
Creación de un motor de traducción automática estadístico (EN>ES) para textos del ámbito farmacéutico. Comparación con otros motores de traducción automática neuronal existentes

Trabajo de fin de máster



Universitat Autònoma de Barcelona

Master Tradumàtica: Tecnologies de la Traducció

Facultat de Traducció i d'Interpretació

Universitat Autònoma de Barcelona

Curso 2018 - 2019

Alumna: Alicia Molina Baños

Director: Gökhan Doğru

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN.....	1
2. OBJETIVOS: PREGUNTAS E HIPÓTESIS	2
3. MARCO TEÓRICO Y ANTECEDENTES	3
3.1. La revolución de las herramientas TAO y la traducción automática en el mundo de la traducción	3
3.2. La traducción automática.....	4
3.2.1 Definición	4
3.2.2. La relevancia de la traducción automática y sus usos	4
3.3. Tipos de traducción automática	6
3.3.1. La traducción automática basada en reglas	6
3.3.2. La traducción automática estadística	7
3.3.3. La traducción automática neuronal.....	10
3.4. La posesición, fundamental en la traducción automática.....	12
3.5. La presencia y el impacto de la traducción automática en la empresa española	14
3.6. KantanMT.....	14
3.6.1. Evaluación de la traducción automática	15
4. METODOLOGÍA: DISEÑO Y OBJETO	16
4.1. Recolección de corpus bilingües	16
4.2. Creación de un corpus bilingüe farmacéutico	18
4.3. Creación de un corpus monolingüe en español	23
4.4. Entrenamiento de un motor de traducción automática estadística	27
4.5. Implementación de un corpus bilingüe adicional.....	32
4.6. Comparación de resultados con dos motores de traducción automática neuronal: Google Translate y DeepL	39
4.6.1. Comparación con Google Translate	41
4.6.2. Comparación con DeepL	43

4.7. Ejemplos de posesición con dicho motor de traducción	45
5. RESULTADOS Y LIMITACIONES	55
6. CONCLUSIONES.....	57
BIBLIOGRAFÍA	60
ANEXOS	64

ÍNDICE DE FIGURAS

<i>Figura 1. Ejecución de un archivo jar para extraer el par de lenguas EN-ES de la memoria de traducción ECDC de la UE.</i>	<i>17</i>
<i>Figura 2. Interfaz del alineador automático del programa TAO en línea Wordfast Anywhere.</i>	<i>20</i>
<i>Figura 3. Resultado de la alineación de dos documentos en la herramienta TAO en línea Wordfast Anywhere.....</i>	<i>20</i>
<i>Figura 4. Interfaz de memoQ en la ventana de alineación de textos.</i>	<i>21</i>
<i>Figura 5. Interfaz de memoQ en la ventana de alineación de textos con un error corregido en la segmentación del texto origen y el texto meta.</i>	<i>22</i>
<i>Figura 6. Interfaz de memoQ en la ventana de LiveDocs, la ventana de creación y gestión de corpus.</i>	<i>23</i>
<i>Figura 7. Interfaz de la herramienta en línea Sketch Engine y su función WebBootCaT para crear un corpus a base de páginas web.</i>	<i>25</i>
<i>Figura 8. Ejecución del comando de unión de archivos TXT en la ventana de terminal de Mac OS.</i>	<i>26</i>
<i>Figura 9. Vista de la pestaña de entrenamiento de KantanMT del motor TAE especializado en farmacia de este trabajo de fin de máster.</i>	<i>28</i>
<i>Figura 10. F-Measure del motor TAE entrenado para este trabajo de fin de máster.</i>	<i>29</i>
<i>Figura 11. Puntuación BLEU del motor TAE entrenado para este trabajo de fin de máster.</i>	<i>30</i>
<i>Figura 12. Puntuación TER del motor TAE entrenado para este trabajo de fin de máster.</i>	<i>31</i>
<i>Figura 13. Recuento de palabras del motor TAE entrenado para este trabajo de fin de máster.</i>	<i>32</i>
<i>Figura 14. F-Measure del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.</i>	<i>35</i>

<i>Figura 15. Puntuación BLEU del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.</i>	<i>37</i>
<i>Figura 16. Puntuación TER del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.</i>	<i>37</i>
<i>Figura 17. Recuento de palabras del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.</i>	<i>38</i>
<i>Figura 18. Vista del comparador automático de sistemas TA de la plataforma Tilde MT, basado en la puntuación BLEU.</i>	<i>40</i>
<i>Figura 19. Plataforma de Interactive BLEU score evaluator de Tilde MT que compara las puntuaciones BLEU de Google Translate, en azul, y de nuestro motor TAE, en verde.</i>	<i>41</i>
<i>Figura 20. Oración 9 del texto AFSTYLA comparada con la traducción humana, la traducción automática de Google Translate (Machine en azul) y la de nuestro motor TAE (Machine en verde).</i>	<i>42</i>
<i>Figura 21. Plataforma de Interactive BLEU score evaluator de Tilde MT que compara las puntuaciones BLEU de DeepL, en azul, y de nuestro motor TAE, en verde.</i>	<i>43</i>
<i>Figura 22. Oración 9 del texto AFSTYLA comparada con la traducción humana, la traducción automática de DeepL (Machine en azul) y la de nuestro motor TAE (Machine en verde).</i>	<i>44</i>
<i>Figura 23. Vista de KantanLQR para configurar un nuevo proyecto.</i>	<i>46</i>
<i>Figura 24. Vista de KantanLQR para asignar al traductor encargado de llevar a cabo el proyecto de evaluación de la calidad.</i>	<i>47</i>
<i>Figura 25. Vista de KantanLQR para introducir los textos que compondrán la evaluación de la calidad.</i>	<i>47</i>
<i>Figura 26. Correo electrónico que recibe el traductor de este experimento de posesión para realizar una evaluación de calidad.</i>	<i>48</i>
<i>Figura 27. Vista de KantanLQR como traductor encargado de la evaluación de calidad del proyecto antes de aceptar la tarea.</i>	<i>49</i>
<i>Figura 28. Vista de KantanLQR como traductor encargado de la evaluación de calidad del proyecto durante la posesión.</i>	<i>49</i>
<i>Figura 29. Porcentaje total de fluency en KantanMT de una de las posesiones del motor 2 (Google Translate).</i>	<i>52</i>
<i>Figura 30. Porcentaje total de adequacy en KantanMT de una de las posesiones del motor 2 (Google Translate).</i>	<i>52</i>

Figura 31. Media de los porcentajes de adequacy y fluency asignados por los traductores del experimento a cada uno de los motores. 53

ÍNDICE DE TABLAS

Tabla 1. Datos de entrenamiento del modelo de traducción de nuestro motor TAE. 27
Tabla 2. Datos de entrenamiento del modelo de lengua de nuestro motor TAE. 27

RESUMEN

En este trabajo de fin de máster se lleva a cabo la creación de un motor de traducción automática estadística (EN>ES) especializado en el ámbito farmacéutico mediante la plataforma KantanMT. Se ofrecen pinceladas de los detalles clave de los diferentes sistemas de traducción automática más populares, así como también se habla de la importancia de la posesición en el mundo de la TA y del uso de sistemas TAE en pequeñas y medianas empresas de traducción españolas. Por otro lado, se muestran, además, los pasos a seguir a la hora de entrenar un motor de traducción automática estadística propio en la nube, incluyendo el proceso de búsqueda y creación de corpus. De esta forma, se pretende comprobar si un motor de traducción automática estadística especializado en el ámbito farmacéutico ofrece mejores resultados en dicho ámbito de especialización que algunos de los motores de traducción automática neuronales de carácter general disponibles en la web. Esta comparación viene dada por la creciente popularidad que ha ganado la traducción automática neuronal en los últimos años.

Palabras clave: *traducción automática, traducción automática estadística, traducción automática neuronal, posesición, textos farmacéuticos, proveedores de servicios lingüísticos.*

TITLE

Creation of a statistical machine translation engine (EN>ES) for texts in the pharmaceutical field. Comparison with other existing neuronal machine translation engines.

ABSTRACT

The aim of this Master's Degree Dissertation is the creation of a statistical machine translation engine (EN>ES) specialised in the pharmaceutical field by means of the KantanMT platform. It presents the key details of the different most popular machine translation systems, as well as the importance of post-editing in the world of MT and the use of MT systems in small and medium-sized Spanish translation companies. On the other hand, it also shows the steps to follow when training a statistical machine translation engine in the cloud, including the corpus search and creation processes. In this way, the aim is to check whether a statistical machine translation engine specialising in the pharmaceutical field offers better results in this area of specialisation than some of the generic neuronal machine translation engines available on the web. This comparison is due to the growing popularity of neural machine translation in recent years.

Keywords: *Machine Translation, Statistical Machine Translation, Neural Machine Translation, Post-editing, Pharmaceutical Texts, Language Service Providers.*

TÍTOL

Creació d'un motor de traducció automàtica estadístic (EN>ES) per a textos de l'àmbit farmacèutic. Comparació amb altres motors de traducció automàtica neuronal existents.

RESUM

En aquest treball de fi de màster es duu a terme la creació d'un motor de traducció automàtica estadística (EN>ES) especialitzat en l'àmbit farmacèutic mitjançant la plataforma KantanMT. S'ofereixen pinzellades dels detalls clau dels diferents sistemes de traducció automàtica més populars, així com també es parla de la importància de la posedició en el món de la TA i de l'ús de sistemes TAE en petites i mitjanes empreses de traducció espanyoles. D'altra banda, es mostren, a més, els passos a seguir a l'hora d'entrenar un motor de traducció automàtica estadística propi en el núvol, incloent el procés de cerca i creació de corpus. D'aquesta forma, es pretén comprovar si un motor de traducció automàtica estadística especialitzat en l'àmbit farmacèutic ofereix millors resultats en aquest àmbit d'especialització que alguns dels motors de traducció automàtica neuronals de caràcter genèric disponibles en la web. Aquesta comparació ve donada per la creixent popularitat que ha guanyat la traducció automàtica neuronal en els últims anys.

Paraules clau: *traducció automàtica, traducció automàtica estadística, traducció automàtica neuronal, posedició, textos farmacèutics, proveïdors de serveis lingüístics.*

1. INTRODUCCIÓN

En este trabajo académico se aborda el entrenamiento de un motor de traducción automática estadística especializado en el ámbito farmacéutico con el objetivo de demostrar que, en la actualidad, el grado de especialización de los motores de traducción automática es clave en la industria de la traducción, ya que estos desempeñan un papel muy importante como herramienta de trabajo y apoyo para el traductor humano.

Asimismo, se pretende desestigmatizar el uso de la traducción automática, ya que muchos traductores humanos de la industria se niegan a emplearlos como método de apoyo alegando que pueden acabar con el matiz humano de las traducciones. Es aquí donde entra en juego la importancia de la posesición, es decir, el factor humano indispensable dentro de un proceso tan automatizado como lo es la traducción automática.

Por otro lado, en el marco teórico de este trabajo se hablará de la revolución que supusieron las herramientas TAO en la industria de la traducción y la posterior evolución hacia la traducción automática. Se hará un breve recorrido por los principales tipos de TA y se hablará de la evolución de los mismos, la cual dio lugar a la traducción automática neuronal, el tipo de TA más innovador y actual. Asimismo, se abordarán las métricas de evaluación automáticas de TA más empleadas y del impacto de la TA en la empresa española.

Es en este contexto y, ante el auge de la novedad de los motores de traducción automática neuronal, se lanza la siguiente pregunta: ¿son realmente mejores los motores TAN que los motores TAE cuando el tema de los textos a traducir se enmarca en un ámbito de especialización concreto?

A continuación, se presentarán los objetivos e hipótesis de este trabajo académico, los cuales están relacionados con las dicotomías traductor automático/traductor humano y traducción automática neuronal/traducción automática estadística en el ámbito especializado.

Posteriormente, en la metodología, se explicará el proceso que se siguió para crear el motor TAE especializado, incluyendo la recopilación de recursos que posteriormente se emplearían para entrenar el motor, como corpus oficiales de la Unión Europea, textos paralelos especializados para la creación de un corpus bilingüe propio, etc.

Tras la creación y entrenamiento del motor, se analizaron los resultados de nuestro motor TAE especializado mediante métricas de evaluación automática, así como se compararon con dos motores de traducción automática general de carácter más genérico: Google Translate y DeepL.

Además, en el apartado de metodología, se incluyó un experimento de posesición para que una muestra de traductores humanos evaluara la calidad de los tres motores en función de la posesición necesaria en un texto farmacéutico de muestra, para así añadir una valoración humana a las evaluaciones automáticas que se obtuvieron.

Tras la metodología, se incluye un breve apartado en el que se presentan los resultados y limitaciones más destacadas del experimento que se ha llevado a cabo en este trabajo.

En este trabajo, mediante la creación de un motor de traducción automática estadística en el ámbito farmacéutico, se pretende demostrar la importancia de la especialización de los motores de traducción automática como herramienta de apoyo para el traductor, así como el carácter esencial de la posesición humana.

2. OBJETIVOS: PREGUNTAS E HIPÓTESIS

En este apartado nos encargaremos de plantear y dar forma a las preguntas a las que queremos dar respuesta y a las hipótesis que queremos demostrar con la realización de este trabajo.

A continuación, se mostrarán las preguntas que hemos considerado pertinente plantear:

- ¿Un motor de traducción automática estadística especializado propio ofrece mejores resultados que algunos de los motores gratuitos y genéricos ya existentes?
- ¿Merece la pena invertir tiempo en la creación de corpus especializados destinados a un motor de traducción automática propio o es más útil dedicar dicho tiempo a la posesición tras traducir con uno de los motores de traducción automática neuronales de carácter general disponibles en la web?
- ¿La creación de un motor de traducción automática especializado en un campo muy concreto puede resultar factible para una empresa?

En cuanto a las hipótesis, mediante la realización de este estudio queremos demostrar lo siguiente:

- La cantidad de posesición necesaria para un texto especializado es mucho menor si el motor de traducción automática que se ha empleado para traducirlo es adecuado a la tipología textual.

Tras la exposición de los objetivos de este trabajo, procederemos a adentrarnos en las nociones teóricas de los conceptos que se van a tratar a lo largo del mismo.

3. MARCO TEÓRICO Y ANTECEDENTES

El marco teórico de este trabajo comenzará con una pequeña introducción sobre qué ha supuesto la llegada de la traducción automática al mundo de la traducción en comparación con las herramientas TAO. Posteriormente, se centrará en explorar la traducción automática y los tipos más destacados. A continuación, se abordará la importancia de la posesión en este campo, así como la presencia e impacto de la traducción automática en las empresas españolas. Finalmente, se hablará de la plataforma de traducción automática estadística que se empleará en este trabajo para el entrenamiento del motor propio: KantanMT y de algunas de las medidas de evaluación de traducción automática más empleadas.

3.1. La revolución de las herramientas TAO y la traducción automática en el mundo de la traducción

En una sociedad cada vez más tecnológica, las empresas de traducción deben saber adaptarse al ritmo frenético de sus clientes. De hecho, como ya bien dijo Parra (2018): «La traducción es, como muchas otras, una profesión supeditada a los avances tecnológicos». (p. 26)

En su momento, la revolución llegó con las herramientas TAO (Traducción Asistida por Ordenador), sobre todo con la gran utilidad de las memorias de traducción, que ayudaron al traductor a ser más rápido y efectivo, ya que permitieron reutilizar traducciones anteriores.

En cuanto a este tema, Arevalillo (2012) ya destacó la productividad que trajeron a las empresas las memorias de traducción, que permitieron no solo trabajar con gran rapidez, sino también aportar una mayor coherencia terminológica a las traducciones de un mismo cliente, por ejemplo. Este autor quiere también marcar una clara diferencia entre la traducción asistida por ordenador y la traducción automática, porque pese a ser ambas herramientas tecnológicas aliadas del traductor, la traducción automática, objeto de estudio de este trabajo, va un paso más allá en el proceso de la traducción.

Por otra parte, Arevalillo (2012) también subraya que la dificultad con la que ambas entraron en la industria de la traducción no es equivalente, ya que, a pesar de que la traducción asistida necesita de un mantenimiento de las memorias de traducción y de la familiarización con una herramienta TAO concreta, la traducción automática conlleva un proceso más complicado. La generación de motores de traducción automática fiables no solo requiere de una gran inversión de tiempo, sino también de mayores gastos en la contratación de otros profesionales indispensables del sector, los poseedores, figura sobre la que hablaremos en apartados posteriores.

Sin embargo, las herramientas TAO y la traducción automática no son tan diferentes como cabría pensar. Collantes et al. (2018) subrayan algunas de sus similitudes más destacadas, como el aumento de la productividad de la tarea de traductores, el rechazo de una parte del sector de la traducción, su perfeccionamiento a lo largo del tiempo y su carácter dependiente de la calidad de la información que utilicen como base.

3.2. La traducción automática

3.2.1 Definición

Sería conveniente comenzar este apartado definiendo la traducción automática. La EAMT (s.f.) acuña la siguiente definición:

«Machine translation (MT) is the application of computers to the task of translating texts from one natural language to another»¹.

Berner (2003), por su parte, aporta una definición más desarrollada:

Machine translation (MT) is the use of computer software to translate text or speech from one natural language into another. Like translation done by humans, MT does not simply involve substituting words in one language for another, but the application of complex linguistic knowledge: morphology (how words are built from smaller units of meaning), syntax (grammar), semantics (meaning), and understanding of concepts such as ambiguity. (Berner, 2003)

3.2.2. La relevancia de la traducción automática y sus usos

Es bien sabido que la principal razón por la que surgió la necesidad de agilizar y automatizar, en la medida de lo posible, el proceso de traducción y, por ende, aumentar la productividad, es debido a la gran cantidad de información que manejan las empresas internacionalmente como consecuencia de la globalización.

Way (2013) resume de forma bastante acertada lo que está sucediendo actualmente en la industria de la traducción:

¹ EAMT - European Association for Machine Translation. Recuperado de <http://www.eamt.org/mt.php>

The key question we all face in the translation industry today is how to help businesses cope with the explosion of content in the global economy, especially given the difficult prevailing economic circumstances. Not only is the internet more multilingual than ever before, but there's a growing demand for very rapid – or even instant – communication. (Way, 2013)

En este escenario, las ventajas de la traducción automática son evidentes, pues permiten a la industria de la traducción, ante tal sobrecargo de información, aceptar encargos que, en otra ocasión, probablemente se hubieran visto obligados a rechazar por cuestiones de magnitud y tiempo.

Tradicionalmente, la traducción automática se ha venido empleando de tres formas diferentes: sin posesición, ligeramente poseída y totalmente poseída. Normalmente la traducción automática sin posesición se usa en el sector de las tecnologías (blogs, chats en línea...), en manuales de instrucciones, en ciertos textos de *marketing* (respuestas a cuestionarios y encuestas...) y en *e-commerce* (información básica de productos, por ejemplo). Sin embargo, normalmente, en *marketing* se prefiere contenido traducido por humanos o textos escritos por expertos, ya que entra en juego una mayor creatividad y originalidad. Además, la TA también se emplea con fines de comunicación interna dentro de una empresa, es decir, para traducir correos electrónicos, para listados de información básica de productos, etc. Por su parte, la traducción automática ligeramente poseída se suele utilizar para manuales de ayuda en línea, documentación de apoyo, estudios de mercado, etc. Finalmente, la traducción automática con posesición completa suele ser la más útil a la hora de traducir manuales cuya mala interpretación puede poner en peligro la seguridad y la salud del individuo, para traducir documentos dirigidos a clientes, contratos, patentes o información detallada sobre un producto en venta (Way, 2013).

Way (2013) también informa de que, a los usos tradicionales comentados en el anterior párrafo, se está sumando, en el caso de la traducción automática sin posesición, la traducción de contenido generado por el usuario, como valoraciones de hoteles o productos, publicaciones en las redes sociales, chats en línea, etc. Además, muchas empresas también emplean la TA para ofrecer páginas web multilingües y, en el caso del ámbito educativo, se usa en muchas ocasiones para traducir los programas de las diferentes formaciones académicas.

Tal y como se ha comprobado a lo largo de este apartado, la traducción automática está presente en muchas industrias actuales y su uso no deja de aumentar. Su relevancia en el mundo

globalizado en que vivimos es innegable. A pesar de que muchas industrias utilizan la traducción automática sin posesición, el mundo de la traducción profesional resiente los resultados no tan ideales de la TA. En relación con este tema, Alcina (2011) se mantiene optimista, ya que, a pesar de que queda un largo camino para mejorar la calidad de la traducción automática, subraya que la esfera investigadora del mundo de la traducción intenta mejorar los resultados por medio de distintas vías, entre las que se encuentran la preedición y posesición de textos, así como la creación de motores de traducción automática especializados en textos según su tipo, tema u otras variables.

A continuación, veremos los principales tipos de traducción automática existentes, para posteriormente centrarnos en los sistemas estadísticos y neuronales, que son los que se compararán en este trabajo.

3.3. Tipos de traducción automática

En la actualidad, a pesar de que existen muchos tipos de traducción automática, se dividen principalmente en las siguientes tres categorías: la TABR o Traducción Automática Basada en Reglas (*RBMT* o *Rule-Based Machine Translation* en inglés), la TAE o Traducción Automática Estadística (*SMT* or *Statistical Machine Translation* en inglés) y la TAN o Traducción Automática Neuronal (*NMT* o *Neural Machine Translation* en inglés), que es la más empleada hoy en día. Antes de centrarnos en los dos últimos tipos de traducción automática mencionados, sería interesante explorar brevemente cómo funciona la TABR.

3.3.1. La traducción automática basada en reglas

Hearne & Way (2011b) explican la TABR como una estrategia que implica conocimiento lingüístico, ya que la oración del texto origen debe analizarse primero para que así se le pueda asociar una anotación lingüística. Esta cadena de texto anotada se traduce entonces utilizando recursos léxicos y estructurales (es decir, gramaticales) y se genera una oración en la lengua meta a partir de esta anotación traducida. Estas operaciones se realizan generalmente haciendo uso de recursos lingüísticos como diccionarios y gramáticas: monolingües, en los casos de análisis y generación, y bilingües en el caso de transferencia.

Ginestí & Forcada (2009) destacan que, entre los sistemas basados en reglas, los más habituales son los de traducción automática por transferencia, que funcionan en tres fases bien diferenciadas:

1. La fase de análisis produce, a partir de la frase en lengua origen, una representación intermedia abstracta (interlingua), en la que se establecen clasificaciones y agrupamientos lingüísticos que permiten la aplicación de reglas generales de traducción.

2. La fase de transferencia convierte la representación intermedia anterior en una nueva representación intermedia en la lengua meta, restableciendo en ella la concordancia y haciendo reordenamientos o cambios léxicos.

3. La fase de generación produce una frase concreta en la lengua meta a partir de esta representación intermedia abstracta.

A pesar de que la TABR genera buenos resultados, podría decirse que, en cierto modo, está algo limitada, ya que como subraya Parra (2018), además de necesitar diccionarios y gramáticas, entre otros, «el sistema no puede traducir estructuras lingüísticas que no estén incluidas en sus gramáticas o reglas de transferencia, ni palabras o expresiones que no se encuentren en sus diccionarios» (p. 21). Esto convierte la TABR en un sistema al que dedicar gran cantidad de tiempo y recursos para poder mantenerlo actualizado, de ahí que en los mundo académico y profesional se haya invertido más en otros sistemas como la TAE y la TAN, ya que su mantenimiento no es tan costoso y, al parecer, suele proporcionar mejores resultados a nivel general y una mayor fluidez en el texto meta.

Como ejemplo de motor de TABR, Parra (2018) destaca la plataforma de código abierto Apertium, que fue desarrollada por el grupo de investigación Transducens, de la Universidad de Alicante y que se emplea, sobre todo, para la traducción entre lenguas similares, como el catalán, el gallego, el portugués o el francés.

Tras este momentáneo paseo por el mundo de la TABR, es hora de hablar de uno de los sistemas de TA que más se ha empleado en los últimos años, la traducción automática estadística.

3.3.2. La traducción automática estadística

Como alternativa a los sistemas TABR, surgieron los sistemas TAE, cuyo entrenamiento es bastante menos costoso, ya que, al contrario que los sistemas TABR, que necesitan disponer de diccionarios, gramáticas, reglas de transferencia, etc., los TAE únicamente necesitan servirse de información para generar una traducción, es decir, no necesitan conocer la gramática ni demás reglas lingüísticas de los idiomas implicados.

Hearne & Way (2011a) destacan que los sistemas TAE pretenden averiguar cuál es la traducción más probable y para ello se sirven de corpus: un gran corpus monolingüe de la lengua de destino, que serviría para inducir el modelo de lengua y un corpus paralelo bilingüe, que induce el modelo de traducción.

En cuanto al modelo de lengua, estos autores (2011a) explican que este proporciona la probabilidad de que una cadena de texto meta sea realmente una oración válida en dicha lengua. Por lo tanto, se puede afirmar que este modelo cuenta con dos características principales: aporta tanto un modelo del corpus de entrenamiento monolingüe como un método para calcular la probabilidad de una cadena de texto meta previamente no vista según ese modelo. Por su parte, el modelo de traducción calcula la probabilidad de que una cadena de texto origen y una cadena de texto meta sean equivalentes. Por lo tanto, se puede afirmar que el modelo de traducción proporciona tanto un corpus bilingüe alineado como un método para calcular la probabilidad de que el texto origen y el texto meta sean equivalentes.

La TAE emplea dos procesos distintos y separados: el entrenamiento y la decodificación. La fase de entrenamiento consiste en extraer un modelo de traducción de un corpus paralelo y un modelo de lengua de un corpus monolingüe (normalmente mucho mayor).

El modelo de traducción comprende un diccionario bilingüe en el que cada posible traducción de una palabra o frase de origen tiene una probabilidad asociada. Sin embargo, este modelo no se parece a un diccionario convencional en el que solo se permiten entradas plausibles, ya que muchas de las entradas representan traducciones que son improbables, pero no imposibles. A estos diccionarios probabilísticos también se les denomina tablas de frases.

El modelo de lengua comprende una base de datos de secuencias de palabras en el idioma de destino (generalmente entre una y siete palabras de longitud), cada una de las cuales también está asociada a una probabilidad. Durante el entrenamiento se puede extraer información adicional más allá de los modelos de lengua y de traducción, como modelos de longitud relativa de las frases, reordenación de palabras, importancia relativa de la traducción frente al modelo de lengua, etc. Estos modelos inducidos se utilizan durante la decodificación, el proceso que realmente produce una traducción. El proceso de decodificación trata la traducción meramente como un problema de búsqueda: dada una frase a traducir, busca todas las traducciones posibles permitidas por el modelo de traducción, y todas las posibles reordenaciones de las mismas, en busca de la que tenga asignada la mayor probabilidad global según los modelos de traducción y lengua. (Hearne & Way, 2011a)

Hearne & Way (2011b) destacan las principales diferencias entre la TABR y la TAE basándose en una distinción entre proceso y producto. Afirman que, mientras que la TABR se centra en el proceso mediante el cual se produce una traducción para cada oración origen, la TAE se centra en cómo determinar cuál es la mejor de dos o más traducciones propuestas para una oración origen. Por lo tanto, la TABR se enfoca en la mejor manera de generar una traducción para cada cadena de texto origen (proceso), mientras que la TAE se centra en generar miles de traducciones hipotéticas para la cadena de texto origen para entonces determinar cuál es la más probable (producto).

Tras ver cómo funcionan ambos sistemas de traducción, podría destacarse también otra gran diferencia que hemos mencionado anteriormente. Tal y como Ginestí & Forcada (2009) apuntan:

Els sistemes basats en regles requereixen més temps de construcció (cal codificar explícitament les dades lingüístiques que utilitzarà el sistema) mentre que els sistemes basats en corpus són més ràpids de construir sempre que es disposi d'un gran volum de corpus de textos alineats; per tant, són difícils d'aplicar a la traducció d'una llengua minoritària amb poca disponibilitat de corpus bilingües digitalitzats.
(p.48)

Esto significa que, aunque la TAE también requiere recursos humanos (creación de corpus, etc.), el volumen de trabajo para crear un sistema TAE no es tan considerable como el de la TABR, ya que codificar los datos lingüísticos del sistema conlleva una gran inversión de tiempo y recursos humanos, por lo que el desarrollo de los sistemas TABR es mucho más costoso que el de los TAE. Sin embargo, cuando hablamos de lenguas minoritarias que apenas cuentan con corpus bilingües, los sistemas TABR son más adecuados que los TAE.

Habiendo comparado ya, a grandes rasgos, ambos tipos de traducción automática, hemos de señalar que, en nuestro caso, al querer centrarnos en un ámbito especializado, la farmacia, el hecho de que nuestro motor TAE esté compuesto de corpus con la misma tipología textual que la de los textos que se van a traducir supone una gran ventaja. En este caso, por razones de tiempo limitado y accesibilidad de recursos, puede afirmarse, por lo tanto, que la elección de un sistema TAE resulta más beneficiosa para nosotros que un sistema TABR.

Como ejemplo referente de los sistemas TAE se encuentra Moses. Parra (2018) informa que se trata de un sistema de traducción automática estadística de código libre que nació gracias a la colaboración derivada de diversos proyectos europeos. Sistemas como este determinaron el

nacimiento de otros proyectos, como Moses for Mere Mortals², que tenían como objetivo acercar el mundo de la traducción automática estadística a traductores profesionales.

En cuanto a Moses³, se trata de un *software* de código abierto y gratuito que sirve para entrenar sistemas de traducción automática estadística y que fue fruto del proyecto EuroMatrix, cuyo objetivo fue el de desarrollar y mejorar los sistemas de traducción automática para todos los pares de lenguas oficiales de la Unión Europea.

Asimismo, Parra (2018) también hace mención a otras soluciones comerciales para entrenar motores TAE propios en la nube que surgieron posteriormente, como KantanMT, basado en Moses, que es la plataforma que se empleará en este trabajo para entrenar un motor de traducción y en la que profundizaremos en un apartado posterior.

En este ámbito, sería interesante señalar la existencia de MTradumàtica⁴, que como bien señalan Doğru et. al (2017), es una alternativa a KantanMT de código abierto y gratuita. Se trata de una plataforma web basada también en Moses que sirve para entrenar y crear motores de traducción automática estadística.

A continuación, y después de haber conocido las principales características y diferencias de los sistemas TABR y TAE, nos adentraremos en el mundo de la traducción automática neuronal, un sistema TA innovador en el que se está empezando a invertir cada vez más y que en los últimos años ha comenzado a desplazar considerablemente a los sistemas TAE.

3.3.3. La traducción automática neuronal

Bahdanau, Cho & Bengio (2014) presentan la traducción automática neural de la siguiente forma: «Neural machine translation is a recently proposed approach to machine translation. Unlike the traditional statistical machine translation, the neural machine translation aims at building a single neural network that can be jointly tuned to maximize the translation performance» (p.1).

Forcada (2017) va más allá y expresa que la traducción automática neuronal es una nueva generación de traducción automática basada en corpus. Se entrena con corpus enormes de pares de segmentos de la lengua origen (generalmente oraciones) y con sus traducciones, es decir, básicamente se compone de enormes memorias de traducción que contienen cientos de miles o

² Moses for Mere Mortals (<https://github.com/jladcr/Moses-for-Mere-Mortals>)

³ Moses (<http://www.statmt.org/moses/>)

⁴ MTradumàtica (<https://mtradumatica.uab.cat/>)

incluso millones de unidades de traducción. En este sentido, es similar a la tecnología de traducción automática estadística que fue la predominante hasta hace muy poco tiempo, pero utiliza un enfoque computacional completamente diferente: las redes neuronales.

Parra (2018) ilustra un ejemplo de cómo funcionan estas redes neuronales artificiales: explica que el sistema es capaz de aprender, por ejemplo, que las palabras «perro» y «gato» están más relacionadas que «perro» y «mesa». El sistema deduce esta información porque en el texto generalmente estos términos estarán rodeados de palabras similares que nos permiten averiguar posibles traducciones.

Sin embargo, esto puede representar un arma de doble filo, ya que como Parra (2018) ejemplifica, estos sistemas neuronales pueden, por ejemplo, cambiar en la traducción el nombre de un país por el de otro y todo por el simple hecho de haber aprendido que, en contextos similares, en una posición determinada de una oración suele aparecer el nombre de ese país.

Asimismo, Parra (2018) afirma que también se ha demostrado que en el caso de frases muy largas los sistemas TAN suelen obtener peores resultados que los sistemas TAE.

Asimismo, Forcada (2017) explica que los sistemas TAN son difíciles de entrenar, incluso más que los sistemas TAE, que ya requerían corpus paralelos que generalmente no están disponibles para traductores autónomos o incluso para agencias de traducción pequeñas. Estos corpus paralelos son mucho más grandes que las memorias de traducción habituales. Se necesita *hardware* específico (como GPU, originalmente empleado como unidades de procesamiento gráfico) y los tiempos de entrenamiento (días, semanas, meses) pueden ser demasiado largos en algunos casos.

Como conclusión de este subapartado, merece la pena mencionar el razonamiento de Parra (2018), que habla de la predominancia de los sistemas TAN en el ámbito académico y de su entrada gradual en el mundo de la traducción profesional. Subraya, además, el hecho de que hay corporaciones como Google, Microsoft, eBay, Booking.com o DeepL que pueden sí que pueden invertir en comprar supercomputadores que les permitan entrenar motores neuronales propios y manejar grandes bases de datos, pero en el mundo de las pequeñas y medianas empresas esto no representa una posibilidad real y muchas siguen empleando motores estadísticos. A pesar de esto, esta autora habla de la existencia de algunas propuestas de traducción automática neuronal que tienen como objetivo disminuir la enorme cantidad de datos necesaria para entrenar motores neuronales. Se denomina *zero-shot translation* y se basa en el uso de una interlingua. Por el momento ya hay algunos proyectos, como el reciente

ModernMT⁵, que pretende impulsar la creación de motores neuronales con recursos más modestos.

En este punto convendría señalar que, en un principio, se consideró la opción de crear un motor TAN para este trabajo de investigación, junto con otro TAE. Sin embargo, fueron varias las razones que hicieron que se descartara: razones computacionales, por el hecho de no contar con un ordenador con una interfaz gráfica de usuario (GUI) lo suficientemente potente; razones técnicas, ya que no disponíamos de las habilidades técnicas necesarias para entrenar un motor de tales características; y, por último, económicas, ya que crear un motor de traducción automática neuronal en una plataforma basada en la nube hubiera sido muy costoso (por ejemplo, Google AutoML).

Seguidamente, tras haber visitado los rincones más destacados de la traducción automática, procederemos a hablar de una tarea inherente a esta tecnología: nos referimos a la posesición.

