

ANÁLISIS DEL ERROR EN LA TRADUCCIÓN AUTOMÁTICA: ALGUNOS EJEMPLOS DE LAS FORMAS *-ING* DEL INGLÉS AL ESPAÑOL

M. Isabel Diéguez M.

Pontificia Universidad Católica de Chile

Karin Riedemann H.

Pontificia Universidad Católica de Chile

Universidad Tecnológica Vicente Pérez Rosales

Resumen

En el presente artículo se expondrá un estudio contrastivo de las formas *-ing* presentes en textos originales en inglés y sus respectivas traducciones al español realizadas con un programa de traducción automática (TA). Para ello se utilizaron como corpus de análisis dos textos originales en inglés, uno del área de ciencias sociales (economía) y el otro de las ciencias exactas (minería). A través del análisis de los errores y aciertos de las traducciones realizadas mediante el programa de TA se pretende en último término extraer conclusiones prácticas para la enseñanza de la traducción técnico-científica.

Abstract

A contrastive study of -ing forms in English and Spanish is presented. For such analysis, two original English texts and their respective Spanish machine translations (MT) were used. One original text has been taken from the social sciences (economy) and the other, from the technical-scientific field (mining). By comparing the machine translations of the -ing forms present in the original texts, conclusions may be drawn regarding the teaching of professional translation.

INTRODUCCIÓN

En los últimos años, en Chile se ha podido apreciar un aumento considerable de los volúmenes de traducción de textos científicos y técnicos, fundamentalmente gracias al desarrollo económico que se

ha alcanzado. Como país que aspira al progreso, Chile requiere acceder en forma urgente a información de diversa índole tal como tecnología de punta, *know-how* y avances científicos. A ello se suma el hecho de encontrarnos en un momento clave de integración a iniciativas de desarrollo económico como Mercosur, Nafta, Apec, etc., las que exigen que se optimice todo tipo de actividades, entre las cuales estimamos que la traducción juega y jugará en el futuro un papel fundamental.

Si bien existen iniciativas como la creación de nuevas oficinas de traducción y una mayor oferta de carreras de traductor profesional a nivel universitario, las cuales han contribuido a dar alguna respuesta a la alta demanda de traducciones, persisten algunos problemas tales como la precaria calidad de ciertas traducciones, el costo y el tiempo que se invierte en traducir, sin mencionar el enorme porcentaje de literatura especializada que queda sin traducirse.

En las últimas décadas, el crecimiento sostenido del mercado de la traducción, hecho que afecta no sólo a Chile sino también al resto del mundo, ha llevado a algunos países desarrollados a investigar e implementar programas de traducción automática (TA) como una posible solución al aumento de la demanda de traducciones. Para ilustrar esto cabe destacar que en Canadá, país eminentemente bilingüe, se utiliza en la televisión un programa computacional de TA diseñado en 1978 por la Universidad de Montreal, llamado TAUM METEO. Este programa permite traducir del inglés al francés los informes meteorológicos, y sus resultados son bastante satisfactorios si tomamos en consideración que los textos de estos informes se caracterizan por una fraseología estandarizada y por oraciones relativamente breves.

En términos generales, existen diferentes programas de TA:

- 1) sistemas autónomos que requieren una edición posterior del texto traducido a cargo de un traductor humano;
- 2) sistemas interactivos, ya sea de TA asistida por el usuario o traducción manual asistida por el computador, y
- 3) sistemas más independientes, entre los cuales podemos destacar los sistemas de gestión de bases de datos terminológicos, los programas de vaciado automático de términos y los programas de TA dotados de un módulo de terminología.

Gracias a la investigación activa en materia de TA y al desarrollo de la inteligencia artificial, en los últimos años los programas de TA existentes en el mercado han mejorado ostensiblemente en comparación con la calidad de los primeros programas. En cuanto a los parámetros más estudiados para evaluar estos programas, que son el

tiempo, el costo y la calidad, existe un volumen considerable de literatura especializada. Respecto al primer parámetro, la literatura más reciente demuestra que los programas de TA, en términos generales, ayudarían al traductor a rendir más en menos tiempo. Respecto del parámetro costo, la mayoría de los estudios realizados se muestran más bien cautos. Vasconcellos destaca que para estimar los costos de aplicación de un método de traducción automática son numerosos los aspectos que deben considerarse, por ejemplo, la inversión inicial en el programa, la mantención del sistema, la constante actualización y alimentación de los glosarios y el trabajo de postedición, entre otros.

En lo referente al parámetro calidad, se han hecho estudios comparativos de diferentes programas de TA cuyos resultados no son del todo alentadores en comparación con la calidad que presentan las traducciones hechas por el traductor humano. Por otra parte, algunos investigadores reconocen que la calidad del producto de un programa de TA depende fundamentalmente de variables tales como la cantidad y grado de confiabilidad de la información terminológica que el programa contenga. En otras palabras, reconocen que la calidad aumenta en la medida en que el programa cuente con un glosario general y con varios microglosarios especializados según áreas temáticas. No obstante, cabe destacar que este parámetro es el más difícil de medir, pues el evaluador debe contar con criterios claros y realistas que permitan una evaluación confiable.

