



Onomázein

ISSN: 0717-1285

onomazein@uc.cl

Pontificia Universidad Católica de Chile
Chile

Diéguez M., María Isabel

ACIERTOS Y ERRORES EN LA TRADUCCIÓN AUTOMÁTICA: METODOLOGÍA DE LA
ENSEÑANZA-APRENDIZAJE DE LA TRADUCCIÓN HUMANA

Onomázein, núm. 6, 2001, pp. 203-221

Pontificia Universidad Católica de Chile
Santiago, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=134518177011>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

**ACIERTOS Y ERRORES
EN LA TRADUCCIÓN AUTOMÁTICA:
METODOLOGÍA DE LA ENSEÑANZA-APRENDIZAJE
DE LA TRADUCCIÓN HUMANA¹**

María Isabel Diéguez M.

Pontificia Universidad Católica de Chile

Resumen

En este artículo se expondrá una metodología de enseñanza-aprendizaje de la traducción humana (TH) basada en el análisis del error de la traducción automática (TA). El corpus parcial del análisis está compuesto por dos textos en inglés que representan diferentes tipologías textuales y áreas temáticas al interior del discurso académico o científico: un breve artículo semiespecializado sobre un instituto de biotecnología y un resumen especializado de una comunicación presentada en un seminario sobre medio ambiente. Para realizar la traducción automática al español de los textos se utilizó el programa de TA en línea Systran. Mediante el análisis de la TA se ha podido proponer una tipología de aciertos y errores más recurrentes como también un análisis de las posibles causas de los errores. El objetivo último es familiarizar al estudiante de traducción con la técnica de postedición de TA, habilidad que se exige cada día más en el mercado de la traducción profesional.

Abstract

(A methodology of teaching-learning human translation (HT) based on error analysis of machine translation (MT) is presented. For such analysis, two original English texts and their respective Spanish MT were used. One original text is a short news article about a biotechnological institute and the other an abstract on environment. The on-line machine translation program, Systran, was used. By analysing the errors of MT, conclusions may be drawn regarding type and cause of recurrent errors and regarding the teaching of professional translation.)

¹ Investigación realizada en el marco del proyecto FONDEDOC 2000 “Aplicación del método de traducción asistida por computador (TAC) al proceso de enseñanza-aprendizaje de la traducción humana”. Investigadora responsable: M. Isabel Diéguez M.

INTRODUCCIÓN

En la actualidad, la formación de traductores profesionales requiere la incorporación de una serie de contenidos y metodologías que apunten al desarrollo de todas las competencias que estos profesionales deben adquirir. Es sabido que un currículum equilibrado y eficiente para la formación de traductores que se desempeñarán en el siglo XXI debe considerar todas las competencias que conforman su quehacer, a saber, la competencia lingüística, cultural, teórica, práctica y, no menos importante, la competencia técnica.

Este último componente de habilidades técnicas ha cobrado gran relevancia en los últimos años fundamentalmente por dos razones. La primera dice relación con el vertiginoso desarrollo de las tecnologías de información que permiten al profesional de cualquier disciplina mejorar ostensiblemente la calidad y eficiencia de su trabajo. La segunda se refiere al desafío que implica “traducir” estas nuevas tecnologías a la formación concreta de un profesional; en este caso específico, del traductor. Es sabido que se requiere un tiempo considerable para poder estar a la par con los nuevos avances de la informática y es por ello que muchas veces se rehúye esta tarea. No obstante, consideramos que por las características del quehacer del traductor profesional en tanto mediador lingüístico-cultural, no podemos obviar este desafío en un mundo cada vez más globalizado. De esta manera, es necesario desarrollar las instancias que permitan a nuestros alumnos adquirir, en el marco de su formación como futuros traductores profesionales, los conocimientos y, lo que es más importante aún, los criterios para manejar una serie de herramientas informatizadas que se han transformado en el último tiempo en el “lápiz y papel” de antaño.

La competencia técnica en el ámbito de la formación de traductores profesionales abarca una serie de habilidades como el manejo del contrato de prestación de servicios y de la orden de trabajo, técnicas de documentación temática y terminológica, técnicas de revisión de textos y la aplicación de sistemas computacionales multimediales. Los docentes del Programa de Traducción de nuestra universidad han ido incorporando paulatinamente estas habilidades técnicas a los contenidos y metodología de los cursos mínimos de la carrera. De esta manera, se ha logrado un equilibrio entre el desarrollo de habilidades tan disímiles como la competencia cultural y teórica, y las actividades de gestión terminológica, de traducción automática (TA) y de traducción asistida por computador (TAC). Sin embargo, estamos conscientes de que el desafío por actualizarse en lo que respecta a la formación de nuestros futuros traductores sigue presio-

nádonos día a día, sobre todo si pensamos que los grandes avances tecnológicos le imprimen gran velocidad a las comunicaciones.

Dentro de los sistemas computacionales multimediales, han adquirido en las últimas décadas gran desarrollo las llamadas aplicaciones informáticas a la lingüística, aplicaciones que se han materializado en procesadores de textos, correctores de ortografía y sintaxis, lectores ópticos y programas robot capaces de resumir textos en fracciones de segundos, especialmente diseñados para el lector ávido de información. A ello debemos sumar el gran desarrollo que ha experimentado desde la década de los 50 la TA tradicional y posteriormente técnicas de traducción asistida (TAC) como las memorias de traducción.