3.4. La posesición, fundamental en la traducción automática

En cuanto al debate de la posesición (PE), Rico (2014) nos pone en contexto y afirma que en los últimos años ha surgido un papel aparentemente nuevo en el mundo de los proveedores de servicios lingüísticos: el del poseedor. El debate sigue abierto a quién, cómo y cuándo se debe hacer la posesición y, aunque los traductores tienen opiniones opuestas, la realidad es que la industria está incorporando progresivamente la traducción automática como un servicio más (y el cliente, además, lo pide). Aclarado esto, podríamos definir la posesición como una tarea específica para la corrección de la producción de traducción generada automáticamente.

Aunque los recientes avances en la traducción automática han situado la posesición en el centro del debate, la verdad es que la posesición como profesión en sí misma se remonta a 1985. La traducción automática se había implementado algunos años antes para la traducción del español al inglés. La posesición se definió entonces como un ajuste del resultado de la máquina para que reflejara con la mayor precisión posible el significado del texto original. Desde entonces, el uso del término «poseedor» y, por ende, su papel específico en la industria de la traducción como profesión en ascenso ha ido ganando impulso.

⁵ ModernMT (<https://www.modernmt.eu/>)

Doherty & Gaspari (2013) definen los objetivos principales de la posedición:

The aim of PE is to improve the output, not necessarily to make it perfect. Post-edited output must become (more) usable or understandable; least possible effort must be applied quickly: the priority is to save time (not to lose the speed gains due to MT) and money; the extent and accuracy of PE are negotiated/specified on a case by case basis, depending on user's needs and requirements.

Por otra parte, Zaretskaya (2017) añade que «the level of acceptance of MT in the translators' community also seems to have raised in the past years and one can say that post-editing has become another regular type of services linguists provide». (p.117)

Zaretskaya (2017) advierte, además, que la posedición de TA es un tipo de servicio relativamente nuevo y muchos lingüistas no son completamente conscientes de lo que implica. Por ello, la formación es esencial y existen varias empresas y organizaciones que imparten formación sobre MTPE (*Machine Translation Post-Editing*). Los cursos más conocidos son probablemente los cursos de formación en línea creados por TAUS.

Cabe destacar que TAUS⁶ ha creado una guía informativa en línea sobre posedición en traducción automática que ofrece recomendaciones muy útiles para realizar posediciones de calidad.

Zaretskaya (2017), por su parte, (2017) da una visión global de la situación actual de la posedición en la industria de la traducción:

The translation industry is changing and it affects all the stakeholders including translators. There is no doubt that MT will play a significant role in the near future. Even though it existed for many decades, MT has found a way into the translators' business and everyday work only recently. For translators, it is a relatively new type of technology and post-editing is a new task, which they have to learn and get used to. (p. 123)

De esta información se desprende que la posedición es una profesión cada vez más en auge debido a la creciente presencia de la TA en el mundo de la traducción profesional y que la formación de traductores en esta disciplina es necesaria. En último lugar, estos

⁶ TAUS (2010). MT Post-editing Guidelines. Recuperado de <https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines>

razonamientos sobre TAE y posesión nos conducen a discutir en el siguiente apartado la presencia de la traducción automática en las empresas españolas.

3.5. La presencia y el impacto de la traducción automática en la empresa española

En este breve apartado, discutiremos la postura general de las pequeñas y medianas empresas españolas del sector de la traducción en términos de traducción automática. ¿Emplean TA las agencias de traducción españolas de pequeño y mediano tamaño? ¿O el ven como un intrusismo que rebaja la calidad de las traducciones?

Torres-Hostench, Presas & Cid-Leal (2016) realizan un análisis de estas preguntas clave. Para ello escogieron una muestra de 187 empresas distribuidas por toda España, de las cuales 55 respondieron a la encuesta:

Casi la mitad de las empresas encuestadas utilizan TA en el flujo de trabajo. Sin embargo, resulta llamativo que el 45,5 % solo la emplean en menos del 10 % de sus proyectos y únicamente el 16 % de las empresas que respondieron dispone de un sistema de traducción automática propio. (p. 27)

Esos datos demuestran que los sistemas de traducción automática todavía no tienen gran presencia en el día a día de las pequeñas y medianas agencias de traducción. En este estudio tenemos como objetivo demostrar que merece la pena invertir en esta tecnología de la traducción, ya que los resultados, poseídos, pueden ser llegar a ser igual de buenos que los de una traducción humana y, además, con un coste inferior de tiempo.

En el siguiente y último apartado del marco teórico de este trabajo, abordaremos el sistema de evaluación automática de motores TA que emplea KantanMT, es decir, el mismo sistema que utilizaremos para evaluar nuestro motor TAE.

3.6. KantanMT

Este apartado está dedicado a explicar qué es KantanMT y qué funcionalidades nos interesan para nuestro trabajo de investigación.

KantanMT es una plataforma en la nube de carácter privado que permite crear motores de traducción automática estadística personalizados y de la que haremos uso para crear nuestro motor TAE especializado en farmacia. Además, esta plataforma genera automáticamente evaluaciones de calidad de los motores entrenados basándose principalmente en tres medidas: las puntuaciones BLEU, F-Measure y TER.

3.6.1. Evaluación de la traducción automática

Según aparece en las diferentes secciones de la página web de KantanMT.com (s.f.), algunos de los principales sistemas de evaluación de la calidad de TA son los siguientes: BLEU, TER y F-Measure. Ahora se procederá a explicar en qué consiste cada uno.

En primer lugar, la métrica BLEU califica las traducciones en una escala de 0 a 1, pero se muestra frecuentemente como un valor porcentual. Cuanto más cerca de 1, más se correlaciona la traducción con una traducción humana. En pocas palabras, la métrica BLEU mide cuántas palabras se superponen en una traducción en comparación con una traducción de referencia.

Según KantanMT (s.f.), la puntuación de BLEU oscila entre 0-100 %. Una puntuación inferior al 15 % significa que el motor no está funcionando de manera óptima y se requerirá un alto nivel de posesición para que los textos tengan la suficiente calidad como para publicarse. Por el contrario, una puntuación superior al 50 % es una puntuación muy buena y se requerirá una posesición significativamente menor.

Por su parte, la medida TER (*Translation Error Rate*), según Snover et al. (2006): «(...) is an error metric for machine translation that measures the number of edits required to change a system output into one of the references». Es decir, se encarga de determinar la cantidad de edición que un traductor tiene que realizar para modificar una traducción de modo que coincida exactamente con la traducción de referencia.

La puntuación TER se encuentra en un rango de 0-1, pero frecuentemente se presenta en porcentaje, al igual que la medida BLEU y, cuanto más bajo sea el porcentaje, mejor será la calidad de la traducción y menor la posesición necesaria. En cuanto a la siguiente y última medida, la F-Measure, se trata de una medición automatizada que determina la precisión y la capacidad de recuperación de un motor KantanMT. Puede utilizarse como guía general para determinar el rendimiento general de calidad de un motor y se muestra también como un valor porcentual. (KantanMT, s.f.)

Tras haber concluido este apartado, en el capítulo siguiente se procederá a explicar la metodología que se ha seguido para el entrenamiento del motor TAE especializado en el ámbito farmacéutico que nos ocupa en este trabajo.

4. METODOLOGÍA: DISEÑO Y OBJETO

En este capítulo, se procederá a la explicación del proceso de recolección de corpus de medicina y de la creación de un corpus propio a base de textos farmacéuticos. Asimismo, también se hablará del procedimiento de entrenamiento y desarrollo de un motor de traducción automática estadística especializado en farmacia mediante la plataforma KantanMT. Posteriormente, evaluaremos dicho motor con la ayuda de las medidas BLEU, TER y F-Measure y lo compararemos con dos motores de traducción automática neuronal: Google Translate y DeepL. Por último, pondremos a prueba los tres motores para comprobar cuál ofrece mejores resultados. Estos serán evaluados mediante métricas de evaluación automática, y, por lo tanto, se determinará cuál necesita menos posesición cuando se trata de traducción farmacéutica.

4.1. Recolección de corpus bilingües

En primer lugar, ha de mencionarse el corpus EMEA (EN-ES), un corpus de textos farmacéuticos de carácter público que se puede descargar gratuitamente de la página web OPUS⁷ en varios formatos. En nuestro caso, lo descargamos en formato TMX, que se trata del formato XML estándar para las memorias de traducción. Este corpus está formado por textos de la Agencia Europea del Medicamento (EMA⁸), es decir, por prospectos farmacéuticos y por fichas técnicas de medicamentos que se comercializan en la Unión Europea. Este corpus es el más interesante que puede ofrecer la Unión Europea para nuestro fin, ya que se compone exactamente del mismo tipo de textos que queremos que se traduzcan con nuestro motor de traducción, además de que se trata de textos de buena calidad que han sido traducidos por traductores humanos.

Esta memoria de traducción cuenta, de acuerdo con el análisis de memoQ, con 366 769 segmentos y 5 302 365 palabras en la lengua origen, es decir, en inglés.

A continuación, mencionaré y haré una breve descripción del resto de corpus y memorias de traducción de la Unión Europea que se han empleado para el entrenamiento del motor de traducción que nos ocupa en este trabajo.

En cuanto a la memoria de traducción de la ECDC⁹ (Centro Europeo para la Prevención y el Control de Enfermedades), representó una fuente de información de gran utilidad dada la

⁷ OPUS (<http://opus.nlpl.eu/>)

⁸ EMA (<https://www.ema.europa.eu/en>)

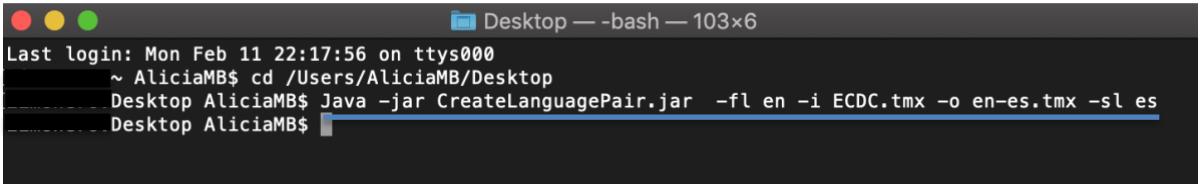
⁹ ECDC (<https://ecdc.europa.eu/en/home>)

temática de la misma. Esta memoria de traducción se construyó sobre la base del sitio web de la ECDC, por lo que la mayor parte de los documentos trata de temas relacionados con la salud, pero también se describen la organización y las funciones de la ECDC, lo cual puede resultar interesante para el aspecto más técnico de los documentos farmacéuticos a traducir.

Para extraer la combinación de lenguas que nos interesaba (EN-ES), fue necesario descargar un archivo comprimido de la página web de la Unión Europea que contenía todos los pares de lenguas. Además, dentro de esta carpeta comprimida se incluía un documento de texto que explicaba cómo hacer la extracción de un par de lenguas concreto y un programa creado con Java. Se empleó el siguiente comando en la ventana de la terminal: `Java -jar CreateLanguagePair.jar -fl First_Language -i Input_File -o Output_File -sl Second_Language`.

«First_Language» es el primer idioma que se quiere tener en el par de idiomas, «Second_language» es el segundo idioma que se quiere tener en el par de idiomas, «Input_File» es la memoria de traducción principal de la que quiere extraer el par de idiomas, es decir, el archivo «ECDC.tmx» y, por último, «Output_File» es el nombre del archivo con los resultados de la extracción, por ejemplo, en nuestro caso fue «en-es.tmx». Es importante mencionar que, al seleccionar los idiomas, es necesario utilizar los códigos ISO de dos letras. Por ejemplo, «en» en el caso del inglés y «es» en el caso de español.

A continuación, se muestra una imagen de la ventana de terminal con el proceso que se siguió para la extracción del par de lenguas EN > ES.



```
Desktop — -bash — 103x6
Last login: Mon Feb 11 22:17:56 on ttys000
~ AliciaMB$ cd /Users/AliciaMB/Desktop
Desktop AliciaMB$ Java -jar CreateLanguagePair.jar -fl en -i ECDC.tmx -o en-es.tmx -sl es
Desktop AliciaMB$
```

Figura 1. Ejecución de un archivo jar para extraer el par de lenguas EN-ES de la memoria de traducción ECDC de la UE.

Esta memoria de traducción nos proporcionó, según el análisis realizado en memoQ, 2502 segmentos y 36 813 palabras en la lengua origen para la combinación de idiomas que nos interesa: inglés-español.

Por último, hemos de mencionar la biblioteca de KantanMT, de la cual hicimos uso del corpus paralelo temático Medical 1 (EN > ES). La biblioteca de KantanMT es un repositorio de datos de entrenamiento de gran calidad que ayuda a mejorar y personalizar rápidamente los motores de KantanMT, la plataforma que emplearemos posteriormente para el entrenamiento

de nuestro motor TAE. Estos datos están a disposición del público y tienen Derechos de Propiedad Intelectual (DPI).

Este corpus Medical 1 está compuesto por 5 369 913 palabras en inglés y sus correspondientes 6 043 134 en español.

A continuación, se procederá a explicar el proceso de creación del corpus bilingüe propio especializado en farmacia que se creó con el objetivo de aumentar la calidad y la cantidad de datos especializados con los que posteriormente entrenaremos nuestro motor.

4.2. Creación de un corpus bilingüe farmacéutico

Este apartado del trabajo está dedicado a la explicación de la creación del corpus farmacéutico propio que se ha utilizado para complementar los corpus del apartado anterior.

Para comenzar, ha de señalarse que para recopilar los textos que posteriormente se alinearían, se utilizó como principal fuente de información la página web de la EMA, por tratarse de textos oficiales de la Unión Europea. Para asegurarnos de que los textos que elegíamos para descargar no estaban repetidos en el corpus de la EMEA que habíamos bajado con anterioridad del sitio web OPUS, decidimos comprobar si los medicamentos más recientes de la página web de la EMA estaban incluidos en dicho corpus. Todos aquellos que buscábamos y no estaban incluidos, decidimos que serían los que formarían parte de nuestro corpus. Este fue el proceso que se siguió para seleccionar qué medicamentos íbamos a escoger.

A continuación, mencionaré los medicamentos que están incluidos en el corpus: Amgevita, Apealea, Axumin, Besponsa, Braftovi, Cablivi, Cabometyx, Cinquaero, Cystadrops, Darzalex, Delstrigo, Diflucan, Feraccru, Flixabi, Fulphila, Halimatoz, Incruse, Influvac, Inhixa, Jinarc, Kalydeco, Kisplyx, Kisqali, Ledaga, Lemtrada, Lokelma, Lumark, Luxturna, Myalepta, Mylotarg, Nyxoid, Ocrevus, Ogivri, Olumiant, Onpattro, Orkambi, Pelmeg, Pelgraz, Pifeltro, Prevymis, Raxone, Ritemvia, Rubraca, Sivextro, Spinraza, Symtuza, Tagrisso, Trelegy, Trumenba, Venclyxto, Verzenios, Viekirax, Vindaqel, Wakix, Zejula, Ziextenzo y Zykadia.

Una vez que sabíamos qué medicamentos concretos nos interesaban, establecimos dos tipos de texto para cada uno. El primero sería el documento oficial de la UE que plasma la información básica del medicamento, es decir, el documento que incluye la misma información básica que un prospecto farmacéutico (resumen del EPAR para el público general), y, el segundo, la ficha técnica del medicamento, cuya extensión es bastante más considerable que la del primer documento, que suele tener entre 2 y 3 páginas.

Para etiquetar los documentos, empleamos la siguiente fórmula: medicamento_ES, medicamento_ES1, medicamento_EN y medicamento_EN1. Por ejemplo, TRELEGY_EN y TRELEGY_EN1. En este caso, el primer documento es el documento en inglés correspondiente al resumen del EPAR para el público general de Trelegy, mientras que TRELEGY_EN1 es el documento en inglés correspondiente a la ficha técnica del medicamento Trelegy. Lo mismo sucede con los documentos en español que, en este ejemplo concreto se etiquetarían de la siguiente forma: TRELEGY_ES y TRELEGY_ES1.

Tras descargar ambos documentos tanto en inglés como en español en formato PDF, único formato disponible en el sitio web de la EMA, procedimos a comenzar el proceso de alineación, que se expondrá más adelante.

Por otro lado, también se seleccionaron 11 artículos de la agencia estadounidense FDA¹⁰(Food and Drug Administration), organismo encargado de la regulación de los medicamentos en EE. UU. Se descargaron los artículos en inglés junto con sus correspondientes traducciones oficiales al español en formato PDF. Estos artículos son, en realidad, comunicaciones sobre la seguridad de los medicamentos en EE. UU., por lo que la temática resulta ideal para nuestro corpus especializado en el ámbito farmacéutico.

Para la alineación de los documentos escogimos la herramienta Wordfast Anywhere¹¹. Se trata de una herramienta completamente gratuita en línea de Wordfast, un programa de traducción asistida por ordenador, que ofrece la posibilidad de alinear textos automáticamente con el único requisito de crearse una cuenta.

Al disponernos a subir los textos para su correspondiente alineación, nos percatamos de que esta herramienta no acepta el formato PDF, por lo que decidimos que la mejor opción sería convertir todos los documentos a formato DOC. Para ello, utilizamos una página web, también gratuita, que se denomina PDF to DOC¹². Este sitio web es muy fácil de usar, ya que simplemente hay que subir los archivos en PDF y tras unos segundos o minutos, dependiendo del peso de los archivos, están listos para descargar en formato DOC.

Una vez que hubimos descargado los archivos DOC, procedimos a subirlos a la plataforma de Wordfast Anywhere. No es necesario determinar la lengua origen y la lengua de llegada de

¹⁰ FDA's Drug Safety Communications (<https://www.fda.gov/Drugs/DrugSafety/ucm263010.htm>)

¹¹ Alineador de textos de Wordfast Anywhere (<https://www.freetm.com/>)

¹² PDF to DOC (<https://pdf2doc.com/es/>)

los textos, ya que esta herramienta los detecta automáticamente, como podemos observar en la siguiente imagen:

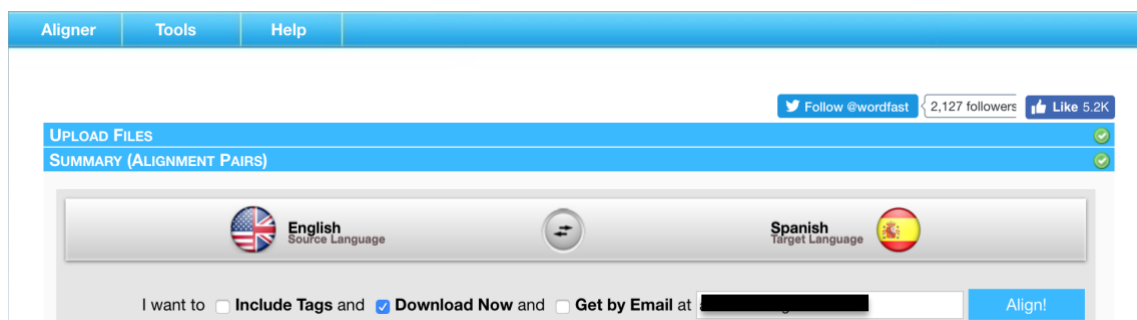


Figura 2. Interfaz del alineador automático del programa TAO en línea Wordfast Anywhere.

Tras detectar las lenguas de los documentos a alinear, Wordfast ofrece la opción de incluir las etiquetas de los documentos y de descargar el archivo alineado en formato TMX de dos formas: directamente desde la página web o desde el correo electrónico.

En nuestro caso, decidimos prescindir de las etiquetas, ya que lo que nos interesa es el contenido del texto, no su forma, y descargamos las alineaciones directamente desde esa misma página web.

El programa, en el peor de los casos, puede tardar un máximo de 10 minutos en tener lista la alineación y, además, tras alinear los archivos nos ofrece un porcentaje que muestra la calidad de la alineación, así como un recuento de las unidades de traducción, tal y como podemos apreciar en la imagen de la página siguiente.

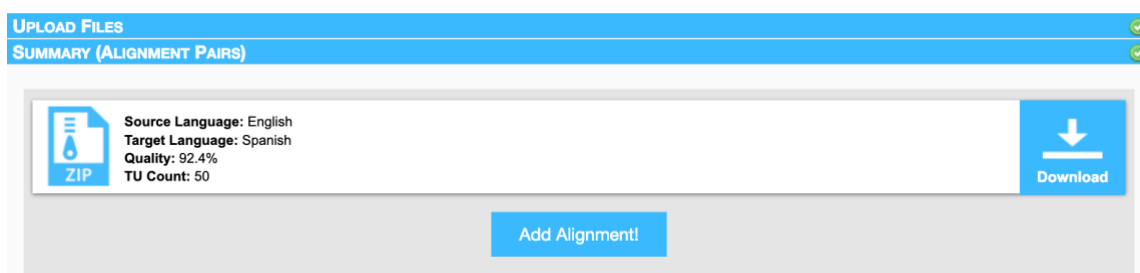


Figura 3. Resultado de la alineación de dos documentos en la herramienta TAO en línea Wordfast Anywhere.

Ha de señalarse que el proceso de alineación de todos los textos fue un poco lento debido a que subimos los archivos por tandas, de dos en dos, ya que se intentó hacerlo con todos los archivos a la vez, pero el programa daba problemas y las alineaciones no presentaban un porcentaje de calidad tan bueno.

A pesar de que las alineaciones automáticas del programa Wordfast tienen muy buena calidad, fue necesario realizar modificaciones. Fue en este escenario cuando entró en juego el *software* TAO memoQ¹³. A pesar de era posible modificar las alineaciones automáticas en Wordfast Anywhere, decidimos cambiar de herramienta a estas alturas del proceso, ya que memoQ ofrece una interfaz más intuitiva y visual que Wordfast y se consideró una gran ventaja el hecho de que memoQ ofrezca la opción de realizar una estadística general que nos permite saber cuántas palabras tienen nuestras memorias de traducción finales.

Por un lado, fue necesario modificar todas las alineaciones de los textos de medicamentos de la página web de la EMA. Sin embargo, en el caso de los artículos de la FDA, al tener la estructura de párrafos exactamente igual y tratarse de documentos no muy extensos, ya que todos cuentan con menos de diez páginas, las alineaciones fueron perfectas. Únicamente se consideró conveniente, por cuestiones de estructura y organización, agrupar los once artículos de la FDA en una sola memoria de traducción que se denominó FDA ARTICLES.

En la siguiente imagen, se puede apreciar en la interfaz de MemoQ un ejemplo de alineación perfecta, la cual corresponde a una muestra del documento ZYKADIA1 y que no hubo necesidad de modificar, ya que los segmentos en la lengua origen (inglés) y en la lengua meta (español) coinciden.

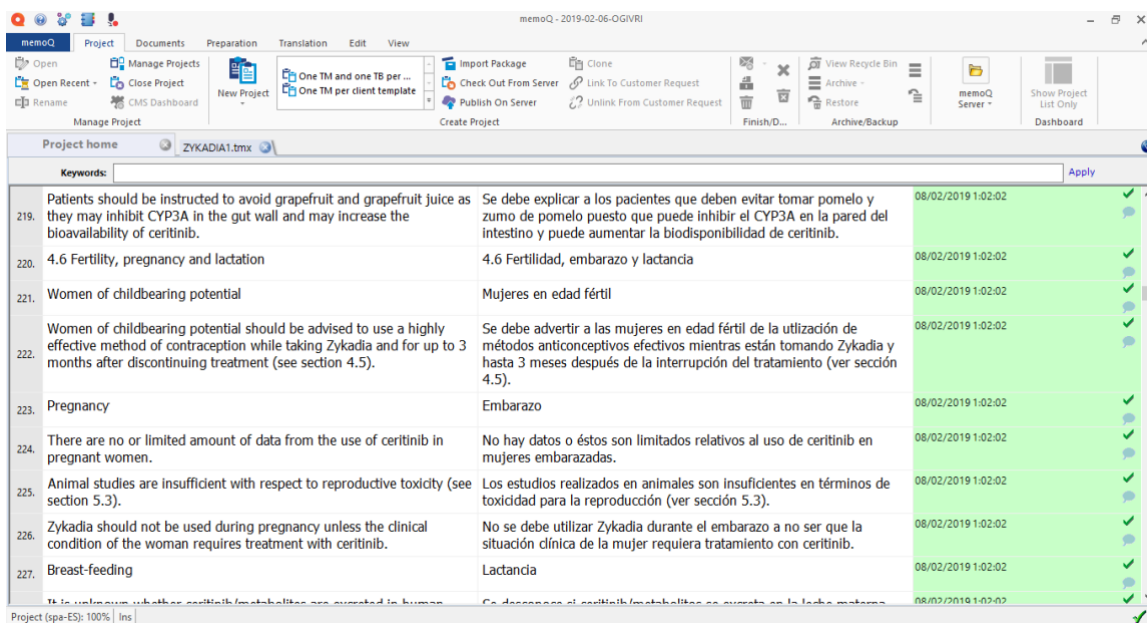


Figura 4. Interfaz de memoQ en la ventana de alineación de textos.

¹³ memoQ (<https://www.memoq.com/es/>)

Sin embargo, en la imagen que se muestra a continuación, aparece otra muestra del documento ZYKADIA1 en la que la alineación sí que tuvo que ser modificada en el segmento 83, ya que ese segmento estaba originalmente dividido dónde aparece el símbolo {1} y, de esa forma, el contenido de la lengua origen no coincidía exactamente con el contenido de la lengua meta.

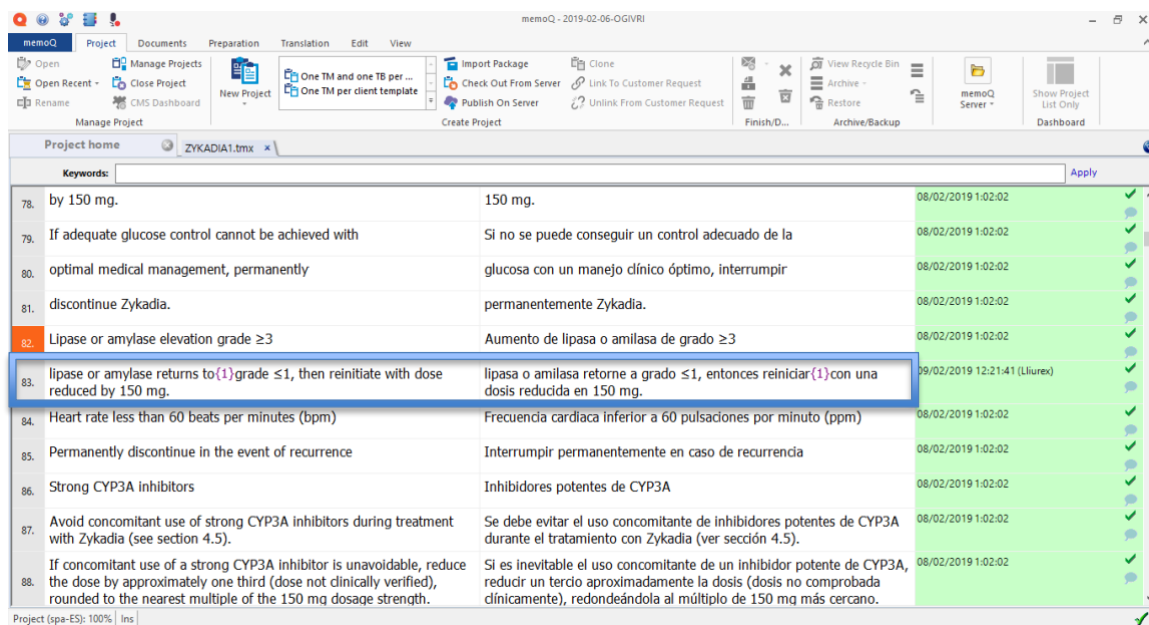


Figura 5. Interfaz de memoQ en la ventana de alineación de textos con un error corregido en la segmentación del texto origen y el texto meta.

Por otro lado, para la creación de la memoria de traducción final en formato TMX que posteriormente emplearemos para el entrenamiento del motor TAE seguimos un mismo proceso que consistió en crear en memoQ un solo corpus por cada medicamento. En cada corpus se añadieron los documentos 1 y 2 del medicamento como archivos pares de alineación y se modificaban los segmentos de las alineaciones en caso de ser necesario. Por ejemplo, creamos el corpus ZYKADIA añadiendo los dos pares de documentos como alineaciones ZYKADIA_EN con ZYKADIA_ES y ZYKADIA_EN1 con ZYKADIA_ES1, como puede apreciarse en la imagen de la siguiente página.

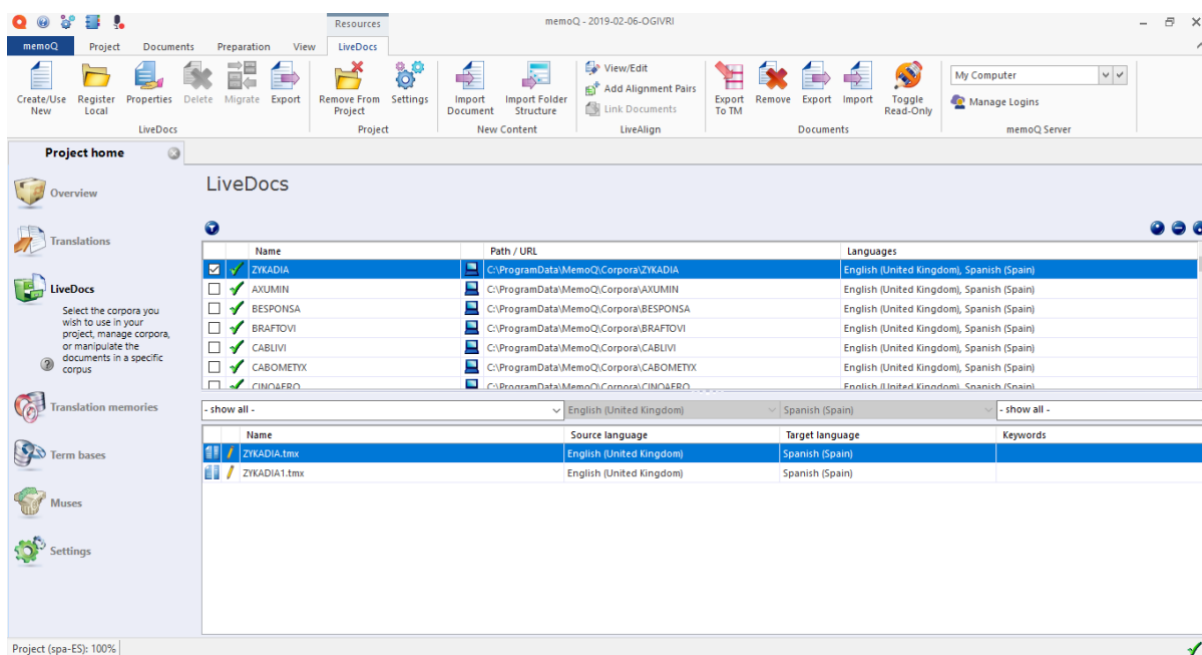


Figura 6. Interfaz de memoQ en la ventana de LiveDocs, la ventana de creación y gestión de corpus.

Una vez realizado este paso, procedimos a la creación de una única memoria de traducción, a la cual se exportarían posteriormente los documentos alineados de todos los corpus, por defecto, en formato mqxliff. Tras crear la memoria de traducción, la cual se denominó PHARMA, procedimos a exportarla en formato TMX. Este paso de creación de una memoria de traducción era estrictamente necesario, ya que KantanMT, la plataforma de entrenamiento de motores de traducción automática que emplearemos posteriormente, no acepta el formato XLIFF, sino TMX.

Adicionalmente, también se llevó a cabo una estadística de memoQ para saber de cuántas palabras constaba nuestra memoria de traducción. Los resultados indican que nuestra memoria de traducción especializada en farmacia está formada por 40 313 segmentos y 501 284 palabras en la lengua origen, así como por 584 595 palabras en la lengua meta.

Seguidamente se procederá a relatar el procedimiento que se ha seguido para crear el corpus monolingüe de nuestro motor de traducción automática.

4.3. Creación de un corpus monolingüe en español

En este apartado, como se ha mencionado anteriormente, se explicará el proceso de creación del corpus monolingüe en español que se empleó en el entrenamiento de nuestro motor TAE.

En primer lugar, lo que más se valoró a la hora de seleccionar las fuentes de información fue la temática especializada y, por lo tanto, nuestro corpus monolingüe está constituido en su

totalidad por textos de temática farmacéutica. Posteriormente, se procedió a escoger las fuentes de información y, para ello, nos servimos de la herramienta en línea Sketch Engine¹⁴. Como gran ventaja de esta herramienta que permite crear corpus en la nube, hemos de destacar, sin duda, la posibilidad que ofrece de crear corpus a partir de la información extraída de páginas web y todo ello gracias a su función WebBootCaT.

Teniendo esto en cuenta, procedimos a la elección de páginas web relacionadas con el ámbito farmacéutico y médico que formarían nuestro corpus monolingüe y finalmente seleccionamos los siguientes sitios web que pueden clasificarse de la siguiente forma:

- **Empresas farmacéuticas:** Novartis y Roche Plus.
- **Organismos de regulación oficiales y sociedades:** Aemps (Agencia Española de Medicamentos y Productos Sanitarios), Healthfinder (página web del Centro Nacional de Información sobre la Salud de EE. UU.), SEFH (Sociedad Española de Farmacia Hospitalaria), NIH (Instituto Nacional de Artritis y Enfermedades Musculoesqueléticas y de la Piel de EE. UU.), NICHD (Instituto Nacional de la Salud Infantil y Desarrollo Humano de EE. UU.), SEFAP (Sociedad Española de Farmacéuticos de Atención Primaria) y SEF (Sociedad Española de Farmacología).
- **Periódicos farmacéuticos en línea y páginas web de farmacias:** El Periódico de la Farmacia, im Farmacias y El farmacéutico.
- **Enciclopedias y vademécums:** enciclopedia médica en línea MedlinePlus y un vademécum en línea de medicamentos denominado Vademecum.

