

Una propuesta de análisis comparativo de traducciones humanas y automáticas de textos especializados: implicaciones para la evaluación

A Proposal of Comparative Analysis of Human and Automatic Translations of Specialized Texts: Implications for Evaluation

Marina Fomicheva

Universidad Nacional Autónoma de México
Centro de Enseñanza de Lenguas Extranjeras
Ciudad Universitaria, 04510, México D.F.
mari.fomicheva@gmail.com

Iria da Cunha

Universitat Pompeu Fabra
Institut Universitari de Lingüística Aplicada
C/ Roc Boronat 138, 08018, Barcelona
iria.dacunha@upf.edu

Gerardo Sierra

Universidad Nacional Autónoma de México
Instituto de Ingeniería, Torre de Ingeniería, basamento,
Ciudad Universitaria, 04510, México D.F.
GSierraM@iingen.unam.mx

Resumen: El objetivo de este artículo es ofrecer una metodología para el análisis comparativo de traducciones automáticas (TAs) y traducciones humanas (THs). Para ello conformamos un corpus paralelo inglés-español de textos especializados del ámbito médico, comparamos la distribución de ciertas unidades lingüísticas (unidades terminológicas, n-gramas de etiquetas POS y relaciones discursivas) en THs y TAs, e identificamos las condiciones en las que se producen las diferencias. Los resultados del estudio muestran que la metodología propuesta resulta útil para detectar las diferencias entre THs y TAs relevantes para la evaluación u optimización de sistemas de TA.

Palabras clave: traducción automática, evaluación, comparación, textos especializados

Abstract: The goal of this study is to offer a methodology for the comparative analysis of machine translations (MTs) and human translations (HTs). We build an English-Spanish parallel corpus of specialized texts from the medical domain, compare the distribution of certain linguistic units (terminological units, POS n-grams and discourse relations) in HTs and MTs, and identify the conditions in which these differences take place. The results of the study show that the methodology is useful to detect differences between HTs and MTs that could be relevant when evaluating or developing MT systems.

Keywords: Machine Translation, Evaluation, Comparison, Specialized Texts

1 Introducción¹

La Traducción Automática (TA) difiere mucho de la Traducción Humana (TH) en términos de

calidad. En este contexto la evaluación tiene un papel estratégico, ya que permite detectar carencias, establecer prioridades y guiar el desarrollo de sistemas de TA. El supuesto de base de la evaluación automática es que “machine translation might be considered a solved problem should it ever become impossible to distinguish automated output from human translation” (Corston-Oliver et al., 2001: 140). Sin embargo, las métricas más utilizadas en la actualidad que calculan la

¹ Este trabajo ha sido parcialmente financiado por los proyectos RICOTERM (FFI2010-21365-C03-01) y APLE (FFI2009-12188-C05-01), por un contrato Juan de la Cierva del Ministerio de Ciencia e Innovación de España (JCI-2011-09665) y por una beca de maestría del CONACYT de México.

similitud TA-TH en términos de coocurrencia de n-gramas (para una revisión detallada, véase Giménez, 2008) no proporcionan información sobre los rasgos distintivos de la TA. Al realizar una comparación directa partiendo de una representación superficial, estas métricas penalizan de la misma manera cualquier diferencia entre la TA y la TH de referencia, sin hacer una distinción entre la variación aceptable y las divergencias que realmente afectan la calidad de la traducción.

Las diferencias TA-TH pueden estar relacionadas tanto con las particularidades de la TA, como con las decisiones del traductor humano, condicionadas no solamente por las diferencias sistémicas entre las lenguas, sino también por el propio proceso de traducción (Baker, 1995) y por las restricciones provenientes de las convenciones de uso de los recursos de la lengua meta en un contexto de situación determinado.

El objetivo de este trabajo es ofrecer una metodología para el análisis comparativo de la TA y la TH, teniendo en cuenta diversos niveles lingüísticos (léxico-terminológico, sintáctico y discursivo). Para ello conformamos un corpus paralelo inglés-español de textos especializados del ámbito médico, y comparamos la distribución de ciertas unidades lingüísticas (unidades terminológicas, n-gramas de etiquetas POS y relaciones discursivas) en las TAs y en las THs realizadas por dos sistemas basados en estrategias diferentes.

