

LA NOCIÓN DE TRADAUTOMATICIDAD EN LA TRADUCCIÓN AUTOMÁTICA

JOAQUIM MORÉ LÓPEZ
Universitat Oberta de Catalunya
jmore@uoc.edu
SALVADOR CLIMENT ROCA
Universitat Oberta de Catalunya
scliment@uoc.edu

RESUMEN

En este artículo presentamos la noción de tradautomaticidad y explicamos su relevancia en la traducción automática. La tradautomaticidad es el 'aroma de la traducción automática'; esto es, los fenómenos lingüísticos que provocan que un texto parezca haber sido traducido por una máquina y no por un ser humano. Esta noción es relevante en la evaluación de sistemas de traducción automática pero lo es también en la corrección de traducciones, en la confirmación de usos fraudulentos, la detección de spam, o la búsqueda de información

El artículo explica el trabajo experimental que realizamos para elaborar una tipología de tradautomaticidad. Analizamos también la apreciación de estos fenómenos por parte de un extenso grupo de informantes, con diferentes niveles de estudios, edades, etc. El experimento nos ha dado datos sobre la influencia de los pares de lenguas y el método del sistema en la generación de ejemplos de tradautomaticidad. Finalmente, hemos hecho unas recomendaciones metodológicas para detectar ejemplos de tradautomaticidad de la forma más objetiva posible.

PALABRAS CLAVE: tradautomaticidad, traducción automática, evaluación, sistemas basados en reglas, sistemas estocásticos, postedición

THE NOTION OF MACHINE TRANSLATIONNESS IN MT

ABSTRACT

In this article we explain the notion of machine translationness and its relevance in machine translation. Machine translationness refers to the linguistic phenomena that reveals a text to be translated by a machine. This notion is relevant for machine translation evaluation, post-editing, confirmation of dishonest use of machine translation engines, spam detection and information search.

The article explains the empirical study we performed in order to elaborate a machine translationness typology. We also analyse the perception of machine translationness types by a large group of informants with varied learning levels, ages, etc. The experiment provided us with data about the influence of language-pairs and the methodology of the translation engine in machine-translationness generation. Finally, we suggest a few methodological hints about how machine translationness can be detected automatically and objectively.

KEY WORDS: machine translationness, machine translation, evaluation, rule-based systems, stochastic systems, postedition

1. INTRODUCCIÓN

En este artículo presentaremos la tradautomaticidad y la importancia de esta noción en la traducción automática. Hemos creado el término tradautomaticidad para referirnos a los fenómenos lingüísticos que provocan que un texto parezca haber sido traducido por una máquina y no por un ser humano. Esta noción surgió de nuestros trabajos sobre evaluación de sistemas de traducción automática (Moré 2006; Moré y Climent, 2007) pero se puede aplicar a otros usos, como la corrección de traducciones automáticas, o bien la detección de usos fraudulentos de los sistemas que están disponibles en la Red.

En el apartado 2 argumentaremos que la detección de casos de tradautomaticidad es cada vez más necesaria, sobretodo en un contexto en el que la traducción automática se ha popularizado. Este contexto es una consecuencia de la presencia de servicios de traducción automática en Internet que, en el momento en que escribimos este artículo, son gratuitos y de libre acceso. En el apartado 3 explicaremos el estudio empírico que realizamos para conocer qué fenómenos lingüísticos son responsables de que una traducción 'suene a máquina'. En el apartado 4 presentaremos una tipología de casos de tradautomaticidad que fue fruto de este estudio, y en el apartado 5 analizaremos la percepción de estos casos. Sobre este análisis propondremos unos criterios para la detección automática de ejemplos de tradautomaticidad y, por tanto, para la distinción de traducciones automáticas deficientes.

Este artículo se centra en la tradautomaticidad en traducciones al español. De todos modos, nuestro trabajo pretende impulsar el análisis de fenómenos de tradautomaticidad en otras lenguas y la comparación interlingüística de estos fenómenos.

2. LA NOCIÓN DE TRADAUTOMATICIDAD Y SU APLICACIÓN

La noción de tradautomaticidad surgió para fundamentar un enfoque metodológico de evaluación de sistemas de traducción automática distinto del que se está realizando en la actualidad (Moré y Climent 2008). De todos modos, el empuje que ha experimentado la traducción automática en estos últimos años ha provocado que la noción de tradautomaticidad también deba ser considerada en ámbitos distintos del de la evaluación. La traducción automática se ha universalizado, en el momento de escribir este artículo está al alcance de todo el mundo de manera gratuita, y es accesible con un simple botón en servicios tan variados como la traducción de páginas web, la subtitulación multilingüe de vídeos colgados en la red, la traducción de la carta de un restaurante captada por un teléfono móvil, o la traducción en tiempo real de una conversación telefónica.

Actualmente, el público ya se ha acostumbrado a que la traducción automática sea un servicio accesible para todo el mundo y gratuito, además de estar integrado a otros servicios de la red. Esta accesibilidad y facilidad de la traducción automática convierte la tradautomaticidad en una cuestión importante a tener en cuenta. En este apartado argumentaremos la importancia de la tradautomaticidad desde cuatro puntos de vista. El primero será el ya mencionado de la evaluación de sistemas de traducción automática. A continuación hablaremos de la tradautomaticidad en la generación y corrección de documentos, en el uso de la red como corpus de uso de una lengua y finalmente en la búsqueda de información.

2.1. La tradautomaticidad y la evaluación de sistemas de traducción automática

Con el fin de ahorrar los costes en tiempo y dinero de las evaluaciones humanas, se ha trabajado en el desarrollo de un método automático de evaluación. Los métodos que se han propuesto tienen un criterio común, que es la Asunción de la Bondad de las Traducciones Humanas (ABTH), que exponemos de este modo: las traducciones que parecen traducciones humanas son buenas. De hecho, BLEU (Papineni et al. 2001), que es la métrica de evaluación que se ha convertido en estándar, calcula la similaridad entre la traducción automática de un original y un conjunto de traducciones representativo de las variantes de traducción correctas de dicho original. Estas traducciones se llaman traducciones de referencia y son traducciones humanas. El criterio es el siguiente: cuanto más parecida sea la traducción automática a una traducción humana, mejor es. Cuando la traducción automática coincide exactamente con alguna de las traducciones de referencia, el valor de BLEU es el más alto.

Según la ABTH, podemos establecer otra asunción, que es la Asunción de la Mala Calidad de la Tradautomaticidad (AMQT), que exponemos de este modo: las traducciones que parecen traducciones automáticas son malas, o como mínimo, mejorables. Por lo tanto, cuantos más rasgos de tradautomaticidad tenga una traducción, peor será respecto a una traducción del mismo original que no tenga ningún rasgo o tenga menos.

Las dos asunciones pueden parecer las dos caras de la misma moneda pero, en realidad, los gastos de una evaluación dependerán mucho de la asunción que uno tome. Según Callison-Burch et al. (2008: 71), la obtención de traducciones humanas para la campaña de evaluación del Workshop on Statistical Machine Translation de la ACL de ese año, que se basó en la ABTH, costó 17.200 euros.¹ Además de la contratación de unos traductores para que generen las traducciones

¹ Traducción de 2000 frases, aproximadamente, para cada lengua de las seis que establecía la campaña de evaluación.

de referencia, otros costes deberían ser considerados. Habría que añadir la formación de los traductores para que interioricen unas prácticas orientadas a la evaluación que son muy diferentes de las prácticas de su oficio. Por ejemplo, mantenerse en la literalidad, siempre que sea posible, y mantener también las estructuras del original siempre que se pueda, porque hay que tener en cuenta que los sistemas de traducción automática no tienen la capacidad creativa, ni estilística de los traductores humanos.

Por otro lado, para estar seguros de que el resultado de la evaluación será fiable, habría que añadir el coste de revisar todas las traducciones de referencia, con el fin de evitar la existencia de traducciones inadecuadas que perjudiquen el resultado de la evaluación de forma injusta. Sin embargo, cuando las traducciones de referencia son de traductores profesionales, esta revisión no se suele realizar, por influencia de lo que Jelinek (2004) llama *mito de la traducción humana*, y es la asunción de que las traducciones humanas son buenas. Sin embargo, Jelinek considera que el mito no siempre se corresponde con la realidad. Algunas de las razones para desconfiar es el hecho de que los traductores humanos también utilizan sistemas de traducción automática y pueden hacer pasar una traducción automática (revisada o no) como una traducción original. Por otro lado, también reconoce que las traducciones humanas pueden ser peores que las automáticas por factores que van más allá de la pericia del traductor. Las decisiones –a veces demasiado conservadoras y otras veces demasiado arriesgadas– que toma el traductor cuando no tiene suficiente conocimiento del contexto, pueden afectar la idoneidad de su traducción. Es el caso de las traducciones de expresiones idiomáticas. Si la expresión idiomática está en el diccionario de equivalencias del sistema, la traducción será mejor que la de un traductor humano que, por no haber entendido el contexto en el que encuentra, la traduzca de forma literal o, por el contrario, de forma demasiado libre y completamente inadecuada.

Comparada con los costes de una evaluación basada en la ABTH, una evaluación basada en la AMQT supone un coste mucho menor. Una evaluación basada en la AMQT se limita a detectar los fenómenos lingüísticos que contribuyen a la tradautomatización de una traducción. Para hacerlo es necesario tener una tipología de fenómenos lingüísticos de tradautomatización, unos algoritmos de detección de estos fenómenos y unos recursos lingüísticos e informáticos para realizar esta tarea. El coste se concentraría en la creación de los recursos, pero una vez creados, las evaluaciones de sistemas y de sus versiones serían completamente automáticas y se podrían realizar tantas evaluaciones como se quisiera sin costes añadidos. Actualmente, debido a la posibilidad de disponer de herramientas de ingeniería lingüística y de recursos lingüísticos gratuitos, es posible plantearse la evaluación de traducciones automáticas con un coste pequeño.