Sketch Engine es una herramienta con una interfaz muy intuitiva. A continuación, se muestra una imagen con la función de WebBootCaT:

¹⁴ Sketch Engine (<https://www.sketchengine.eu/>)

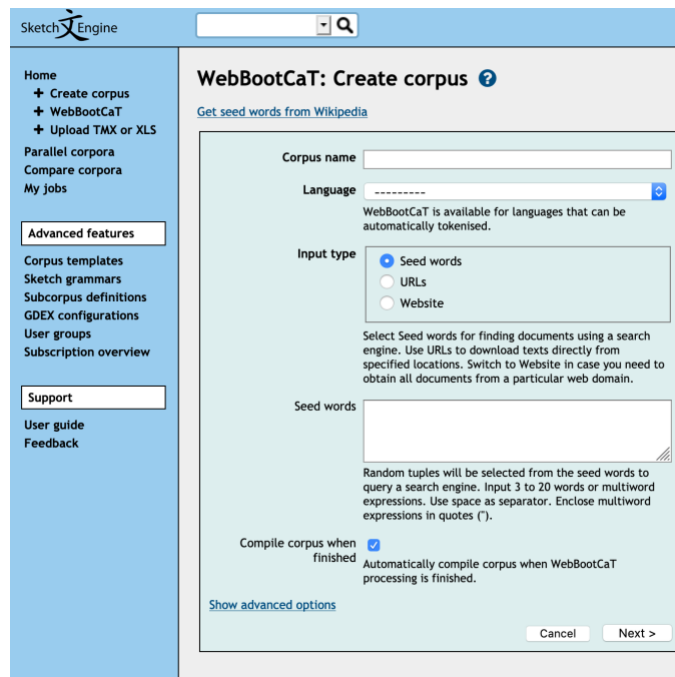


Figura 7. Interfaz de la herramienta en línea Sketch Engine y su función WebBootCaT para crear un corpus a base de páginas web.

Como podemos apreciar en la imagen anterior, la función WebBootCaT permite elegir el idioma del corpus que vamos a crear y también podemos escoger a partir de qué fuente de información queremos que se cree el corpus. En nuestro caso, tendríamos que seleccionar la opción «Website».

Debido a la restricción impuesta a las cuentas académicas de Sketch Engine, solo pueden compilarse un máximo de 1 000 000 palabras entre todos los corpus, por lo que se decidió ir creando un corpus por cada página web, descargarlos en formato de texto simple e ir eliminándolos de la plataforma para poder seguir compilando nuevos corpus.

Una vez tuvimos todos los corpus en formato TXT, ejecutamos un comando en la ventana del terminal para unir todos los corpus monolingües y crear un único TXT. Esta acción se realizó en un ordenador con sistema operativo Mac OS, por lo que el comando aquí empleado difiere del de otros sistemas operativos como podría ser el de Windows.

El comando consiste en escribir «cat» en el terminal e ir arrastrando los archivos a la ventana del terminal, lo que hace que se copie la ruta del documento. En nuestro caso, los documentos se encontraban en una carpeta del escritorio denominada «CORPUS ESPAÑOL MONOLINGÜE» y, seguidamente, escribimos el símbolo «>» seguido del nombre que se le quiere dar al archivo final. El nombre escogido para nuestro corpus monolingüe fue

source.utf8.trg.mono, denominación concienzuda que se explicará en apartados posteriores y que se puede apreciar en la siguiente imagen:

```

Desktop -- -bash -- 80x24
Last login: Mon Feb 25 18:08:45 on ttys000
AliciaMB$ cd /Users/AliciaMB/Desktop
AliciaMB$ cat /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/Aemps.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/Elfarmaceutico.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/Elperiodicodelafarmacia.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/healthfinder.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/medlineplus.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/nih1.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/nih2.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/novartis.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/rocheplus.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/sef.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/sefap.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/sefh.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/vademecum_medicamentos.txt /Users/AliciaMB/Desktop/CORPUS\ ESPAÑOL\ MONOLINGÜE/web_farmacia.txt > source.utf8.trg.mono.txt
AliciaMB$

```

Figura 8. Ejecución del comando de unión de archivos TXT en la ventana de terminal de Mac OS.

Por último, tras compilar el corpus monolingüe final en un solo archivo, se procedió a realizar un análisis estadístico del mismo con memoQ, del que obtuvimos como resultado que dicho corpus consta de 1 238 443 segmentos y 14 024 114 palabras.

A continuación, se muestran unas tablas que recogen, a modo de resumen, todos los datos relevantes empleados para el entrenamiento del motor TAE, datos que han sido extraídos de los análisis de los corpus realizados en memoQ y que podemos encontrar en los **anexos**.

Modelo de traducción

Corpus bilingües (EN > ES)	Fuente	Número de palabras del archivo original (Source Word Count)	Número de palabras del archivo traducido (Target Word Count)
PHARMA	Corpus propio con textos procedentes de la web de la EMA	501 284	584 595
EMEA	OPUS (textos recopilados de la web de la EMA)	5 302 365	5 983 422
ECDC	UE	36 813	44 244

Tabla 1. Datos de entrenamiento del modelo de traducción de nuestro motor TAE.

<u>Modelo de lengua</u>		
Corpus monolingüe (ES)	Segmentos	Palabras
source.utf8.trg.mono	1 238 443	14 024 114

Tabla 2. Datos de entrenamiento del modelo de lengua de nuestro motor TAE.

En el apartado posterior se explicará el proceso de entrenamiento del motor de traducción automática estadística que nos ocupa en este trabajo.

4.4. Entrenamiento de un motor de traducción automática estadística

Ahora procederemos a explicar cómo hemos creado el motor de traducción automática estadística en la plataforma de KantanMT con todos los corpus y datos recopilados y desmenuzados en los apartados anteriores.

El primer paso es escoger el nombre que queremos darle al motor y seguidamente hemos de seleccionar la combinación lingüística. En nuestro caso, la lengua de partida es inglés de Reino Unido, mientras que la lengua de llegada es español de España. Esta elección de variedades lingüísticas está justificada por el hecho de que la mayor parte de los datos usados para entrenar el motor provienen de textos de la Unión Europea, de ahí que se hayan querido utilizar las variedades europeas de estas dos lenguas.

En estos primeros pasos de la creación de un motor, KantanMT ofrece la posibilidad de elegir de antemano si queremos activar alguna de las bibliotecas o repositorios que almacenan en su base de datos. En este caso particular y, como ya vimos con anterioridad, nuestro motor se ha valido de la utilidad del corpus EN>ES de KantanMT especializado en medicina, por lo que marcamos esa opción. Después de todos estos pasos, el motor ya estaría creado y activo.

Seguidamente, hemos de añadir todos los datos y corpus que emplearemos para entrenar dicho motor. A continuación, se mostrará una imagen en la que se puede apreciar la interfaz de KantanMT en la pestaña de entrenamiento de nuestro motor personalizado.

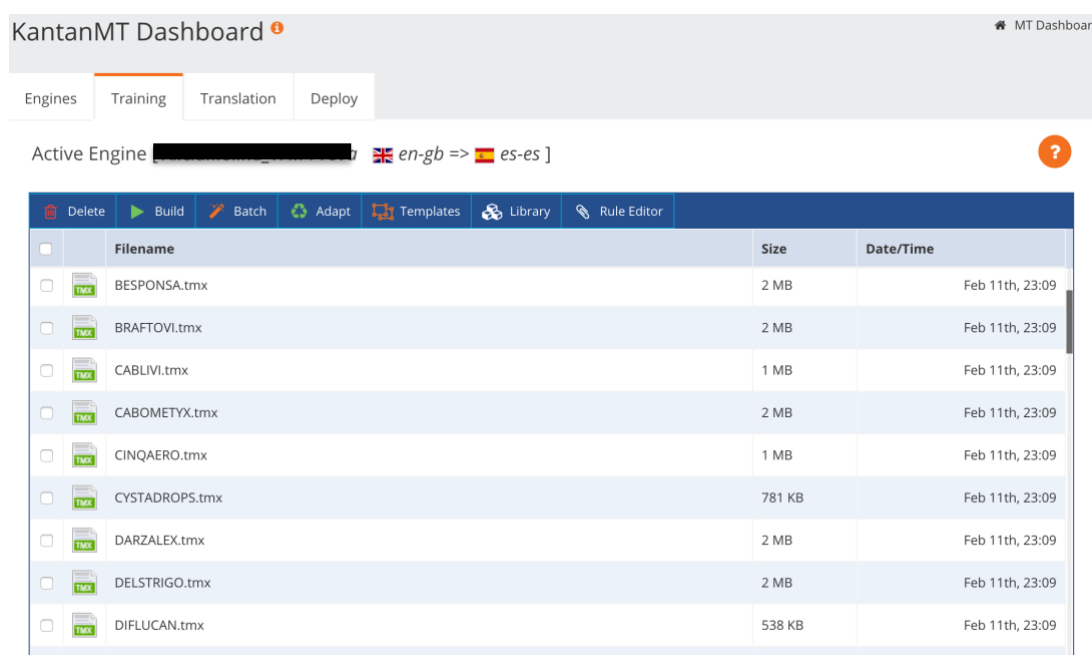


Figura 9. Vista de la pestaña de entrenamiento de KantanMT del motor TAE especializado en farmacia de este trabajo de fin de máster.

Para subir los archivos a la plataforma, tan solo hemos de arrastrarlos. En esta etapa del proceso, procedimos a subir todos los corpus que analizamos en apartados anteriores:

- Corpus bilingües: PHARMA, ECDC y EMEA en formato TMX
- Corpus monolingüe: source.utf8.trg.mono en formato TXT

En el caso del corpus monolingüe, fue necesario darle ese nombre debido a que KantanMT no lo reconocerá como tal y, por lo tanto, no lo procesará, si no tiene esa denominación. En este paso es muy importante asegurarse de que el nombre del archivo no contiene la extensión, ya que, en tal caso, la plataforma tampoco será capaz de detectarlo como corpus monolingüe. Asimismo, también es primordial que el archivo esté codificado en UTF-8.

En este paso también se pueden incluir glosarios o bases de datos terminológicas tanto en formato TBX como XLSX, aunque en nuestro caso no nos servimos de ninguno de estos recursos.

Tras subir los archivos, hemos de clicar en «Build» para que la plataforma comience a entrenar el motor. Este proceso puede tardar más o menos tiempo en función de la cantidad de información involucrada. En este caso particular, el motor estuvo listo en menos de 24 horas.

A continuación, nos dispondremos a comentar las puntuaciones que obtuvo el motor TAE según los análisis automáticos de KantanMT y realizaremos un análisis en profundidad de todas las variables implicadas basándonos en las referencias de calidad orientativas que proporciona KantanMT.

Antes de comenzar, conviene señalar que las variables que se tendrán en cuenta son las siguientes: F-Measure, BLEU, TER, el recuento de fuentes bilingües, el de palabras únicas y el número de palabras comprendidas en el corpus monolingüe.

En primer lugar, analizaremos la puntuación de F-Measure que, en el caso del motor que aquí nos ocupa, es de 78 %. Esto significa que nuestro motor cuenta con un buen potencial para recordar términos de manera precisa durante el proceso de traducción. Según los estándares de KantanMT, un buen motor de traducción automática tiene que contar al menos con un porcentaje igual o superior al 70 % en la medida de F-Measure si pretende alcanzar una buena calidad en sus traducciones.

A continuación, podemos ver una imagen extraída de la plataforma de KantanMT en la que aparece el porcentaje de F-Measure de nuestro motor TAE.

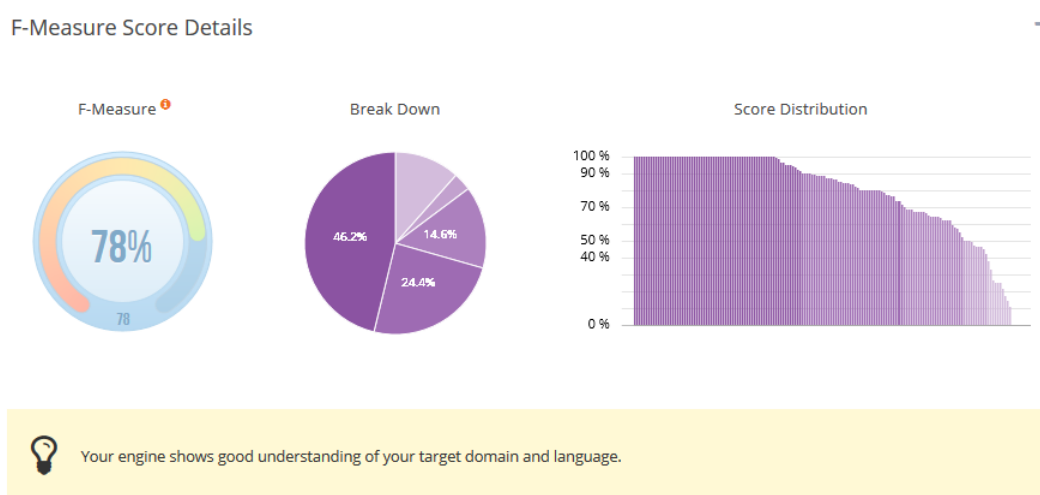


Figura 10. F-Measure del motor TAE entrenado para este trabajo de fin de máster.

En segundo lugar, nos ocuparemos de la puntuación BLEU, una de las más significativas. En cuanto a nuestro motor, se ha obtenido un porcentaje del 73 %, lo que significa que vamos a obtener resultados de traducción fluidos en la lengua meta, es decir, en español.

Según los estándares de KantanMT, un motor TA debe tener al menos una puntuación igual o superior al 60 % para considerarse efectivo y de buena calidad. Por lo tanto, la puntuación de nuestro motor es buena, aunque esto no impide que, a pesar de ello, produzca traducciones con errores gramaticales o relacionados con la terminología en ciertas ocasiones.

En la siguiente imagen, se pueden apreciar los gráficos extraídos de la plataforma de KantanMT en la que se refleja el porcentaje de BLEU de nuestro motor TAE.

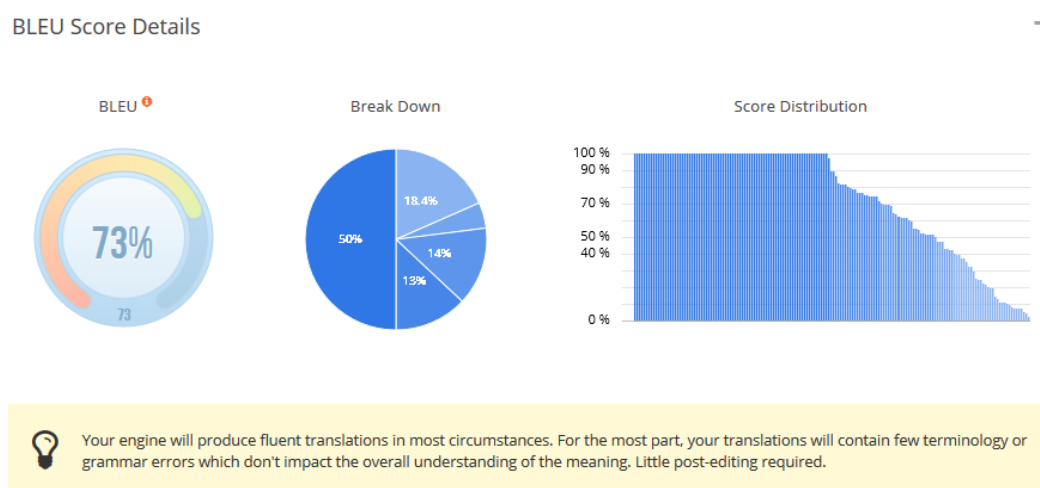


Figura 11. Puntuación BLEU del motor TAE entrenado para este trabajo de fin de máster.

En tercer lugar, se discutirá la puntuación TER, que mide la cantidad de posesición que necesitará la traducción que resulte del motor TA. En realidad, mide cuánta posesición necesitará para que coincida exactamente con la traducción que toma como referencia, tal y como mencionamos en el marco teórico de este trabajo.

En nuestro caso particular, se ha obtenido 34 %, el cual, podemos afirmar, es un buen porcentaje, ya que según los estándares de KantanMT, un motor TA efectivo ha de tener un porcentaje igual o inferior al 40 % de puntuación TER.

A continuación, aparece una imagen que ha sido extraída de la plataforma de KantanMT y en la que podemos observar el porcentaje TER de nuestro motor TAE.

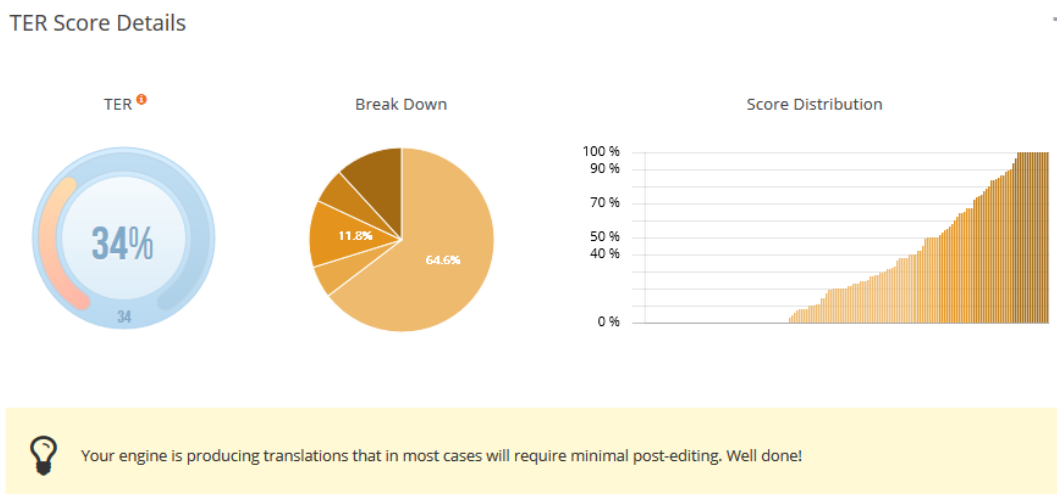


Figura 12. Puntuación TER del motor TAE entrenado para este trabajo de fin de máster.

Por último, comentaremos el análisis del recuento de fuentes bilingües, de palabras únicas y de palabras comprendidas en el corpus monolingüe.

En cuanto al recuento de palabras procedentes de fuentes bilingües, nuestro motor cuenta con 10 millones de palabras aproximadamente. Habría que tener en cuenta que en este recuento van incluidas únicamente las palabras comprendidas en los segmentos de la lengua de origen. Según los estándares de KantanMT, son necesarias al menos 5 millones de palabras para contar con un motor de buena calidad y nuestro motor dobla dicha cantidad. Sin embargo, la calidad de un motor TA está más determinada por la variable que vamos a analizar a continuación: el recuento de palabras únicas.

En lo referente al recuento de palabras únicas, nuestro motor cuenta con 122 874 palabras. Este es el número de palabras únicas que nuestro motor TAE puede procesar y determina la cantidad de vocabulario con el que cuenta el mismo.

Según las referencias de KantanMT, se trata de un motor especializado en un tema concreto, basta con 300 000 palabras para lograr una buena calidad, aunque si se trata de motores de carácter general, puede que sean necesarias 500 000 o más palabras. En este sentido, se puede afirmar que nuestro motor no alcanzaría el mínimo de esta variable según los estándares de KantanMT, lo que se traduce en una falta de vocabulario por parte de nuestro motor. Esto significa que es probable que, a la hora de traducir un texto, encontremos palabras sin traducir, ya que el motor las desconoce.

Por otra parte, el recuento de palabras monolingües de nuestro motor, extraído de nuestro corpus monolingüe en español, sería de 9 millones aproximadamente. De acuerdo con los estándares de KantanMT, son necesarias entre 2 y 3 millones de palabras para que un motor TA

tenga fluidez a la hora de traducir. En nuestro caso particular, se cumple con mucho margen este umbral de calidad.

Posteriormente puede visualizarse la imagen extraída de la plataforma de KantanMT, la cual muestra la relación de porcentajes y las cifras de las tres variables discutidas en los párrafos anteriores.

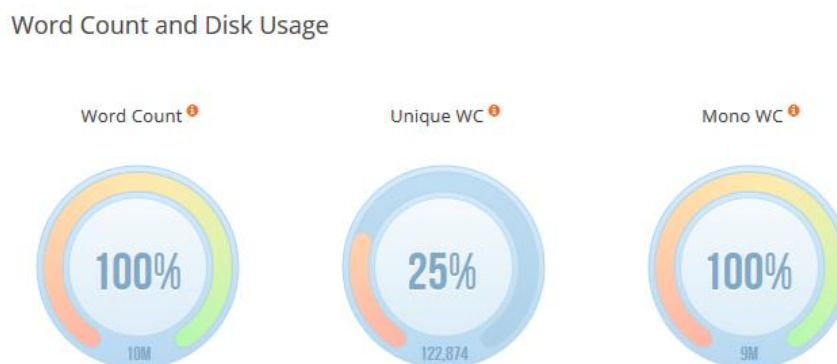


Figura 13. Recuento de palabras del motor TAE entrenado para este trabajo de fin de máster.

En el próximo apartado, seremos testigos de la ampliación de los archivos bilingües de nuestro motor, con una consecuente mejora del mismo, gracias a la colaboración con TAUS¹⁵, que proporcionó un corpus especializado en la combinación de lenguas EN > ES.

4.5. Implementación de un corpus bilingüe adicional

En esta sección del trabajo, se hablará de la colaboración que se estableció con TAUS para mejorar la calidad del motor de este estudio.

Ante de comenzar con los detalles de esta colaboración, sería interesante aclarar qué es TAUS y por qué nos ha cedido un corpus para este trabajo académico.

TAUS es una organización que proporciona servicios de datos a compradores y proveedores de servicios lingüísticos y de traducción y que, además, cuenta con múltiples colaboraciones con diferentes empresas de la industria de la traducción. Asimismo, uno de sus servicios principales consiste en recolectar, seleccionar y proporcionar datos pertinentes que los proveedores lingüísticos puedan aprovechar para mejorar sus motores de traducción automática.

En nuestro caso, TAUS decidió apoyar el desarrollo de este trabajo académico cediéndonos un corpus que crearon especialmente adaptado a nuestras necesidades a partir de segmentos de

¹⁵ TAUS (<https://md.taus.net/corpus.php>)

su base de datos. Este corpus cuenta con 79 502 segmentos, con 1 160 584 palabras en la lengua de origen, es decir, en inglés y con 1 309 024 palabras en la lengua de destino, es decir, en español.

En cuanto al proceso que se siguió, en primer lugar, nos pusimos en contacto con TAUS, les informamos del tipo de datos que necesitábamos, textos especializados en el ámbito farmacéutico y médico, la combinación lingüística que estábamos buscando (EN>ES) y, por último, le proporcionamos nuestro corpus PHARMA.

TAUS, teniendo todos estos factores en cuenta, procedió al cotejo de datos. Tomando como referencia nuestro corpus PHARMA, esta organización procedió a identificar los mejores datos coincidentes, a nivel de segmento, de su nube de datos y con todos esos datos creó un corpus de las dimensiones que hemos mencionado en párrafos anteriores.

La nube de datos de TAUS está formada por las memorias de traducción que han ido aportando sus miembros y usuarios durante los últimos diez años y por datos públicos puestos a su disposición por parte de grandes organizaciones e instituciones.

TAUS se encarga de «limpiar» y conservar esos datos para asegurarse de que sean de buena calidad. Cuando TAUS realiza la comparación de datos para crear un nuevo corpus, se lleva a cabo una búsqueda semántica y se busca en todos los dominios y documentos para obtener los segmentos que estén «más cerca» de las propiedades semánticas del corpus que le hemos proporcionado.

Este enfoque permitió que pudiéramos recibir datos de calidad y relevancia para el dominio concreto de nuestro motor TAE, el farmacéutico, lo que nos permite afinar el proceso de entrenamiento del mismo.

Ahora se procederá a realizar un análisis más exhaustivo del tipo de datos que contiene el corpus de TAUS.

En cuanto a la procedencia concreta de este corpus, podemos señalar que los proveedores principales de datos han sido Alibaba, Michael Beijer y la empresa de traducción Pangeanic, aunque la mayoría de los datos provienen de los dos primeros.

A continuación, hemos agrupado una selección representativa de los segmentos en lengua origen según su proveedor:

- Alibaba:
 - *Determination of the pH.*
 - *Children: 18 mg per daily dose as recommended by the manufacturer*
 - *Normally a single dose will be administered once daily; where for example a chemical is a local irritant, it may be possible to maintain the daily dose-rate by administering it as a split dose (twice a day).*
 - *SUMMARY OF PRODUCT CHARACTERISTICS, LABELLING AND PACKAGE LEAFLET.*
 - *Virus strain of vaccine.*
- Michael Beijer:
 - *1,38 mg astaxanthin dimethyl disuccinate is equivalent to 1 mg astaxanthin.*
 - *Recommended doses.*
 - *Always consult your doctor to make sure that this restraint system is suitable for your child.*
 - *Doctor currently treating the patient ...*
 - *Microporous polypropylene film of a thickness of not more than 100 µm.*
- Pangeanic:
 - *At Week 1, median change from baseline was 34% for adult patients treated with ELIDEL compared with 4% for patients treated with placebo cream.*
 - *At Day 8, median change in EASI from baseline was 47% for pediatric ELIDEL patients compared with 2% for placebo cream patients.*
 - *Can you tell me why it is that you think that this product will not meet your needs?*
 - *If the connection is successful, you will see “Disconnect” button instead of “Connect” and you own user name under the buddy list.*

Tras analizar los datos del corpus, podemos afirmar que tanto los segmentos de Alibaba como los de Michael Beijer son en su mayoría relevantes para el ámbito farmacéutico y los que no están directamente relacionados con este son técnicos, por lo que se trata de datos igualmente

aprovechables, ya que la tipología de textos a la que está destinado nuestro motor emplea mucha terminología técnica e incluso reglamentos jurídicos de la UE.

Sin embargo, los procedentes de Pangeanic, en su mayoría, están más relacionados con el *marketing* y la industria electrónica. Un claro ejemplo de ello serían los dos últimos segmentos que aparecen en la enumeración anterior. Habría de señalarse también que la mayoría de los segmentos procede de Alibaba y son los más pertinentes para mejorar nuestro motor TAE especializado en farmacia.

A continuación, nos dispondremos a realizar la comparación entre el primer motor que creamos con los corpus públicos de la UE y el nuestro, PHARMA y el motor final con la adición del corpus de TAUS. Asimismo, conviene señalar que tendremos en cuenta las variables de medición de calidad del apartado anterior.

En primer lugar, nos ocuparemos de la puntuación otorgada para la medida F-Measure que, en el caso del motor con la adición del corpus de TAUS, es de 78 %, por lo que podemos afirmar que este porcentaje no ha experimentado ningún cambio en comparación con el que obtuvimos en el primer entrenamiento de nuestro motor.

A continuación, aparece una imagen que ha sido extraída de la plataforma de KantanMT y en la que podemos observar la puntuación F-Measure de nuestro motor TAE definitivo.

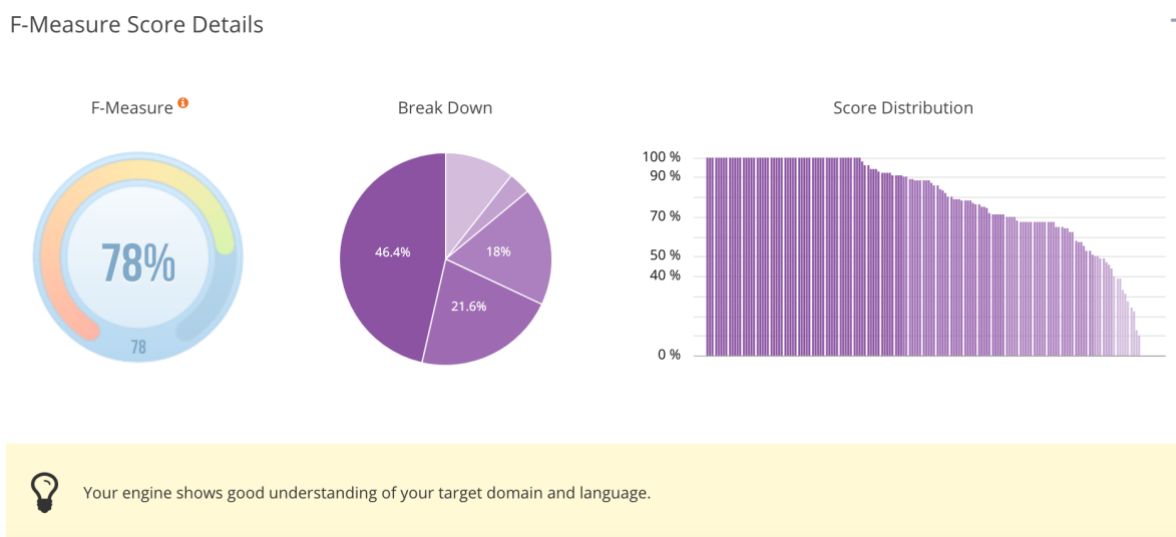


Figura 14. F-Measure del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.

En segundo lugar, analizaremos la puntuación BLEU. En cuanto a nuestro motor, se ha obtenido un porcentaje del 70 %, lo que significa que vamos a obtener resultados de traducción fluidos en la lengua meta.

A diferencia del porcentaje de F-Measure, la puntuación BLEU sí que ha sido diferente de la de obtenida en el primer entrenamiento del motor sin el corpus de TAUS. En dicho primer entrenamiento, nuestro motor obtuvo un porcentaje del 73 %, mientras que, en este segundo entrenamiento, que incluye la adición del corpus proporcionado por TAUS, se ha obtenido un 70 %, es decir, la cifra ha bajado un 3 %.

Siguiendo esta línea de discusión, ha de subrayarse que la obtención de esta diferencia negativa no tiene por qué significar una disminución de la calidad de las traducciones de nuestro motor. De hecho, Way (2018) ya señaló que las puntuaciones BLEU más altas suelen ser calificadas con puntuaciones más bajas por parte de traductores humanos y, aunque el análisis de los resultados sea lento, la evaluación humana de la producción de la TA sigue siendo crucial para mejorar los motores. Por lo tanto, para cerciorarse de la mejora real, o no, del motor habría que realizar una comparación entre el motor sin el corpus de TAUS y con el corpus mediante posesición humana. Sin embargo, el objetivo de este trabajo de investigación no es tal.

Además, KantanMT afirma que la adición de material bilingüe ayuda a mejorar la calidad del motor, siempre que sea de buena calidad y, en este caso, hemos comprobado que el corpus de TAUS se adapta a las necesidades de nuestro motor, tanto en calidad como en especialización. Por lo tanto, este porcentaje no tendría por qué traducirse en una calidad inferior en este caso concreto, aunque, como hemos señalado anteriormente, sería necesario realizar un experimento con traductores humanos para poder evaluar la calidad real del motor.

En la siguiente página, podemos ver una imagen extraída de la plataforma de KantanMT en la que aparece el porcentaje BLEU de nuestro motor TAE tras añadir el corpus de TAUS.

BLEU Score Details

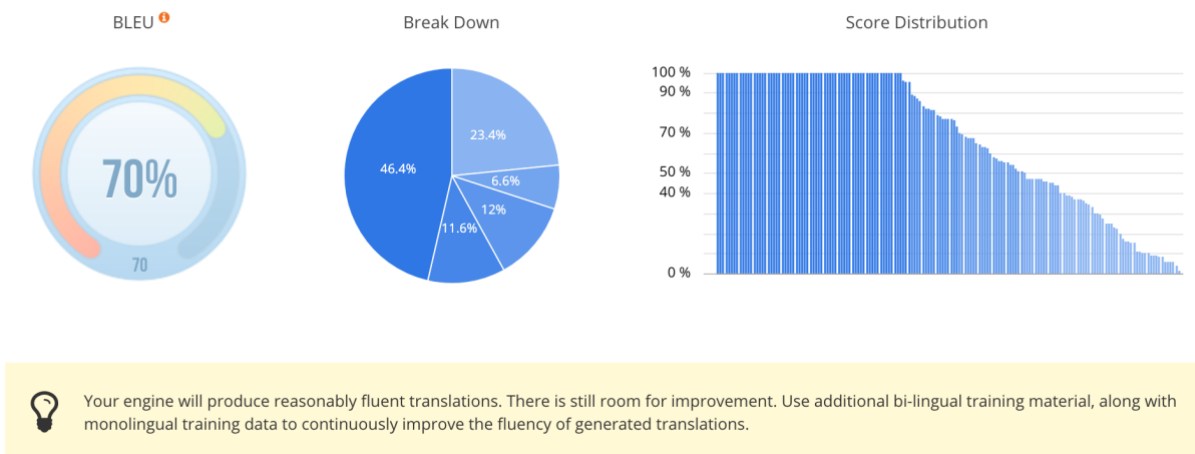


Figura 15. Puntuación BLEU del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.

En tercer lugar, encontramos la puntuación TER, que, tras la adición del corpus de TAUS, la cifra ha descendido a un 36 %. Se ha experimentado únicamente una diferencia negativa del 2 % en comparación con el primer entrenamiento del motor, es decir, nada significativo. Esta disminución del 2 % está relacionada con la disminución en la puntuación BLEU, ya que ambas variables están directamente relacionadas.

Seguidamente se puede apreciar la imagen extraída de la plataforma de KantanMT en la que se refleja el porcentaje de TER de nuestro motor TAE definitivo.

TER Score Details

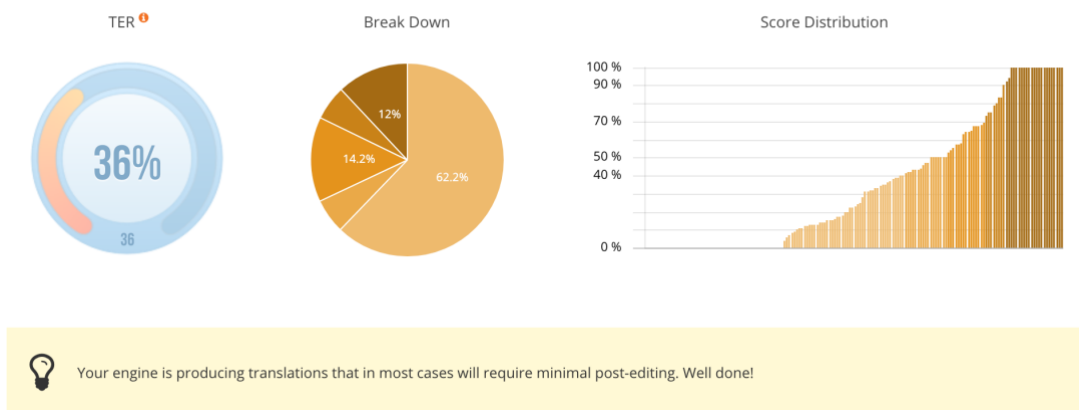


Figura 16. Puntuación TER del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.

En cuanto a los recuentos de palabras, los únicos valores que han cambiado han sido el recuento de palabras bilingües totales y el de palabras únicas. Las palabras totales han pasado de ser 10 millones a ser 11 tras la adición del corpus de TAUS, mientras que el recuento de palabras únicas ha pasado de 122 874 a 167 221.

