

TRABAJO DE FIN DE GRADO

El par alemán-español en la plataforma de traducción automática Apertium

Análisis de dificultades

Héctor Bolorinos Ruiz
hrbr@alu.ua.es

Grado en Traducción e Interpretación de Alemán

Universidad de Alicante

AUTOR

TUTOR

Héctor Bolorinos Ruiz

Mikel L. Forcada

Resumen:

Apertium es una plataforma de traducción automática de código abierto para la que a día de hoy no se ha implementado el par de lenguas español–alemán. El presente Trabajo de Fin de Grado tiene por objetivo elaborar una relación de aquellas diferencias lingüísticas entre ambos idiomas que supongan un problema a la hora de realizar una traducción automática del alemán al español. Para ello, se llevan a cabo dos traducciones de un texto original alemán: una traducción palabra por palabra realizada mediante un sistema automático (modelo cero) y otra realizada por un traductor. Mediante la comparación de ambas traducciones se identificarán en el texto original aquellos elementos o factores lingüísticos que causen los errores de traducción que genera el modelo cero. A continuación se realizará un análisis cuantitativo de los mismos, con el que se pretende establecer una prioridad con respecto a qué problemas convendría abordar primero si se quiere introducir el par de lenguas en Apertium. Además, se ofrecerá una propuesta de solución para cada uno de los problemas.

Abstrakt:

Apertium ist eine Open Source Plattform für maschinelle Übersetzung, in der heutzutage das Sprachpaar Spanisch-Deutsch noch nicht implementiert wurde. Das Ziel der vorliegenden Abschlussarbeit ist es, diejenige Sprachunterschiede in einer Liste zu bearbeiten, die ein Problem für die maschinelle Übersetzung darstellen. Dafür werden zwei Übersetzungen aus einem deutschen Originaltext durchgeführt: eine Wort für Wort maschinelle Übersetzung (Nullmodell) und eine von einem Übersetzer bearbeitete Übersetzung. Mittels des Vergleichs beider Übersetzungen werden diejenigen sprachlichen Elemente bzw. Faktoren im Originaltext identifiziert, die im Nullmodell Übersetzungsfehler verursachten. Anschließend wird eine quantitative Analyse von solchen Elementen durchgeführt, die dazu dienen, eine Lösungspriorität zu setzen. Zu jedem Element wird außerdem einen Lösungsvorschlag vorgelegt.

Abstract:

Apertium is an open-source machine translation platform for which the language pair Spanish-German has not yet been implemented. This bachelor thesis aims to list the linguistic differences between both languages that pose a problem when performing automatic translation from German into Spanish. For this purpose two translations of an original German text will be carried out –an automatic translation word by word (model zero) and a translation carried out by a translator. By comparing the two translations it will be possible to identify the text's linguistic elements or factors that cause translation errors with the model zero. These elements will be the subject of a quantitative analysis that is intended to establish a priority over what problems should be addressed first when implementing the language pair in Apertium. In addition, a solution for each of the problems will be proposed.

Índice

1.	Introducción.....	1
2.	Tipos de TA y sus características	3
2.1	Traducción <i>directa</i>	3
2.2	Traducción basada en reglas (TABR)	4
2.3	Traducción basada en corpus (TABC)	6
3.	Apertium.....	7
3.1	Arquitectura de Apertium	8
4.	Análisis del corpus	9
4.1	El modelo cero.....	9
4.2	Método de análisis y evaluación.....	11
4.3	Resultados del análisis.....	11
5.	Propuestas de solución en Apertium	15
6.	Conclusión.....	28
	Anexo	29
7.	Bibliografía.....	35

1. Introducción

La idea de crear un sistema de traducción que permitiera traducir textos entre dos lenguas ha estado presente desde la primera mitad del siglo XX y tuvo un importante respaldo en el auge informático generado a partir de la Segunda Guerra Mundial, cuando los primeros ordenadores ayudaron a descifrar los códigos que encriptaban las comunicaciones enemigas. No obstante, a diferencia de la criptografía, en la que la codificación está basada en relaciones unívocas entre los elementos de salida y los de llegada, el lenguaje humano reviste una complejidad muchísimo mayor, adornada con infinidad de sutilidades y de relaciones intertextuales difíciles de procesar mediante ordenadores (Arnold 1995:12).

