partículas de plástico de menos de 5 mm de Tamaño. Esta definición incluye comúnmente piezas de</p>

Comentado [MA43]: Registro inadecuado, demasiado informal ID

Comentado [MA44]: Error en el término LEX

Comentado [MA45]: Término incorrecto, LEX

Comentado [MA46]: Error de formato no mayúscula FOR

	plástico en la escala nano , de término l μm de Tamaño .
The extent of plastic degradation depends on factors including polymer type, age, and environmental conditions like weathering, temperature, irradiation, and pH. Over time, plastic particles contaminate the marine ecosystem and the food chain, including foodstuffs intended for human consumption. In vivo studies have demonstrated that nanoplastics can translocate to all organs. Evidence is evolving regarding relationships between micro- and nanoplastic exposure, toxicology, and human health.	La extensión de la degradación del plástico depende de factores como el tipo de polímero, la edad y las condiciones ambientales como la intemperie , la temperatura, la irradiación y el pH. Con el tiempo, las partículas plásticas acompañan al ecosistema marino y a la cadena alimentaria, incluidos los alimentos destinados al consumo humano. Los estudios In vivo han demostrado que las nanopartículas se pueden traslocar a todos los órganos. La evidencia está evolucionando con respecto a las relaciones entre la exposición a micro y nano plástico , la toxicología y la salud humana.
Nutritional authorities advise Americans to double their seafood consumption; however, awareness or concerns about microplastics in seafood could lead consumers to reduce their consumption. Research to understand and reduce human health risks is critical in order to	Las autoridades nutricionales aconsejan a los estadounidenses que dupliquen su consumo de mariscos ; sin embargo, la conciencia o las preocupaciones sobre los microplásticos en los mariscos podría llevar a los consumidores a reducir su consumo. La investigación para

Comentado [MA47]: Error de léxico LEX

Comentado [MA48]: Adición que es incorrecta AD

Comentado [MA49]: Omisión de *menos de* OM

Comentado [MA50]: Error de formato FOR

Comentado [MA51]: Error de término LEXI

Comentado [MA52]: Errata y léxico incorrecto SINT

Comentado [MA53]: T. erróneo, falso sentido FS

Comentado [MA54]: Sin mayúsculas y en cursiva FOR

Comentado [MA55]: Error de palabra FS

Comentado [MA56]: Error de término, calco LEX

Comentado [MA57]: Error de término, calco ID

Comentado [MA58]: Debería ser en plural ID

Comentado [MA59]: Término inadecuado LEX

Comentado [MA60]: LEX

simultaneously protect consumers and support their nutritional health.	comprender y reducir los riesgos para la salud humana es fundamental para declarar simultáneamente a los consumidores y apoyar su salud nutricional.
--	--

- Comentado [MA61]: Error LEX
- Comentado [MA62]: Error de colocación SINT

ERRORES	
SIN SENTIDO (SS)	3
FALSO SENTIDO (FS)	2
NO MISMO SENTIDO (NMS)	—
ANACOLUTO (SINT)	5
ORTOGRAFÍA (ORT)	—
ORTOTIPOGRAFÍA (OTP)	—
FORMATO (FOR)	5
TERMINOLOGÍA (LEX)	12
IDIOMATICIDAD (ID)	7
OMISIÓN (OM)	3
ADICIÓN (AD)	2
TOTAL	39

TEXTO ORIGINAL	TRADUCCIÓN APERTIUM
<p>Since the 1960s, plastic production has increased by approximately 8.7% annually, evolving into a \$600 billion global industry. Approximately eight million metric tons of plastics enter the oceans annually, and conservative estimates suggest 5.25 trillion plastic particles currently circulate in ocean surface waters. While some plastics enter oceans from maritime operations, 80% is suspected to originate from land-based sources.</p>	<p>Desde el 1960s, la producción plástica ha aumentado por aproximadamente 8.7% anualmente, evolucionando a un \$600 mil millones industria global. Aproximadamente ocho millones de toneladas métricas de plásticos introducen los océanos anualmente, y las estimaciones conservadoras sugieren 5.25 billones de partículas plásticas actualmente circulan en aguas de superficie del océano. Mientras algunos plásticos introducen océanos de operaciones marítimas, 80% está sospechado para originar de tierra-basó fuentes.</p>
<p>Discarded plastic materials enter the marine environment as trash, industrial discharge, or litter through inland waterways, wastewater outflows, and transport by winds or tides. Waste generation and waste leakage are inextricably linked and proportionally associated with economic development,</p>	<p>Discarded Los materiales plásticos introducen el entorno marino cuando basura, caudal industrial, o ensuciar a través de inland waterways, wastewater outflows, y transporte por vientos o mareas. Generación de residuos y escape de residuos son inextricably enlazado y proporcionalmente asociado con</p>