Esto significa que dichos valores, a pesar de haber aumentado más de 40 000 palabras, nuestro motor seguiría estando por debajo de los estándares de KantanMT en cuanto a conocimiento de vocabulario, estándares que se sitúan en un umbral mínimo entre 300 000 y 500 000 como mínimo. Sin embargo, se trata de un incremento positivo, ya que cuanto más vocabulario tenga a su disposición el motor, más comprensivo será este. Además, esto provocará una disminución de las situaciones OOV (fuera de vocabulario, en inglés *Out of Vocabulary*), situaciones que se producen cuando el motor no es capaz de reconocer una palabra.

Ha de señalarse que también, al tratarse de un motor TA especializado en un ámbito concreto, resulta más complicado alcanzar una buena cantidad de palabras únicas que si se tratara de un motor TA general, ya que aquí hablamos de vocabulario técnico y específico y, por lo tanto, más reducido, y no de vocabulario básico general.

En la siguiente imagen, extraída de la plataforma de KantanMT, se puede apreciar la relación de porcentajes y cifras de las variables discutidas anteriormente.

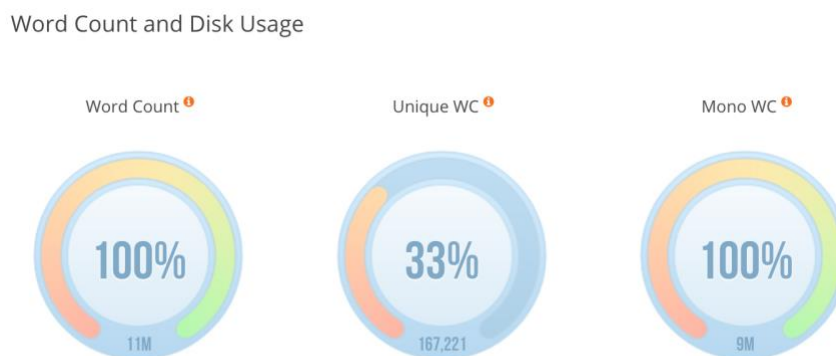


Figura 17. Recuento de palabras del motor TAE con la adición del corpus de TAUS entrenado para este trabajo de fin de máster.

4.6. Comparación de resultados con dos motores de traducción automática neuronal: Google Translate y DeepL

En este apartado nos dispondremos a realizar una valoración de la calidad de nuestro motor de traducción definitivo en el marco comparativo, motor que incluye, como hemos mencionado antes, el corpus de TAUS. Para ello, nos serviremos de la plataforma en línea Tilde MT¹⁶. Esa plataforma, al igual que KantanMT, ofrece a los usuarios la opción de entrenar sus propios sistemas de traducción automática, entre otros servicios. En nuestro caso, nos interesa una de sus funcionalidades gratuitas en línea. Se trata de un evaluador automático de motores TA denominado Interactive BLEU Score evaluator, basado en la puntuación BLEU.

Para realizar esta evaluación automática, el sistema se sirve de un texto de muestra traducido con los motores que se quieren comparar, a partir del cual determina una puntuación BLEU para cada motor.

En primer lugar, escogimos un texto de muestra para poder realizar dicha comparación. Seleccionamos un texto de la página web de la EMA que no estuviera en ninguno de los corpus que empleamos para entrenar nuestro motor TAE. El texto escogido fue un resumen del EPAR para el público general del fármaco AFSTYLA, texto compuesto de 941 palabras y 61 frases. Escogido el texto, se procedió a la descarga en PDF del mismo, en versión original, y de la traducción oficial en español. Posteriormente, realizamos su conversión al formato TXT, ya que es el único formato aceptado por la plataforma Tilde MT para realizar la evaluación automática de los motores TA. Esta conversión de formato se realizó con una herramienta en línea gratuita denominada PDF2Go¹⁷.

A continuación, procedimos a traducir dicho texto con todos los motores de traducción automática implicados en este trabajo académico: nuestro motor TAE, Google Translate y DeepL. Una vez hubimos traducido los textos, disponibles en los **anexos**, nos dispusimos a subirlos a Tilde MT. La vista general de la plataforma para comparar motores se puede apreciar en la imagen de la página siguiente.

¹⁶ Interactive BLEU Score evaluator de Tilde MT (<https://www.letsmt.eu/Bleu.aspx>)

¹⁷ PDF2Go (<https://www.pdf2go.com/es/pdf-a-text>)

Interactive BLEU score evaluator

Perform comparative quality evaluations of files translated with one or more MT systems. This allows you to compare MT output with human translations and compare the BLEU scores of various MT systems. [Click here to learn more.](#)

Step 0: Pick source file (Optional) No file selected. .txt

Step 1: Pick human translated file No file selected. .txt

Step 2: Pick machine translated file No file selected. .txt

Step 3: Pick second machine translated file (Optional) No file selected. .txt

Calculate BLEU Display

Lowercase Tokenized Difference highlighting

Figura 18. Vista del comparador automático de sistemas TA de la plataforma Tilde MT, basado en la puntuación BLEU.

En la imagen anterior podemos observar que existe la opción de subir cuatro archivos en formato TXT. En el primero de ellos, correspondiente a *Step 0* en la imagen, podemos subir el archivo original, que en nuestro caso sería el texto de AFSTYLA en inglés, aunque este paso es opcional. En el segundo, correspondiente a *Step 1*, debemos subir la traducción oficial al español del texto anterior, traducción que previamente nos habíamos descargado de la página web de la EMA. En el tercero, correspondiente a *Step 2*, subimos la traducción del texto original AFSTYLA del motor de traducción automática que queramos comparar. En el cuarto y último, correspondiente a *Step 3*, subimos también la traducción del texto original AFSTYLA de otro motor de traducción automática.

En este caso particular, se decidió comparar, en primer lugar, el motor TAN genérico Google Translate con nuestro motor TAE especializado en farmacia y, en segundo lugar, el motor TAN genérico DeepL también con nuestro motor. Así seremos capaces de observar si hay diferencias significativas en la calidad de los motores para este texto en concreto, texto escogido concienzudamente basándonos en el tipo de textos que comprenden los corpus con los que entrenamos nuestro motor TAE.

Sería conveniente subrayar que, en los **anexos**, podemos encontrar los informes comparativos de BLEU completos extraídos de la plataforma de Tilde MT, ya que en los siguientes apartados solo compararemos un par de ejemplos.

En el siguiente subapartado, analizaremos con detalle los resultados obtenidos de la primera comparación de motores.

4.6.1. Comparación con Google Translate

En esta primera comparación, representada en la imagen que aparece a continuación, la puntuación BLEU en azul corresponde al motor TAN genérico Google Translate, mientras que la puntuación en verde es la asociada a nuestro motor TAE especializado.

BLEU:	51.58					53.97			
Precision x brevity:	51.58 x 100.00					53.97 x 100.00			
Type	1-gram	2-gram	3-gram	4-gram		1-gram	2-gram	3-gram	4-gram
Individual	74.28	56.29	45.54	37.17		76.42	57.89	47.71	40.20
Cumulative	74.28	64.66	57.53	51.58		76.42	66.51	59.54	53.97

Figura 19. Plataforma de Interactive BLEU score evaluator de Tilde MT que compara las puntuaciones BLEU de Google Translate, en azul, y de nuestro motor TAE, en verde.

En este caso, podemos afirmar que nuestro motor recibe una puntuación BLEU más alta que Google Translate, aunque la diferencia no es muy grande, ya que se trata de poco más de dos puntos: 51,58/100 (Google Translate) vs. 53,97/100 (nuestro motor).

Por lo tanto, según la evaluación automática de Tilde MT, nuestro motor TAE especializado produce mejores traducciones que el motor TAN genérico de Google Translate cuando se trata de un texto farmacéutico.

Asimismo, con la herramienta BLEU interactiva de Tilde MT, también permite ver los resultados de la comparación de la evaluación de sistemas TA segmento a segmento. Para ilustrar esto, hemos escogido una de las oraciones, la nueve, tal y como podemos apreciar en la imagen de la siguiente página. Esta funcionalidad de Tilde MT nos permite comparar la puntuación BLEU que asigna dicha plataforma a cada uno de los segmentos traducidos por los dos sistemas TA tomando como referencia la traducción humana oficial.

Por otro lado, es necesario añadir que, en la siguiente imagen, las palabras que aparecen en rojo serían las que difieren de la traducción oficial de la UE, las que están destacadas en color azul son los aciertos de Google Translate que nuestro motor TAE ha fallado, mientras que las verdes, al contrario, representan los aciertos de nuestro motor TAE que Google Translate no ha acertado, todo ello tomando siempre como referencia la traducción oficial.

Sentence 9	BLEU	Length ratio	Text
Source	-	-	This is a summary of the European public assessment report (EPAR) for Afstyla. It explains how the Agency assessed the medicine to recommend its authorisation in the EU and its conditions of use. It is not intended to provide practical advice on how to use Afstyla.
Human	100.00	1.00	El presente documento resume el Informe Público Europeo de Evaluación (EPAR) de Afstyla . En él se explica cómo la Agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la UE y sus condiciones de uso . No está destinado a proporcionar consejos prácticos sobre cómo utilizar Afstyla .
Machine	30.26	0.78	Este es un resumen del informe de evaluación pública europea (EPAR) para Afstyla . Explica cómo la Agencia evaluó el medicamento para recomendar su autorización en la UE y sus condiciones de uso . No pretende proporcionar consejos prácticos sobre cómo usar Afstyla .
Machine	90.78	1.00	El presente documento resume el Informe Público Europeo de Evaluación (EPAR) para Symtuza . En él se explica cómo la Agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la UE y sus condiciones de uso . No está destinado a proporcionar consejos prácticos sobre cómo utilizar Axumin .

Figura 20. Oración 9 del texto AFSTYLA comparada con la traducción humana, la traducción automática de Google Translate (Machine en azul) y la de nuestro motor TAE (Machine en verde).

En el ejemplo de la imagen anterior, podemos apreciar que nuestro motor ha producido una mejor traducción para ese segmento concreto, la nueve. Mientras que Tilde MT le otorga una puntuación BLEU de 30,26/100 a la traducción de Google Translate, nuestro motor TAE sobresale con un 90,78/100. Esta puntuación tan alta se debe a que, con respecto a la traducción oficial de la UE, nuestro motor ha producido un texto que tan solo difiere en 3 palabras, tal y como se puede comprobar en la anterior imagen.

El ejemplo de la figura 20 nos muestra una traducción más adecuada en el caso de nuestro motor TAE, ya que la estructura y terminología empleada es casi completamente igual a la de la traducción humana oficial de la UE. Se emplea una terminología más formal y adecuada que en la traducción de Google Translate, que incluso llega a cometer un error en la denominación oficial del tipo de documento, ya que no se trata de un «informe de evaluación pública europea», sino de un «Informe Público Europeo de Evaluación».

Es cierto que la puntuación BLEU final en ambos casos, a pesar de que la de nuestro motor sea más alta, no se diferencia mucho, pero el hecho de que nuestro motor aporte traducciones tan fieles a las que produciría un traductor humano hace que el poseedor ahorre más tiempo y sea más productivo en este tipo de textos que si tuviera que trabajar con una traducción de Google Translate. Esto se debe a que los cambios a realizar en el caso de nuestro motor serían mínimos, mientras que con Google Translate la documentación previa habría de ser mayor, por ejemplo, para poseer correctamente el nombre del informe EPAR.

En el siguiente subapartado procederemos con la segunda comparación, la del motor TAN de DeepL, como ya explicamos al principio del apartado 4.6.

4.6.2. Comparación con DeepL

En esta segunda comparación, representada en la siguiente imagen, la puntuación BLEU en azul corresponde al motor TAN genérico DeepL, mientras que la puntuación en verde es la asociada a nuestro motor TAE especializado, tal y como sucede en la primera comparación.

BLEU:	50.54	53.97
Precision x brevity:	50.92 x 99.25	53.97 x 100.00
Type	1-gram 2-gram 3-gram 4-gram	1-gram 2-gram 3-gram 4-gram
Individual	74.33 55.52 44.88 36.29	76.42 57.89 47.71 40.20
Cumulative	73.78 63.76 56.58 50.54	76.42 66.51 59.54 53.97

Figura 21. Plataforma de Interactive BLEU score evaluator de Tilde MT que compara las puntuaciones BLEU de DeepL, en azul, y de nuestro motor TAE, en verde.

En este caso, podemos afirmar que nuestro motor recibe una puntuación BLEU más alta que DeepL, aunque la diferencia, al igual que sucedía en el caso anterior con Google Translate, no es muy grande, ya que se trata de poco más de tres puntos: 50,54/100 (DeepL) vs. 53,97/100 (nuestro motor). En comparación con Google Translate, se puede destacar que, a pesar de ser mínima, existe una diferencia positiva a favor del motor Google Translate con respecto al motor DeepL.

En la siguiente imagen, podemos apreciar la puntuación BLEU que Tilde MT concede a las dos traducciones del segmento once.

Al igual que sucedía en el subapartado anterior, en la imagen que aparece a continuación, las palabras que aparecen en rojo serían las que difieren de la traducción oficial de la UE, las que están destacadas en color azul son los aciertos de DeepL que nuestro motor TAE ha fallado, mientras que las verdes, al contrario, representan los aciertos de nuestro motor TAE que DeepL no ha acertado, teniendo siempre como referencia la traducción oficial.

Sentence 11	BLEU	Length ratio	Text
Source	-	-	For practical information about using Afstyla, patients should read the package leaflet or contact their doctor or pharmacist.
Human	100.00	1.00	Para más información sobre el tratamiento con Afstyla , el paciente debe leer el prospecto (también incluido en el EPAR) o consultar a su médico o farmacéutico .
Machine	20.12	0.87	Para obtener información práctica sobre el uso de Afstyla , los pacientes deben leer el prospecto o ponerse en contacto con su médico o farmacéutico .
Machine	90.96	1.00	Para más información sobre el tratamiento con Lokelma , el paciente debe leer el prospecto (también incluido en el EPAR) o consultar a su médico o farmacéutico .

Figura 22. Oración 9 del texto AFSTYLA comparada con la traducción humana, la traducción automática de DeepL (Machine en azul) y la de nuestro motor TAE (Machine en verde).

En la imagen anterior, se puede observar cómo nuestro motor ha generado una traducción más acertada para ese segmento concreto, el once. Mientras que Tilde MT le otorga una puntuación BLEU de 20,12/100 a la traducción de DeepL, nuestro motor TAE sobresale con un 90,96/100. Esta puntuación tan alta se debe a que, con respecto a la traducción oficial de la UE, nuestro motor ha producido un texto que tan solo difiere en una palabra.

Ha de subrayarse que, como podemos observar en la figura anterior, la calidad de la traducción producida por nuestro motor se acerca mucho más a la traducción humana oficial de la UE que la producida por DeepL, sobre todo teniendo en cuenta la estructura cerrada y la terminología característica de este tipo de documento.

A pesar de que los resultados son muy similares en cuanto al porcentaje BLEU, hemos podido comprobar que la personalización de motores de traducción automática es esencial cuando se trata de documentos oficiales con una estructura muy definida, como es el caso de las fichas técnicas de medicamentos y de los resúmenes del EPAR para el público general de la UE. Esto asegura contar con una estructura idéntica en la traducción de todos los textos, así como una coherencia necesaria en cuanto a terminología. Además, cuando se trata de documentos que contienen información confidencial resulta esencial y necesario contar con un motor de traducción propio para no poner en peligro dicha confidencialidad, ya que realizar las traducciones en línea podría tener consecuencias legales, sobre todo, en este tipo de traducción. Además, las grandes farmacéuticas trabajan con información muy valiosa y confidencial, ya que los estudios para crear nuevos fármacos pueden durar muchos años hasta dar resultados fructíferos y la confidencialidad es indispensable para no dar pistas a las demás empresas, ya que se trata de un mundo es muy competitivo.

En el próximo apartado, se evaluarán las traducciones obtenidas con traductores humanos, todo ello con el objetivo de establecer una comparación entre los tres motores TA en función

del nivel de posesición necesario. Consideramos necesario realizar una evaluación humana para intentar dar respuesta a las dudas que han surgido como consecuencia de los resultados de las métricas de evaluación automáticas que hemos abordado en los anteriores apartados.

4.7. Ejemplos de posesición con dicho motor de traducción

En este apartado, nos dispondremos a comparar el nivel de posesición que necesita cada uno de los textos generados con los tres motores de traducción automática que nos ocupan en este trabajo de investigación.

Antes de comenzar el experimento, es necesario establecer cuál es el objetivo, así como las variables implicadas y los conceptos que se van a utilizar para medir la calidad de las traducciones automáticas.

En cuanto al texto muestra a poseer, ha de señalarse que se trata del mismo que se empleó en el apartado anterior para realizar la comparación automática de la puntuación BLEU en Tilde MT: el texto AFSTYLA, que cuenta con 941 palabras. Como ya se mencionó anteriormente, se trata de un texto farmacéutico de un fármaco comercializado en la Unión Europea y, más concretamente, es un resumen del EPAR para el público general.


El objetivo de este experimento es evaluar la calidad de los tres motores de traducción automática de este trabajo mediante posesición con traductores humanos teniendo en cuenta los siguientes indicadores de rendimiento o KPI (*Key Performance Indicators*): *adequacy* y *fluency*, que explicaremos posteriormente.

Por otro lado, la plataforma empleada para realizar este experimento es, de nuevo, KantanMT, que cuenta con una funcionalidad denominada KantanLQR. Esta opción nos permite actuar como gestores de proyectos para poder asignar evaluaciones de calidad de traducciones a traductores humanos.

Como gestor de proyectos, la vista de dicha plataforma para crear proyectos sería la siguiente.

Create New Project ✕

1 Project name	2 Languages	3 Settings	4 Finish
Enter new project name and select type.	Select source and target language.	Select project start and end date.	Check the summary of your new project.

 KantanLQR is used for checking the quality of your engine output. You can customise the kind of quality review you want to do, add reviewers and manage the progress of the review.

Name: <input type="text" value="TFM_AliciaMolina_Motor1"/>	Project Type: <input type="text" value="Quality Evalua..."/>	Group: <input type="text"/>
---	---	--------------------------------


◀ Previous Next ▶ ✕ Cancel  Create

Figura 23. Vista de KantanLQR para configurar un nuevo proyecto.

En el paso número uno, se establece el nombre del proyecto y el tipo de proyecto que, en este caso, se trata de una evaluación de la calidad. En el paso número dos, se configuran las lenguas del proyecto que, para este proyecto, serán el inglés de Reino Unido y el español de España. En el paso número tres, se establece la fecha límite para realiza dicha tarea y en el último paso, el cuatro, se crea el proyecto.

Tras haber creado y configurado el proyecto, el gestor debe asignar al poseedor mediante correo electrónico, tal y como se muestra en la imagen que aparece en la siguiente página.

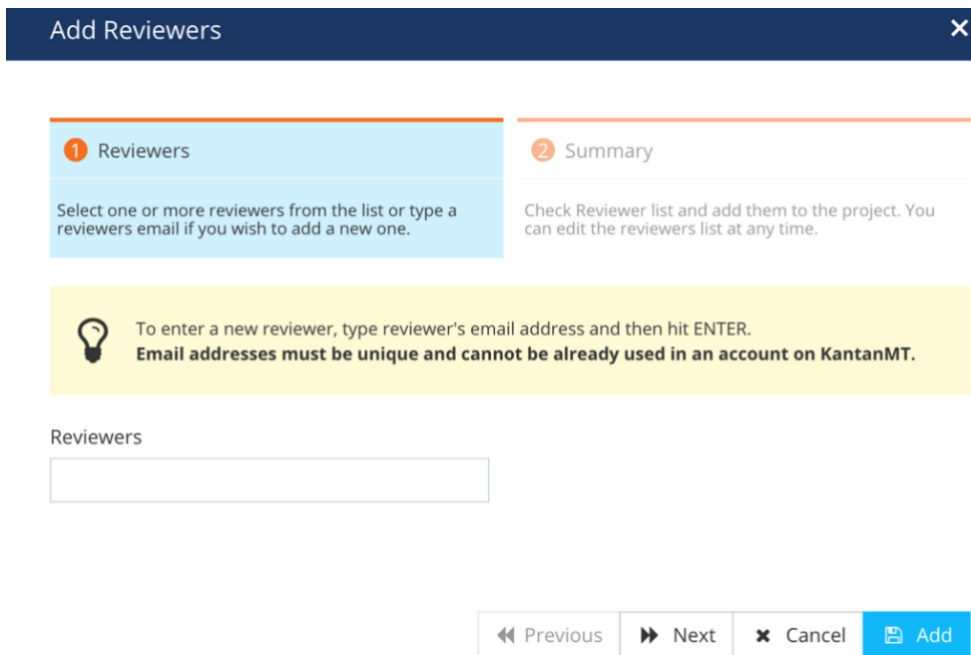


Figura 24. Vista de KantanLQR para asignar al traductor encargado de llevar a cabo el proyecto de evaluación de la calidad.

Por último, el gestor debe añadir los textos que formarán parte de la evaluación de la calidad, es decir, tanto el texto en lengua origen como el texto en lengua meta que será objeto de la posesición, el español.

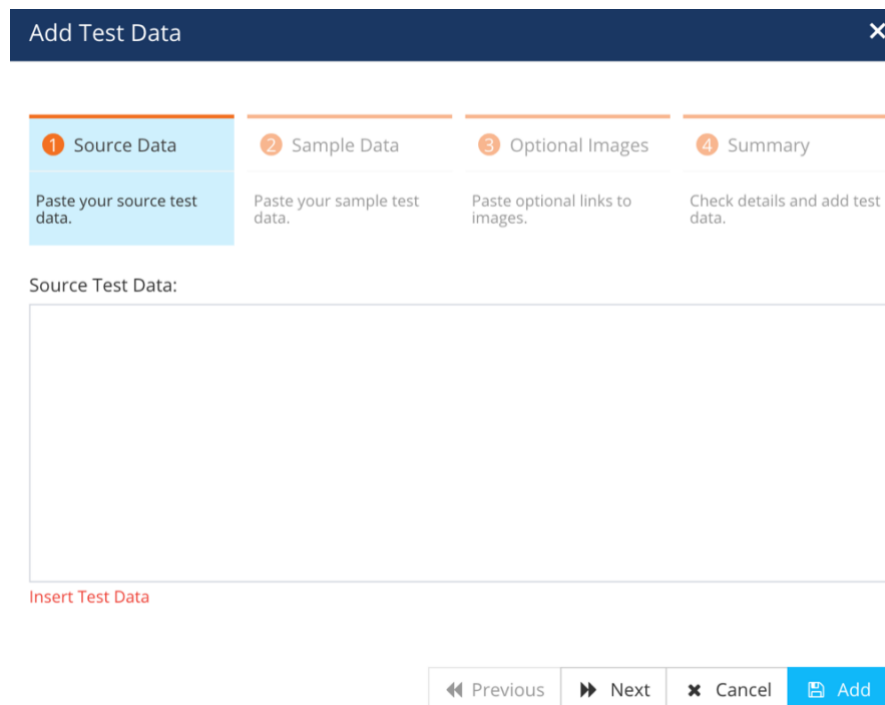


Figura 25. Vista de KantanLQR para introducir los textos que compondrán la evaluación de la calidad.

Aunque los traductores no dispusieran de una cuenta de KantanMT, al asignarlos el gestor de proyectos como poseedores para un proyecto concreto, estos reciben un correo electrónico que les indicaba los pasos a seguir para crear una cuenta de usuario que les permitiría acceder a la plataforma de KantanLQR en la que solo serán capaces de visualizar el proyecto que les ha asignado.

A continuación, se muestra el correo electrónico que recibe el traductor asignado para la posesición para poder acceder a la plataforma KantanLQR.

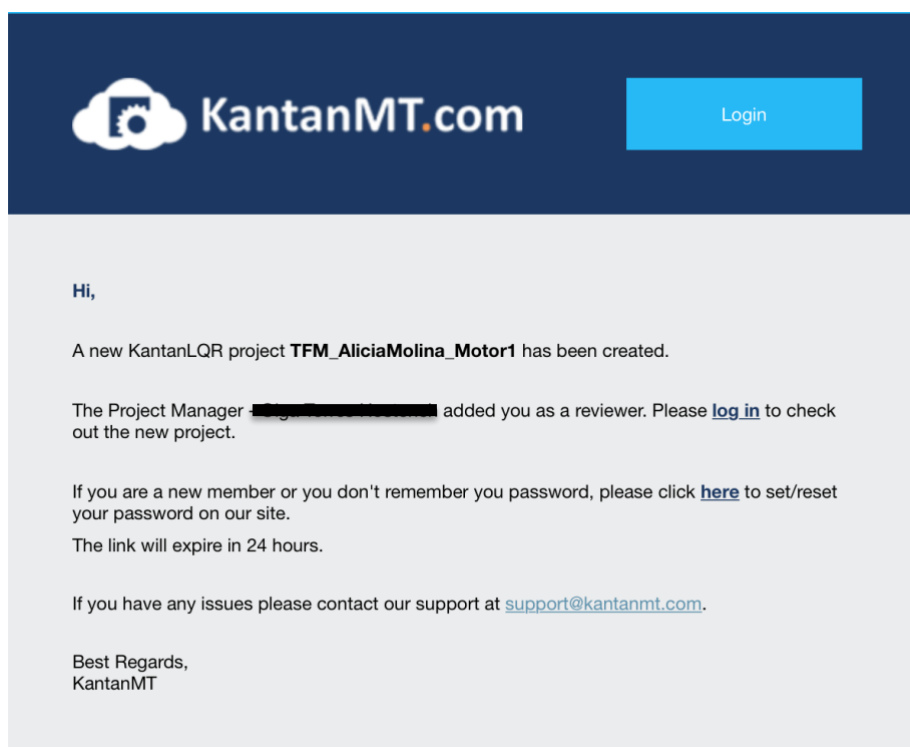


Figura 26. Correo electrónico que recibe el traductor de este experimento de posesición para realizar una evaluación de calidad.

Tras acceder a la plataforma, la vista del traductor, la cual podemos ver en la siguiente imagen, muestra un signo de interrogación en el que debería clicar para aceptar la tarea que le ha sido asignada. Una vez aceptada la tarea, aparecerá el símbolo de un lápiz donde antes se encontraba el signo de interrogación, en que el traductor también tendrá que clicar para comenzar la tarea de evaluación de calidad.

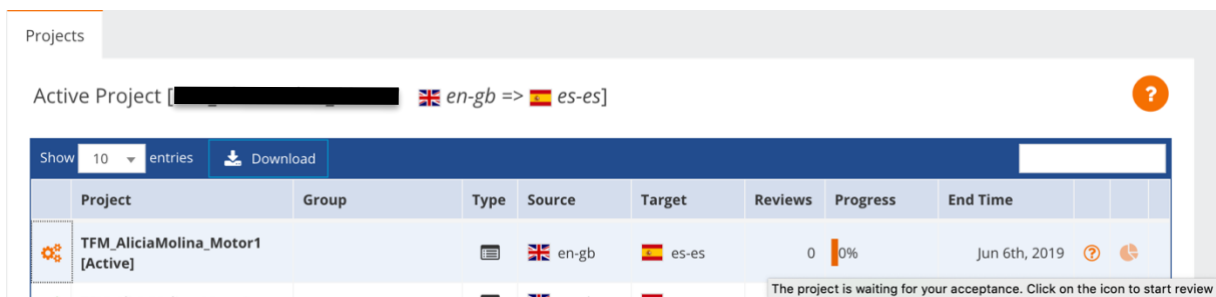


Figura 27. Vista de KantanLQR como traductor encargado de la evaluación de calidad del proyecto antes de aceptar la tarea.

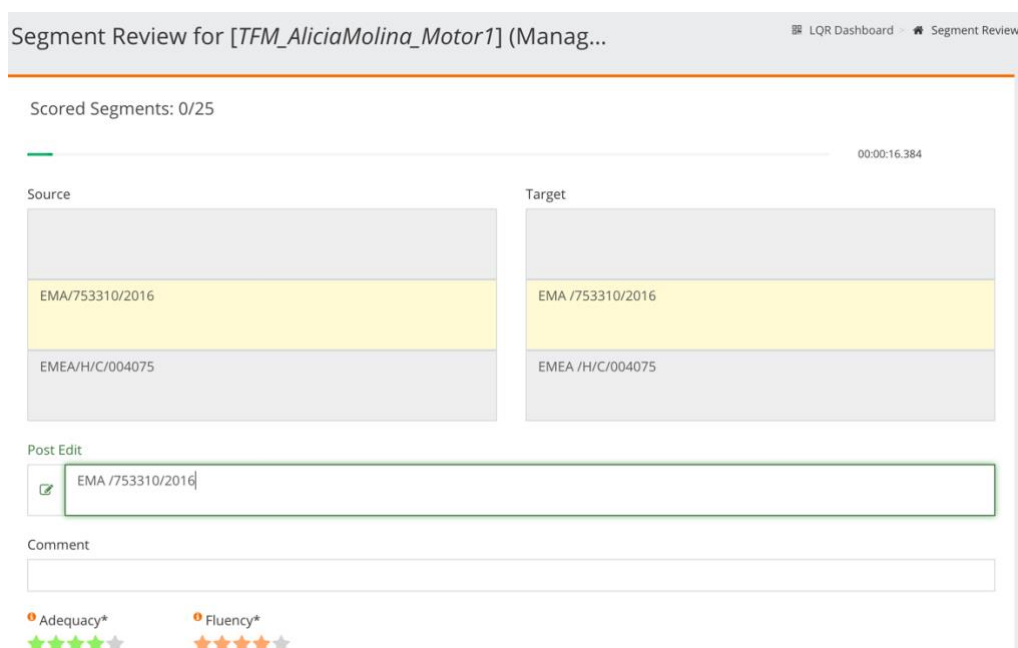


Figura 28. Vista de KantanLQR como traductor encargado de la evaluación de calidad del proyecto durante la posesición.

Seguidamente, definiremos la muestra de traductores que se encargan de realizar la posesición, así como el texto que se va a utilizar en este experimento de posesición.

En cuanto a la muestra de traductores, está formada por nueve personas que cuentan con un perfil similar. Todas ellas tienen formación educativa en traducción e interpretación, así como en tecnologías de la traducción (localización, traducción automática, herramientas TAO, gestión de proyectos, posesición...), y, además, cuatro de ellas tienen experiencia como traductores profesionales, aunque en ningún caso supera los dos años. Los nueve traductores son, además, nativos de español. Sin embargo, ninguno de los sujetos de esta muestra está especializado en traducción médica o farmacéutica.

Esta evaluación de la calidad se organizó de forma que los nueve traductores quedaron divididos en tres grupos, por lo que cada grupo consta de tres traductores. Asimismo, se crearon también tres tipos de proyectos en KantanLQR que se asignaron de la siguiente forma:

- **Grupo 1:** proyecto *TFM_AliciaMolina_Motor1*.
- **Grupo 2:** proyecto *TFM_AliciaMolina_Motor2*.
- **Grupo 3:** proyecto *TFM_AliciaMolina_Motor3*.

Los traductores del grupo 1 se encargaron de realizar la posesición de la traducción generada por nuestro motor TAE especializado en farmacia, los de grupo 2 de la de Google Translate y, por último, los del grupo 3 de la de DeepL.

Todos los traductores fueron informados de que iba a poseer un texto traducido por un motor de traducción automática. El nombre de los motores de traducción automática no se indicó a los traductores humanos en ningún momento para no influenciar las puntuaciones, por lo que desconocían de qué motor procedía el texto que iban a poseer. Igualmente, los nombres de los traductores humanos que participaron en este experimento permanecerán anónimos en todo momento.

Una vez asignados los nueve proyectos, los traductores humanos se encargaron de llevar a cabo sus correspondientes posesiciones, para las que se tuvieron que basar en los anteriormente mencionados indicadores de rendimiento *adequacy* y *fluency*. Para cada segmento, 25 en total, los traductores tuvieron que asignar una puntuación de *adequacy* y de *fluency* mediante la elección de estrellas. Esto les permitía calificar las traducciones con un máximo de cinco estrellas para cada uno de los indicadores, siendo 1 un valor muy negativo y 5 un valor muy positivo.

Para facilitar esta tarea, KantanMT ofrece una serie de consideraciones a tener en cuenta a la hora de valorar estos dos indicadores de rendimiento en una traducción:

Adequacy

- **Una estrella:** ninguno de los significados expresados en el fragmento origen se expresa en el fragmento traducido.
- **Dos estrellas:** el significado del fragmento origen apenas se expresa en el fragmento traducido.

- **Tres estrellas:** gran parte del significado del fragmento origen se expresa en el fragmento traducido.
- **Cuatro estrellas:** la mayor parte del significado del fragmento origen se expresa en el fragmento traducido.
- **Cinco estrellas:** todos los significados expresados en el fragmento origen aparecen en el fragmento traducido.

Fluency:

- **Una estrella:** no tiene fluidez. Totalmente antigramatical y, en su mayor parte, no tiene sentido. La traducción tiene que ser reescrita desde cero.
- **Dos estrellas:** poca fluidez. Mala elección de palabras, mala gramática y estructura sintáctica. Se requiere mucha posesición.
- **Tres estrellas:** bastante fluido. Aproximadamente la mitad de las traducciones contienen errores y requieren posesición.
- **Cuatro estrellas:** fluidez casi nativa. Pocos errores terminológicos o gramaticales que no afectan la comprensión general del significado. Requiere poca posesición.
- **Cinco estrellas:** dominio del idioma nativo. No hay errores gramaticales, buena elección de palabras y estructura sintáctica. No se requiere posesición.

Así, en función de las calificaciones asignadas por los traductores humanos, KantanMT genera automáticamente, por proyecto, un porcentaje global para cada indicador de rendimiento.

A continuación, en la página siguiente, aparecen dos imágenes como ejemplo ilustrativo de dichos porcentajes automáticos de KantanMT.