2 Trabajo relacionado

Uno de los problemas fundamentales de las métricas basadas en n-gramas es la falta de poder descriptivo, debido a que la comparación se realiza a un nivel concreto y superficial. Para solucionar este problema, Amigó et al. (2006) proponen calcular la similitud entre la TA y la TH de referencia a partir de un análisis más fino combinando varias métricas relacionadas con distintos niveles de la lengua en un marco integrador que proporcione una evaluación fundamentada lingüísticamente.

Otra propuesta enfocada a los rasgos distintivos de las TAs se basa en la extracción de patrones lingüísticos a partir de un corpus grande de TAs y THs. Gamon et al. (2005) utilizan el algoritmo de clasificación textual *Support Vector Machines* (SVM) entrenado

para distinguir entre THs y TAs. La calidad de una traducción se define en función de la probabilidad de su pertenencia a la clase de THs o TAs. La clasificación con SVM da muy buenos resultados, pero no ofrece información sobre la naturaleza de los rasgos lingüísticos pertinentes para la comparación de las versiones traducidas.

Al investigar las diferencias entre la TA y la TH se deben tomar en cuenta tanto la naturaleza de los textos producidos por sistemas de TA, como los rasgos prototípicos de la TH. Estos han sido estudiados en el marco de los estudios de traducción basados en corpus (Baker, 1995), que han demostrado que los textos traducidos (TTs) presentan una distribución de unidades (léxicas, sintácticas o discursivas) sistemática y significativamente distinta con respecto a los textos escritos originalmente en la lengua de llegada (en este sentido se reconoce la existencia del "lenguaje de traducción" o *translationese*).

Al describir las propiedades de los TTs, es necesario determinar cuáles son las posturas regularmente adoptadas por los traductores en un contexto determinado y qué factores condicionan sus decisiones (Tourey, 2004). Una manera de abordar esta tarea es por medio de la noción de *translation shifts*² ("desviaciones del texto original [TO] en la traducción"). El interés por este fenómeno se debe a la contradicción entre la expectativa prototípica de la similitud máxima entre el TT y el TO, y la expectativa de la "naturalidad" del TT (Szymańska, 2011). De acuerdo con Fernández Polo (1999), para describir y explicar las características de los TTs deben "identificarse las 'desviaciones' manifestadas en la traducción con respecto a un ideal de 'adecuación' determinado por el propio texto original (TO), que actúa como 'invariante de comparación'. Descubiertos los lugares en los que la traducción se desvía del original, podrá comprobarse que algunas de estas desviaciones son obligatorias, motivadas por las diferencias sistémicas entre las dos lenguas, mientras que otras son opcionales, resultado de la elección consciente o inconsciente del traductor" (Fernández Polo, 1999: 43-44). Las diferencias TA-TH relacionadas con los *translation shifts*

² Utilizamos el término en inglés, ya que no existe un consenso sobre su equivalente en español.

en las THs y las diferencias vinculadas a los errores de la TA no tienen el mismo impacto en la calidad de la traducción.

3 Metodología

En relación con el estudio de las diferencias TA-TH en el contexto de la evaluación, partimos de las siguientes consideraciones generales. En primer lugar, siguiendo a Toury (2004), consideramos que, para identificar las regularidades en el comportamiento de los traductores humanos, los TTs objeto de estudio deben compartir un contexto de situación específico; por tanto, para llegar a conclusiones válidas sobre las características de las traducciones debe partirse de un corpus homogéneo en términos de género y tipo textual. En segundo lugar, el punto de partida para el estudio de los rasgos distintivos del lenguaje de la traducción son los TTs, con lo cual se identifica primero un patrón lingüístico (distribución específica de unidades o fenómenos lingüísticos) en la lengua meta y, a continuación, se investiga con qué tipo de contextos de los TOs se relaciona de manera sistemática. En tercer lugar, la calidad de la traducción como objeto de evaluación tiene varios aspectos relacionados con los niveles de la lengua, y no existe una manera trivial de ponderar dichos aspectos en términos de su efecto en la calidad global de los TTs. Por tanto, en un primer acercamiento al análisis cualitativo de las diferencias TA-TH, conviene considerar dichas diferencias con detalle en cada nivel por separado.