Comentado [MA63]: Anacoluto SINT

Comentado [MA64]: Error de formato FOR

Comentado [MA65]: Error de formato FOR

Comentado [MA66]: Sin sentido SS

Comentado [MA67]: Calco ID

Comentado [MA68]: Adición no necesaria AD

Comentado [MA69]: Error de término LEX

Comentado [MA70]: Sugieren que SINT

Comentado [MA71]: Formato FOR

Comentado [MA72]: Término incorrecto LEX

Comentado [MA73]: Error SS

Comentado [MA74]: Sin sentido, error SS

Comentado [MA75]: Error traducción SS

Comentado [MA76]: Término equivocado LEX

Comentado [MA77]: Término incorrecto LEX

Comentado [MA78]: Término incorrecto LEX

Comentado [MA79]: Error SS

Comentado [MA80]: Error SS

Comentado [MA81]: Falta reformulación ID

Comentado [MA82]: Falta artículo SINT

Comentado [MA83]: Repetición de palabra ID

Comentado [MA84]: Error SS

<p>local infrastructure, and legislation.</p> <p>Today, uncollected waste accounts for 75% of these land-based discharges, while the remaining 25% comes from within the waste management system.</p>	<p>desarrollo económico, infraestructura local, y legislación. Hoy, uncollected cuentas de residuos para 75% de esta tierra-basó caudales, mientras el restantes 25% proviene dentro del sistema de administración de los residuos.</p>
<p>When plastics enter the ocean, the rate of degradation and persistence of plastics varies by polymer, shape, density, and the purpose of the plastic itself. These characteristics also govern where in the water column plastics may be found. For example, more buoyant plastics are more likely to be carried by ocean currents and wind across the environment.</p>	<p>Cuándo los plásticos introducen el océano, el índice de la degradación y la persistencia de plásticos varía por polímero, forma, densidad, y el propósito del plástico él. Estas características también gobiernan donde en los plásticos de columna del agua pueden ser encontrados. Por ejemplo, más buoyant los plásticos son más probablemente para ser llevados por corrientes de océano y viento a través del entorno.</p>
<p>Additionally, when plastics are exposed to natural forces like sunlight and wave action, plastics will degrade into microplastics—defined as plastic particles under 5 mm in size. This definition commonly includes plastic pieces in the nano-scale, < 1 µm in size.</p>	<p>Además, cuándo los plásticos están expuestos a a fuerzas naturales les gusta el sol y acción ondulatoria, los plásticos degradarán a microplastics—definió partículas tan plásticas debajo 5 mm en medida. Esta definición generalmente incluye piezas plásticas en el nano-escala, < 1 µm en medida.</p>