Dos investigaciones han permitido al Programa de Traducción incursionar en el mundo de la TA y de la TAC. En primer lugar, se realizó entre 1996 y 1997 un Proyecto Fondecyt que apuntaba a determinar criterios confiables de toma de decisiones frente a la TA y la TH como métodos de traducción, con miras a optimizar el tiempo, costo y calidad de la traducción profesional (véase Diéguez y Cabrería, 1997 y 1998). En la segunda investigación se realizó un análisis del error inglés-español de la TA de formas *-ing* (véase Diéguez y Riedemann, 1998). Ambas investigaciones arrojaron resultados de gran interés y proyecciones para futuras investigaciones y, a su vez, permitieron visualizar la posibilidad real de utilizar los programas de TA y el método de TAC como herramientas didácticas que permiten complementar el proceso de enseñanza-aprendizaje de la traducción profesional.

Por todo lo anterior, en este estudio se pretende demostrar, mediante el análisis de un corpus parcial de textos, que los ejercicios de postedición de textos traducidos mediante programas de TA en línea pueden constituir una metodología complementaria válida de enseñanza-aprendizaje de la traducción humana (TH).

A continuación se presenta el marco teórico y la metodología de trabajo de esta investigación. Luego se presenta una selección de ejercicios de postedición de textos pertenecientes al corpus parcial de trabajo. Por último, se proponen conclusiones basadas en los resultados obtenidos a partir de la postedición realizada.

MARCO TEÓRICO

Luego del término de la Segunda Guerra Mundial, la traducción se ha transformado en un instrumento de comunicación interlingual extremadamente importante. Prueba de ello son las investigaciones

que se inician en la década del 50 en torno a la TA. A juicio de Wilss (1982), la ciencia de la traducción probablemente habría necesitado aún más tiempo para evolucionar como una disciplina dentro de la lingüística, con su propio marco de referencia, si no fuera por la emergencia de la TA. Esta última permitió exemplificar con mucha claridad las grandes dificultades metodológicas que debe enfrentar cualquier intento orientado hacia el habla, tendiente a describir en forma exhaustiva los factores que intervienen en la transferencia interlingual y derivar a partir de esos factores una serie de reglas para su desarrollo futuro.

Cabe agregar que la TA, como rama de la lingüística aplicada, es importante desde el punto de vista científico, pues sirve como campo experimental de la lingüística y la informática, especialmente en el ámbito del procesamiento y análisis automático del lenguaje natural. Esta disciplina aplicada permite establecer vínculos con otras disciplinas de la lingüística aplicada como la traductología, la terminología, la sicolingüística y la pragmática, entre otras.

La relación con la traductología se establece al relacionar las etapas que intervienen en el proceso de TH y las correspondientes al proceso de TA, con el fin de obtener equivalencias lingüísticas y de sentido. En el ámbito terminológico, cabe destacar la uniformidad terminológica que garantiza la TA, facilitando así la labor de los traductores profesionales especializados en áreas técnico-científicas. La conexión que se establece entre TA y sicolingüística dice relación con el proceso comunicativo, que implica la comprensión de un mensaje emitido mediante un texto en una lengua con una intención comunicativa determinada y la reexpresión de ese mensaje mediante otro texto en otra lengua con la misma intención comunicativa, que garantice la comprensión del mensaje por parte del receptor. La pragmática también se relaciona con la TA en el sentido de que los significados pertinentes que se realizan en los enunciados se deben determinar no solo lingüísticamente, sino también considerando los aspectos ilocutivos, situacionales y socioculturales, entre otros (Espi, 1999).

Ahora bien, para satisfacer la alta demanda de traducciones que impone el mundo globalizado, se pensó hace ya varias décadas en la TA, cuyo desarrollo no ha estado exento de dificultades. Sabemos que los programas de TA no son capaces de abordar todo tipo de textos y que su producto requiere, en el mejor de los casos, una postedición por parte del traductor. Además, la mayoría de los programas de TA, como se comprobará en esta investigación, presentan problemas de ambigüedad a nivel léxico, sintáctico y semántico. Hasta la fecha, los métodos empleados en la elaboración de sistemas

de TA no han podido resolver muchas de las ambigüedades que presentan las lenguas naturales, lo que se refleja en la precaria calidad de las traducciones que se producen mediante estos sistemas (Diéguéz y Riedemann, 1998).

Junto a la evolución que ha experimentado la TA en las últimas décadas, han surgido algunas iniciativas tendientes a aplicar la TA al método de enseñanza-aprendizaje de la TH. En 1984, Wilss y Thome editan una obra que aborda el tema de la teoría de la traducción y su aplicación en la didáctica de la traducción y de la interpretación. En dicha obra, Thome expone un método de trabajo en el aula de clases con textos de TA. Para los efectos de nuestra investigación, nos hemos basado en el modelo didáctico de Thome y hemos adoptado también su tipología de errores de TA, la cual hemos ampliado, incorporando otros niveles de error como, por ejemplo, el léxico, el léxico-sintáctico, el sintáctico-semántico y el léxico-sintáctico-semántico (Wilss y Thome, 1984). El método didáctico elaborado a partir de Thome ha sido aplicado con algunas variaciones en diferentes cursos de la carrera de traducción y ha dado lugar a publicaciones que exponen las ventajas de dicho método (véanse Diéguéz y Riedemann, 1998 y 1999).

Para llevar a cabo la investigación propuesta se parte de una serie de supuestos teóricos en los que se enmarca el proceso de traducción involucrado, sea este automático, asistido por computador o humano. En primer lugar, la labor de traducción se entiende como un fenómeno lingüístico-cultural cuyas bases se sientan en el análisis del discurso y en la teoría funcionalista de la traducción. En efecto, estos principios teóricos han demostrado en las últimas décadas una especial relevancia para describir y analizar la traducción en general y la traducción del discurso científico, en particular.