Si damos un vistazo a nuestra realidad, podemos constatar que los traductores, con miras a optimizar el quehacer de la traducción, emplean, cada vez más, programas de TA para realizar su trabajo. En este punto, se presenta la siguiente situación: hemos podido observar que el traductor chileno utiliza en forma indiscriminada estos programas de TA, puesto que los aplica a cualquier tipo de texto y a cualquier área temática. Planteamos entonces en este estudio la hipótesis de que la traducción automática dista aún mucho de ser una solución aceptable para contribuir al trabajo del traductor profesional. Si bien a nivel terminológico se ha avanzado bastante en las técnicas de alimentación de diccionarios temáticos que mejoran ostensiblemente la calidad de la TA, el soporte sintáctico-semántico de dichos programas es aún muy precario. A este respecto y a pesar de los avances de la inteligencia artificial, lo que más resalta es la incapacidad del programa de TA de reconocer contexto, condición fundamental para discernir tanto la carga semántica como las categorías gramaticales del léxico del texto original y de su traducción.

Por lo tanto, el objetivo del presente trabajo será evaluar la calidad de la traducción automática (TA) mediante un estudio contrastivo de las formas *-ing* presentes en textos originales en inglés

y sus respectivas traducciones al español realizadas por un programa de TA disponible en el mercado nacional. Mediante el análisis de los errores y aciertos de este programa, en el marco de la traducción automática de las formas *-ing* del inglés al español, se pretende en último término extraer conclusiones prácticas para la enseñanza de la traducción técnico-científica.

Este estudio se realizará sobre la base de dos textos: uno perteneciente a las ciencias exactas, específicamente un informe técnico del área de minería, y el otro, un artículo especializado del área de economía perteneciente a las ciencias sociales. Se pretende con ello analizar contrastivamente en diferentes áreas temáticas y tipos de textos los aciertos y errores de la traducción automática, centrándonos para ello en las formas *-ing* presentes en el original y la traducción respectiva. Creemos necesario destacar lo novedoso que resulta un trabajo de esta naturaleza, dado que en Chile el tema de la traducción automática es prácticamente desconocido, y por ello aspiramos a que este estudio sea un aporte más a lo que viene realizando desde 1995 el Programa de Traducción de nuestra Universidad en este tema.

MARCO TEÓRICO

A la luz de la literatura estudiada, podemos afirmar que la traducción automática se ha visualizado en los últimos años como un método informático nuevo que puede contribuir al trabajo del traductor moderno. En 1988 Hutchins dio una breve definición de TA: es la traducción generada por un computador con o sin interferencia humana. Los mecanismos básicos de la TA son el texto de entrada, los diccionarios electrónicos y el programa que desarrolla el aspecto lógico o algoritmos.

La interacción hombre-máquina puede ser de tres tipos: preedición, edición interactiva y postedición. En la fase de preedición, el traductor revisa el texto que se va a traducir antes de que éste ingrese a la máquina con el fin de resolver ambigüedades que puedan existir en el texto. La edición interactiva consiste en realizar la traducción de manera tal que durante el proceso la máquina entregue opciones al traductor para que éste determine la mejor solución en cada caso. Finalmente, la postedición consiste en “corregir” el producto resultante del programa de TA, y es una de las fases más usadas en los diferentes programas que existen en el mercado.

En cuanto a los mecanismos generales de un programa de TA, éstos se ordenan en tres fases principales. La primera es la fase de análisis, en la cual se analiza el texto fuente tanto a nivel morfológico,

sintáctico como semántico en la lengua original. La transferencia corresponde a la segunda fase, llamada también intermedia, y consiste en transformar la estructura obtenida en la fase anterior en otra similar en la lengua de llegada. Finalmente, la fase de generación o síntesis permite pasar, de la representación formal transformada en la segunda fase, a una oración correcta en la lengua de llegada.

Los componentes lingüísticos de un sistema de TA, según Lehrberger y Bourbeau, se dan a nivel léxico, morfológico, sintáctico y semántico. El componente léxico está dado por el o los diccionarios que posee el sistema y se puede tratar de diccionarios unilingües y bilingües. A su vez, en una entrada léxica se puede contemplar una serie de informaciones como la categoría gramatical, el campo semántico, las reglas sintácticas, entre otros. Además, en un diccionario se pueden incorporar unidades semánticas o modismos.

En cuanto al aspecto morfológico, podemos destacar que la forma escrita de una entrada léxica constituye la clave de acceso para ingresar al diccionario. Para evitar crear una entrada en el diccionario para cada forma de una palabra, se introduce la forma básica de la palabra y se usa un componente morfológico para establecer relaciones entre las diferentes formas que una palabra adopta en un determinado texto y la forma base de la palabra que aparece en el diccionario. En el caso de los verbos, se suele introducir el infinitivo como forma base del diccionario. El análisis morfológico se desarrolla a nivel de flexiones o conjugaciones de las palabras, a nivel de derivaciones, y finalmente a nivel de composición. A modo de ejemplo, Vasconcellos destaca que en el programa SPANAM, que se utiliza en la Organización Panamericana de la Salud, el análisis morfológico permite reconocer la persona, el número y el tiempo de los verbos.

El componente sintáctico de un programa de TA tiene como función analizar las oraciones y sus partes constituyentes en la lengua fuente y generar oraciones sintácticamente correctas en la lengua meta. Puede tratarse de oraciones simples o complejas. Por ejemplo, una frase sustantiva al interior de una oración compleja obliga al traductor automático a identificar las valencias del verbo puesto que esa es la única forma de segmentar la oración y poder representar su estructura en forma adecuada.