La metodología consta de tres fases. Primero, se identifican y se cuantifican las diferencias sistemáticas entre las THs y las TAs. Segundo, se detectan los contextos de los TOs con los que se relacionan dichas diferencias. Tercero, las diferencias observadas se clasifican de acuerdo con su origen. Las diferencias que se relacionan con la TA indican una selección equivocada de unidades del léxico o estructuras sintácticas por parte de los sistemas, la cual en algunos casos (que documentamos en el análisis en el nivel discursivo) afecta la coherencia global de del TT. Las diferencias relacionadas con la TH indican modificaciones opcionales realizadas por los traductores con respecto a la forma y el contenido del original. Elaboramos una propuesta de clasificación de dichas

modificaciones tomando como base los trabajos de Catford (1965), Leuven-Zwart (1989) y Cyrus (2006), la cual se resume en la Tabla 1.

Nivel léxico	<i>Especificación</i> (la unidad del léxico en la TH tiene un significado más específico en relación con la unidad correspondiente del TO)
	<i>Generalización</i> (la unidad del léxico en la TH tiene un significado más general en relación con la unidad correspondiente del TO)
	<i>Modificación</i> (las unidades del léxico de los TOs y las THs tienen una relación diferente a las que se indican arriba)
Nivel sintáctico	<i>Cambios relacionados con la función y propiedades de la cláusula</i> (voz, modo, cláusulas con verbos conjugados frente a cláusulas infinitivas, etc.)
	<i>Cambios relacionados con la función y la posición de los constituyentes</i> (adverbial vs. predicativo, traducción de una cláusula con una frase preposicional, etc.)
Nivel discursivo	<i>Explicitación</i> (la relación discursiva no está marcada en el TO, pero sí en el TT)
	<i>Implicitación</i> (la relación discursiva está marcada en el TO, pero no en el TT)
	<i>Modificación</i> (la relación discursiva que existe en el TO no se preserva en el TT)

Tabla 1: Clasificación de *translation shifts*

Aplicamos este procedimiento general para analizar un corpus de textos especializados del ámbito médico, concretamente del género artículo de divulgación científica. El corpus se compone de 40 textos: 10 artículos publicados en inglés en la revista *Scientific American* entre los años 1994 y 2000 (24.053 palabras), 10 traducciones de estos textos al español publicadas en la versión española de dicha revista (*Investigación y Ciencia*)³ (25.305 palabras), 10 TAs inglés-español realizadas por el sistema estadístico Google Translator [Google] (27.547 palabras), y 10 TAs inglés-español realizadas por el sistema basado en reglas Lucy Translator [Lucy] (26.206 palabras). Para este experimento tomamos las

³ La revista no aporta información sobre el número de traductores, pero sabemos que "Los traductores no forman parte, pues, del personal de la propia revista, sino que están adscritos a un centro de enseñanza o de investigación y son invitados a colaborar esporádicamente con *Investigación y Ciencia*" (Fernández Polo, 1999: 99).

traducciones de estos sistemas, ya que una de las preguntas que nos interesa discutir de cara a la comparación de las TAs con las THs es qué tipo de modificaciones realizadas por el humano son capaces de modelar los sistemas de TA basados en estrategias diferentes.

