

Trabajo de Fin de Máster en “Tecnologías de la Información y la Comunicación en la Enseñanza y el Tratamiento de Lenguas”

Análisis de la utilidad de herramientas de traducción automática en traducción especializada

La traducción inglés-español de sentencias del Tribunal Europeo de Derechos Humanos

Núria Ferrer Añó

Tutora: D. ^a Esther Juan Oliva

Facultad de Filología

UNED

Convocatoria de septiembre del curso 2018-2019

Agradecimientos

A mi tutora, Esther, por su paciencia y guía, tanto en la concepción del trabajo como durante su largo proceso de elaboración.

Al Dr. Vidal, por su desinteresada colaboración y participación. Espero que esta aportación pueda ser de utilidad en el continuo proceso de traducción de la jurisprudencia del TEDH que lleva a cabo con sus alumnos.

A mi familia, por su atención en los momentos de duda, aunque no entendieran de lo que hablaba.

Y, por último, a los investigadores que he leído y citado, por sus contribuciones en este apasionante campo de la traducción. Espero no haber malinterpretado sus palabras.

Resumen

Debido a una necesidad real de traducción de la jurisprudencia del Tribunal Europeo de Derechos Humanos del inglés al español, en el presente trabajo se analiza la utilidad de dos sistemas de traducción automática en la traducción especializada jurídica. Se pone a prueba tanto Google Translate integrado en la consola de traducción del software memoQ como Amazon Translate de Amazon Web Services, dos sistemas de traducción por redes neuronales. Las traducciones de un reducido corpus de jurisprudencia se analizan cualitativamente y se comparan cuantitativamente con traducciones humanas de referencia haciendo uso de la métrica de evaluación BLEU.

Palabras clave

traducción automática especializada, traducción jurídica, Google Translate, Amazon Translate, BLEU

Abstract

This paper is built on an actual need to translate the European Court of Human Rights' case law from English to Spanish and analyses the usefulness of two machine translation systems in specialized legal translation. The two systems tested are Google Translate, used within the translator's console of the software memoQ, and Amazon Translate from Amazon Web Services, both neural machine translation systems. Translations of a reduced case law corpus are analysed qualitatively, and compared quantitatively with human reference translations through the evaluation metric BLEU.

Key words

specialised machine translation, legal translation, Google Translate, Amazon Translate, BLEU

Índice

1. Introducción.....	7
1.1. Justificación.....	7
2. Marco teórico.....	9
2.1. La traducción jurídica de derecho internacional inglés-español	9
2.1.1. El lenguaje jurídico.....	10
2.1.2. Los ordenamientos jurídicos.....	11
2.1.3. El inglés jurídico y el español jurídico	13
2.2. El Tribunal Europeo de Derechos Humanos (TEDH).....	14
2.3. Las herramientas TAO y los sistemas de traducción automática	18
2.3.1. Traducción Asistida por Ordenador vs. Traducción Automática	18
2.3.2. La integración de TAO y TA.....	24
2.3.3. Herramientas del estudio	25
3. Metodología.....	27
3.1. Objetivos e hipótesis.....	28
3.2. Diseño de la investigación.....	29
3.3. Características del corpus	31
3.3.1. Otgon v. The Republic of Moldova.....	31
3.3.2. İlbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey	32
3.3.3 A v. Croatia	33
4. Descripción y análisis de los resultados	35
4.1. Google Translate integrado en memoQ.....	35
4.1.1. Usabilidad.....	36
4.1.2. Comentario de errores	37
4.1.3. La traducción con y sin terminología	45
4.1.4. Terminología jurídica	47

4.2. Amazon Translate en la consola de Amazon Web Services	48
4.2.1. Usabilidad.....	49
4.2.2. Comentario de errores	50
4.2.3. La traducción con y sin terminología	57
4.2.4. Terminología jurídica	58
4.3. Comentario cualitativo de los resultados.....	58
4.4. Evaluación con métricas cuantitativas.....	60
4.4.1. Comparación de los resultados	60
4.4.2. Comentario cuantitativo de los resultados.....	63
5. Conclusiones.....	64
5.1. Validación de las hipótesis	64
5.2. Consideraciones éticas.....	65
5.3. Limitaciones	66
5.4. Futuras líneas de investigación.....	66
6. Bibliografía.....	68
7. Apéndices	75
Apéndice 1	75
Apéndice 2.....	76

Lista de figuras

Figura 1. Competencia del TEDH y admisibilidad de demandas.....	15
Figura 2. Representación de un sistema de traducción automática mediante transferencia	20
Figura 3. Relación entre sistemas mediante transferencia y mediante lengua intermedia ..	21
Figura 4: Editor de base de datos terminológica de memoQ.....	35
Figura 5: Panel de traducción de memoQ y panel derecho de resultados de traducción	36
Figura 6: Características de la terminología en Amazon Translate.....	48
Figura 7: Consola de Amazon Translate en Amazon Web Services.....	49
Figura 8: Evaluación BLEU del asunto Otgon c. la República de Moldavia.....	61
Figura 9: Evaluación BLEU del asunto Kemalöglu c. Turquía.....	62
Figura 10: Evaluación BLEU del asunto A c. Croacia.....	62

Lista de tablas

Tabla 1. Errores de redacción en GT: ortografía y puntuación	38
Tabla 2. Errores de redacción en GT: léxico	41
Tabla 3. Errores sobre el traslado del sentido en GT: sin sentido	42
Tabla 4. Errores sobre el traslado del sentido en GT: falso sentido	43
Tabla 5. Errores sobre el traslado del sentido en GT: no mismo sentido	43
Tabla 6. No mismo sentido en GT: predisposición por el género masculino.....	44
Tabla 7. Diferencias entre traducción con y sin terminología en GT.....	45
Tabla 8. Errores provocados por la introducción de la terminología en GT	46
Tabla 9. Errores de redacción en AT: ortografía y puntuación	50
Tabla 10. Errores sobre el traslado del sentido en AT: omisión y adición.....	51
Tabla 11. Errores sobre el traslado del sentido en AT: no mismo sentido	53
Tabla 12. Errores sobre el traslado del sentido en AT: falso sentido	54
Tabla 13. Errores sobre el traslado del sentido en AT: sin sentido	54
Tabla 14. No mismo sentido en AT: predisposición por el género masculino.....	56
Tabla 15. Errores provocados por la introducción de la terminología en AT	57

Lista de abreviaturas y siglas

A continuación, se recogen las siglas que aparecen en el trabajo y su significado. En los casos en que se utilicen las siglas en español, se aporta también en calidad informativa las siglas y la nomenclatura en inglés:

HAMT	Traducción Automática con Ayuda Humana <i>Human-Assisted Machine Translation</i>
MT	Memoria de Traducción <i>Translation Memory (TM)</i>
NMT	Traducción Automática por Redes Neuronales <i>Neural Machine Translation</i>
RBMT	Traducción Automática Basada en Reglas <i>Rule Based Machine Translation</i>
SMT	Traducción Automática Estadística <i>Statistical Machine Translation</i>
TA	Traducción Automática <i>Machine Translation (MT)</i>
TAO	Traducción Asistida por Ordenador <i>Computer-Assisted Translation (CAT)</i>
TEDH	Tribunal Europeo de Derechos Humanos <i>European Court of Human Rights (ECHR)</i>
GT	Google Translate
AT	Amazon Translate
AWS	Amazon Web Services

1. Introducción

1.1. Justificación

En este trabajo se analiza la posibilidad de traducir sentencias del Tribunal Europeo de Derechos Humanos (TEDH) del inglés al español mediante sistemas de traducción automática de fácil uso y acceso abierto. El TEDH es un organismo internacional vinculado al Consejo de Europa, institución constituida en 1949 y formada por el conjunto de estados de Europa, a excepción de Bielorrusia, Kazajistán y la Ciudad del Vaticano. Son un total de 47 países miembros, debido a la concepción de Europa en su mayor amplitud geográfica.

El TEDH, como organismo cuya jurisprudencia afecta a todos los países miembros, incluido España, dispone de una base de datos abierta accesible por Internet de todas las sentencias dictaminadas. Sin embargo, las lenguas de trabajo del Tribunal son el inglés y el francés y la mayoría de sentencias que no implican a España directamente no se encuentran traducidas al español. Dado que el Tribunal no asume este trabajo de traducción al resto de idiomas de los países miembros, esta carga ha recaído durante años en el Ministerio de Justicia, sin haberse alcanzado a traducir toda la jurisprudencia ya publicada ni, mucho menos, el constante flujo que proviene de las salas del TEDH. A raíz de un acuerdo entre el citado Ministerio, el Tribunal Europeo de Derechos Humanos y el Vicerrectorado de Investigación e Internacionalización de la Universidad Nacional de Educación a Distancia (UNED), los alumnos del Máster en Unión Europea han empezado a traducir sentencias, tomando el relevo del Ministerio y considerando no solo las sentencias que involucran a España, sino también a otros países. Estas sentencias son revisadas y posteriormente publicadas por ambos organismos.

Debido al carácter de lenguaje de especialización del lenguaje jurídico, se plantea que algunas de sus características, como la rigidez de estructuras o la repetitividad, permitan llevar a cabo una traducción automática de calidad de los textos que, posteriormente, puede ser poseditada por estos alumnos. Se lograría así una mayor rapidez en el proceso de traducción de la jurisprudencia. Se trata, por tanto, de una necesidad real en un contexto concreto de uso, ya que la inexistencia de traducciones al español de las sentencias provoca un vacío en legislación que afecta directamente a los derechos humanos de los españoles e informa la legislación nacional del país.

Es por eso que las condiciones del trabajo son claras y el enfoque definido, ya que no se pretende un estudio extrapolable a la traducción jurídica en general. Hay que exponer las características propias de las variables en juego en la investigación y guiar el análisis de los resultados desde la perspectiva que estas otorgan, con el objetivo de determinar la utilidad de incluir estas herramientas en el proceso de traducción, no como sustitutas completas del traductor humano, sino como complementos de su tarea, tomando estas la carga repetitiva del trabajo y agilizando la labor del traductor.

2. Marco teórico

Para situar en contexto la investigación realizada sobre la traducción de sentencias del Tribunal Europeo de Derechos Humanos con herramientas de traducción automática, se presentan a continuación las características del corpus estudiado mediante una síntesis del funcionamiento del organismo internacional en cuestión, tanto en términos legales como en relación a su régimen lingüístico; de las características del lenguaje jurídico y de la traducción jurídica de derecho internacional penal público en la combinación inglés-español; y del estado actual de las herramientas TAO en combinación con traducción automática.

2.1. La traducción jurídica de derecho internacional inglés-español

La traducción jurídica, como uno de los campos de traducción especializada más antiguos que se conocen, cuenta con una serie de características, retos y demandas intrínsecos al tipo de textos y a la temática que se trabaja. Su cualidad principal, y la que la distingue de la traducción general y literaria, es la limitación a una serie de preferencias funcionales dentro de la gran cantidad de opciones lingüísticas que se pueden dar en un contexto de comunicación profesional (König, 1998).

La traducción jurídica en instituciones internacionales con legislación vinculante para sus miembros es un elemento inherente de la armonización legal de sus sistemas (Sammur, 2018). Estos organismos dependen de la traducción para que se lleve a cabo la diseminación e implementación legal de las sentencias a los niveles nacionales y regionales (Prieto Ramos, 2014). El lector objetivo de los documentos en este contexto es dual, ya que incluye tanto a profesionales legales de los países afectados por la legislación, como a la población general de dichos países, para quienes la documentación debe ser accesible. Se requiere, por tanto, además de conocimientos lingüísticos y técnicos, conocimientos jurídicos básicos de derecho internacional, de la estructura y características del derecho propio de los países involucrados y del organismo en concreto al que alude el texto (Policastro Ponce, 2015).

Como indica Pietro Ramos (2014), en estos contextos se enfatiza la concordancia interlingüística formal, la armonización terminológica y la consistencia intratextual e intertextual con las convenciones establecidas en la institución. Así, podemos destacar que este tipo de traducción levanta una expectativa de precisión muy elevada, para asegurar la univocidad semántica del texto origen y del texto meta (Prieto Ramos, 2014).

Lograr esta armonización terminológica depende en gran medida de la interiorización de los sistemas jurídicos particulares asociados a cada lengua de trabajo y de los diferentes procedimientos y agentes que dan forma a los términos comunes en cada tradición legal. Esto provoca que la traducción de los términos se deba ajustar al máximo al término del documento original y que se reserve su interpretación para la aplicación de la ley (König, 1998). Si el traductor se aleja de la palabra escrita en el original, corre el riesgo de enmendar el texto y provocar distintas interpretaciones dependiendo de la zona de aplicación. Esta indicación, en un principio tan sencilla, se encuentra con el escollo de que el lenguaje es un instrumento vivo y característico de las distintas comunidades, por lo que normalmente refuerza la cultura y la tradición legal del sistema jurídico de su país de origen.

La combinación entre diferentes ordenamientos jurídicos que presenta una falta de correspondencia biunívoca entre ambas lenguas (inglés y español) y la propia dificultad que el lenguaje jurídico entraña, como consecuencia de su reiteración terminológica y la ambigua terminología de sus textos, dificultan los intentos por hacer del lenguaje jurídico un lenguaje preciso, unívoco y lógicamente riguroso.

(Policastro Ponce, 2015)

2.1.1. El lenguaje jurídico

Como hemos adelantado, el lenguaje jurídico es la lengua de especialidad que caracteriza a los textos jurídicos, objeto de la traducción jurídica. Un lenguaje de especialidad, también llamado sublenguaje o lenguaje especializado, se distingue de la lengua general por el contenido de especialidad, los términos que designan conceptos propios y el tipo de discurso empleado (Giráldez Ceballos-Escalera, 2007). Se trata de un refinamiento de los términos, conceptos y saberes de un área del conocimiento, algo que ya mencionaba Saussure como «lenguas especiales». Así, podemos hablar del lenguaje jurídico como un sublenguaje que presenta unas tendencias sintácticas y estilísticas y un vocabulario propio diferentes de los de la lengua común y que se emplea en unos determinados géneros profesionales propios e inconfundibles (Alcaraz & Hughes, 2002).

Siguiendo a Fernández (1999) podemos destacar una serie de características comunes que comparten los lenguajes especializados:

1. se emplean como instrumento de comunicación entre especialistas de una determinada materia;
2. utilizan la gramática de la lengua común, aunque con ciertas formas y usos específicos;
3. se destinan a contextos formales;

4. se centran en la eficacia comunicativa, la precisión, concisión y claridad ante la estética;
5. se sirven, principalmente, de un canal escrito.

En el caso del lenguaje jurídico, se pueden considerar subdivisiones según las distintas ramas del derecho, por lo que observamos una terminología propia del derecho civil, del derecho penal, etc. Sin embargo, la presencia de un término en un texto jurídico no lo convierte inmediatamente en un término jurídico, ya que se da una coexistencia de términos de uso exclusivamente jurídico y de términos polisémicos usados en la lengua general que se emplean también en el lenguaje jurídico (Giráldez Ceballos-Escalera, 2007). De hecho, el léxico jurídico se apoya en la lengua común y la mayor parte de los términos jurídicos no son vocablos completamente alejados del vocabulario general, sino significados diferentes y exclusivamente jurídicos de términos de la lengua común (Mattila, 2016; Ruellan, 2010). La doble pertenencia de los términos, que se da en menor medida en otros lenguajes de especialidad, crea problemas de comprensión frecuentes (Ruellan, 2010).

Esta base en la lengua común lleva a algunos autores a dudar de la caracterización del lenguaje jurídico como un lenguaje de especialidad. Sin embargo, Cornu (1990) señala la presencia de una terminología exclusivamente jurídica, que sitúa en unos 400 términos en el caso del francés, la lengua que estudia, como rasgo indiscutible de la especificidad del lenguaje jurídico.

Además del vocabulario, el lenguaje jurídico se caracteriza por un estilo propio, marcado principalmente por la opacidad y falta de naturalidad, a causa del excesivo formulismo y la embrollada sintaxis (Alcaraz & Hughes, 2002), a pesar de que en teoría el discurso jurídico se dirige no solo a los especialistas, sino también al ciudadano profano (Harvey, 2002). Estas decisiones estilísticas vienen marcadas por las distintas jurisdicciones y las características del género en cuestión. Por ejemplo, en la exposición de los hechos el estilo suele ser descriptivo con alta presencia del lenguaje común, mientras que los fallos siguen una estructura predeterminada y contienen más términos técnicos.

En apartados posteriores se indican características propias del lenguaje jurídico específicas a la combinación lingüística que nos ocupa.

2.1.2. Los ordenamientos jurídicos

Es importante destacar, como ya se ha adelantado, que en la traducción jurídica de derecho penal internacional en la combinación inglés-español se produce una transposición entre los dos sistemas jurídicos asentados en Europa: el derecho anglosajón (en inglés, *common*

law) y el derecho continental europeo (en inglés, *civil* o *continental law*). Por una parte, el derecho anglosajón emana de la jurisprudencia creada por los jueces y tribunales en Inglaterra desde el siglo XII. La fuente fundamental del derecho son las decisiones judiciales vinculantes anteriores (Orts Llopis, 2015). Es, por tanto, una legislación empírica, que también se conoce como derecho de creación judicial (*judge-made law*) o derecho del caso (*case law*). Por su parte, el derecho continental es un derecho codificado resultado del esfuerzo de los Estados europeos por crear una serie de normas para gobernar todos los aspectos legales de la vida de los ciudadanos. Es un derecho especulativo que legisla con anterioridad a cualquier demanda específica que pueda surgir en el futuro a través de un código de leyes escrito y promulgado (Sureda, s. f.; Zupancic, 2016). En este ordenamiento jurídico, los precedentes judiciales solo tienen un papel secundario con valor persuasivo.

Estos dos ordenamientos jurídicos conllevan estilos de redacción muy distintos, más detallado en el caso anglosajón, ya que no se pueden producir errores de comprensión de las provisiones establecidas, y más vago en el derecho continental, dado que se trata de legislación creada con anterioridad, en términos genéricos, que los tribunales han de interpretar (Jacobs, 2003). Las decisiones de traducción en este contexto conllevan un ejercicio de análisis legal comparativo que identifique los grados de correspondencia entre conceptos que existen en ambas tradiciones (Piris, 2005).

Prieto Ramos (2014) recoge, por tanto, dos indicaciones en relación con el uso de terminología jurídica en instituciones internacionales:

- 1) Concepts borrowed from particular legal traditions (not only common law) must not be taken at face value; their legal scope within the shared system is defined through the relevant authoritative interpretation and is not necessarily the same as in national law.
- 2) Efforts are made to avoid country-specific terms and adopt more neutral or generic ones, particularly as they emerge from comparative law or conceptual hybrids in the negotiation of harmonisation-oriented instruments.

Es decir, se recomienda precaución a la hora de discernir el significado de ciertos términos asociados a una tradición legal concreta, dado que pueden contener matices distintos a aquellos que los caracterizan en el sistema jurídico nacional, y se intenta evitar términos específicos asociados a un determinado país para adoptar así otros más genéricos que a menudo provienen de la realización de un ejercicio de derecho comparado, o que son híbridos conceptuales que han surgido de un esfuerzo de armonización. La dificultad recae, además de en la existencia de términos exclusivamente jurídicos, en el empleo de palabras del lenguaje cotidiano con acepciones típicamente jurídicas (Policastro Ponce, 2015).

Es por esto que ciertos autores se cuestionan si, en contextos como la traducción de legislación en la Unión Europea, donde se presume que los textos traducidos conllevan el mismo valor legal que el original, los traductores llevan a cabo un trabajo mayor que la traducción y se pueden llegar a considerar autores de la legislación (Kjær, 2015).