El componente semántico es el que menos ha sido desarrollado, por la complejidad que reviste, pero ninguno de los expertos en TA niega su gran importancia en el futuro desarrollo de esta disciplina. El análisis semántico se puede dar a nivel de palabra, cuyo significado se puede encontrar en el o los diccionarios del programa. Sin embargo, suele ocurrir que la entrada de la palabra en el diccionario contiene más información morfológica y sintáctica que semántica. Para que la entrada tenga más información semántica, se le pueden

asignar rasgos semánticos a la palabra, como por ejemplo: concreto, abstracto, objeto físico, acción, etc. La homografía también debe resolverse con datos semánticos que permitan distinguir por ejemplo la palabra “walk” como sustantivo o verbo. En el caso de los verbos, se suelen asignar las categorías semánticas de causativo, factivo, incoativo, entre otras.

A modo de resumen, podemos destacar algunas de las técnicas lingüísticas más empleadas en el ámbito de la TA, las cuales apuntan al análisis morfológico, a la resolución de la polisemia, a rasgos sintácticos y semánticos, y a la síntesis del verbo en español. Dentro del análisis morfológico, el algoritmo permite reconocer las terminaciones del plural y el femenino de sustantivos y pronombres y adjetivos, y las terminaciones de persona, número y tiempo de los verbos. En cuanto a los aspectos sintácticos y semánticos, en el caso de los verbos se especifica su número de objetos o valencia, sus complementos y si posee una forma infinitiva marcada o no, entre otros. Por último, en el caso de la síntesis del verbo español se registran todos los verbos regulares e irregulares, el tiempo, el modo y la persona. Se ingresa la raíz de los verbos y sólo se ingresa la forma completa en el caso de los verbos auxiliares.

Ahora bien, para acotar el tema objeto de esta investigación se presentará una contrastación de las funciones gramaticales más frecuentes que pueden adoptar las formas *-ing* del inglés y sus posibilidades de traducción al español desde la perspectiva de la traducción humana. Para ello nos hemos basado en varias gramáticas (Solé y Solé 1977; Quirk y Greenbaum 1978; Close 1975; Thompson y Martinet 1980; Swan 1995; Quirk et al. 1985) y la experiencia de las autoras en la enseñanza práctica de la traducción. Hemos optado en este trabajo por el término más general, formas o construcciones *-ing*, puesto que al consultar una serie de gramáticas descubrimos que los autores no coinciden en la terminología que emplean para referirse en forma más específica a las diferentes funciones de las formas *-ing*. En todo caso, nos adscribimos al planteamiento de Swan, quien señala que en general en inglés se habla de *gerund* cuando se remite a la función nominal de las formas *-ing*, y se habla de *present participle* cuando las formas *-ing* denotan una función verbal, adjetiva y adverbial. Por ello hemos considerado pertinente para este estudio presentar el siguiente paradigma contrastivo de las formas *-ing* del inglés.

1) Paradigma verbal

1.1 Verbo con aspecto progresivo: *More and more underground mining is being conducted using the longwall method*. Se tradu-

ce por la forma verbal equivalente en español ser + participio, que indica por una parte la voz pasiva y tiene además un carácter dinámico (está siendo realizado/se está realizando).

- 1.2 Verbo con función adverbial: *More and more underground mining is being conducted **using** the longwall method.* (usando/mediante el uso) Este es el único caso en que se traduce por el gerundio español. Según la RAE 1992, el gerundio en español comunica a la acción verbal carácter durativo, puede referirse a cualquier tiempo, género y número, según el sentido de la frase del que forma parte, y agrega que tiene generalmente carácter adverbial y puede expresar modo, condición, motivo o circunstancia. En definitiva, se trata, a juicio de Solé y Solé, de la construcción -ando/-iendo cuyo carácter adverbial remite a una acción secundaria que coexiste o es anterior a la acción principal; vale decir, no puede ser nunca una acción subsecuente respecto de la acción principal. Cabe destacar que este uso adverbial de las formas *-ing* en inglés es menos productivo que el gerundio en español. En inglés se suelen preferir frases adverbiales introducidas por *while*, *when*, o cláusulas adverbiales encabezadas por *because*, *if*, *as*, etc.

2) Paradigma nominal:

- 2.1 Sujeto de la oración: *More and more underground **mining** is being conducted using the longwall method.* Se traduce por un sustantivo (la minería) o por un sustantivo + adjetivo (la explotación minera).
- 2.2 Objeto del verbo: *If possible, the areas of the mine that will have secondary recovery of pillars or longwall panels should avoid **mining** beneath perennial streams in areas of thin oberburden.* Al igual que en el caso anterior, se traduce por un sustantivo (minería) o por un sustantivo + adjetivo (explotación minera).
- 2.3 Objeto de una preposición: *Determining the impacts from **mining** on surface and ground water.* En este caso se debe utilizar un sustantivo (de la minería), un sustantivo + adjetivo (explotación minera) o una cláusula, como se ilustra en el ejemplo siguiente: *I am looking forward to seeing you* (tengo ganas de verte). En algunos casos también puede traducirse por un infinitivo.

3. Paradigma atributivo

Puede estar representado por un lexema simple o una secuencia de lexemas (dos o más lexemas independientes o unidos por guión o barra).

Mining regulations have, for a number of years, required that mining companies collect hydrologic data in and around the area to be mined. Se traduce por el adjetivo correspondiente (mineras) o por cláusulas adjetivas en el caso de secuencias de lexemas.

