

## Informe final:

### Hacia un uso didáctico del traductor automático DeepL.

Código proyecto: ID2018/200

## 1 Observaciones generales

La acción de este proyecto se adscribe a la de Innovación de metodologías docentes para las clases teóricas y prácticas. Participaban en él cuatro miembros del Departamento de Traducción e Interpretación y un miembro del Departamento de Informática. Para la composición del equipo era importante contar con especialistas en diferentes ámbitos.

**Prof. Emilio Rodríguez Vázquez de Aldana:** estudioso de la lingüística computacional, imparte las diferentes asignaturas vinculadas con las tecnologías de la traducción en el Grado en Traducción e Interpretación de la Universidad de Salamanca.

**Prof. Jesús Torres del Rey:** reconocido experto en formación en tecnología de la traducción, accesibilidad de documentos digitales y posesición. Imparte asignaturas de este ámbito en el mismo Grado (combinación lingüística: inglés-español // francés-español), en el Doble Grado de Traducción y Derecho y en el Máster de Traducción y Mediación lingüística.

**Prof<sup>a</sup>. Goedele de Sterck:** reconocida experta en traducción directa e inversa, traductora profesional (combinación lingüística: francés-español // neerlandés-español, inglés-español). Imparte clases en el Grado de Traducción e Interpretación, en el Doble Grado de Traducción y Derecho y en el Máster de Traducción y Mediación lingüística.

**Prof<sup>a</sup>. Petra Zimmermann González:** reconocida experta en didáctica y en traducción inversa (español-alemán), traductora profesional. Imparte clases en el Grado de Traducción e Interpretación y en el Doble Grado de Traducción y Derecho y en el Máster de Traducción de la Universidad de Granada.

**Prof<sup>a</sup>. Silvia Roiss** (directora del proyecto) y reconocida experta en didáctica, traducción inversa (español-alemán) y traducción jurídica (alemán-español//español-alemán). Imparte clases en el Grado de Traducción e Interpretación, en el Doble Grado de Traducción y Derecho así como en el Máster de Traducción y Mediación lingüística, el Máster de Traducción de la Universidad de Granada y en el Doble Máster de Traducción con la Universidad de Heidelberg, Alemania.

## 2 Consecución de los objetivos

Para este PID nos hemos propuesto cinco objetivos concretos que vamos a enumerar a continuación:

1. Fomentar la capacidad crítica del estudiante a la hora de trabajar con el traductor automático DeepL;
2. Elaborar un extenso corpus multilingüe basado en clases textuales para el posterior análisis de errores cometidos por el programa DeepL;
3. Investigar los mecanismos y la potencialidad que el programa DeepL ofrece;
4. Elaborar material didáctico para el alumno para la comprensión de la traducción automática neuronal;
5. Recopilar posibles técnicas de posesición de traducción automática.

### 2.1 Fomentar la capacidad crítica del estudiante a la hora de trabajar con el traductor automático (TA) DeepL:

En los últimos años venimos constatando un uso casi indiscriminado del traductor automático DeepL por parte de nuestros estudiantes a la hora de afrontar su trabajo de traducción. El discente, sobre todo al inicio de sus estudios de Grado en Traducción e Interpretación cae fácilmente en la tentación de considerar válida una serie de incorrecciones provocadas por esta herramienta. De esta experiencia acumulada en nuestra actividad docente en relación con el programa y de las lecturas acerca de la traducción automática y de la posesición, pensamos que alertar al alumno sobre las ventajas y desventajas acerca de DeepL podría redundar en una serie de beneficios para el mismo, en consonancia con los principios metodológicos empleados en las asignaturas de traducción en nuestro Grado. De esta manera, fomentamos la capacidad del análisis crítico y, en consecuencia, el aprendizaje activo. Otra razón de peso para trabajar con DeepL de manera crítica es fomentar la motivación del alumno e implicarle en su propio proceso de aprendizaje, puesto que debe tomar decisiones sobre aspectos relevantes en relación con el uso del programa como, por ejemplo, tener en cuenta la clase textual, discernir los errores de distinta procedencia, efectuar un análisis textual multinivel etc.

Para poder evaluar de manera eficaz si un texto o su traducción cumplen los requisitos de calidad, ha sido necesario acordar previamente determinados criterios que se tenían que tener en cuenta para el análisis. Nos hemos guiado por la clasificación que ofrece el modelo de evaluación de la calidad MQM (sistema métrico de calidad multidimensional) del Centro Alemán de Investigación en Inteligencia Artificial (véase <http://www.qt21.eu/mqm-definition/definition-2015-12-30.html#default-values>) y lo hemos adaptado a nuestro propósito. Los errores aquí clasificados recogen los tipos de falta que se pueden detectar en las traducciones originadas por el TA.

#### MODELO DE EVALUACIÓN

##### 1. TRA (TRANSFERENCIA o PRECISIÓN)

Errores producidos en el proceso de traducción o por falta de correspondencia en sentido o función entre la traducción y el original.

###### 1.1. Adición inapropiada

- 1.2. Omisión inapropiada
- 1.3. No traducción: cuando se ha mantenido, de forma inapropiada, el fragmento en la lengua original.
- 1.4. Traducción incorrecta:
  - 1.4.1. A nivel léxico (mala selección léxica)
  - 1.4.2. Por literalidad léxica
  - 1.4.3. Por literalidad estructural
  - 1.4.4. Por presencia incorrecta de elementos numéricos, técnicos, nombres propios, etc.
  - 1.4.5. Por mala interpretación o confusión de elementos o sistemas técnicos o conceptuales
  - 1.4.6. Que ha introducido ambigüedad estructural o léxica

## 2. FLU (FLUIDEZ o INCORRECCIÓN LINGÜÍSTICA / NORMATIVA o DE TIPO MECÁNICO -

Es **independiente del proceso de traducción**, lo que significa que es aplicable a textos no traducidos. En nuestro caso, vendría acompañado de un error de transferencia si el error no estaba ya en el original.