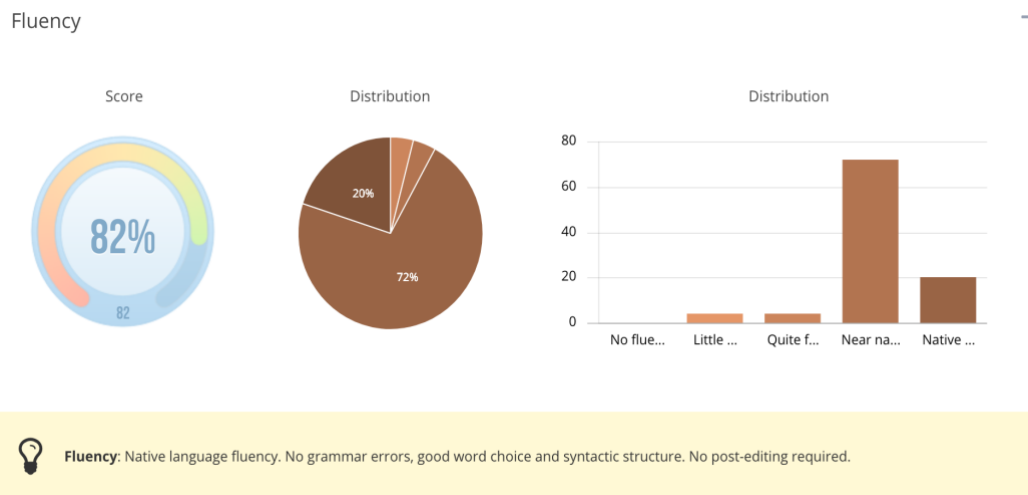


Figura 29. Porcentaje total de fluency en KantanMT de una de las posesiciones del motor 2 (Google Translate).

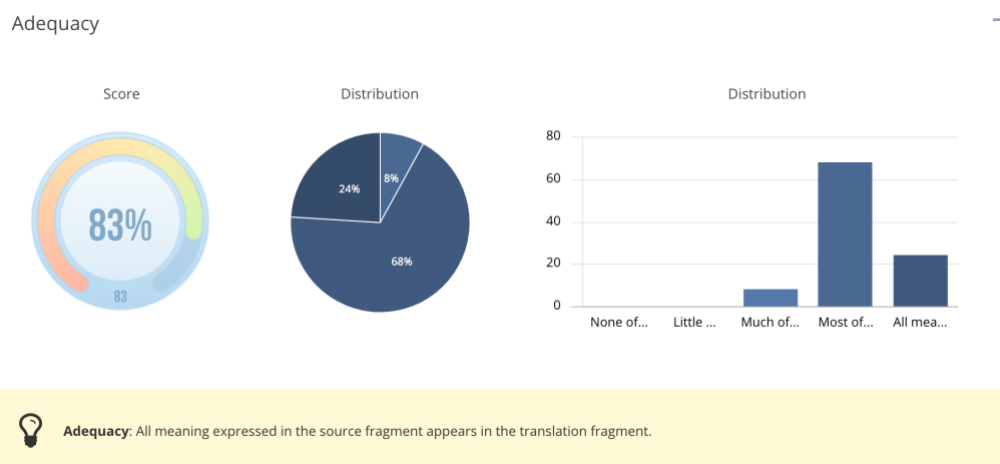


Figura 30. Porcentaje total de adequacy en KantanMT de una de las posesiciones del motor 2 (Google Translate).

Tras las posesiciones llevadas a cabo por los traductores, nos dispusimos a analizar los resultados obtenidos, que se muestran a continuación, y para los cuales se ha creado un gráfico:

- Adequacy
 - Motor TAE: 79,66 %
 - Google Translate: 85,33 %
 - DeepL: 88,33 %

- Fluency

- Motor TAE: 74,66 %
- Google Translate: 82,33 %
- DeepL: 86,66 %

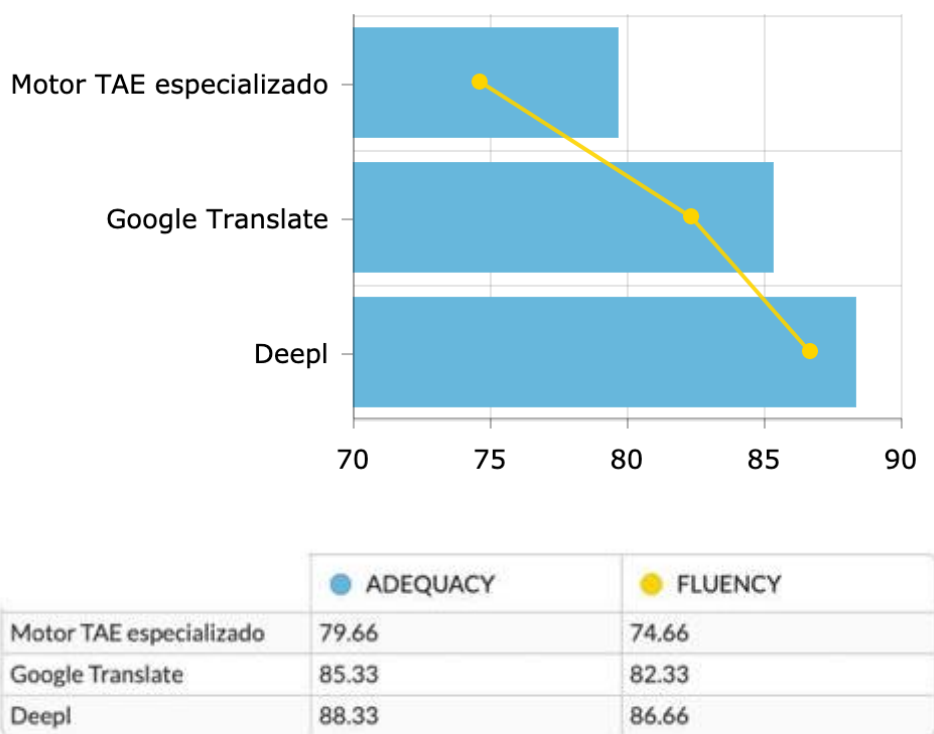


Figura 31. Media de los porcentajes de adecuacy y fluency asignados por los traductores del experimento a cada uno de los motores.

Como se puede observar en el gráfico anterior, los traductores humanos del experimento consideran que DeepL es el mejor motor de traducción automática teniendo en cuenta los valores de *adequacy* y *fluency* a la hora de poseer el texto de prueba AFSTYLA. En segundo lugar, se encuentra Google Translate y, por último, nuestro motor TAE especializado en farmacia. Ha de señalarse que la diferencia entre DeepL y Google Translate es muy poca, ya que no supera en ninguno de los dos casos el 5 %. Sin embargo, en el caso de nuestro motor especializado, la diferencia es algo mayor, ya que llega a alcanzar, en el caso del valor *fluency*, un 12 %, mientras que, en el caso de *adequacy*, un 9 %.

También sería interesante mencionar que, en los **anexos** que se encuentran al final de este trabajo, están recopiladas los porcentajes correspondientes a los valores *adequacy* y *fluency* que cada traductor del experimento asignó a cada motor mediante la posesión en la plataforma de KantanMT.

Por otro lado, el tiempo medio de posesición de los traductores, en milisegundos, fue el siguiente:

- Traductor 1 (Motor 1 - Motor TAE especializado): 59 936 ms
- Traductor 2 (Motor 1 - Motor TAE especializado): 42 501,44 ms
- Traductor 3 (Motor 1 - Motor TAE especializado): 13 651,28 ms
- Traductor 4 (Motor 2 - Google Translate): 20 053, 23 ms
- Traductor 5 (Motor 2 - Google Translate): 13 076,04 ms
- Traductor 6 (Motor 2 - Google Translate): 31 367,24 ms
- Traductor 7 (Motor 3 - DeepL): 77 691,20 ms
- Traductor 8 (Motor 3 - DeepL): 20 268,96 ms
- Traductor 9 (Motor 3 - DeepL): 25 537,36 ms

Por consiguiente, el tiempo medio empleado por los traductores para realizar la posesición de la traducción resultante de cada uno de ellos es la siguiente:

- Tiempo medio de posesición - Motor 1 (**Motor TAE especializado**): **38 696,24 ms**
- Tiempo medio de posesición - Motor 2 (**Google Translate**): **21 498,84 ms**
- Tiempo medio de posesición - Motor 3 (**DeepL**): **41 165,84 ms**

Como hemos podido constatar, los revisores han invertido menos tiempo en poseer el texto traducido por Google Translate, mientras que nuestro motor TAE especializado se posiciona en segundo lugar y DeepL se sitúa en última posición por poca diferencia. DeepL, a pesar de haber sido el motor calificado como el mejor por los mismos traductores en términos de *adequacy* y *fluency*, no es el mejor en cuanto al nivel de posesición se refiere.

En el siguiente apartado, expondremos los resultados finales de este trabajo de investigación, así como las limitaciones a las que nos hemos enfrentado en el mismo.

5. RESULTADOS Y LIMITACIONES

Como resultado de este trabajo de investigación y tras haber realizado la evaluación automática de los tres motores que nos ocupan en este trabajo y el experimento de posesición con traductores humanos, procedemos a realizar una breve recapitulación, así como una exposición de las limitaciones que hemos encontrado a la hora de realizar este estudio.

En cuanto a las métricas de evaluación automática, más concretamente el porcentaje BLEU, los resultados mostraron que nuestro motor TAE especializado es el que mejores puntuaciones obtiene, seguido de DeepL y, en último lugar, Google Translate.

Sin embargo, posteriormente en el experimento de posesición, los traductores humanos califican nuestro motor, por un margen muy pequeño, como el menos acertado en cuanto a valores de *adequacy* y *fluency* se refiere, otorgando la mayor puntuación a DeepL, seguido de Google Translate. Asimismo, el tiempo empleado para la posesición de las traducciones de cada uno de los motores es mayor en el caso de DeepL, seguido de nuestro motor y de Google Translate. Estos datos muestran inconsistencias, ya que no tiene sentido que DeepL haya sido calificado como el mejor motor TA en términos de *adequacy* y *fluency*, pero haya sido, a su vez, aquel cuya posesición ha requerido más tiempo por parte del traductor humano. Esto puede haberse debido a que los poseedores no han seguido unos criterios rigurosos a la hora de llevar a cabo la posesición o que no hayan seguido las instrucciones dadas, ya que se les indicó que la posesición debía hacerse sin pausas, de ahí esa diferencia de tiempos. Esa es la razón por la que, dada la inconsistencia de dichos tiempos con respecto a la evaluación, no puede afirmarse que sean muy significativos.

Además, otro factor a tener en cuenta podría ser la «calidad percibida» en cuanto a los motores TAN, ya que las traducciones producidas por este tipo de motores parecen más fluidas y esto puede desconcertar a los revisores. Hablamos de la «falsa ilusión» de naturalidad que acompaña a los motores TAN, terreno donde entran en juego los conceptos de «calidad percibida» y «calidad real», ya que, la calidad percibida, a diferencia de la real, se considera subjetiva. Esta «falsa ilusión», sumada a la poca experiencia y nula especialización de los poseedores, podría haber sido la responsable de estos resultados.

Por otro lado, también resulta curioso que, a pesar de que el motor TAE especializado recibió la peor puntuación, los revisores emplearon menos tiempo en poseer su traducción que la producida por el motor de traducción automática neuronal DeepL.

Sin embargo, ha de subrayarse que este estudio ha contado con diversas limitaciones. Una de ellas ha sido el poco tiempo disponible, ya que, si se hubiera dispuesto de más tiempo, se hubieran obtenido más datos para poder reentrenar nuestro motor especializado en farmacia y, así, mejorarlo.

Por otro lado, la cantidad de textos de muestra a la hora de hacer el experimento de posesición con traductores humanos ha sido insuficiente, ya que se empleó un único texto de muestra, un resumen del EPAR para el público general correspondiente a Afstylá. Este experimento hubiera sido interesante realizarlo con más textos de otros medicamentos, ya hubieran sido resúmenes del EPAR para el público general o fichas técnicas.

Por último, hemos de añadir que el experimento de posesición hubiera sido más riguroso si hubiéramos podido contar con la colaboración de un mayor número de traductores y que, además, estos traductores fueran todos profesionales especializados en traducción farmacéutica, ya que la valoración de las traducciones habría sido más precisa y acertada.

En cuanto a la potencia de los tres motores de traducción automática, también podría señalarse que, a pesar de ser motores genéricos, DeepL y Google Translate cuentan con una gran cantidad de datos de entrenamiento que, en un ámbito especializado y con limitados recursos a nuestro alcance en nuestro caso, es imposible igualar. Con esto se quiere resaltar el hecho de que las condiciones en las que hemos realizado la comparación no son equitativas, ya que nuestro motor especializado cuenta con unos datos de entrenamiento mucho más reducidos que aquellos de DeepL y Google Translate.

Con respecto a otros estudios relacionados, Peña (2019) ya llegó a la conclusión de que, en el caso del entrenamiento de un motor TAE propio para el par de lenguas español y rumano, especializado en medicina y farmacia, los resultados de la evaluación automática de la TA en ambas direcciones demuestran que la calidad de dicho motor propio es superior, y con diferencia, a Google Translate y Yandex. El motor TAE propio, entrenado con MTradumàtica, muestra unos resultados BLEU superiores a Google. Finalmente, el estudio sostiene que la especialización de los motores TAE es una práctica que da buenos resultados, dado que se consigue una calidad superior que la de los traductores privados no especializados.

En nuestro caso, sucede algo similar, ya que podemos afirmar que con el par de lenguas inglés y español y habiendo usado KantanMT para entrenar nuestro motor TAE, este ofrece resultados de mejor calidad que los motores TA privados Google Translate y DeepL. Sin embargo, nuestro estudio es innovador en el sentido de que está especializado únicamente en

textos farmacéuticos, sobre todo, en prospectos farmacéuticos y fichas técnicas de medicamentos, y la comparación se hace exclusivamente con dos motores genéricos de traducción automática neuronal. Además, nuestro estudio está centrado en el par de lenguas EN-ES.

Por otro lado, Costa-jussà et al. (2012) realizaron un estudio para el que llevaron a cabo una serie de experimentos con traductores automáticos populares, como el traductor de Google, con el fin de evaluar si los textos médicos pueden ser traducidos usando estas herramientas. Sin embargo, una evaluación automática y humana realizada en siete pares de idiomas concluyó que el rendimiento de estos sistemas todavía no es lo suficientemente bueno en un ámbito en el que se requiere el 100 % de precisión.

Nuestro estudio, sin embargo, a pesar de que está centrado en la traducción farmacéutica, desmiente que los sistemas de traducción populares en la web, como Google Translate, sean de insuficiente calidad, ya que hemos comprobado que, a pesar de tener que realizar posesición, los resultados obtenidos son de calidad. Ha de señalarse, por otro lado, que Google Translate está utilizando actualmente un sistema de traducción automática neuronal, mientras que en 2012 aún estaba haciendo uso de la traducción automática estadística, por lo que es probable que los resultados de la traducción automática del estudio de Costa-jussà et al. (2012) no fuera de la misma calidad que hoy en día puede proporcionar este motor.

Por último, en el siguiente apartado, dedicado a las conclusiones, abordaremos las posibles causas de los resultados obtenidos en este trabajo, así como el impacto de los mismos en nuestras hipótesis y objetivos establecidos al principio del trabajo y las posibles líneas de investigación futuras.

6. CONCLUSIONES

Como conclusión de este trabajo de investigación, podemos afirmar que hemos cumplido con los objetivos propuestos, aunque con ciertas puntualizaciones, así como también hemos demostrado las hipótesis planteadas.

En cuanto a los objetivos, en primer lugar, afirmamos que la creación de un motor de traducción automática estadística especializado propio ofrece mejores resultados que los motores gratuitos y genéricos Google Translate y DeepL según las métricas de evaluación automáticas. Sin embargo, el experimento con traductores humanos mostró que el tiempo dedicado a la posesición de la traducción de nuestro motor es mayor que el empleado en la posesición de la traducción del motor TA genérico Google Translate.

En segundo lugar, también consideramos que sí que merece la pena invertir tiempo en la creación de corpus especializados destinados a un motor de traducción automática propio, más que dedicar dicho tiempo a la posesición tras traducir con uno de los motores de traducción automática neuronales de carácter genérico disponibles en la web, tanto por cuestiones de terminología como de privacidad de datos.

En tercer lugar, la creación de un motor de traducción automática especializado en un campo muy concreto, como lo es el farmacéutico, también puede resultar factible para una empresa, ya que aumentaría la productividad del traductor.

En lo referente a las hipótesis, se demuestra que la estigmatización de la traducción automática es injustificada, ya que, bien empleada, puede ser una gran aliada del traductor. Como podemos ver, los resultados de la traducción automática son buenos en todos los casos, por lo que el hecho de poseer únicamente el texto, en lugar de traducirlo desde cero, aumentaría la productividad del traductor, como ya hemos destacado en el párrafo anterior.

Por otro lado, no se ha podido demostrar que la cantidad de posesición necesaria para un texto especializado sea mucho menor si el motor de traducción automática que se ha empleado para traducirlo es adecuado a la tipología textual. El experimento de posesición no lo muestra así, pero puede deberse, como hemos mencionado en apartados anteriores de este TFM, a la limitación en cuanto a la recogida de datos para entrenar nuestro motor y en cuanto a la muestra de poseedores escogida para el experimento, ya que las métricas automáticas lo posicionan como el mejor de los tres y, además, la calidad terminológica y estructural de la traducción de nuestro motor en comparación con DeepL y Google Translate es evidente, como se muestra en los ejemplos de Tilde MT.

Sin embargo, no se esperaba que los resultados del experimento de posesición no otorgaran el primer puesto en cuanto a calidad (según los valores de *adequacy* y *fluency*) a nuestro motor TAE, por lo que no respaldan las métricas de evaluación automática de KantanMT.

En cuanto a posibles futuras líneas de investigación, se plantean las siguientes cuestiones y situaciones:

- ¿Un motor de traducción automática neuronal ofrecería más calidad que uno de traducción estadística si ambos estuvieran especializados en el mismo ámbito y entrenados con los mismos datos?

- También sería interesante realizar el mismo experimento de posesición con un mayor número de participantes, en este caso traductores profesionales con experiencia y especializados en el ámbito farmacéutico. Asimismo, se podría también, a su vez, utilizar un mayor número de textos de muestra, como fichas de técnicas de medicamentos.

En conclusión, podemos afirmar que, en este trabajo de investigación, hemos podido comprobar que un motor TAE especializado en el ámbito farmacéutico puede llegar a ofrecer mejores resultados que un motor TAN no especializado cuando se trata de textos farmacéuticos, aunque habría que realizar un experimento más detallado con una muestra considerable de poseedores profesionales y especializados en traducción farmacéutica para poder dar una afirmación consistente. De la misma forma, hemos comprobado que las métricas de evaluación automáticas no tienen por qué corresponderse siempre con el juicio de traductores humanos. Además, también hemos sido testigos de la importancia que ha adquirido la traducción automática en la industria de la traducción, ya que se trata de un instrumento que puede aumentar en gran medida la productividad del traductor humano en gran medida. Esta situación ha dado lugar al nacimiento de la figura del poseedor, profesión que cobrará vital importancia en los próximos años como consecuencia directa de la evolución de las tecnologías de la traducción. Por tanto, la estigmatización de la traducción automática es injustificada, ya que, bien empleada, puede ser una gran aliada del traductor.

BIBLIOGRAFÍA

- Alcina, M.A. (2011). Los traductores automáticos en la red. El español en el mundo. Anuario del Instituto Cervantes 2010-2011. Recuperado de https://cvc.cervantes.es/lengua/anuario/anuario_10-11/alcina/p01.htm
- Arevalillo, J.J. (2012). La traducción automática en las empresas de traducción. *Revista Tradumàtica: tecnologies de la traducció*, 10, 179-184. Recuperado de https://ddd.uab.cat/pub/tradumatica/tradumatica_a2012n10/tradumatica_a2012n10p179.pdf
- Bahdanau, D., Cho, K. & Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate. *Computing Research Repository*, doi: [arXiv:1409.0473](https://arxiv.org/abs/1409.0473). Recuperado de <https://arxiv.org/pdf/1409.0473.pdf>
- Berner, S. (10-12 septiembre, 2003). *Lost in translation: cross-lingual communication, and virtual academic communities*. Trabajo presentado en *5th Annual Conference on World Wide Web Applications*. Durban, Sudáfrica. Recuperado de <http://www.samberner.com/documents/KM/lost.pdf>
- Collantes, C., Mallo, J., Parra, C., Quiñones, H. & Serrano, R. (2018). Pásate al lado oscuro: ventajas de la traducción automática para el traductor profesional. *La Linterna del Traductor*, 17, 33-39. Recuperado de <http://lalinternadeltraductor.org/n17/ventajas-traduccion-automatica.html>
- Costa-jussà, M. R., Farrús, M. & Serrano, J. (3-7 diciembre, 2012). Machine Translation in Medicine. A quality analysis of statistical machine translation in the medical domain. Trabajo presentado en *Proceedings in Advanced Research in Scientific Areas. The 1st Virtual International Conference*. Recuperado de <https://www.upf.edu/documents/2981382/2985826/Machine+Translation+in+Me.pdf/01f096b5-6a54-4568-aab7-c2b6f6481f69>
- Doherty, S. & Gaspari, F. (5 diciembre, 2013). *Effective Post-Editing in Human & Machine Translation Workflows: Critical Knowledge & Techniques*. Centre for Next Generation

Localisation. Dublín, Irlanda. Recuperado de <http://www.qt21.eu/downloads/QTLaunchPad-Post-Editing-Webinar.pdf>

Doğru, G., Martín-Mor, A. & Ortiz-Rojas, S. (2017). *MTradumàtica: Free Statistical Machine Translation Customisation for Translators*. Trabajo presentado en *20th Annual Conference of the European Association for Machine Translation*. Praga, República Checa. Recuperado de https://ddd.uab.cat/pub/posters/2017/174910/MTradumatica_EAMT_May_2017.pdf

EAMT - European Association for Machine Translation. (s.f.). Machine Translation. Recuperado el 15 enero 2019 de <http://www.eamt.org/mt.php>

Forcada, M. (2017). Making sense of neural machine translation. *Translation Spaces. A Multidisciplinary, Multimedia, And Multilingual Journal Of Translation*, 6(2), 291-309, doi: 10.1075/ts.6.2.06for. Recuperado de <https://www.dlsi.ua.es/~mlf/docum/forcada17j2.pdf>

Ginestí, M. & Forcada, M. (2009). La traducció automàtica en la pràctica: aplicacions, dificultats i estratègies de desenvolupament. *Caplletra: Revista Internacional de Filologia*, 46, 43-59. Recuperado de <http://roderic.uv.es/handle/10550/48542>.

Hearne, M. & Way, A. (2011a). Statistical Machine Translation: A Guide for Linguists and Translators. *Language And Linguistics Compass*, 5(5), 205-226, doi: 10.1111/j.1749-818x.2011.00274.x. Recuperado de <https://www.computing.dcu.ie/~away/CA446/SMTforLinguists.pdf>

Hearne, M. & Way, A. (2011b). On the Role of Translations in State-of-the-Art Statistical Machine Translation. *Language And Linguistics Compass*, 5(5), 227-248, doi: 10.1111/j.1749-818x.2011.00275.x. Recuperado de https://www.researchgate.net/publication/228380065_A_state-of-the-art_statistical_machine_translation_system_based_on_moses

KantanMT. (s.f.). What is BLEU score? Recuperado el 19 enero 2019 de <https://www.kantanmt.com/whatisbleuscore.php>

- KantanMT. (s.f.). What is F-Measure? Recuperado el 19 enero 2019 de <https://www.kantanmt.com/whatisfmeasure.php>
- Parra, C. (2018). ¿Cómo ha evolucionado la traducción automática en los últimos años? *La Linterna del Traductor*, 16, 20-27. Recuperado de http://www.lalinternadeltraductor.org/pdf/lalinterna_n16.pdf
- Peña, V. (2019). *Entrenament de motors de traducció automàtica estadística entre el castellà i el romanés especialitzats en farmàcia i medicina* (Trabajo de Fin de Máster). Universitat Autònoma de Barcelona, Bellaterra, España. Recuperado de https://ddd.uab.cat/pub/trerecpro/2017/hdl_2072_289460/victor_pena_irls_tfm.pdf
- Presas, M., Cid-Leal, P. & Torres-Hostench, O. (2016). Machine translation implementation among language service providers in Spain: A mixed methods study. Informe ProjeTA 2015. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 3 (1), 126-144. Recuperado de https://ddd.uab.cat/pub/artpub/2016/176378/jouresdes_a2016v3n1.pdf
- Rico, C. (30 octubre - 1 noviembre, 2014). *Editor, Translator or Post-editor? Or how Machine Translation is here to stay*. Trabajo presentado en *14th Meeting of the Mediterranean Editors and Translators Association* (MET). El Escorial, Madrid, España. Recuperado de <https://es.slideshare.net/celiarico/editor-translator-or-posteditor-by>
- Torres-Hostench, O., Presas, M. & Cid-Leal, P. (coords.). (2016). El uso de traducción automática y posesición en las empresas de servicios lingüísticos españolas: Informe de investigación ProjeTA 2015. Bellaterra, España. Recuperado de <https://ddd.uab.cat/record/148361>
- Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A Study of Translation Edit Rate with Targeted Human Annotation. Trabajo presentado en *Proceedings of the 7th Biennial Conference of the Association for Machine Translation in the Americas (AMTA-2006)*, Cambridge, EE. UU. Recuperado de <https://www.semanticscholar.org/paper/A-Study-of-Translation-Edit-Rate-with-Targeted-Snover-Dorr/3d07b5087e53c6f7c228b3c7e769494527be228e>

- Way, A. (2013). Traditional and Emerging Use-Cases for Machine Translation. *Proceedings of Translating and the Computer*, 35. Recuperado de https://www.computing.dcu.ie/~away/PUBS/2013/Way_AS LIB_2013.pdf
- Way, A. (2018). Quality Expectations of Machine Translation. En: Moorkens, J., Castilho, S., Gaspari, F. & Doherty, S. (eds.). *Translation Quality Assessment: From Principles to Practice*, 159-178. Berlín, Alemania. Recuperado de https://doi.org/10.1007/978-3-319-91241-7_8
- Zaretskaya, A. (2017) Machine Translation Post-Editing at TransPerfect – the ‘Human’ Side of the Process. *Tradumàtica: tecnologies de la traducció*, 15, 116-123, doi: <https://doi.org/10.5565/rev/tradumatica.201>. Recuperado de https://ddd.uab.cat/pub/tradumatica/tradumatica_a2017n15/tradumatica_a2017n15p11_6.pdf

ANEXOS

Anexo 1. Texto de muestra escogido en versión original (inglés) de la EMA para hacer la comparativa entre las traducciones de Google Translate, DeepL y del motor creado para este trabajo

EMA/753310/2016

EMA/H/C/004075

EPAR summary for the public

Afstyla

lonoctocog alfa

This is a summary of the European public assessment report (EPAR) for Afstyla. It explains how the Agency assessed the medicine to recommend its authorisation in the EU and its conditions of use. It is not intended to provide practical advice on how to use Afstyla.

For practical information about using Afstyla, patients should read the package leaflet or contact their doctor or pharmacist.

What is Afstyla and what is it used for?

Afstyla is a medicine used for the treatment and prevention of bleeding in patients with haemophilia A (an inherited bleeding disorder caused by lack of a clotting protein called factor VIII). It contains the active substance lonoctocog alfa.

How is Afstyla used?

Afstyla is available as a powder and solvent used to make a solution for injection. The injection is given into a vein over several minutes. The dose and frequency of the injections depend on whether Afstyla is used to treat or prevent bleeding, the severity of the patient's factor VIII deficiency, the extent and location of the bleeding and the patient's condition and bodyweight.

Afstyla can only be obtained with a prescription and treatment should be under the supervision of a doctor who has experience in the treatment of haemophilia. For further information, see the package leaflet.

How does Afstyla work?

Patients with haemophilia A lack factor VIII, a protein needed for normal clotting of the blood, and as a result, they bleed readily.

30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom

Telephone +44 (0)20 3660 6000 Facsimile +44 (0)20 3660 5555

Send a question via our website www.ema.europa.eu/contact

An agency of the European Union

© European Medicines Agency, 2017. Reproduction is authorised provided the source is acknowledged.

The active substance in Afstyla, lonoctocog alfa, is a shorter version of human factor VIII that works in the body in the same way as human factor VIII. It replaces the missing factor VIII, thereby helping the blood to clot and giving temporary control of bleeding.

What benefits of Afstyla have been shown in studies?

Afstyla has been shown to be effective at preventing and treating bleeding episodes in two main studies in patients with severe haemophilia A who were previously treated with other factor VIII products.

The first study involved 173 patients aged 12 years and above. A total of 848 bleeding episodes were recorded during the study, 94% of which resolved with one or two injections of Afstyla. Afstyla was rated as ‘excellent’ or ‘good’ at treating 92% of the bleeding episodes.

In the 16 surgical procedures that occurred during the study, Afstyla was rated as ‘excellent’ or ‘good’ in preventing bleeding episodes when used two to three times a week. Among those patients who received Afstyla to prevent bleeding, an average of 1.14 bleeding episodes per year was recorded for each patient which was lower than the 19.64 episodes per year in patients not receiving Afstyla for prevention.

The second study involved 83 patients below 12 years of age. Afstyla was rated as ‘excellent’ or ‘good’ at treating 96% of the 347 bleeding episodes recorded during the study; 96% of bleeding episodes resolved with one or two injections of Afstyla. Among those patients who received Afstyla to prevent bleeding, the average number of bleeding episodes per year was 2.30 for patients receiving Afstyla three times a week, and 4.37 for those receiving Afstyla twice a week.

What are the risks associated with Afstyla?

Hypersensitivity (allergic) reactions are common with Afstyla, affecting up to 1 in 10 people. They may include: angioedema (swelling of tissues under the skin), burning and stinging at the injection site, chills, flushing, itchy rash over the whole body, headache, hives, hypotension (low blood pressure), lethargy, nausea (feeling sick), restlessness, tachycardia (rapid heartbeat), tightness of the chest, tingling, vomiting and wheezing. In some cases these reactions can become severe.

There is also a risk with factor VIII medicines that some patients will develop inhibitors (antibodies) against factor VIII, causing the medicine to stop working and resulting in a loss of bleeding control.

Afstyla must not be used in patients with known allergy to hamster proteins.

For the full list of all side effects and restrictions, see the package leaflet.

Why is Afstyla approved?

Afstyla has been shown to be effective in both preventing and treating bleeding episodes. Regarding safety, the reported side effects are in line with those expected for a factor VIII product, although hypersensitivity reactions occurred more frequently with Afstyla. Further safety data should be provided by ongoing studies.

The Agency's Committee for Medicinal Products for Human Use (CHMP) decided that Afstyla's benefits are greater than its risks and recommended that it be approved for use in the EU.

Afstyla

EMA/753310/2016

Page 2/3

What measures are being taken to ensure the safe and effective use of Afstyla?

Recommendations and precautions to be followed by healthcare professionals and patients for the safe and effective use of Afstyla have been included in the summary of product characteristics and the package leaflet.

Other information about Afstyla

The European Commission granted a marketing authorisation valid throughout the European Union for Afstyla on 4 January 2017.

The full EPAR for Afstyla can be found on the Agency's website: [ema.europa.eu/Find medicine/Human medicines/European public assessment reports](http://ema.europa.eu/Find/medicine/Human%20medicines/European%20public%20assessment%20reports). For more information about treatment with Afstyla, read the package leaflet (also part of the EPAR) or contact your doctor or pharmacist.

This summary was last updated in 01-2017.

Afstyla

EMA/753310/2016

Page 3/3

Anexo 2. Traducción oficial del texto de muestra escogido de la EMA para hacer la comparativa entre los resultados de Google Translate, DeepL y del motor creado para este trabajo

EMA/753310/2016

EMEA/H/C/004075

Resumen del EPAR para el público general

Afstyla

lonoctocog alfa

El presente documento resume el Informe Público Europeo de Evaluación (EPAR) de Afstyla. En él se explica cómo la Agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la UE y sus condiciones de uso. No está destinado a proporcionar consejos prácticos sobre cómo utilizar Afstyla.

Para más información sobre el tratamiento con Afstyla, el paciente debe leer el prospecto (también incluido en el EPAR) o consultar a su médico o farmacéutico.

¿Qué es Afstyla y para qué se utiliza?

Afstyla es un medicamento que se utiliza para el tratamiento y la prevención de las hemorragias en pacientes con hemofilia A (un trastorno congénito de la coagulación causado por la falta de una proteína llamada factor VIII). Contiene el principio activo lonoctocog alfa.

¿Cómo se usa Afstyla?

Afstyla se presenta como un polvo y un disolvente para preparar una solución inyectable. La inyección se administra en una vena durante varios minutos. La dosis y la frecuencia de las inyecciones dependen de si Afstyla se utiliza para el tratamiento y la prevención de hemorragias, del grado de deficiencia de factor VIII en el paciente, del grado y la localización de la hemorragia y del estado y el peso corporal del paciente.

Afstyla solo se podrá dispensar con receta médica y el tratamiento debe iniciarse bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia. Para más información, consulte el prospecto.

¿Cómo actúa Afstyla?

Los pacientes con hemofilia A carecen del factor VIII, una proteína necesaria para la coagulación normal de la sangre y, como resultado, sangran fácilmente.

30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom

Telephone +44 (0)20 3660 6000 Facsimile +44 (0)20 3660 5555

Send a question via our website www.ema.europa.eu/contact

An agency of the European Union

© European Medicines Agency, 2017. Reproduction is authorised provided the source is acknowledged.

El principio activo de Afstyla, lonoctocog alfa, es una versión más corta del factor VIII humano que actúa en el organismo de la misma forma que el factor VIII humano. Sustituye al factor VIII ausente, con lo que ayuda a la sangre a coagular y permite el control temporal de la hemorragia.

¿Qué beneficios ha demostrado tener Afstyla en los estudios realizados?

Se ha demostrado que Afstyla es eficaz para la prevención y el tratamiento de episodios hemorrágicos en dos estudios principales realizados en pacientes con hemofilia A grave que habían recibido tratamiento previo con otros productos que contenían factor VIII.

El primer estudio se realizó en 173 pacientes de 12 o más años de edad. Se registraron un total de 848 episodios hemorrágicos durante el estudio, 94 % de los cuales remitieron con una o dos inyecciones de Afstyla. La eficacia de Afstyla se consideró «excelente» o «buena» para el tratamiento del 92 % de los episodios hemorrágicos.

En los 16 procedimientos quirúrgicos que tuvieron lugar durante el estudio, la eficacia de Afstyla se consideró «excelente» o «buena» para la prevención de episodios hemorrágicos cuando se utilizó dos o tres veces a la semana. Entre los pacientes que recibieron Afstyla para prevenir hemorragias, se registró un promedio de 1,14 episodios hemorrágicos al año en cada paciente, un número inferior a los 19,64 episodios al año registrados en pacientes que no recibieron Afstyla para fines de prevención.