2.1.3. El inglés jurídico y el español jurídico

Siguiendo a Alcaraz y Hughes (2014), destacamos algunos rasgos distintivos del inglés jurídico: 1) presencia de vocablos de origen francés o normando; 2) presencia de latinismos; 3) registro formal y arcaizante; 4) presencia de adverbios y sintagmas preposicionales arcaicos; 5) redundancias frecuentes en forma de dobletes y tripletes; 6) verbos performativos; y 7) uso simultáneo de eufemismos (*'detention during Her Majesty's pleasure'*) y coloquialismos contemporáneos (*'stalking, money laundering'*). Algunas características morfológicas y sintácticas comunes son: 1) oraciones inusualmente largas; 2) abundancia de conectores restrictivos (*'notwithstanding', 'under', 'subject to', 'having regard to'*); 3) uso frecuente de la voz pasiva; 4) formulaciones hipotéticas y condicionales; y 5) el empleo de sufijos que marcan las relaciones activas y pasivas en términos legales (*-er/-or vs. -ee*).

Por su parte, el español en su uso en el contexto legal se caracteriza por: 1) lenguaje denotativo y unívoco, con inclinación hacia la nominalización; 2) presencia de arcaísmos y latinismos, mostrando un gran gusto por lo altisonante; 3) estructura rígida, preestablecida de antemano; 4) apego a fórmulas estereotipadas y presencia de fraseología; 5) limitada creatividad, expresividad y subjetividad, a la vez que se observa una gran audacia en la creación de nuevos términos, (*«predisponente»* por «la parte contractual que predispone»); 6) uso de tecnicismos y eufemismos; y 7) redundancia expresiva léxica, colocando juntas palabras de significaciones muy aproximadas en forma de dobletes o tripletes (Alcaraz & Hughes, 2002; Policastro Ponce, 2015).

Asimismo, como las lenguas dan forma a distintos enfoques para ver el mundo, es común argumentar que la red conceptual expresada por la lengua predominante de comunicación en la institución ejerce una influencia considerable en el lenguaje legal internacional (Prieto Ramos, 2014). Tanto a nivel mundial como europeo, esta lengua es el inglés. Sin embargo, es un inglés como *lingua franca*, utilizado por autores de antecedentes diferentes y con distintas lenguas maternas, por lo que su uso idiomático suele quedar reducido. Además, suele contener ciertas rarezas o interferencias en el vocabulario y la sintaxis (Frame, 2005).

Por tanto, los traductores de legislación en instituciones internacionales deben conocer a fondo la institución y su funcionamiento para comprender plenamente los textos, así como los componentes del material original, muchas veces escrito en una variedad internacional del inglés. Este desafío doble demanda un alto nivel de especialización profesional en los traductores, quienes se erigen como filtro lingüístico de neologismos legales en inglés, garantía de calidad tanto del texto traducido como del original y desambiguadores de terminología vaga (Prieto Ramos, 2014).

Como último apunte, es necesario destacar que los argumentos por el uso del inglés como *lingua franca* en estas instituciones internacionales, un modelo que eliminaría las complicaciones del multilingüismo y permitiría una gestión jurídica más fluida, han sido refutados tanto en términos económicos, ya que la carga económica pasa a los ciudadanos que deben aprender inglés, como en términos legales y políticos, dado que el multilingüismo garantiza el acceso del público a la legislación y permite invocarla en tribunales nacionales (Gazzola & Grin, 2013; Pacho Aljanati, 2015).

2.2. El Tribunal Europeo de Derechos Humanos (TEDH)

El Tribunal Europeo de Derechos Humanos (TEDH), también conocido como el Tribunal de Estrasburgo por la ciudad donde se sitúa, es el organismo único destinado a «enjuiciar, bajo determinadas circunstancias, las posibles violaciones de los derechos reconocidos en el Convenio Europeo de Protección de los Derechos Humanos y de las Libertades Fundamentales (CEDH) y en sus Protocolos» (Fundación Acción Pro Derechos Humanos, s. f.). Fue fundado a tiempo completo en 1998 por el Consejo de Europa en sustitución de un antiguo sistema mixto y actúa sobre los 47 Estados miembros del Consejo y firmantes del Convenio. Es importante destacar que el TEDH y el Consejo de Europa son independientes de la Unión Europea y sus organismos.

El TEDH se compone de un juez por cada estado miembro elegido por la Asamblea Parlamentaria del Consejo de Europa, quien ejerce por un periodo de nueve años no renovable. El Tribunal se distribuye en cinco secciones, de composición estable por un periodo de tres años, y la Gran Sala, de diecisiete jueces, también constituida para tres años (Zupancic, 2016). Las sentencias se adoptan por mayoría simple (más del 50% de unanimidad) (Ruiloba Alvariño, 2006).

El Tribunal no juzga cualquier presunta vulneración del Convenio, sino que se deben cumplir una serie de requisitos para que una causa sea admisible por el TEDH. Aun así, en 2014 el Tribunal emitió 2 388 sentencias y tomó 83 675 decisiones de admisibilidad, las

cuales suele dictar un juez único (Brannan, 2016). Solo los asuntos más importantes se juzgan en la Gran Sala. Estas condiciones que deben cumplir las demandas que se presentan el TEDH son de dos tipos, materiales y formales, y sirven para determinar la admisibilidad de las demandas y si el TEDH tiene competencia sobre el tema:

Condiciones materiales	
<i>Ratione loci</i>	La violación alegada del Convenio ha tenido lugar en la jurisdicción del Estado demandado o en territorio controlado efectivamente por este Estado.
<i>Ratione materiae</i>	El derecho invocado por el demandante está protegido por el Convenio y sus Protocolos en vigor.
<i>Ratione temporis</i>	Solo se cubre el periodo posterior a la ratificación del Convenio o de sus Protocolos por el Estado demandado.
<i>Ratione personae</i>	Requiere que la violación alegada del Convenio haya sido cometida por un Estado Contratante o que le sea imputable de una u otra manera.
Condiciones formales	
«Al Tribunal no podrá recurrirse sino después de agotar las vías de recurso internas».	
«Al Tribunal no podrá recurrirse [sino] en el plazo de seis meses a partir de la fecha de la resolución interna definitiva».	
«El Tribunal no admitirá ninguna demanda individual presentada (...) cuando sea anónima».	
«El Tribunal no admitirá ninguna demanda individual (...) cuando sea esencialmente la misma que una demanda examinada anteriormente por el Tribunal o ya sometida a otra instancia internacional de investigación o de acuerdo, y no contenga hechos nuevos».	
«El Tribunal declarará inadmisibile cualquier demanda individual presentada (...) si considera que la demanda es (...) abusiva. »	

Figura 1. Competencia del TEDH y admisibilidad de demandas

(Consejo de Europa, 1950; Tribunal Europeo de Derechos Humanos, 2014)

Como hemos adelantado anteriormente, existen dos sistemas jurídicos principales en Europa y, a pesar de contar con un texto escrito con anterioridad al hecho como es el Convenio y sus Protocolos, el TEDH opera claramente en términos de derecho anglosajón. El propio texto del Convenio, por el momento histórico en que se redactó, después de la

Segunda Guerra Mundial, contiene características del sistema legal británico y estadounidense. Asimismo, con su jurisprudencia, el TEDH «ha hecho del Convenio un instrumento vivo, ampliando, de este modo, los derechos garantizados y ha permitido su aplicación a situaciones que resultaban imprevisibles en el momento de la adopción del Convenio» (Unidad de Relaciones Públicas del Tribunal, 2012). Se justifica argumentando que en pocas ocasiones solamente el Convenio y el resto de Protocolos han proporcionado base suficiente como para dictar sentencias. La acumulación de sentencias del Tribunal con el tiempo ha permitido que actualmente los jueces se refieran a la jurisprudencia del propio Tribunal (Zupancic, 2016).

Como apunte, cabe destacar que tanto el TEDH como los Tribunales de la Unión Europea han motivado cambios en países con sistemas de derecho continental europeo, donde, en ocasiones contadas, se ha empezado a aludir a precedentes del hecho (Prieto Ramos, 2014). Otra característica del derecho anglosajón que es aparente en el TEDH es el grado de predictibilidad de las sentencias. La consigna de que las causas similares deben ser juzgadas de forma similar (*stare decisis*) lleva a que el derecho anglosajón, aunque basado en la analogía, sea mucho más predecible que el derecho continental, basado en la lógica y el silogismo (Zupancic, 2016). Según algunos estudios, se ha observado hasta un 79% de precisión en la predicción de sentencias del TEDH mediante el análisis del lenguaje empleado en la sentencia (Aletras, Tsarapatsanis, Preoțiu-Pietro, & Lampos, 2016).

En su forma actual, las sentencias del TEDH se estructuran en los siguientes artículos (White, 2009):

1. Lista de jueces y fecha
2. Procedimiento
3. Antecedentes de hecho
 - a. Circunstancias del caso
 - b. Derecho interno relevante
4. Fundamentos jurídicos
 - a. Supuesta vulneración
 - i. Alegaciones de las partes
 - ii. Alegaciones por parte del Tribunal
5. Aplicación del Artículo 41 del Convenio (si procede)
6. Sentencia, donde se indica si la decisión del Tribunal ha sido unánime o no
7. Opiniones separadas o disconformes

El Artículo 41 recoge la consideración de una satisfacción equitativa si ha habido vulneración del Convenio.

Las sentencias del Tribunal se publican en inglés o francés, las dos lenguas de trabajo del TEDH, dependiendo de si la deliberación se ha llevado a cabo en una lengua o la otra. En ocasiones, las sentencias se publican tanto en inglés como en francés, indicando que se trata en ambos casos de documentos originales. Estas sentencias se pueden consultar abiertamente en la base de datos del TEDH llamada HUDOC¹, donde se clasifican los asuntos en tres niveles de importancia: 1) alta, es decir, las sentencias que contribuyen significativamente al desarrollo, aclaración o modificación de la jurisprudencia del Tribunal o de un Estado miembro; 2) media, o las sentencias sin aportación significativa pero que no se han limitado simplemente a aplicar jurisprudencia; y 3) baja, en el caso de las sentencias de poco interés legal y/o acuerdos amistosos (White, 2009). Las sentencias solo publicadas en una de las dos lenguas de trabajo suelen ser de baja o media importancia.

El apoyo lingüístico propio del TEDH incluye una división de francés compuesta por 6 traductores y 2 revisores y una división inglesa con 4 traductores y 2 revisores (Brannan, 2016). Esta situación en la que ambos textos, tanto en inglés como en francés, se consideran originales ha llevado en ocasiones a situaciones problemáticas en las que se han hallado matices diferentes en las versiones de estos textos dependiendo de si se ha tomado la versión en francés o la versión en inglés como original (Derlén, 2015; König, 1998).

Cualquier traducción o documento publicado en otra lengua que no sea inglés o francés se lleva a cabo de forma externa. Esto suele ocurrir con sentencias de alta importancia, para que puedan alcanzar al público de todos los países miembros. En 2012, se inició un programa de información, formación y divulgación de la jurisprudencia del Tribunal (*the Court's case-law information, training and outreach program*²) destinado a mejorar la accesibilidad y comprensión de la jurisprudencia del TEDH a nivel nacional. El programa incluye una serie de publicaciones realizadas en conjunto con gobiernos nacionales y otros socios externos que incluyen colecciones impresas y digitales de las principales sentencias, la base de datos HUDOC, un programa de traducción de jurisprudencia y otras publicaciones, como resúmenes anuales de las sentencias dictadas, guías de jurisprudencia o perfiles por país.

¹ <http://echr.coe.int/echr/en/hudoc>

² https://www.echr.coe.int/Documents/Case_law_info_training_outreach_2015_ENG.pdf

En nuestro caso, es especialmente interesante el programa de traducción de jurisprudencia, que se llevó a cabo entre 2012 y 2016 a través del Fondo Fiduciario de Derechos Humanos (HRTF, por sus siglas en inglés). Con esta financiación, se tradujeron más de 21 000 textos (sentencias y resúmenes legales) en treinta y un idiomas distintos al inglés y al francés. Se centró, sobre todo, en sentencias de alta importancia. Por otra parte, el Tribunal acepta traducciones o resúmenes llevados a cabo por agencias externas y los publica en la web con fines informativos. Se establecen unos criterios de admisibilidad que deben cumplir las solicitudes enviadas. Estas traducciones están protegidas por derechos de autor y no pueden reproducirse o republicarse electrónicamente, en forma impresa o en cualquier otro formato sin la aprobación previa del titular de los derechos de autor.

Para evitar posibles situaciones de divergencia en las versiones traducidas se recomienda, como ya se ha indicado en el apartado anterior, realizar un ejercicio comparativo de ambas versiones para aclarar dudas de traducción aparentemente ocultas si solo se consulta una versión (Dzehtsiarou, 2010), pegarse al texto original y mantener su lógica, utilizar la terminología de forma consistente, alejarse de conceptos específicos del derecho anglosajón, ajustarse al estilo y lenguaje del TEDH, encontrar conceptos equivalentes funcionales en el área legal correspondiente (si no es posible, utilizar un término genérico o el original) y alejarse de terminología legal y colocaciones típicas del lenguaje jurídico (Brannan, 2016).

2.3. Las herramientas TAO y los sistemas de traducción automática

2.3.1. Traducción Asistida por Ordenador vs. Traducción Automática

Las herramientas conocidas como herramientas de Traducción Asistida por Ordenador (TAO) engloban una serie de programas de software destinados a asistir al traductor en el proceso de traducción, bien sea de forma directa o indirecta, e independientemente del momento en que se utilicen, tanto antes, como durante o después del proceso (Melby, 1998). En inglés, se conocen como *CAT Tools*, las siglas de *Computer-Assisted Translation Tools*.

Entre este tipo de herramientas, las más comunes son (Bermúdez Bausela, 2014; Champollion, 2003):

- a) Memorias de traducción (MT), que utilizan traducciones previamente almacenadas para sugerir equivalentes según el porcentaje de correspondencia con el segmento en el que está trabajando el traductor.
- b) Gestores terminológicos.

- c) Herramientas de alineación de textos, comúnmente asociadas a proporcionar *input* para la memoria de traducción.
- d) Herramientas de ayuda lexicográfica o explicación de corpus, para realizar concordancias, buscar agrupaciones de palabras, etc.
- e) Herramientas de reconocimiento óptico de caracteres (OCR, por sus siglas en inglés).
- f) Herramientas de reconocimiento de voz.

Existen también herramientas específicas del campo de la traducción audiovisual, como herramientas de subtitulación o localización, que no nos conciernen en el contexto de este trabajo. Algunos autores consideran las herramientas TAO como un concepto más amplio, que incluye programas de revisión ortográfica o de procesamiento de textos (Zampieri & Vela, 2014), pero por su importancia y especialización nos centramos en aquellas herramientas mencionadas anteriormente.

También es importante destacar en este punto la diferencia entre la traducción automática (TA) y la traducción asistida por ordenador (TAO), términos estrechamente relacionados que definen dos disciplinas distintas con orígenes distintos (Sánchez Castany & Díaz Pore, 2011).

La traducción automática (TA o MT por las siglas en inglés de *Machine Translation*) es un área de la lingüística computacional que investiga el uso de software para traducir entre lenguas. Aúna distintas áreas de investigación, como la inteligencia artificial, la traductología, el procesamiento del lenguaje natural, la informática y la lingüística contrastiva. La completa mecanización de la traducción ha sido un sueño histórico del ser humano que empezó a materializarse a mitad del siglo XX con aspiraciones muy optimistas, aunque estas fueron reduciéndose después de algunos varapalos en los años 60 (Hutchins, 2007). Como vemos, la TA se centra en la creación de un sistema que pueda traducir un texto sin intervención humana, mientras que las herramientas TAO son una ayuda computarizada al traductor humano para mejorar su productividad y facilitar su tarea.

En la clasificación de sistemas de TA se habla sobre todo de las distintas arquitecturas que los conforman. Históricamente, los primeros sistemas basados en reglas de TA se regían por una arquitectura directa, mientras que la arquitectura mediante transferencia y mediante lengua intermedia han dominado el panorama en las últimas décadas (Hutchins, 2007). Las dos últimas, conocidas como arquitecturas de conocimiento lingüístico, producen sistemas con amplios conocimientos de ambas lenguas y su relación.

Los sistemas de arquitectura mediante transferencia cuentan con tres tareas principales: análisis, para crear una representación abstracta de las características lingüísticas dejando a un lado la forma del texto origen; transferencia, para buscar la representación de la lengua origen en la lengua meta; y síntesis, para transformar la representación abstracta en oraciones de la lengua meta (Kit & Wong, 2008). Así, el análisis sintáctico de la lengua de partida produce una representación en árbol que es transferida a la lengua de llegada. El texto final se genera a partir de esta representación (*Figura 2*).

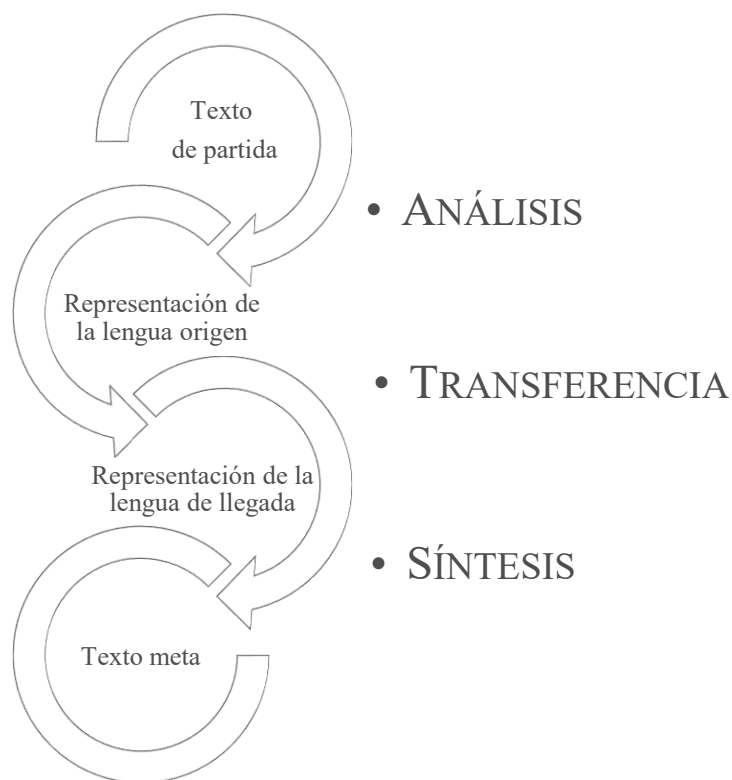


Figura 2. Representación de un sistema de traducción automática mediante transferencia

La diferencia entre las dos arquitecturas de conocimiento lingüístico recae en el nivel de abstracción de dichas representaciones (*Figura 3*). En un sistema mediante transferencia, la representación está ligada a las dos lenguas involucradas. En un sistema mediante lengua intermedia, esta representación sería más profunda, hasta llegar a ser independiente de las dos lenguas, utilizando para ello una interlengua. Sin embargo, la representación abstracta de la semántica y la sintaxis neutraliza las similitudes entre lenguas de las que se podría aprovechar el sistema en un contexto de transferencia con gramática contrastiva (Hutchins, 2007).