- 2.1. Incorrección gramatical:
  - 2.1.1. Por mala elección de elementos gramaticales/funcionales (artículos, preposiciones, conjunciones, verbos auxiliares...)
  - 2.1.2. Por incorrección morfosintáctica (género, número, categoría gramatical, tiempo, modo, aspecto...)
  - 2.1.3. Por orden incorrecto de palabras (que vaya contra la norma, no tanto por idiomática o estilo).
- 2.2. Ortografía: mayúsculas, tildes, otros diacríticos...
- 2.3. Tipografía: símbolos de puntuación, espacios en blanco ...
- 2.4. Inconsistencia: no hay coherencia en el uso de un mismo significante o conjunto de normas para un referente único: distintas formas de abreviatura para el mismo elemento; distinta forma de nombrar un elemento que se muestra en una imagen de referencia; se usa una norma bibliográfica distinta de la que se ha indicado...
- 2.5. Coherencia: hay problemas en la lógica del texto a nivel interoracional o del texto en su conjunto.
- 2.6. Cohesión: no se emplean correctamente conectores gramaticales o lógicos, referentes textuales (anáfora, catáfora, elipsis, deixis...). Es un error a nivel intraoracional que puede causar inconsistencia o incoherencia a nivel interoracional.
- 2.7. Repetición incorrecta: una palabra o parte de frase duplicada.
- 2.8. Registro gramatical: registro incorrecto en el uso de algún elemento gramatical específico (verbo, pronombre, etc.)
- 2.9. Ambigüedad no introducida en el proceso de traducción (puede ser por referente poco claro).
- 2.10. Ininteligible: oración ininteligible por razones distintas a las anteriores.

## 3. EST (ESTILO)

No suponen errores claros mecánicos o de la norma, pero son textos estilísticamente mejorables, o que no siguen determinada guía de estilo.

- 3.1. No idiomático: uso de elementos de la lengua de forma poco natural, malas colocaciones.
  - 3.2. Enrevesado: en su formulación gramatical/textual, lo que dificulta la lectura.
  - 3.3. No seguimiento de estilo especificado
  - 3.4. Estilo inconsistente
  - 3.5. Registro: problema de registro de tipo léxico, no gramatical.
- 4. TERM (TERMINOLOGÍA)**
- 4.1. No sigue base de datos específica.
  - 4.2. No sigue la terminología del campo.
  - 4.3. Uso inconsistente de terminología.
- 5. REFS (REFERENCIAS A LA CULTURA Y LA REALIDAD DE DESTINO)**
- 5.1. Referencia cultural inapropiada.
  - 5.2. Referencia al contexto de destino inapropiada: p.ej. en la traducción se hace referencia a elementos que se ofrecen en el contexto de partida, pero no en el de destino.
  - 5.3. Incompleto de partida: se han omitido pasos o elementos de una lista en el texto de origen.
  - 5.4. Inapropiado para el usuario de destino
  - 5.5. Requisitos legales de destino no seguidos
- 6. LOC (CONVENCIONES MATERIALES DEL CONTEXTO O CULTURA DE DESTINO)**: formatos de direcciones, teléfonos, calendarios, puntuación, formatos numéricos, etc.
- 7. DIS (FORMATO, DISEÑO MATERIAL)**: gráficos, guionización, solapamientos, etc..

## 2.2 Elaborar un extenso corpus multilingüe basado en clases textuales para el posterior análisis de los desajustes provocados por el programa DeepL:

A este respecto, cada profesor ha elaborado su propio corpus dependiendo de la combinación lingüística en la que es experto. A continuación, se inserta la tabla Excel en la que se recogen los criterios que se tomaban en cuenta para su selección y que consta de los siguientes elementos: 1) combinación lingüística 2) clase textual, 3) grado de especificidad 4) número de secuencias analizado.

investigador	combinación lingüística	clase textual	grado de especialidad	número de secuencias
Zimmermann	ES-DE	Condiciones generales de compraventa	jurídico especializado	12
Zimmermann	ES-DE	Condiciones generales de compraventa	jurídico especializado	27
Zimmermann	ES-DE	Catálogo de preguntas y respuestas	semi-divulgativo	16
Zimmermann	ES-DE	Artículo de revista odontológica	semi-divulgativo	33

Zimmermann	ES-DE	Receta de cocina	divulgativo	27
Roiss	DE-ES	Apostilla	jurídico especializado	18
Roiss	DE-ES	Acreditación de tutela	jurídico especializado	11
Roiss	DE-ES	Contrato de servicio (extracto)	jurídico especializado	13
Roiss	DE-ES	Certificado de herencia	jurídico especializado	21
Roiss	DE-ES	Contrato de arrendamiento (extracto)	jurídico especializado	29
Roiss	ES-DE	Contrato de arrendamiento (extracto)	jurídico especializado	40
Roiss	ES-DE	Escritura de donación	jurídico especializado	34
Torres del Rey	EN-ES	Artículo periodístico	divulgativo	37
Torres del Rey	EN-ES	Informe sobre política sanitaria y su comunicación	científico-técnico divulgativo	34
De Sterck	FR-ES	Definición de diabetes	científico-técnico divulgativo	11
De Sterck	FR-ES	Definición de diabetes	científico-técnico especializado	12
De Sterck	FR-ES	Folleto publicitario medidor de glucosa	científico-técnico divulgativo	37
De Sterck	FR-ES	Manual usuario medidor de glucosa	científico-técnico divulgativo	13
De Sterck	FR-ES	Prospecto diabetes	científico-técnico divulgativo	18
De Sterck	NL-ES	Definición de diabetes	científico-técnico divulgativo	14
De Sterck	NL-ES	Definición de diabetes	científico-técnico especializado	11
De Sterck	NL-ES	Folleto publicitario medidor de glucosa	científico-técnico divulgativo	33
De Sterck	NL-ES	Manual usuario medidor de glucosa	científico-técnico divulgativo	18
De Sterck	NL-ES	Prospecto diabetes	científico-técnico divulgativo	19

El Prof. Rodríguez de Aldana ha facilitado una metodología sencilla con alguna herramienta accesible que permitía alinear de manera rápida y anotable un texto (Word) y su traducción automática con DeepL, de modo que se podía establecer un diagnóstico sobre los errores que comete el programa teniendo en cuenta los siguientes variables: clase textual (Doc-Id) / Secuencias (S-T) / Direccionalidad / Texto original / Traducción de DeepL / Tipo de error / Gravedad del error / Comentarios.