Por otro lado, el tiempo de realización de una evaluación basada en la AMQT puede ser significativamente menor. Según Florence Reeder (2001), del mismo

modo que no es necesario leer textos largos para saber si el autor es un hablante nativo o está aprendiendo la lengua, tampoco es necesario leer textos largos para saber si la traducción es automática o no. Esto significa que se pueden realizar evaluaciones rápidas de textos cortos porque los rasgos de tradautomaticidad son lo suficientemente reveladores de la calidad del sistema. Sólo en el caso de que las evidencias no fueran lo suficientemente claras, se podría plantear una evaluación más exhaustiva.

La reducción de costes en tiempo y dinero es una consecuencia positiva de una asunción metodológica que, a nuestro juicio, es más consecuente que la asunción de la ABTH. La evaluación de las traducciones automáticas teniendo en cuenta lo que son –traducciones automáticas– y no en función de su parecido con lo que no son –traducciones humanas– y el criterio según el cual una traducción que parece haber sido hecha por una máquina es mala, no plantean tantas objeciones como las que ha provocado la asunción de la bondad según el parecido con una traducción humana.

2.2. La tradautomaticidad y la creación y corrección de documentos y mensajes

Una consecuencia de la accesibilidad, inmediatez y gratuidad de la traducción automática para el público en general ha sido su uso para fines no apropiados. Por ejemplo, los profesores de idiomas admiten tener, en ocasiones, la impresión de estar leyendo una traducción automática y no una redacción escrita por su alumno directamente en la lengua extranjera (Sommers, Gaspari y Nino 2006). Por eso, del mismo modo que existen herramientas de detección de plagio, también deberían existir herramientas de detección de tradautomaticidad. Otro uso no apropiado es el de la generación masiva de mensajes de phishing. Precisamente, tal como dice Pullum (2011), una característica importante para su detección es la ininteligibilidad del mensaje y el caos lingüístico producido por una traducción automática sin supervisar. Por lo tanto, un detector de tradautomaticidad sería también apropiado como filtro de spam y phishing.

En el uso profesional, ya hemos comentado que los traductores utilizan la traducción automática. Cuando las lenguas origen y destino son próximas, por ejemplo el catalán y el español, el traductor automático realiza unas traducciones con la suficiente calidad como para que el esfuerzo se concentre en la postedición, que es la corrección de los errores de traducción del sistema y su adecuación estilística. Además, para otros profesionales, la traducción automática permite aprovechar el trabajo ya hecho. Por ejemplo, un profesor de una universidad catalana que tiene que preparar un curso para estudiantes hispanoamericanos, puede reutilizar algunos de sus materiales escritos en catalán para el curso, traduciéndolos automáticamente al español.

Sin embargo, la cercanía entre las lenguas origen y destino y una virtud de los sistemas, que es la conservación de la maquetación del documento original,

pueden ser el motivo de que muchos errores pasen desapercibidos, no sólo en una primera lectura sino en una segunda y tercera lecturas. Por ejemplo, algunos ejemplos de errores que suelen pasar desapercibidos en traducciones del catalán al español son la presencia de la conjunción copulativa catalana *i* en vez de la conjunción *y*, formas del pretérito imperfecto catalán sin acento, como *corria* en vez de *corría*, o la ausencia de la preposición *a* delante de un complemento directo de persona.

Los errores que pasan desapercibidos, a causa de la cercanía entre las lenguas origen y destino, no son el único problema que tiene el traductor humano que usa un sistema de traducción automática. Es frecuente que el traductor de un artículo científico tenga que traducir una cita en una lengua que no conoce. Cuando el traductor recurre a la traducción automática y no postedita la traducción, algo que suele ocurrir porque el mismo traductor desconoce la lengua, es posible que la publicación provoque las quejas de los lectores que sí la conocen (Fluent Historian 2011).

Traductores, profesores y otros profesionales que deben ser muy sensibles a la corrección lingüística de sus escritos deben dedicar un tiempo de postedición minucioso, atento, que no debe limitarse a una sola lectura. Por esta razón, además de los correctores ortográficos y gramaticales, una herramienta automática de postedición que detectara fenómenos de tradautomatización les sería muy útil. Este detector de tradautomatización iría más allá que los correctores tradicionales. Por ejemplo, la traducción errónea de *chip* en *implantar una patata frita*.

2.3. La tradautomatización y la red como corpus de uso de una lengua

Para los especialistas en lingüística computacional, la red se ha considerado un corpus representativo del uso de la lengua. Existe una aplicación, el *WebCorp*², de la Birmingham City University (Renouf, Kahoe i Banerjee 2005), que tiene un buscador que muestra contextos de uso de palabras, o combinaciones de palabras sacados de la Red. Es particularmente interesante cuando las palabras son muy nuevas o bien queremos ver el uso real de combinaciones de palabras que no están tratadas en un diccionario o en una gramática normativa.

Trabajos como Moré y Climent (2004) y Sjöbergh (2006) presentaban correctores gramaticales que tomaban la Red como corpus de referencia del uso de la lengua, aunque ya advertían que, debido a la falta de control en la calidad lingüística de los contenidos colgados en la Red, había que ser prudente y establecer unas condiciones de idoneidad de los ejemplos de uso.

La evolución de Internet en estos últimos años ha contribuido a la explosión de publicaciones informales, inmediatas y colaborativas, como blogs, wikis, tweets, comentarios en redes sociales, etc. En estas publicaciones se puede apreciar

² <http://www.webcorp.org.uk/>

cómo un público de procedencia lingüística diversa hace comentarios en la lengua del autor, y el autor también responde a esos comentarios no en su lengua sino en la lengua nativa del comentarista. La comodidad con la que se puede tener acceso a la traducción automática, con un abanico de pares de lenguas cada vez más amplio, anima a los participantes a comunicarse superando las barreras lingüísticas. Actúan valorando la inmediatez de la respuesta, la posibilidad de llegar a un público amplio y lingüísticamente diverso, y también actúan por razones de cortesía y ciberetiqueta. Por ejemplo, un blog en español puede tener seguidores franceses. Estos seguidores leen los posts utilizando un servicio de traducción automática español-francés en línea. Cuando quieren publicar un comentario lo traducen automáticamente del francés al español y lo publican al cabo de unos minutos. Tras leer el comentario, el autor del blog, correspondiendo al esfuerzo del seguidor por escribir en su lengua, traduce automáticamente la respuesta del español al francés y la publica.

Es evidente que la postedición es nula cuando el emisor usa la traducción automática para comunicarse de forma espontánea en una lengua que le es desconocida. La postedición tampoco será usual cuando se generalice el uso de la traducción automática para enviar tweets a comunidades multilingües. Es impensable que los seguidores se dediquen a corregir los tweets que se generan a un ritmo trepidante.

Internet, como entorno de comunicación inmediata, espontánea, y multilingüe, es un espacio actualmente abierto y vulnerable a los efectos de la tradautomaticidad. Los desarrolladores de aplicaciones que toman la Red como modelo de uso son cada vez más conscientes de ello. Un ejemplo es la aplicación *Linguee*.³ *Linguee* es una herramienta muy útil para personas que tienen que escribir textos en una lengua que no dominan lo suficiente. Si no saben cómo expresar en la lengua destino una expresión de la lengua original, el buscador de *Linguee* busca en la Red segmentos que contienen esta expresión, paralelizados con su correspondiente traducción. La solución de traducción más probable aparece marcada en negrita. Lo interesante es que la aplicación filtra los resultados según un algoritmo que distingue buenos y malos contextos. Las traducciones automáticas sin posteditar son malos contextos que el buscador ha aprendido a distinguir mediante métodos de aprendizaje automático. Entre las cosas que 'aprende' es que un sitio web suele estar traducido automáticamente cuando encuentra la palabra Wordpress. De todos modos, este criterio discrimina un número importante de blogs y tiene una desventaja clara para los traductores, profesores, etc. que pueden encontrar en los blogs un término equivalente de uso común entre especialistas, y que todavía no está recogido en un diccionario bilingüe de referencia. Recientemente ha aparecido una herramienta, el *Machine*

³ <http://www.linguee.es/>

*Translation Detector*⁴, que detecta de forma rápida si un texto es una traducción automática. Sin embargo, se limita a determinar si la traducción ha sido generada por alguno de estos tres sistemas: Google Translate, Yahoo!, Babel Fish y Bing pero está orientado a detectar usos inadecuados y no indica errores de traducción. Un criterio con fundamento lingüístico de detección de tradautomaticidad, como el que presentaremos aquí, tiene también su utilidad en la corrección, como hemos indicado arriba.

La publicación inmediata, sin revisión lingüística previa, puede provocar una situación paradójica que explicaremos tomando como ejemplo el traductor de Google. El sistema de traducción de Google también toma la red como un corpus representativo del uso de la lengua. Aprende a traducir tomando como referencia un número ingente de documentos paralelizados con sus versiones traducidas. La paradoja está en que si toma también como referencia un buen número de documentos traducidos automáticamente, sin supervisar, también está aprendiendo a traducir como las máquinas, cuando de lo que se trata es de aprender a traducir como los humanos. En realidad, la presencia de la tradautomaticidad en la red se está convirtiendo en un problema que afectará el futuro de la comunicación y la búsqueda de información, un aspecto que tratamos seguidamente.

2.4. La tradautomaticidad y la búsqueda de información

Actualmente, el interés por obtener información se está focalizando en los espacios de comunicación informales, colaborativos e inmediatos. Destacaremos dos razones. La primera es que los blogs y los wikis se están convirtiendo en los espacios de referencia para estar al día de un tema, profundizar en él y compartir conocimientos. La segunda razón es el interés comercial por el *sentiment analysis*; esto es, conocer las opiniones del público acerca de un producto, un servicio, etc. expresados en las valoraciones de los clientes en la página web de un hotel, o en una red social. Por tanto, hay un interés creciente en la búsqueda de información en fuentes multilingües. Estas fuentes son susceptibles de tener, como ha hemos explicado antes, tradautomaticidad.

Aunque el esfuerzo colaborativo controle y asegure la calidad de una publicación, de momento es posible que momentáneamente la publicación lleve a engaño por una traducción automática no supervisada. Es el caso posible de una entrada de la Wikipedia en una lengua que es la traducción automática de la misma entrada en una lengua distinta. Si la entrada no ha sido revisada, hasta el momento que alguien la corrija, los elementos de tradautomaticidad pueden llevar a confusión. En realidad, ahora es posible que algún internauta lea los *Bocados* en

⁴ <http://www.translatordetector.com>

un blog, wicky o tweet, refiriéndose a los *Mossos* (de Escuadra)⁵, y crea que hay que llamarlos así en español.