METODOLOGÍA

Para realizar este estudio contrastivo se utilizaron como corpus dos textos originales en inglés, uno proveniente del área minera (técnico-científico) y el otro de economía (área de ciencias sociales). Cabe destacar que ambos textos forman parte de un corpus más amplio de estudio en el marco del proyecto Fondecyt N° 1960441, titulado “Traducción automática versus traducción humana: variables que inciden en la elección de uno u otro método de traducción con miras a optimizar el tiempo, costo y calidad de la traducción”, coordinado por una de las autoras del presente estudio.

En cuanto al programa de TA empleado en esta investigación, debemos aclarar que se trata de un programa disponible en el mercado nacional cuya característica principal se adscribe al tipo de traducción automática asistida por el usuario. Esto significa que produce traducciones en borrador de textos que se preeditan y posteriormente se posteditan hasta llegar a su forma final. El programa trabaja a nivel oracional, vale decir, analiza cada oración en el documento original, busca las palabras y frases respectivas en sus diccionarios y traduce la oración completa como un todo. Mediante este programa se pueden traducir diferentes tipos de documentos pertenecientes a una variedad de áreas temáticas.

Para efectos de este estudio, las autoras destacan que no se pretende abarcar aquí todos los tópicos relacionados con el amplio y complejo tema de la evaluación de la TA, sino que decidieron reducir el objeto de esta investigación al ámbito de la evaluación lingüística del producto de la TA realizada por el usuario/traductor del programa. Esto último significa que el evaluador no tiene acceso a los componentes internos del programa de TA. En palabras de Hutchins y Somers, se trata en definitiva de una evaluación de “caja negra” en la cual sólo se pueden observar los resultados del proceso de TA. Dado que aun así el tema resultaba extremadamente amplio, se optó por realizar una evaluación lingüística de las formas *-ing* del inglés y sus equivalencias al español arrojadas por el programa de TA. Cabe agregar que se optó por el análisis de estas construcciones dado que en la traducción humana constituyen un ámbito que ofrece gran dificultad tanto a los alumnos de traducción como a los traductores profesionales por la diversidad de posibilidades de traducción que

ofrece esta forma tan productiva en la lengua inglesa. Precisamente la diversidad de funciones gramaticales a que pueden remitir las formas *-ing* en inglés conduce en muchas ocasiones a errores de sentido en las traducciones al español. Se pretende por tanto que el análisis de los errores cometidos por el programa de TA en este ámbito permita dar luces para la didáctica de la traducción humana.

A continuación, se analizarán algunas formas *-ing* del inglés presentes en cada corpus y las traducciones al español que entregó el programa de TA tomando como criterios tanto la frecuencia de errores como el tipo de error presente en el texto traducido. Para efectos de este análisis, nos basaremos en el paradigma de formas *-ing* planteado en el marco teórico del presente estudio. Se analizarán los aciertos y errores más recurrentes referidos a cada función gramatical de las formas *-ing* presentes en el corpus. Cabe hacer la salvedad de que se empleó un criterio bastante amplio de corrección, pues se tomó en consideración sólo la aparición del equivalente en español de la forma *-ing*, sin considerar en toda su amplitud el contexto oracional. Ello se debió fundamentalmente al hecho de que prácticamente todas las oraciones del texto traducido por el programa de TA presentaban errores sintácticos graves, como se podrá observar en los ejemplos.

ANÁLISIS DEL CORPUS

El corpus I corresponde a un informe técnico del área minera titulado “*Determining the Impacts from Mining on Surface and Ground Water*”, publicado por C. M. McCulloch *et al.* en enero de 1996 en la revista especializada *Mining Engineering*. El número total de palabras del texto asciende a 2.871, y el número de formas *-ing* presentes en el texto, a 123, lo que arroja un índice de densidad de un 4,28%. A su vez, el índice de formas *-ing* traducidas correctamente por el programa de TA es de un 40,65% y las formas *-ing* traducidas incorrectamente ascienden a un 59,34%. Ahora bien, del total de formas *-ing* presentes en el texto, un 7,31% corresponde a preposiciones que para efectos de este estudio no fueron consideradas. Un 12,19% de las formas *-ing* correspondían al paradigma verbal, y un 35,77% al paradigma nominal. Cabe destacar que la función atributiva es la más frecuente en el texto, con un índice de 44,71%.

El corpus II corresponde a un artículo presentado en enero de 1996 a la CVIII reunión anual de la American Economic Association y publicado en *The American Economic Review* en mayo del mismo año. Su autor es Ronald Wonnacott. El número de formas *-ing* encontradas en el texto es de 87, lo que corresponde a un 2,85% del total de

número de palabras del texto, el cual asciende a 3.045. Las formas *-ing* traducidas con cierto grado de corrección asciende al 30,82%; las traducidas erróneamente, a un 69,18%. Del total de formas *-ing* presentes en el texto, 5,93% corresponden a formas con uso preposicional que no fueron analizadas. Un 13,12%, al paradigma verbal, y un 33,83%, al paradigma nominal. La función atributiva es en este texto la de mayor recurrencia y asciende a un 47,12%.

A continuación se presenta un análisis selectivo de las traducciones de las formas *-ing* realizadas por el programa de TA, atendiendo principalmente a la diversidad de aciertos y errores detectados en el corpus.

1. Paradigma verbal

1.1 Verbo con aspecto progresivo

1.1.1(TO) *More and more underground mining is being conducted using the longwall method.*

(TA) Más y más minería de subterráneo **está siendo conducido** usando el método de tajo largo.