En total, se han estudiado 24 textos de extensión variada, pertenecientes a diversos ámbitos, mayoritariamente científico-técnico y jurídico. El rango de especificidad cubre tanto textos divulgativos y semi-divulgativos como especializados. Las combinaciones lingüísticas estudiadas han sido: español-alemán (Zimmermann, Roiss) / alemán-español (Roiss) / inglés-español (Torres del Rey) / francés-español (de Sterck) / neerlandés-español (de Sterck).

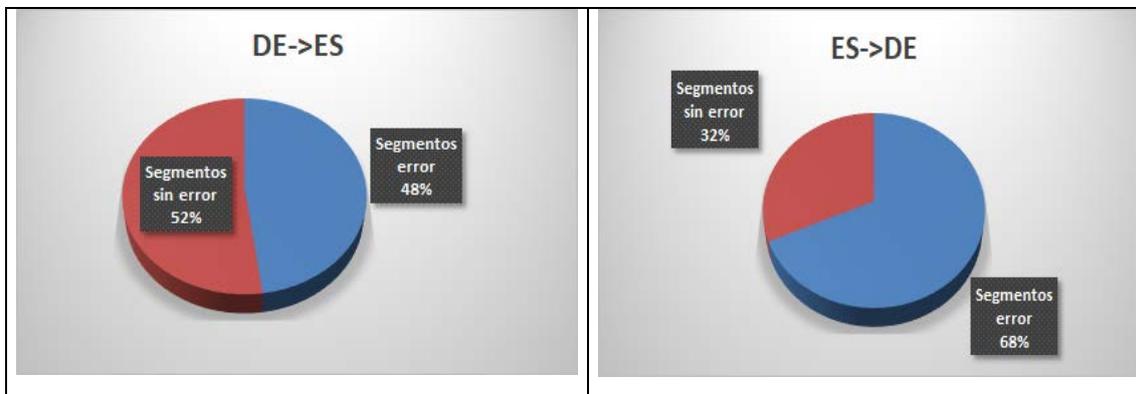
La base de datos que hemos elaborado en las distintas combinaciones lingüísticas, y a la que se hará referencia a continuación, sirve para ilustrar al alumno en las diferentes asignaturas de

Traducción, que ofrece el Plan de estudios, muestras representativas de errores en versión PowerPoint. Se presentan fragmentos de textos originales con la respectiva traducción efectuada por DeepL. Los errores están marcados previamente. El alumno ha de sistematizar los errores según la tabla MQM. El posterior debate sirve para tomar conciencia de las limitaciones que tiene el TA a la hora del trasvase.

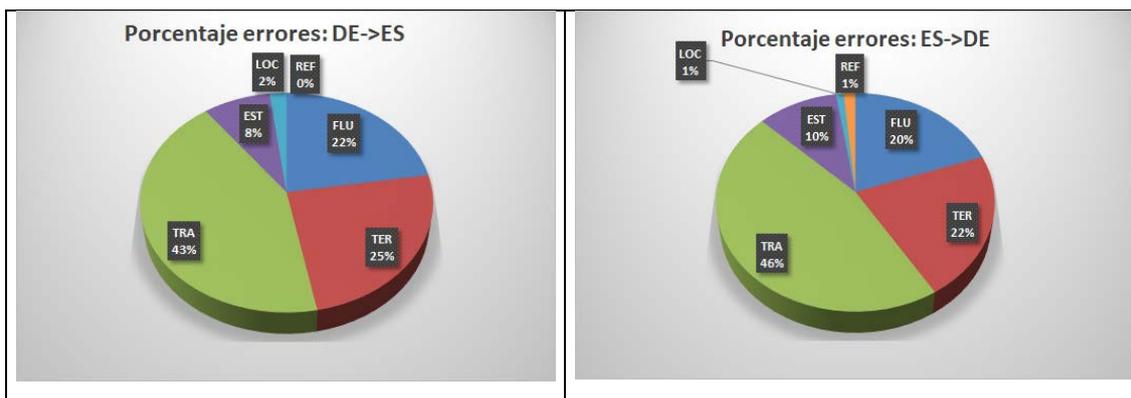
Para que el alumno tome conciencia de la repercusión que puede tener un uso irresponsable del TA, es imprescindible mostrar las estadísticas que hemos podido elaborar mediante el análisis minucioso de ciertas variables. A continuación, insertamos en este informe los gráficos referidos a las diferentes combinaciones lingüísticas, comentados.

### 2.2.1 Alemán-español y español-alemán

En la combinación lingüística español (ES)/alemán (DE) se ha analizado un total de 12 textos originales (5 en lengua alemana y 7 en lengua española). Los 5 textos redactados en lengua alemana pertenecen todos al ámbito notarial/administrativo con un grado de especialidad alto, jurídico. Los 7 textos en lengua española pertenecen a campos temáticos y clases textuales diversos. Queremos destacar la marcada diferencia proporcional de errores en ambos gráficos. Consideramos que un factor determinante en este resultado reside en el ámbito temático único del conjunto de los textos originales analizados en lengua alemana. Parece haber indicios de que la calidad de traducción con DeepL en la dirección DE-ES y el ámbito jurídico-administrativo es más alta. Este resultado sugiere futuras investigaciones en el mismo ámbito, en la combinación ES-DE.