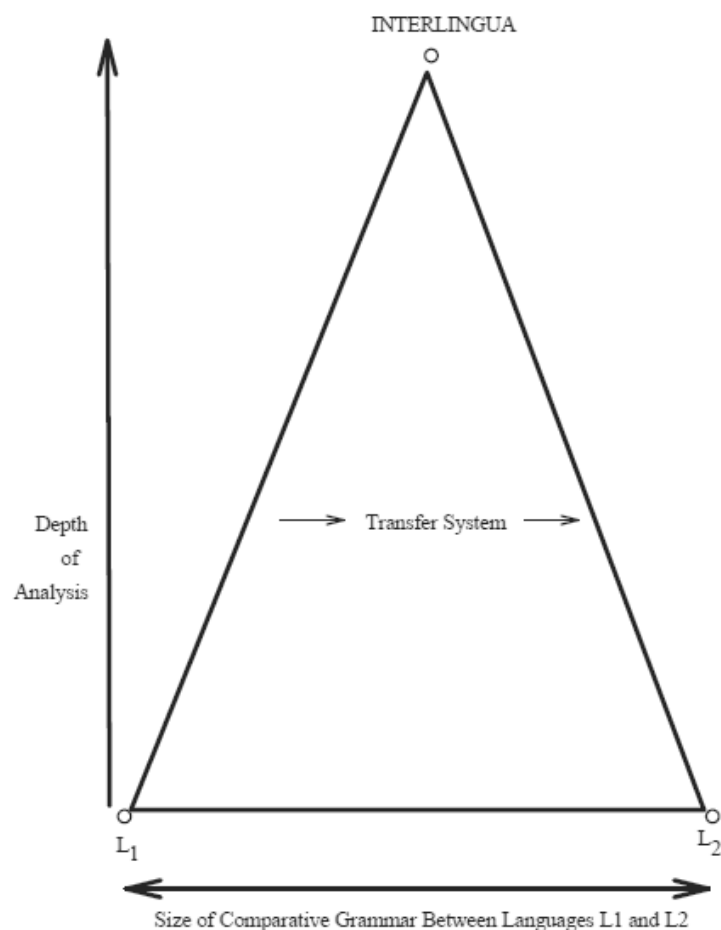


Figura 3. Relación entre sistemas mediante transferencia y mediante lengua intermedia
(extraído de Arnold, Balkan, Meijer, Humphreys, & Sadler, 1994)

El consenso actual es que los sistemas mediante transferencia son la mejor base para avanzar en el desarrollo de la TA, aunque siguen encontrando dificultades en captar el conocimiento lingüístico humano mediante reglas precisas (Nolla & Abril, 2017). Es por ello que han surgido nuevos enfoques que se pueden aplicar dentro de distintas estructuras, como el análisis estadístico de modelos probabilísticos en un corpus de entrenamiento, una de las líneas de investigación principales en la actualidad (Kit & Wong, 2008). Estos modelos asignan una probabilidad alta a las frases destino del corpus que son traducciones reales de la frase dada y una probabilidad baja a las que no lo son (Nolla & Abril, 2017). Sin embargo, una limitación de la traducción automática estadística es que solo toma unas pocas palabras anteriores y posteriores como contexto, lo que lleva a una gran variabilidad de calidad según la longitud del texto.

A partir de 2016, una nueva tecnología de aprendizaje automático llamada *Deep Learning* o *Deep Neural Networks*, que trata de imitar cómo funciona el cerebro humano, se

convirtió en una opción viable de aplicación en la traducción automática gracias a avances en la potencia de los ordenadores y procesadores, ya que es necesario el uso de procesadores gráficos que permitan realizar operaciones matriciales a gran velocidad (Bentivogli, Bisazza, Cettolo, & Federico, 2016; Nolla & Abril, 2017).

Actualmente, pues, conviven los sistemas de traducción automática basados en reglas (RBMT por sus siglas en inglés: *Rule Based Machine Translation*), los que utilizan información estadística (SMT por sus siglas en inglés: *Statistical Machine Translation*) y este último avance, basado en redes neuronales (NMT por sus siglas en inglés: *Neural Machine Translation*) (Parra Escartín, 2018). Estos dos últimos necesitan datos para ser entrenados, los cuales se proporcionan en forma de corpus paralelos. La diferencia entre el método estadístico y el de redes neuronales es el método de computación: las redes neuronales artificiales, de forma similar a nuestras neuronas, asocian la información recibida con otra información subyacente lingüística y no lingüística para formar conexiones y generar traducciones. En pruebas a partir de 2016, los sistemas basados en redes neuronales han superado a los sistemas estadísticos, en forma de menores errores morfológicos, léxicos y de orden de palabras, así como un menor esfuerzo en posesición (Bentivogli et al., 2016; Parra Escartín, 2018). Sin embargo, estos resultados se obtienen tras el entrenamiento del sistema con cantidades ingentes de información, por lo que el elevado coste que supone este entrenamiento ha llevado a que solo grandes corporaciones como Google, Microsoft o Amazon sean capaces de implantarlos. Asimismo, la curva de aprendizaje para los traductores y/o usuarios del sistema es mayor que en sistemas estadísticos (Koehn & Knowles, 2017).

La completa automatización con TA sigue estando lejos de la realidad actual. En la práctica, los textos producidos han de ser revisados (poseditados) por traductores humanos o restringidos en el momento de redacción (sublenguajes) para producir un resultado adecuado. No obstante, no hay que perder de vista la utilidad de textos traducidos con TA que no sean adecuados como traducción final, pero que puedan servir para dar una visión del contenido y ser un factor en la decisión final de invertir o no en una mejor traducción. Asimismo, se puede considerar el texto obtenido por la TA como primer borrador para traducciones humanas oficiales (Gotti, Farzindar, Lapalme, & Macklovitch, 2008). Este borrador ha de ser posteriormente editado por traductores profesionales. Aun así, varios estudios han demostrado que se consigue un aumento de la productividad del traductor en oposición a una traducción realizada desde cero por humanos (Parra Escartín & Arcedillo, 2015).

La TA también se ve afectada por la calidad del texto origen, ya que la presencia de argot, palabras mal escritas, mala puntuación y oraciones complejas o largas pueden hacer que el texto sea traducido incorrectamente (Kit & Wong, 2008).

El mayor volumen de traducción humana actual se deriva de textos científicos, técnicos o comerciales, donde abundan las repeticiones y se exige una gran precisión terminológica y consistencia a lo largo del texto o textos que el traductor produce. Es en este trabajo en el que se centra la investigación en TA y en el que puede brillar la existencia de un software fiable, coherente y con capacidad de almacenar traducciones anteriores y reutilizarlas. Las herramientas TAO también son especialmente útiles para traducir materiales repetitivos, ya que el grado de correspondencia con la memoria de traducción es elevado.

Ya existen estudios que han investigado este potencial de la TA para la traducción jurídica, aprovechando su carácter repetitivo y formulaico. Sin embargo, la traducción jurídica presenta muchos de los problemas mencionados anteriormente. Por ejemplo, la presencia de oraciones largas y complejas con un orden de palabras inusual o el uso de terminología dependiente del sistema legal y cultural (Kit & Wong, 2008). Sin embargo, la TA puede ser una buena solución en ciertas situaciones, incluso en este contexto. Al considerar la usabilidad de la TA, debemos tener en cuenta, como hemos mencionado anteriormente, que es relativa a las expectativas de los usuarios. Según Hutchins (2001), las demandas se pueden clasificar en cuatro tipos de acuerdo con su propósito y la calidad de la traducción requerida para servir tales propósitos, desde la necesidad de textos con una calidad suficiente para ser publicados con propósitos de difusión, hasta traducciones de calidad inferior para garantizar el acceso a la información. Asimismo, la producción con TA de un texto de calidad suficiente para ser eficientemente poseído por un traductor humano es una práctica constructiva y común en muchas circunstancias, incluso dentro de la Unión Europea. Se conoce como traducción automática con ayuda humana (HAMT, por las siglas en inglés de *Human-Assisted Machine Translation*) (Kit & Wong, 2008).

Entre los estudios que se han enfocado a evaluar la usabilidad de sistemas de TA en la traducción jurídica destacamos a Kit y Wong (2008), quienes se centraron en averiguar qué sistema de TA en línea podría ser adecuado para la traducción de textos jurídicos desde distintas lenguas al inglés. Seleccionaron seis sistemas conocidos: Babel Fish, Google, ProMT, SDL, Systran, y WorldLingo, y observaron que los sistemas que utilizan la tecnología Systran (Systran, Babel Fish y Google) habían obtenido resultados comparativamente mejores que los otros, aun sin detectar ningún sistema que fuera sustancialmente mejor en todos los pares de idiomas. Gotti, Farzindar, Lapalme, &

Macklovitch (2008), sin embargo, se encargaron de entrenar un sistema de TA con la creación de un corpus de referencia a partir de las traducciones anteriores de los tribunales canadienses para experimentar sobre la traducción automática de francés a inglés y de inglés a francés de sentencias de los tribunales. A pesar de obtener resultados de calidad no publicable, las traducciones generadas se podían utilizar para la posesición humana con resultados alentadores en términos de productividad.

Estos estudios de evaluación sobre los resultados obtenidos con traducción automática se han llevado a cabo tanto con herramientas de evaluación manual como de evaluación automática, como la conocida métrica BLEU (Papineni, Roukos, Ward, & Zhu, 2002). BLEU mide el número de coincidencias entre una traducción de referencia y la traducción generada con TA y penaliza diferencias importantes en la longitud de las oraciones. Por tanto, una buena TA deberá tener una alta puntuación BLEU, ya que indica su mayor cercanía a una traducción humana profesional (Gotti et al., 2008). Otras propuestas que se han centrado en la evaluación humana, tanto con pruebas similares a pruebas de comprensión escrita, como con evaluaciones humanas subjetivas, se han considerado sensibles al sesgo personal y a alteraciones según las capacidades del evaluador (Kit & Wong, 2008).

2.3.2. La integración de TAO y TA

Una de las tareas más arduas a la hora de utilizar las herramientas TAO es la producción de la memoria de traducción, que está inicialmente en blanco, y es solo realmente útil y precisa cuando contiene una gran cantidad de segmentos traducidos (Zampieri & Vela, 2014). Con el tiempo, la MT se rellena con las traducciones producidas por el traductor en el programa de traducción. Sin embargo, se ha investigado recientemente sobre la integración de sistemas de TA en software de traducción para producir la memoria de traducción con segmentos traducidos generados por el sistema informático (Federico, Cattelan, & Trombetti, 2012; Guerberof, 2009; Zampieri & Vela, 2014). En el primer caso, se utilizaron correspondencias de Google Translate en un entorno de traducción con el software Trados Studio y se midieron indicadores de productividad como la velocidad en posesición y el esfuerzo que supuso para los traductores humanos, contabilizado el número de cambios que se aplicaron a las sugerencias de la MT. Los resultados fueron positivos en ambos parámetros, con una mejora del 27% en la velocidad y una reducción del 54,6% en el esfuerzo.

En Zampieri y Vela (2014), la comparación de tres situaciones, una sin uso de MT, otra con MT de segmentos producidos del corpus KDE sin ser modificados, y otra con MT de

segmentos del corpus KDE y modificados por un humano condujo a mejoras en la productividad en los dos últimos casos, sobre todo con segmentos no modificados. Aun así, cabe destacar que la MT de corpus no editado tenía una cobertura un 4,93% mayor que en la situación 3.

Guerberof (2009) comparó productividad y calidad de correspondencias imperfectas de TA y de segmentos de MT, con resultados que indican más rapidez en la posesición de segmentos generados por traducción automática y un 80-90% menos de errores en el uso de correspondencias con un corpus de traducción automática (*fuzzy matches*) que con uno de memoria de traducción de segmentos traducidos por humanos.

La medición de calidad del output en estos tres experimentos se realizó con la métrica BLEU (Papineni et al., 2002). Sin embargo, la mayoría de estos experimentos se llevaron a cabo mediante el uso de Trados Studio SDL y, como sugieren Zampiere y Vela (2014), la alta complejidad de manejo del sistema puede provocar pérdidas en la evaluación de la velocidad y productividad.

2.3.3. Herramientas del estudio

A continuación, se describen las herramientas utilizadas en el desarrollo de este análisis, las cuales han sido seleccionadas según criterios de accesibilidad, sencillez de uso y funcionalidad.

memoQ es una herramienta TAO de completas funcionalidades, ya que cuenta con listas de auto extracción, extracción de terminología y creación de memorias de traducción y bases de datos terminológicas. Se permite trabajar con todo tipo de formatos de documentos y, una vez finalizada la traducción, se exporta el texto meta en el formato del documento original. Algunas ventajas de su uso incluyen la aceleración en la edición del texto, la reducción del tiempo de documentación, la mejora en la consistencia de las traducciones y el aumento en la productividad el usuario, ya que evita errores y permite trabajar en varios documentos a la vez (Sánchez Castany & Díaz Pore, 2011). Asimismo, uno de sus aspectos más interesantes es la integración de sistemas de TA en línea a través *plug-ins* que se pueden activar en el software. Esto permite llevar a cabo una pretraducción del texto mediante TA para ser poseditada por un traductor humano aprovechando las herramientas que incluye.

De estos *plug-ins* integrables en memoQ se ha seleccionado **Google Translate**, por ser este el sistema más conocido y extendido de traducción automática, por su fácil acceso y configuración y por el bajo gasto que supone, con un coste de 20 dólares por millón de

caracteres traducidos. Además, desde 2016, Google Translate utiliza una arquitectura por redes neuronales.

La otra herramienta utilizada en este estudio es **Amazon Translate**, un servicio de traducción automática neuronal del gigante comercial estadounidense. El pago en Amazon Translate también se determina por el uso de la API de forma mensual, aunque permite una capa gratuita de dos millones de caracteres al mes durante los primeros 12 meses de uso. Con el servicio en Amazon Web Services se obtiene acceso a distintos servicios de informática en la nube incluyendo la traducción automática.

Se ofrece la posibilidad de traducir en la consola online del traductor que, al igual que la de Google Translate, se limita a 5 000 caracteres. Asimismo, tras una instalación, se permite traducir a través de una interfaz de línea de comando (CLI). La traducción en el entorno CLI está asimismo limitada, siendo la descarga del entorno JDK la única opción para traducir documentos largos sin haber de ser segmentados. También se permite el uso de una SDK para integrar el sistema en aplicaciones propias. En la actualidad, Amazon Translate soporta 25 idiomas y 595 combinaciones de traducción.

Amazon Web Services se fundó en 2006 y el sistema de traducción Amazon Translate fue anunciado a finales de 2017 y entró en uso a principios de 2018, por lo que es un sistema de creación reciente. Se ha seleccionado como herramienta para el estudio también por la facilidad de uso y registro y por ser un sistema de traducción automática por redes neuronales con una corta trayectoria, pero que ya se ha instaurado en aplicaciones conocidas como hoteles.com y que ha sido bien reseñado.

3. Metodología

Entendemos por metodología de investigación el conjunto de procedimientos y técnicas que se aplican de manera ordenada y sistemática en la realización del estudio que conduce al objetivo propuesto. Estas decisiones están guiadas por las hipótesis de partida que el estudio pretende probar y el marco en que se sitúa la investigación.

Debido a su relativa novedad e interdisciplinariedad, las propuestas teóricas en traductología se encuentran aún en proceso de consolidación (Albí & Belén, 2003). Como indica Orozco Jutorán (2001), la investigación científica en traducción no está exenta de problemas, ya que existe una falta de tradición y formación en metodologías de la investigación y una falta de instrumentos de medida específicos para la traductología. Históricamente, los estudios de traducción se han enmarcado en disciplinas con mayor tradición, como la lingüística o la literatura comparada, o se ha intentado crear un nuevo marco que los englobe junto a otros estudios culturales. Sin embargo, Pym (1999) aboga por recalcar el carácter interdisciplinar de la traducción, que ha sido influida por y ha influido las disciplinas relacionadas.

Asimismo, la delicada balanza en la priorización de validez interna (experimento controlado y, por tanto, riguroso y objetivo) o validez externa (capacidad de generalización del experimento) y la existencia de gran cantidad de variables externas al estudio que son difíciles de controlar, como el uso de diferentes estrategias de traducción según el perfil del traductor o la experiencia previa que tenga en un tema concreto, crean un campo de estudio de difícil limitación y que bebe de distintas disciplinas a la hora de establecer metodologías propias.

En la actualidad, existen varias clasificaciones posibles respecto a los tipos de investigación existentes, ya que se pueden llevar a cabo infinidad de diseños experimentales de la investigación. Una de las clasificaciones tradicionales distingue entre dos extremos, la investigación cualitativa y la investigación cuantitativa. Los métodos cuantitativos son deductivos y parten de afirmaciones teóricas que después se contrastan empíricamente. La teoría que se pretende probar se sustenta en evidencias de hechos y en la búsqueda de generalizaciones, es decir, la posibilidad de aplicar los resultados obtenidos con una muestra representativa de una población determinada a toda la población. Para construir los instrumentos de medición en la investigación cuantitativa se buscan métodos rigurosos y se aplican análisis estadísticos con exigencias de cuantificación, verificación y control.

Sin embargo, existen diferentes diseños y tipos de investigación dentro de cada método, que oscilan desde extremos muy estructurados y con un alto grado de control hasta operaciones de naturaleza observacional y/o introspectiva dentro de métodos cualitativos (Orozco Jutorán, 2001). Como indica esta misma autora, además, «algunos de los nombres que se dan a los métodos también pueden utilizarse para referirse a las técnicas de recogida de datos».

En traducción es necesario una combinación de las facetas de investigación teórica y de investigación aplicada, ya que el objetivo es el estudio y la aplicación de una serie de soluciones respecto a los problemas interlingüísticos que plantea la actividad traductora (Albí & Belén, 2003).

3.1. Objetivos e hipótesis

La realización de esta investigación tiene como objetivo el análisis de la utilidad de herramientas de traducción automática en la traducción especializada, más concretamente, en la traducción jurídica. El ámbito en el que se encuadra la investigación parte de una necesidad real de los alumnos del Máster Universitario en Unión Europea de la UNED. A través de un acuerdo entre el Vicerrectorado de Investigación e Internacionalización de la UNED, el Ministerio de Justicia y el TEDH se ha establecido un marco de cooperación para que los estudiantes de la UNED traduzcan al español la jurisprudencia más relevante del Tribunal Europeo de Derechos Humanos, tanto concerniente al Reino de España como a otros estados, así como otras publicaciones escogidas del TEDH, con vistas a conseguir la más amplia difusión posible.

Este convenio de colaboración ha llevado a algunos alumnos del Máster Universitario en Unión Europea a elaborar Trabajos de Fin de Máster en los que se analiza la jurisprudencia del TEDH y se realizan traducciones de sentencias relevantes. Debido a las características únicas de la traducción jurídica como traducción de especialidad que se han recogido en el apartado 2.1., la introducción de la traducción automática en un marco de software TAO se justifica por el carácter repetitivo y formulaico de los textos, la necesidad de un aumento de la productividad en el proceso de traducción y la posibilidad de posesición por parte de traductores versados en el lenguaje jurídico y en la jurisprudencia del Tribunal.

Con la investigación aquí diseñada se pretende establecer en qué medida la traducción automática con posesición humana mejora el proceso que llevan a cabo estos alumnos en términos de calidad y tiempo invertido. Es importante destacar que, debido al carácter

práctico de la investigación, las herramientas que se plantean utilizar son sencillas y de uso libre o con coste reducido, ya que se dirigen a estudiantes de Máster de la UNED.

Asimismo, se plantea la comparación de los dos sistemas de traducción automática utilizados, Google Translate en memoQ y Amazon Translate, para determinar si existen ventajas cuantificables de uno sobre el otro. Debido a la relativa novedad que supone Amazon Translate respecto a la mayor trayectoria de Google Translate, se presuponen mejores resultados de este último. Amazon Translate fue anunciado a finales de 2017 y entró en uso a principios de 2018, mientras que el sistema de Google Translate se lanzó en 2006 como sistema basado en reglas, en 2007 como traducción automática estadística y en 2016 incorporó un sistema de traducción por redes neuronales.

Por tanto, las hipótesis planteadas se resumen como:

- H1: la traducción automática es útil para la traducción jurídica de sentencias del TEDH de inglés a español.
- H2: Google Translate produce traducciones de mejor calidad que Amazon Translate para la traducción de sentencias del TEDH de inglés a español.

3.2. Diseño de la investigación

La investigación se desarrolla con la traducción de los textos del corpus en las dos herramientas seleccionadas. Como se ha expuesto anteriormente, estas han sido elegidas por su sencillez de uso, accesibilidad y bajo coste, ya que se pretende simular una situación de necesidad real de la traducción de sentencias.

En el diseño de la investigación se ha planteado el uso de una base de datos terminológica en los sistemas de traducción para asegurar la correcta traducción de términos ya establecidos en la lengua meta, a saber, el Tribunal Europeo de Derechos Humanos o las siglas TEDH, por ejemplo. Sin embargo, también se cuestionó si la introducción de esta injerencia humana podría enmascarar posibles ventajas de un sistema sobre otro, dado que no se podría determinar si cierto sistema otorgaba la traducción social e institucionalmente acuñada por sugerencia de la base de datos o por propio conocimiento del sistema. Es por ello por lo que se ha decidido realizar cuatro experimentos distintos con cada texto, uno en cada sistema con y sin la base de datos terminológica. Los términos incluidos en esta base de datos se pueden consultar en el apéndice 1. Se han seleccionado únicamente traducciones acuñadas de terminología propia del tribunal.