A partir de 1950, ese empuje de los avances en la informática propiciaron la idea de que la inteligencia artificial permitiría automatizar la traducción mediante la creación de sistemas informáticos capaces de traducir texto de manera rápida y fiable. Sin embargo, estas ideas sufrieron un fuerte varapalo a raíz de la publicación en 1966 del informe del famoso comité ALPAC (Automatic Language Processing Advisory Committee), al que se encargó la labor de evaluar la situación de la TA en EE. UU. En su informe, el comité ALPAC concluyó que la TA no resultaba rentable, ya que era lenta, cara y de poca calidad en comparación con la traducción humana. Estos resultados provocaron que las agencias militares y de inteligencia suprimieran la financiación de la que la TA había disfrutado hasta el momento y que se transmitiera al público general un mensaje claro: no hay futuro en la automatización del proceso de traducción (Hutchins 2003:131).

Por fortuna, la situación en el resto del mundo no era igual a la que atravesaba EE. UU., a pesar de que el desánimo creado por el informe del ALPAC logró en parte atravesar las fronteras estadounidenses. La investigación sobre la TA continuó durante la década de 1960 en Canadá y en lo que entonces empezaba a ser la Comunidad Europea. Con el tiempo, estas investigaciones dieron como resultado sistemas como Météo, desarrollado por el proyecto de TA de la Universidad de Montreal, y posibilitaron la aparición de empresas como SYSTRAN, que ha trabajado muy estrechamente con la UE y la Fuerza Aérea de los Estados Unidos (USAF). Además, en 1988 se elaboró en Japón el informe JEIDA, en el cual se consideraba que la conclusión principal del informe ALPAC era

errónea, ya que las circunstancias habían cambiado y las investigaciones en TA demostraban que se habían producido numerosas mejoras. El informe JEIDA animaba a seguir con la investigación, en especial sobre el procesamiento de lenguas naturales y sobre la mejora y desarrollo de datos lingüísticos.

Con el tiempo, la realidad ha mostrado ser bien distinta a lo que parecía pronosticarse en EE. UU. en la década de los años 50. Tanto desde el punto de vista laboral como el económico, la TA se ha mostrado como una herramienta capaz de mejorar en determinadas circunstancias la productividad de los traductores y, por consiguiente, ha adquirido un papel muy relevante tanto a nivel comercial como institucional. Un claro ejemplo de este último ámbito lo mencionaba ya Arnold (1994:5) al aseverar que

“it has been estimated that some 40-45% of the running costs of European Community institutions are ‘language costs’, of which translation and interpreting are the main element. This would give a cost of something like £300 million per annum. This figure relates to translations actually done, and is a tiny fraction of the cost that would be involved in doing all the translations that could, or should be done”.

Es decir, si tenemos en consideración únicamente aquellos organismos internacionales en los que se hace necesario traducir entre distintas lenguas (como por ejemplo la propia UE), el volumen de traducciones que se realiza es tal, que la mejora e implementación de la traducción automática puede suponer un aumento importante de la productividad, con el consiguiente ahorro económico. Más aún, los esfuerzos actuales a este respecto dentro de la UE no se centran exclusivamente en la utilidad de la TA para la traducción de documentos a las lenguas oficiales, sino también en eliminar las barreras lingüísticas existentes entre las distintas administraciones nacionales y entre estas y la administración central europea, de manera que los agentes públicos puedan usar las infraestructuras europeas de servicios digitales y comunicarse en su propia lengua con sus contrapartes en el resto de países. Este es el objetivo en el que trabaja actualmente la Plataforma de Traducción Automatizada del programa europeo MCE (Mecanismo “Conectar Europa”), CEF.AT por sus siglas en inglés (Connecting Europe Facility. Automated Translation) (<http://lr-coordination.eu/es/cef>).