El enfoque funcionalista de la traducción plantea como primer paso un análisis del texto fuente y del texto meta, análisis que el traductor debe realizar a la luz de la interpretación de la orden de trabajo que este elabora a partir de la información que obtiene del cliente y del texto fuente. Este primer paso permite al traductor determinar la estrategia de traducción más adecuada y las características que debe cumplir el texto meta. En este contexto, adquieren especial relevancia la intención (*skópos*) de la traducción y los destinatarios de la misma (Nord, 1988). En definitiva, el traductor trabaja en un contexto comunicativo y por ello debe tomar en consideración todos los factores internos y externos al texto fuente y texto meta para poder elaborar una traducción que cumpla con las exigencias de la situación comunicativa especificada por el cliente.

Junto con tomar en consideración la intención de la traducción y sus destinatarios, el traductor debe evaluar el tipo de texto fuente que enfrenta en términos de tipo de discurso que este presenta. Se debe determinar la o las funciones comunicativas (Bühler, 1965 y Reiss, 1983) que caracterizan al texto fuente y su grado de especialización (especializado, semiespecializado y de difusión). El grado de especialización se revela en características lingüísticas y retóricas como la simplificación o no de la terminología, una sintaxis más o menos compleja, la redundancia conceptual y lingüística, el uso o no del recurso de explicitación y la selección léxica, entre otros (Harvey, 1994). Estas características deben ser tomadas en cuenta por el traductor al hacer la traducción y considerarlas en el texto meta, previa consideración de la intención de la traducción y las características de sus destinatarios. Como ya señalamos, en este estudio se han utilizado dos tipos de textos: uno es un breve artículo semiespecializado sobre un instituto de biotecnología y el otro un resumen especializado de medio ambiente. Como se verá en el análisis que se presenta a continuación, la TA del texto especializado se diferencia de la del texto semiespecializado en términos de aciertos y errores, y ello es relevante para desarrollar en el alumno técnicas de postedición de textos traducidos en forma automática.

METODOLOGÍA

Como ya señalamos anteriormente, el método general de trabajo utilizado con los alumnos en el marco de esta investigación fue propuesto y aplicado en dos investigaciones anteriores (véase Diéguéz y Riedemann, 1998 y 1999). Este método ha permitido lograr una traducción correcta desde el punto de vista lingüístico, pragmático y funcional, e introducir y familiarizar al alumno con las técnicas de postedición de textos traducidos mediante programas de TA. Destacamos la pertinencia de practicar este tipo de ejercicios con los alumnos, en primer lugar, porque es una realidad que muchos clientes solicitan a los traductores que utilicen los programas de TA que ellos han adquirido y esperan que el traductor posteñe dichos textos y, en segundo lugar, porque el método ofrece a los alumnos la posibilidad de optimizar el proceso de aprendizaje de la traducción humana al detectar, describir y evaluar los aciertos y errores de la TA.

Para complementar esta metodología y con el fin de facilitar el análisis se propuso a los alumnos una tipología de errores ampliada (Wilss y Thome, 1984) atendiendo al criterio funcionalista del impacto del error en el texto y en el eventual destinatario de la traducción. Este criterio se rige por patrones lingüísticos y pragmáticos.

- 1) **Error léxico:** el término se mantiene en inglés en la traducción; sin embargo, por tratarse de un término transparente o motivado, no afecta el sentido general del texto; el programa de TA opta por una acepción estilísticamente poco adecuada para el contexto, pero no se altera gravemente el sentido de la oración.
- 2) **Error sintáctico:** si bien todos los términos y palabras están correctamente traducidos, la oración no respeta el orden sintáctico normal del español (traducción literal que no afecta el sentido de la oración).
- 3) **Error léxico-sintáctico:** se usa una acepción inadecuada de un término y, a su vez, este está inserto en una oración mal construida.
- 4) **Error léxico-semántico:** la acepción que ha entregado el programa de TA es incorrecta y por ello altera el sentido de la oración.
- 5) **Error sintáctico-semántico:** si bien las palabras y los términos están en general bien traducidos, el orden sintáctico es tan ajeno al genio del español que altera el sentido.
- 6) **Error léxico-sintáctico-semántico:** en este caso, el error se produce por la mala elección de una o más acepciones de términos, a lo cual se suma una sintaxis incorrecta en español, lo cual redundaría en un problema grave de sentido.

El corpus parcial de análisis está compuesto por dos textos que corresponden a dos tipologías textuales: un artículo semiespecializado sobre un instituto de biotecnología y un resumen especializado de una comunicación en un seminario de medio ambiente. Como complemento a los resultados obtenidos, se pretende comparar en términos generales el desempeño del programa de TA en ambos tipos de textos, si bien este objetivo excede los límites de este estudio.

En el análisis se presenta una selección de ejemplos del análisis de la TA en ambos textos originales, los cuales fueron traducidos al español con el programa de TA en línea Systran. Este programa se encuentra disponible en el sitio web de la empresa Systran, la cual viene desarrollando programas de TA desde la década del 70. Puesto que se trata de un programa en línea, no permite una preedición del texto original ni otro tipo de interactividad durante el proceso de TA. El usuario sólo puede posteditar el texto traducido una vez terminado el proceso de TA.