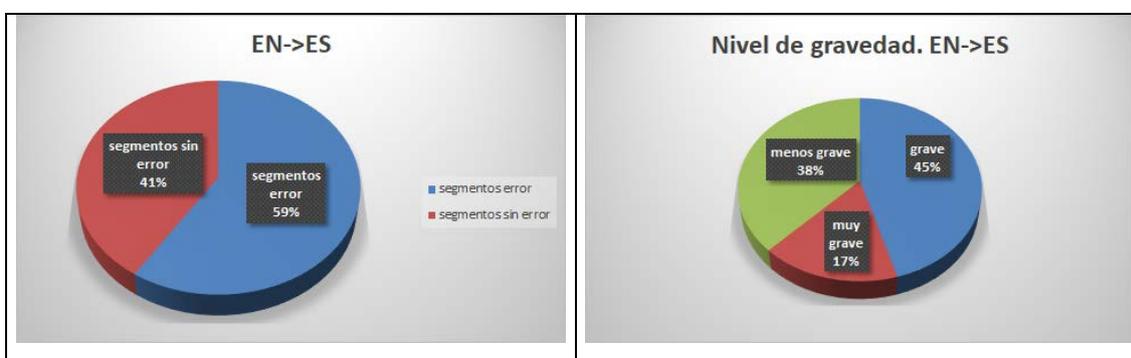
En cuanto a la proporcionalidad de los tipos de errores en las dos direcciones destaca la similitud de los porcentajes. Este hecho parece indicar que los tipos de errores no están ligados ni a clases textuales ni a ámbitos temáticos concretos ni a la dirección dentro de la combinación lingüística aquí tratada. Llama la atención el alto porcentaje de errores producidos por falta de correspondencia en sentido o función entre la traducción y el original (TRA). En un porcentaje menor siguen los tipos de error en cuanto a terminología (TER) y fluidez (FLU). Los errores restantes suman solo un 10% y 12%, respectivamente.



## 2.2.2 Inglés-español y español-inglés

En cuanto a la combinación lingüística inglés-español, se han escogido dos textos: uno de carácter divulgativo-periodístico, y otro semi-divulgativo, procedente del ámbito científico (políticas de salud pública). En ambos casos, se trata de documentos con una alta probabilidad de ser traducidos. Estas tipologías constan de un amplio corpus *online* bilingüe en dicha combinación de idiomas, si bien desconocemos si el motor DeepL los ha incorporado específicamente.

Una primera lectura de la traducción de DeepL en ambos textos da la sensación de un alto grado de fluidez y de logros de transferencia que serían complicados (o que no se conseguirían fácilmente) salvo en el caso de profesionales muy cualificados. No obstante, el análisis más minucioso de los resultados muestra, como se destaca en los gráficos, que los errores son numerosos, y que no pocos son de tipo grave y muy grave:

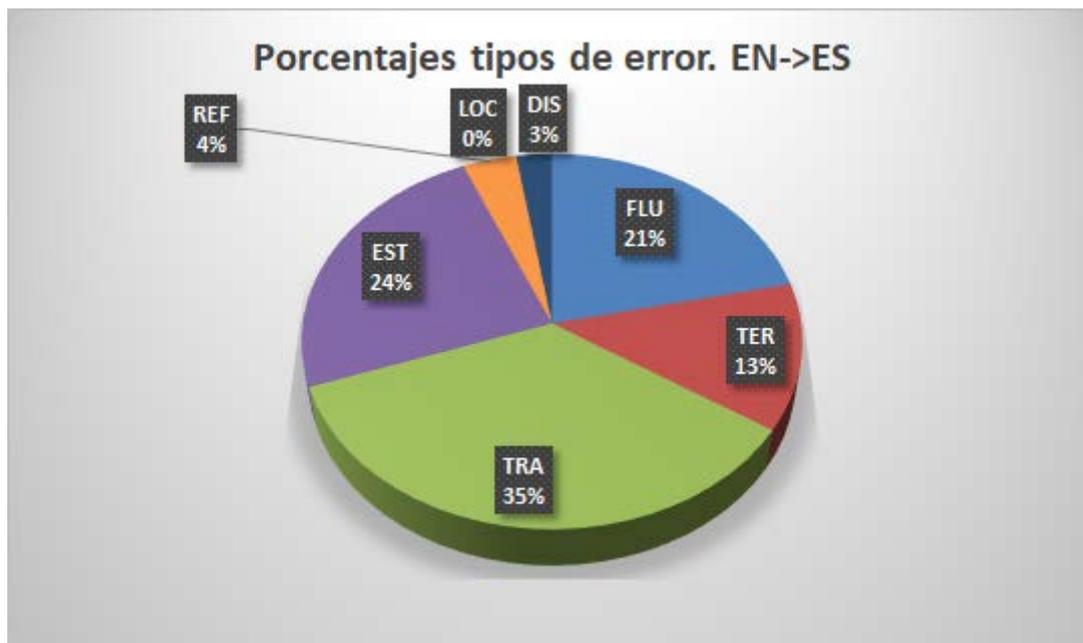


Por otro lado, se puede establecer una correlación entre la longitud del segmento y el porcentaje de errores. Si bien es cierto que, probabilísticamente, un segmento más largo tiene más posibilidades de contener algún error, también se deriva de nuestro análisis que hay una evidente relación directa entre mayor longitud y número de errores:

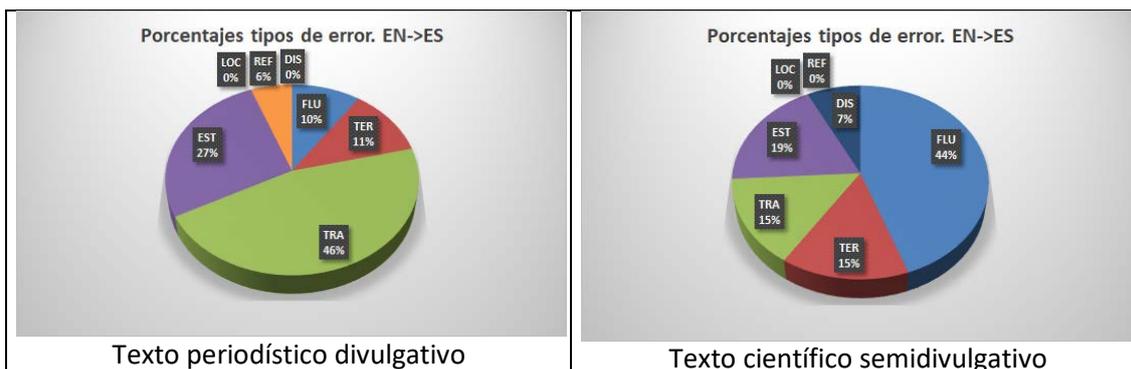
Por otro lado, se puede establecer una correlación entre la longitud del segmento y el porcentaje de errores. Si bien es cierto que, probabilísticamente, un segmento más largo tiene más posibilidades de contener algún error, también se deriva de nuestro análisis que hay una evidente relación directa entre mayor longitud y número de errores:



En cuanto al tipo de error preponderante, mostramos a continuación la media de ambos textos, que presenta una distribución uniforme entre los principales tipos de error, aunque destacan los errores de transferencia, como es habitual en la traducción automática neuronal:



No obstante, hay diferencias fundamentales entre el texto periodístico de tipo divulgativo y el texto científico (políticas de salud pública) de carácter semi-divulgativo:



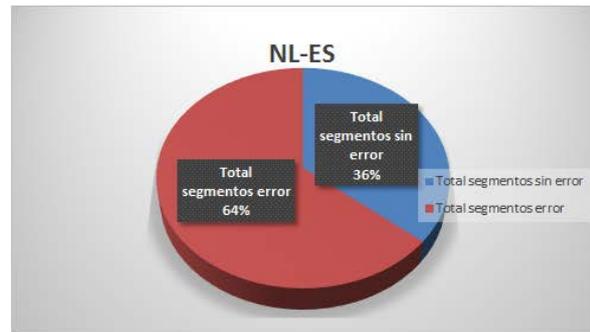
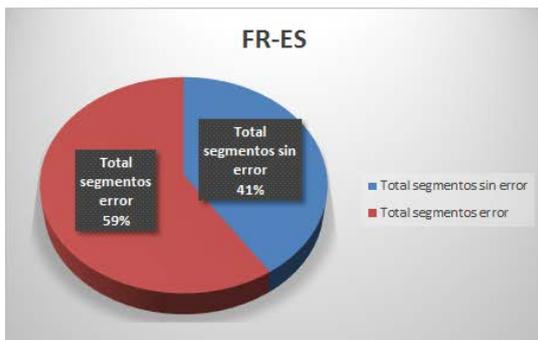
Vemos cómo el texto más cercano al lenguaje general pero, a la vez, con mayores figuras expresivas típicas del lenguaje periodístico, así como con mayor necesidad de conocimiento y contextualización con la realidad, presenta más errores de transferencia y, en segundo lugar, de estilo. Sin embargo, hay menos errores de fluidez, dado el reaprovechamiento de expresiones hechas y fórmulas periodísticas habituales que se pueden encontrar en los corpus generales. Por último, hay un porcentaje pequeño pero representativo de errores de referencia con la cultura y con el contexto actual.

Por el contrario, el documento más científico presenta una mayor abundancia de errores de fluidez y algo más de errores terminológicos. Esto puede deberse a que el propio texto original tiene una calidad lingüística inferior, y a la mayor presencia de términos específicos. La menor cantidad de errores de estilo puede deberse a un proceso de normalización del texto por parte de DeepL, aspecto que sería interesante investigar y poner en relación con los ya clásicos estudios de Laviosa y Baker al respecto. Por otro lado, desaparecen los errores referidos a la realidad y la cultura, y aparecen algunos errores relacionados con el formato y el diseño, debido a la mayor diversidad de formato del texto origen. Dichos errores provocan problemas de comprensión, por lo que deben tenerse en cuenta también.

### 2.2.3 Francés-español y neerlandés-español

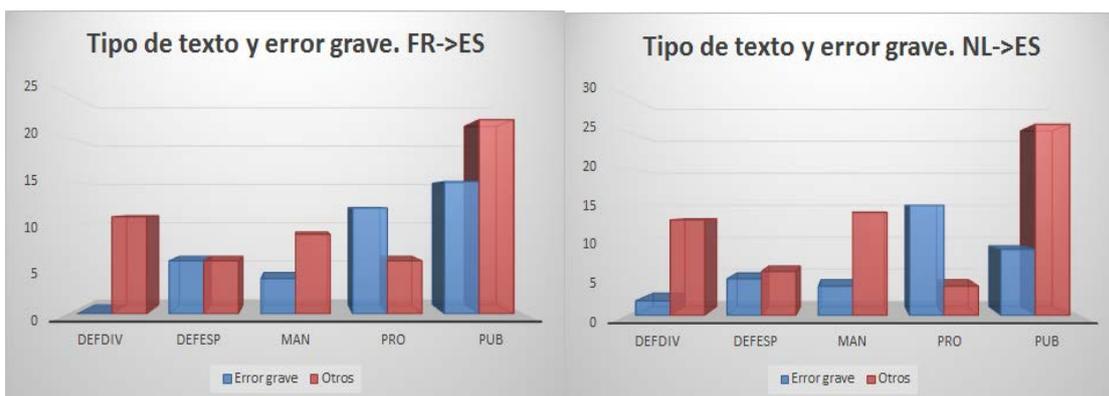
El estudio referente a las combinaciones lingüísticas francés-español y neerlandés-español se centra exclusivamente en textos de carácter científico-técnico. Con el fin de unificar al máximo la muestra en ambas lenguas se ha optado en cada caso por cinco fragmentos tomados de cinco tipos de textos diferentes, todos ellos relacionados con la diabetes. En concreto, se trata para cada idioma de una descripción especializada de la enfermedad, una descripción divulgativa de la enfermedad, un folleto publicitario con ficha técnica, un manual para el paciente y un prospecto.

Al interpretar los gráficos basados en el análisis de errores salta a la vista la gran coincidencia entre los resultados obtenidos para el francés y el neerlandés. En ambos casos, el número total de fragmentos con error ronda el 60 % frente al 40 % de fragmentos sin error.



Aunque no cabe duda de que se trata de un balance francamente positivo, conviene preparar a los alumnos para que sean capaces de detectar y enmendar ese 60 % de segmentos con errores, sobre todo si tenemos en cuenta que, en ambas muestras, los errores graves constituyen casi el 50 % del total. En un contexto didáctico habrá que prestar especial atención a los errores graves, ya que dan lugar a una traducción de todo punto inaceptable.

Tal y como se demuestra en los gráficos siguientes, la distribución de los errores graves varía según el tipo de texto.