En este primer ejemplo se puede observar que el programa de TA reconoció la forma progresiva en inglés de la forma *-ing*, la cual fue traducida por la forma verbal correcta en español, que mantiene la voz pasiva y el carácter de dinamismo de la forma verbal en inglés.

1.1.2(TO) (*e.g., U.S. import purchases being switched to NAFTA partner Mexico from an outside Caribbean source*) *seem to indicate trade diversion, but are instead welfare decreasing diversing reversals.*

(TA) (por ejemplo el u.S. las adquisiciones de importación **siendo cambiadas**) a TLC México de socio desde la fuente Caribeña externa parece indicar la desviación del comercio, pero son en vez bienestar-crecientes revocaciones de desviación.

En este segundo caso, el computador tradujo bien la forma progresiva, pero el español requiere la forma del verbo “estar”, que en el original está elidida; el programa de TA no tiene la posibilidad, aparentemente, de discriminar esa necesidad de la lengua meta.

1.2 Verbo con función adverbial

1.2.1(TO) *More and more underground mining is being conducted using the longwall method.*

(TA) Más y más minería de subterráneo está siendo conducido **usando** el método de tajo largo.

En este caso, la función adverbial de la forma *-ing*, que es poco usual en inglés, fue reconocida por el programa de TA y traducida correctamente por el gerundio en español (usando).

1.2.2(TO) *This in turn could impact surface drainage from the area and divert it into the mine, thereby **altering** surface water flow in the stream and also **increasing** infiltration into the mine.*

(TA) Esto a la vez podría incidir el drenaje de superficie desde el área y desvía lo en la mina, por medio de eso **alterando** la agua de superficie fluye en el arroyo e infiltración también **creciente** en la mina.

En este ejemplo, el programa de TA sólo reconoce y traduce bien una de las formas *-ing* con función adverbial (*altering/alterando*); sin embargo, comete error en el segundo caso e interpreta la forma *-ing* como un adjetivo (*increasing/creciente*). Una razón de este error puede radicar en el hecho de que en el segundo caso la preposición *by* está elíptica en la oración y al no “leerla” antepuesta a la forma *-ing* el programa de TA aplica otra regla sintáctica. Esta última consiste en que una forma *-ing* antepuesta a un sustantivo (*infiltration*) tiene función atributiva.

2. Paradigma nominal

2.1 Sujeto de la oración

2.1.1(TO) *This **warning** also applies to G.S.P. (the Generalized System of Preferences) the Lomé Convention, and other preferential arrangements.*

(TA) Esta **advertencia** también aplica a G.S.P. (el Sistema Generalizado de Preferencias, el Lomé, de Convenio, y otros arreglos preferentes).

En este caso se aprecia claramente que cuando la estructura del texto original obedece a reglas sintácticas simples, el traductor automático no comete errores en la identificación de la categoría gramatical; aquí claramente se trata de un sustantivo.

2.1.2(TO) ***Longwall mining** requires large blocks of contiguous reserves to layout the panels for development.*

(TA) **El tajo largo minando** requiere bloques grandes de reservas contiguas al esquema los paneles para el desarrollo.

Resulta bastante curioso el error que comete el programa de TA en esta oración pues a simple vista es incapaz de reconocer una

forma *-ing* compuesta del tipo sustantivo con función atributiva + forma *-ing* con función sustantiva. Este ejemplo viene a ratificar lo señalado por Hutchins y Somers en el sentido de que una secuencia de sustantivos yuxtapuestos sin conectores constituye una estructura compleja para el programa de TA. De hecho, el programa es incapaz de dilucidar las funciones gramaticales que cumple cada una de las palabras que componen esta forma compuesta traduciendo la primera palabra como un sustantivo y la forma *-ing* como un gerundio.

2.2 Objeto del verbo

2.2.1(TO) *If possible, the areas of the mine that will have secondary recovery of pillars or longwall panels should avoid **mining** beneath perennial streams in areas of thin overburden.*

(TA) Si posible, las áreas de la mina que tendrá recuperación secundaria de los paneles de tajo largo o pilares deberían evitar **minar** arroyos abajo perennes en áreas de delgadas sobrecargar. En este caso, el error cometido por el programa de TA resulta menos grave, pues el infinitivo puede ser una buena alternativa de traducción para esta forma *-ing*; sin embargo en esta oración tan compleja la traducción más acertada habría sido la explotación minera, lo cual permite una mejor comprensión del mensaje por parte del destinatario.

2.3 Objeto de la preposición

2.3.1(TO) ***Prior to initiating** any program, the mining company, or their consultant, should obtain a specific list from that state.*

(TA) **Con anterioridad a iniciar** cualquier programa, la compañía minera, o su consultor, debería obtener una lista específica desde el estado.

Este ejemplo ilustra cómo el programa de TA fue capaz de reconocer la función gramatical de la forma *-ing* como objeto de la preposición en inglés y por lo tanto pudo traducirla correctamente al español por un infinitivo.

2.3.2(TO) *Since the effect of diversion is thus apparently limited to the Vinerian terms-of-trade loss **from purchasing** from a more expensive source, it seems that trade diversion is welfare-decreasing by definition.*

(TA) Desde el efecto de desviación se así aparentemente limita al Vinerian los plazos-de- comercial pérdida **desde comprar** desde una fuente más cara, parece esa que desviación del comercio sea el bienestar disminuyendo por la definición.

En este ejemplo, el programa de TA reconoce la primera acepción de la preposición *from* (desde), que, sin embargo, en este contexto, seguida de la forma *-ing*, no corresponde al sentido de la oración. El programa de TA adopta correctamente la forma del infinitivo verbal, si bien traduce la preposición en forma semánticamente incorrecta.