Las pruebas con Google Translate se han llevado a cabo en el contexto del software memoQ, en la versión de prueba del paquete profesional. Existe una versión gratuita del

software, pero esta no permite la integración de traducción automática ni de terminología ni la importación de más de un texto por proyecto. Se han creado dos proyectos distintos, uno para las traducciones sin base de datos terminológica y otro con ella.

En el caso de Amazon Translate, se ha traducido en la consola online debido a la dificultad de configuración del entorno JDK, que se ha determinado por encima de las posibilidades del público objetivo del estudio. Sin embargo, esto ha imposibilitado la traducción entera del documento en bloque, siendo la limitación de la consola 5 000 caracteres.

Tras la obtención de las traducciones, se procede a su evaluación cualitativa y cuantitativa. La evaluación cuantitativa se lleva a cabo a través de la métrica BLEU, mientras que el análisis cualitativo tiene en cuenta elementos como la usabilidad de sistema, los errores del sistema, el uso de la base de datos terminológica y el tratamiento de la terminología jurídica.

BLEU (*bilingual evaluation understudy*) es una métrica de evaluación de traducciones automáticas desarrollada por Papineni *et al.* (2002) que se calcula por segmentos con la media geométrica de la precisión en n-gramas entre la traducción del sistema de traducción automática y una traducción de referencia hecha por un humano. Esta precisión, sin embargo, está modificada en la fórmula para evitar deficiencias como, por ejemplo, una penalización por brevedad o un límite de contabilización del mismo término.

Los errores son solo una parte de la tarea de evaluación de la calidad de la traducción. Según la definición de Hurtado Albir (2001, pp. 636, 637), los errores de traducción son «una equivalencia inadecuada para la tarea traductora encomendada» y «se determinan según criterios textuales, contextuales y funcionales». Estos errores pueden afectar al sentido del texto origen y a la redacción en la reformulación a la lengua de llegada, una clasificación relacionada con las dos fases del proceso de traducción, la comprensión y la reexpresión. Asimismo, propone subcategorías de esta distinción, a pesar de que reconoce la falta de fronteras claras entre estas. En los errores que se refieren al texto origen se encuentran casos de adición, omisión, contrasentido, falso sentido, sin sentido, no mismo sentido, referencia extralingüística mal solucionada e inadecuación de variación lingüística. En la reformulación a la lengua de llegada se incluye ortografía y puntuación, gramática, léxico, aspectos textuales y redacción.

3.3. Características del corpus

El corpus de este estudio está delimitado por las características de la investigación y consiste en sentencias del Tribunal Europeo de Derechos Humanos redactadas originalmente en inglés.

Se han seleccionado para ello Trabajos de Fin de Máster de alumnos de ediciones anteriores del Máster de la Unión Europea, para poder establecer una comparativa de las traducciones obtenidas con el estudio y aquellas realizadas por los alumnos.

Los asuntos seleccionados son: *Otgon v. The Republic of Moldova* (2017), de 3 068 palabras; *İlbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey* (2012), de 5 201 palabras; y *A. v. Croatia*, de 11 555 palabras; para crear un corpus total de 19 824 palabras.

3.3.1. Otgon v. The Republic of Moldova³

El asunto Otgon c. la República de Moldavia, firme en fecha 26 de enero de 2017, plantea la interpretación de la noción de vida privada y su relación con el derecho a un medioambiente sano, medida no recogida explícitamente en el Convenio Europeo de Derechos Humanos. A través de la demanda de Svetlana Otgon por una indemnización mayor a la recibida por los tribunales moldavos tras un ingreso hospitalario a causa de la ingesta de agua potable contaminada, el TEDH proporcionó orientación sobre el significado y alcance del concepto de vida privada a efectos del artículo 8 del Convenio.

Por un lado, indicó que la noción de «vida privada» es vasta y difícil de limitar y, entre otros, incluye la integridad física y moral de una persona. Por otro lado, consideró que el primer párrafo del artículo 8 y casos anteriores (como *López Ostra contra España* de 1994 y *Hatton contra el Reino Unido* de 2003) permiten considerar que el bienestar de una persona puede verse afectado de forma negativa por las condiciones medioambientales inseguras o perjudiciales. El artículo 8 puede aplicarse en casos medioambientales si la contaminación es causada directamente por el Estado o si la responsabilidad del Estado surge de la negligencia (como en este caso) o de la incapacidad de regular las actividades del sector privado.

El TEDH condenó al Estado moldavo por incumplimiento del artículo 8 del Convenio Europeo de Derechos Humanos al no ofrecer una compensación justa por la reparación de daños causados a la salud de la señora Otgon y su hija. Teniendo en cuenta el periodo relativamente corto de hospitalización y la ausencia de pruebas sobre secuelas en la salud de la demandante, le concedió una indemnización de 4 000 euros. Era un importe menor al

³ El resumen de la sentencia está asistido por el comentario de Ostrowska (2018).

que reclamaba la demandante, pero mayor al que le habían ofrecido los tribunales nacionales.

Según el manual del TEDH sobre el artículo 8 del Convenio Europeo de Derechos Humanos (2018), el respeto a la vida privada también debe comprender las relaciones con otras personas e incluir actividades profesionales o comerciales. A pesar de que el Convenio Europeo de Derechos Humanos no cubre el derecho a un medio ambiente sano, la jurisprudencia del TEDH ha demostrado que el Convenio tiene un carácter abierto y en evolución: los jueces realizan una interpretación de sus disposiciones, para adaptarlas a los nuevos problemas y desafíos.

No obstante, hay quienes critican esta tendencia, evocando las dificultades interpretativas que suscita la inclusión del derecho a un medio ambiente sano en el Convenio. De hecho, uno de los jueces discrepó con la invocación del artículo 8 en este caso, considerando que se trataba más bien de una demanda civil por daños y perjuicios, ya que la demandante no había mencionado la injerencia en su vida privada en sus demandas ante los tribunales nacionales. En la opinión de este juez, haber bebido agua contaminada no tuvo repercusiones en la vida privada de la demandante porque su salud general no había empeorado notoriamente y esta situación se reducía a un único incidente.

La calidad de vida ambiental es un bien jurídico abstracto y difícil de delimitar, dado que la comunidad internacional aún no ha identificado el umbral para la degradación medioambiental que pueda ser considerada una violación del derecho humano a un medio ambiente sano.

3.3.2 İlbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey⁴

El asunto İlbeyi Kemaloğlu y Meriye Kemaloğlu c. Turquía, firme en fecha 10 de julio de 2012, interpreta y aplica dos artículos del Convenio, el artículo 2.1 sobre el derecho a la vida y el artículo 6.1 sobre el derecho a un proceso justo. El artículo 2 vincula a los Estados tanto de forma negativa (prohíbe atentar contra la vida) como de forma positiva (obliga a establecer sistemas efectivos de protección jurídica de la vida). En referencia al artículo 6, se apela al derecho de acceso a un Tribunal, no previsto originariamente en el Convenio, pero desarrollado a través de la jurisprudencia del TEDH.

En este caso, tras la muerte del niño turco de 7 años Atalay Kemaloğlu, quien regresaba a casa solo durante una tormenta de nieve por el cambio en el horario de cierre de la escuela, los padres buscan responsabilizar a las autoridades por su conducta negligente y su

⁴ El resumen de la sentencia está asistido por el comentario de Martín Caño (2018).

incapacidad de proteger el derecho a la vida de su hijo, además de no haber esclarecido los hechos de forma eficaz, no haber identificado responsabilidades, ni haber establecido una reparación adecuada. Además, reclaman que no se les proporcionó una respuesta legal ni se les concedió el beneficio de la asistencia jurídica gratuita para el pago de las tasas judiciales a pesar de su situación de pobreza, lo cual vulneró su derecho a un proceso justo. El TEDH consideró que la justicia turca no protegió el derecho a la vida del hijo de los demandantes, que no proporcionó los medios adecuados para responsabilizar a los culpables de la muerte del niño, ni una reparación adecuada a sus padres debido a la larga duración de los procedimientos, así como a la denegación de la asistencia jurídica gratuita. Estimó que la tasa era desproporcionada en relación con la capacidad económica de los demandantes y que, por tanto, se había restringido el derecho de acceso a un tribunal.

Los demandantes solicitaron 182.000€ y 160.900€, respectivamente, en concepto de daños morales. El Tribunal Europeo de Derechos Humanos les concedió, en conjunto, la cantidad de 50.000€ en este concepto. Además, estableció que, si así lo solicitaban los demandantes, debía reiniciarse el procedimiento administrativo, el cual no fue capaz de resolver el caso de indemnización por haberseles denegado la asistencia jurídica gratuita.

3.3.3 A v. Croatia⁵

El asunto A c. Croacia, firme en fecha 14 de octubre de 2010, recoge la demanda de una ciudadana croata, quien denuncia que las autoridades pertinentes no la protegieron de la violencia doméstica de su ex marido a pesar de ser conocedores de sus repetidos ataques físicos y verbales y amenazas de muerte. La demandante, identificada como A, alega que se incumplieron las medidas interpuestas por sentencias nacionales de entrada en prisión y tratamiento psiquiátrico. Por tanto, las autoridades croatas fallaron en implementar medidas de protección positivas e infringieron así el Artículo 8 del Convenio, el cual recoge el respeto a la vida privada y familiar.

El ex marido de la demandante, identificado como B, sufría trastornos mentales severos tras haber sido torturado en un campo de concentración durante la Guerra de Croacia en 1992. Entre noviembre de 2003 y junio de 2006, hubo episodios de violencia del ex marido hacia la demandante. Entre 2004 y 2009, A inició una serie de procedimientos judiciales distintos contra B, por lo que se ordenaron ciertas medidas de protección, como prisión preventiva, tratamiento psiquiátrico, órdenes de alejamiento e incluso tiempo en prisión. Las órdenes de alejamiento y la prisión preventiva se cumplieron, pero no el tratamiento

⁵ El resumen de la sentencia está asistido por los comentarios de Arribas Gómez (2018) y García Toledano (2018).

psiquiátrico ni el tiempo en prisión. El ex marido fue sentenciado en octubre de 2009 a tres años en prisión por amenazas de muerte a uno de los jueces encargado de los procedimientos penales abiertos por violencia doméstica. En el momento de la demanda ante el TEDH se desconocía dónde estaba retenido o si estaba recibiendo tratamiento psiquiátrico.

La demandante alegó violaciones de los siguientes artículos: Artículo 2 (derecho a la vida), 3 (prohibición de la tortura, los tratamientos inhumanos y los tratamientos humillantes), 8 (respeto a la vida privada y familiar), 13 (derecho a una compensación justa) y 14 (prohibición de la discriminación).

En la sentencia del TEDH por violación del Artículo 8 se determinó que la protección a la víctima podría haber sido más efectiva si las autoridades hubieran tratado con único procedimiento judicial que les proporcionara una mejor visión de la situación, en lugar de varios procedimientos separados. La no implementación de las medidas anteriormente mencionadas supuso, por tanto, una vulneración del derecho a la vida privada de la demandante. El Tribunal consideró que no había fundamento para considerar una posible vulneración de los artículos 2, 3 y 13, y que la demanda sobre el artículo 14 era inadmisibile. La compensación ofrecida a la demandante por daños no pecuniarios fue 9 000 euros, además de 4 470 euros por gastos y costes incurridos.

Este caso se centra en la consideración de la violencia doméstica como una violación de derechos humanos. A pesar de no estar reconocido como tal en el Convenio, por cuestiones del momento en que se redactó el documento, el TEDH ha interpretado casos de violencia de género desde una visión de los derechos humanos en numerosas ocasiones. Esta evolución de la interpretación del Convenio permite considerar que la prohibición a la violencia doméstica es inherente en el documento. Se imponen así obligaciones positivas sobre los estados miembros para la protección de las mujeres cuando estén en una posición que les permita evitar la violencia de género. En la práctica, esto significa que se puede considerar responsables a los estados por actos cometidos por individuos privados al haber fallado en sancionar a estos individuos.

4. Descripción y análisis de los resultados

La información obtenida en el estudio se analiza mediante la evaluación de las traducciones y la comparación de los resultados obtenidos en las cuatro variables, es decir, traducción con Google Translate en el entorno memoQ con y sin base de datos terminológica del TEDH y traducción con Amazon Translate en la consola online con y sin base de datos terminológica del TEDH. Las traducciones resultantes de los tres textos del corpus se pueden consultar en el apéndice 2. En este análisis se procede a comentar, en distintas subsecciones, los resultados obtenidos con las traducciones automáticas de Amazon Translate y Google Translate.

4.1. Google Translate integrado en memoQ

Como se ha comentado en repetidas ocasiones en este trabajo, Google Translate se ha explotado desde el software de traducción memoQ en la versión de 30 días de prueba de su paquete profesional. Este sistema permite la importación de bases de datos terminológicas y la activación, bajo licencia del usuario, de distintos sistemas de traducción automática, entre ellos Google Translate.

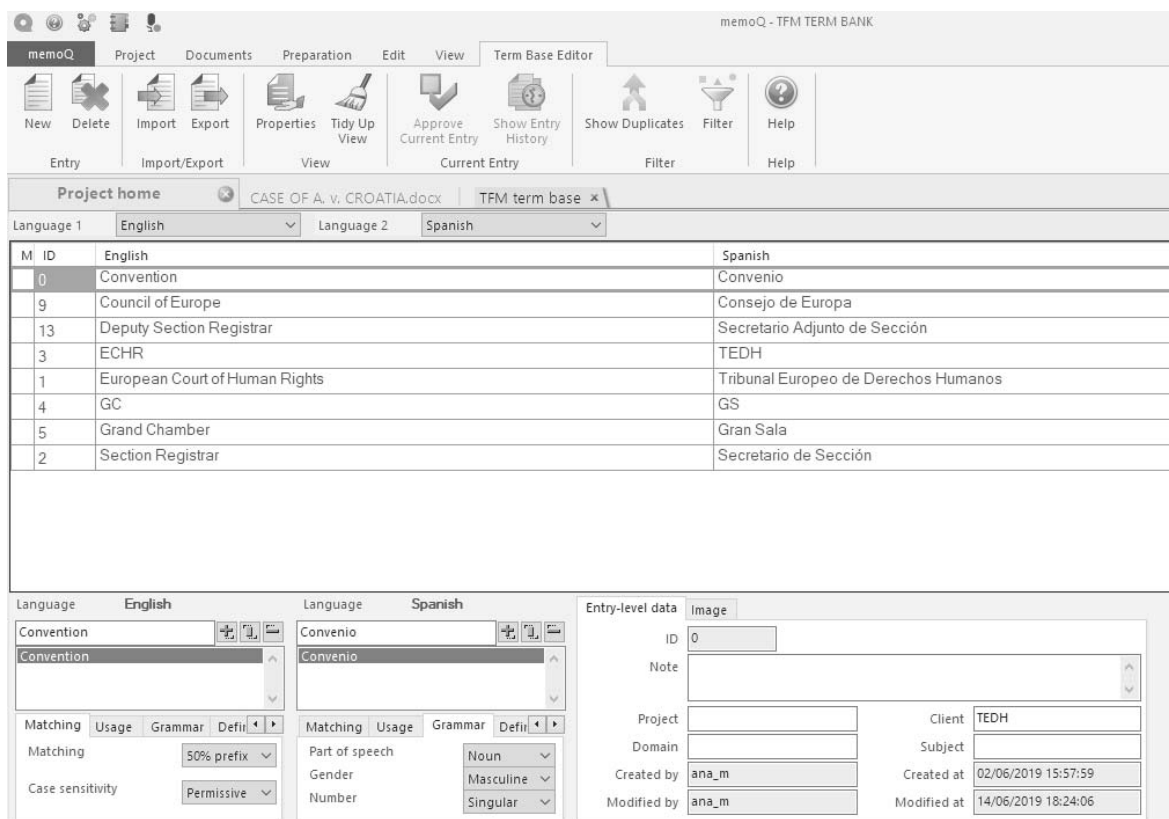


Figura 4: Editor de base de datos terminológica de memoQ

En la figura 4 se puede apreciar la base de datos terminológica utilizada en la investigación en su entorno de configuración. Una vez configurado e instalado el sistema de traducción automática con Google Translate y la terminología (en el experimento en el que se requiere), en el entorno de traducción de memoQ (figura 5) aparecen las sugerencias de Google Translate y de la terminología en el panel derecho, al igual que aparecerían sugerencias de la memoria de traducción si esta estuviera activa. Las sugerencias en azul provienen de la base de datos terminológica, mientras que la sugerencia de color naranja es el resultado del sistema de traducción automática. En color púrpura, memoQ sugiere una combinación de las distintas sugerencias de la terminología, memoria de traducción, etc.

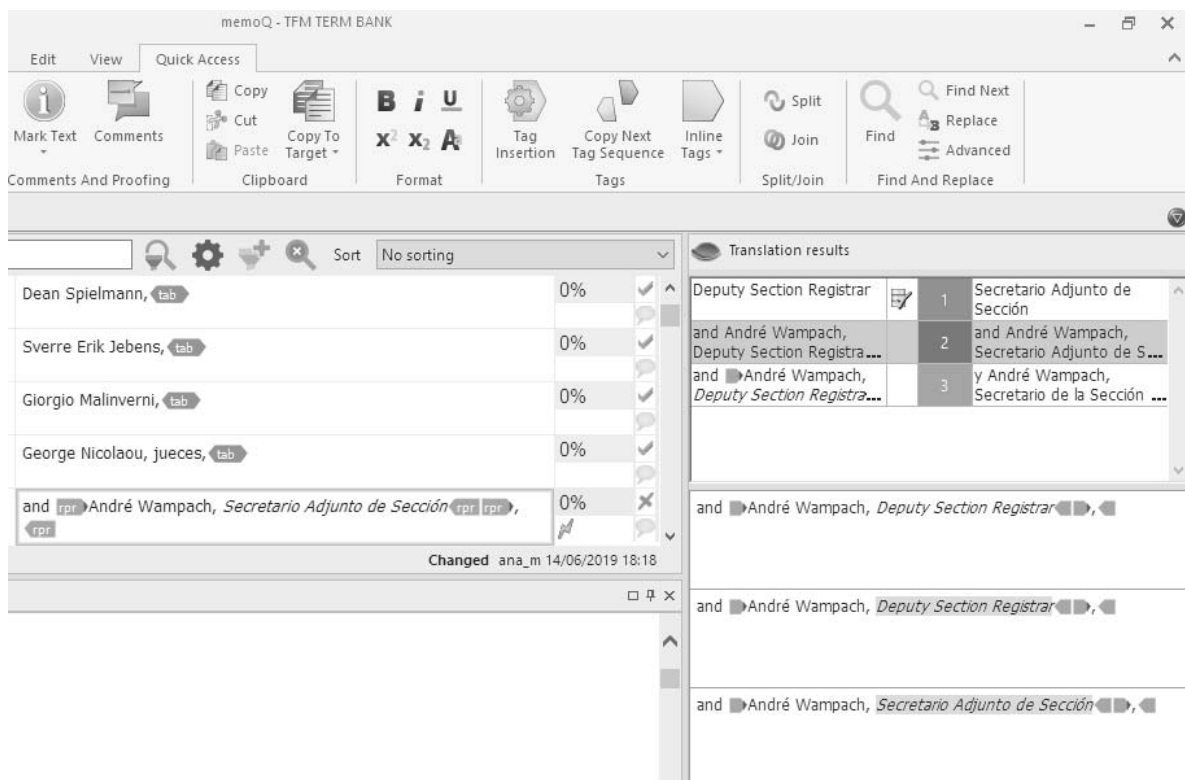


Figura 5: Panel de traducción de memoQ y panel derecho de resultados de traducción

Una vez aclarado el funcionamiento de memoQ, se procede al análisis de las traducciones obtenidas con este sistema.