Se pueden formular algunas conclusiones provisionales con independencia de la combinación lingüística:

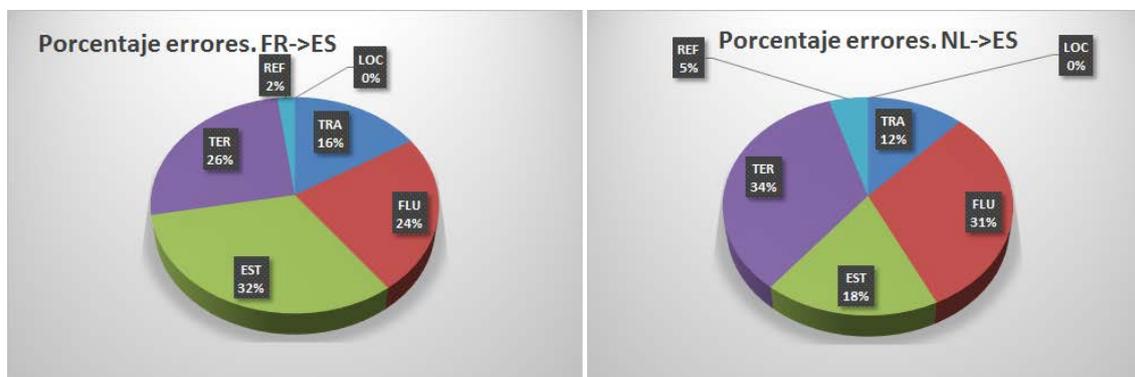
- DeepL comete poquísimos errores graves en un texto puramente informativo como puede ser la descripción divulgativa de una enfermedad;
- tanto en los fragmentos de los manuales como en los textos publicitarios con ficha técnica el número de segmentos con errores graves es relativamente bajo.

Todo ello parece indicar que, en estos casos, el uso de DeepL compensa, puesto que la versión inicial es más que aceptable y requiere una intervención mínima del traductor humano.

En cambio, los resultados de este pequeño estudio no parecen aconsejar el uso de DeepL para textos muy especializados y prospectos:

- en la descripción especializada de la diabetes el número de segmentos con errores graves ronda el 50 %;
- en el caso de los prospectos, el número de segmentos con errores graves adquiere proporciones alarmantes, entre otras razones porque el programa no “reflexiona” y no aplica las convenciones textuales y terminológicas recogidas en la plantilla oficial de la UE.

La clasificación por tipos de errores no arroja resultados demasiado significativos, por lo que habrá que ampliar el corpus y afinar los criterios de clasificación. Una de las posibles conclusiones es que la terminología constituye un problema en ambas combinaciones lingüísticas.



En francés predominan los errores estilísticos, en tanto que en neerlandés falla sobre todo la fluidez. Dado que el aprendizaje del traductor automático se basa en traducciones humanas, no hay que descartar que “calque” los vicios de estas. Es sabido que, bajo la influencia del texto fuente, el traductor de dos lenguas tan cercanas como el francés y el español tiende a dar por buenos giros y estructuras “posibles” pero quizá no del todo idiomáticas en la lengua meta (contaminación). Por otro lado, el número relativamente alto de incorrecciones gramaticales parece indicar que, en la combinación lingüística neerlandés-español, el programa “se nutre” en buena parte de traducciones inversas realizadas por traductores con un dominio no del todo satisfactorio de la lengua meta. Por todo lo anterior parecería altamente recomendable que los traductores automáticos aprendan no a base de traducciones paralelas sino de corpus comparables.

## 2.3 Investigar los mecanismos y la potencialidad que el programa DeepL ofrece:

Este objetivo se ha alcanzado mediante la elaboración de un decálogo que servirá para sensibilizar al alumno ante un uso responsable de la herramienta. En él se recogen las especificidades que hemos ido recogiendo mediante el análisis del comportamiento del TA. Queremos facilitar así una serie de advertencias en relación con el trabajo con DeepL que vamos a exponer a continuación.

### 2.3.1 Decálogo

1) Podemos decir que la inmediatez a la que se llega hoy en día a la información ha creado la **cultura del click** que condiciona las habilidades de pensamiento cotidianas. Como afirma Pozo (2016: 75) en relación con las tecnologías digitales: «[estas] se convierten en verdaderas prótesis mentales, de forma que extienden, modifican o reconstruyen nuestras capacidades de aprendizaje».

2) Aunque la propia compañía **DeepL** califica el traductor en su blog como «**el mejor del mundo**» (véase: <https://www.deepl.com/quality.html>) ha de tomarse esta afirmación con absoluto **escepticismo**.

3) La **gravedad de los errores** cometidos por el traductor automático varía de entre muy grave o crítico (que convierte el contenido en inutilizable o peligroso), grave (que afecta a la usabilidad o comprensibilidad, pero no hace el contenido inutilizable) y menos grave (que no afecta a la usabilidad o comprensibilidad).

4) La gravedad no está ligada a un solo **tipo de error**, sino a una amplia **variedad** de ellos: a) TRANSFERENCIA o PRECISIÓN; b) FLUIDEZ o INCORRECCIÓN LINGÜÍSTICA / NORMATIVA o DE TIPO MECÁNICO; c) ESTILO; d) TERMINOLOGÍA; e) REFERENCIAS A LA CULTURA Y LA REALIDAD DE DESTINO; f) CONVENCIONES MATERIALES DEL CONTEXTO O CULTURA DE DESTINO; g) DISEÑO y FORMATO.

5) Sobre la base de las muestras estudiadas no se puede establecer una relación fiable entre las variables textuales (clase textual, tema textual, grado de especialidad) y los tipos/gravedad de error. Sin embargo, el análisis de los resultados obtenidos para las combinaciones lingüísticas francés-español y neerlandés-español parece aportar alguna hipótesis interesante. Por tanto, conviene ampliar el corpus y seguir investigando.

6) La gran diversidad de tipos de errores dificulta el establecimiento de pautas para el uso crítico del TA para el traductor aprendiz. Un mero catálogo de errores resulta de poca utilidad. Por eso, incidimos en la necesidad de la aplicación de las destrezas de **análisis textual multinivel como requisito imprescindible para un uso provechoso del TA** del cual se derivan las siguientes acciones: elegir las estrategias de traducción requeridas según el encargo, el texto en su conjunto y el contexto, filtrar convenciones textuales en el TO y actuar en consecuencia, refinar las técnicas de documentación y comprobar el grado de corrección gramatical y de sentido del TO.

7) Dado que DeepL traduce oraciones y no textos, **el nivel discursivo requiere especial atención**, en particular la consistencia interoracional (coherencia y cohesión). La falta de consistencia está íntimamente ligada a errores como traducción incorrecta a nivel léxico, terminología inconsistente, registro inadecuado.