En cualquier caso, desde el punto de vista de los profesionales involucrados en la creación de sistemas de TA, para ofrecer resultados de calidad los equipos de desarrollo deben ser multidisciplinares. Como ya apuntaba Sager (1994:312), el papel de los traductores, los terminólogos y los lingüistas también es decisivo para mejorar la calidad de los sistemas de TA, ya que pueden aportar un papel activo en su desarrollo, por ejemplo, añadiendo especificaciones sobre la traducción a las que el sistema deba ajustarse o para las cuales deban crearse nuevos módulos. Este tipo de perfil profesional es especialmente relevante, como veremos en el siguiente punto, en el desarrollo de sistemas de traducción automática basada en reglas de transferencia (TABR).

2. Tipos de TA y sus características

El término *traducción automática* designa el proceso de traducción de un texto contenido en un soporte informático de una lengua a otra llevado a cabo por una máquina, es decir, sin intervención humana. Dentro de ese proceso automático encontramos tres aproximaciones que repasaremos brevemente a continuación: la traducción *directa*, la traducción basada en reglas (TABR) y la traducción basada en corpus.

Además, las aplicaciones de la TA se pueden agrupar en dos clases: las de *asimilación* y las de *diseminación*. En el caso de la asimilación, lo que se pretende es acceder a la información que contiene el texto original (TO), para lo cual no es estrictamente necesario que dicho texto esté perfectamente traducido sino simplemente que el lector pueda entender lo suficiente como para hacerse una idea de las ideas contenidas en el texto. Con la diseminación, por el contrario, se busca producir un texto en la lengua meta (LM) que pueda distribuirse y que tenga la suficiente calidad como para que lo único que haga falta sea que una persona se encargue de realizar una corrección profesional, lo que se conoce como *postedición* (Hutchins 1995:195).

2.1 Traducción *directa*.

Este método de traducción es, en teoría, el más simple que puede concebirse y es el que se usaba en los primeros sistemas allá por la década de 1950. Aunque a priori la denominación *directa* pueda hacer pensar que se trataba de un método de traducción

palabra por palabra sin ningún tipo de estadio intermedio entre el TO y el texto meta (TM), en realidad este tipo de sistemas sí que realizaba un análisis mínimo para poder realizar la traducción, por lo que, en realidad, se trataba ya de un sistema de TABR. El método de traducción consistía en llevar a cabo un análisis morfológico de las palabras tal cual aparecen en el texto, flexionadas (formas superficiales), e identificar para cada una de ellas su forma léxica correspondiente, es decir, su lema (su forma básica) junto con la categoría léxica y la información morfológica que contiene. Por ejemplo, la forma léxica que el sistema reconocería para la forma superficial *traduje* sería: *traducir*, verbo, 1ª persona del singular, pretérito perfecto simple. Este pequeño análisis le servía al sistema para poder consultar las palabras del texto en un diccionario bilingüe, ya que en los diccionarios normalmente solo se recogen los lemas. En palabras de Hutchins (1995:115), el producto de este sistema era “una traducción ‘palabra por palabra’ a la que se aplican algunos ajustes de carácter local que afectan al orden de las palabras”.

Este método era tan sencillo que equivalía a encargarle la traducción a “una persona que contara únicamente con un diccionario bilingüe muy barato y con un conocimiento muy rudimentario de la gramática de la lengua meta”, por lo que el resultado de la traducción contenía “frecuentes errores de índole léxica [...] y estructuras sintácticas inapropiadas que reflejaban, en gran medida, las estructuras propias de la lengua de origen” (Hutchins, 1995:115). Por lo tanto, para mejorar este método se necesita un análisis más profundo del TO, que permita identificar las relaciones sintácticas de los elementos en la frase y la estructura propia de la LO.

2.2 Traducción basada en reglas (TABR)

Debido a estas limitaciones, a partir de 1980 se empezó a desarrollar sistemas de traducción más sofisticados, a los que se les denominó *indirectos*, en los que se ampliaba el conocimiento lingüístico sobre las lenguas de trabajo. El objetivo que se persigue con estos sistemas indirectos es el de realizar un análisis del TO con la profundidad suficiente para poder obtener una representación abstracta de la LO (RALO). Esta es una abstracción que permite establecer generalizaciones que facilitan la traducción. Posteriormente se realiza un proceso de transferencia, pero no del TO al TM, sino de la RALO a la representación abstracta del texto en LM (RALM), similar a

la RALO. A continuación se realiza el proceso inverso al de análisis: la generación del TM a partir de la RALM.