En el segundo estudio participaron 83 pacientes menores de 12 años. La eficacia de Afstyla se consideró «excelente» o «buena» para el tratamiento del 96 % de los 347 episodios hemorrágicos registrados durante el estudio; el 96 % de los episodios hemorrágicos remitieron con una o dos inyecciones de Afstyla. Entre los pacientes que recibieron Afstyla para la prevención de hemorragias, el número promedio de episodios hemorrágicos al año fue de 2,30 en los pacientes que recibieron Afstyla tres veces por semana y de 4,37 en los que recibieron Afstyla dos veces por semana.

¿Cuál es el riesgo asociado a Afstyla?

Las reacciones de hipersensibilidad (alérgicas) son frecuentes con Afstyla y afectan a hasta 1 de cada 10 personas. Estas reacciones pueden consistir en: angioedema (inflamación de los tejidos debajo de la piel), ardor y hormigueo en el lugar de la inyección, escalofríos, sofocos, erupción cutánea con picor en todo el cuerpo, dolor de cabeza, sarpullido, hipotensión (presión arterial baja), letargo, náuseas, inquietud, taquicardia (latido cardíaco rápido), opresión torácica, hormigueo, vómitos y pitidos al respirar (sibilancias). En algunos casos, estas reacciones pueden llegar a ser graves.

Existe también un riesgo asociado a los medicamentos que contienen factor VIII de que algunos pacientes desarrollen inhibidores (anticuerpos) contra el factor VIII, lo que hace que el medicamento deje de funcionar, con la consiguiente pérdida del control de las hemorragias.

Afstyla no debe utilizarse en pacientes con alergia conocida a proteínas de hámster.

Para consultar la lista completa de efectos adversos notificados y las restricciones de Afstyla, consultar el prospecto.

¿Por qué se ha aprobado Afstyla?

Se ha demostrado que Afstyla es eficaz tanto para la prevención como para el tratamiento de episodios hemorrágicos. En cuanto a la seguridad, los efectos secundarios comunicados son similares a los Afstyla esperados para un producto que contiene factor VIII, aunque las reacciones de hipersensibilidad fueron más frecuentes con Afstyla. Los estudios en curso deberían facilitar datos adicionales sobre la seguridad.

El Comité de Medicamentos de Uso Humano (CHMP) de la Agencia decidió que los beneficios de Afstyla son mayores que sus riesgos y recomendó autorizar su uso en la UE.

Afstyla

EMA/753310/2016

Página 2/3

¿Qué medidas se han adoptado para garantizar un uso seguro y eficaz de Afstyla?

Las recomendaciones y precauciones que deben seguir los profesionales sanitarios y los pacientes para un uso seguro y eficaz de Afstyla se han incluido en la Ficha Técnica o Resumen de las Características del Producto y el Prospecto.

Otras informaciones sobre Afstyla

El EPAR completo de Afstyla se puede consultar en la página web de la Agencia: [ema.europa.eu/Find medicine/Human medicines/European public assessment reports](http://ema.europa.eu/Find%20medicine/Human%20medicines/European%20public%20assessment%20reports). Para mayor información sobre el tratamiento con Afstyla, lea el prospecto (también incluido en el EPAR) o consulte a su médico o farmacéutico.

Afstyla

EMA/753310/2016

Página 3/3

Anexo 3. Traducción resultante de Google Translate del texto de muestra escogido de la EMA para hacer la comparativa entre los resultados de Google Translate, DeepL y del motor creado para este trabajo

EMA / 753310/2016

EMA / H / C / 004075

Resumen del EPAR para el público.

Afstyla

lonoctocog alfa

Este es un resumen del informe de evaluación pública europea (EPAR) para Afstyla. Explica cómo la Agencia evaluó el medicamento para recomendar su autorización en la UE y sus condiciones de uso. No pretende proporcionar consejos prácticos sobre cómo usar Afstyla.

Para obtener información práctica sobre el uso de Afstyla, los pacientes deben leer el prospecto o ponerse en contacto con su médico o farmacéutico.

¿Qué es Afstyla y para qué se utiliza?

Afstyla es un medicamento utilizado para el tratamiento y la prevención del sangrado en pacientes con hemofilia A (un trastorno hemorrágico hereditario causado por la falta de una proteína de coagulación llamada factor VIII). Contiene el principio activo lonoctocog alfa.

¿Cómo se usa Afstyla?

Afstyla está disponible como un polvo y un disolvente utilizado para hacer una solución inyectable. La inyección se administra en una vena durante varios minutos. La dosis y la frecuencia de las inyecciones dependen de si se usa Afstyla para tratar o prevenir el sangrado, la gravedad de la deficiencia de factor VIII del paciente, la extensión y la ubicación del sangrado y la condición y el peso corporal del paciente.

Afstyla solo puede obtenerse con receta médica y el tratamiento debe estar bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia. Para más información, consulte el prospecto.

¿Cómo funciona Afstyla?

Los pacientes con hemofilia A carecen de factor VIII, una proteína necesaria para la coagulación normal de la sangre y, como resultado, sangran fácilmente.

30 Churchill Place • Canary Wharf • Londres E14 5EU • Reino Unido

Teléfono +44 (0) 20 3660 6000 Fax +44 (0) 20 3660 5555

Enviar una pregunta a través de nuestro sitio web www.ema.europa.eu/contact

Una agencia de la unión europea

© Agencia Europea de Medicamentos, 2017. La reproducción está autorizada siempre que se cite la fuente.

El principio activo en Afstyla, lonoctocog alfa, es una versión más corta del factor VIII humano que funciona en el cuerpo de la misma manera que el factor VIII humano. Reemplaza el factor VIII faltante, lo que ayuda a que la sangre se coagule y proporciona un control temporal del sangrado.

¿Qué beneficios ha demostrado tener Afstyla en los estudios realizados?

Afstyla ha demostrado ser eficaz para prevenir y tratar los episodios de sangrado en dos estudios principales en pacientes con hemofilia A grave que fueron tratados previamente con otros productos de factor VIII.

El primer estudio involucró a 173 pacientes de 12 años y más. Durante el estudio se registraron un total de 848 episodios de sangrado, el 94% de los cuales se resolvieron con una o dos inyecciones de Afstyla. Afstyla fue calificada como "excelente" o "buena" en el tratamiento del 92% de los episodios de sangrado.

En los 16 procedimientos quirúrgicos que se produjeron durante el estudio, Afstyla se calificó de "excelente" o "buena" en la prevención de episodios de sangrado cuando se usó de dos a tres veces por semana. Entre los pacientes que recibieron Afstyla para prevenir el sangrado, se registró un promedio de 1.14 episodios de sangrado por año para cada paciente, que fue inferior a los 19.64 episodios por año en pacientes que no recibieron Afstyla para la prevención.

El segundo estudio involucró a 83 pacientes menores de 12 años. Afstyla fue calificada como "excelente" o "buena" en el tratamiento del 96% de los 347 episodios de sangrado registrados durante el estudio; El 96% de los episodios de sangrado se resolvieron con una o dos inyecciones de Afstyla. Entre los pacientes que recibieron Afstyla para prevenir el sangrado, el

número promedio de episodios de sangrado por año fue de 2.30 para los pacientes que recibieron Afstyla tres veces por semana, y de 4.37 para los que recibieron Afstyla dos veces por semana.

¿Cuáles son los riesgos asociados con Afstyla?

Las reacciones de hipersensibilidad (alérgicas) son comunes con Afstyla y afectan a 1 de cada 10 personas. Pueden incluir: angioedema (hinchazón de los tejidos debajo de la piel), ardor y picazón en el lugar de la inyección, escalofríos, enrojecimiento, erupción con picor en todo el cuerpo, dolor de cabeza, urticaria, hipotensión (presión arterial baja), letargo, náuseas (sensación de malestar) , inquietud, taquicardia (latidos rápidos del corazón), opresión en el pecho, hormigueo, vómitos y sibilancias. En algunos casos estas reacciones pueden llegar a ser graves.

También existe un riesgo con los medicamentos del factor VIII de que algunos pacientes desarrollen inhibidores (anticuerpos) contra el factor VIII, lo que hace que el medicamento deje de funcionar y provoque una pérdida del control del sangrado.

Afstylo no debe usarse en pacientes con alergia conocida a las proteínas de hámster.

Para ver la lista completa de todos los efectos secundarios y restricciones, consulte el prospecto.

¿Por qué se aprueba Afstylo?

Afstylo ha demostrado ser eficaz tanto en la prevención como en el tratamiento de los episodios de sangrado. En cuanto a la seguridad, los efectos secundarios informados están en línea con los esperados para un producto de factor VIII, aunque las reacciones de hipersensibilidad ocurrieron con más frecuencia con Afstylo. Los estudios en curso deben proporcionar más datos de seguridad.

El Comité de Medicamentos de Uso Humano (CHMP) de la Agencia decidió que los beneficios de Afstylo son mayores que sus riesgos y recomendó que se aprobara para su uso en la UE.

Afstylo

EMA / 753310/2016

Página 2/3

¿Qué medidas se han adoptado para garantizar un uso seguro y eficaz de Afstylo?

Las recomendaciones y precauciones que deben seguir los profesionales de la salud y los pacientes para el uso seguro y eficaz de Afstyla se han incluido en el resumen de las características del producto y en el prospecto.

Otras informaciones sobre Afstyla.

La Comisión Europea emitió una autorización de comercialización válida en toda la Unión Europea para Afstyla el 4 de enero de 2017.

El EPAR completo de Afstyla se puede encontrar en el sitio web de la Agencia: [ema.europa.eu/Encuentre medicamentos / Medicamentos humanos / Informes públicos europeos de evaluación](http://ema.europa.eu/Encuentre_medicamentos/Medicamentos_humanos/Informes_p%C3%BAblicos_europeos_de_evaluaci%C3%B3n). Para obtener más información sobre el tratamiento con Afstyla, lea el prospecto (que también forma parte del EPAR) o consulte a su médico o farmacéutico.

Este resumen fue actualizado por última vez en 01-2017.

Afstyla

EMA / 753310/2016

Página 3/3

Anexo 4. Traducción resultante de DeepL del texto de muestra escogido de la EMA para hacer la comparativa entre los resultados de Google Translate, DeepL y del motor creado para este trabajo

EMA/753310/2016

EMEA/H/C/004075

Resumen de EPAR para el público

Afstyla

alfa lonococog

Este es un resumen del informe de evaluación pública europea (EPAR) de Afstyla. Explica cómo la Agencia evaluó el medicamento para recomendar su autorización en la UE y sus condiciones de uso. No tiene por objeto proporcionar consejos prácticos sobre cómo utilizar Afstyla.

Para obtener información práctica sobre el uso de Afstyla, los pacientes deben leer el prospecto o ponerse en contacto con su médico o farmacéutico.

¿Qué es Afstyla y para qué sirve?

Afstyla es un medicamento utilizado para el tratamiento y prevención de hemorragias en pacientes con hemofilia A (un trastorno hemorrágico hereditario causado por la falta de una proteína coagulante llamada factor VIII). Contiene la sustancia activa lonoctocog alfa.

¿Cómo se usa Afstyla?

Afstyla está disponible en polvo y disolvente para hacer una solución inyectable. La inyección se administra en una vena durante varios minutos. La dosis y la frecuencia de las inyecciones dependen de si Afstyla se utiliza para tratar o prevenir el sangrado, la gravedad de la deficiencia de factor VIII del paciente, el grado y la ubicación del sangrado y la condición y el peso corporal del paciente.

Afstyla sólo puede obtenerse con receta médica y el tratamiento debe realizarse bajo la supervisión de un médico que tenga experiencia en el tratamiento de la hemofilia. Para más información, véase el prospecto.

¿Cómo funciona Afstyla?

Los pacientes con hemofilia A carecen del factor VIII, una proteína necesaria para la coagulación normal de la sangre, y como resultado, sangran fácilmente.

30 Churchill Lugar • Canary Wharf • Londres E14 5EU • Reino Unido

Teléfono:+44 (0) 20-3660-6000 Facsimile +44 (0) 20 en 3660 5555

Envíe una pregunta a través de nuestra página web www.ema.europa.eu/contact

Una agencia de la Unión Europea

© Agencia Europea de Medicamentos, 2017. Reproducción autorizada con indicación de la fuente bibliográfica.

El principio activo de Afstyla, lonoctocog alfa, es una versión más corta del factor VIII humano que funciona en el cuerpo de la misma manera que el factor VIII humano. Reemplaza el factor VIII que falta, ayudando así a la sangre a coagularse y dando un control temporal de la hemorragia.

¿Qué beneficios de Afstyla han sido demostrados en estudios?

Se ha demostrado que Afstyla es eficaz para prevenir y tratar los episodios hemorrágicos en dos estudios principales en pacientes con hemofilia A grave que fueron tratados previamente con otros productos de factor VIII.

En el primer estudio participaron 173 pacientes de 12 años o más. Se registraron un total de 848 episodios de sangrado durante el estudio, el 94% de los cuales se resolvieron con una o dos inyecciones de Afstyla. Afstyla fue calificado como "excelente" o "bueno" en el tratamiento del 92% de los episodios hemorrágicos.

En los 16 procedimientos quirúrgicos que ocurrieron durante el estudio, Afstyla fue calificado como "excelente" o "bueno" en la prevención de episodios de sangrado cuando se utilizó dos o tres veces por semana. Entre los pacientes que recibieron Afstyla para prevenir hemorragias, se registró un promedio de 1,14 episodios de hemorragia por año para cada paciente, lo que fue inferior a los 19,64 episodios por año en los pacientes que no recibieron Afstyla para la prevención.

En el segundo estudio participaron 83 pacientes menores de 12 años de edad. Afstyla fue calificado como "excelente" o "bueno" en el tratamiento del 96% de los 347 episodios hemorrágicos registrados durante el estudio; el 96% de los episodios hemorrágicos se resolvieron con una o dos inyecciones de Afstyla. Entre los pacientes que recibieron Afstyla para prevenir el sangrado, el número promedio de episodios hemorrágicos por año fue de 2.30 para los pacientes que recibieron Afstyla tres veces por semana, y de 4.37 para los que recibieron Afstyla dos veces por semana.

¿Cuáles son los riesgos asociados con Afstyla?

Las reacciones de hipersensibilidad (alérgicas) son comunes con Afstyla, afectando hasta 1 de cada 10 personas. Pueden incluir: angioedema (hinchazón de los tejidos debajo de la piel), ardor y escozor en el sitio de la inyección, escalofríos, rubor, sarpullido con picazón en todo el cuerpo, dolor de cabeza, urticaria, hipotensión (presión arterial baja), letargo, náuseas (sensación de náuseas), inquietud, taquicardia (latidos cardíacos rápidos), opresión en el pecho, hormigueo, vómito y sibilancias. En algunos casos, estas reacciones pueden llegar a ser graves.

También existe el riesgo de que algunos pacientes desarrollen inhibidores (anticuerpos) contra el factor VIII, lo que provoca que el medicamento deje de funcionar y resulte en una pérdida de control de la hemorragia.

Afstyla no debe usarse en pacientes con alergia conocida a las proteínas del hámster.

Para obtener una lista completa de todos los efectos secundarios y restricciones, consulte el prospecto.

¿Por qué está aprobado Afstyla?

Se ha demostrado que Afstyla es eficaz tanto en la prevención como en el tratamiento de los episodios hemorrágicos. En cuanto a la seguridad, los efectos secundarios reportados están en línea con los esperados para un producto de factor VIII, aunque las reacciones de hipersensibilidad ocurrieron con mayor frecuencia con Afstyla. Los estudios en curso deben proporcionar más datos de seguridad.

El Comité de Medicamentos de Uso Humano de la Agencia (CHMP) decidió que los beneficios de Afstyla son mayores que sus riesgos y recomendó que se aprobara su uso en la UE.

Afstyla

EMA/753310/2016

Página 2/3

¿Qué medidas se están tomando para garantizar el uso seguro y eficaz de Afstyla?

En el resumen de las características del producto y en el prospecto se han incluido recomendaciones y precauciones a seguir por los profesionales sanitarios y los pacientes para un uso seguro y eficaz de Afstyla.

Otra información sobre Afstyla

El 4 de enero de 2017, la Comisión Europea concedió a Afstyla una autorización de comercialización válida en toda la Unión Europea.

El EPAR completo de Afstyla puede consultarse en el sitio web de la Agencia: ema.europa.eu/Find medicine/Human medicines/European public assessment reports. Para más información sobre el tratamiento con Afstyla, lea el prospecto (que también forma parte de la EPAR) o póngase en contacto con su médico o farmacéutico.

Este resumen fue actualizado por última vez en 01-2017.

Afstyla

Anexo 5. Traducción resultante de nuestro motor especializado en farmacia del texto de muestra escogido de la EMA para hacer la comparativa entre los resultados de Google Translate, DeepL y del motor creado para este trabajo

EMA /753310/2016

EMA /H/C/004075

RESUMEN DEL EPAR PARA EL PÚBLICO GENERAL

Afstyla

lonoctocog alfa

El presente documento resume el Informe Público Europeo de Evaluación (EPAR) para Syntuza. En él se explica cómo la Agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la UE y sus condiciones de uso. No está destinado a proporcionar consejos prácticos sobre cómo utilizar Axumin.

Para más información sobre el tratamiento con Lokelma, el paciente debe leer el prospecto (también incluido en el EPAR) o consultar a su médico o farmacéutico.

¿Qué es Afstyla y para qué se utiliza?

Afstyla es un medicamento que se utiliza para el tratamiento y prevención de hemorragias en pacientes con hemofilia A (trastorno hereditario de la coagulación causado por la carencia de una proteína denominada factor VIII de la coagulación de la sangre). Contiene el principio activo lonoctocog alfa.

¿Cómo se usa Afstyla?

Afstyla se presenta en forma de polvo y disolvente para preparar una solución para inyección. La inyección se administra en una vena durante varios minutos. La dosis y la frecuencia de la inyección dependerán de si Afstyla se usa para tratar o prevenir la hemorragia, la gravedad de la deficiencia del factor VIII, de la localización y el grado de la hemorragia y del estado del paciente y de su peso corporal.

Afstyla solo se podrá dispensar con receta médica y el tratamiento deberá administrarse bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia. Para más información, consulte el prospecto.

¿Cómo actúa Afstyla?

Pacientes con hemofilia a carecen del factor VIII, una proteína necesaria para el normal de la coagulación de la sangre, y como resultado, pueden sangrar fácilmente.

30 Churchill lugar • Canary Wharf • Londres E14 5EU • Reino Unido

Teléfono:+44 (0) 20-3660-6000 Facsimile +44 (0) 20 en 3660 5555

Enviar una pregunta a través de nuestro sitio web [www.ema.europa.eu /contact](http://www.ema.europa.eu/contact)

Un organismo de la Unión Europea

© European Medicines Agency, 2016. Reproduction is authorised provided the source is acknowledged.

El principio activo de Afstyla, lonoctocog alfa, es una versión humana máxima del factor VIII, que actúa en el organismo de la misma forma que la del factor VIII humano. Se sustituye el factor VIII que falta, lo que contribuye a coagular la sangre y administración de un control temporal de la hemorragia.

¿Qué beneficios ha demostrado tener Apealea en los estudios realizados?

Afstyla ha demostrado ser eficaz para prevenir y tratar episodios hemorrágicos en dos estudios principales en pacientes con hemofilia A grave que hayan sido previamente tratados con otros medicamentos de factor VIII.

En el primer estudio participaron 194 pacientes de 12 años y mayores. Un total de 848 episodios hemorrágicos fueron registrados durante el estudio, el 94% de los cuales se resolvió con una o dos inyecciones de Afstyla. Afstyla fue calificada como "excelente" o "buena" en el tratamiento del 92% de los episodios hemorrágicos.

En 16 procedimientos quirúrgicos que se produjeron durante el estudio, Afstyla fue calificada como "excelente" o "buena" en la prevención de episodios de hemorragia cuando se utilizan dos o tres veces a la semana. Entre los pacientes que recibieron Afstyla para prevenir la hemorragia, una media de 1,14 episodios hemorrágicos al año. Se registró para cada paciente

que fue menor que la de episodios 19.64 al año en los pacientes que no recibieron Afstyla para la prevención.

En el segundo estudio participaron un total de 83 pacientes menores de 12 años de edad. Afstyla fue calificada como "excelente" o "buena" en el tratamiento de 96% de los 353 episodios hemorrágicos registrados durante el estudio; El 96% de los episodios de hemorragia se solucionaron con una o dos inyecciones de Afstyla. Entre los pacientes que recibieron Afstyla para prevenir la hemorragia, el número medio de episodios hemorrágicos al año fue 2,30 para los pacientes que recibieron Afstyla tres veces por semana, 4,37 para los pacientes que recibieron Afstyla dos veces a la semana.

¿Cuáles son los riesgos asociados a Apealea?

Reacciones alérgicas (hipersensibilidad) son frecuentes con Afstyla, afectan hasta 1 de cada 10 personas. Glucosa en sangre: angioedema (hinchazón de tejidos debajo de la piel), quemazón y picor en el lugar de inyección, escalofríos, enrojecimiento de la piel, erupción cutánea con picor en todo el cuerpo, dolor de cabeza, urticaria, hipotensión (presión arterial baja), letargo, náuseas (ganas de vomitar), cansancio, taquicardia (palpitaciones), opresión en el pecho, hormigueo, vómitos y respiración sibilante. En algunos casos, estas reacciones pueden ser graves.

También hay un riesgo de factor VIII medicamentos que algunos pacientes que desarrollan anticuerpos (inhibidores) contra el factor VIII, causando el medicamento dejar el trabajo y dando como resultado la pérdida de control de la hemorragia.

Afstyla no debe utilizarse en pacientes con alergia conocida a las proteínas de hámster.

La lista completa de efectos adversos y restricciones puede consultarse en el prospecto.

¿Por qué se ha aprobado Afstyla?

Afstyla ha demostrado ser eficaz en la prevención y tratamiento de episodios hemorrágicos. En cuanto a la seguridad, los efectos adversos se encuentran en línea con los esperados para un medicamento de factor VIII, aunque las reacciones de hipersensibilidad se produjeron con mayor frecuencia con Afstyla. Nuevos datos de seguridad debe ser suministrada por ensayos en marcha.

La Agencia en el Comité de Medicamentos de Uso Humano (CHMP) decidió que los beneficios de Afstyla son mayores que sus riesgos y recomendó autorizar su uso en la UE.

Afstyla

¿Qué medidas se han adoptado para garantizar un uso seguro y eficaz de Rubraca?

Las recomendaciones y precauciones que deben seguir los profesionales sanitarios y los pacientes para un uso seguro y eficaz de Rubraca se han incluido en la Ficha Técnica o Resumen de las Características del Producto y el prospecto.

Otras informaciones sobre Afstyla

La Comisión Europea emitió una autorización de comercialización válida en toda la Unión Europea para el medicamento Afstyla el 4 de enero de 2017.

El texto completo del EPAR de Afstyla puede consultarse en la página web de la Agencia Española: ema.europa.eu /Find medicine /Find medicines /Human public assessment reports. Para mayor información sobre el tratamiento con Lokelma, lea el prospecto (también incluido en el EPAR) o consulte a su médico o farmacéutico.

Fecha de la última actualización del presente resumen 01-2007.

Afstyla

Anexo 6. Informe de análisis del motor creado para este trabajo, extraído de KantanMT, antes de la adición del corpus de TAUS KantanMT

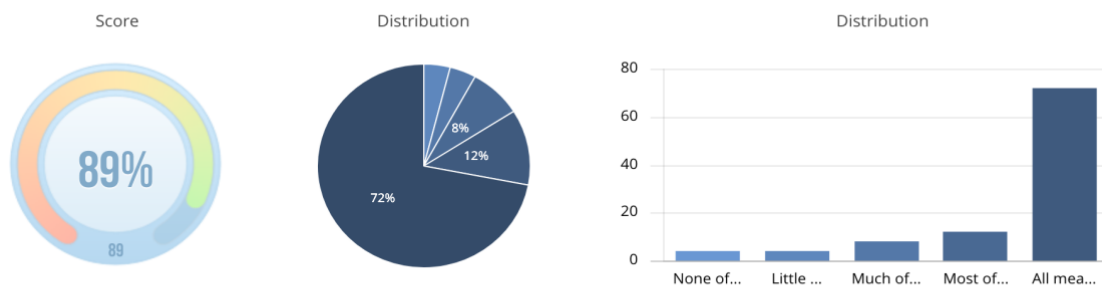
Automated Scores Distribution - Summary Report for Job #135122					
Engine Name	Source	Target	Word Count	Unique WC	Mono WC
AliciaMolina_Without_TAUS	en-gb	es-es	10388225	122874	8749015
F-Measure					78%
Interval	<40%	40-49%	50-69%	70-89%	90-100%
Segments	58	16	73	122	231
BLEU					73%
Interval	<40%	40-49%	50-69%	70-89%	90-100%
Segments	92	23	70	65	250
TER					34%
Interval	<40%	40-49%	50-69%	70-89%	90-100%
Segments	323	28	59	31	59

Anexo 7. Informe de análisis del motor creado para este trabajo, extraído de KantanMT, tras la adición del corpus de TAUS

Automated Scores Distribution - Summary Report for Job #136343					
Engine Name	Source	Target	Word Count	Unique WC	Mono WC
PruebaTAUS	en-gb	es-es	11466504	167221	8749015
F-Measure					78%
Interval	<40%	40-49%	50-69%	70-89%	90-100%
Segments	53	17	90	108	232
BLEU					70%
Interval	<40%	40-49%	50-69%	70-89%	90-100%
Segments	117	33	60	58	232
TER					36%
Interval	<40%	40-49%	50-69%	70-89%	90-100%
Segments	311	29	71	29	60

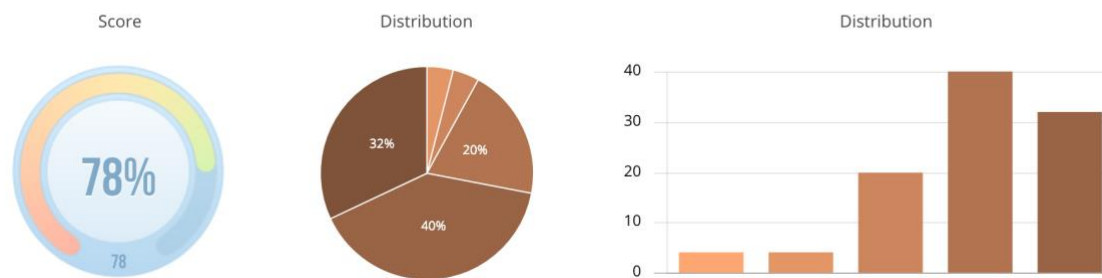
Anexo 8. Resultados de adequacy y fluency en la posesición del texto AFSTYLA del traductor 1 (correspondiente al motor TAE especializado en farmacia).

Adequacy



Adequacy: All meaning expressed in the source fragment appears in the translation fragment.

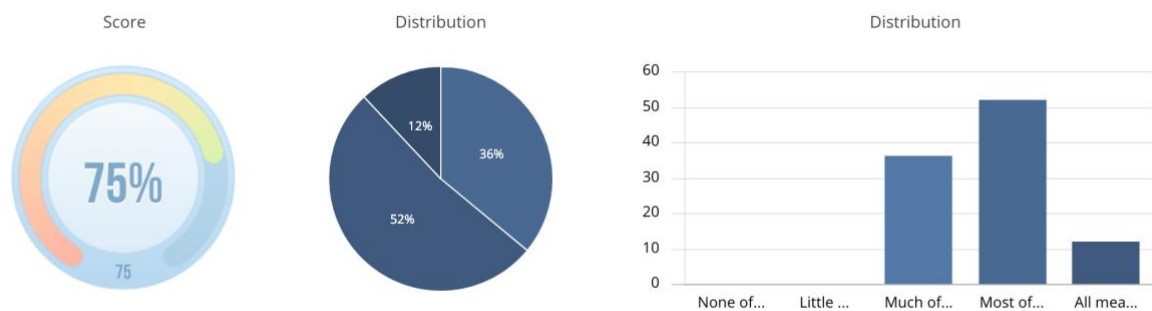
Fluency



Fluency: Near native fluency. Few terminology or grammar errors which don't impact the overall understanding of the meaning. Little post-editing required.

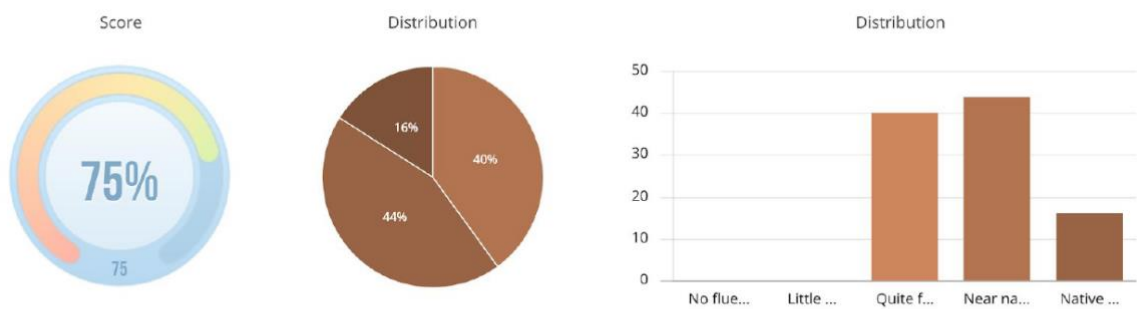
Anexo 9. Resultados de adequacy y fluency en la posesición del texto AFSTYLA del traductor 2 (correspondiente al motor TAE especializado en farmacia).

Adequacy



Adequacy: Most of the source fragment meaning is expressed in the translation fragment.

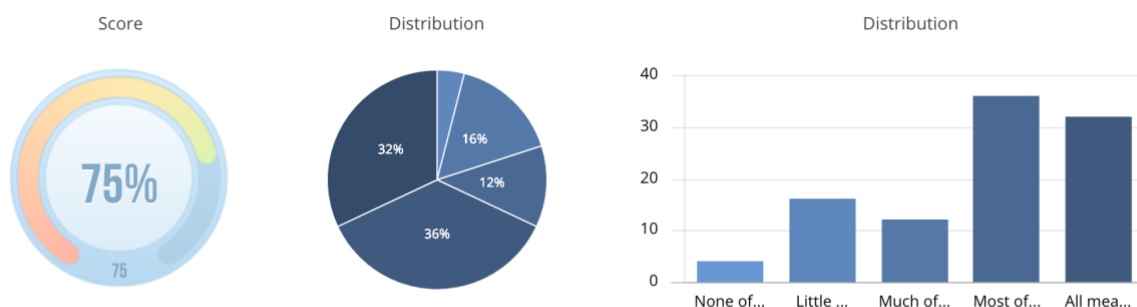
Fluency



Fluency: Near native fluency. Few terminology or grammar errors which don't impact the overall understanding of the meaning. Little post-editing required.

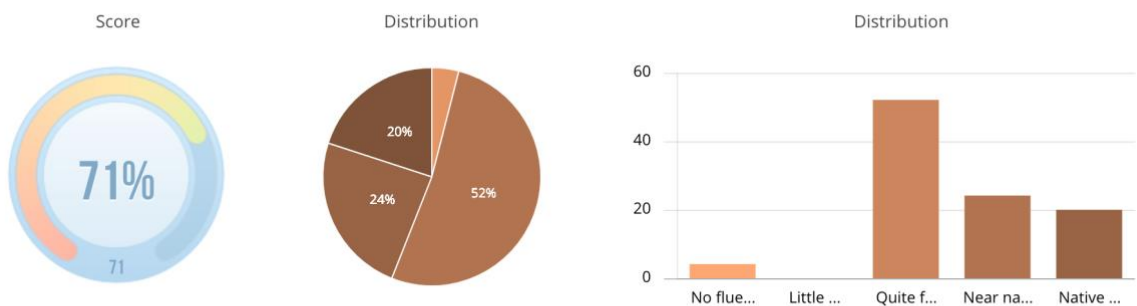
Anexo 10. Resultados de adecuacy y fluency en la posesición del texto AFSTYLA del traductor 3 (correspondiente al motor TAE especializado en farmacia).

Adequacy



Adequacy: Most of the source fragment meaning is expressed in the translation fragment.

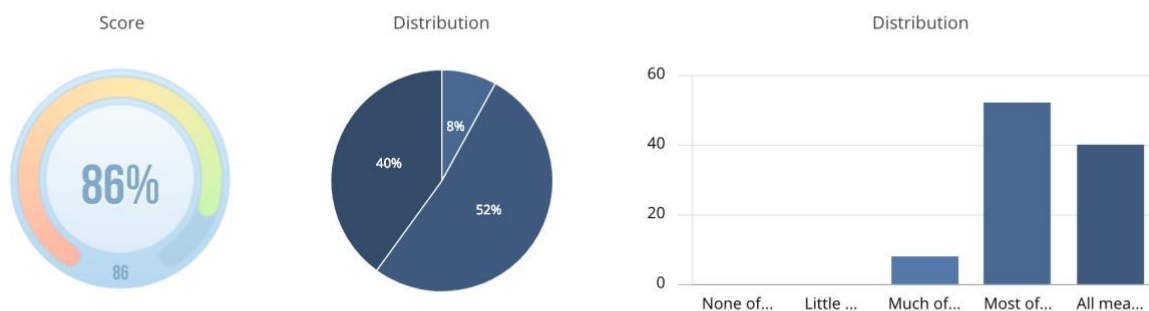
Fluency



Fluency: Near native fluency. Few terminology or grammar errors which don't impact the overall understanding of the meaning. Little post-editing required.

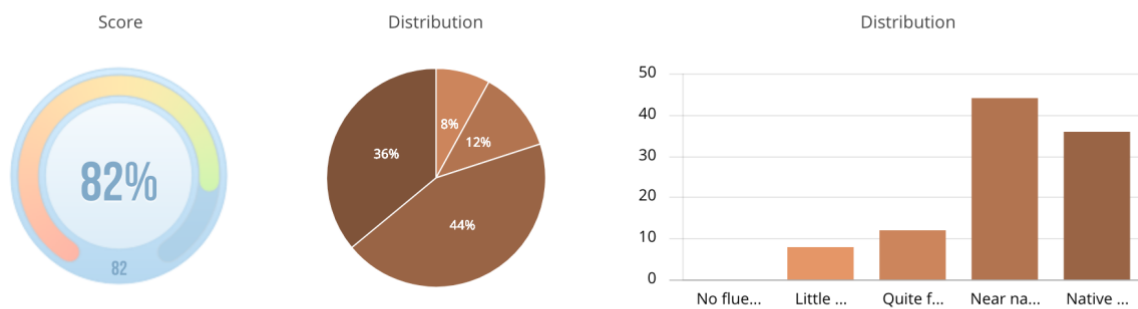
Anexo 11. Resultados de adecuacy y fluency en la posesición del texto AFSTYLA del traductor 4 (correspondiente al motor Google Translate).