4.1.1. Usabilidad

Como se observa en las capturas de pantalla insertadas en el apartado anterior, memoQ y, en concreto, del panel de traducción es una herramienta de uso relativamente sencillo. La instalación de los módulos adicionales como la base de datos terminológica o el sistema de

traducción automática están claramente explicados en la guía de uso de memoQ que se encuentra en línea⁶.

La gran ventaja del uso de traducción automática en un entorno de software de traducción es que las sugerencias del sistema de TA se pueden editar durante el proceso de traducción y no solamente en una fase final de posesición humana. memoQ segmenta el texto de forma que el traductor debe aceptar las sugerencias de Google Translate segmento a segmento, lo que podría verse como una mayor inversión de tiempo que obtener una traducción de todo el texto en un lote, pero permite un mayor control sobre el texto meta resultante.

Adicionalmente, la integración con memoQ permite preservar el formato del texto origen e insertar la traducción en este formato que posteriormente se exporta en un documento de extensión “doc”, por lo que se reduce el tiempo de maquetación en la posesición considerablemente. Dado que el sistema permite la edición durante la traducción, se permiten en esta fase cambios de formato, como puedan ser cursivas que el sistema no ha detectado, con lo que también se reduce el tiempo de posesición y la búsqueda posterior de las partes del texto que no estén correctamente formateadas.

4.1.2. Comentario de errores

En este apartado, localizamos y tipificamos algunos errores de traducción del sistema según la clasificación anterior (ver apartado 3.2.) con la intención de encontrar sus causas y determinar el impacto que han tenido sobre la calidad de la traducción y el esfuerzo de posesición que requerirían. Se analizan, por este orden, errores de redacción (ortografía y puntuación y léxico) y errores sobre el traslado del sentido (falso sentido, sin sentido y no mismo sentido), destacando entre ellos una predisposición general al uso del género masculino en elementos del texto origen sin marca de género.

En este análisis se presentan ejemplos representativos de los errores identificados en los textos traducidos. El comentario de corrección proviene de la normativa recogida en el Diccionario Panhispánico de Dudas (Real Academia Española, 2005).

En primer lugar, un error común de Google Translate en los textos del corpus ha sido el calco ortotipográfico del inglés al español, especialmente en las comas denominadas de tipo Oxford ante la conjunción que cierra una enumeración, el uso de comillas altas con cierre después del punto final de oración o la separación de cifras con comas. También se ha observado que las mayúsculas del texto origen no se mantienen homogéneamente en el

⁶ <https://help.memoq.com/8-5/en/Welcome/welcome-memoq-help.html>

texto meta, como ocurre constantemente en el caso de citar artículos del Convenio. En inglés, se citan como *Article*, indicando o no el número del artículo, pero siempre en mayúsculas, mientras que en español se ha mantenido la mayúscula en ocasiones y en otras se ha escrito con minúscula.

Asimismo, en el proceso de traducción se insertan mayúsculas que no había en origen, particularmente tras puntos de abreviaturas. Por el contrario, en *A c. Croacia* los puntos de las abreviaturas confunden al sistema y en finales de oración no los considera signos de puntuación. Otro añadido del proceso son los puntos finales en títulos de la sentencia. Por lo que respecta a la acentuación se observan varias omisiones claras y un patrón de no tildar las mayúsculas.

Otros errores ortotipográficos se observan de forma inconsistente. Ya en la propia traducción del nombre del asunto se observa disparidad en el empleo de distintas soluciones para la traducción de *v.*, abreviatura de *versus*. Desde mantenerla como en inglés hasta traducirlo como *c.* o *contra*, sin abreviar. De modo similar, la abreviatura de número se mantiene como *no.*, la abreviatura propia del inglés, en algunas ocasiones, mientras que en otras se proporcionan alternativas españolas. como *n^o* o *núm.*

En estos casos, los errores se producen por la falta de reglas al respecto en Google Translate. Dado que es un sistema por redes neuronales y anteriormente era estadístico, cabe inferir que no cuenta con una gran variedad de reglas ortotipográficas y el uso adecuado de las opciones en español viene determinado por la copia de textos en su base de datos donde sí se utiliza la alternativa correcta.

Se trata de errores de redacción en la lengua de llegada, en la que las convenciones de acentuación, signos de puntuación y tipografía son considerablemente distintas a la lengua de partida. A pesar de que no imposibilitan la comprensión del texto, en el contexto de esta investigación supondrían un mayor trabajo de posesición para el traductor, sobre todo por la inconsistente aparición de algunos de estos patrones.

Tabla 1. Errores de redacción en GT: ortografía y puntuación

Texto origen	Traducción Google Translate	Comentario
106. The applicant claimed 20,000 euros (EUR) in respect of non-pecuniary damage. (A v. Croatia)	106. El solicitante reclamó 20,000 euros (EUR) por daños no pecuniarios.	Los números se agrupan de tres en tres y separan con un espacio en blanco.

<p>“If the Court finds (...), the Court shall, if necessary, afford just satisfaction to the injured party.” (A v. Croatia)</p>	<p>“Si la Corte considera (...), la Corte deberá, si es necesario, brindar una justa satisfacción a la parte lesionada.”</p>	<p>La preferencia en español es por las comillas angulares, con cierre de comillas anterior al signo de puntuación.</p>
<p>3. On 3 September 2009 the President of the First Section decided to communicate to the Government the complaints concerning the lack of adequate positive measures under Articles 2, 3 and 8 of the Convention, the complaint concerning the lack of an effective remedy under Article 13 and the complaint under Article 14 that the applicant was discriminated against on the basis of her gender. (A v. Croatia)</p>	<p>3. El 3 de septiembre de 2009, el Presidente de la Sección Primera decidió comunicar al Gobierno las quejas sobre la falta de medidas positivas adecuadas en virtud de los artículos 2, 3 y 8 del Convenio, la queja sobre la falta de un recurso efectivo en virtud del Artículo 13 y la denuncia en virtud del artículo 14 de que la demandante fue discriminada por su género.</p>	<p>Los artículos o decretos de un documento legal se escriben en minúscula si no se cita el artículo completo. Además, el uso de mayúscula y minúscula es inconsistente a lo largo del texto.</p>
<p>(...) filed an indictment with the Ümraniye Criminal Court against S.Ç (the deputy headmaster of the school), S.Ö. (Atalay’s class teacher) and Ö.Ö (the school headmaster), (...) (İlbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)</p>	<p>(...) presentó una acusación ante el Tribunal Penal de Ümraniye contra S.Ç (el director adjunto de la escuela), S.Ö. (El maestro de la clase de Atalay) y Ö.Ö (el director de la escuela), (...)</p>	<p>El punto de las abreviaturas no exige el uso de una mayúscula posterior.</p>
<p>1. The submissions of the parties (İlbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)</p>	<p>1. Las presentaciones de las partes.</p>	<p>Nunca se escribe punto tras los títulos cuando aparecen aislados y son el único texto del renglón.</p>

<p>The extended indictment alleged that between November 2003 and February 2006 B had continually abused the applicant (...) and had on several occasions punched and kicked C. (A v. Croatia)</p>	<p>La acusación extendida alegaba que entre noviembre de 2003 y febrero de 2006, B había abusado continuamente del solicitante (...) y en varias ocasiones había golpeado y pateado a C</p>	<p>Falla en detectar el punto final tras C como un signo de puntuación. Asimismo, falta una coma de inicio de inciso tras la conjunción <i>que</i>.</p>
<p>II. RELEVANT DOMESTIC LAW Section 6 (A v. Croatia)</p>	<p>II. LEY DOMESTICA PERTINENTE Seccion 6</p>	<p>Las letras mayúsculas deben llevar tilde si así les corresponde según las reglas de acentuación.</p>
<p>51. She also argued that the requirement for her to prove an immediate risk to her life in order to have a protective measure of prohibition on harassing and stalking a victim of violence applied put a disproportionate burden on her as the victim of violent acts (see paragraph 35 above). (A v. Croatia)</p>	<p>51. También argumentó que el requisito para que ella demuestre un riesgo inmediato para su vida a fin de tener una medida protectora de la prohibición de acosar y acosar a una víctima de violencia aplicada le pone una carga desproporcionada como víctima de actos violentos (véase el párrafo 35). encima).</p>	<p>Se ha añadido un cierre de paréntesis y un punto y seguido innecesario.</p>
<p>(...) had had to hide in the shelter, together with C, and had also had to move to a secret address. (A v. Croatia)</p>	<p>(...) tuvo que esconderse en el refugio, junto con C, y También había tenido que mudarse a una dirección secreta.</p>	<p>Inserción de mayúscula a mitad de oración.</p>
<p>In the case of Otgon v. the Republic of Moldova, (...) and G.B. and R.B. v. the Republic of Moldova, (...) (Otgon v. the Republic of Moldova)</p>	<p>En el caso de Otgon v. República de Moldova, (...) y GB y RB c. La República de Moldova, (...)</p>	<p>Aunque versus ya se considera una palabra propia del español, la abreviatura correcta es vs.</p>

1. The case originated in an application (no. 22743/07) against the Republic of Moldova (...) (Otgon v. the Republic of Moldova)	1. El caso se originó en una solicitud (no. 22743/07) presentada ante la Corte (...)	La abreviatura de número es <i>n.º</i> , <i>núm.</i> o <i>nro.</i>
--	--	--

Más allá de las convenciones ortotipográficas, la traducción con Google Translate presenta otros patrones erróneos de redacción en la lengua de llegada. En la traducción de nombres de países se han observado problemas e inconsistencias léxicas.

Los errores que a continuación se presentan son errores de reformulación en la lengua de llegada que, además, se presentan de forma inconsistente a lo largo del texto. Aunque tampoco hacen el texto traducido incomprensible, esta falta de consistencia supone un escollo más en la tarea de posesición del traductor.

Tabla 2. Errores de redacción en GT: léxico

Texto origen	Traducción Google Translate	Comentario
In the case of Otgon v. the Republic of Moldova, (Otgon v. the Republic of Moldova) (...) (see L.C.B. v. the United Kingdom , 9 June 1998, § 36, Reports of Judgments and Decisions 1998-III). (Ilbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)	En el caso de Otgon v. República de Moldova , (...) (ver LCB v. United States Unido , 9 de junio de 1998, § 36, Informes de Sentencias y Decisiones 1998-III).	Se desaconseja el uso en español de la forma original vernácula <i>Moldova</i> . La denominación española es <i>el Reino Unido</i> .

Al centrarse en los errores sobre el traslado del sentido, se observa que hay siglas que el sistema no reconoce y, por tanto, no traduce, como es el caso de *PTSD*, expresión que sí reconoce y traduce cuando las siglas están desarrolladas, incluso añadiendo las siglas españolas correspondientes, aunque en el texto origen no estuvieran. También se observan fallos en el reconocimiento de nombres de instituciones, principalmente en *A v. Croatia* por la confusión que provocan las abreviaciones para mantener el anonimato de las personas implicadas y la ciudad en la que ocurrieron los hechos.

La clasificación a continuación se basa en la distinción entre sin sentido, un error que crea incomprensión en el texto; falso sentido, la atribución de un sentido diferente al original; y no mismo sentido, una apreciación inadecuada de un matiz del original que introduce ambigüedad, aunque las fronteras entre las tres subcategorías son en ocasiones difusas.

Tabla 3. Errores sobre el traslado del sentido en GT: sin sentido

Texto origen	Traducción Google Translate	Comentario
The expert stated that, owing to his difficult war experiences, B suffered from PTSD ; (A v. Croatia)	El experto afirmó que, debido a sus difíciles experiencias de guerra, B sufría de PTSD ;	No reconoce las siglas.
(...) under the expert supervision of the Z. Social Welfare Centre . (A v. Croatia)	(...) bajo la supervisión de un experto de Z. Social Welfare Centre .	Z, identificativo de la ciudad donde ocurrieron los hechos, confunde al sistema, que no traduce la parte posterior.
“ B is guilty on the grounds that (A v. Croatia)	" SEGUNDO es culpable con base en que	B (el acusado), es interpretado como segundo elemento de una enumeración.
Registrar President (Ilbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)	Presidente del registrador	No los detecta como títulos independientes.
(...) (see L.C.B. v. the United Kingdom , 9 June 1998, § 36, Reports of Judgments (...). (Ilbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)	(...) (ver LCB v. United States Unido , 9 de junio de 1998, § 36, Informes de Sentencias (...).	La denominación española es <i>el Reino Unido</i> .

Tabla 4. Errores sobre el traslado del sentido en GT: falso sentido

Texto origen	Traducción Google Translate	Comentario
(...) respectively for the court fees within one month in order for the proceedings to be continued . (Ilbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)	(...) respectivamente por los gastos judiciales dentro de un mes para el procedimiento continuará .	<i>To</i> es una preposición de finalidad en este caso.
The applicant's partner opened and the man at the door introduced himself as a private detective hired by B to find out the whereabouts of the applicant and C. (A v. Croatia)	El compañero del solicitante se abrió y el hombre de la puerta se presentó como un detective privado contratado por B para averiguar el paradero del solicitante y C.	No es reflexivo, se refiere a que abrió la puerta.

Tabla 5. Errores sobre el traslado del sentido en GT: no mismo sentido

Texto origen	Traducción Google Translate	Comentario
30. (...) the respondent State had failed to protect the right to life of their child . (Ilbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)	30. (...) el Estado demandado no había protegido el derecho a la vida de sus hijos .	<i>Child</i> es singular.
That hearing was also adjourned because B did not appear . (A v. Croatia)	Esa audiencia también fue aplazada porque B no apareció .	En contextos jurídicos, la equivalencia en español será <i>comparecer</i> (presentarse personalmente o por poder ante un órgano público, especialmente ante un juez o tribunal).

Como un caso concreto de no mismo sentido, es también un patrón destacable del sistema la predisposición por el género masculino cuando no existen marcas de género en el texto

origen. Las demandantes en el corpus son en ocasiones mujeres y, aun habiendo establecido su género en los primeros párrafos del texto, se prefieren fórmulas como *el demandante* en la gran mayoría de ocasiones en las que aparecen en el texto. El análisis de la distribución del artículo masculino y femenino en los textos revela que la traducción del sistema se lleva a cabo segmento a segmento y que el sistema no es capaz de “recordar” que la persona a la que se refiere el texto es una mujer más allá de un par de frases después de haberlo establecido así.

Tabla 6. No mismo sentido en GT: predisposición por el género masculino

Texto origen	Traducción Google Translate	Comentario
<p>The indictment alleged that from 12 November 2003 to 21 August 2005 B had verbally insulted and threatened the applicant, prevented her from leaving the house and physically assaulted her; on 12 November 2003 he had physically assaulted her by punching her in the belly, throwing her on the floor and continuing to hit and kick her in the body and head; on 7 August 2005 he had hit the applicant in the face, back and hands, causing lacerations; and on 21 August 2005 he had kicked her in the leg.</p> <p>(A v. Croatia)</p>	<p>La acusación alega que desde el 12 de noviembre de 2003 hasta el 21 de agosto de 2005, B insultó verbalmente y amenazó a la demandante, le impidió salir de la casa y la agredió físicamente; el 12 de noviembre de 2003 la había agredido físicamente golpeándola en el vientre, tirándola al suelo y continuando golpeándola y pateándola en el cuerpo y la cabeza; el 7 de agosto de 2005 había golpeado al solicitante en la cara, la espalda y las manos, causando laceraciones; y el 21 de agosto de 2005 la había pateado en la pierna.</p>	<p><i>The applicant</i> no está marcado por género, pero sí los pronombres posteriores en la oración, por lo que el sistema identifica correctamente que se trata de una demandante.</p>

Como aspecto positivo del sistema, es capaz de identificar que Françoise Tulkens, presidenta de la Sala en el asunto İlbeyi Kemaloğlu y Meriye Kemaloğlu c. Turquía, es una

mujer, a pesar de que el texto traducido por humanos y publicado por el Ministerio de Justicia la identifica como *presidente*⁷.

Estos errores presentados en las tablas anteriores sí provocan alteraciones en el significado del texto y podrían conllevar incompreensión por parte del lector del mismo, por lo que se requiere una tarea de posesición con lectura detallada de la traducción y corrección de los términos empleados por el sistema.

Cabe destacar que llegados a este punto se comenta por primera vez en este análisis la traducción de terminología jurídica de términos polisémicos con significado también en la lengua común (más en 4.1.4.). En este caso el sistema presenta inconsistencias, ya que la polisemia no existe en español para este término (*aparecer – comparecer*), siendo *comparecer* el término adecuado en contextos jurídicos, al contrario que en el inglés, donde se utiliza el mismo vocablo para ambos significados. A lo largo del texto se traduce indistintamente como uno u otro sin que se pueda identificar un patrón que explique este comportamiento.

4.1.3. La traducción con y sin terminología

Este apartado corresponde a la decisión de llevar a cabo dos traducciones, una de ellas con base de datos terminológica y otra sin. Con las traducciones resultantes de ambos procesos, las diferencias entre las versiones se centran sobre todo en dos términos concretos, el Convenio Europeo de Derechos Humanos y la nomenclatura del secretariado de las salas. A través de la terminología, además, se fuerza al sistema a emplear convenciones tipográficas del español, como las minúsculas en títulos y cargos.

Tabla 7. Diferencias entre traducción con y sin terminología en GT

Texto origen	Traducción Google Translate sin terminología	Traducción Google Translate con terminología
and Stanley Naismith, Section Registrar , (Ilbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)	y Stanley Naismith, Registrador de Sección ,	y Stanley Naismith, secretario de sección ,
and André Wampach, Deputy Section Registrar , (A v. Croatia)	y André Wampach, Secretario de la Sección Adjunta ,	y André Wampach, secretario adjunto de sección ,

⁷ Aunque «presidente» puede usarse como común en cuanto al género («el/la presidente»), es preferible hoy usar el femenino «presidenta», documentado en español desde el siglo XV y registrado en el diccionario académico desde 1803. Fuente: <https://www.rae.es/noticias/es-la-presidenta-o-la-presidente>

Cabe destacar, sin embargo, que la introducción de terminología ha provocado errores en la traducción no presentes en la versión sin terminología. Por ejemplo, el sistema, sin la información proporcionada por la base de datos terminológica ha traducido indistintamente el Convenio Europeo de Derechos Humanos como *el Convenio* y como *la Convención*. En la versión con la base de datos terminológica, a pesar de haber indicado el género masculino del sustantivo, se ha introducido forzosamente en la traducción, por lo que se mantiene el artículo femenino en los casos en los que el sistema habría elegido el término *la Convención*.

Por otro lado, en caso de toparse con ellas, el sistema ha interpretado las siglas ECHR, que en el texto se refieren a *European Court of Human Rights*, como *European Convention of Human Rights*, ya que en inglés ambas nomenclaturas comparten siglas. Se ha traducido así al español como CEDH (Convenio Europeo de Derechos Humanos), cambiado en la versión con terminología al adecuado TEDH (Tribunal Europeo de Derechos Humanos). No obstante, se plantea el problema de haber creado un error por mala identificación de las siglas en la versión con terminología en caso de que sí hubieran aparecido en alguno de los textos las siglas del Convenio que el sistema identifica correctamente y las modificara por las del Tribunal por sugerencia de la base de datos.

Es algo similar a lo que ocurre en *A v. Croatia*, en el que el uso de la base de datos terminológica crea un error que no estaba en la versión propia del sistema. En este caso, la nomenclatura correcta del documento es *Convención*.

Tabla 8. Errores provocados por la introducción de la terminología en GT

Texto origen	Traducción Google Translate sin terminología	Traducción Google Translate con terminología
She relied on Articles 2, 3 and 8 of the Convention , the relevant parts of which read: (A v. Croatia)	Se basó en los artículos 2, 3 y 8 de la Convención , cuyas partes relevantes decían:	Se basó en los artículos 2, 3 y 8 de la Convenio , cuyas partes relevantes decían:
(...) and prepared by the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) (...) (A v. Croatia)	(...) y preparados por el Comité de la Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (CEDAW), (...)	(...) y preparados por el Comité de la Convenio sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (CEDAW), (...)