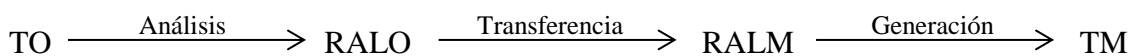


Fig. 2.1: Las tres fases de un sistema de traducción por transferencia

En la figura 2.2 se representa el triángulo de Vauquois, que ilustra los distintos niveles de análisis que pueden darse en la relación TO-TM. En él se muestran los tres estadios de todo proceso de traducción: análisis, transferencia y generación. Como puede observarse, cuanto más se asciende por el lado izquierdo del triángulo, más profundo es el análisis y, por lo tanto, mayor la abstracción que se consigue. En teoría, estos sistemas ofrecen la posibilidad de traducir prácticamente cualquier oración usando un número limitado de reglas, si bien la cantidad de reglas que deban implementarse para conseguir un texto adecuado dependerá de las diferencias y similitudes entre ambas lenguas de trabajo.

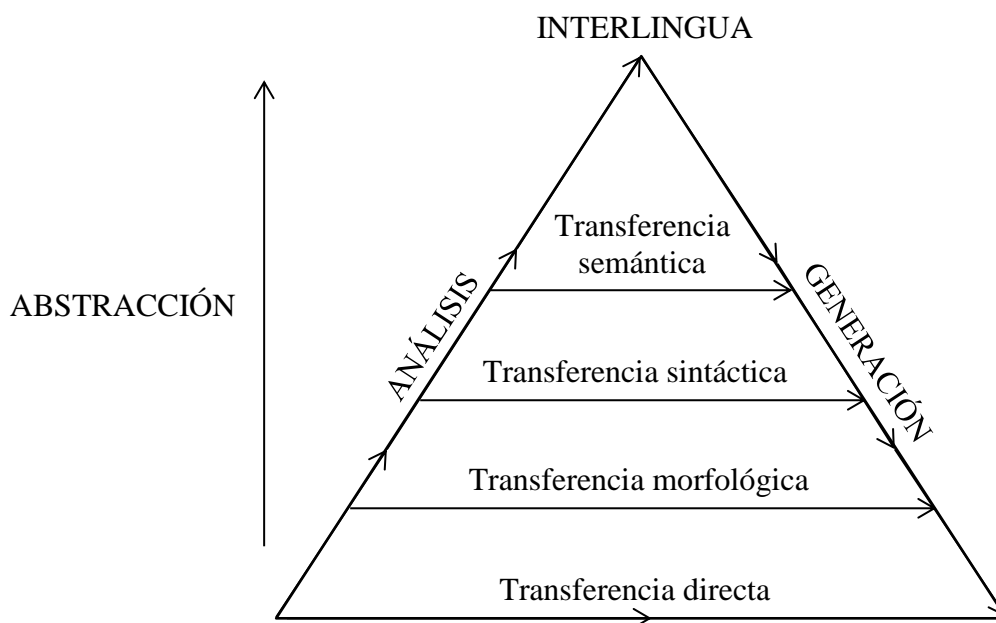


Fig. 2.2: Triángulo de Vauquois

El vértice superior del triángulo de Vauquois representa la interlingua, que supone conseguir el mayor nivel de abstracción posible, de manera que se obtiene un modelo común a todas las lenguas. De esta manera se elimina la fase de transferencia y se consigue una única representación abstracta del significado del TO. Debido a la complejidad de los sistemas de interlingua, hoy en día los esfuerzos de la investigación en TA están orientados a la traducción automática estadística, si bien durante los años noventa se llevaron a cabo varios proyectos de interlingua, uno de los cuales sigue vigente hoy en día: el UNL (Universal Networking Language), dirigido actualmente por la organización sin ánimo de lucro UNDL Foundation, organismo independiente creado en 2001 por la United Nations University.