Adequacy



Adequacy: All meaning expressed in the source fragment appears in the translation fragment.

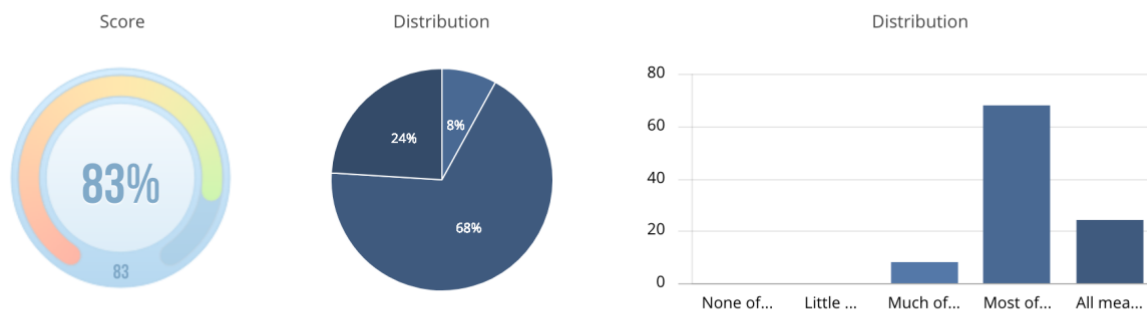
Fluency



Fluency: Native language fluency. No grammar errors, good word choice and syntactic structure. No post-editing required.

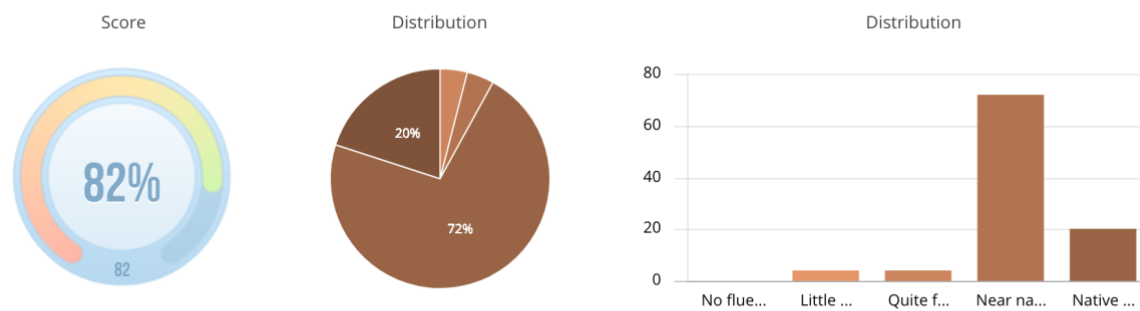
Anexo 12. Resultados de adecuacy y fluency en la posesición del texto AFSTYLA del traductor 5 (correspondiente al motor Google Translate).

Adequacy



Adequacy: All meaning expressed in the source fragment appears in the translation fragment.

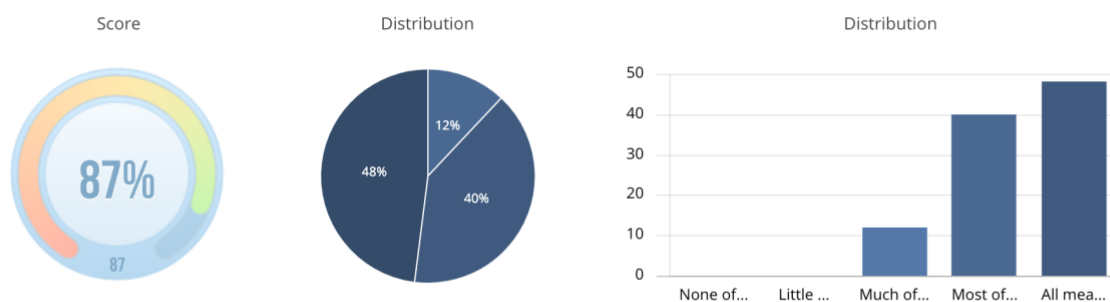
Fluency



Fluency: Native language fluency. No grammar errors, good word choice and syntactic structure. No post-editing required.

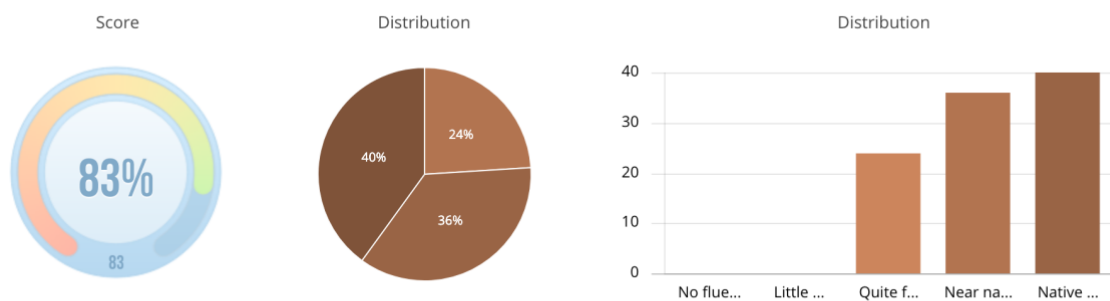
Anexo 13. Resultados de adecuacy y fluency en la posesición del texto AFSTYLA del traductor 6 (correspondiente al motor Google Translate).

Adequacy



Adequacy: All meaning expressed in the source fragment appears in the translation fragment.

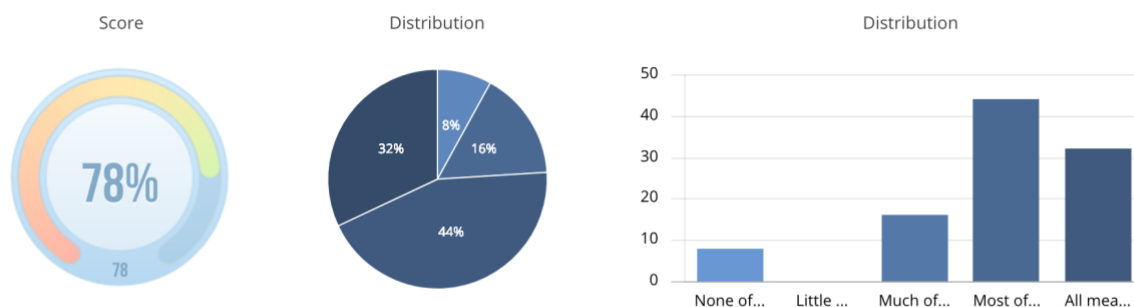
Fluency



Fluency: Native language fluency. No grammar errors, good word choice and syntactic structure. No post-editing required.

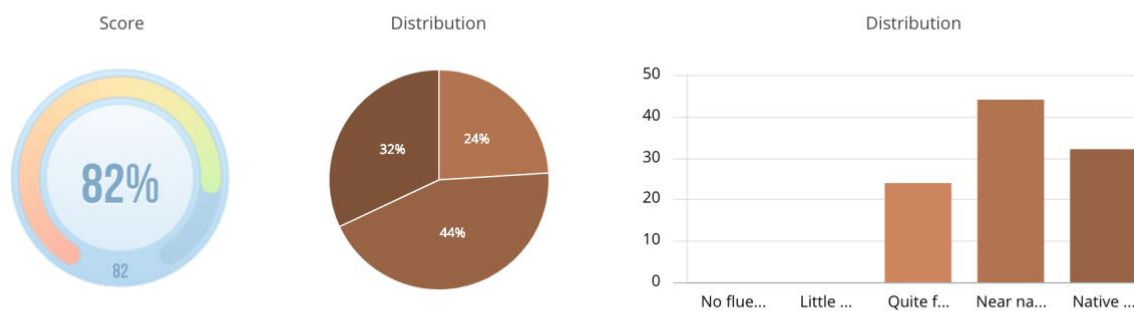
Anexo 14. Resultados de adecuacy y fluency en la posesición del texto AFSTYLA del traductor 7 (correspondiente al motor DeepL).

Adequacy



Adequacy: Most of the source fragment meaning is expressed in the translation fragment.

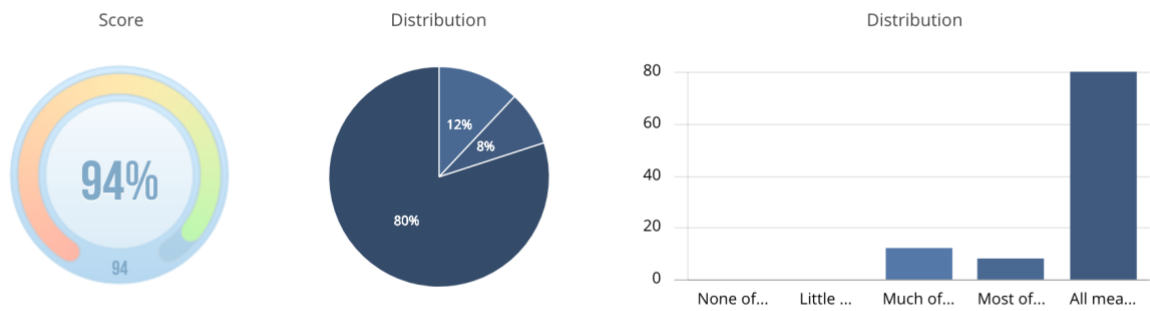
Fluency




Fluency: Native language fluency. No grammar errors, good word choice and syntactic structure. No post-editing required.

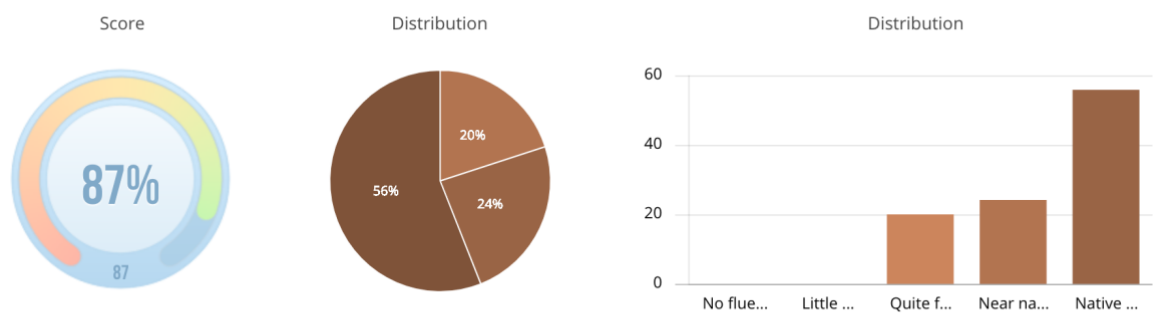
Anexo 15. Resultados de adecuacy y fluency en la posesición del texto AFSTYLA del traductor 8 (correspondiente al motor DeepL).


Adequacy



 **Adequacy:** All meaning expressed in the source fragment appears in the translation fragment.

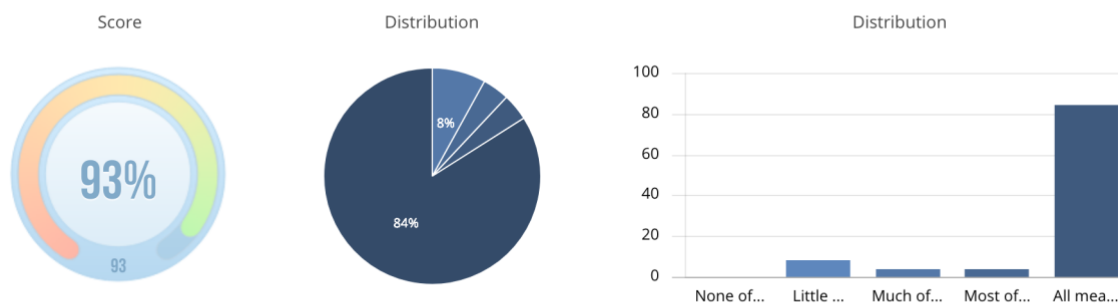
Fluency




 **Fluency:** Native language fluency. No grammar errors, good word choice and syntactic structure. No post-editing required.

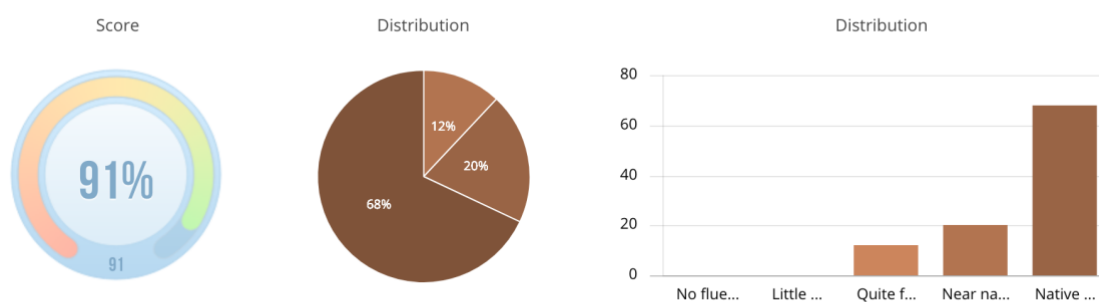
Anexo 16. Resultados de adecuacy y fluency en la posesición del texto AFSTYLA del traductor 9 (correspondiente al motor DeepL).


Adequacy



 **Adequacy:** All meaning expressed in the source fragment appears in the translation fragment.

Fluency



 **Fluency:** Native language fluency. No grammar errors, good word choice and syntactic structure. No post-editing required.

Anexo 17. Resultados de Interactive BLEU Score de Tilde MT que compara Google Translate (First Machine) con nuestro motor TAE especializado (Second Machine).

Sentence Nr	Source sentence	Human translated sentence	First machine translated sentence	Second Machine translated sentence	First machine translated sentence bleu score	Second machine translated sentence bleu score
1	EMA/753310/2016	ema / 753310 / 2016	ema / 753310 / 2016	ema / 753310 / 2016	100.000000	100.000000
2	EMA/H/C/004075	emea / h / c / 004075	emea / h / c / 004075	emea / h / c / 004075	100.000000	100.000000
3					100.000000	100.000000
4	EPAR summary for the public	resumen del epar para el público general	resumen del epar para el público .	resumen del epar para el público general	80.910671	100.000000
5					100.000000	100.000000
6	Afstyla	afstyla	afstyla	afstyla	100.000000	100.000000
7	lonococog alfa	lonococog alfa	lonococog alfa	lonococog alfa	100.000000	100.000000
8					100.000000	100.000000
9	This is a summary of the European public assessment report (EPAR) for Afstyla. It explains how the Agency assessed the medicine to recommend its authorisation in the EU and its conditions of use. It is not intended to provide practical advice on how to use Afstyla.	el presente documento resume el informe público europeo de evaluación (epar) de afstyla . en él se explica cómo la agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la ue y sus condiciones de uso . no está destinado a proporcionar consejos prácticos sobre cómo utilizar afstyla .	este es un resumen del informe de evaluación pública europea (epar) para afstyla . explica cómo la agencia evaluó el medicamento para recomendar su autorización en la ue y sus condiciones de uso . no pretende proporcionar consejos prácticos sobre cómo usar afstyla .	el presente documento resume el informe público europeo de evaluación (epar) para symtuza . en él se explica cómo la agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la ue y sus condiciones de uso . no está destinado a proporcionar consejos prácticos sobre cómo utilizar axumin .	30.255748	90.776162
10					100.000000	100.000000
11	For practical information about using Afstyla, patients should read the package leaflet or contact their doctor or pharmacist.	para más información sobre el tratamiento con afstyla , el paciente debe leer el prospecto (también incluido en el epar) o consultar a su médico o farmacéutico .	para obtener información práctica sobre el uso de afstyla , los pacientes deben leer el prospecto o ponerse en contacto con su médico o farmacéutico .	para más información sobre el tratamiento con lokelma , el paciente debe leer el prospecto (también incluido en el epar) o consultar a su médico o farmacéutico .	20.120074	90.959306
12					100.000000	100.000000
13	What is Afstyla and what is it used for?	¿qué es afstyla y para qué se utiliza ?	¿qué es afstyla y para qué se utiliza ?	¿qué es afstyla y para qué se utiliza ?	100.000000	100.000000
14					100.000000	100.000000
15	Afstyla is a medicine used for the treatment and prevention of bleeding in patients with haemophilia A (an inherited bleeding disorder caused by lack of a clotting protein called factor	afstyla es un medicamento que se utiliza para el tratamiento y la prevención de las hemorragias en	afstyla es un medicamento utilizado para el tratamiento y la prevención del sangrado en	afstyla es un medicamento que se utiliza para el tratamiento y prevención de hemorragias en	64.513514	60.829280

	VIII). It contains the active substance lonoctocog alfa.	pacientes con hemofilia a (un trastorno congénito de la coagulación causado por la falta de una proteína llamada factor viii) . contiene el principio activo lonoctocog alfa .	pacientes con hemofilia a (un trastorno hemorrágico hereditario causado por la falta de una proteína de coagulación llamada factor viii) . contiene el principio activo lonoctocog alfa .	pacientes con hemofilia a (trastorno hereditario de la coagulación causado por la carencia de una proteína denominada factor viii de la coagulación de la sangre) . contiene el principio activo lonoctocog alfa .		
16					100.000000	100.000000
17	How is Afstyla used?	¿cómo se usa afstyla ?	¿cómo se usa afstyla ?	¿cómo se usa afstyla ?	100.000000	100.000000
18					100.000000	100.000000
19	Afstyla is available as a powder and solvent used to make a solution for injection. The injection is given into a vein over several minutes. The dose and frequency of the injections depend on whether Afstyla is used to treat or prevent bleeding, the severity of the patient's factor VIII deficiency, the extent and location of the bleeding and the patient's condition and bodyweight.	afstyla se presenta como un polvo y un disolvente para preparar una solución inyectable . la inyección se administra en una vena durante varios minutos . la dosis y la frecuencia de las inyecciones dependen de si afstyla se utiliza para el tratamiento y la prevención de hemorragias , del grado de deficiencia de factor viii en el paciente , del grado y la localización de la hemorragia y del estado y el peso corporal del paciente .	afstyla está disponible como un polvo y un disolvente utilizado para hacer una solución inyectable . la inyección se administra en una vena durante varios minutos . la dosis y la frecuencia de las inyecciones dependen de si se usa afstyla para tratar o prevenir el sangrado , la gravedad de la deficiencia de factor viii del paciente , la extensión y la ubicación del sangrado y la condición y el peso corporal del paciente .	afstyla se presenta en forma de polvo y disolvente para preparar una solución para inyección . la inyección se administra en una vena durante varios minutos . la dosis y la frecuencia de la inyección dependerán de si afstyla se usa para tratar o prevenir la hemorragia , la gravedad de la deficiencia del factor viii , de la localización y el grado de la hemorragia y del estado del paciente y de su peso corporal .	53.484286	43.856807
20					100.000000	100.000000
21	Afstyla can only be obtained with a prescription and treatment should be under the supervision of a doctor who has experience in the treatment of haemophilia. For further information, see the package leaflet.	afstyla solo se podrá dispensar con receta médica y el tratamiento debe iniciarse bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia . para más información , consulte el prospecto .	afstyla solo puede obtenerse con receta médica y el tratamiento debe estar bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia . para más información , consulte el prospecto .	afstyla solo se podrá dispensar con receta médica y el tratamiento deberá administrarse bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia . para más información , consulte el prospecto .	80.070402	89.667734
22					100.000000	100.000000
23	How does Afstyla work?	¿cómo actúa afstyla ?	¿cómo funciona afstyla ?	¿cómo actúa afstyla ?	35.355339	100.000000
24					100.000000	100.000000
25	Patients with haemophilia A lack factor VIII, a protein needed for normal clotting of the blood, and as a result, they bleed readily.	los pacientes con hemofilia a carecen del factor viii , una proteína necesaria para la coagulación normal de la sangre y	los pacientes con hemofilia a carecen de factor viii , una proteína necesaria para la coagulación normal de la sangre y	pacientes con hemofilia a carecen del factor viii , una proteína necesaria para el normal de la coagulación de la sangre , y como	90.251398	56.944723

		, como resultado , sangran fácilmente .	, como resultado , sangran fácilmente .	resultado , pueden sangrar fácilmente .		
26					100.000000	100.000000
27	30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom	30 churchill place • canary wharf • london e14 5eu • united kingdom	30 churchill place • canary wharf • londres e14 5eu • reino unido	30 churchill lugar • circus canary wharf londres e14 5eu • reino unido	57.835699	17.474336
28	Telephone +44 (0)20 3660 6000 Facsimile +44 (0)20 3660 5555	telephone + 44 (0) 20 3660 6000 facsimile + 44 (0) 20 3660 5555	teléfono + 44 (0) 20 3660 6000 fax + 44 (0) 20 3660 5555	teléfono : + 44 (0) 20 - 3660 - 6000 facsimile + 44 (0) 20 en 3660 5555	77.781112	56.333168
29	Send a question via our website www.ema.europa.eu/contact	send a question via our website www . ema . europa . eu / contact	enviar una pregunta a través de nuestro sitio web www . ema . europa . eu / contact	enviar una pregunta a través de nuestro sitio web www . ema . europa . eu / contact	46.248926	46.248926
30					100.000000	100.000000
31	An agency of the European Union	an agency of the european union	una agencia de la unión europea	un organismo de la unión europea	4.058349	4.058349
32					100.000000	100.000000
33	© European Medicines Agency, 2017. Reproduction is authorised provided the source is acknowledged.	© european medicines agency , 2017 . reproduction is authorised provided the source is acknowledged .	© agencia europea de medicamentos , 2017 . la reproducción está autorizada siempre que se cite la fuente .	© european medicines agency , 2016 . reproduction is authorised provided the source is acknowledged .	8.562365	81.535510
34					100.000000	100.000000
35	The active substance in Afstyla, lonoctocog alfa, is a shorter version of human factor VIII that works in the body in the same way as human factor VIII. It replaces the missing factor VIII, thereby helping the blood to clot and giving temporary control of bleeding.	el principio activo de afstyla , lonoctocog alfa , es una versión más corta del factor viii humano que actúa en el organismo de la misma forma que el factor viii humano . sustituye al factor viii ausente , con lo que ayuda a la sangre a coagular y permite el control temporal de la hemorragia .	el principio activo en afstyla , lonoctocog alfa , es una versión más corta del factor viii humano que funciona en el cuerpo de la misma manera que el factor viii humano . reemplaza el factor viii faltante , lo que ayuda a que la sangre se coagule y proporciona un control temporal del sangrado .	el principio activo de afstyla , lonoctocog alfa , es una versión humana máxima del factor viii , que actúa en el organismo de la misma forma que la del factor viii humano . se sustituye el factor viii que falta , lo que contribuye a coagular la sangre y administración de un control temporal de la hemorragia .	46.256408	53.549314
36					100.000000	100.000000
37	What benefits of Afstyla have been shown in studies?	¿qué beneficios ha demostrado tener afstyla en los estudios realizados ?	¿qué beneficios ha demostrado tener afstyla en los estudios realizados ?	¿qué beneficios ha demostrado tener apealea en los estudios realizados ?	100.000000	70.168794
38					100.000000	100.000000
39	Afstyla has been shown to be effective at preventing and treating bleeding episodes in two main studies in patients with severe haemophilia A who were previously treated with other factor VIII products.	se ha demostrado que afstyla es eficaz para la prevención y el tratamiento de episodios hemorrágicos en dos estudios principales realizados en pacientes con hemofilia a grave que habían recibido tratamiento previo con otros productos	afstyla ha demostrado ser eficaz para prevenir y tratar los episodios de sangrado en dos estudios principales en pacientes con hemofilia a grave que fueron tratados previamente con otros productos de factor viii .	afstyla ha demostrado ser eficaz para prevenir y tratar episodios hemorrágicos en dos estudios principales en pacientes con hemofilia a grave que hayan sido previamente tratados con otros medicamentos de factor viii .	29.413358	33.237605

		que contenían factor viii .				
40					100.000000	100.000000
41	The first study involved 173 patients aged 12 years and above. A total of 848 bleeding episodes were recorded during the study, 94% of which resolved with one or two injections of Afstyla. Afstyla was rated as 'excellent' or 'good' at treating 92% of the bleeding episodes.	el primer estudio se realizó en 173 pacientes de 12 o más años de edad . se registraron un total de 848 episodios hemorrágicos durante el estudio , 94 % de los cuales remitieron con una o dos inyecciones de afstyla . la eficacia de afstyla se consideró «excelente» o «buena» para el tratamiento del 92 % de los episodios hemorrágicos .	el primer estudio involucró a 173 pacientes de 12 años y más . durante el estudio se registraron un total de 848 episodios de sangrado , el 94% de los cuales se resolvieron con una o dos inyecciones de afstyla . afstyla fue calificada como " excelente " o " buena " en el tratamiento del 92% de los episodios de sangrado .	en el primer estudio participaron 194 pacientes de 12 años y mayores . un total de 848 episodios hemorrágicos fueron registrados durante el estudio , el 94% de los cuales se resolvió con una o dos inyecciones de afstyla . afstyla fue calificada como " excelente " o " buena " en el tratamiento del 92% de los episodios hemorrágicos .	35.168906	36.267166
42					100.000000	100.000000
43	In the 16 surgical procedures that occurred during the study, Afstyla was rated as 'excellent' or 'good' in preventing bleeding episodes when used two to three times a week. Among those patients who received Afstyla to prevent bleeding, an average of 1.14 bleeding episodes per year was recorded for each patient which was lower than the 19.64 episodes per year in patients not receiving Afstyla for prevention.	en los 16 procedimientos quirúrgicos que tuvieron lugar durante el estudio , la eficacia de afstyla se consideró «excelente» o «buena» para la prevención de episodios hemorrágicos cuando se utilizó dos o tres veces a la semana . entre los pacientes que recibieron afstyla para prevenir hemorragias , se registró un promedio de 1,14 episodios hemorrágicos al año en cada paciente , un número inferior a los 19,64 episodios al año registrados en pacientes que no recibieron afstyla para fines de prevención .	en los 16 procedimientos quirúrgicos que se produjeron durante el estudio , afstyla se calificó de " excelente " o " buena " en la prevención de episodios de sangrado cuando se usó de dos a tres veces por semana . entre los pacientes que recibieron afstyla para prevenir el sangrado , se registró un promedio de 1.14 episodios de sangrado por año para cada paciente , que fue inferior a los 19.64 episodios por año en pacientes que no recibieron afstyla para la prevención .	en 16 procedimientos quirúrgicos que se produjeron durante el estudio , afstyla fue calificada como " excelente " o " buena " en la prevención de episodios de hemorragia cuando se utilizan dos o tres veces a la semana . entre los pacientes que recibieron afstyla para prevenir la hemorragia , una media de 1,14 episodios hemorrágicos al año . se registró para cada paciente que fue menor que la de episodios 19.64 al año en los pacientes que no recibieron afstyla para la prevención .	40.521949	43.024393
44					100.000000	100.000000
45	The second study involved 83 patients below 12 years of age. Afstyla was rated as 'excellent' or 'good' at treating 96% of the 347 bleeding episodes recorded during the study; 96% of bleeding episodes resolved with one or two injections of Afstyla. Among those patients who received Afstyla to prevent bleeding, the average number of bleeding episodes per year was 2.30 for patients receiving Afstyla three times a week, and	en el segundo estudio participaron 83 pacientes menores de 12 años . la eficacia de afstyla se consideró «excelente» o «buena» para el tratamiento del 96 % de los 347 episodios hemorrágicos registrados durante el estudio ; el 96 % de los episodios hemorrágicos	el segundo estudio involucró a 83 pacientes menores de 12 años . afstyla fue calificada como " excelente " o " buena " en el tratamiento del 96% de los 347 episodios de sangrado registrados durante el estudio ; el 96% de los episodios de sangrado se resolvieron con una o dos inyecciones de	en el segundo estudio participaron un total de 83 pacientes menores de 12 años de edad . afstyla fue calificada como " excelente " o " buena " en el tratamiento de 96% de los 353 episodios hemorrágicos registrados durante el estudio ; el 96% de los episodios de hemorragia se	51.239214	47.945240

	4.37 for those receiving Afstyla twice a week.	remitieron con una o dos inyecciones de afstyla . entre los pacientes que recibieron afstyla para la prevención de hemorragias , el número promedio de episodios hemorrágicos al año fue de 2,30 en los pacientes que recibieron afstyla tres veces por semana y de 4,37 en los que recibieron afstyla dos veces por semana .	afstyla . entre los pacientes que recibieron afstyla para prevenir el sangrado , el número promedio de episodios de sangrado por año fue de 2.30 para los pacientes que recibieron afstyla tres veces por semana , y de 4.37 para los que recibieron afstyla dos veces por semana .	solucionaron con una o dos inyecciones de afstyla . entre los pacientes que recibieron afstyla para prevenir la hemorragia , el número medio de episodios hemorrágicos al año fue 2,30 para los pacientes que recibieron afstyla tres veces por semana , 4,37 para los pacientes que recibieron afstyla dos veces a la semana .		
46					100.000000	100.000000
47	What are the risks associated with Afstyla?	¿cuál es el riesgo asociado a afstyla ?	¿cuáles son los riesgos asociados con afstyla ?	¿cuáles son los riesgos asociados a apealea ?	11.044796	6.567275
48					100.000000	100.000000
49	Hypersensitivity (allergic) reactions are common with Afstyla, affecting up to 1 in 10 people. They may include: angioedema (swelling of tissues under the skin), burning and stinging at the injection site, chills, flushing, itchy rash over the whole body, headache, hives, hypotension (low blood pressure), lethargy, nausea (feeling sick), restlessness, tachycardia (rapid heartbeat), tightness of the chest, tingling, vomiting and wheezing. In some cases these reactions can become severe.	las reacciones de hipersensibilidad (alérgicas) son frecuentes con afstyla y afectan a hasta 1 de cada 10 personas . estas reacciones pueden consistir en : angioedema (inflamación de los tejidos debajo de la piel) , ardor y hormigueo en el lugar de la inyección , escalofríos , sofocos , erupción cutánea con picor en todo el cuerpo , dolor de cabeza , sarpullido , hipotensión (presión arterial baja) , letargo , náuseas , inquietud , taquicardia (latido cardíaco rápido) , opresión torácica , hormigueo , vómitos y pitidos al respirar (sibilancias) . en algunos casos , estas reacciones pueden llegar a ser graves .	las reacciones de hipersensibilidad (alérgicas) son comunes con afstyla y afectan a 1 de cada 10 personas . pueden incluir : angioedema (hinchazón de los tejidos debajo de la piel) , ardor y picazón en el lugar de la inyección , escalofríos , enrojecimiento , erupción con picor en todo el cuerpo , dolor de cabeza , urticaria , hipotensión (presión arterial baja) , letargo , náuseas (sensación de malestar) , inquietud , taquicardia (latidos rápidos del corazón) , opresión en el pecho , hormigueo , vómitos y sibilancias . en algunos casos estas reacciones pueden llegar a ser graves .	reacciones alérgicas (hipersensibilidad) son frecuentes con afstyla , afectan hasta 1 de cada 10 personas . glucosa en sangre : angioedema (hinchazón de tejidos debajo de la piel) , quemazón y picor en el lugar de inyección , escalofríos , enrojecimiento de la piel , erupción cutánea con picor en todo el cuerpo , dolor de cabeza , urticaria , hipotensión (presión arterial baja) , letargo , náuseas (ganas de vomitar) , cansancio , taquicardia (palpitaciones) , opresión en el pecho , hormigueo , vómitos y respiración sibilante . en algunos casos , estas reacciones pueden ser graves .	61.755099	51.560921
50					100.000000	100.000000
51	There is also a risk with factor VIII medicines that some patients will develop inhibitors (antibodies) against factor VIII, causing the medicine to stop working and resulting in a loss of bleeding control.	existe también un riesgo asociado a los medicamentos que contienen factor viii de que algunos pacientes desarrollen inhibidores (anticuerpos) contra el factor viii , lo que hace que el	también existe un riesgo con los medicamentos del factor viii de que algunos pacientes desarrollen inhibidores (anticuerpos) contra el factor viii , lo que hace que el	también hay un riesgo de factor viii medicamentos que algunos pacientes que desarrollan anticuerpos (inhibidores) contra el factor viii , causando el medicamento dejar el	60.655865	18.844559

		medicamento deje de funcionar , con la consiguiente pérdida del control de las hemorragias .	medicamento deje de funcionar y provoque una pérdida del control del sangrado .	trabajo y dando como resultado la pérdida de control de la hemorragia .		
52					100.000000	100.000000
53	Afstyla must not be used in patients with known allergy to hamster proteins.	afstyla no debe utilizarse en pacientes con alergia conocida a proteínas de hámster .	afstyla no debe usarse en pacientes con alergia conocida a las proteínas de hámster .	afstyla no debe utilizarse en pacientes con alergia conocida a las proteínas de hámster .	57.735027	80.032032
54					100.000000	100.000000
55	For the full list of all side effects and restrictions, see the package leaflet.	para consultar la lista completa de efectos adversos notificados y las restricciones de afstyla , consultar el prospecto .	para ver la lista completa de todos los efectos secundarios y restricciones , consulte el prospecto .	la lista completa de efectos adversos y restricciones puede consultarse en el prospecto .	21.063358	32.762648
56					100.000000	100.000000
57	Why is Afstyla approved?	¿por qué se ha aprobado afstyla ?	¿por qué se aprueba afstyla ?	¿por qué se ha aprobado afstyla ?	32.159351	100.000000
58					100.000000	100.000000
59	Afstyla has been shown to be effective in both preventing and treating bleeding episodes. Regarding safety, the reported side effects are in line with those expected for a factor VIII product, although hypersensitivity reactions occurred more frequently with Afstyla. Further safety data should be provided by ongoing studies.	se ha demostrado que afstyla es eficaz tanto para la prevención como para el tratamiento de episodios hemorrágicos . en cuanto a la seguridad , los efectos secundarios comunicados son similares a los afstyla esperados para un producto que contiene factor viii , aunque las reacciones de hipersensibilidad fueron más frecuentes con afstyla . los estudios en curso deberían facilitar datos adicionales sobre la seguridad .	afstyla ha demostrado ser eficaz tanto en la prevención como en el tratamiento de los episodios de sangrado . en cuanto a la seguridad , los efectos secundarios informados están en línea con los esperados para un producto de factor viii , aunque las reacciones de hipersensibilidad ocurrieron con más frecuencia con afstyla . los estudios en curso deben proporcionar más datos de seguridad .	afstyla ha demostrado ser eficaz en la prevención y tratamiento de episodios hemorrágicos . en cuanto a la seguridad , los efectos adversos se encuentran en línea con los esperados para un medicamento de factor viii , aunque las reacciones de hipersensibilidad se produjeron con mayor frecuencia con afstyla . nuevos datos de seguridad debe ser suministrada por ensayos en marcha .	42.720202	35.584607
60					100.000000	100.000000
61	The Agency's Committee for Medicinal Products for Human Use (CHMP) decided that Afstyla's benefits are greater than its risks and recommended that it be approved for use in the EU.	el comité de medicamentos de uso humano (chmp) de la agencia decidió que los beneficios de afstyla son mayores que sus riesgos y recomendó autorizar su uso en la ue .	el comité de medicamentos de uso humano (chmp) de la agencia decidió que los beneficios de afstyla son mayores que sus riesgos y recomendó que se aprobara para su uso en la ue .	la agencia en el comité de medicamentos de uso humano (chmp) decidió que los beneficios de afstyla son mayores que sus riesgos y recomendó autorizar su uso en la ue .	83.851300	87.628711
62					100.000000	100.000000
63	Afstyla	afstyla	afstyla	afstyla	100.000000	100.000000
64	EMA/753310/2016	ema / 753310 / 2016	ema / 753310 / 2016	ema / 753310 / 2016	100.000000	100.000000
65					100.000000	100.000000

66	Page 2/3	página 2 / 3	página 2 / 3	página 2 / 3	100.000000	100.000000
67					100.000000	100.000000
68	What measures are being taken to ensure the safe and effective use of Afstyly?	¿qué medidas se han adoptado para garantizar un uso seguro y eficaz de afstyly ?	¿qué medidas se han adoptado para garantizar un uso seguro y eficaz de afstyly ?	¿qué medidas se han adoptado para garantizar un uso seguro y eficaz de rubraca ?	100.000000	86.664157
69					100.000000	100.000000
70	Recommendations and precautions to be followed by healthcare professionals and patients for the safe and effective use of Afstyly have been included in the summary of product characteristics and the package leaflet.	las recomendaciones y precauciones que deben seguir los profesionales sanitarios y los pacientes para un uso seguro y eficaz de afstyly se han incluido en la ficha técnica o resumen de las características del producto y el prospecto .	las recomendaciones y precauciones que deben seguir los profesionales de la salud y los pacientes para el uso seguro y eficaz de afstyly se han incluido en el resumen de las características del producto y en el prospecto .	las recomendaciones y precauciones que deben seguir los profesionales sanitarios y los pacientes para un uso seguro y eficaz de rubraca se han incluido en la ficha técnica o resumen de las características del producto y el prospecto .	66.845352	93.183848
71					100.000000	100.000000
72	Other information about Afstyly	otras informaciones sobre afstyly	otras informaciones sobre afstyly .	otras informaciones sobre afstyly	66.874030	100.000000
73					100.000000	100.000000
74	The European Commission granted a marketing authorisation valid throughout the European Union for Afstyly on 4 January 2017.	el epar completo de afstyly se puede consultar en la página web de la agencia : ema . europa . eu / find medicine / human medicines / european public assessment reports . para mayor información sobre el tratamiento con afstyly , lea el prospecto (también incluido en el epar) o consulte a su médico o farmacéutico .	la comisión europea emitió una autorización de comercialización válida en toda la unión europea para afstyly el 4 de enero de 2017 .	la comisión europea emitió una autorización de comercialización válida en toda la unión europea para el medicamento afstyly el 4 de enero de 2017 .	0.570845	0.660492
75					100.000000	100.000000
76	The full EPAR for Afstyly can be found on the Agency's website: ema.europa.eu/FindMedicine/HumanMedicines/European public assessment reports . For more information about treatment with Afstyly, read the package leaflet (also part of the EPAR) or contact your doctor or pharmacist.	afstyly	el epar completo de afstyly se puede encontrar en el sitio web de la agencia : ema . europa . eu / encuentre medicamentos / medicamentos humanos / informes públicos europeos de evaluación . para obtener más información sobre el tratamiento con afstyly , lea el prospecto (que también forma parte del epar) o consulte a su médico o farmacéutico .	el texto completo del epar de afstyly puede consultarse en la página web de la agencia española : ema . europa . eu / find medicine / find medicines / human public assessment reports . para mayor información sobre el tratamiento con lokelma , lea el prospecto (también incluido en el epar) o consulte a su médico o farmacéutico .	0.574979	0.584486

Anexo 18. Resultados de Interactive BLEU Score de Tilde MT que compara DeepL (First Machine) con nuestro motor TAE especializado (Second Machine).