8) La brecha abierta entre la traducción automática y la traducción humana solo se supera gracias a una **competencia traductora fortalecida** por una amplia capacidad de análisis crítico.

9) El programa de TA **DeepL**, sin duda, es una **herramienta valiosa** para la actividad traductora. No obstante, no es y **nunca podrá ser el sustituto del trabajo realizado por la mente humana**.

10) Es preciso explicar a los alumnos que **DeepL, lejos de ser la panacea**, no es sino una de las herramientas disponibles. Hay que situarlo en el lugar que le corresponde, junto a las memorias de traducción y los programas para la creación y gestión de corpus, por nombrar algunos ejemplos. En este sentido interesa investigar cuáles son las ventajas y desventajas de usar una u otra herramienta (o una combinación de todas ellas) en cada contexto para así mejorar la calidad de las traducciones y aumentar la productividad del traductor.

## **2.4 Seleccionar material didáctico para el alumno para la comprensión de la traducción automática neuronal.**

Sin duda, la entrada en escena en 2017 del motor DeepL ha centrado la atención de toda la comunidad profesional de la traducción. Se le puede considerar, en estos momentos, como el motor de traducción de referencia. A partir de ahí, todas las compañías que ofrecían servicios de traducción automática a través de un motor propio (Microsoft, Systran, Baidu, Facebook, entre otros), comenzaron a poner en explotación su nueva versión neuronal. Los profundos y vertiginosos cambios han tenido, naturalmente, su reflejo en nuestra labor docente. Por un lado, en las asignaturas vinculadas estrictamente con la traducción en las diferentes combinaciones de lenguas que se imparten en la Facultad, pues una gran parte de los estudiantes utilizan un motor de traducción automática como versión inicial a partir de la cual completar nuestros ejercicios propuestos de traducción. Por otro lado, y este es el aspecto que en este apartado nos interesa, esta revolución en el paradigma de la traducción automática también ha significado una renovación profunda de los créditos dedicados a la traducción automática en la asignatura optativa “Postedición” (código 104644).

La asignatura optativa Postedición, de 3 créditos, se distribuye en dos partes: una dedicada a la traducción automática y otra dedicada a las técnicas de posesición de traducción automática. A cada bloque le decidamos 1,5 créditos que suponen, a su vez, 15 horas presenciales entre clases teóricas y prácticas. Nos centramos aquí en los créditos y horas dedicados a la traducción automática.

En dicha asignatura, en el curso académico 2017-2018, recién aparecido DeepL, ya introducimos en nuestros contenidos educativos una lección y unas lecturas básicas sobre traducción automática neuronal. En el curso a punto de finalizar, 2018-2019, ya la traducción automática neuronal ha pasado a ser el eje central de la asignatura, desplazando las horas dedicadas a la explicación y comprensión de la traducción automática estadística, que fue el bloque básico y fundamental sobre el que se construía la asignatura hasta el curso 2016-17.

Durante este año académico, como parte de la necesaria renovación docente ante los cambios acaecidos antes mencionados hemos procedido a la selección de material didáctico para la ayuda en la comprensión de la traducción automática neuronal. Esta selección es la que presentamos a continuación.

La selección de material didáctico que hemos considerado apropiada para el estudio y comprensión de la traducción automática neuronal por un estudiante del Grado en Traducción e Interpretación que presentamos a continuación, la dividimos en los bloques que nombramos, por un lado, como material de estudio y lectura y, por otro, como recursos para el aprendizaje a través la práctica. Sobra decir que todo el material presentado en la fecha de elaboración de este informe estaba disponible en línea.

### 2.4.1 Material de estudio: bibliografía y fuentes

- Casacuberta Nolla, F. y Peris Abril, A. (2017): “Traducción automática neuronal”. *Revista Tradumàtica. Tecnologies de la Traducció*, 15, 66-74. [En línea: [https://ddd.uab.cat/pub/tradumatica/tradumatica\\_a2017n15/tradumatica\\_a2017n15p66.pdf](https://ddd.uab.cat/pub/tradumatica/tradumatica_a2017n15/tradumatica_a2017n15p66.pdf)]
- Chaudhary, J.R. and Patel, A.C. (2018): Machine Translation Using Deep Learning: A Survey, *International Journal of Scientific Research in Science, Engineering and Technology*, 4 (2), 145-150. [En línea: [https://pdfs.semanticscholar.org/e888/85410f47f9fab48d9b5f3102e1650cbae728.pdf?\\_ga=2.210442582.2131314044.1561540196-1750898839.1544438044](https://pdfs.semanticscholar.org/e888/85410f47f9fab48d9b5f3102e1650cbae728.pdf?_ga=2.210442582.2131314044.1561540196-1750898839.1544438044)]
- Cho, K.(2015, May 27): Introduction to Neural Machine Translation with GPUs. [Mensaje en un Blog] NVIDIA Developer Blog. [Recuperado de: <https://devblogs.nvidia.com/introduction-neural-machine-translation-with-gpus/>]
- Forcada, M. L. (2017): “Making Sense of neural machine translation”, *Translation Spaces* 6:2, 291-309. [En línea: <https://www.dlsi.ua.es/~mlf/docum/forcada17j2.pdf>]
- López, R. (2014, 7 de mayo): ¿Qué es y como funciona “Deep Learning”? [Mensaje en un Blog]. [Recuperado de: <https://rubenlopezg.wordpress.com/2014/05/07/que-es-y-como-funciona-deep-learning/>]
- Torres, J. (2018): Deep Learning. Introducción práctica con Keras. Primera parte. [En línea: <https://torres.ai/deep-learning-inteligencia-artificial-keras/>]
- Sato, K. (July 26, 2016): Understanding neural networks with TensorFlow Playground. [Mensaje en un Blog]. Google Cloud: AI & Machine Learning. [Recuperado de: <https://cloud.google.com/blog/products/gcp/understanding-neural-networks-with-tensorflow-playground>]
- Luong, T.; Cho, K. and Manning, C. (2016) : Neuronal Machine Translation: Tutorial of Association for Computational Linguistic. [En línea: <https://sites.google.com/site/acl16nmt/>]

### 2.4.2 Recursos para el aprendizaje a través de la práctica

- Google Open Source (2016): TensorFlow Playground. [Aplicación web]. [En línea: <https://opensource.google.com/projects/tensorflow-playground>]
- Google Colab: [entorno de ejecución web de cuadernos Jupyter]. [En línea: <https://colab.research.google.com/notebooks/welcome.ipynb>].