4 Análisis y resultados

4.1 Nivel léxico-terminológico

En el *nivel léxico* nos centramos en el tratamiento de la terminología por parte de los traductores humanos y sistemas de TA, ya que en textos de especialidad son los términos quienes vehiculan el conocimiento especializado (Cabré, 1999). Para extraer automáticamente las unidades terminológicas, empleamos la herramienta propuesta por Vivaldi y Rodríguez (2010). Este extractor obtiene términos de un dominio de especialidad utilizando las estructuras de páginas y categorías de la Wikipedia, y proporciona para cada candidato a término un coeficiente de dominio [CD] (de 0 a 1) que indica su grado de pertenencia al ámbito de especialidad (en nuestro caso "Biología" y "Medicina"). Nos limitamos a analizar los candidatos a términos con un CD mayor a 0.5. Al realizar la extracción, calculamos el número total de candidatos a términos y el número de candidatos a términos con CDs altos (de 0.8 a 1) en cada grupo de TTs. Esta información nos proporciona una caracterización general del tratamiento de la terminología en la TH y la TA. A continuación, detectamos las unidades léxicas correspondientes en los TOs. A partir de estos datos, detectamos los patrones que reflejan el tratamiento de la terminología en los TTs (por ejemplo, sustitución de una unidad que no pertenece al dominio de especialidad en el TO por una unidad terminológica en el TT). Finalmente, clasificamos las diferencias sistemáticas TA-TH de acuerdo con la tipología incluida en la Tabla 1.

Obtenemos los siguientes resultados cuantitativos generales. El número total de candidatos a términos extraídos en las THs es 567, en las TAs de Google 532 y en las TAs de Lucy 434. En las traducciones de Google se observa un 24% de diferencias, mientras que en las traducciones de Lucy el porcentaje es mayor, un 46%. En la Tabla 2 ejemplificamos las tendencias que se presentan en el

tratamiento de la terminología por parte de los traductores y los sistemas de TA.

Ej.	TO	TH (CD)	Google (CD)	Lucy (CD)
1	disorder	patología (0.7)	trastorno (0.2)	desorden (-1)
2	breed	progenie (1)	raza (0)	raza (0)
3	patient	organismo (0.8)	paciente (0.1)	paciente (0.1)
4	eating	deglución (0.5)	comer (N/A)	comiendo (N/A)
5	treating	tratamiento (1)	tratamiento (1)	tratar (N/A)
6	motor coordina- tion	coordina- ción motora (0.5)	coordina- ción motora (0.5)	coordina- ción de motor (N/A)
7	stroke	accidente cerebro- vascular (0.6)	carrera (0)	golpe (0)
8	stroke	accidente cerebro- vascular (0.6)	derrame cerebral (0.6)	golpe (0)
9	stroke	accidente cerebro- vascular (0.6)	accidente cerebro- vascular (0.6)	golpe (0)

Tabla 2: Ejemplos de diferencias en el tratamiento de la terminología en THs y TAs

En los ejemplos 1, 2, 3 y 4, observamos que los traductores humanos tienden a utilizar unidades terminológicas con un CD más alto y, por tanto, con una pertenencia al ámbito de especialidad mayor que en las TAs. Esta tendencia entra en contradicción con las propiedades del discurso de divulgación científica, el cual se caracteriza por una reducción del vocabulario de especialidad y una simplificación a nivel léxico-terminológico. Suponemos que ello se explica por una tendencia propia de la traducción en general (Baker, 1995) a exagerar los rasgos prototípicos del tipo de textos a traducir (en este caso, textos especializados).

Otra razón por la cual el número de los candidatos a términos es mayor en las THs es que el extractor de Vivaldi y Rodríguez (2010) sólo detecta las unidades nominales con función referencial y no asigna ningún coeficiente a los verbos ("N/A" en la Tabla 2), de manera que el número elevado de nominalizaciones en las THs tiene un impacto en los resultados (ejemplos 4 y 5). El uso de nominalizaciones en el discurso especializado conlleva una mayor densidad léxica y una implicación de las relaciones semánticas al

pasar de las construcciones clausales a frases nominales. Al igual que en el caso anterior, observamos una tendencia de la TH a exagerar los rasgos propios del dominio de especialidad. En el ejemplo 5, Google realiza la misma selección que el traductor humano debido a que dicha selección es la más frecuente en este contexto.