A continuación, se presenta la oración del texto original (TO), la traducción automática (TA) y luego el análisis de los aciertos y

errores para cada oración. También se clasifican los errores según la tipología del error propuesta anteriormente y se determinan las posibles causas de dichos errores. Con el fin de favorecer la progresión didáctica, los ejercicios de postedición se fueron graduando a medida que los alumnos avanzaban en el análisis del corpus completo de esta investigación, que está constituido por cinco textos². Por razones de espacio, se presentará sólo una selección de oraciones analizadas correspondientes a los dos tipos de textos que conforman el corpus parcial de esta investigación.

ANÁLISIS DEL CORPUS

Ejercicio I

Texto: EUROGENTEC IN FRESH WATER

Fuente: Wallonie/Bruxelles, agosto 1995.

Actividades de postedición:

1. Analice la TA siguiendo los pasos del método general de análisis propuesto en Diéguez y Riedemann 1998.
2. Destaque y describa los aciertos y errores de traducción que surgen por orden de aparición.
3. Determine las posibles causas de los errores de la TA.

(TO) *Eurogentec, founded in 1985 as a spin-off of the University of Liège, specialises in biotechnological products.*

(TA) Eurogentec, fundado en 1985 como efecto de la universidad de Lieja, se especializa en productos biotechnological.

ACIERTOS

1. Traduce correctamente el nombre de la ciudad (lexicón bien alimentado).
2. Oración sintácticamente correcta.

ERRORES

1. Acepción errónea de “spin-off” que en este contexto significa “entidad derivada”. Causa: presencia de guión.
2. Mal uso de minúsculas y mayúsculas. En este caso, “universidad” debe ir con mayúsculas en español, como figura en el texto original.
3. No tiene alimentado en su diccionario el término “biotechnological”. Causa: diccionario mal alimentado.

² Los otros tres textos que formaron parte de la investigación son un contrato legal de representación de ventas y dos artículos especializados, uno de minería y el otro de economía.

(TO) *Its market niche is the development of veterinary products from genetically modified micro-organisms.*

(TA) Su lugar de mercado es el desarrollo de productos veterinarios de microorganismos genéticamente modificados.

ACIERTOS

1. Reconocimiento del guión como signo que permite formar un término compuesto en inglés (“micro-organisms”) y omisión correcta de este signo en español.
2. Reconoce función adjetiva de la palabra “veterinary” y hace el cambio sintáctico en español.
3. Reconoce función adjetiva de la frase “genetically modified” y hace el cambio sintáctico en español.

ERRORES

1. Terminología poco precisa (“market niche”). Causa: lexicón mal alimentado.
2. Error sintáctico en “genetically” (usa adjetivo en español en lugar de adverbio).
Causa: algoritmo sintáctico deficiente.
3. Uso repetitivo de la preposición “de”.

(TO) *After having identified the protein factor responsible for this adaptation, the researchers at Eurogentec (in cooperation with the University of Liège and other foreign teams) have developed a protein, based on a genetically modified bacteria, with which young salmon are injected.*

(TA) Despues identificando el factor de la proteína responsable de esta adaptación, los investigadores en Eurogentec (en cooperación con la universidad de Lieja y de otros equipos no nativos) han desarrollado una proteína, basada en las bacterias genéticamente modificadas, con las cuales se inyectan los salmones jóvenes.

ACIERTOS

1. Sintaxis bastante aceptable para una oración relativamente compleja en inglés.
2. Respeto del paréntesis.
3. Reconoce función adjetiva de “protein”, “foreign” y “young”, y hace el cambio sintáctico en español.
4. Reconoce función adjetiva de la frase “genetically modified” y hace el cambio sintáctico en español.

ERRORES

1. No reconoce la forma verbal “having identified” y usa en español un gerundio en forma errónea. Causa: algoritmo sintáctico deficiente.
2. Mal uso de acepción para “foreing” (no nativo). Causa: lexicón mal alimentado.

3. No reconoce el artículo indeterminado singular de “bacteria” y lo traduce incorrectamente por el artículo determinado plural (las bacterias).
4. Error sintáctico de “genetically” (usa un adjetivo en lugar de un adverbio). Causa: algoritmo sintáctico deficiente.

(TO) *This osmoregulation factor (Smoltine) enables them to enter a marine environment without harm, because they are less stressed.*

(TA) Este factor del osmoregulation (Smoltine) les permite incorporar un ambiente marina sin daño, porque los tensionan menos.

ACIERTO

1. Reconoce función adjetiva de la palabra “marine” que modifica a “environment” y hace el cambio sintáctico en español.

ERRORES

1. El lexicón no cuenta con el término “osmoregulation”.
2. Error de régimen preposicional (incorporar a). Causa: algoritmo sintáctico deficiente.
3. Error de uso de verbo reflexivo y no reflexivo (incorporarse/estresarse). Causa: algoritmo sintáctico deficiente.
4. Error de concordancia de género (ambiente marina).

(TO) *Apart from the fish world, Eurogentec is developing a recombinant vaccine which protects against infection by the BVD virus (which is well known among cattle breeders), participates in various European research programmes (BRIDGE, FLAIR, etc.) and markets products intended for molecular biology laboratories.*

(TA) Aparte de el mundo de los pescados, Eurogentec está desarrollando una vacuna recombinant que proteja contra la infección por el virus de BVD (que es bien sabido entre los ganaderos), participa en los varios programas de investigación europeos (PUENTE, FLAIR, los etc.) y los productos de los mercados pensaron para los laboratorios moleculares de la biología.