3. Paradigma atributivo

3.1 *Lexema simple*

3.1.1(TO) *In the past, a company would collect a limited amount of data from a few streams and local water wells, write up a short description, state that mining would not impact these areas but, if they did, they (the **mining** company) would “repair” any significant damage.*

(TA) En el pasado, una compañía cobraría una cantidad limitada de datos desde los arroyos y los pozos locales de agua, escriben arriba una descripción corta, estado que la minería no incidiría estas áreas pero, si ellos hicieron, ellos (la compañía **minera**) repararían ningún “importantes” danarían.

En este caso, el programa de TA reconoció la función adjetiva de *mining* dando para ella también el equivalente correcto en español (minera).

3.1.2(TO) *There is no clear expectation for a heavily **diverting** FTA, because diversion may decrease or increase welfare.*

(TA) Aquí es ninguna expectativa clara para un pesadamente FTA **entretenido** porque la desviación puede disminuir o aumentar bienestar.

En este segundo ejemplo el programa de TA reconoce el adjetivo, pero semánticamente el español exige que se explicita la atribución de la siguiente manera: que se diversifica fuertemente o fuertemente diversificado. De esto se puede inferir que el programa de TA está planteado sólo a nivel sintáctico y que aparentemente no tiene incorporadas reglas de decodificación semántica, pues opta por una acepción del término inadecuada para el contexto.

3.2 *Secuencia de lexemas*

3.2.1(TO) *The regulations that pertain to the collection of hydrologic information for mining permits can be found in the **various** state **mining** (usually surface mining) regulations.*

(TA) Las regulaciones que pertenecen al recaudo de información hidrológica para permisos mineros pueden encontrarse en que el diverso estado **mine** (comúnmente superficie minando) regulaciones.

Este ejemplo resulta más complejo que los casos anteriores pues se trata de una secuencia de lexemas compuesta por adjetivo + sustantivo con función atributiva + forma *-ing* con función atributiva, paréntesis con información anexa + sustantivo al cual modifica toda la secuencia. La frase intercalada sería un elemento sintáctico distractor que impide al programa discernir la función gramatical correcta de la forma *-ing*. Esta afirmación está nuevamente avalada por Hutchins y Somers, quienes destacan que ciertos estilos pueden ocasionar dificultades en la TA, como por ejemplo las oraciones intercaladas entre paréntesis.

3.2.2(TO) *Traditional supply-switching effect.*

(TA) **Abastecimiento tradicional conmutador efecto.**

En este ejemplo se puede observar la cadena adjetivo + sustantivo-forma *-ing* con función atributiva + sustantivo. El programa de TA no reconoce la cadena y comete los siguientes errores: el primer adjetivo “traditional” se lo adjudica a “supply”, no reconociendo que “supply” es parte del segundo adjetivo compuesto por sustantivo + forma *-ing*. “Switching” lo separa de “supply” y lo usa para adjetivar a “effect”, lo que produce error sintáctico y semántico a la vez, al transformar la cadena real en: sustantivo + adjetivo + sustantivo + sustantivo.

CONCLUSIONES

Al comparar los resultados obtenidos del análisis de ambos corpus cabe destacar en primer lugar que en el texto económico se presenta un margen del error relativamente mayor (69,18%) en lo que respecta a la traducción de las formas *-ing* al español que en el texto de minería (59,34%). Este hecho corrobora la hipótesis planteada por numerosos autores en el sentido de que aquellos textos que pertenecen a las ciencias exactas, en comparación con aquellos pertenecientes a las ciencias sociales, pueden ser traducidos en forma más correcta por un programa de TA. En efecto, existen ciertas características de los textos pertenecientes a las ciencias exactas, como un campo semántico más restringido, un grado menor de interdisciplinariedad, una terminología más unívoca y que alude más a conceptos concretos que abstractos, entre otras, las cuales en general permiten un mejor desempeño de los programas de TA cuando se enfrentan a este tipo de textos.

Por otra parte, pudimos observar que en ambos corpus se da una coincidencia en lo que respecta al índice de densidad de las funciones gramaticales que adquieren las formas *-ing* en los textos estudiados. En efecto, en ambos corpus se pudo verificar una secuencia de menor a mayor incidencia de la función verbal (12,19% minería/13,12% economía), nominal (35,77%/33,83) y atributiva (44,71%/47,12%), respectivamente. Ello permite concluir que la incidencia de las diferentes funciones gramaticales que conllevan las formas *-ing* no depende de las áreas temáticas ni de los tipos de texto involucrados, sino más bien responde a un patrón general del genio de la lengua inglesa.

En un plano más específico, se puede observar en ambos corpus que el programa de TA rinde buenos resultados cuando se enfrenta a oraciones breves y de una sintaxis simple. Este hecho, que suele ser reconocido por la mayoría de los estudiosos del tema, pone de manifiesto las limitaciones actuales de los programas de TA y la necesidad de seguir trabajando para desarrollar algoritmos más sofisticados que permitan traducir en forma automática oraciones de mayor grado de complejidad, que son por cierto las más recurrentes en textos especializados, como los analizados en este estudio.