Comentado [MA85]: Faltan artículos SINT

Comentado [MA86]: Error SS

Comentado [MA87]: Error SS

Comentado [MA88]: Error SS

Comentado [MA89]: Falta preposición SINT

Comentado [MA90]: Error de concordancia SINT

Comentado [MA91]: Falta artículo SINT

Comentado [MA92]: Léxico no adecuado LEX

Comentado [MA93]: Sin acento ORT

Comentado [MA94]: Término incorrecto LEX

Comentado [MA95]: Falta artículo SINT

Comentado [MA96]: Falta artículo SINT

Comentado [MA97]: Error SS

Comentado [MA98]: Término incorrecto LEX

Comentado [MA99]: Error de formulación SS

Comentado [MA100]: Error de término SS

Comentado [MA101]: Error de formulación SS

Comentado [MA102]: Repetición AD

Comentado [MA103]: Error SS

Comentado [MA104]: Falta la preposición SINT

Comentado [MA105]: Error ortografía SINT

Comentado [MA106]: Error del tiempo verbal SINT

Comentado [MA107]: Form. incorrecta y errores SS

Comentado [MA108]: Error de concordancia SINT

Comentado [MA109]: Error de preposición SINT

<p>The extent of plastic degradation depends on factors including polymer type, age, and environmental conditions like weathering, temperature, irradiation, and pH. Over time, plastic particles contaminate the marine ecosystem and the food chain, including foodstuffs intended for human consumption. In vivo studies have demonstrated that nanoplastics can translocate to all organs. Evidence is evolving regarding relationships between micro- and nanoplastic exposure, toxicology, and human health.</p>	<p>La extensión de degradación plástica depende de factores incluyendo tipo de polímero, edad, y condiciones medioambientales como weathering, temperatura, irradiación, y pH. Con el tiempo, las partículas plásticas contaminan el ecosistema marino y la cadena alimentaria, incluyendo foodstuffs pretendido para consumo humano. En vivo los estudios han demostrado que nanoplastics puede translocate a todos los órganos. La evidencia está evolucionando con respecto a relaciones entre micro- y nanoplastic exposición, toxicología, y salud humana.</p>
<p>Nutritional authorities advise Americans to double their seafood consumption; however, awareness or concerns about microplastics in seafood could lead consumers to reduce their consumption. Research to understand and reduce human health risks is critical in order to simultaneously protect consumers and support their nutritional health.</p>	<p>Las autoridades nutritivas aconsejan americanos para plegar su seafood consumo; aun así, concienciación o preocupaciones sobre microplastics en seafood podría dirigir consumidores para reducir su consumo. Búsqueda para entender y reducir riesgos de salud humana es críticos para simultáneamente proteger consumidores y apoyar su salud nutritiva.</p>

- Comentado [MA110]:** Calco ID
- Comentado [MA111]:** Faltan artículos SINT
- Comentado [MA112]:** Falta traducción SS
- Comentado [MA113]:** Faltan artículos SINT
- Comentado [MA114]:** Falta traducción SS
- Comentado [MA115]:** Término incorrecto LEX
- Comentado [MA116]:** Traducción incorrecta LEX
- Comentado [MA117]:** Falta de traducción SS
- Comentado [MA118]:** Falta de traducción SS
- Comentado [MA119]:** Doble espaciación OTP
- Comentado [MA120]:** Calco ID
- Comentado [MA121]:** Falta de traducción SS
- Comentado [MA122]:** Formulación incorrecta SS
- Comentado [MA123]:** Faltan artículos SINT
- Comentado [MA124]:** Error de término ID
- Comentado [MA125]:** Error SS
- Comentado [MA126]:** Falta de traducción SS
- Comentado [MA127]:** Faltan artículos SINT
- Comentado [MA128]:** Falta traducción SS
- Comentado [MA129]:** Falta traducción SS
- Comentado [MA130]:** Traducción incorrecta NMS
- Comentado [MA131]:** Error de artículo SINT
- Comentado [MA132]:** Artículo SINT y t. inadecuado LEX
- Comentado [MA133]:** Falta artículo SINT
- Comentado [MA134]:** Error de concordancia SINT
- Comentado [MA135]:** Falta preposición SINT
- Comentado [MA136]:** Falta artículo SINT
- Comentado [MA137]:** Error de palabra ID

ERRORES	
SIN SENTIDO (SS)	26
FALSO SENTIDO (FS)	—
NO MISMO SENTIDO (NMS)	1
ANACOLUTO (SINT)	24
ORTOGRAFÍA (ORT)	1
ORTOTIPOGRAFÍA (OTP)	1
FORMATO (FOR)	3
TERMINOLOGÍA (LEX)	10
IDIOMATICIDAD (ID)	7
OMISIÓN (OM)	—
ADICIÓN (AD)	2
TOTAL	75