La tradautomaticidad creciente en la red también está planteando un problema para los buscadores. La calidad de los resultados de la búsqueda depende de la calidad de los documentos publicados en la red. Por eso, en el contexto actual de competencia entre buscadores (Google, Yahoo, Bing, etc.), la evaluación de la calidad de los documentos encontrados es fundamental. De hecho, Google se ha dado cuenta de que la API que permite acceder a su servicio de traducción automática desde cualquier aplicación está fomentando su uso abusivo. La publicación indiscriminada de documentos, publicidad, etc. con traducción automática sin supervisión supone para la compañía tirar piedras en su propio tejado. Como dice Vashee (2011), no es fácil determinar si el contenido publicado ha sido traducido por una máquina o por un humano, por lo que Google, en la fecha de redacción de este artículo, ha decidido no fomentar el uso no controlado de su sistema de traducción, y por el contrario, quiere tener un control sobre cómo y cuándo se traduce y quién lo hace.

En el campo del seguimiento automático del impacto de un producto, un candidato político, etc., los algoritmos de detección de opiniones se basan en el reconocimiento de palabras que connotan sentimientos positivos de aprecio, satisfacción o bien sentimientos negativos de rechazo, decepción, etc. Aquí también la tradautomaticidad puede inducir a error. Imaginemos que alguien lee *la comisión aprobó mandar a hacer puñetas el proyecto presentado* cuando lo que pasó en realidad fue que la comisión aprobó que se iniciara el proyecto.⁶ Un lector puede tener la idea equivocada de que el proyecto fue evaluado de forma desastrosa por la comisión, con las consecuencias negativas que puede tener esta interpretación para los integrantes del proyecto y para las entidades que lo financian.

En resumen, dotar al buscador de un componente de detección de tradautomaticidad sería muy interesante para poder presentar los resultados más óptimos y también frenar, a medio y largo plazo, el uso abusivo y nocivo de la traducción automática.

3. ESTUDIO EXPERIMENTAL SOBRE LA TRADAUTOMATICIDAD

En el apartado anterior hemos presentado la tradautomaticidad como una noción que hay que tener cada vez más en cuenta, sobre todo cuando se abordan los efectos de un uso abusivo y sin supervisión de la traducción automática. Para frenar los efectos negativos de la tradautomaticidad es necesario conocer los

⁵ Bocados sería la traducción de *mossos*, como plural de *mos* que en catalán significa 'bocado'

⁶ 'mandar a hacer puñetas' es una traducción de *engegar*, que también significa 'iniciar'

rasgos que lo caracterizan, pero hasta ahora no ha habido todavía un estudio sobre los rasgos de la tradautomaticidad que puedan ser detectados de forma objetiva.

A primera vista, parece que la tradautomaticidad es algo difícil de establecer y delimitar. Que una traducción 'suene a máquina' es todavía más una impresión vaga e imprecisa que una evidencia que pueda explicarse de forma objetiva. A pesar de que hay mucha literatura sobre tipologías de errores de traducción automática, estas tipologías se han confeccionado con unos objetivos que no son el de explicar por qué una traducción parece haber sido hecha por una máquina y no por un ser humano. Por tanto, ha sido necesario obtener datos sobre la percepción real de la tradautomaticidad de forma experimental. Una tarea que no se había hecho en el campo de la traducción automática.

Objetivos del experimento

Los objetivos principales del experimento fueron los siguientes:

- Demostrar que los rasgos de tradautomaticidad pueden tipificarse
- Demostrar que la tradautomaticidad existe independientemente del método del sistema evaluado y de los pares de lenguas elegidos.

La tipificación de los rasgos de tradautomaticidad es fundamental. La tipología aporta el fundamento teórico para discriminar de forma automática las traducciones que suenan a máquina. Por otro lado, también es importante demostrar que la tradautomaticidad es un criterio objetivo de evaluación que puede aplicarse independientemente del par de lenguas y del método del traductor. Esto no niega, sin embargo, que los pares de lenguas y los métodos puedan favorecer la presencia de unos tipos de tradautomaticidad sobre otros.

Criterios metodológicos

A continuación exponemos los criterios metodológicos del trabajo experimental.

■ Definición del objeto de estudio

Nuestro objeto de estudio fue la apreciación de los fenómenos lingüísticos de tradautomaticidad.

■ Población

Para tener datos sobre la apreciación de la tradautomaticidad decidimos entrevistar a 100 personas de distintos niveles de estudios y comprensión lectora. De todos modos, para evitar en lo posible que las circunstancias personales de los informantes condicionaran los resultados, descartamos a gente que conociera los fundamentos de la traducción automática y sus limitaciones y también descartamos dos tipos de gente por su capacidad de análisis de la lengua, tanto

por exceso (traductores, profesores de lengua, etc.) como por defecto (gente que no sabía leer o que no podía entender el significado de las traducciones).

Clasificamos la población en grupos diferenciados porque teníamos la asunción metodológica de que los fenómenos de tradautomaticidad pertinentes tenían que ser detectados por personas de distintos grupos. Los grupos se formaron según tres criterios: edad, nivel de comprensión lectora y género.

Organizamos la población por edades, en intervalos de 15 años. Este margen fue lo suficientemente estrecho como para incluir a individuos con competencias lectoras adquiridas en sistemas educativos parecidos. Incluimos en el mismo intervalo a personas a partir de los 61 años porque los sistemas educativos no fueron muy distintos para las personas de estas edades. No establecimos una edad límite para participar en la entrevista, pero sí que establecimos una edad mínima de 16 años, porque consideramos que las personas por debajo de esta edad todavía están desarrollando sus competencias lectoras.

Para distribuir la población según los niveles de comprensión lectora tuvimos en cuenta su nivel de estudios y sus hábitos de lectura por motivos laborales y también de ocio. Asumimos que el nivel de estudios era revelador de las capacidades de una persona para comprender un texto y, en el caso de los niveles universitarios, también se tuvieron en cuenta los tipos de estudios ya que el contenido de los textos puede diversificar la comprensión lectora. No creíamos que un estudiante de historia, por ejemplo, tuviera una comprensión lectora de un artículo de física como la que tendría un estudiante de ciencias.

Finalmente, aunque no creímos que el sexo tuviera que ver con la apreciación de la tradautomaticidad, decidimos elegir la misma cantidad de hombres y mujeres para evitar cualquier interpretación sesgada de los resultados relacionados con el género.

En la tabla 1 presentamos los datos concretos relativos a la población

Muestra inicial: 100 personas mayores de 16 años						
Distribución de la muestra por nivel y tipos de estudios						
Sin estudios	Estudios primarios o secundarios obligatorios finalizados o en curso	Estudios secundarios no obligatorios finalizados o en curso	Estudios universitarios (técnicos) finalizados o en curso	Estudios universitarios (cc naturales) finalizados o en curso	Estudios universitarios (cc sociales) finalizados o en curso	Estudios universitarios (humanidades) finalizados o en curso
25	25	25	6	6	7	6
Distribución de la muestra per edades						
De 16 a 30 años		De 31 a 45		De 46 a 60		De 61 años en adelante
25		25		25		25
Distribución de la muestra por género						
Hombres			Mujeres			
50			50			

Tabla 1. Datos relativos a la población del trabajo experimental.

■ La entrevista

La entrevista tuvo dos partes diferenciadas. Primero el informante tuvo que responder un breve cuestionario. Este cuestionario sirvió para ubicarlo según la distribución que hemos explicado. Posteriormente se le explicó la prueba que tenía que hacer. La prueba consistía en leer unas frases que eran traducciones al español. Las traducciones podían haber estado hechas por personas o por máquinas. Para cada traducción, el informante tenía que escribir si le parecía una traducción hecha por una máquina o por una persona, Si creía que la había hecho una máquina, tenía que subrayar el segmento (o segmentos) de la traducción que, según su criterio, lo demostraba.

■ Corpus de traducciones

Para confeccionar la tipología de fenómenos de tradautomaticidad, decidimos que las traducciones no fueran textos enteros sino frases descontextualizadas. De este modo, los informantes tenían que interpretar los textos en las mismas condiciones que las máquinas, sin poder recurrir al contexto del discurso para interpretar el sentido de una frase.

Para comprobar si la apreciación de la tradautomaticidad dependía de las lenguas origen y destino, y ante la imposibilidad de realizar una prueba universal, decidimos recopilar las traducciones de sistemas de traducción automática en un par de lenguas representativas de lenguas cercanas, y también en un par de lenguas representativas de lenguas más alejadas. Elegimos el par catalán-español como representativo de las lenguas cercanas, y el par inglés-español como representativo de las lenguas alejadas.

La lengua destino fue el español por razones prácticas. Los informantes provenían del área de Barcelona y un buen número de ellos pertenecían a una generación escolarizada en español. Por tanto, estaban más acostumbrados a leer en español y tenían una familiaridad con esta lengua suficiente como para juzgar si una traducción les sonaba a máquina o no con fiabilidad.

Para comprobar si una evaluación basada en la AMQT favorece un método de traducción sobre otro recopilamos, para cada par de lenguas, las traducciones de un sistema basado en reglas y de un sistema estadístico. Los sistemas basados en reglas y los estadísticos (o estocásticos) son los paradigmas metodológicos que están en vigor en la actualidad.

Se recopilaron 1.000 traducciones automáticas, distribuidas de la siguiente manera: 250 traducciones de un sistema basado en reglas catalán-español, 250 de un sistema estocástico catalán-español, 250 de un sistema basado en reglas inglés-español y 250 traducciones de un sistema estocástico inglés-español. Los sistemas basados en reglas y estocásticos tradujeron los mismos originales.

Para comprobar si los fenómenos de tradautomaticidad eran percibidos por la mayoría de informantes, replicamos tres veces el corpus, porque queríamos que

tres personas de distinta edad y nivel de estudios vieran cada una de las traducciones. La coincidencia de dos informantes, como mínimo, al subrayar el mismo segmento de la traducción era la evidencia de que apreciaban el mismo caso de tradautomaticidad. Así pues, el corpus tuvo un volumen total de 3000 traducciones automáticas.