En cuanto a las posibles causas más recurrentes de errores que comete el programa de TA en la traducción de las formas *-ing* al español, podríamos consignar que en ambos corpus se verificaron las siguientes: elipsis de partes de la oración como método de economía de lenguaje, secuencias de dos o más lexemas sin conectores o unidos por guión o barra, frases intercaladas entre comas o entre paréntesis que rompen la estructura sintáctica normal de la oración. En todos los casos anteriores, el programa de TA se vio enfrentado a serias dificultades y la mayoría de las veces incurrió en errores sintácticos, e incluso semánticos, graves.

Además, podemos destacar que en ambos corpus se verificó una tendencia general del programa de TA a traducir las formas *-ing* por el gerundio en español en aquellos casos en que no era capaz de determinar la verdadera función gramatical de dichas estructuras en los textos originales. Podemos aventurar al respecto que el programa de TA, junto con presentar un nivel de análisis sintáctico de la oración bastante pobre, recurre al nivel morfológico y simplemente asigna el equivalente en español a la forma *-ing* que se da en este nivel (forma *-ing* = forma *-ando*, *-iendo*).

Finalmente, cabe destacar que en ambos corpus se verificaron una serie de errores en el plano semántico, puesto que el programa de TA no entregaba la acepción correcta de determinados términos polisémicos. Lo anterior demuestra que a pesar de que el programa de TA traduce a nivel de oración, es incapaz en algunos casos de

extraer a partir del contexto intraoracional el verdadero sentido de los términos polisémicos. Éste es quizás uno de los niveles más complejos y desafiantes para la traducción automática, puesto que el reconocimiento de contexto implica aspectos extralingüísticos, como el conocimiento de mundo, que sólo se pueden abordar por medio de la inteligencia artificial. Precisamente a esa tarea se encuentran abocados investigadores y diseñadores de programas de TA en los últimos años.

APRECIACIONES GENERALES

En este estudio se ha pretendido realizar una evaluación de la calidad lingüística de la traducción automática mediante un estudio contrastivo de las formas *-ing* del inglés y su correspondiente traducción al español. A partir de los dos corpus analizados, podemos obtener algunas conclusiones de interés para los traductores profesionales, para la didáctica de la traducción y para aquellos usuarios de programas de TA que piensan que pareciera ya haberse reemplazado el traductor humano por el traductor automático.

En términos generales, podemos deducir claramente que el programa de TA, sin una clara orientación de tipo sintáctico y semántico, no sabe realmente qué hacer. Este hecho supone que, al menos por el momento y seguramente por bastante tiempo más, habrá que investigar y trabajar intensamente para incorporarles a los programas de TA algoritmos sintácticos y semánticos que permitan optimizar su rendimiento. En todo caso, a través de los ejemplos expuestos en el corpus de formas *-ing* deben haber quedado más o menos claras las enormes limitaciones que presenta el programa de TA evaluado para el par de lenguas inglés-español. Esto plantea un claro desafío para los diseñadores de futuros programas de TA.

En cuanto a los ejemplos analizados en este estudio, todos ellos nos demuestran que la ausencia de un mecanismo para aprehender el contexto oracional impide al programa de TA analizar las relaciones sintácticas contiguas al interior de la oración y percibir las distintas posibilidades de interpretación de la polisemia presente en el corpus. Este hecho por cierto no es sólo relevante para la TA, y es por ello que nuestro estudio apunta a lograr conclusiones de interés para la didáctica de la traducción. De hecho, las formas *-ing* del inglés ofrecen serias dificultades al traductor profesional y más aún al alumno de traducción. Es por ello que el análisis de los errores cometidos por el programa de TA respecto de las traducciones al español de las formas *-ing* nos sugiere la posibilidad de realizar ejercicios de postedición de TA con los alumnos de manera de sensibilizarlos más

con esta problemática. En algunos casos las soluciones que da el programa de TA son tan aberrantes, que el futuro traductor profesional puede sacar de ellas conclusiones en forma más rápida al hacer el análisis comparativo del original en inglés y las correspondientes traducciones del programa de TA, que si realiza la traducción del original en forma humana, cometiendo errores que en algunos casos no le quedaría muy en claro por qué los cometió.

De más está decir que, aunque hemos restringido el objeto de estudio de esta investigación a la traducción automática al español de las formas *-ing* del inglés, el análisis de las otras formas verbales y de otros aspectos textuales del corpus ofrecen un gran espectro de posibilidades de estudio que ayudarán, por una parte, a perfeccionar los programas de TA y, por otra, a valerse de sus resultados con el fin de contribuir a la didáctica de la traducción y optimizar con ello la traducción humana.

Queremos finalizar este estudio haciendo una proposición concreta a la didáctica de la traducción postulando en términos generales algunas estrategias que se desprenden de la experiencia adquirida mediante el presente estudio. Ante todo quisiéramos destacar que para realizar este aporte nos hemos apoyado en un estudio de Gisela Thome de la Universidad de Saarland, Alemania Federal, quien propone una metodología de trabajo con el alumno de traducción utilizando traducciones realizadas mediante un programa de TA. A nuestro juicio, postulamos la imperiosa necesidad de familiarizar al alumno con textos traducidos mediante programas de TA por varias razones. 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 del traductor que postedita dichos textos, y en segundo lugar porque ofrece a los alumnos la posibilidad de optimizar el proceso de aprendizaje de la traducción humana al detectar, describir y buscar la causa de los errores de traducción que comete el programa de TA y luego corregir dichos errores.