ACIERTOS

1. Reconocimiento de paréntesis.
2. Reconoce función adjetiva de la sigla “BVD” y de las frases “European research” y “molecular biology” y hace el cambio sintáctico en español.

ERRORES

1. No tiene incorporado el término “recombinant” en su lexicón.
2. Acepción incorrecta del verbo “know” (usa “saber” en lugar de “conocer”). Causa: polisemia.
3. Uso incorrecto de artículo determinado antes de “etc.”.
4. Confunde el verbo “markets” en tercera persona singular con un sustantivo plural (mercados). Causa: homografía.

5. No reconoce la cláusula “intended for” y la traduce por el verbo “pensaron para” en lugar de “destinados a”. Causa: algoritmo sintáctico deficiente.

Ejercicio II

Texto: RESUMEN DE COMUNICACIÓN EN SEMINARIO SOBRE CONTAMINACIÓN ATMOSFÉRICA

Fuente: Seminario sobre Contaminación Atmosférica en Santiago y en el Mundo, Centro de Extensión, Pontificia Universidad Católica de Chile, 1996.

Actividades de postedición:

1. Analice la TA siguiendo los pasos del método general de análisis propuesto en Diéguez y Riedemann 1998.
2. Destaque y describa los aciertos y errores de la TA utilizando la tipología del error propuesta y priorice el tipo de error (del más grave o léxico-sintáctico-semántico al menos grave o léxico).
3. Determine las posibles causas de los errores de la TA.

(TO) *Source Assessment of Urban VOC: LPG Leakage versus Traffic, Results from the June 1996 Santiago Urban Study.*

(TA) Gravamen de la fuente de VOC urbano: Salida del LPG contra el tráfico, resultados del estudio urbano de junio de 1996 de Santiago.

ACIERTOS

1. Reconoce una frase adjetiva compleja (“June 1996 Santiago Urban”) que modifica al sustantivo “study” y hace el cambio sintáctico en español.
2. Reconoce función adjetiva de “Urban” antes de la sigla “VOC” y de la sigla “LPG” antes del término “Leakage” y hace el cambio sintáctico en español.

ERRORES

1. Mal uso de la acepción del sustantivo “assessment”, que en este contexto corresponde a “evaluación” (error léxico-semántico).
2. Mal uso de la acepción del sustantivo “leakage”, que en este contexto significa “filtración” (error léxico).
3. No traduce la sigla “VOC”, que corresponde a “COV” (compuestos orgánicos volátiles) en español (error léxico).
4. No traduce la sigla LPG, que corresponde a GLP (gas licuado de petróleo) (error léxico).
5. Mala traducción del término “versus” (error léxico).
6. Las causas de estos errores pueden ser dos: 1) lexicón pobre en cuanto a acepción de léxico y a inclusión de terminología específica sobre el tema. 2) No reconocimiento de contexto, lo cual le

impide optar por una acepción más adecuada (siempre y cuando estas estén alimentadas en el lexicón) (algoritmo sintáctico y semántico deficiente).

(TO) *Between June 1 and June 8, 1996, one hundred and forty four Santiago, Chile whole air samples were collected in evacuated 2-liter stainless steel canisters.*

(TA) Entre de junio el 1 y de junio cuatro Santiago el 8, 1996, ciento y cuarenta, las muestras de aire enteras de Chile fueron recogidos en cajas evacuadas del acero inoxidable 2-liter.

ERRORES

1. Oración, en general, muy mal traducida tanto desde el punto de vista sintáctico como semántico.
2. Si bien se trata de una frase adjetiva larga y muy compleja (“one hundred and fourty four Santiago, Chile whole air”) la cual modifica a “samples”, el programa no la reconoce y comete serios errores (error léxico-sintáctico-semántico). La coma entre “Santiago” y “Chile” puede ser la causa del error, y además un algoritmo sintáctico deficiente.
3. No reconoce la frase adjetiva “evacuated 2-liter stainless steel” que modifica a “canisters” (error léxico-sintáctico-semántico). En este caso, la presencia del guión podría ser una de las causas del error, y además un algoritmo sintáctico deficiente.
4. No reconoce fechas como “June 1” y “June 8, 1996” (error sintáctico-semántico). Causa: algoritmo sintáctico deficiente.
5. No reconoce el numeral “hundred and fourty four” (error sintáctico-semántico). Causa: lexicón mal alimentado.
6. Mala acepción de la palabra “evacuated”, que en este contexto significa “al vacío” (error léxico-semántico). Causa: lexicón mal alimentado.
7. No traduce la palabra “liter” (error léxico). Causa: lexicón mal alimentado y presencia de guión.
8. No hay concordancia de género (muestras - recogidos) (error sintáctico).

(TO) *Santiago air exhibits extremely high concentrations for many hydrocarbon and some halocarbon species when compared to the background sample taken off the Pacific coast near San Antonio, Chile.*

(TA) Las concentraciones extremadamente altas de los objetos expuestos del aire de Santiago para muchos el hidrocarburo y una cierta especie del halocarbon cuando estaba comparado a la muestra del fondo sacaron la costa pacífica cerca de San Antonio, Chile.