Para comprobar que los fenómenos de tradautomaticidad detectados eran realmente característicos de las traducciones automáticas, añadimos al corpus 750 traducciones humanas. El volumen de traducciones humanas fue equivalente a una cuarta parte del volumen de traducciones automáticas.

Un aspecto importante fue la procedencia de los originales. Para no favorecer los sistemas estocásticos que se entrenan con traducciones de un solo dominio temático, recopilamos traducciones de originales de dominios variados. Nos interesaba saber si el conocimiento de un dominio por parte del entrevistado tenía influencia en la apreciación de la tradautomaticidad. Por esta razón, recopilamos traducciones de dominios no especializados (diarios, revistas de turismo) y de dominios más especializados (revistas de informática, artículos de economía, documentos jurídicos oficiales, como el Diario Oficial de la Generalitat de Cataluña (DOGC), o los discursos parlamentarios del Parlamento Europeo⁷). Ahora bien, para garantizar que los textos especializados no exigieran una capacidad de comprensión lectora muy alta, que hubiera podido condicionar los resultados, elegimos textos dirigidos al público en general.

4. TIPOLOGÍA DE FENÓMENOS DE TRADAUTOMATICIDAD

A continuación presentamos los catorce tipos de tradautomaticidad que detectamos a partir de los segmentos subrayados por, al menos, dos informantes. Los tipos van seguidos de un código que servirá para referirnos a ellos a partir de ahora. En los ejemplos de cada tipo, la parte en negrita corresponde exactamente al segmento de la traducción subrayado por uno de los entrevistados.⁸

Presentaremos los tipos según seis clases:

- léxicos
- sintácticos
- semánticos
- contextuales
- de codificación
- orto-tipográficos

⁷ Los discursos del Parlamento Europeo son el corpus de entrenamiento de algunos sistemas estocásticos (Moses). Si discriminábamos las traducciones de este corpus podíamos llegar a conclusiones poco ajustadas sobre la tradautomaticidad en los sistemas estadísticos.

⁸ Los códigos son los que presentamos en Moré y Climent (2008) y corresponden a la descripción en inglés de la tipología. Así, AGR corresponde a Agreement (concordancia), POS a Part of Speech, etc.

Clase 1. Léxicos

Tipo 1.1 Palabra que no es de la lengua destino (NO-L2)

Palabra que no se ha traducido a la lengua destino. Por ejemplo, en (1) no se ha traducido la palabra catalana *Missatges*.

(1) Protocolo de acceso IMAP (Internet Message Access Protocolo o Protocolo de Acceso de **Missatges** de Internet)

Curiosamente, también vimos casos del fenómeno contrario; esto es, se subrayaron traducciones de instituciones, productos, etc. que por convención no se traducen. Es el caso de la traducción del periódico *The Times* por *Los Tiempos* o el sistema operativo *Windows* por *Ventanas*. Nuestra tipología no distingue estos casos, porque su grado de acuerdo no llegó a ser significativo. De todos modos, esto no quiere decir que este fenómeno no deba tenerse en cuenta.

Clase 2. Sintácticos

Tipo 2.1. Errores de concordancia (W-AGR)

En este tipo se encuentran los errores de concordancia entre el sujeto y el verbo, entre el determinante y el nombre, el adjetivo y el nombre, una expresión anafórica y su antecedente, o bien entre el sujeto y el predicado nominal. Presentamos unos ejemplos de cada caso.

a) Sujeto-verbo

(2) Estoy menos satisfecho con la idea expresada en el apartado 2 que las ayudas estatales no **debe seguir** adelante .

b) Determinante-nombre

(3) En el pasado , en el curso de las relaciones económicas entre el desarrollo de las grandes potencias y los países pobres de Asia , en nombre de « socios iguales » para el futuro es , en todo caso , **un piadosas** esperanzas , más probablemente , un medio de ocultar explotative relaciones con pseudo-democratic y pseudo-humanitarian frases .

c) Adjetivo-nombre

(4) En primer lugar , respecto al método , una cansada, lifeless procedimiento para la opinión pública , que en realidad marginalises las instituciones de la Unión Europea y que los gobiernos son víctimas de sus **propios laberinto**

d) Expresión anafórica y su antecedente

(5) En vista de la doble desafío de la ampliación , que debe permitir a Europa para llegar a términos consigo mismo, la reforma institucional , que debe reforzar la democracia y la eficacia de las instituciones , la Unión Europea ya , como se suele hace , se sitúa en un

importante cambio de rumbo en su historia .

e) Sujeto-predicado nominal

(6) “¿Por qué los hombres no son **fantásticas?**, dice, por ejemplo.”

Tipo 2.2. Palabra con una categoría gramatical extraña (I-POS)

Palabra con una categoría gramatical que el lector no espera en el contexto en el que aparece. A continuación, veamos algunos ejemplos.

a) Adjetivo inadecuado

(7) He **concreto** mencionado algunos de los factores que nos permita para determinar inmediatamente si la reforma es suficiente o no .

b) Verbo inadecuado

(8) Señor presidente , he votado a favor del Informe Souchet que pretende hacer más facilitar el proceso de cambio de productos agrícolas en envasado de productos agrícolas .

c) Sustantivo inadecuado

(9) Para acabar **feudo** click a Aceptar.

(10) Se traen por ir a la playa, pero también por salir **a cena**.

Tipo 2.3. Forma verbal inadecuada (I-VERBF)

A este tipo pertenecen los verbos en formas no finitas, que tendrían que estar en forma finita, o a la inversa. En el siguiente ejemplo, el verbo no puede aparecer en forma no finita.

(11) Queremos , a petición de que se **diferida** por tercera vez por razones políticas.

En cambio, en (12), el verbo 'ayudar' debería haber aparecido en infinitivo por ir detrás de una preposición.

(12) Creo que el impulso debe llegar desde abajo y que se trata de **ayuda** en un problema en su conjunto .

También pertenecen a este tipo los verbos en forma finita que tiene un modo (subjuntivo o indicativo) no adecuado al contexto. También los verbos en formas no finitas no adecuadas, como infinitivos en vez de participios, o gerundios en vez de infinitivos, y viceversa.

En (13) podemos ver un ejemplo en el cual están juntas formas finitas y no finitas extrañas al contexto. Por un lado, el verbo 'ser' debería de estar en

subjuntivo y el verbo 'reparar' tendría que estar en participio y no en infinitivo, por estar la oración en voz pasiva.

(13) A fin de preservar el equilibrio entre las instituciones , creo que debemos actuar para que este defecto **es reparar**.

Tipo 2.4. Orden inadecuado de los constituyentes (I-ORD)

En este tipo agrupamos los ejemplos de tradautomaticidad provocados por una colocación de los constituyentes de la oración que al lector le parece extraña. Esto afecta a las posiciones de los siguientes constituyentes:

a) Nombre y adjetivo

En (14), el adjetivo ocupa una posición prenominal errónea.

(14) Señora presidenta , hay más víctimas de ETA , pero también más víctimas de la **española represión**: más detenciones , más personas torturados , nuevas dispersals de presos políticos jailed en el País Vasco .

b) Determinantes y nombre

En el ejemplo siguiente, la posición del sustantivo 'tarde', alejada de su determinante 'una' hace que la frase sea difícil de entender. La comprensión se complica porque el sustantivo 'domingo' aparece antes que el determinante 'un'.

(15) Señor presidente , Señor Comisario , señorías , **una domingo tarde un** hace quince días en el sur de Francia , Johan , envejecido 17 . Robert , envejecido 16 , asesinada su infancia amigo , Romain, un firearm y una baseball bat , por ningún motivo o intención , como en la televisión .

En el siguiente ejemplo, el determinante numeral que tendría que haber estado en posición pre nominal, aparece en posición postnominal.

(16) Corp. de Máquinas de Negocio Internacional, Armonk, N.Y., que acompañaba **la caída segundo mayor** en la satisfacción, declinada hacer comentarios sobre la encuesta.

Fijémonos que en estos dos ejemplos también hay problemas de concordancia.

c) Nombre y complemento nominal preposicional

En el ejemplo siguiente, el complemento nominal preposicional ocupa una posición prenominal. La comprensión es difícil porque la posición que ocupa 'de Katrina' hace que el lector crea que es un complemento de 'evacuados' y no de 'huracán'.

(17) Con solamente 22 evacuados **de Katrina de Huracán** dejados en Arena de Reunion, el Red Cross americano el miércoles había esperado cerrar su refugio de Dallas en un día o dos.

También hay casos, como en (18), donde el complemento nominal ocupa una posición prenominal pero la preposición no está presente.

(18) En un incidente que he visto con mis propios ojos , **una ambulancia conductor** ha asesinado en el pie de un 10 centimetre serie como ha aspira a llegar a la salvamento de algunos . lo más sorprendente de todos es que los israelíes policía y el ejército utilizado fragmentación balas .

La forma correcta tendría que haber sido *un conductor de ambulancia*.

d) Verbos y clíticos

En este ejemplo vemos un pronombre colocado en posición pre-verbal cuando el verbo es no finito⁹

(19) En compensación , los estados miembros y la comisión opinamos que las autoridades y el sector privado operadores económicos **a los sufragar** los costes del paso en sus respectivos ámbitos .

e) Constituyentes de una colocación y constituyentes sintácticos con posiciones fijas

El lector queda desconcertado por la disposición diferente de las palabras en colocaciones, como en el siguiente ejemplo, en el cual el lector espera leer *esta mañana temprano*.

(20) **Temprano** esta mañana que embarcamos en un catamarán grande y navegamos por el río de hielo de Ultima Esperanza Fjord, donde las cascadas incontables salen de acantilados escarpados y bosque denso.

El ejemplo siguiente también desconcierta al lector porque el sujeto ocupa una posición pre-verbal, que no es la posición que le corresponde al sujeto en una frase en discurso directo.

(21) Intel no ve nada ahora mismo para reconsiderar sus expectativas para el cuarto cuarto, **Otellini decía**

f) Posición partículas negativas

En el siguiente ejemplo, el lector marcó la posición incorrecta de la partícula negativa.