En los casos en los que la diferencia TA-TH se relaciona con los errores de TA, detectamos diversas tendencias. Como se ve en el ejemplo 6, los errores de Lucy se producen en un contexto de ambigüedad sintáctica o semántica de los TOs, problema que Google es capaz de resolver en la mayoría de los casos. En este ejemplo la relación semántica entre las unidades que conforman el término poliléxico no tiene ninguna marca explícita en inglés, lo cual lleva a un error en la traducción proporcionada por el sistema basado en reglas. Las traducciones de Google se acercan más a la selección léxica de las THs, pero presentan una desventaja importante en comparación con la traducción del sistema basado en reglas, a saber: una falta de continuidad en el tratamiento de la terminología. En los ejemplos 7, 8 y 9 observamos que Lucy, aun realizando una selección poco apropiada, al ofrecer "golpe" como equivalente de *stroke*, lo hace siempre de la misma manera. En cambio, Google realiza la traducción en función del contexto lingüístico inmediato, y ofrece varias opciones para el mismo término ("derrame cerebral", "accidente cerebrovascular", "carrera"), con lo cual el TT pierde la coherencia a nivel del léxico.

4.2 Nivel sintáctico

Para realizar el análisis a nivel sintáctico, anotamos los TTs con el pos-tagging de Freeling-2.2 (Carreras et al., 2004). A continuación, calculamos las frecuencias de aparición de unigramas, bigramas y trigramas de etiquetas POS en las THs y TAs. Utilizamos una representación más o menos detallada en función del tipo de n-gramas. Así, calculamos la frecuencia de aparición de unigramas tomando en cuenta la categoría y algunas de las subcategorías que proporciona el POS-tagging de Freeling-2.2, y, para los bigramas y trigramas, sólo tomamos en cuenta la categoría y la primera subcategoría de las etiquetas POS. Por medio del test de proporciones del

software estadístico R Project identificamos las diferencias estadísticamente significativas (p -valor < 0.05) en la frecuencia de aparición de n-gramas de etiquetas POS en las THs y las TAs. Extraemos del corpus una muestra aleatoria de contextos de aparición (a nivel de oración) de los n-gramas que presentan diferencias significativas en su distribución en las TAs y las THs, e identificamos los fragmentos correspondientes en los TOs. A partir de estos datos, detectamos las regularidades en el comportamiento de los traductores y los sistemas de TA al enfrentarse con una construcción sintáctica del original determinada. Finalmente, clasificamos las diferencias siguiendo las categorías presentadas en la Tabla 1.

En las TAs de Google un 23% de n-gramas están sobre-/sub-representados con respecto a las THs. En las traducciones de Lucy, esto ocurre en un 37% de los casos.

n-gramas	Frec. TH	Frec. Goo.	p -valor TH-Google	Frec. Lucy	p -valor TH-Lucy
pp30	594	486	$9.34*10^{-3}$	451	$1.29*10^{-6}$
pd00	67	26	$7.28*10^{-3}$	33	$6.10*10^{-4}$
nc sp nc	579	710	0.06	1160	$2.2*10^{-16}$
pr vs aq	17	61	$9.48*10^{-6}$	36	0.01
vmii	216	52	$2.20*10^{-16}$	374	$5.624*10^{-10}$
vmis	138	178	0.11	1	$2.2*10^{-16}$
vs vm	12	92	$2.56*10^{-10}$	83	$2.665*10^{-12}$

Tabla 3⁴. Ejemplo de identificación de diferencias estadísticamente significativas TA-TH a nivel sintáctico

En la Tabla 3 presentamos algunos de los unigramas, bigramas y trigramas de etiquetas POS que presentan diferencias significativas en sus frecuencias de aparición en las THs y las TAs. Organizamos todas las diferencias identificadas en términos de funciones discursivas básicas: referencia, modificación y predicación (Croft, 1991). En cuanto a los recursos que se utilizan en la TH en relación con la función referencial, observamos, por

⁴ aq: adjetivo calificativo; nc: nombre común; pr: pronombre relativo; pd00: pronombre demostrativo; pp30: pronombre personal, 3ª persona; sp: preposición; vs: verbo semiauxiliar; vm: verbo principal; vmii: verbo principal, indicativo, imperfecto; vmis: verbo principal, indicativo, pasado.