En este entorno de ejecución Google Colab, trabajaremos con diferentes cuadernos Jupyter (Jupyter Notebooks). Guiaremos al alumno para que analice, aprenda a ejecutar e interactuar con el código Python, de diversos ejemplos de cuadernos que construyen, entrenan, evalúan y prueban diversas redes neuronales. En concreto, aprenderemos sobre redes neuronales y traducción automática neuronal interactuando con los siguientes cuadernos:

- TensorFlow Tutorials: *Train your first neural network: basic classification*. [En línea: [https://www.tensorflow.org/tutorials/keras/basic\\_classification](https://www.tensorflow.org/tutorials/keras/basic_classification)]

- ○ TensorFlow Tutorials: Text classification with movie reviews. [En línea: [https://www.tensorflow.org/tutorials/keras/basic\\_text\\_classification](https://www.tensorflow.org/tutorials/keras/basic_text_classification)]
- ○ TensorFlow Tutorials: Neural Machine Translation with Attention. [En línea: [https://www.tensorflow.org/beta/tutorials/text/nmt\\_with\\_attention](https://www.tensorflow.org/beta/tutorials/text/nmt_with_attention)]

## 2.5. Recopilar posibles técnicas de posesición de traducción automática.

La traducción automática neuronal (TAN) trae consigo un grado superior de fluidez del texto pretraducido automáticamente. A la vez, dicha naturalidad en las reformulaciones enmascara errores motivados semánticamente que pueden ser graves.

Es cierto que las relaciones multidimensionales que la TAN establece entre los elementos de las oraciones dan la impresión que el motor "entiende" el sentido (intraoracionalmente) e incluso algunas intenciones de los autores: de hecho, nosotros, al producir lenguaje, vinculamos de manera dinámica todo ese conjunto de características de cada palabra y serie de palabras con diversos aspectos: sonoridad, categoría gramatical, fórmulas memorables y convencionales, asociación con producciones lingüísticas de otros y recuerdos en los medios, culturales, co-locaciones típicas, "sentimiento" positivo o negativo de cada palabra, longitud de oraciones... Ahora bien, la TAN sigue sin poder tener en cuenta el contexto, el co-texto, la creatividad y otras cuestiones pragmáticas que tan importantes son en determinados tipos de texto, lo que hace que debemos prestar especial atención si queremos reproducir un texto de alta calidad, efectivo, y que tenga en cuenta todos los matices y el potencial del lenguaje humano.

Hay otras consideraciones acerca de la naturaleza de las traducciones realizadas automáticamente mediante el método neuronal que están en estudio y que se aportan al final de este apartado en modo de tabla. Proceden de estudios recientes que comparan la traducción automática estadística con la neuronal. Cabría, según dichos estudios y nuestras pruebas, adaptar las recomendaciones para las pautas de posesición de traducción automática neuronal, de la siguiente manera:

De entrada, siempre debe analizarse una muestra de traducción automática neuronal en el ámbito y tipo de texto en cuestión, con el motor escogido. En muchos casos, la TA será de una calidad alta, pero habrá ámbitos donde las especificidades gramaticales, estilísticas o léxicas puedan suponer un descenso pronunciado de la calidad. En esos casos, es muy importante seguir el siguiente proceso:

- Analizar y categorizar los tipos de errores de una muestra representativa.
- Plantearse, a partir del análisis anterior, la posibilidad de un proceso de preedición (sobre todo, si los problemas son de origen estructural o estilístico), y repetir el análisis; de manera alternativa, utilizar dicha categorización como ayuda a la posesición así como, si procede, a la sustitución de patrones basada en expresiones regulares.

Estar particularmente atentos a los errores de transferencia, cultivando unas técnicas adecuadas de lectura comparativa, para lo que se recomienda probar con la direccionalidad

inversa (TT-TO) o alterna (TT-TO, TO-TT, TT-TO, TO-TT, etc.). Aprovechar al máximo los aciertos de la pretraducción, en particular en términos de fraseología y estilo. No cambiar cosas por cambiar si no hay errores o se aprecian mejoras evidentes.

Consultar con los clientes o iniciadores la posibilidad de que se pida o pueda ofrecer una posesición de calidad humana (superior incluso a la tradicionalmente conocida como "full post-editing" o posesición completa). Esto supondría, sobre todo, ampliar el número de relecturas, en particular para detectar cuestiones de coherencia y cohesión textual, colocaciones léxicas, y, crucialmente, de documentación temática.

- El punto anterior apunta a una posible reorganización de las formas de traducción apoyada en traducción automática y asistida por ordenador, donde la posesición no sea sino una técnica más, activable en circunstancias en las que sea útil (por requerir, generalmente, poco tiempo y esfuerzo) frente a la traducción desde cero y/o el apoyo de correspondencias de memoria de traducción (útil sobre todo en tipos de texto muy repetitivos y convencionales, o en ámbitos muy regulados o donde se requiera un control preciso de fórmulas de traducción para tipos expresivos/legales concretos). Desde el punto de vista terminológico, se podría reservar el término "posesición" cuando no se estipule una calidad igual o superior a la que realizaría el ser humano (en condiciones favorables), si bien es cierto que en una gran parte de los casos, dicha traducción humana estaría también alimentada por posesición de traducción automática.
- Así, es probable que en casi todas las circunstancias de traducción (en empresas o de forma autónoma) se use (o se evalúe inicialmente el uso) de TAN, por lo que las opciones resultantes, en función de dicha evaluación inicial podrían ser:
  1. Traducción con calidad humana (incluidos aspectos de estilo);
  2. Posedición de alta calidad;
  3. Posedición semiautomática (con baja intervención humana);
  4. Traducción automática en bruto (para sobrevuelo general de contenido/recoger la idea general).

Aprovechar las características de las herramientas y sistemas de gestión de la edición y la traducción (o traducción asistida por ordenador -TAO) y gestión terminológica.