⁹ Hay que decir también que el clítico no está en el caso correcto (dativo). Tendría que ser sufragarles.

(22) En empresa con una serie de otros diputados , he valorado mucho del presidente Prodi declaraciones y he tomado nota de señor Moscovici del modo en que , con diplomática prudencia , él también **advertido no nos a esperar tanto** en la forma de resultados en los próximos dos o tres meses .

g) Posición de adverbios y constituyentes con función adverbial

Hay casos en los que la posición del adverbio hace que el lector interprete que éste modifica una palabra que no le corresponde modificar. En este ejemplo, 'aproximadamente' parece complementar un sustantivo ('acuerdo') y no el numeral. Tendría que ser *el acuerdo de aproximadamente 150 páginas*.

(23) Firmaban el alcalde Robert Cluck, el propietario de equipo Jerry Jones y el vicepresidente de Vaqueros Stephen Jones **el aproximadamente acuerdo de 150 página**.

En el siguiente ejemplo, un constituyente adverbial hace referencia a una fecha que ocupa una posición entre el auxiliar y el verbo principal que no le corresponde.

(24) Éste es un país donde un tercio del territorio ya ha 1974 sido ocupado por el ejército turco y que , por un particular impervious línea de demarcación , los principios fundamentales de la Unión Europea no es aplicable en todo el territorio .

h) Complementos preposicionales desplazados

En este ejemplo, hay un constituyente ('la agenda') que hace de sujeto del verbo 'ser', lo cual desconcierta al lector porque las oraciones iniciadas por 'es preciso' no tienen un sujeto pre-verbal.

(25) Debemos poder hacer que se está avanzando , este avance debe ser mensurables y **la agenda social es preciso contar con** el centro del escenario en el debate político a nivel europeo .

En (26), son la preposición 'a' y su complemento los que no ocupan la posición que les corresponde, (*su política respecto a los estados unidos*), y se interponen entre dos sintagmas nominales ('su política' y 'subcontinente') rompiendo toda coherencia lógica.

(26) A fin de evitar una catástrofe , es fundamental que Europa es su política a los estados unidos subcontinente respecto .

Tipo 2.5. Sobregeneración de palabras (OVER-WRD)

Este tipo se caracteriza por tener más palabras que las que el lector espera. Por ejemplo, en (27i) sobra un verbo, en (27ii) sobra un pronombre y en (27iii) sobra una conjunción.

(27)

- i) Para mí , esto **es** también tiene una gran atención .
- ii) Para consultar la clasificación temática de una página web y /o inscriure'n una nueva os podéis **dirigirse a** Internet Content Rating Association.
- iii) Temprano esta mañana **que embarcamos en un catamarán** grande y navegamos por el río de hielo de Ultima Esperanza Fjord, donde las cascadas incontables salen de acantilados escarpados y bosque denso.

El lector aprecia estos casos de tradautomaticidad en combinaciones de palabras que tienen formas morfosintácticas que no son combinables según la gramática de la lengua destino. Por ejemplo, en (28) la combinación *preposición+preposición+conjunción* le resulta extraña al lector.

- (28) Si la comisión acoge con satisfacción el amplio esencial la propuesta de la comisión , el informe presenta una serie de enmiendas que deben hacerse **en para que** la propuesta separación de poderes trabajo .

Tipo 2.6. Repetición de palabras (WRD-REP)

Dos palabras de la misma forma y categoría gramatical se repiten en el mismo constituyente sintáctico o en constituyentes muy cercanos

(29)

- i) Por consiguiente , negociamos en **un un** mínima base y tenemos una mínima carta , en particular respecto de los derechos sociales .
- ii) El sistema educativo **no** es **no** mejor en satisfacer a Eliza teórico correcto a terapia de discurso.
- iii) Repetir hasta que vuelva el problema y así identificaréis el **dispositivo** o driver de **dispositivo** que causa el problema
- iv) El **paseo** resulta un duro **paseo** en la mesa de trabajo de Mac o dentro de Mi Ordenador en ordenadores de Windows.

Tipo 2.7. Ausencia de una palabra (SYNT-GAP)

En esta categoría se encuentran los ejemplos de tradautomaticidad que el lector ha detectado por faltar una palabra. Es el caso de (30i), (30ii) y (30iii), en los que falta un verbo (*que **hay** en el norte del país*), un pronombre relativo (*una propuesta de directiva de la comisión **que** abarque estas cuestiones*) y un clítico (***los** retorna al Congreso*) respectivamente.

(30)

- i) He visitado Rumania a finales de julio para ver en la primera parte , el funcionamiento de uno de los mayores orfanatos que <Marca de l'entrevistat> en el norte del país .
- ii) Los servicios de la comisión está preparando actualmente una propuesta de directiva de la comisión **abarque** estas cuestiones y también algodón semillas.
- iii) El Senado veta los presupuestos y **retorna al Congreso** por primera vez en la democracia.

Los ejemplos de SYNT-GAP, como los de OVER-WRD, suelen detectarse cuando aparecen combinaciones de palabras con categorías gramaticales y formas que no se pueden combinar en la lengua destino. Es el caso de (31), donde se combina un verbo en gerundio, que no puede tener la función de auxiliar, con un verbo en infinitivo.

(31) Actualmente , el consejo está **hablando** incorporar esos mecanismos en el artículo 7 .

En (32) es extraño ver un verbo en infinitivo precedido por un sustantivo ('autoridad') y no por una partícula preposicional como 'al', por ejemplo (*acojo con satisfacción su autoridad **al** mencionar...*).

(32) Quisiera felicitar al presidente de la comisión en su intervención y acojo con satisfacción su autoridad **mencionar**, en este hemicycle ayer , el papel , ollegial naturaleza y dirección capacidad de la comisión

También es extraño ver una combinación adjetivo+determinante+nombre, como en (33). Se echa en falta una preposición como 'en' que dé contenido semántico al complemento circunstancial (*en todos los vuelos domésticos e internacionales*).

(33) Los mecheros están ahora oficialmente prohibidos **todos los vuelos domésticos e internacionales**.

Clase 3. Semánticos

Tipo 3.1. Vacíos semánticos (SEM-GAP)

El lector echa en falta constituyentes que completen el significado de un verbo, sustantivo, etc. Entre los ejemplos hemos encontrado casos en los que el verbo no tiene un sujeto (35i) o en los que al adjetivo le falta un sustantivo para complementar (35ii).

(35)

i) La ciudad fue fundada por españoles el 1538 y se conoció durante el periodo colonial **sea** Chuquisaca.

ii) Por último , necesitamos una definición más precisa de **la relevantes** del mercado , porque cada vez más , el mercado no es el mercado nacional .

En el siguiente ejemplo, el entrevistado no encontró la referencia a la cosa o a la persona excluida.

(36) Estas zonas sería destacado por la aplicación de criterios para definir este valor añadido en términos de creación **rankings y la exclusión**.

Tipo 3.2. Incoherencia semántica (SEM-INCOH)

En este tipo se encuentran los casos en los cuales los constituyentes no cumplen las restricciones semánticas del nombre o del verbo. Por ejemplo, en (37i), el complemento indirecto del verbo 'contar' no tiene el rasgo semántico de ser humano que debería tener. Por su parte, el sujeto del verbo 'detectar' en (37ii) tampoco tiene este rasgo semántico.

(37)

- i) **¿No deberíamos contar cocaína a la verdad?**
- ii) **Los Bocados detectan** al Vallès una reavivada de asaltos nocturnos a viviendas.

En ocasiones, la presencia de un complemento verbal con unos determinados rasgos semánticos llama la atención sobre una selección verbal o nominal. Por ejemplo, en (37ii) la traducción *Bocados* del original *Mossos*, refiriéndose a los Mossos d'Esquadra, provoca una incoherencia semántica con el verbo *detectar* que no pasa desapercibida. Si se hubiera mantenido la forma original *Mossos* no se habría percibido dicha incoherencia.

En cuanto a las restricciones semánticas de los nombres, veamos los siguientes ejemplos

(38)

- i) Los de la Generalitat Valenciana practican la política del **valenciano escondido**.
- ii) A lo largo de la guerra larga del IRA contra nosotros, dos generaciones de políticos británicos indicaban que habría siempre la ""grieta"" suelta ""en el sistema"" por la cual **el terrorista determinado** se iría.

En (38i), el adjetivo *escondido* no se percibió como el más adecuado para complementar *valenciano*. Por otro lado, los rasgos definido/no definido del sintagma nominal también tienen que mantener una coherencia con el adjetivo, como se muestra en (38ii), donde si en vez de leer *el terrorista determinado*, hubiera leído *un terrorista determinado*, posiblemente el entrevistado la habría dado por buena.

Tipo 3.3. Segmentos distorsionadores (NOI-SEG)

Son segmentos de la traducción que añaden ruido a la interpretación global de la frase y la hacen completamente incomprensible. No se clasifican en un tipo definido porque a menudo estos segmentos son el resultado de la confluencia de diferentes decisiones en la resolución sintáctica y léxica del sistema, y es muy difícil llegar a conocer los pasos concretos que han llevado a generarlos. Además, estos procesos no se replican en otras traducciones. A continuación presentamos algunos ejemplos

(39)

- i) "Robert era un golfista ávido, cántara de softball, y **bombín quienes ganaban** un Campeonato de Estado en 1951.
- ii) El problema más difícil ante esta nueva carta no es su contenido **y si ha jurídica**.
- iii) Los precios de indumentaria eran hacia abajo de 2% de un año más temprano en septiembre, pero resultados de satisfacción **para la rosa sectorial por uno punto a una lectura de índice de 80**, el nivel más alto desde 1995.

Clase 4. Contextuales

Tipo 4.1. Traducción inadecuada al contexto (CON-INCOH)

Los ejemplos de este tipo se caracterizan por tener una palabra o todo un constituyente sintáctico que no se adecua semánticamente al contexto en el que se encuentra. A diferencia de los casos de incoherencia semántica (SEM-INCOH), de los vacíos semánticos (SEM-GAP) y de los segmentos que añaden ruido a la interpretación de la frase (NOI-SEG), la palabra o el constituyente es inteligible y coherente por sí mismo, pero produce la perplejidad al lector en el contexto en el que se encuentra. En el siguiente ejemplo ilustramos la diferencia entre CON-INCOH y SEM-INCOH.