ejemplo, la preferencia por la referencia anafórica frente a la referencia nominal en los casos en que esta última conlleva la repetición léxica. Ello se manifiesta en la frecuencia de aparición de los pronombres personales y demostrativos, la cual es significativamente más baja en las TAs. En este caso, en la TH se observa un *translation shift* a nivel sintáctico condicionado tanto por la adecuación a la lengua meta (diferencia entre el inglés y el español en el uso de mecanismos de cohesión: referencia anafórica frente a repetición léxica), como por una tendencia universal de la traducción: evitar la repetición.

En relación con la modificación, en las TAs de Lucy el recurso más frecuente es la frase preposicional, mientras que en las THs y las TAs de Google el uso de esta construcción es menos frecuente. Estas diferencias se dan en el contexto de la traducción de los sintagmas nominales complejos del inglés, que se traducen de diversas maneras por los traductores humanos y por el sistema estadístico. En cambio, el sistema basado en reglas las traduce siempre por medio de una frase preposicional, lo cual en algunos casos conlleva una falta de naturalidad en el discurso o una anomalía semántica.

Otro ejemplo de diferencia a nivel sintáctico es la sobre-representación en las TAs del n-grama que corresponde a las subordinadas de relativo con predicación atributiva. Al realizar el análisis cualitativo de una muestra de los contextos de aparición de este n-grama, observamos que dicha diferencia se produce al traducir esta misma construcción del inglés, en cuyo caso los traductores humanos muestran una preferencia por la modificación adjetival debido a que esta tiene un mayor grado de concisión, mientras que los sistemas de TA ofrecen una traducción literal.

En cuanto a los mecanismos de predicación, dado que inglés es una lengua con una morfología verbal pobre en comparación con el español, ambos sistemas de TA presentan problemas en su tratamiento. En el caso de los verbos, las formas en futuro, subjuntivo e imperfecto están subrepresentadas en las TAs de Google, lo cual concuerda con la observación confirmada en numerosos estudios sobre los problemas que presentan los sistemas estadísticos al traducir a una lengua con morfología flexiva rica (Lee, 2004). En las

TAs de Lucy también se observan diferencias en la traducción de la morfología verbal, pero debido sobre todo a la necesidad de explicitar los rasgos gramaticales que no se marcan en inglés de la misma manera que en español. Un ejemplo es la sobre-representación de los verbos en imperfecto y la sub-representación de los verbos en pretérito indefinido al tratar las formas verbales del pasado simple del inglés. En las TAs de Google se detecta la tendencia contraria, que, suponemos, se debe a la frecuencia de estas formas en su corpus de entrenamiento. Otra diferencia en cuanto a las construcciones con función predicativa es la sobre-representación de los n-gramas que corresponden a la construcción pasiva analítica en las TAs, frente al uso de las construcciones medias o activas en las THs. En este caso no se trata de un error gramatical, sino del aspecto pragmático-funcional de estas construcciones, así como de los factores de frecuencia de uso (la construcción media es más frecuente en español que la pasiva analítica).

4.3 Nivel discursivo

A nivel discursivo anotamos los TTs con las relaciones discursivas de la *Rhetorical Structure Theory* (Mann y Thompson, 1988) con la herramienta RSTTool (O'Donnell, 2000) siguiendo la metodología desarrollada en da Cunha e Iruskieta (2010). La anotación se divide en tres fases: segmentación discursiva, detección de relaciones discursivas (por ejemplo, relaciones de tipo Núcleo-Satélite como CAUSA, ANTÍTESIS o RESULTADO, y Multinucleares como LISTA, SECUENCIA o CONTRASTE) y construcción de árboles retóricos. En los casos en los que es imposible interpretar una relación discursiva debido a la falta de coherencia en las TAs, introducimos la marca "?", que indica que la unidad discursiva mínima (EDU) en cuestión no se relaciona con ninguna otra EDU del texto. Para obtener resultados cuantitativos generales, calculamos el número de EDUs en cada grupo de textos. Asimismo, calculamos la frecuencia de aparición de los diferentes tipos de relaciones con el fin de obtener una descripción general de la organización del discurso en la TA y la TH. Realizamos la alineación de TOs, THs y TAs a nivel de EDU y registramos el estatus de cada EDU en términos de su relación discursiva con otro segmento del texto. A

partir de esta alineación, identificamos las diferencias sistemáticas entre las THs y las TAs a nivel discursivo y las clasificamos de acuerdo con la Tabla 1.