ACIERTO

1. Reconoce función adjetiva de la palabra “halocarbon” y hace el cambio sintáctico en español.

ERRORES

1. No reconoce la función verbal de la palabra “exhibits” y con ello altera seriamente el orden sintáctico de la oración (error sintáctico-semántico). Causa: homografía.
2. Mala traducción del régimen preposicional en el caso de “taken off”, que en este contexto significa “extraer desde un lugar”. Falta la preposición, que exige el verbo de la oración (error sintáctico-semántico). Causa: algoritmo sintáctico deficiente.
3. Mala acepción de la palabra “background sample”, que en este contexto significa “muestra de control” (error léxico-semántico). Causa: al parecer el diccionario no contiene terminología especializada como, por ejemplo, términos compuestos.
4. Traducción literal de la expresión “when compared to” (error sintáctico). Causa: algoritmo sintáctico deficiente.
5. Mala traducción de la preposición “for” en la cláusula “for many hydrocarbon and some halocarbon”, que en este contexto sería “de” y no “para” (error sintáctico). Causa: algoritmo sintáctico deficiente.

(TO) *Many of these elevated mixing ratios can be attributed to leakages of liquefied petroleum gas (LPG) and industrial stockpile, and incomplete combustion from vehicular usage.*

(TA) Muchos de éstos las relaciones de transformación que se mezclan elevadas se pueden atribuir a las salidas del gas licuefecho del petróleo (LPG) y de la reserva industrial, y a la combustión incompleta del uso de vehículos.

ACIERTO

1. Reconoce función adjetiva de “industrial”, “incomplete” y “vehicular” y hace el cambio sintáctico en español.

ERRORES

1. No reconoce la frase adjetiva antepuesta al sustantivo “ratios” (error sintáctico-semántico). Causa: algoritmo sintáctico deficiente.
2. Confunde el artículo definido plural “these” con un pronombre (error sintáctico-semántico). Causa: homografía.
3. No traduce correctamente el término “liquefied petroleum gas” (error léxico-semántico). Causa: el diccionario no posee terminología especializada como, por ejemplo, términos compuestos.
4. Acepción incorrecta de “stockpile”, que en este contexto significa “material acumulado” (error léxico). Causa: diccionario mal alimentado.

(TO) *The concentrations of propane, i-butane and n-butane showed little change between 5 a.m. and 9 a.m., indicating the contribution from leakage was much greater than vehicular usage.*

(TA) Las concentraciones del propano, yo-butano y n-butano mostraron pequeño el cambio entre 5 mañanas y 9 mañanas, indicar la contribución de la salida era mucho uso de vehículos mayor que.

ERRORES

1. No reconoce la expresión “much greater than” y con ello altera gravemente la sintaxis de la oración (error sintáctico-semántico). Causa: algoritmo sintáctico deficiente.
2. Mala traducción de la forma *-ing* en “indicating”, que en este caso tiene función adverbial y se puede traducir por el gerundio español “indicando que” (error sintáctico-semántico). Causa: algoritmo sintáctico deficiente.
3. Traducción literal del segmento “showed little change” (error sintáctico-semántico). Causa: algoritmo sintáctico deficiente.
4. Confunde la “i” minúscula con el pronombre “I” y lo traduce como “yo-butano” (error sintáctico-semántico). Causa: lexicón mal alimentado (terminología).
5. Mal uso del artículo determinado antes de “propano” y de “cambio”.
6. Repite el error con las horas y con la palabra “leakage”.

CONCLUSIONES

Si bien en este artículo se presenta una muestra de solo dos de los cinco textos que conforman el corpus total de análisis, hemos considerado pertinente exponer aquí los resultados que se desprenden del análisis del corpus original entero.

Aciertos más recurrentes

El análisis del corpus entero permite en primer lugar indicar como los aciertos más recurrentes de la TA los siguientes:

- reconocimiento de la función adjetiva de palabras y frases más extensas antepuestas al sustantivo que modifican y adecuación al orden sintáctico del texto meta
- acepción correcta del léxico
- uso de terminología correcta
- reconocimiento del guión y del paréntesis
- reconocimiento de la forma verbal compuesta separada en inglés por un segmento de la oración y de la categoría gramatical de algunas formas *-ing*
- en general, reconoce también nombres propios y direcciones.

Errores más recurrentes

En cuanto a los errores más recurrentes de la TA podemos mencionar los siguientes:

- uso incorrecto de artículos determinados e indeterminados (elección incorrecta, eliminación o uso inadecuados)
- orden sintáctico incorrecto (adjetivo-sustantivo, frase adjetiva-sustantivo, orden incorrecto de las palabras en la oración)
- acepción errónea del léxico o terminología en un contexto específico
- no traduce léxico y terminología
- no reconoce y traduce erróneamente términos compuestos (terminología y fraseología)
- no reconoce sintagmas que conforman unidades cuyo sentido va más allá de la suma de los sentidos individuales de las palabras que los componen (en estos casos suele hacer una traducción literal: “phrasal verbs” o verbos compuestos verbo+preposición, expresiones fijas, etc.)
- preposiciones mal traducidas (elección incorrecta, eliminación o uso inadecuados)
- traducción incorrecta de formas verbales (tiempo, modo, aspecto, reflexivo/irreflexivo)
- mala traducción de régimen preposicional (verbo+preposición)
- cambio de categoría gramatical
- traducción incorrecta de nombres propios
- traducción incorrecta de pronombres, de formas -ing (categoría gramatical) y de siglas
- concordancia incorrecta de género y número
- no reconoce paréntesis ni guión
- no reconoce función adjetiva de una palabra antepuesta al sustantivo que modifica
- no reconoce una frase adjetiva de más de dos palabras antepuestas al sustantivo que modifica
- no reconoce, en general, fechas, numerales ni horas.