- (40) Qué traer: botas de excursionismo, marcha de lluvia, sombrero, sunblock, **insecto repelente**, prismáticos, paquete de día, que nadan vestimenta.

Insecto repelente es un caso de traducción inadecuada porque es un constituyente inteligible pero no se adecua en absoluto al contexto, que es un listado de cosas para llevar en una actividad excursionista. En cambio, *nadan vestimenta* no es un caso de traducción inadecuada sino de SEM-INCOH, porque la combinación del verbo y el complemento directo no constituyen un grupo semánticamente coherente.

Muchas traducciones inadecuadas lo son de palabras que pueden tener más de una traducción, pero el sistema ha generado la solución menos adecuada al contexto. Por ejemplo, en el siguiente ejemplo, la traducción de *bloody* como 'sangrante' y no como 'sangrienta' no es la más adecuada al contexto.

- (41) Una cincuentena de muertes en otro **día sangrante** en Irak.

El resultado incoherente de un equivalente de traducción mal elegido es evidente en la traducción del verbo *to be* en inglés como *ser* o *estar*, o en la selección de la preposición castellana para traducir la preposición catalana *a*, que puede tener tanto un sentido locativo como de dirección. En (42) encontramos un ejemplo de mala selección de la traducción al catalán del verbo *ser* y también de la preposición *a*.

- (42) La versión 4.76 del Navegador en catalán **ya es al** mercado.

También consideramos como traducciones inadecuadas al contexto otros casos que, si bien el contexto no afecta el significado, sí afecta la forma de la palabra. Así consideramos como traducciones inadecuadas los errores de apocopación (43i), el uso inadecuado de 'de' o 'del' (43ii) o errores en la forma clítica o no del pronombre (43iii) y (43iv).

(43)

- i) **El primero** ministro de Ucrania impugna las elecciones y reitera que no dimitirá.
- ii) Si el billete onward termina en México, Canadá, Islas Bermudas o las Islas **de** Caribe, los viajeros deben ser residentes permanentes legales de esos países.
- iii) Yo **los** estaría muy agradecido, porque no sé como hacerlo.
- iv) Necesitamos respuestas políticas a cuestiones políticas y debemos adoptar ellos todo bien

Clase 5. Codificación

Tipo 5.1. Presencia de caracteres extraños (STR-CHAR)

En este tipo se encuentran los caracteres extraños que aparecen debido a un error de codificación del documento original, como en el siguiente ejemplo

(44)Una vez que Windows ha arrancado, apagueu Windows y espereu que **se?** apague.

En este caso, el sistema no ha reconocido un apóstrofe porque está mal codificado en el documento original.

Clase 6. Orto-tipográficos

Tipo 6.1. Errores ortotipográficos (TYPO-E)

En este tipo se encuentran los casos en los que el lector considera que no se ha hecho un uso adecuado de las mayúsculas y minúsculas, como en (45)

(45) He leído en **El** enlace que configurando el acceso remoto de CESCA, vía el proxy, la conexión debería ser mejor.

O bien el lector considera que se ha hecho un uso inadecuado de un signo de puntuación, como los dos puntos en (46)

(46) La memoria RAM: puede ser insuficiente para alojar todos los programas residentes.

También se encuentran en este tipo los casos en los que el informante señaló la falta de un signo de puntuación, como en (47), donde el entrevistado marcó el inicio de la frase para señalar que faltaba el signo abierto de interrogación.

(47)<Marca del entrevistado>Se atrevería a Francia a poner en cuestión el control alimentario español?

5. ANÁLISIS DE LA PERCEPCIÓN DE LA TRADAUTOMATICIDAD

Del corpus de traducciones del experimento, se obtuvieron 1.250 traducciones que no eran iguales a ninguna otra. De estas 1.250, 863 (69%) tenían algún ejemplo de tradautomaticidad como mínimo para un entrevistado.

En la figura 1 mostramos la distribución de ejemplos de tradautomaticidad según el traductor. Llama la atención que un 13% de las traducciones percibidas como automáticas fueran traducciones humanas. En cuanto a las traducciones automáticas, se ve que no hubo un par de lenguas que provocara mucha más tradautomaticidad que otra; aunque sí que hay más casos en la dirección inglés-español (EN-ES), debido a que es un par más alejado que el catalán-español (CA-ES). Por otro lado, no se ve que un método provoque significativamente más tradautomaticidad.

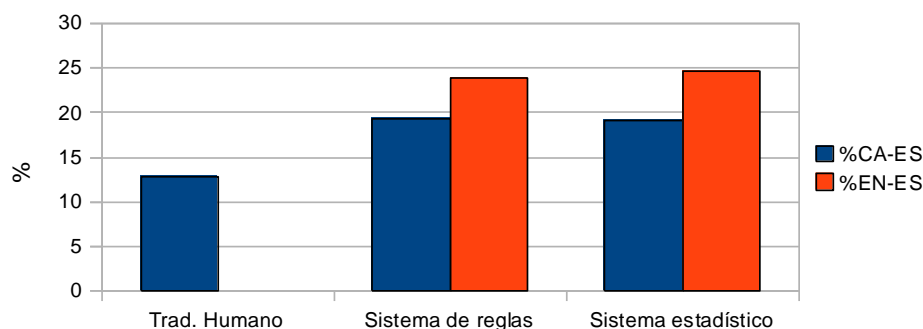


Figura 1. Distribución de casos de tradautomaticidad según el traductor

Aunque hay segmentos en las traducciones humanas que no sabemos por qué fueron subrayados, otros segmentos apuntaban a que el entrevistado había identificado casos de nuestra tipología por razones que dependían de su criterio estilístico, su conocimiento del tema o su capacidad en reconocer la palabra de la lengua origen. Estos son factores particulares y, de hecho, el grado de acuerdo entre los informantes, que no llegó al 2%, demuestra que la apreciación fue muy subjetiva.

El desconocimiento del tema puede explicar que una palabra se considerara como no traducida, como en (48).

(48) **Anonymity Proxy** para Windows es de dominio público y lo podéis bajar.

Fue sorprendente ver 'fichero adjunto' subrayado como ejemplo de tradautomaticidad. El desconocimiento del tema explica que el entrevistado no

encontrara una coherencia semántica entre la palabra 'fichero' y la palabra 'adjunto'.

También hay que tener en cuenta el criterio subjetivo sobre cuándo es pertinente o no traducir una palabra según el contexto. En el siguiente ejemplo, el informante consideró que el traductor tendría que haber traducido 'seny', que es un criterio opuesto al del traductor profesional que prefirió mantener la palabra original.

(49) 'El **seny** Benach'

La interferencia lingüística en los entrevistados bilingües explica que algunos informantes consideraran erróneamente que faltaba o sobraba una palabra funcional, simplemente por interferencia sintáctica con su lengua nativa.

La impresión de que faltan o sobran palabras también se debe al desconocimiento de criterios estilísticos que están relacionados con un género de escritura. Por ejemplo, la ausencia de un artículo definido para las referencias genéricas en el lenguaje jurídico y administrativo.

(50)

- i) EDICTO del Ayuntamiento de la Vall de Bianya, **sobre contratación** de personal.
- ii) ANUNCIO del IES Gallecs, sobre **pérdida** de un título académico.

Sorprendió la presencia de un caso en que toda la traducción humana fue subrayada, como si fuera un NOI-SEG, posiblemente por faltarle contexto para entenderla. De todos modos, fue un solo caso y no hubo acuerdo entre informantes para considerarla así.

La subjetividad causada por las diferencias de conocimiento, criterios estilísticos y conocimiento de la lengua origen influye también en las traducciones automáticas. Por ejemplo, veamos dos casos de CON-INCOH

(51)

- i) **La pestilencia**, también, está llegando, dicen, en las alas de pájaros que emigran, en forma de gripe avian.
- ii) Así pues, os **convidamos** a descubrir la luz y el color de un paisaje que ha enamorado a tantos artistas.

El criterio estilístico que prefiere *peste* a *pestilencia* podría explicar el por qué esta palabra fue subrayada. Este criterio puede ser muy estricto en una persona que no utiliza la traducción automática y que no es consciente de sus limitaciones y, por tanto, cualquier expresión distinta a cómo la diría de forma natural le suena a máquina. En cambio, otra persona más familiarizada con la traducción automática puede tener un criterio más laxo y marcar sólo los casos que afectan la inteligibilidad de la frase. En cuanto al ejemplo (51ii), la similitud de *convidamos*

con el verbo catalán *convidar* pudo haber alertado a un entrevistado catalanoparlante.

Veamos ahora la distribución de tipos de tradautomaticidad en las traducciones automáticas. En las figuras 2 y 3 mostramos esta distribución en los sistemas basados en reglas y en los sistemas estadísticos respectivamente.