En cuanto a la segmentación discursiva las TAs de Google presentan un 23% de diferencias; en cuanto a las relaciones discursivas, un 47%. En las TAs de Lucy el número de diferencias en la segmentación es mayor (27%), pero el número de diferencias en las relaciones discursivas es menor (40%).

Las diferencias que se originan en las traducciones de Lucy se dan en un contexto de ambigüedad léxica o sintáctica del TO. En cuanto a las diferencias relacionadas con las TAs de Google, se deben a las omisiones/adiciones que realiza este sistema al seleccionar la opción más probable. El ejemplo (1) ilustra esta última observación:

- (1) TH: [La biología consistía ahora en el estudio de la información almacenada en ADN - ristas de cuatro letras, A, T, G y C, símbolos de las bases adenina, timina, guanina y citosina - y de las transformaciones que esa información experimenta en el interior de la célula.]NÚCLEO [¡Aquí había matemáticas!]SATÉLITE_INTERPRETACIÓN

Google: [Biología era ahora el estudio de la información almacenada en las cadenas de ADN de cuatro letras: A, T, G y C para las bases denine, timina, guanina y citosina - y de las transformaciones que sufre la información en la celda.] NÚCLEO [No era la matemática aquí!]?

Las diferencias que se originan en las THs son un reflejo de la interpretación del original por parte del traductor. A nivel discursivo, este proceso afecta el orden de las EDUs y la selección de marcadores del discurso y conlleva modificaciones en la estructura discursiva del original en el TT. El ejemplo (2) ilustra esta situación:

- (2) TH: [Los parientes de autistas pueden presentar algunos síntomas,]NÚCLEO [aunque no el cuadro completo que justifique el diagnóstico de la enfermedad.]SATÉLITE_CONCESIÓN

Google: [Los familiares de las personas con autismo pueden no

cumplir todos los criterios para el trastorno,]NÚCLEO [pero todavía tienen algunos de sus síntomas.]SATÉLITE_ANTÍTESIS

Lucy: [Los familiares de gente con el autismo pueden fracasar en encontrar todos los criterios para el desorden]NÚCLEO [pero todavía tener algunos de sus síntomas.]SATÉLITE_ANTÍTESIS

5 Conclusiones

En este artículo hemos presentado una metodología que permite comparar las diferencias lingüísticas existentes entre TAs y THs, en tres niveles de la lengua (léxico-terminológico, sintáctico y discursivo), usando un corpus paralelo inglés-español de textos especializados del ámbito médico.

En términos de rasgos lingüísticos cuantitativos a nivel léxico y sintáctico (frecuencia de unidades léxicas o n-gramas de etiquetas POS), las TAs del sistema estadístico tienen más similitud con las THs que las traducciones del sistema basado en reglas, debido a que Google realiza la tarea mediante modelos de lengua y de traducción tomando en cuenta el factor de frecuencia y haciendo generalización a partir de datos reales. Sin embargo, a nivel discursivo el número de diferencias entre la TA de Google y la TH es mayor, ya que, aun cuando las TAs de Google muestran una aparente naturalidad del discurso a nivel del sintagma u oración, presentan graves errores que afectan la estructura discursiva y la coherencia global del TT.