Posibles causas de los errores

Antes de comentar las posibles causas de los errores, quisiéramos aclarar que ello resulta bastante difícil para el traductor que realiza una postedición, puesto que en este caso él se ve enfrentado a lo que se suele llamar una evaluación de “caja negra” del programa de TA. Este tipo de evaluación se caracteriza por limitarse al acceso de aquellos mecanismos del programa que operan, con lo que entra o

sale del programa. En otras palabras, se evalúa el programa solo a través de su funcionamiento. Por lo tanto, como no se puede acceder a todos los mecanismos del mismo, lo que sería una evaluación de “caja de cristal” reservada para los que crean y desarrollan los programas de TA, las afirmaciones sobre las posibles causas de los errores sólo se pueden plantear a nivel de hipótesis, pues no se pueden comprobar empíricamente.

Algunas de las posibles causas de los errores más recurrentes pueden ser las siguientes:

- presencia de término compuesto unido con guión (no reconoce el guión)
- el lexema o término no figura en lexicón (lexicón mal alimentado)
- presencia de sintagma de dos o más términos, uno de los cuales está unido con guión (algoritmo sintáctico deficiente)
- presencia de forma verbal compuesta separada por parte de la oración (algoritmo sintáctico deficiente)
- elipsis del verbo principal o de lexemas (algoritmo sintáctico deficiente)
- ambigüedad semántica, polisemia, homografía y fraseología
- uso de lenguaje figurado
- espacios en blanco
- ausencia de conectores
- listado de términos descontextualizados
- algoritmos sintáctico y semántico deficientes.

Antes de concluir, quisiéramos destacar que si bien no se ha realizado un estudio estadístico cuantitativo de los errores presentes en cada texto del corpus parcial, sí podemos señalar que se observa un margen mayor de error en el resumen especializado sobre medio ambiente que en el texto semiespecializado sobre el instituto de biotecnología. Creemos que este hecho refleja una vez más los resultados observados en investigaciones anteriores (véanse Diéguez y Riedemann, 1998 y Diéguez y Cabrera, 1998), los cuales demuestran que, ante la presencia de oraciones relativamente breves y de sintaxis simple, el programa de TA reduce considerablemente su margen de error. Esto se puede observar, en general, en las oraciones analizadas en el texto sobre biotecnología.

Por otra parte, frente a oraciones sintácticamente condensadas, el margen de error es mucho mayor, como se puede observar en el resumen sobre medio ambiente. Por razones de economía de lenguaje, en este último texto abundan, por ejemplo, las oraciones con secuencias de dos o más lexemas yuxtapuestos y antepuestos al

sustantivo que modifican, construcciones que resultan muy difíciles de traducir para el programa de TA, debido probablemente a la falta de algoritmos sintácticos eficientes. Ejemplos de estas estructuras son: "...one hundred and fourty four Santiago, Chile whole air..." que modifica a "samples"; "...evacuated 2-liter stainless steel..." que modifica a "canisters" y "...elevated mixing..." que modifica a "ratios".

Asimismo, en esta investigación hemos comprobado una vez más que el programa de TA comete errores ante algunas formas *-ing* ("having identified" e "indicating") y frente a homógrafos ("markets" y "exhibits"), lo que una vez más demuestra la deficiencia de los algoritmos sintácticos. Podríamos concluir entonces que, contrariamente a lo que se suele señalar, el programa de TA se desempeña mejor ante un texto semiespecializado como el del instituto de biotecnología que ante un texto altamente especializado como el resumen de medio ambiente, dado que en este último se presentan estructuras sintácticas más complejas.

Por todo lo anterior, podemos afirmar que en este estudio se confirma nuestra hipótesis inicial en el sentido de que los ejercicios de postedición de textos traducidos mediante programas de TA pueden constituir una metodología válida de enseñanza-aprendizaje de la TH. Nuestro objetivo en esta investigación ha sido, como ya se señaló, aplicar la TA al proceso de enseñanza-aprendizaje de la TH y con ello ofrecer nuevas estrategias didácticas para su enseñanza-aprendizaje.

Estamos conscientes de que se trata de un tema complejo y que puede ser profundizado aún más. De hecho, la gran variedad de tipos de errores presentes en la TA permite profundizar este enfoque en muchos niveles lingüísticos como el sistema verbal, adverbial, preposicional, de conectores, entre otros. Esperamos que este aporte pueda servir de ayuda a todos los interesados en este tema y sea el inicio de una serie de investigaciones futuras en materia de didáctica de la traducción humana.

BIBLIOGRAFÍA

- ARNOLD, D. et al. (1995). *Machine Translation: An Introductory Guide*, [en línea] (URL: http://www.georgetown.edu/cball/ling361/ling361_mt_links.html)
- ARNTZ, R. (1988). *Textlinguistik und Fachsprache*, AILA Symposium, Hildesheim: Olms Verlag.
- BALKAN, L. et al. (1997). *Tools and Technics for Machine Translation Teaching: a Survey*, [en línea] (balka@essex.ac.uk)