Sentence Nr	Source sentence	Human translated sentence	First machine translated sentence	Second Machine translated sentence	First machine translated sentence bleu score	Second machine translated sentence bleu score
1	EMA/753310/2016	ema / 753310 / 2016	ema / 753310 / 2016	ema / 753310 / 2016	100.000000	100.000000
2	EMA/H/C/004075	emea / h / c / 004075	emea / h / c / 004075	emea / h / c / 004075	100.000000	100.000000
3					100.000000	100.000000
4	EPAR summary for the public	resumen del epar para el público general	resumen de epar para el público	resumen del epar para el público general	45.480190	100.000000
5					100.000000	100.000000
6	Afstyla	afstyla	afstyla	afstyla	100.000000	100.000000
7	lonoctocog alfa	lonoctocog alfa	alfa lonoctocog	lonoctocog alfa	84.089642	100.000000
8					100.000000	100.000000
9	This is a summary of the European public assessment report (EPAR) for Afstyla. It explains how the Agency assessed the medicine to recommend its authorisation in the EU and its conditions of use. It is not intended to provide practical advice on how to use Afstyla.	el presente documento resume el informe público europeo de evaluación (epar) de afstyla . en él se explica cómo la agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la ue y sus condiciones de uso . no está destinado a proporcionar consejos prácticos sobre cómo utilizar afstyla .	este es un resumen del informe de evaluación pública europea (epar) de afstyla . explica cómo la agencia evaluó el medicamento para recomendar su autorización en la ue y sus condiciones de uso . no tiene por objeto proporcionar consejos prácticos sobre cómo utilizar afstyla .	el presente documento resume el informe público europeo de evaluación (epar) para symtuza . en él se explica cómo la agencia ha evaluado el medicamento y ha emitido un dictamen favorable a la autorización de comercialización en la ue y sus condiciones de uso . no está destinado a proporcionar consejos prácticos sobre cómo utilizar axumin .	39.736962	90.776162
10					100.000000	100.000000
11	For practical information about using Afstyla, patients should read the package leaflet or contact their doctor or pharmacist.	para más información sobre el tratamiento con afstyla , el paciente debe leer el prospecto (también incluido en el epar) o consultar a su médico o farmacéutico .	para obtener información práctica sobre el uso de afstyla , los pacientes deben leer el prospecto o ponerse en contacto con su médico o farmacéutico .	para más información sobre el tratamiento con lokelma , el paciente debe leer el prospecto (también incluido en el epar) o consultar a su médico o farmacéutico .	20.120074	90.959306
12					100.000000	100.000000
13	What is Afstyla and what is it used for?	¿qué es afstyla y para qué se utiliza ?	¿qué es afstyla y para qué sirve ?	¿qué es afstyla y para qué se utiliza ?	62.401954	100.000000
14					100.000000	100.000000
15	Afstyla is a medicine used for the treatment and prevention of bleeding in patients with	afstyla es un medicamento que se utiliza para el	afstyla es un medicamento utilizado para el	afstyla es un medicamento que se utiliza para el	52.053701	60.829280

	haemophilia A (an inherited bleeding disorder caused by lack of a clotting protein called factor VIII). It contains the active substance lonoctocog alfa.	tratamiento y la prevención de las hemorragias en pacientes con hemofilia a (un trastorno congénito de la coagulación causado por la falta de una proteína llamada factor viii) . contiene el principio activo lonoctocog alfa .	tratamiento y prevención de hemorragias en pacientes con hemofilia a (un trastorno hemorrágico hereditario causado por la falta de una proteína coagulante llamada factor viii) . contiene la sustancia activa lonoctocog alfa .	tratamiento y prevención de hemorragias en pacientes con hemofilia a (trastorno hereditario de la coagulación causado por la carencia de una proteína denominada factor viii de la coagulación de la sangre) . contiene el principio activo lonoctocog alfa .		
16					100.000000	100.000000
17	How is Afstyla used?	¿cómo se usa afstyla ?	¿cómo se usa afstyla ?	¿cómo se usa afstyla ?	100.000000	100.000000
18					100.000000	100.000000
19	Afstyla is available as a powder and solvent used to make a solution for injection. The injection is given into a vein over several minutes. The dose and frequency of the injections depend on whether Afstyla is used to treat or prevent bleeding, the severity of the patient's factor VIII deficiency, the extent and location of the bleeding and the patient's condition and bodyweight.	afstyla se presenta como un polvo y un disolvente para preparar una solución inyectable . la inyección se administra en una vena durante varios minutos . la dosis y la frecuencia de las inyecciones dependen de si afstyla se utiliza para el tratamiento y la prevención de hemorragias , del grado de deficiencia de factor viii en el paciente , del grado y la localización de la hemorragia y del estado y el peso corporal del paciente .	afstyla está disponible en polvo y disolvente para hacer una solución inyectable . la inyección se administra en una vena durante varios minutos . la dosis y la frecuencia de las inyecciones dependen de si afstyla se utiliza para tratar o prevenir el sangrado , la gravedad de la deficiencia de factor viii del paciente , el grado y la ubicación del sangrado y la condición y el peso corporal del paciente .	afstyla se presenta en forma de polvo y disolvente para preparar una solución para inyección . la inyección se administra en una vena durante varios minutos . la dosis y la frecuencia de la inyección dependerán de si afstyla se usa para tratar o prevenir la hemorragia , la gravedad de la deficiencia del factor viii , de la localización y el grado de la hemorragia y del estado del paciente y de su peso corporal .	55.080716	43.856807
20					100.000000	100.000000
21	Afstyla can only be obtained with a prescription and treatment should be under the supervision of a doctor who has experience in the treatment of haemophilia. For further information, see the package leaflet.	afstyla solo se podrá dispensar con receta médica y el tratamiento debe iniciarse bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia . para más información , consulte el prospecto .	afstyla sólo puede obtenerse con receta médica y el tratamiento debe realizarse bajo la supervisión de un médico que tenga experiencia en el tratamiento de la hemofilia . para más información , véase el prospecto .	afstyla solo se podrá dispensar con receta médica y el tratamiento deberá administrarse bajo la supervisión de un médico con experiencia en el tratamiento de la hemofilia . para más información , consulte el prospecto .	63.001219	89.667734
22					100.000000	100.000000
23	How does Afstyla work?	¿cómo actúa afstyla ?	¿cómo funciona afstyla ?	¿cómo actúa afstyla ?	35.355339	100.000000
24					100.000000	100.000000
25	Patients with haemophilia A lack factor VIII, a protein needed for normal clotting of	los pacientes con hemofilia a carecen del factor viii , una proteína necesaria	los pacientes con hemofilia a carecen del factor viii , una proteína necesaria	pacientes con hemofilia a carecen del factor viii , una proteína necesaria	88.073793	56.944723

	the blood, and as a result, they bleed readily.	para la coagulación normal de la sangre y , como resultado , sangran fácilmente .	para la coagulación normal de la sangre , y como resultado , sangran fácilmente .	para el normal de la coagulación de la sangre , y como resultado , pueden sangrar fácilmente .		
26					100.000000	100.000000
27	30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom	30 churchill place • canary wharf • london e14 5eu • united kingdom	30 churchill lugar • canary wharf • londres e14 5eu • reino unido	30 churchill lugar • circus canary wharf londres e14 5eu • reino unido	31.170907	17.474336
28	Telephone +44 (0)20 3660 6000 Facsimile +44 (0)20 3660 5555	telephone + 44 (0) 20 3660 6000 facsimile + 44 (0) 20 3660 5555	teléfono : + 44 (0) 20 - 3660 - 6000 facsimile + 44 (0) 20 en 3660 5555	teléfono : + 44 (0) 20 - 3660 - 6000 facsimile + 44 (0) 20 en 3660 5555	56.333168	56.333168
29	Send a question via our website www.ema.europa.eu/contact	send a question via our website www . ema . europa . eu / contact	envíe una pregunta a través de nuestra página web www . ema . europa . eu / contact	enviar una pregunta a través de nuestro sitio web www . ema . europa . eu / contact	46.248926	46.248926
30					100.000000	100.000000
31	An agency of the European Union	an agency of the european union	una agencia de la unión europea	un organismo de la unión europea	4.058349	4.058349
32					100.000000	100.000000
33	© European Medicines Agency, 2017. Reproduction is authorised provided the source is acknowledged.	© european medicines agency , 2017 . reproduction is authorised provided the source is acknowledged .	© agencia europea de medicamentos , 2017 . reproducción autorizada con indicación de la fuente bibliográfica .	© european medicines agency , 2016 . reproduction is authorised provided the source is acknowledged .	9.672650	81.535510
34					100.000000	100.000000
35	The active substance in Afstyla, lonoctocog alfa, is a shorter version of human factor VIII that works in the body in the same way as human factor VIII. It replaces the missing factor VIII, thereby helping the blood to clot and giving temporary control of bleeding.	el principio activo de afstyla , lonoctocog alfa , es una versión más corta del factor viii humano que actúa en el organismo de la misma forma que el factor viii humano . sustituye al factor viii ausente , con lo que ayuda a la sangre a coagular y permite el control temporal de la hemorragia .	el principio activo de afstyla , lonoctocog alfa , es una versión más corta del factor viii humano que funciona en el cuerpo de la misma manera que el factor viii humano . reemplaza el factor viii que falta , ayudando así a la sangre a coagularse y dando un control temporal de la hemorragia .	el principio activo de afstyla , lonoctocog alfa , es una versión humana máxima del factor viii , que actúa en el organismo de la misma forma que la del factor viii humano . se sustituye el factor viii que falta , lo que contribuye a coagular la sangre y administración de un control temporal de la hemorragia .	57.526452	53.549314
36					100.000000	100.000000
37	What benefits of Afstyla have been shown in studies?	¿qué beneficios ha demostrado tener afstyla en los estudios realizados ?	¿qué beneficios de afstyla han sido demostrados en estudios ?	¿qué beneficios ha demostrado tener apealea en los estudios realizados ?	9.993744	70.168794
38					100.000000	100.000000
39	Afstyla has been shown to be effective at preventing and treating bleeding episodes in two main studies in patients with severe haemophilia A who were previously treated with other factor VIII products.	se ha demostrado que afstyla es eficaz para la prevención y el tratamiento de episodios hemorrágicos en dos estudios principales realizados en pacientes con hemofilia a grave que	se ha demostrado que afstyla es eficaz para prevenir y tratar los episodios hemorrágicos en dos estudios principales en pacientes con hemofilia a grave que fueron tratados previamente con	afstyla ha demostrado ser eficaz para prevenir y tratar episodios hemorrágicos en dos estudios principales en pacientes con hemofilia a grave que hayan sido previamente tratados	49.177667	33.237605

		habían recibido tratamiento previo con otros productos que contenían factor viii .	otros productos de factor viii .	con otros medicamentos de factor viii .		
40					100.000000	100.000000
41	The first study involved 173 patients aged 12 years and above. A total of 848 bleeding episodes were recorded during the study, 94% of which resolved with one or two injections of Afstyla. Afstyla was rated as 'excellent' or 'good' at treating 92% of the bleeding episodes.	el primer estudio se realizó en 173 pacientes de 12 o más años de edad . se registraron un total de 848 episodios hemorrágicos durante el estudio , 94 % de los cuales remitieron con una o dos inyecciones de afstyla . la eficacia de afstyla se consideró «excelente» o «buena» para el tratamiento del 92 % de los episodios hemorrágicos .	en el primer estudio participaron 173 pacientes de 12 años o más . se registraron un total de 848 episodios de sangrado durante el estudio , el 94% de los cuales se resolvieron con una o dos inyecciones de afstyla . afstyla fue calificado como " excelente " o " bueno " en el tratamiento del 92% de los episodios hemorrágicos .	en el primer estudio participaron 194 pacientes de 12 años y mayores . un total de 848 episodios hemorrágicos fueron registrados durante el estudio , el 94% de los cuales se resolvió con una o dos inyecciones de afstyla . afstyla fue calificada como " excelente " o " buena " en el tratamiento del 92% de los episodios hemorrágicos .	42.558505	36.267166
42					100.000000	100.000000
43	In the 16 surgical procedures that occurred during the study, Afstyla was rated as 'excellent' or 'good' in preventing bleeding episodes when used two to three times a week. Among those patients who received Afstyla to prevent bleeding, an average of 1.14 bleeding episodes per year was recorded for each patient which was lower than the 19.64 episodes per year in patients not receiving Afstyla for prevention.	en los 16 procedimientos quirúrgicos que tuvieron lugar durante el estudio , la eficacia de afstyla se consideró «excelente» o «buena» para la prevención de episodios hemorrágicos cuando se utilizó dos o tres veces a la semana . entre los pacientes que recibieron afstyla para prevenir hemorragias , se registró un promedio de 1,14 episodios hemorrágicos al año en cada paciente , un número inferior a los 19,64 episodios al año registrados en pacientes que no recibieron afstyla para fines de prevención .	en los 16 procedimientos quirúrgicos que ocurrieron durante el estudio , afstyla fue calificado como " excelente " o " bueno " en la prevención de episodios de sangrado cuando se utilizó dos o tres veces por semana . entre los pacientes que recibieron afstyla para prevenir hemorragias , se registró un promedio de 1,14 episodios de hemorragia por año para cada paciente , lo que fue inferior a los 19,64 episodios por año en los pacientes que no recibieron afstyla para la prevención .	en 16 procedimientos quirúrgicos que se produjeron durante el estudio , afstyla fue calificada como " excelente " o " buena " en la prevención de episodios de hemorragia cuando se utilizan dos o tres veces a la semana . entre los pacientes que recibieron afstyla para prevenir la hemorragia , una media de 1,14 episodios hemorrágicos al año . se registró para cada paciente que fue menor que la de episodios 19.64 al año en los pacientes que no recibieron afstyla para la prevención .	52.824322	43.024393
44					100.000000	100.000000
45	The second study involved 83 patients below 12 years of age. Afstyla was rated as 'excellent' or 'good' at treating 96% of the 347 bleeding episodes recorded during the study; 96% of bleeding episodes resolved with one or two injections of Afstyla. Among those patients who received Afstyla to prevent bleeding, the average number of bleeding episodes per year was	en el segundo estudio participaron 83 pacientes menores de 12 años . la eficacia de afstyla se consideró «excelente» o «buena» para el tratamiento del 96 % de los 347 episodios hemorrágicos registrados durante el	en el segundo estudio participaron 83 pacientes menores de 12 años de edad . afstyla fue calificado como " excelente " o " bueno " en el tratamiento del 96% de los 347 episodios hemorrágicos registrados durante el estudio ; el 96% de	en el segundo estudio participaron un total de 83 pacientes menores de 12 años de edad . afstyla fue calificada como " excelente " o " buena " en el tratamiento de 96% de los 353 episodios hemorrágicos registrados durante el	59.282790	47.945240

	2.30 for patients receiving Afstyla three times a week, and 4.37 for those receiving Afstyla twice a week.	estudio ; el 96 % de los episodios hemorrágicos remitieron con una o dos inyecciones de afstyla . entre los pacientes que recibieron afstyla para la prevención de hemorragias , el número promedio de episodios hemorrágicos al año fue de 2,30 en los pacientes que recibieron afstyla tres veces por semana y de 4,37 en los que recibieron afstyla dos veces por semana .	los episodios hemorrágicos se resolvieron con una o dos inyecciones de afstyla . entre los pacientes que recibieron afstyla para prevenir el sangrado , el número promedio de episodios hemorrágicos por año fue de 2.30 para los pacientes que recibieron afstyla tres veces por semana , y de 4.37 para los que recibieron afstyla dos veces por semana .	estudio ; el 96% de los episodios de hemorragia se solucionaron con una o dos inyecciones de afstyla . entre los pacientes que recibieron afstyla para prevenir la hemorragia , el número medio de episodios hemorrágicos al año fue 2,30 para los pacientes que recibieron afstyla tres veces por semana , 4,37 para los pacientes que recibieron afstyla dos veces a la semana .		
46					100.000000	100.000000
47	What are the risks associated with Afstyla?	¿cuál es el riesgo asociado a afstyla ?	¿cuáles son los riesgos asociados con afstyla ?	¿cuáles son los riesgos asociados a apealea ?	11.044796	6.567275
48					100.000000	100.000000
49	Hypersensitivity (allergic) reactions are common with Afstyla, affecting up to 1 in 10 people. They may include: angioedema (swelling of tissues under the skin), burning and stinging at the injection site, chills, flushing, itchy rash over the whole body, headache, hives, hypotension (low blood pressure), lethargy, nausea (feeling sick), restlessness, tachycardia (rapid heartbeat), tightness of the chest, tingling, vomiting and wheezing. In some cases these reactions can become severe.	las reacciones de hipersensibilidad (alérgicas) son frecuentes con afstyla y afectan a hasta 1 de cada 10 personas . estas reacciones pueden consistir en : angioedema (inflamación de los tejidos debajo de la piel) , ardor y hormigueo en el lugar de la inyección , escalofríos , sofocos , erupción cutánea con picor en todo el cuerpo , dolor de cabeza , sarpullido , hipotensión (presión arterial baja) , letargo , náuseas , inquietud , taquicardia (latido cardíaco rápido) , opresión torácica , hormigueo , vómitos y pitidos al respirar (sibilancias) . en algunos casos , estas reacciones pueden llegar a ser graves .	las reacciones de hipersensibilidad (alérgicas) son comunes con afstyla , afectando hasta 1 de cada 10 personas . pueden incluir : angioedema (hinchazón de los tejidos debajo de la piel) , ardor y escozor en el sitio de la inyección , escalofríos , rubor , sarpullido con picazón en todo el cuerpo , dolor de cabeza , urticaria , hipotensión (presión arterial baja) , letargo , náuseas (sensación de náuseas) , inquietud , taquicardia (latidos cardíacos rápidos) , opresión en el pecho , hormigueo , vómito y sibilancias . en algunos casos , estas reacciones pueden llegar a ser graves .	reacciones alérgicas (hipersensibilidad) son frecuentes con afstyla , afectan hasta 1 de cada 10 personas . glucosa en sangre : angioedema (hinchazón de tejidos debajo de la piel) , quemazón y picor en el lugar de inyección , escalofríos , enrojecimiento de la piel , erupción cutánea con picor en todo el cuerpo , dolor de cabeza , urticaria , hipotensión (presión arterial baja) , letargo , náuseas (ganas de vomitar) , cansancio , taquicardia (palpitaciones) , opresión en el pecho , hormigueo , vómitos y respiración sibilante . en algunos casos , estas reacciones pueden ser graves .	57.186109	51.560921
50					100.000000	100.000000
51	There is also a risk with factor VIII medicines that some patients will develop inhibitors (antibodies) against factor VIII, causing the medicine to stop working and resulting in a loss of bleeding control.	existe también un riesgo asociado a los medicamentos que contienen factor viii de que algunos pacientes desarrollen inhibidores (también existe el riesgo de que algunos pacientes desarrollen inhibidores (anticuerpos) contra el factor viii , lo que provoca que el	también hay un riesgo de factor viii medicamentos que algunos pacientes que desarrollan anticuerpos (inhibidores) contra	45.831479	18.844559

		anticuerpos) contra el factor viii , lo que hace que el medicamento deje de funcionar , con la consiguiente pérdida del control de las hemorragias .	medicamento deje de funcionar y resulte en una pérdida de control de la hemorragia .	el factor viii , causando el medicamento dejar el trabajo y dando como resultado la pérdida de control de la hemorragia .		
52					100.000000	100.000000
53	Afstyla must not be used in patients with known allergy to hamster proteins.	afstyla no debe utilizarse en pacientes con alergia conocida a proteínas de hámster .	afstyla no debe usarse en pacientes con alergia conocida a las proteínas del hámster .	afstyla no debe utilizarse en pacientes con alergia conocida a las proteínas de hámster .	45.788314	80.032032
54					100.000000	100.000000
55	For the full list of all side effects and restrictions, see the package leaflet.	para consultar la lista completa de efectos adversos notificados y las restricciones de afstyla , consultar el prospecto .	para obtener una lista completa de todos los efectos secundarios y restricciones , consulte el prospecto .	la lista completa de efectos adversos y restricciones puede consultarse en el prospecto .	14.810600	32.762648
56					100.000000	100.000000
57	Why is Afstyla approved?	¿por qué se ha aprobado afstyla ?	¿por qué está aprobado afstyla ?	¿por qué se ha aprobado afstyla ?	32.159351	100.000000
58					100.000000	100.000000
59	Afstyla has been shown to be effective in both preventing and treating bleeding episodes. Regarding safety, the reported side effects are in line with those expected for a factor VIII product, although hypersensitivity reactions occurred more frequently with Afstyla. Further safety data should be provided by ongoing studies.	se ha demostrado que afstyla es eficaz tanto para la prevención como para el tratamiento de episodios hemorrágicos . en cuanto a la seguridad , los efectos secundarios comunicados son similares a los afstyla esperados para un producto que contiene factor viii , aunque las reacciones de hipersensibilidad fueron más frecuentes con afstyla . los estudios en curso deberían facilitar datos adicionales sobre la seguridad .	se ha demostrado que afstyla es eficaz tanto en la prevención como en el tratamiento de los episodios hemorrágicos . en cuanto a la seguridad , los efectos secundarios reportados están en línea con los esperados para un producto de factor viii , aunque las reacciones de hipersensibilidad ocurrieron con mayor frecuencia con afstyla . los estudios en curso deben proporcionar más datos de seguridad .	afstyla ha demostrado ser eficaz en la prevención y tratamiento de episodios hemorrágicos . en cuanto a la seguridad , los efectos adversos se encuentran en línea con los esperados para un medicamento de factor viii , aunque las reacciones de hipersensibilidad se produjeron con mayor frecuencia con afstyla . nuevos datos de seguridad debe ser suministrada por ensayos en marcha .	53.815507	35.584607
60					100.000000	100.000000
61	The Agency's Committee for Medicinal Products for Human Use (CHMP) decided that Afstyla's benefits are greater than its risks and recommended that it be approved for use in the EU.	el comité de medicamentos de uso humano (chmp) de la agencia decidió que los beneficios de afstyla son mayores que sus riesgos y recomendó autorizar su uso en la ue .	el comité de medicamentos de uso humano de la agencia (chmp) decidió que los beneficios de afstyla son mayores que sus riesgos y recomendó que se aprobara su uso en la ue .	la agencia en el comité de medicamentos de uso humano (chmp) decidió que los beneficios de afstyla son mayores que sus riesgos y recomendó autorizar su uso en la ue .	71.210431	87.628711
62					100.000000	100.000000
63	Afstyla	afstyla	afstyla	afstyla	100.000000	100.000000

64	EMA/753310/2016	ema / 753310 / 2016	ema / 753310 / 2016	ema / 753310 / 2016	100.000000	100.000000
65					100.000000	100.000000
66	Page 2/3	página 2 / 3	página 2 / 3	página 2 / 3	100.000000	100.000000
67					100.000000	100.000000
68	What measures are being taken to ensure the safe and effective use of Afstyla?	¿qué medidas se han adoptado para garantizar un uso seguro y eficaz de afstyla ?	¿qué medidas se están tomando para garantizar el uso seguro y eficaz de afstyla ?	¿qué medidas se han adoptado para garantizar un uso seguro y eficaz de rubraca ?	53.036246	86.664157
69					100.000000	100.000000
70	Recommendations and precautions to be followed by healthcare professionals and patients for the safe and effective use of Afstyla have been included in the summary of product characteristics and the package leaflet.	las recomendaciones y precauciones que deben seguir los profesionales sanitarios y los pacientes para un uso seguro y eficaz de afstyla se han incluido en la ficha técnica o resumen de las características del producto y el prospecto .	en el resumen de las características del producto y en el prospecto se han incluido recomendaciones y precauciones a seguir por los profesionales sanitarios y los pacientes para un uso seguro y eficaz de afstyla .	las recomendaciones y precauciones que deben seguir los profesionales sanitarios y los pacientes para un uso seguro y eficaz de rubraca se han incluido en la ficha técnica o resumen de las características del producto y el prospecto .	57.712444	93.183848
71					100.000000	100.000000
72	Other information about Afstyla	otras informaciones sobre afstyla	otra información sobre afstyla	otras informaciones sobre afstyla	31.947155	100.000000
73					100.000000	100.000000
74	The European Commission granted a marketing authorisation valid throughout the European Union for Afstyla on 4 January 2017.	el epar completo de afstyla se puede consultar en la página web de la agencia : ema . europa . eu / find medicine / human medicines / european public assessment reports . para mayor información sobre el tratamiento con afstyla , lea el prospecto (también incluido en el epar) o consulte a su médico o farmacéutico .	el 4 de enero de 2017 , la comisión europea concedió a afstyla una autorización de comercialización válida en toda la unión europea .	la comisión europea emitió una autorización de comercialización válida en toda la unión europea para el medicamento afstyla el 4 de enero de 2017 .	0.624264	0.660492
75					100.000000	100.000000
76	The full EPAR for Afstyla can be found on the Agency's website: ema.europa.eu/FindMedicine/HumanMedicines/European public assessment reports . For more information about treatment with Afstyla, read the package leaflet (also part of the EPAR) or contact your doctor or pharmacist.	afstyla	el epar completo de afstyla puede consultarse en el sitio web de la agencia : ema . europa . eu / find medicine / human medicines / european public assessment reports . para más información sobre el tratamiento con afstyla , lea el prospecto (que también forma parte de la epar) o póngase en contacto con su	el texto completo del epar de afstyla puede consultarse en la página web de la agencia española : ema . europa . eu / find medicine / find medicines / human public assessment reports . para mayor información sobre el tratamiento con lokelma , lea el prospecto (también incluido en el epar) o consulte a su	0.574979	0.584486

			médico o farmacéutico .	médico o farmacéutico .		
--	--	--	-------------------------	-------------------------	--	--

Anexo 19. Informe de análisis del corpus PHARMA en memoQ.

Statistics for file(s) [eng-GB-spa-ES]PHARMA.tmx

Counts

Scope Selected documents; number of documents: 1

Type	Segments	Source words	Source chars	Source tags	Percent	Target words	Target chars	Target tags
All	40313	501284	2797024	805	100,00	584595	3181373	803
Repetition	7947	46091	269491	4	9,19	52075	309240	4
X-translated / double context	0	0	0	0	0,00	0	0	0
Not started	0	0	0	0	0,00	0	0	0
Pre-translated	0	0	0	0	0,00	0	0	0
Fragments	0	0	0	0	0,00	0	0	0
Edited	0	0	0	0	0,00	0	0	0
Rejected	0	0	0	0	0,00	0	0	0
Translator confirmed	40313	501284	2797024	805	100,00	584595	3181373	803
Reviewer 1 confirmed	0	0	0	0	0,00	0	0	0
Reviewer 2 confirmed	0	0	0	0	0,00	0	0	0
Locked	0	0	0	0	0,00	0	0	0

Anexo 20. Informe de análisis del corpus ECDC en memoQ.

Statistics for file(s) [eng-GB-spa-ES]en-es(ECDC).tmx

Counts

Scope Selected documents; number of documents: 1

Type	Segments	Source words	Source chars	Source tags	Percent	Target words	Target chars	Target tags
All	2502	36813	212907	5427	100,00	44244	247979	5498
Repetition	0	0	0	0	0,00	0	0	0
X-translated / double context	0	0	0	0	0,00	0	0	0
Not started	0	0	0	0	0,00	0	0	0
Pre-translated	0	0	0	0	0,00	0	0	0
Fragments	0	0	0	0	0,00	0	0	0
Edited	0	0	0	0	0,00	0	0	0
Rejected	0	0	0	0	0,00	0	0	0
Translator confirmed	2502	36813	212907	5427	100,00	44244	247979	5498
Reviewer 1 confirmed	0	0	0	0	0,00	0	0	0
Reviewer 2 confirmed	0	0	0	0	0,00	0	0	0
Locked	0	0	0	0	0,00	0	0	0

Anexo 21. Informe de análisis del corpus de TAUS en memoQ.

Statistics for file(s) [eng-GB-spa-ES]TAUS.tmx

Counts

Scope Selected documents; number of documents: 1

Type	Segments	Source words	Source chars	Source tags	Percent	Target words	Target chars	Target tags
All	79502	1160584	6026196	0	100,00	1309024	6758515	0
Repetition	3182	42840	221293	0	3,69	49226	252275	0
X-translated / double context	0	0	0	0	0,00	0	0	0
Not started	0	0	0	0	0,00	0	0	0
Pre-translated	0	0	0	0	0,00	0	0	0
Fragments	0	0	0	0	0,00	0	0	0
Edited	0	0	0	0	0,00	0	0	0
Rejected	0	0	0	0	0,00	0	0	0
Translator confirmed	79502	1160584	6026196	0	100,00	1309024	6758515	0
Reviewer 1 confirmed	0	0	0	0	0,00	0	0	0
Reviewer 2 confirmed	0	0	0	0	0,00	0	0	0
Locked	0	0	0	0	0,00	0	0	0

Anexo 22. Informe de análisis del corpus EMEA en memoQ.

Statistics for file(s) [eng-GB-spa-ES]CORPUS EMA (EN-ES).tmx

Counts

Scope *Selected documents; number of documents: 1*

Type	Segments	Source words	Source chars	Source tags	Percent	Target words	Target chars	Target tags
All	366769	5302365	28864989	0	100,00	5983422	32393568	0
Repetition	74750	875034	4849669	0	16,50	1041873	5687875	0
X-translated / double context	0	0	0	0	0,00	0	0	0
Not started	0	0	0	0	0,00	0	0	0
Pre-translated	0	0	0	0	0,00	0	0	0
Fragments	0	0	0	0	0,00	0	0	0
Edited	0	0	0	0	0,00	0	0	0
Rejected	0	0	0	0	0,00	0	0	0
Translator confirmed	366769	5302365	28864989	0	100,00	5983422	32393568	0
Reviewer 1 confirmed	0	0	0	0	0,00	0	0	0
Reviewer 2 confirmed	0	0	0	0	0,00	0	0	0
Locked	0	0	0	0	0,00	0	0	0