4.1.4. Terminología jurídica

Dado que uno de los objetivos de esta investigación es el análisis de utilidad de la traducción automática en traducción especializada, más concretamente en traducción jurídica, se plantea este apartado de análisis. La terminología jurídica no propia del TEDH no se ha incluido en la base de datos terminológica para así comprobar la actuación del sistema en casos de polisemia de terminología con significados jurídicos y comunes.

En el caso de Google Translate, se han identificado una serie de puntos clave que determinan la actuación del sistema. En primer lugar, se observa que se ha traducido indistintamente *court* tanto como *tribunal* como *corte*. Según el Diccionario del Español Jurídico (2016), *corte* se reserva para los tribunales internacionales, por lo que su uso para referirse a cualquier otro tribunal sería incorrecto.

La traducción de *judgement* como *juicio* en los primeros párrafos en el sentido de *sentencia* decidida por el tribunal sería un caso de falso sentido, ya que se trata de una palabra polisémica con un significado común y otro jurídico y el sistema ha optado por el significado común. Otros ejemplos que se dan en la traducción de términos con polisemia jurídica en Google Translate son: *application* como *solicitud* y no *demanda* (también en la nomenclatura de las personas que han presentado el acto jurídico como *solicitantes* y no *demandantes*); *final* como *final* y no *firme*; *merits* como *méritos* y no *fondo*; *award* como *premio* y no *indemnización*; *holds* como *sostiene* y no *considera*; o *costs* como *costos*, cuando existe el término *costas*, que se refiere a los gastos relacionados con el juicio.

Se observan, además, calcos del inglés que producen sinsentidos en español, como es el caso de *alegada violación* (del artículo del Convenio) en el asunto de Turquía, traducción de *alleged violation*, cuando en español el adjetivo correcto podría ser tanto *presunta* como *supuesta* (presentes en los otros dos asuntos). *Alegar* sí es un término empleado comúnmente en el lenguaje jurídico, pero con el significado «argumentar oralmente o por escrito hechos y derechos en la defensa de una pretensión» (Real Academia Española, 2016).

Además de estas elecciones del lenguaje común en situaciones de polisemia, se echa en falta terminología y fraseología jurídica establecida en el sistema español, como el «fundamento de derecho», «antecedentes de hecho», «valoración del tribunal» o «argumentos de las partes», traducidos como «la ley», «los hechos», «observaciones de las partes» y «evaluación de la corte», respectivamente. Más allá de la terminología, también el estilo de redacción de sentencias judiciales en España correspondiente al ordenamiento jurídico es menos perceptible debido a las características propias del TEDH.

4.2. Amazon Translate en la consola de Amazon Web Services

Por la dificultad que entraña la instalación, configuración y uso de JDK para la traducción de documentos largos con Amazon Translate comentada anteriormente, las pruebas en este trabajo se han realizado en la consola en línea del servicio. El uso de Amazon Translate se rige bajo una licencia de la API, con un coste proporcional al uso que se haga de ella. La consola en línea, a pesar de la limitación de caracteres a 5000 bytes, sí permite la importación y uso de bases de datos terminológicas, tanto desde un formato “csv” como “tmx”.

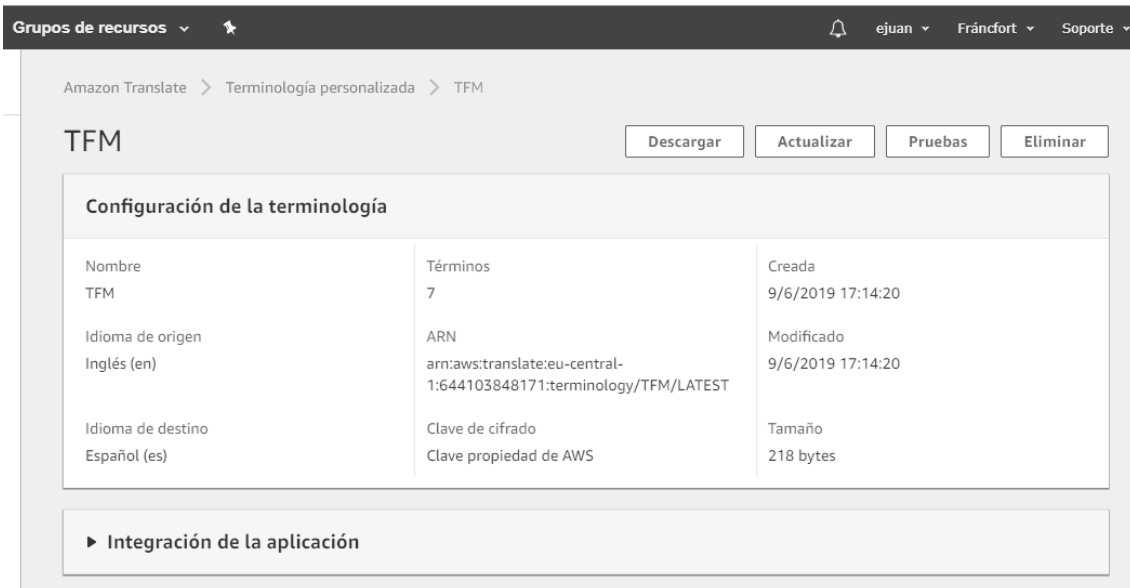


Figura 6: Características de la terminología en Amazon Translate

En la consola se muestra la posibilidad de usar la terminología en la traducción que se esté llevando a cabo y la selección de los idiomas implicados en la traducción. Como se observa en la figura 7, se trata de una interfaz muy sencilla y similar a consolas online conocidas, como puede ser el propio Google Translate.

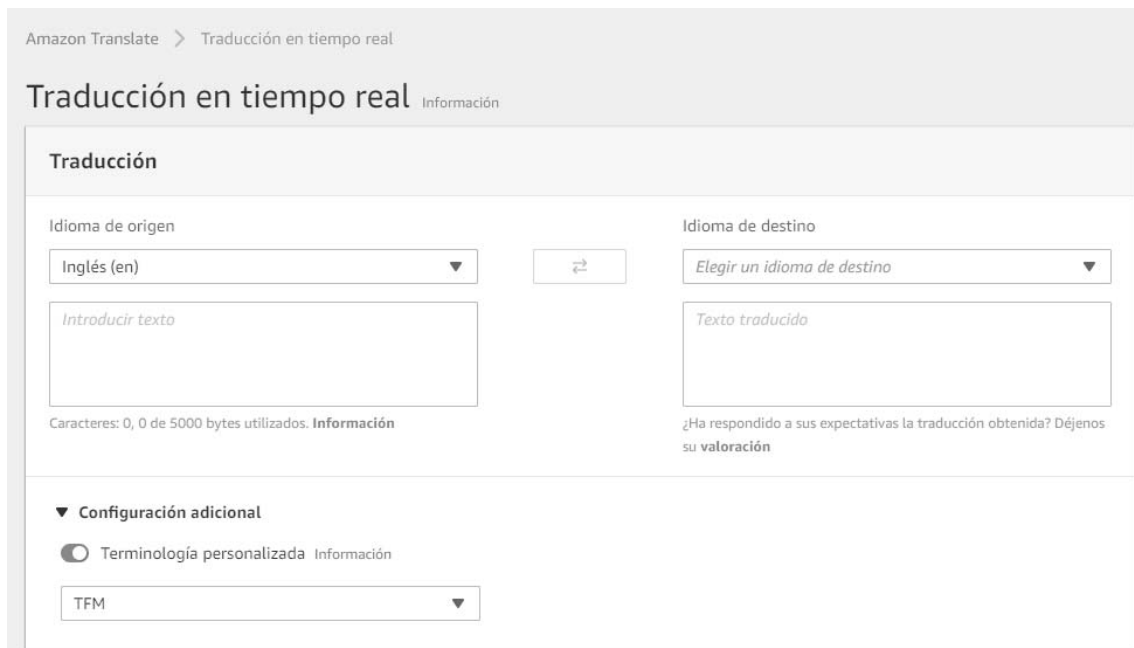


Figura 7: Consola de Amazon Translate en Amazon Web Services

Se procede así también al análisis del uso de Amazon Translate para la traducción de sentencias del TEDH.

4.2.1. Usabilidad

A pesar de la sencillez de la consola, en comparación con el uso de Google Translate, la limitación de caracteres supone un escollo en la traducción fluida del texto. Esta limitación se encuentra también en la interfaz de línea de comando y es solo salvable con el uso de Java. Esto significa que la traducción de textos ha sido segmentada en secciones de aproximadamente 5000 caracteres, con el empleo de tiempo y recursos que eso conlleva.

El sistema no permite la creación de terminología *in situ*, siendo necesaria la creación de un documento “csv” o “tmx” para posteriormente subirlo a la consola. En el caso de este experimento, se ha exportado la terminología de memoQ y subido a Amazon Web Services para contar una terminología homogénea en ambos experimentos. La inserción de terminología en caso de estar activada la opción es automática, por lo que no se brinda la posibilidad de editar la traducción durante el proceso. La posesición en este caso debe ser sobre el texto finalizado.

Debido al uso de la consola online, el formato de salida del texto no es relevante, ya que es el usuario el que debe copiarlo y pegarlo. Para mantener el formato del documento propio del TEDH como en el caso de Google Translate y memoQ, el trabajo de maquetación recae también sobre el usuario en el momento de posesición.

4.2.2. Comentario de errores

Al igual que con Google Translate, se procede a localizar y comentar de forma representativa algunos errores de traducción de Amazon Translate. Se analizan, por este orden, errores de redacción (ortografía y puntuación) y errores sobre el traslado del sentido (adición y omisión, falso sentido, sin sentido y no mismo sentido), destacando entre ellos una predisposición general al uso del género masculino en elementos del texto origen sin marca de género.

En Amazon Translate se observa también el calco de la puntuación del inglés al español, sobre todo en el uso de comas y de comillas altas. La separación de las cifras no se ha calcado del inglés, pero se han usado puntos para separar los grupos de tres cifras, lo que tampoco es recomendado por la Academia de la Lengua. En este caso, la tipografía de *Artículo* es consistente a lo largo del texto, ya que se escribe en minúscula.

De forma similar a Google Translate, estos errores ortotipográficos suponen una mayor tarea de posesición para el traductor, además incrementada por la inconsistencia con la que se presentan estos patrones, lo cual dificulta las tareas de detección y reemplazo.

Tabla 9. Errores de redacción en AT: ortografía y puntuación

Texto origen	Trad. Amazon Translate	Comentario
The Government argued that in Croatia the protection of victims of domestic violence was ensured through the mechanisms of criminal law, and in particular the Protection against Domestic Violence Act. (A c. Croatia)	El Gobierno sostuvo que en Croacia la protección de las víctimas de la violencia doméstica se garantizaba mediante los mecanismos del derecho penal, y en particular la Ley de protección contra la violencia doméstica.	La puntuación española sitúa <i>en particular</i> entre comas. Se eliminaría la coma calcada anterior a y.
106. The applicant claimed 20,000 euros (EUR) in respect of non-pecuniary damage. (A v. Croatia)	106. La demandante reclamó 20.000 euros (euros) por daños no pecuniarios.	Los números se agrupan de tres en tres y separan con un espacio en blanco.
Işıl Karakaş, President ,	Işıl Karakaş, Presidente ,	Los títulos y cargos se escriben con minúscula.

(...) in practical terms, remove the risk to his environment.” (g) seizure of objects intended for or used in the commission of a minor offence.” (A v. Croatia)	(...) en términos prácticos, eliminaría el riesgo para su medio ambiente.” g) incautación de objetos destinados o utilizados en la comisión de un delito menor “.	La preferencia en español es por las comillas angulares, con cierre de comillas anterior al signo de puntuación.
She relied on Article 13 of the Convention, which provides: (A v. Croatia)	Se basó en el artículo 13 de la Convenio, que dispone lo siguiente:	Los artículos o decretos de un documento legal se escriben en minúscula si no se cita el artículo completo.

Por lo que respecta a patrones léxicos como los comentados con Google Translate, se observa el mismo uso de *Moldova* en la traducción del país conocido como Moldavia en España.

En los errores sobre el traslado del sentido incluimos adiciones y supresiones tanto léxicas como de ortografía y puntuación, entre ellas, la eliminación del paréntesis de apertura de los elementos introductorios en algunas enumeraciones. A pesar de que este uso es correcto, también lo sería el paréntesis doble del texto original en inglés. Asimismo, se detectan de dos puntos detrás de la exposición de artículos en el asunto de *A c. Croacia* que no están en el texto original en inglés. A pesar de no se trataría de un error ortotipográfico en español, ya que los dos puntos se utilizan para preceder una cita, su inserción en el texto es inconsistente. También se comentan otras adiciones no presentes en el texto origen y sustracciones de palabras o frases.

Tabla 10. Errores sobre el traslado del sentido en AT: omisión y adición

Texto origen	Trad. Amazon Translate	Comentario
50. The applicant argued that the State authorities had (...). She maintained that although the national courts, (...)	50. La demandante alegó que las autoridades estatales habían (...). La autora sostuvo que, aunque los tribunales nacionales, (...)	Se añade <i>la autora</i> y <i>la oradora</i> , no presente en el texto origen.

<p>In that connection she stressed that the training of experts (...) (A v. Croatia)</p>	<p>A ese respecto, la oradora subraya que la capacitación de los expertos (...)</p>	
<p>Article 75 (A v. Croatia)</p>	<p>Artículo 75:</p>	<p>Preceden a la reproducción de citas o palabras textuales. La adición es inconsistente a lo largo del texto.</p>
<p>(c) The fourth set of proceedings (A v. Croatia)</p>	<p>c) Cuarto conjunto de procedimientos</p>	<p>Las letras o números que introducen elementos de una clasificación o enumeración pueden escribirse entre paréntesis o, más frecuentemente, seguidas solo del paréntesis de cierre.</p>
<p>(...) are still pending before the Court of Cassation, although eight years have elapsed since the incident.</p> <p>53. The Government argued that in Croatia the protection of victims (...) (Ilbeyi Kemaloglu and Meriye Kemaloglu v. Turkey)</p>	<p>(...) siguen pendientes ante el Tribunal de Casación, aunque han transcurrido ocho años desde el incidente - Sí.</p> <p>53 - Sí. El Gobierno sostuvo que en Croacia la protección de las víctimas (...)</p>	<p>Se inserta - <i>Sí</i>. en dos oraciones sin estar esta expresión presente en el texto origen.</p>

– physical assault irrespective of whether it has caused injury; (A v. Croatia)	— agresión física independientemente de que	Falta parte de la oración.
--	--	----------------------------

Se presenta una selección de sentidos diferentes al texto original, algunas de ellas creando sinsentidos. Estos son errores de reformulación que sí pueden suponer un mayor escollo para el traductor durante la tarea de posesición por producir incompreensión en el usuario y alterar el significado del texto.

Tabla 11. Errores sobre el traslado del sentido en AT: no mismo sentido

Texto origen	Trad. Amazon Translate	Comentario
19. Further to that, on 30 June 2006 B was arrested and detained on suspicion of the criminal offence of making death threats against the applicant and a police officer, I.G. (A v. Croatia)	19. Además, el 30 de junio de 2006 B fue detenido y detenido bajo sospecha de haber cometido amenazas de muerte contra el solicitante y un agente de policía. I.G.	Se cambia la coma anterior a <i>I.G.</i> por un punto.
8. On 1 March 2006 the Călărași District Court found in her favour. (Otgon v. the Republic of Moldova)	8. El 1º de marzo de 2006, el Tribunal de Distrito de Călărași falló a su favor.	Se presenta la fecha con número cardinal y con números ordinales en el resto del texto.
– verbal assaults, insults, cursing, calling names or other forms of serious harassment; (A v. Croatia)	— agresiones verbales, insultos, insultos, insultos, insultos u otras formas de acoso grave;	Se pierden las connotaciones y matices con la traducción de todos los términos como <i>insultos</i> . Además, se añade un elemento más a la enumeración.

Tabla 12. Errores sobre el traslado del sentido en AT: falso sentido

Texto origen	Trad. Amazon Translate	Comentario
Section 1 (A v. Croatia)	Artículo 1:	Dado que se trata de las secciones de los artículos del Convenio, el cambio en la traducción a <i>artículo</i> crea una confusión innecesaria.
97. The Court notes at the outset (...) (A v. Croatia)	97. El Tribunal de Primera Instancia señala, en primer lugar, (...).	<i>Court</i> aquí se refiere al TEDH. Se ha añadido erróneamente <i>de Primera Instancia</i> .

Tabla 13. Errores sobre el traslado del sentido en AT: sin sentido

Texto origen	Trad. Amazon Translate	Comentario
VIOLENT BEHAVIOUR WITHIN THE FAMILY (A v. Croatia)	COMPOTACIÓN VIOLENCIA EN LA FAMILIA	No existe constancia de uso del término <i>compotación</i> en la lengua española.
PURPOSE OF A PRISON TERM (A v. Croatia)	FINALIDAD DE LA PLAZ DE PRISIÓN	Se presupone que el término a insertar era <i>plazo</i> , pero falta la <i>o</i> final.
97. The Court notes at the outset (...) (A v. Croatia)	97. El Tribunal de Primera Instancia señala, en primer lugar, (...).	<i>Court</i> aquí se refiere al TEDH. Se ha añadido erróneamente <i>de Primera Instancia</i> .

<p>50. The applicant argued that the State authorities had failed (...). She maintained that although the national courts, (...)</p> <p>In that connection she stressed that the training of experts (...) (A v. Croatia)</p>	<p>50. La demandante alegó que las autoridades estatales habían incumplido (...). La autora sostuvo que, aunque los tribunales nacionales, (...)</p> <p>A ese respecto, la oradora subraya que la capacitación de los expertos (...)</p>	<p>Se añade <i>la autora</i> y <i>la oradora</i>, no presente en el texto origen.</p>
<p>(...) he approached A on the street in front of the building of the Social Welfare Centre and whispered in her ear (...) (A v. Croatia)</p>	<p>(...) se acercó a A en la calle frente al edificio del Bienestar Social Centro y susurró en su oído (...)</p>	<p>El orden de palabras no es correcto, ya que Centro es el primer elemento en español.</p>

Al igual que con Google Translate, la primera elección del sistema de Amazon Translate en términos no marcados por género en el texto original es también el género masculino, como se puede apreciar en los siguientes ejemplos. Más allá de la predisposición mayoritaria por el género masculino, existe una gran inconsistencia a lo largo de los textos, ya que se selecciona el género femenino en ocasiones sin estar el original así marcado. Aunque se podría pensar que se debe a la cercanía a otro elemento marcado, no es siempre así, por lo que no se puede determinar que sea un patrón consistente. Tampoco se ha detectado que la limitación de caracteres del sistema suponga un corte en la identificación del género del usuario en cuestión, habiendo inconsistencias dentro del mismo segmento traducido.