El objetivo último de la TA es lograr resultados comparables a la TH en términos de calidad. Precisamente, para identificar los aspectos más problemáticos que deben optimizarse, consideramos que es necesario caracterizar las TAs en oposición a las THs, mediante un análisis lingüístico, y tener en cuenta que no todas las diferencias entre la TH y la TA reflejan la calidad global de esta última de la misma manera. Además, dada la naturaleza literal de la TA, la comparación TA-TH permite observar con más claridad las diferencias entre las lenguas y las decisiones del traductor. En cuanto al desarrollo de sistemas de la TA, el análisis de las estrategias de TH da luz sobre los aspectos de esta que podemos/debemos modelar en la TA.

Somos conscientes de que el presente trabajo tiene varias limitaciones. En primer lugar, utilizamos sistemas de TA generales, no adaptados a la tarea de traducción de textos especializados del ámbito médico. En segundo lugar, no realizamos la comparación de las THs con los textos originalmente escritos en español, que aportaría datos interesantes sobre las características distintivas de los TTs. Nos planteamos como trabajo futuro solventar estas carencias, así como aplicar la metodología a otro tipo de textos. Además, trataremos de automatizar la metodología el máximo posible.

Bibliografía

- Amigó, E., J. Giménez, J. Gonzalo, y Ll. Márquez. 2006. MT Evaluation: Human-Like vs. Human Acceptable. En *Proceedings of the 44th Annual Meeting of COLINGACL*, páginas 17-24.
- Baker, M. 1995. Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target*, 7(2):223-243.
- Cabré, M.T. 1999. *La Terminología. Representación y comunicación*. IULA-UPF, Barcelona.
- Carreras, X., I. Chao, L. Padró, y M. Padró. 2004. Freeling: An open-source suite of language analyzers. En *Proceedings of the 4th International Conference LREC '04*.
- Catford, J.C. 1965. *A Linguistic Theory of Translation: an Essay on Applied Linguistics*. Oxford Univer. Press, London.
- Corston-Oliver, S., M. Gamon, y C. Brockett. 2001. A Machine Learning Approach to the Automatic Evaluation of Machine Translation. En *Proceedings of the 39th ACL*, páginas 140-147.
- Croft, W. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago University Press, Chicago.
- Cyrus, L. (2006). Building a resource for studying translation shifts. En *Proceedings of the 5th LREC*, páginas 1240-1245.
- da Cunha, I. y M. Iruskieta. 2010. Comparing rhetorical structures of different languages: The influence of translation strategies. *Discourse Studies*, 12(5):563-598.
- Fernández Polo, F.J. (1999). Traducción y retórica contrastiva: A propósito de la traducción de textos de divulgación científica del inglés al español. Universidad de Santiago de Compostela, Santiago de Compostela.
- Gamon, M., A. Aue, y M. Smets. 2005. Sentence-Level MT evaluation without reference translations: beyond language modeling. En *EAMT 2005 Conference Proceedings*, páginas 103-111.
- Giménez, J. 2008. *Empirical Machine Translation and its Evaluation*. Universitat Politècnica de Catalunya, Barcelona. [Tesis doctoral]
- Lee, Y.-S. 2004. Morphological analysis for statistical machine translation. En *Proceedings of HLT-NAACL*.
- Leuven-Zwart, K.M. van. 1989. Translation and original: Similarities and dissimilarities. *Target*, 1(2):151-181.
- Mann, W.C. y S.A. Thompson. 1988. Rhetorical Structure Theory: Towards a functional theory of text organization. *Text*, 8(3):243-281.
- O'Donnell, M. 2000. RSTTOOL 2.4 - A markup tool for rhetorical structure theory. En *Proceedings of the Natural Language Generation Conference*, páginas 253-256.
- Szymańska, I. 2011. *Mosaics. A Construction-Grammar-Based Approach to Translation*. Semper, Warszawa.
- Toury, G. 2004. Probabilistic Explanations in Translation Studies: Welcome as they are, would they Qualify as Universals? En Mauranen, A. y P. Kujamäki (eds), *Translation Universals. Do they Exist?* Benjamins, Amsterdam, páginas 15-32.
- Vivaldi, J. y H. Rodríguez. 2010. Using Wikipedia for term extraction in the biomedical domain: first experiences. *Procesamiento del lenguaje natural*, 45:251-254