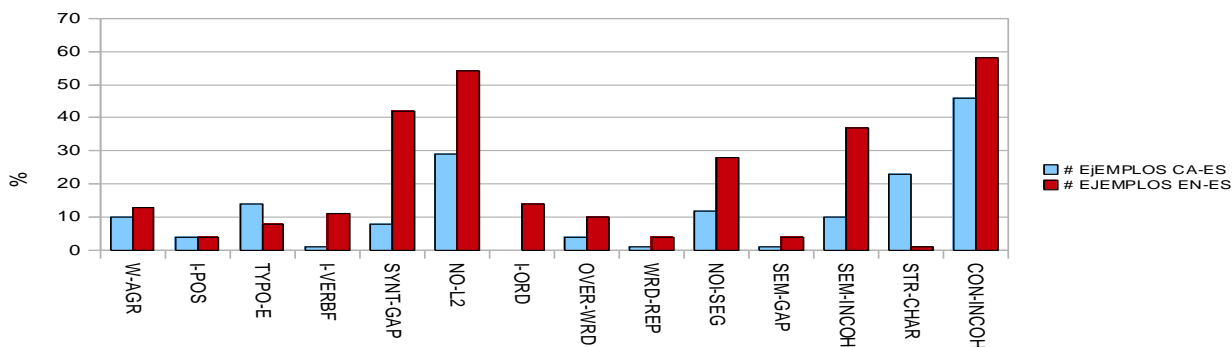


Figura 2. Casos de tradautomaticidad por tipos en los sistemas basados en reglas

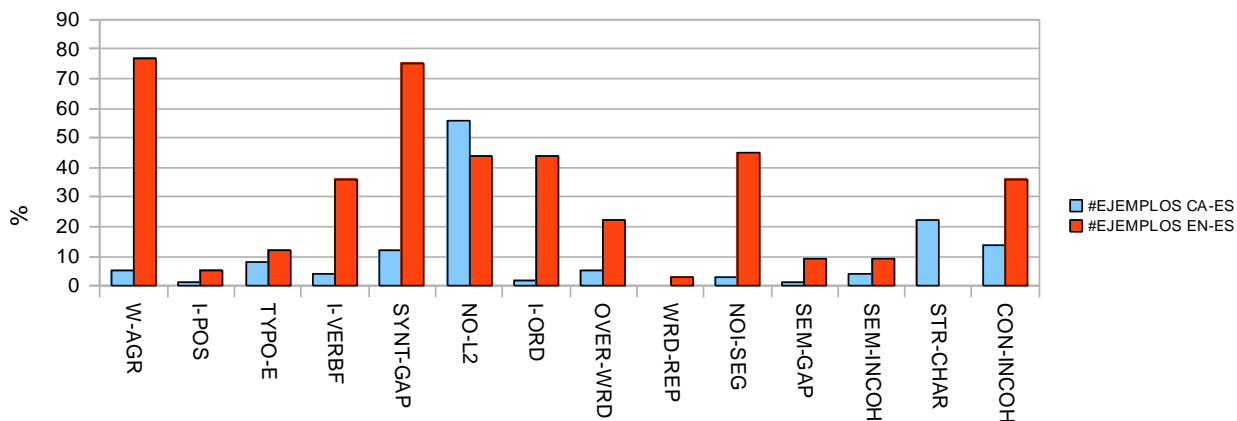


Figura 3. Casos de tradautomaticidad por tipos en los sistemas estadísticos

Por su parte, en la figura 4 mostramos los tipos de tradautomaticidad según su acuerdo.

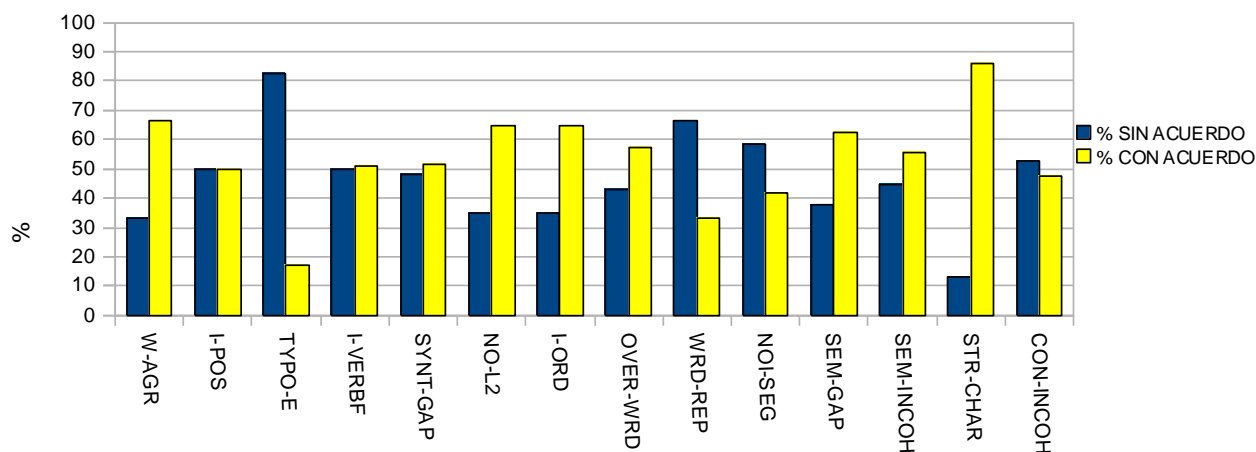


Figura 4. Acuerdos tipos de tradautomaticidad

La proximidad entre lenguas afecta al número de casos de SYNT-GAP, I-ORD, I-VERBF, NOI-SEG y SEM-INCOH, que en la dirección inglés-español superan a la dirección catalán-español.¹⁰ Esta diferencia era previsible por el hecho de que el catalán y el español son lenguas próximas.

Otro factor importante es el método del sistema de traducción automática. Por ejemplo, el sistema estadístico inglés-español generó más ejemplos de W-AGR que el lingüístico, y menos ejemplos del mismo tipo que el lingüístico en el par catalán-español. El tipo CON-INCOH tiene más presencia en los sistemas basados en reglas que en los sistemas estadísticos, mientras que SYNT-GAP, I-ORD están más presentes en los sistemas estadísticos que en los lingüísticos. Esto es consecuente con la metodología de los sistemas estadísticos. El menor conocimiento lingüístico declarado en forma de reglas sintácticas, a favor de un método de aprendizaje del uso real de la lengua, se manifiesta en un aumento de los ejemplos más sintácticos respecto al número de ejemplos sintácticos de los sistemas basados en reglas. En cambio, los sistemas estadísticos no generaron tantos ejemplos relacionados con la incoherencia según el uso de la lengua (CON-INCOH y SEM-INCOH). Por su parte, el sistema estadístico catalán-español generó más ejemplos de NO-L2 que el lingüístico, lo cual pudo ser consecuencia de la menor cobertura léxica de su diccionario.

En cuanto a los acuerdos, advertimos que factores derivados del procedimiento de la entrevista, por ejemplo la falta de atención por fatiga, habían influido en la percepción de los casos de tradautomaticidad y por consiguiente, tuvieron una consecuencia en el recuento de acuerdos. También advertimos que los acuerdos dependieron en gran medida de si el fenómeno lingüístico de tradautomaticidad era 'más visible' o, por el contrario, era más discreto. Por ejemplo, los constituyentes funcionales breves (preposiciones, artículos, clíticos,

¹⁰ STR-CHAR sólo aparece en el par catalán-español debido a un problema de identificación de la codificación del fichero original en catalán.

etc.) no siempre se detectan a primera vista. Se produce un fenómeno parecido al que describe la Gestalt, según el cual, en la percepción de una línea discontinua, el observador llena los huecos y ve una recta. Como comentábamos en el apartado 2.b, muchos correctores que corrigen traducciones automáticas reconocen que la ausencia de estas palabras es imperceptible porque no son imprescindibles para interpretar correctamente la frase: son capaces de relacionar semánticamente los constituyentes sin necesidad de la verificación visual de las palabras que hacen de nexo entre ellas. Por ejemplo, en (52) el verbo 'disgustar' tiene una estructura argumental lo suficientemente establecida como para que la ausencia de la preposición *a* no impida la comprensión de la frase.

(52) Decía que hoy en día encontramos que nosotros mismos éramos clavados en ""un Reino de Aburrimiento"" y que **disgustaba los jóvenes**, especialmente, esto: el papa les había enseñado una salida de este reino.

Esto puede explicar, por tanto, la poca distancia entre los casos con acuerdo y sin acuerdo en SYNT-GAP. En el fenómeno contrario, la presencia de palabras funcionales de más, que serían casos de OVER-WRD, comprobamos que no todos los entrevistados marcaban las preposiciones o verbos auxiliares que sobran si no les provocaban confusión. En cambio, si se repetían palabras con un peso sintáctico y semántico importante, como una partícula de negación, entonces el acuerdo era considerable.

En principio, cuanto más alejados estén los pares de lenguas, los fenómenos de tradautomaticidad son más visibles. En el caso de ejemplos de NO-L2 es claro. Una palabra de la lengua origen no traducida es más evidente, sobre todo si no tiene parecido con una palabra de la lengua destino. Por otro lado, en casos como I-VERBF, son más visibles las formas finitas allí donde se esperaba una forma no finita, y viceversa. Por el contrario, características como el modo verbal o el rasgo imperfecto o perfecto del verbo, son menos visibles y dependen también de criterios individuales estilísticos.

(53)

- i) Para táctica, pero también con el convencimiento de que, si **es así**, la reforma no tendría el calado necesario.
- ii) Hace un mes, cuatro terroristas **vijaban** a Londres.

Los casos más visibles fueron también palabras muy diferentes en forma y significado a la palabra que se habría esperado, como en (54i), donde se esperaba 'haced' y no 'feudo'. En cambio, si la diferencia era de un acento (e.g. 'donde'/'dónde') o si la palabra tenía una forma parecida o un significado relacionado con la palabra esperada (e.g. 'eclipses-nombre'/'eclipses-verbo'; 'facilitar-verbo'/'fácil-adjetivo, como en (54ii)), el grado de acuerdo variaba, a pesar de de que la categoría gramatical fuera inadecuada (I-POS).

(54)

- i) Para acabar **feudo** click a Aceptar.
- ii) Señor presidente , he votado a favor del Informe Souchet que pretende hacer más **facilitar** el proceso de cambio de productos agrícolas en envasado de productos agrícolas ..

Los fenómenos de tradautomaticidad más visibles fueron los errores de codificación (STR-CHAR), seguidos de la falta de concordancia (W-AGR) y el orden de constituyentes (I-ORD), que son tipos sintácticos, y la presencia de palabras que no son de la lengua origen (NO-L2), de tipo léxico. Por el contrario, el caso con menos acuerdo fue E-TYPO, sea porque los signos de puntuación son menos visibles o porque también en su apreciación influye un criterio estilístico subjetivo.

Esperábamos que los tipos semánticos tuvieran más acuerdo, sobre todo NOI-SEG porque añade ruido a la interpretación de la frase. Sin embargo, comprobamos que el desacuerdo no fue tal por esta razón: algunos entrevistados marcaron un segmento que englobaba diversos constituyentes sintácticos con nula coherencia semántica. En cambio otros se concentraron en un segmento concreto, tipificándolo de otra manera, sin entretenerse en comprobar si lo que venía a continuación tenía coherencia. Por ejemplo, en (55) un entrevistado se concentró en marcar el caso de STR-CHAR, mientras que otro subrayó hasta el final, porque consideró que toda la traducción era incomprensible. También hay que decir, como se ve en (55iii), que algunos entrevistados, ante traducciones que les resultaban incomprensibles, no se molestaban en marcar nada.