Tabla 14. No mismo sentido en AT: predisposición por el género masculino

Texto origen	Trad. Amazon Translate	Comentario
<p>5. The applicant was born in 1963 and lives in Călărași.</p> <p>6. On 26 October 2005 the applicant and her daughter drank water from taps in their apartment and shortly thereafter they felt unwell.</p> <p>(Otgon v. the Republic of Moldavia)</p>	<p>5. El solicitante nació en 1963 y vive en Călărași.</p> <p>6. El 26 de octubre de 2005, la demandante y su hija bebieron agua de los grifos de su apartamento y poco después se sintieron mal.</p>	<p>En el párrafo 5 no existe marca de género y se traduce en masculino, mientras que <i>her</i> en el pár. 6 determina el género femenino.</p>
<p>18. Accordingly, on 4 October 2004 the Ümraniye Public Prosecutor filed an indictment with the Ümraniye Criminal Court against S.Ç (the deputy headmaster of the school), S.Ö. (Atalay's class teacher) and Ö.Ö (the school headmaster) (...)</p> <p>Atalay's class teacher, Ms S.Ö., maintained that a day before the incident she had (...)</p> <p>(İlbeyi Kemaloğlu and Meriye Kemaloğlu v. Turkey)</p>	<p>18. En consecuencia, el 4 de octubre de 2004, el fiscal de Ümraniye presentó una acusación ante el Tribunal Penal de Ümraniye contra S.Ç (director adjunto de la escuela), S.Ö. (profesor de la clase de Atalay) y Ö.Ö (director de la escuela) (...)</p> <p>La maestra de la clase de Atalay, la Sra. S.Ö., sostuvo que un día antes del incidente había (...)</p>	<p>En esta ocasión es el título <i>Ms</i> el que permite que el sistema reconozca el sexo de la persona a la que hace referencia.</p>
<p>5. On 21 April 2001 the applicant married B and on 14 May 2001 a daughter, C, was born of the marriage. On 13 December 2005 the applicant brought a civil action in the Z. Municipal Court (Opéinski sud u Z.), seeking a divorce from B. On 7 November 2006 the court dissolved the marriage of the applicant and B.</p> <p>(A v. Croatia)</p>	<p>5. El 21 de abril de 2001 el solicitante se casó con B y el 14 de mayo de 2001 nació una hija, C. El 13 de diciembre de 2005, la demandante interpuso una acción civil ante el Tribunal Municipal de Z. (Opéinski sud u Z.), solicitando el divorcio de B. El 7 de noviembre de 2006, el tribunal disolvió el matrimonio de la demandante y B.</p>	<p>Se observa la inconsistente selección de género en tres ocasiones durante el mismo párrafo en las que el original contiene la misma expresión no marcada, <i>the applicant</i>.</p>

4.2.3. La traducción con y sin terminología

El impacto del uso de la base de datos terminológica en Amazon Translate ha sido similar a lo comentado en el apartado correspondiente con Google Translate. Sin embargo, se destacan una serie de características propias del sistema que nos ocupa, como el hecho de que en ocasiones no detecta la terminología que debe sustituir si esta está escrita en mayúsculas y cuando la detecta inserta el término en minúsculas, como está en la base de datos terminológica, sin adaptarlo al formato.

Como se observa también en el ejemplo presentado, Amazon Translate presenta problemas de identificación del género de la terminología, pero, al contrario que Google Translate, no cuenta con un sistema de explicitación de las características del término. La introducción del artículo como parte de la terminología se descarta por no ser buena práctica al dificultar la distinción de artículos definidos e indefinidos y la construcción de contracciones.

La mala identificación de terminología también se produce en Amazon Translate en el mismo ejemplo presentado anteriormente, así como la traducción de las siglas ECHR como CEDH (Convenio Europeo de Derechos Humanos).

Tabla 15. Errores provocados por la introducción de la terminología en AT

Texto origen	Traducción Amazon Translate sin terminología	Traducción Amazon Translate con terminología
I. ALLEGED VIOLATION OF ARTICLE 2 OF THE CONVENTION (Ilbeyi Kemaloglu and Meriye Kemaloglu v. Turkey)	I. PRESUNTA VIOLACIÓN DEL ARTÍCULO 2 DE LA CONVENCION	I. PRESUNTA VIOLACIÓN DEL ARTÍCULO 2 DE LA CONVENCION
(...) and prepared by the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) (...) (A v. Croatia)	(...) y preparados por el Comité de la Convención sobre la eliminación de todas las formas de discriminación contra la mujer (...)	(...) y preparados por el Comité de la Convenio sobre la eliminación de todas las formas de discriminación contra la mujer (...)
(...) M.C. v. Bulgaria, no. 39272/98, §§ 150 and 152, ECHR 2003-XII;	(...) M. C. c. Bulgaria, N° 39272/98, § § 150 y 152, CEDH 2003 XII; Bevacqua	(...) M. C. c. Bulgaria, N° 39272/98, § § 150 y 152, TEDH 2003 XII; Bevacqua

Bevacqua and S. v. Bulgaria, no. 71127/01, § 65, 12 June 2008; and Sandra Janković, cited above, § 45). (A v. Croatia)	y S. c. Bulgaria, N° 71127/01, § 65, 12 de junio de 2008; y Sandra Janković, citada supra, § 45).	y S. c. Bulgaria, N° 71127/01, § 65, 12 de junio de 2008; y Sandra Janković, citada supra, § 45).
---	---	---

4.2.4. Terminología jurídica

Al igual que en los apartados comentados anteriormente, se observan grandes similitudes en el tratamiento de la terminología jurídica de los textos en relación con Google Translate. Aparece, por tanto, también un uso indistinto de *corte* y *tribunal* para identificar al TEDH, aunque sí se observa que los tribunales nacionales suelen identificarse de esta manera, como tribunales. En relación a este punto, otro patrón destacable en el asunto de *A v. Croatia* es la denominación de *Minor Offences Court* tanto como *Tribunal de Infracciones Menores* como *Tribunal de Delitos Menores* a lo largo del mismo texto refiriéndose al mismo tribunal.

Algunas traducciones que se apreciaban en Google Translate también presentes en Amazon Translate son: *solicitud* por *demanda* y *solicitantes* por *demandantes*, *juicio* por *sentencia*, *final* por *firme*, *sostener* y *no considerar* en el fallo de la sentencia, así como *gastos* por *costas*. Es interesante destacar que el *default interest*, los intereses de demora que se recogen en el fallo de las sentencias, ha sido traducido de dos formas diferentes en los textos, como *intereses por incumplimiento* en Turquía y Moldavia e *intereses por impago* en Croacia.

En Amazon Translate tampoco se encuentra gran variedad ni cantidad de fraseología jurídica propia del español.

4.3. Comentario cualitativo de los resultados

Como se ha comprobado, el análisis de errores de ambos sistemas ha revelado patrones similares de calco, adiciones, ausencias y preferencias. Los errores ortotipográficos se podrían explicar por la configuración de ambos sistemas por redes neuronales sin reglas de uso, ya que se estarían copiando patrones observados en los textos de los que dispone el sistema para aprender.

Se observa que el número de sinsentidos aumenta con la longitud del texto, siendo *A v. Croatia* el que cuenta con más ejemplos, también a causa de la confusión creada por el uso de A, B y C para identificar a las partes implicadas en el asunto. Asimismo, ambos

sistemas demuestran una preferencia por el género masculino en términos no marcados en la lengua origen, lo que supondría un lastre en la posesición de textos con actores femeninos y es, además, un sesgo importante de los sistemas.

La creación de la base de datos terminológica ha permitido comprobar los términos establecidos que los sistemas ya conocían, como han sido el Consejo de Europa y la propia denominación del Tribunal, sin revelar un desempeño ventajoso de un sistema sobre el otro. Cabe comentar que los términos en los que no ha afectado el uso de la terminología son los nombres de dos instituciones europeas ampliamente reconocidas y citadas en textos de diferentes ámbitos, por lo que se entiende que sean los que el sistema reconoce. Es destacable también, que ambos sistemas han interpretado las siglas del tribunal en inglés, ECHR, que comparte siglas con el Convenio en inglés, como este último y las han traducido por CEDH (Convenio Europeo de Derechos Humanos).

Se puede cuestionar si el uso de terminología ha servido para avanzar en la calidad de las traducciones o si, por el contrario, los problemas que ha creado con el género de los artículos anteriores en el caso de la dicotomía *Convenio-Convención* es más un lastre en la posesición que una ayuda. En el caso de Google Translate en el que se puede ir corrigiendo a medida que traduce el sistema parece obvio que el uso de la terminología supondría un avance, ya que se pueden detectar y corregir errores según aparecen. Su uso en Amazon Translate podría dejarse a elección del traductor encargado dadas sus preferencias de posesición.

En relación con la terminología jurídica, se destacan aspectos ventajosos de Amazon Translate, que ha tenido un mejor desempeño en terminología y fraseología jurídica, como demuestra el uso de *cuestiones de fondo* en el caso de *merits* y del sustantivo *indemnización* o el verbo *conceder* en el caso de *award*. La mayoría de errores relacionados con la terminología jurídica se detectan en terminología polisémica, en casos en los que los sistemas han elegido la traducción del lenguaje común y no del término con connotación jurídica. Los sistemas han demostrado mejores actuaciones en relación con términos jurídicos inequívocos (*respondent State – Estado demandado*).

Sin embargo, cabe destacar que no se trata de textos cargados de gran terminología jurídica y que, además, por las características del TEDH como organismo internacional independiente de los países miembros con jurisprudencia propia, las sentencias siguen un esquema y fraseología propias del Tribunal. Por tanto, no se ha tratado la evaluación de traducciones hacia el ordenamiento jurídico español de forma completa, dado que se trata de un formato de sentencia distinto para el que no existe un consenso claro en la traducción

de algunos términos, lo que permite una mejor ejecución por parte de los sistemas de traducción automática.

4.4. Evaluación con métricas cuantitativas

La métrica a utilizar para la evaluación cuantitativa de los resultados de las traducciones del corpus es BLEU (Papineni et al., 2002), un método de evaluación de la traducción automática en comparación con traducciones realizadas por humanos. Estas traducciones de referencia se han obtenido de los trabajos de fin de máster citados anteriormente. Cabe destacar que, aunque el alumnado está versado en materias jurídicas, no se trata de traductores profesionales, por lo que algunos aspectos ortotipográficos y sintácticos comentados anteriormente tampoco se recogen en los textos de referencia.

BLEU es una de las métricas más utilizadas en la evaluación de la calidad de una traducción por su sencillez de uso y por haberse probado en varias ocasiones que tiene una buena correlación con los juicios humanos de calidad (Papineni et al., 2002). Las puntuaciones obtenidas con BLEU se interpretan generalmente como traducciones comprensibles a partir de una puntuación de 30 y traducciones buenas y fluidas a partir de 50 (Lavie, 2010; Zhang, Vogel, & Waibel, 2004).

Para la evaluación con la métrica BLEU en este estudio se ha utilizado la herramienta de evaluación online interactiva del sistema de traducción automática TILDE⁸, la cual calcula la puntuación BLEU de documentos que el usuario puede subir en formato “txt”. Existe la posibilidad de comparar a la vez hasta dos traducciones obtenidas con sistemas de traducción automática con una traducción de referencia. En este estudio, las traducciones de referencia son las traducciones realizadas anteriormente por alumnos de la UNED, accesibles en la web del Ministerio de Justicia⁹. Existe también la posibilidad de cargar el texto origen para que el sistema lo tenga de referencia. Una vez realizada la puntuación, se proporciona una interfaz gráfica interactiva de la puntuación por segmento y la posibilidad de descargar este desglose en formato “csv”.

4.4.1. Comparación de los resultados

La evaluación de los resultados se ha llevado a cabo con la referencia del texto original y la traducción del alumnado de la UNED tanto para las traducciones del sistema con el uso de

⁸ <https://www.letsmt.eu/Bleu.aspx>

⁹ <https://www.mjusticia.gob.es/cs/Satellite/Portal/es/areas-tematicas/area-internacional/tribunal-europeo-derechos>

la base de datos terminológica, como sin ella. A continuación, se presentan las puntuaciones obtenidas en cada uno de los textos y sistemas:

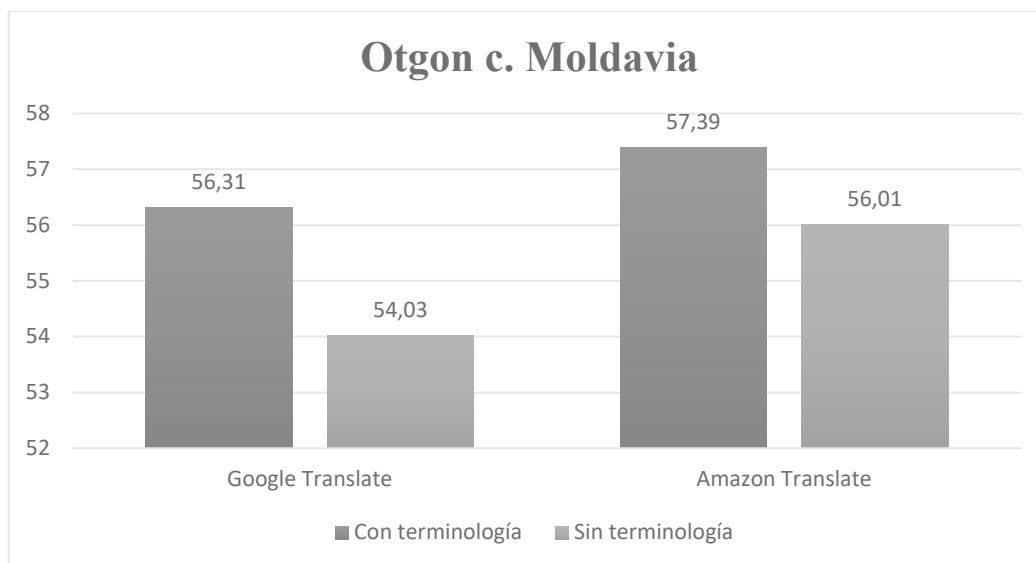


Figura 8: Evaluación BLEU del asunto *Otgon c. la República de Moldavia*

En el asunto de *Otgon c. la República de Moldavia*, el texto más corto del corpus, se aprecia resultado ligeramente mejor de Amazon Translate. La introducción de la terminología supone una mejora de 1'38 entre la versión sin terminología y la versión con terminología en Amazon Translate, en oposición a una mejora de 2'28 en Google Translate. Asimismo, la traducción sin terminología con Amazon Translate se sitúa muy cerca, solo a 0'3 puntos, de la traducción con terminología de Google Translate.

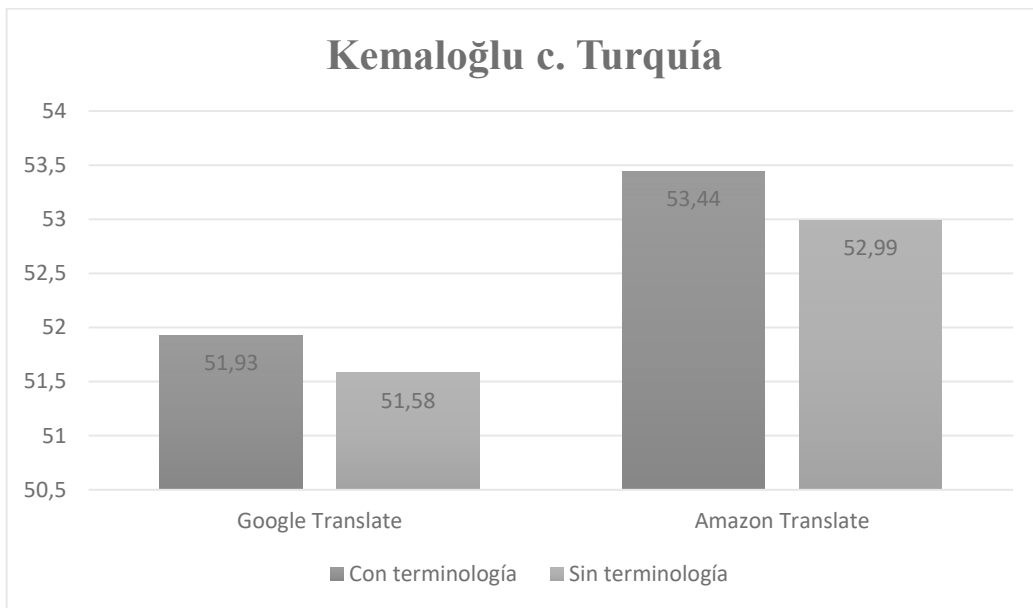


Figura 9: Evaluación BLEU del asunto Kemaloğlu c. Turquía

En el asunto de Kemaloğlu c. Turquía también se observa un mejor desempeño por parte de Amazon Translate, esta vez con diferencias menores entre las versiones con terminología y sin terminología en ambos sistemas, 0'45 en Amazon Translate y 0'35 Google Translate. Amazon Translate se sitúa por encima también en ambas versiones de los resultados de Google.

Los resultados de los dos textos anteriores revelan unas buenas traducciones de los sistemas, comprensibles y fluidas, ya que los resultados se sitúan por encima de 50.

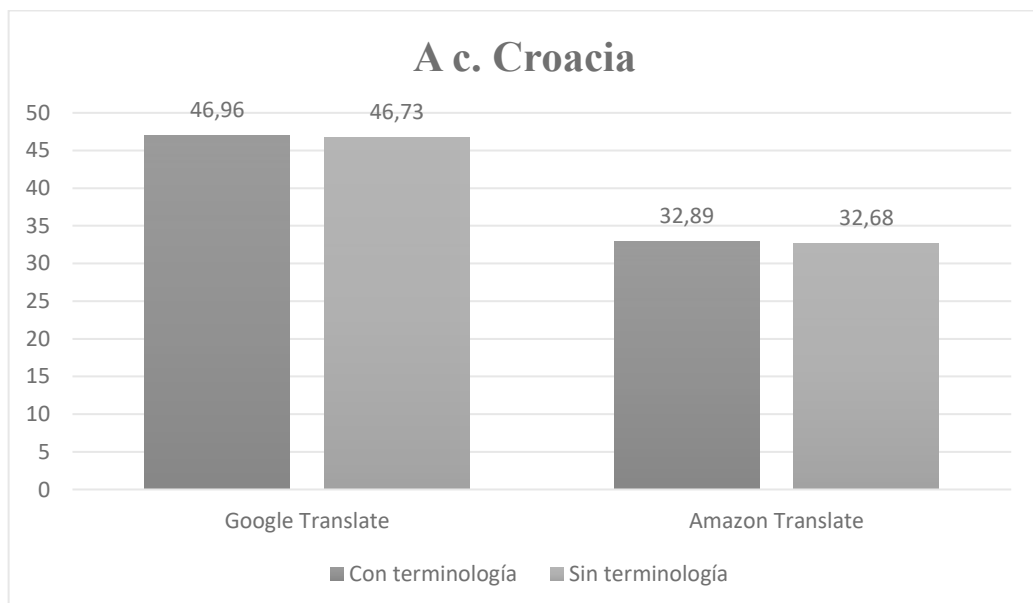


Figura 10: Evaluación BLEU del asunto A c. Croacia

En el asunto de A c. Croacia encontramos un giro en los resultados, siendo Google Translate el sistema con mejor desempeño de los dos, pero sin superar ningún texto el umbral de los 50 puntos. La inserción de terminología tiene un efecto irrisorio en ambos, con solo una mejora del 0'23 en Google Translate y de 0'21 en Amazon Translate. También tiene que ver que se trata de un texto mucho más largo que sus predecesores, por lo que las mejoras que se introducen con la terminología son relativamente menores que en un texto más corto. Amazon Translate se presenta con una puntuación al borde de la comprensibilidad determinada con el límite de 30 puntos. El análisis de errores realizado en el apartado anterior revelaba una mayor presencia de sinsentidos, adiciones y ausencias en este texto que en los dos anteriores.

2.4.2. Comentario cuantitativo de los resultados

Como se ha hecho anteriormente, se procede a comentar desde una perspectiva global los resultados cuantitativos del estudio. La percepción anterior de mejor desempeño de Amazon Translate se confirma con la métrica utilizada, siendo sus traducciones más similares a la traducción humana de referencia, excepto en el caso de *A c. Croatia*. Este último, puede que por la confusión que crea en el sistema el uso de letras para identificar a los actores y las ciudades y/o por la mayor longitud del texto, es el único en el que se producen unos resultados comprensibles, pero no de calidad publicable. Se requeriría aquí una mayor posesición y habría que cuestionar si el tiempo y trabajo invertido en la posesición compensa y justifica el uso de los sistemas.

Como ya se anticipaba también, se comprueba que el efecto del uso de la terminología es relativamente pequeño y menor a medida que aumenta la longitud del texto, ya que cada palabra individual pierde peso en el cómputo total. Sin embargo, se considera un acierto su introducción y la comparación de ambos resultados, teniendo en cuenta que los términos eran limitados pero que la introducción de terminología menos específica crearía mayores problemas en la traducción que soluciones.

Cabe destacar, una vez más, que las traducciones de referencia no han sido realizadas por traductores profesionales y también contienen inadecuaciones ortotipográficas y/o de referencias, por lo que aciertos del sistema que sean erróneos en la traducción de referencia se considerarán errores del sistema de traducción y comportarán una menor puntuación (como el caso de la presidenta Françoise Tulkens).