Proponemos entonces un método general que se puede desglosar en los siguientes pasos:

1. Revisión exhaustiva del texto original y la traducción realizada mediante el programa de TA y detección de los aciertos y errores cometidos por dicho programa. Este punto permite determinar las fortalezas y debilidades del programa de TA.
2. Detección de los problemas de traducción, lo cual implica marcar y clasificar los errores de traducción cometidos por el programa de TA.
3. Determinación de una tipología del error atendiendo a criterios morfológicos, lexemáticos y sintácticos, como también semánticos.

4. Determinación y clasificación de las causas de los errores cometidos.
5. Corrección y posterior descripción y clasificación de los errores cometidos.

El método propuesto permitirá lograr una traducción correcta desde el punto de vista lingüístico y pragmático e introducir y familiarizar al alumno con las estrategias de posesición de textos traducidos mediante programas de TA.

Nuestro objetivo en este estudio ha sido, mediante una evaluación lingüística restringida (formas *-ing* del inglés al español) de un corpus traducido mediante un programa de TA, ofrecer nuevas estrategias didácticas para la enseñanza de la traducción humana. 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 las traducciones automáticas permitiría profundizar este enfoque en muchos otros niveles lingüísticos, como el sistema verbal, adverbial, preposiciones, 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 en pro de la didáctica de la traducción.

BIBLIOGRAFÍA

- ARNTZ, R. (Hrsg.) (1988). *Textlinguistik und Fachsprache*, AILA-Symposium, Olms Verlag, Hildesheim.
- BLATT, A. *et al.* (1985). *Computer und Übersetzen. Eine Einführung*, Olms Verlag, Hildesheim.
- CLOSE, R.A. (1975). *A Reference Grammar for Students of English*, Longman Group Limited, London.
- DE SOPEÑA, L. (1990). "Nuevas perspectivas de la traducción", en: *La lingüística aplicada*, Barcelona: Universidad de Barcelona, pp. 63-71.
- Diccionario de la Lengua Española* (1992). Real Academia Española, 21a edición.
- GUTIÉRREZ, A., M. L. (1978). *Estructuras sintácticas del español actual*, Sociedad General Española de Librería S.A., Madrid.
- HALLIDAY, M.A.K. (1990). *Spoken and Written Language*, Oxford, U.K.
- HUTCHINS, W. J. & Somers, H.L. (1995). *Introducción a la traducción automática*, Visor Distribuciones S.A., Madrid.
- LEHRBERGER, J. & Bourbeau, L. (1988) *Machine Translation. Linguistic Characteristics of MT Systems and General Methodology of Evaluation*, John Benjamins Pu. Co., Amsterdam/Philadelphia.
- MCCULLOCH, C. M. *et al.* (1996). "Determining the Impacts from Mining on Surface and Ground Water", en: *Mining Engineering*, January.
- NAGAO, M. (1989). *Machine Translation: How far can it go?*, Oxford University Press.

- QUIRK, R. & Greenbaum, S. (1977). *A University Grammar of English*, Longman Group Limited, London.
- Revista *Terminómetro*, N° 16, octubre 1994, París: Unión Latina.
- SAGER, J. C. (1992). *La industria de la lengua, la lingüística computacional, los trabajos del UMIST, la traducción especializada y su técnica*, Barcelona: Universidad de Barcelona
- _____. (1993). *Language Engineering and Translation. Consequences of Automation*, John Benjamins, Philadelphia.
- SOLÉ, Y. & Solé, C. (1977). *Modern Spanish Syntax. A Study in Contrast*, University of Texas, Austin.
- STACHOWITZ, R. (1973). *Voraussetzungen für maschinelle Übersetzung. Probleme, Lösungen, Aussichten*, Athenäum Verlag, Frankfurt am Main.
- SWAN, M. (1995) *Practical English Usage*, Oxford University Press, Oxford.
- THOMPSON, A. J. & Martinet, A. V. (1980). *A Practical English Grammar*, Oxford University Press, Oxford.
- VASCONCELLOS, M. & León, M. "SPANAM and ENGSPAN: Machine Translation at the Pan American Health Organization", *Machine Translation Systems*, Slocum, J. (ed), Cambridge University Press, s/f.
- _____. (1986). "Functional Considerations in the Postediting of Machine-Translated Output. Computers and Translation 1", Paradigm Press.
- _____. (1986). "Post-editing on-screen: Machine Translation from Spanish into English. Chief, Terminology and Machine Translation Program", *Translating and the Computer 8*, Picken, C. (ed), CBI Conference Centre, London.
- _____. (1987). "The Contribution of Machine Translation: Present Status and Future Perspectives", *Microcomputer Applications in Education and Training for Developing Countries*, Westview Press, London.
- _____. *et al.* (1987). "Factors in the Evaluation of MT: formal vs. functional Approaches", PAHO, Washington.
- _____. "Machine Translation and the Language Barrier", PAHO, Washington, s/f.
- _____. "The Advantages and Limitations of Machine Translation", PAHO, Washington, s/f.
- WILLIAMS, M. "The Assessment of Professional Translation Quality: Creating Credibility out of Chaos", Department of the Secretary of State of Canada, s/f.
- WILSS, W. & Thome, G. (Hrsg.) (1984). *Die Theorie des Übersetzens und ihr Aufschlusswert für die Übersetzungs- und Dolmetschdidaktik*, Gunter Narr Verlag, Tübingen.
- WILSS, W. (1988). *Kognition und Übersetzen*, Niemeyer Verlag, Tübingen.
- WONNACOTT, R. (1996). "Free-Trade Agreements: For Better or Worse?", en: *The American Economic Review*, Vol. 86, N° 2.