(55)

- i) Finalmente quién **al?** última pestanya Opcions avanzadas.
- ii) Finalmente quién **al? última pestanya Opcions avanzadas**
- iii) Finalmente quién **al?** última pestanya Opcions avanzadas.

En cuanto a SEM-GAP, el número de ejemplos fue muy escaso en las dos lenguas, y los casos de SEM-INCOH tuvieron más acuerdo cuanto más alejadas eran las lenguas. Por otro lado, los grados de acuerdo en SEM-INCOH también dependen del conocimiento del tema. Sin un conocimiento del dominio temático y de su terminología es posible que se consideren como traducciones incoherentes términos que son normales en su dominio, simplemente porque no tienen un significado composicional.

Finalmente, decir que la poca distancia entre acuerdos y no acuerdos en las traducciones adecuadas al contexto (CON-INCOH) se debe a criterios subjetivos, ya explicados al hablar sobre las traducciones humanas, sobre qué solución de traducción es más adecuada al contexto.

6. CRITERIOS DE DETECCIÓN AUTOMÁTICA DE TRADAUTOMATICIDAD

Hay unos tipos que sólo han sido detectados en las traducciones automáticas. Estos son: W-AGR, I-ORD, I-VERBF, I-POS y STR-CHAR. En principio, una detección automática de la tradautomaticidad debería priorizar la búsqueda de ejemplos de estos tipos, aunque sin descuidar la detección de ejemplos de NO-L2 y CON-INCOH, dos de los tipos con más ejemplos. Aunque los ejemplos de SEM-GAP fueron muy escasos, también deberían ser considerados. Sin embargo, los recursos de ingeniería lingüística para detectar ejemplos de este tipo en español todavía están en desarrollo. Por el contrario, la detección automática no debería considerar casos de E-TYPO porque tiene un grado muy elevado de subjetividad.

Quedarían por considerar qué ejemplos con un grado de acuerdo por debajo del 60% (ver figura 4) deberían ser detectados. Para resolver esta cuestión, reflexionamos sobre cuáles podían ser las causas por las cuales un individuo marca unos ejemplos de tradautomaticidad y no marca otros. Creemos que un ejemplo de tradautomaticidad es un segmento de la traducción que llama la atención del individuo porque contradice su modelo de lengua, tanto a nivel sintáctico, morfológico como semántico y estilístico. Un modelo que ha ido elaborando y que es fruto de unas lecturas, de una educación, etc. Como que los modelos de lengua pueden ser distintos entre individuos, la detección de ejemplos de tradautomaticidad independientes de criterios subjetivos debería hacerse según la probabilidad de que unas determinadas combinaciones de formas, categorías gramaticales y significados estén presentes en un corpus representativo del uso de la lengua destinado a un público general. Por ejemplo, no debería marcar *fichero adjunto* pero sí *Los Bocados* cuando debería ser *Los Mossos* (d' Esquadra), o *para acabar feudo* *click a Aceptar*.

Por otro lado, los ejemplos de tradautomaticidad son más visibles si evocan una forma correcta a partir de una forma incorrecta. Esto afecta sobre todo a los casos de CON-INCOH que, aun siendo inteligibles, evocan la solución de traducción correcta. Esto explicaría la detección de ejemplos como *el primero ministro*, *Islas de Caribe*, o *ya es al mercado*. El entrevistado consideró que *primero ministro* o *Islas de Caribe* se desvían de las formas que está acostumbrado a ver en sus lecturas (*primer ministro*, *Islas del Caribe*, *ya está en el mercado*). Por tanto, sería conveniente tener en cuenta los fenómenos lingüísticos que plantean más de una solución de traducción (*primer/primero*, *ser/estar* o *a/en* para la traducción de la preposición *a* en catalán, etc.) y comparar la probabilidad de cada solución, según el corpus representativo del uso de la lengua que hacíamos referencia. Las soluciones con una probabilidad nula o muy baja respecto a otra solución deberían considerarse ejemplos de tradautomaticidad. De este modo se detectaría *primero ministro* y no, en cambio, *os convidamos* ni otros ejemplos que se marcaron en las traducciones humanas.

7. CONCLUSIÓN

En este artículo hemos presentado una definición de la noción de tradautomaticidad y hemos explicado su relevancia no sólo en la evaluación de sistemas de traducción automática sino en la postedición, la detección de usos inadecuados, spam, recuperación de información, *sentiment analysis*, etc. También hemos explicado el trabajo experimental a partir del cual hemos establecido una tipología de ejemplos de tradautomaticidad.

El trabajo experimental se realizó con un número considerable de traducciones humanas y traducciones automáticas realizadas por dos sistemas con métodos distintos (basado en reglas, estadístico). Los pares de lenguas fueron el catalán-español como representativo de lenguas próximas, y el inglés-español como representativo de lenguas alejadas. El experimento nos ha dado datos sobre la influencia de los pares de lenguas y el método empleado en la generación de ejemplos de tradautomaticidad. Hemos visto que los sistemas basados en reglas no generan más tradautomaticidad que los estadísticos, aunque sí es cierto que varía la proporción de ejemplos de cada tipo según el método empleado. Lo mismo ocurre según el par de lenguas elegido.

El trabajo experimental ha confirmado que una evaluación de traducciones automáticas, basada en la asunción de la bondad de las traducciones humanas, depende de apreciaciones muy subjetivas. En realidad, hemos comprobado que hay factores que afectan la apreciación de la calidad de una traducción. Algunos de estos factores son consecuencia del procedimiento de la evaluación, como el cansancio y, por tanto, la falta de atención. Además, los grados de acuerdo disminuyen cuando los fenómenos lingüísticos son poco visibles, sobre todo para un evaluador que, de forma inconsciente, lee la frase como debería ser y no como realmente es. Otros factores son el conocimiento del tema, el nivel de comprensión lectora, o la exigencia lingüística y estilística que son fruto de la experiencia de cada lector. Aún así, los resultados han demostrado que los niveles de acuerdo de percepción de tradautomaticidad en las traducciones humanas es significativamente bajo, comparado con las traducciones automáticas.

Finalmente, hemos hecho unas recomendaciones metodológicas para detectar ejemplos de tradautomaticidad de la forma más objetiva posible. La primera recomendación es que primero se detecten combinaciones de formas, categorías gramaticales, y contenidos semánticos que tienen una probabilidad casi nula o nula de aparecer en un modelo de uso de la lengua, elaborado a partir de un corpus representativo del uso por un público general. El segundo procedimiento sería localizar los fenómenos lingüísticos de la lengua destino que suponen para el sistema decidir entre dos o más soluciones de traducción (apocopación, orden de los constituyentes de un sintagma nominal, palabras con varias traducciones posibles, traducción o no de un neologismo, etc.) y calcular la

probabilidad de cada una de ellas según el modelo de lengua, comprobando si la solución del traductor ha sido la menos probable.

BIBLIOGRAFÍA

- MORÉ LÓPEZ, J. (2006), "La tradautomatización: un concepto aplicado a la evaluación de sistemas de traducción automática", *Revista de Procesamiento del Lenguaje Natural*, 37, 233-240.
- MORÉ LÓPEZ, J. y CLIMENT ROCA, S. (2007), "A Cheap Evaluation Method Based on the Notion of Machine-Translationness", *Proceedings of the Metis-II Workshop New Approaches to Machine Translation*, 83-90.
- _____ (2008), "A Machine Translationness Typology for MT evaluations", *Proceedings of the twelfth conference of the European Association for Machine Translation*, 130-139.
- PAPINENI, K., ROUKOS, S., WARD, T. y ZHU, W, S. (2002), "BLEU: a method for automatic evaluation of machine translation", *ACL-2002 40th Annual meeting of the Association for Computational Linguistics*, 311-318.
- CALLISON-BURCH, C., FORDYCE, C., KOEHN, P., MONZ Y C., SCHROEDER, J. (2008), "Further Meta-Evaluation of Machine Translation", *Proceedings of the Third Workshop on Statistical Machine Translation, ACL*, 70-106.
- JELINEK, R. (2004), "Modern MT Systems and the Myth of Human Translation: Real World Status Quo", *Proceedings of the International Conference Translating and the Computer*.
- REEDER, F. (2001), "In One Hundred Words or Less", *MT Evaluation Workshop MT Summit VIII*.
- SOMERS, H., GASPARI, F. y NIÑO, A. (2006), "Detecting Inappropriate Use of Free Online Machine Translation by Language Students- a Special Case of Plagiarism Detection", *Proceedings of the eleventh conference of the European Association for Machine Translation*, 41-48.
- PULLUM, G.K.. *The Barley is their Goal* [en línea]. Language Goal 7 julio 2011 [Consulta: 25 julio 2011] Disponible en: <<http://languagelog.ldc.upenn.edu/nll/?p=3253>> *Why You Should Not Rely on Machine Translation* [en línea]. Fluent Historian 5 julio 2011 [Consulta: 25 julio 2011] Disponible en: <<http://blog.fluenthistorian.com/2011/07/05/why-you-should-not-rely-on-machine-translation/>>
- RENOUF, A., KEHOE, A. y BANERJEE, J. (2005), "The WebCorp Search Engine: a holistic approach to text Search", *Electronic Proceedings of CL2005*
- MORÉ LÓPEZ, J. y CLIMENT ROCA, S. (2004), "A Grammar and Style Checker Based on Internet Searches", *Proceedings of LREC*, 1931-1934.
- SJÖBERGH, J. (2006), "The internet as a normative corpus. Grammar checking with a search engine", *Technical Report TRITA-CSC-TCS 2006:3, School of Computer Science and Communication, the Royal Institute of Technology, Stockholm, Sweden*.
- VASHEE, K., *The Google Translate API Furor: Analysis of the Impact on the Professional Translation Industry. Part 1* [en línea]. Empty Pages julio 2011 [Consulta: 27 julio 2011] Disponible en: <<http://kv-emptypages.blogspot.com/2011/07/google-translate-api-furor-analysis-of.html>>