5. Conclusiones

5.1. Validación de las hipótesis

Se recogen a continuación las dos hipótesis propuestas anteriormente para comentarlas en relación con los resultados obtenidos. La hipótesis general recogía la idea de que «la traducción automática es útil para la traducción jurídica de sentencias del TEDH de inglés a español». A pesar de que determinar la utilidad de una herramienta no es una medida cuantitativa que permita analizar el estudio como un éxito al superar un cierto número y un fracaso por no superarlo, el análisis cualitativo y cuantitativo de la calidad de las traducciones obtenidas permite afirmar que los textos producidos por los dos sistemas estudiados son de calidad suficiente para permitir una posesición humana que no suponga una carga mayor que la traducción humana entera y así obtener una versión de los textos publicable. Con la excepción del último texto, en el que sí se encuentran sinsentidos y errores de mayor envergadura, una revisión ortotipográfica, de terminología jurídica y de detección de errores serían suficientes para pulir los textos obtenidos.

Asimismo, como demuestran estudios anteriores, existe la posibilidad de que un uso reiterado de los sistemas para este propósito permita mejoras en los resultados al producirse un entrenamiento de los sistemas de traducción automática (Gotti et al., 2008; Sennrich, Haddow, & Birch, 2015; Wu et al., 2016). Dado que el número de textos utilizados en este estudio no es significativo, no se ha podido detectar dicha mejora, pero no es una posibilidad que descartar en el futuro.

De las características y tiempo en el mercado de ambas herramientas de traducción automática, se desprendía también la hipótesis secundaria de que «Google Translate produce traducciones de mejor calidad que Amazon Translate para la traducción de sentencias del TEDH de inglés a español». Google Translate es probablemente la API de traducción más conocida del mundo, por lo que cuenta con millones de ejemplos de textos a los que han contribuido los usuarios durante años para aprender y mejorar las conexiones de las redes neuronales. Amazon, aunque también es un sistema por redes neuronales, es de creación más reciente y menos conocido. Google Translate se puede usar en su versión web gratuitamente y sin registro, mientras que Amazon Translate solo es accesible bajo registro en Amazon Web Services. Por esta trayectoria más larga y el mayor número de usuarios, se suponía un mejor desempeño por parte de Google Translate, hipótesis que no ha sido verificada en el estudio. Amazon Translate ha producido mejores resultados que

Google en dos de los tres textos y ha tenido un mejor desempeño en la traducción de terminología jurídica ambigua.

Cabe destacar, sin embargo, que la mejora es limitada y que el entorno de trabajo con Google Translate tal y como se ha presentado en el estudio es más adecuado para permitir una posesición fluida y una maquetación del texto en el formato propio del TEDH. Por ello, no se puede concluir que para los objetivos del trabajo que nos ocupa, Amazon Translate sea proporcionalmente superior a Google Translate.

5.2. Consideraciones éticas

Del apartado anterior se discierne un aspecto que se ha creído necesario comentar: la alimentación de estos sistemas a través de los textos insertados por los usuarios. Los sistemas de traducción automática por redes neuronales aprenden del corpus de textos del que disponen para realizar conexiones entre ellos, por lo que un mayor corpus permitirá un mayor avance en la calidad del sistema. Se presuponía que Google Translate disponía de un mayor corpus debido a su larga trayectoria primero como sistema basado en reglas, posteriormente como sistema estadístico y desde 2016 como sistema por redes neuronales. La versión online gratis de Google Translate es usada diariamente por millones de usuarios, mientras que la API está en uso en aplicaciones de otras empresas con sus propios millones de usuarios, como es el caso de la plataforma de alquileres vacacionales Airbnb.

Cabe, por tanto, plantearse el uso que Google hace de toda la información y textos que obtiene a través de su API de traducción y hasta qué punto la privacidad de los usuarios está protegida por la plataforma si se implementara este estudio en las futuras traducciones del TEDH. Como informa memoQ a sus usuarios (memoQ, s. f.), existe preocupación sobre la posibilidad de que estos textos insertados por los usuarios no estén protegidos y sean accesibles al público en general. A pesar de que los textos del Tribunal ya son accesibles públicamente desde su base de datos, la introducción en el sistema de Google por parte de los traductores supondría una diseminación mayor de su contenido, por lo que habrían de considerarse las implicaciones éticas que eso conlleva.

El caso de Amazon es menos conocido y no accesible sin registro, como hemos comentado, pero se trata de una situación similar en el sentido de que también es una empresa multimillonaria haciendo uso de datos de los usuarios para mejorar su sistema.

5.3. Limitaciones

Como se ha comentado desde el principio del estudio, el objetivo perseguido era responder a una situación y necesidad real de los alumnos del Máster Universitario en Unión Europea de la UNED y del convenio alcanzado entre los organismos citados anteriormente, por lo que la validez externa no ha sido el objetivo de la investigación. Como tal, se encuentran escollos a la generalización del experimento más allá del TEDH, dado que las características del Tribunal como organismo jurídico propio le otorgan un carácter jurídico propio, así como a la generalización del uso de las herramientas, por las limitaciones en el uso del entorno JDK de Amazon Translate, que podría haber mejorado la usabilidad, la posesición y el maquetado de la traducción con Amazon Translate.

Asimismo, el tamaño del corpus es limitado y no se considera extrapolable como lo sería un corpus mayor. Sí se han introducido textos de distintas características para poner a prueba los sistemas en distintas circunstancias, con lo que se han mostrado diferentes desempeños por la longitud del texto o por la introducción de elementos confusos para los sistemas como las abreviaturas para proteger identidades, elemento común en las sentencias judiciales.

Por último, cabe considerar que la evaluación cuantitativa del estudio se ha llevado a cabo con BLEU, una métrica ni mucho menos perfecta. Como se ha comentado durante años y como se ha observado también en el estudio, las puntuaciones de BLEU son relativas a la calidad de la traducción de referencia, que puede contener errores. Además, la traducción humana, como ciencia no exacta, puede contener gran variabilidad dentro de un rango de traducciones consideradas aceptables, por lo que la comparación con otra traducción de referencia podría producir resultados muy diferentes a los obtenidos aquí. Por otra parte, BLEU no contiene herramientas para identificar y considerar los sinónimos, por lo que todo aquello que no sea una identificación exacta con el texto de referencia no se considera adecuado (Vashee, 2019).

5.4. Futuras líneas de investigación

En el marco de un trabajo como este se abre el área de investigación futura a la traducción automática de textos jurídicos. Por una parte, una vía de investigación derivada del uso de sistemas de traducción enfocados a la traducción jurídica o entrenados para ese propósito, ya que los sistemas empleados en este estudio eran de fácil acceso y uso, pero enfocados a la traducción general.

Por otra parte, debido a las limitaciones comentadas anteriormente, es inevitable desear explorar un mayor corpus para observar las fluctuaciones en la comparación de los sistemas. Se podría también añadir la combinación francés-español, ya que algunas sentencias del TEDH están redactadas en francés, y así obtener una comparativa sobre la traducción del inglés al español y la del francés al español. Asimismo, el uso de otras métricas de evaluación de la traducción automática, como NIST, METEOR o TER, permitiría una visión más amplia de desde distintas perspectivas de la calidad de traducción, dadas las limitaciones de BLEU comentadas anteriormente.

6. Bibliografía

- Albí, B., & Belén, A. (2003). La investigación en traducción jurídica. En *Panorama actual de la investigación en traducción e interpretación*. Granada: Atrio.
- Albir, A. H. (2001). *Traducción y traductología*. Madrid: Cátedra.
- Alcaraz, E., & Hughes, B. (2002). *El español jurídico*. Barcelona: Ariel.
- Alcaraz, E., & Hughes, B. (2014). *Legal Translation Explained*. Abingdon-on-Thames: Routledge.
- Aletras, N., Tsarapatsanis, D., Preoțiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. *PeerJ Computer Science*, 2, e93. <https://doi.org/10.7717/peerj-cs.93>
- Arnold, D., Balkan, L., Meijer, S., Humphreys, R., & Sadler, L. (1994). *Machine Translation: An Introductory Guide*. Oxford: NCC Blackwell.
- Arribas Gómez, D. (2018). *Traducción y comentario de la sentencia A c. Croacia del Tribunal Europeo de Derechos Humanos*. UNED, Madrid.
- Bentivogli, L., Bisazza, A., Cettolo, M., & Federico, M. (2016). Neural versus phrase-based machine translation quality: A case study. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 257–267. <https://doi.org/10.18653/v1/D16-1>
- Bermúdez Bausela, M. (2014). Aplicación de las herramientas de traducción asistida por ordenador (TAO) a la investigación en traducción. *Estudios De Traducción*, 4, 129–143. <https://doi.org/10.5209/rev ESTR.2014.v4.45372>
- Brannan, J. (2016, abril 8). *Translation at the European Court of Human Rights*. Recuperado de https://eulita.eu/wp/wp-content/uploads/files/Strasbourg%202016_Brannan.ppt

- Champollion, Y. (2003). Convergence in CAT: Blending MT, TM, OCR & SR to boost productivity. *Translating and the Computer*, 15–15.
- Consejo de Europa. (1950). *Convenio Europeo para la Protección de los Derechos Humanos y de las Libertades Fundamentales*. Recuperado de https://www.echr.coe.int/Documents/Convention_SPA.pdf
- Cornu, G. (1990). *Linguistique juridique*. Paris: Montchrestien.
- Derlén, M. (2015). A single text or a single meaning: Multilingual interpretation of EU legislation and CJEU case law in national courts. *Language and culture in EU law: multidisciplinary perspectives*, 53–72.
- Dzehtsiarou, K. (2010). Comparative Law in the Reasoning of the European Court of Human Rights. *University College Dublin Law Review*, 10, 109-140.
- Federico, M., Cattelan, A., & Trombetti, M. (2012). Measuring user productivity in machine translation enhanced computer assisted translation. *Proceedings of the Tenth Conference of the Association for Machine Translation in the Americas (AMTA)*, 44–56. AMTA Madison, WI.
- Frame, I. (2005). Linguistic oddities in European Union legislation: Don't shoot the translator. *Clarity—Journal of the International Association Promoting Plain Legal Language*, 53, 21–24.
- Fundación Acción Pro Derechos Humanos. (s. f.). Tribunal Europeo de Derechos Humanos (TEDH). Recuperado 20 de noviembre de 2018, de [DerechosHumanos.net](https://www.derechoshumanos.net) website: <https://www.derechoshumanos.net/tribunales/TribunalEuropeoDerechosHumanos-TEDH.htm>
- García Toledano, L. P. (2018). *Traducción y comentario de la sentencia A c. Croacia del Tribunal Europeo de Derechos Humanos*. UNED, Madrid.

- Gazzola, M., & Grin, F. (2013). Is ELF more effective and fair than translation? An evaluation of the EU's multilingual regime. *International Journal of Applied Linguistics*, 23(1), 93–107.
- Giráldez Ceballos-Escalera, J. (2007). *Las colocaciones léxicas en el lenguaje jurídico del derecho civil francés* (PhD Thesis). Universidad Complutense de Madrid, Madrid.
- Gotti, F., Farzindar, A., Lapalme, G., & Macklovitch, E. (2008). Automatic translation of court judgments. *Proceedings of the Eighth Conference of the Association for Machine Translation in the Americas*, 370–379. Waikiki, Hawaii, USA.
- Guerberof, A. (2009). Productivity and quality in MT post-editing. *MT Summit XII-Workshop: Beyond Translation Memories: New Tools for Translators MT*. Presentado en Ottawa, Ontario, Canada. Ottawa, Ontario, Canada.
- Harvey, M. (2002). What's so special about legal translation? *Meta: Translators' Journal*, 47(2), 177–185.
- Hutchins, J. (2001). Machine translation and human translation: In competition or in complementation. *International Journal of Translation*, 13(1-2), 5–20.
- Hutchins, J. (2007). Machine translation: A concise history. *Journal of Translation Studies*, 13, 29–70.
- Jacobs, F. (2003, octubre 20). *How to interpret legislation which is equally authentic in twenty language*. Recuperado de http://ec.europa.eu/dgs/legal_service/seminars/agjacobs_summary.pdf
- Kit, C., & Wong, T. M. (2008). Comparative evaluation of online machine translation systems with legal texts. *Law Library Journal*, 100, 299.
- Kjær, A. L. (2015). Theoretical aspects of legal translation in the EU: The paradoxical relationship between language, translation and the autonomy of EU law. *Language and Culture in EU Law: Multidisciplinary Perspectives*, 91–107.

- Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. *Proceedings of the First Workshop on Neural Machine Translation*, 28-39. Vancouver, BC, Canada.
- König, F. (1998). Juristisch notwendige Einschränkungen bei der Übersetzung von internationalen Rechtstexten am Beispiel der Europäischen Menschenrechtskonvention. *Traduction de textes juridiques: problèmes et méthodes*.
- Lavie, A. (2010). Evaluating the output of machine translation systems. *AMTA Tutorial*, 86.
- Lawlor, R. C. (1963). What computers can do: Analysis and prediction of judicial decisions. *American Bar Association Journal*, 337–344.
- Martín Caño, E. M. (2018). *Traducción y comentario de la Sentencia İlbeyi Kemaloğlu y Meriye Kemaloğlu c. Turquía del Tribunal Europeo de Derechos Humanos*. UNED, Madrid.
- Mattila, H. E. (2016). *Comparative Legal Linguistics: Language of Law, Latin and Modern Lingua Francas*. Abingdon-on-Thames: Routledge.
- Melby, A. K. (1998). Eight types of translation technology. *American Translators Association*, 4–9.
- memoQ. (s. f.). What is Google MT? Recuperado 13 de julio de 2019, de <https://memoq.com/en/integration-with-machine-translation/google-mt>
- Moreno Fernández, F. (1999). Lenguas de especialidad y variación lingüística. En *Lenguas para fines específicos (VI). Investigación y enseñanza* (pp. 3–14). Alcalá de Henares.: Universidad de Alcalá.
- Nolla, F. C., & Abril, Á. P. (2017). Traducció automàtica neuronal. *Revista Tradumàtica: tecnologies de la traducció*, (15), 66–74.

- Orozco Jutorán, M. (2001). Métodos de investigación en traducción escrita: ¿qué nos ofrece el método científico? *Sendebarr*, 12, 95–115.
- Orts Llopis, M. Á. (2015). El sistema legal inglés y su hermenéutica: La importancia del lenguaje en el derecho anglosajón. *Revista de Lenguas para Fines Específicos*, 8, 259–280.
- Ostrowska, J. M. (2018). *Traducción y comentario de la sentencia Otgon c. La República de Moldavia del Tribunal Europeo de Derechos Humanos*. UNED, Madrid.
- Pacho Aljanati, L. (2015). *The court of justice of the European Union's case law on linguistic divergences (2007-2013): Interpretation criteria and implications for the translation of EU legislation*. Barcelona: Universitat Autònoma de Barcelona.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: A method for automatic evaluation of machine translation. *Proceedings of the 40th annual meeting on association for computational linguistics*, 311–318. Philadelphia, PA, USA: Association for Computational Linguistics.
- Parra Escartín, C. (2018, abril). ¿Cómo ha evolucionado la traducción automática en los últimos años? Recuperado 11 de marzo de 2019, de La Linterna del Traductor website: <http://www.lalinternadeltraductor.org/n16/traduccion-automatica.html>
- Parra Escartín, C., & Arcedillo, M. (2015). Living on the edge: Productivity gain thresholds in machine translation evaluation metrics. *4th Workshop on Post-Editing Technology and Practice (WPTP4)*, 46. Miami, Florida, USA.
- Piris, J.-C. (2005). The legal orders of the European Community and of the Member States: Peculiarities and influences in drafting. *Amicus Curiae*, 2005(58), 21–28.
- Plitt, M., & Masselot, F. (2010). A productivity test of statistical machine translation post-editing in a typical localisation context. *The Prague bulletin of mathematical linguistics*, 93, 7–16.

- Policastro Ponce, G. (2015). El derecho comparado y la terminología jurídica en traducción. *Skopos. Revista Internacional de Traducción e Interpretación*, 6, 155–172.
- Prieto Ramos, F. (2014). International and supranational law in translation: From multilingual lawmaking to adjudication. *The Translator*, 20(3), 313–331.
- Pym, A. (1999). Why Translation Studies should learn to be homeless. *Tradução e Multidisciplinaridade*, 35–51.
- Real Academia Española. (2005). DPD 1.^a edición, 2.^a tirada. Recuperado 28 de junio de 2019, de <http://buscon.rae.es/dpd>
- Real Academia Española. (2016). Diccionario del Español Jurídico. Recuperado 28 de junio de 2019, de <https://dej.rae.es/>
- Ruellan, S. (2010). La especificidad del léxico jurídico: Un obstáculo para la traducción. *Epos: Revista de filología*, (26), 421.
- Ruiloba Alvariño, J. (2006). El Tribunal Europeo de Derechos Humanos: Organización y funcionamiento. *Anuario de la Escuela de Práctica Jurídica*, (1).
- Sammut, I. (2018). The effect of Multilingualism in the EU on European Legal Translation. *International Journal for the Rule of Law*, 30-52.
- Sánchez Castany, R., & Díaz Pore, S. (2011). La traducció per al sistema de les Nacions Unides: Propostes tecnològiques. *Fòrum de Recerca*, (16), 615–628.
- Sennrich, R., Haddow, B., & Birch, A. (2015). Improving neural machine translation models with monolingual data. *arXiv preprint arXiv:1511.06709*.
- Sureda, A. (s. f.). *Seminario Traducción Jurídica*. Recuperado de <http://conf-dts1.unog.ch/1%20SPA/Tradutek/Recursos%20Juridicos/Curso%20Sureda%202007.htm>

- TEDH. (2018, diciembre 31). *Guide on Article 8 of the European Convention on Human Rights*. Recuperado de https://www.echr.coe.int/Documents/Guide_Art_8_ENG.pdf
- Tribunal Europeo de Derechos Humanos. (2014). *Guía Práctica sobre la Admisibilidad*. Recuperado de https://www.echr.coe.int/Documents/Admissibility_guide_SPA.pdf
- Unidad de Relaciones Públicas del Tribunal. (2012). *El TEDH en 50 preguntas*. Recuperado de https://www.echr.coe.int/Documents/50Questions_SPA.pdf
- Vashee, K. (2019, abril 12). Understanding MT Quality: BLEU Scores. Recuperado 13 de julio de 2019, de SDL website: <https://www.sdl.com/blog/understanding-mt-quality-bleu-scores.html>
- White, R. C. A. (2009, julio 17). *Judgments in the Strasbourg Court: Some Reflections*. <https://doi.org/10.2139/ssrn.1435197>
- Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... Macherey, K. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.
- Zampieri, M., & Vela, M. (2014). Quantifying the influence of MT output in the translators' performance: A case study in technical translation. *Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation*, 93–98. <https://doi.org/10.3115/v1/W14-03>
- Zhang, Y., Vogel, S., & Waibel, A. (2004). Interpreting BLEU/NIST scores: How much improvement do we need to have a better system? *LREC*.
- Zupancic, B. M. (2016, noviembre 17). In the Context of the Common Law: The European Court of Human Rights in Strasbourg. Recuperado 18 de noviembre de 2018, de <https://www.gresham.ac.uk/lectures-and-events/in-the-context-of-the-common-law-the-european-court-of-human-rights-in-strasbourg>

7. Apéndices

Apéndice 1

Base de datos terminológica

en	es
Convention (European Convention of Human Rights)	Convenio (Convenio Europeo de Derechos Humanos)
European Court of Human Rights	Tribunal Europeo de Derechos Humanos
ECHR	TEDH
Council of Europe	Consejo de Europa
Grand Chamber	Gran Sala
GC	GS
Deputy Section Registrar	secretario adjunto de sección
Section Registrar	secretario de sección

Apéndice 2

Traducciones de los textos del corpus

Se puede acceder al corpus utilizado en este trabajo con las traducciones obtenidas por los sistemas, las traducciones humanas de referencia y los textos originales segmentados y en formato “txt” a través del siguiente enlace: <http://bit.ly/2XUoLDW>.