- BLATT, A. et al. (1985). *Computer und Übersetzen. Eine Einführung*, Hildesheim: Olms Verlag.
- BÜHLER, A. et al. (1965). *Sprachtheorie. Die Darstellungsfunktion der Sprache*, Stuttgart: Gustav Fischer.
- Diccionario de la Lengua Española*, (1992). Real Academia Española, 21^a. edición.
- DIÉGUEZ M.I. & K. RIEDEMANN (1998). “Análisis del error en la traducción automática: algunos ejemplos de las formas -ing del inglés al español”, en: *Onomazein*, Vol. 3, Instituto de Letras, PUC de Chile, pp. 211-229.
- _____. (1999) “La traducción de metáforas: ¿Un acto de rebeldía permanente?”, en: *Onomazein*, Vol. 4, Instituto de Letras, PUC de Chile, pp. 345-369.
- DIÉGUEZ, M.I. & I. CABRERA. (1996). “Estado de avance del Proyecto de investigación FONDECYT N° 1960441”, V Simposio Iberoamericano de Terminología, Ciudad de México [en línea] (<http://www.unilat.org/dtil/MEXICO/dieguez.html>)
- _____. (1997). “Traducción humana o traducción automática : variables para una toma de decisiones eficiente”, *Onomazein*, Vol. 2, Instituto de Letras, PUC de Chile, pp. 227-242.
- _____. (1998). “Traducción automática y traducción mixta: resultados de un estudio comparativo”, *Onomazein*, Vol. 3, Instituto de Letras, PUC de Chile, pp. 305-314.
- ESPI, R. (1999). *El tratamiento de la ambigüedad léxico-sintáctica en la traducción automática del español al inglés*, Tesis en opción al grado científico de Doctor en Ciencias Filológicas, Facultad de Lenguas, Universidad de La Habana, Cuba.
- GUTIÉRREZ A., M. L. (1978). *Estructuras sintácticas del español actual*, Madrid: Sociedad General Española de Librería, S.A.
- HALLIDAY, M.A.K. (1990). *Spoken and Written Language*, Oxford, U.K.
- HARVEY, A. (1994). “La comunicación de resultados de investigación: Un estudio comparativo de un discurso académico y de prensa”, en: *Lenguas Modernas* 21, pp. 59-79.
- HUTCHINS, W. J. & H. L. SOMERS (1995). *Introducción a la traducción automática*, Madrid: Visor Distribuciones S.A.
- LEHRBERGER, J. & L. BOURBEAU (1988). *Machine Translation. Linguistic Characteristics of MT Systems and General Methodology of Evaluation*, Amsterdam/Philadelphia: John Benjamins.
- NAGAO, M. (1989). *Machine Translation: How far can it go?*, Oxford University Press.
- NEUBERT, A. (1991). *Computer-Aided Translation: Where are the Problems?*, Amsterdam: John Benjamins.
- NORD, Ch. (1988). *Textanalyse und Übersetzen*, Heidelberg: Julius Groos Verlag.
- SYSTRAN. Programa de traducción automática en línea Systran (URL: <http://www.onlinetrans.com/systran.cgi>)
- QUIRK, R. & S. GREENBAUM (1977). *A University Grammar of English*, London: Longman Group Limited.
- REISS, K. & H. VERMEER (1984). *Grundlegung einer allgemeinen Translationstheorie*, Tübingen: Niemeyer Verlag.
- REISS, K. (1983). *Texttyp und Übersetzungsmethode. Der operative Text*, Heidelberg: Julius Groos Verlag.

- _____. (1986). *Möglichkeiten und Grenzen der Übersetzungskritik*, München: Max Hueber Verlag.
- SAGER, J. C. (1992). *La industria de la lengua, la lingüística computacional, los trabajos del UMIST, la traducción especializada y su técnica*, Barcelona: Servei de Llengua Catalana, Universitat de Barcelona.
- _____. (1993). *Language Engineering and Translation. Consequences of Automation*, Amsterdam/Philadelphia: Benjamins Translation Library.
- SOLÉ, Y. & C. SOLÉ (1977). *Modern Spanish Syntax. A Study in Contrast*, Austin: University of Texas.
- STACHOWITZ, R. (1973). *Voraussetzungen für maschinelle Übersetzung. Probleme, Lösungen, Aussichten*, Frankfurt am Main: Athenäum Verlag.
- SWAN, M. (1995). *Practical English Usage*, Oxford: Oxford University Press.
- THOMPSON, A. J. & A. V. MARTINET (1980). *A Practical English Grammar*, Oxford: Oxford University Press.
- VASCONCELLOS, M. & M. LEÓN. "SPANAM and ENGSPAN: Machine Translation at the Pan American Health Organization", en: *Machine Translation Systems*, Slocum, J. (ed.), Cambridge University Press, s/f.
- _____. "The Advantages and Limitations of Machine Translation", PAHO, Washington, s/f.
- _____. (1986). "Functional Considerations in the Postediting of Machine-Translated Output", en: *Computers and Translation* 1, Paradigm Press.
- _____. (1986). "Post-editing on-screen: Machine Translation from Spanish into English. Chief, Terminology and Machine Translation Program", en: *Translating and the Computer* 8, Picken, C. (ed), London: CBI Conference Centre.
- _____. (1987). "The Contribution of Machine Translation: Present Status and Future Perspectives", en: *Microcomputer Applications in Education and Training for Developing Countries*, London: Westview Press.
- VASCONCELLOS, M. et al. (1987). "Factors in the Evaluation of MT: formal vs. functional Approaches", Washington: PAHO.
- _____. "Machine Translation and the Language Barrier", Washington: PAHO, s/f.
- WILSS, W. & G. THOME (Hrsg.) (1984). *Die Theorie des Übersetzens und ihr Aufschlusswert für die Übersetzungs- und Dolmetschdidaktik*. Tübingen: Gunter Narr Verlag.
- WILSS, W. (1988). *Kognition und Übersetzen*, Tübingen: Niemeyer Verlag.
- _____. (1996). *Knowledge and Skills in Translation Behavior*, Amsterdam/Philadelphia: Benjamins Translation Library.
- WILSS, W. (1982). *The Science of Translation. Problems and Methods*, Tübingen: Gunter Narr Verlag.