2.3 Traducción basada en corpus (TABC)

Los sistemas de TABC son aquellos que usan como fuente de información para traducir el TO corpus bilingües formados por textos en la LO alineados oración a oración con sus traducciones en LM. Esto ya marca una diferencia importante con respecto a la traducción basada en reglas, ya que los textos traducidos mediante reglas tienden a ofrecer resultados consistentes, pero no se caracterizan por su naturalidad, mientras que los sistemas basados en corpus sí que ofrecen esa naturalidad, debido a que la información de base no son reglas de traducción, sino un corpus bilingüe enorme con textos ya traducidos y modelos de LM entrenados sobre corpus monolingües adicionales.

Dentro de la traducción basada en corpus destaca la traducción automática estadística (TAE). Los sistemas de TAE basan su funcionamiento en la aplicación de modelos estadísticos y probabilísticos al corpus bilingüe, y casi siempre también a modelos monolingües de la lengua meta. Los sistemas estadísticos ofrecen su traducción atendiendo a la probabilidad de que, por ejemplo, las palabras que integran la oración aparezcan en un determinado orden o de que la traducción de una palabra o de un grupo de palabras sea una y no otra según el contexto en el que aparezcan. Los datos de probabilidad que se usan se obtienen mediante el análisis previo del corpus. Para ello, se buscan determinados aspectos del corpus bilingüe que puedan ser fácilmente medibles, como por ejemplo la concurrencia de dos o más palabras en un determinado segmento del TO y su correspondiente en el TM, la posición relativa de las palabras en las

oraciones, la longitud de las oraciones, etc. (Trujillo 1999:210). Este tipo de TA ha ido ganando peso en los últimos años, sobre todo gracias al extraordinario incremento de la capacidad de procesamiento de los ordenadores actuales y de la cantidad de datos y textos bilingües y monolingües disponibles electrónicamente gracias a Internet.

3. Apertium

Apertium es una plataforma de traducción automática desarrollada en sus inicios por el grupo Transducens del Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Alicante. Se trata de un sistema de transferencia superficial pensado para lenguas afines, como el español y el catalán, aunque con el tiempo se ha ido expandiendo para poder trabajar con lenguas no tan afines entre sí.

Apertium ofrece una serie de ventajas con respecto a los sistemas de traducción estadística. Según Forcada et al. (2011:128-129), si bien es cierto que estos últimos pueden ofrecer traducciones más naturales, también lo es que, por lo general suele existir un riesgo de que ofrezcan traducciones muy dispares para una misma palabra, ya que la elección la realizan basándose en la probabilidad de que el conjunto de la oración sea correcto. Esto conlleva dos problemas: por un lado, un cambio pequeño en el TO puede dar lugar a un gran cambio en el TM; por otro, el peso del modelo de la LM puede hacer que la naturalidad prevalezca sobre la fidelidad.

Por el contrario, Apertium, al ser un sistema de TABR, puede producir traducciones que, a pesar de ser más mecánicas y repetitivas, son mucho más consistentes desde el punto de vista terminológico. Además, la información lingüística en Apertium está codificada en lenguaje XML y desligada del motor de traducción. Esto permite que el conocimiento lingüístico esté disponible para ser usado en otros pares de lenguas, que pueda operarse con ella en otros entornos que utilicen el mismo lenguaje y que el sistema pueda transformarse y mantenerse más fácilmente e incluso que se pueda usar para construir otros sistemas de TA. El hecho de que el lenguaje XML sea un formato estándar facilita aún más el llevar a cabo todas estas posibilidades.

Además de estas ventajas, Apertium se distribuye con la licencia pública general GNU, una licencia de software libre o de código abierto, por lo que cualquier persona con los

conocimientos lingüísticos o informáticos suficientes puede contribuir a su expansión de forma desinteresada o crear nuevos productos a partir de él.

3.1 Arquitectura de Apertium

El sistema consta de una serie de módulos que procesan el texto. Como se puede ver en la figura 2.3, cada uno de estos módulos realiza una función concreta y están dispuestos como si fuera una cadena de montaje:

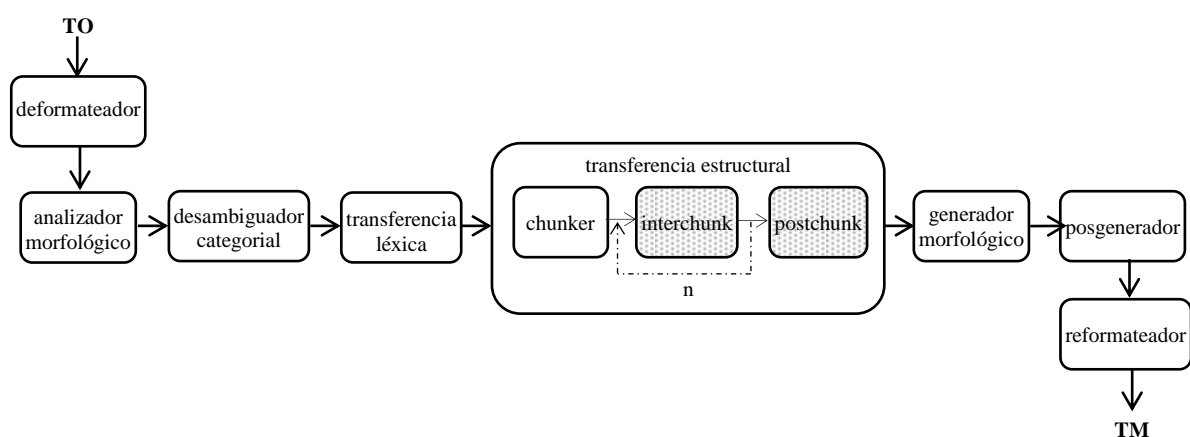


Fig. 2.3: Esqueleto de Apertium. Los módulos sombreados son opcionales y están pensados para pares de lenguas poco relacionadas entre sí.

- El **desformateador** se encarga de encapsular las etiquetas que dan formato al texto para que no sean tenidas en cuenta por el resto de módulos.
- El **analizador morfológico** divide el texto en sus palabras y asigna a cada una de ellas (forma superficial) una forma léxica, o varias en caso de que la forma superficial sea ambigua. Cada forma léxica está compuesta por su lema, su categoría léxica y su información morfológica. Apertium obtiene la información necesaria para realizar esta operación de un diccionario morfológico monolingüe de la lengua del texto original, codificado en formato XML.
- El **desambiguador léxico categorial** entra en funcionamiento cuando una forma superficial genera un caso de ambigüedad. Este módulo resuelve cuál es la forma léxica apropiada mediante un modelo estadístico. Con él se intenta decidir cuál de las formas léxicas que corresponden a la forma superficial funciona mejor dentro del contexto en el que aparece.
- El **módulo de transferencia léxica** utiliza un diccionario bilingüe para asociar a cada forma léxica en la lengua de partida a una forma léxica en la de llegada.

- El **módulo de transferencia estructural** se encarga de detectar y procesar patrones de palabras. Sirve para intentar solucionar las divergencias gramaticales entre ambas lenguas (cambio en el género y número de las palabras, cambio en la posición, cambio de preposición, etc.) y puede constar de hasta tres submódulos: *chunker*, *interchunk* y *postchunk*. Cuando las lenguas entre las que se traduce son similares (como el español y el catalán) se usa exclusivamente el primero. Los otros dos solo se utilizan cuando las lenguas son muy diferentes y el primer submódulo no es suficiente para conseguir una traducción adecuada de la frase que se intenta procesar.
 - El *chunker* se encarga de detectar los distintos patrones de palabras y los encapsulan en segmentos de una o más palabras, denominados *chunks*, dentro de los cuales se reordenan las palabras. Dentro del *chunk*, por ejemplo, la estructura determinante-adjetivo-sustantivo del alemán pasaría al español como determinante-sustantivo-adjetivo.
 - En el *interchunk* se realizan operaciones entre los propios *chunks*, es decir, las operaciones de reordenamiento tienen un alcance mayor.
 - En el *postchunk* se realizan operaciones finales dentro de cada *chunk* por separado y se desencapsula la información que contienen para generar una secuencia de formas léxicas que serán introducidas en el siguiente módulo.
- El **generador morfológico** flexiona las formas léxicas que le llegan y produce las formas superficiales en la LM.
- El **posgenerador** se encarga de realizar las operaciones ortográficas necesarias, como por ejemplo la contracción de *de* y *el* para formar *del*.
- Por último, el **reformateador** desencapsula la información de formato que se había encapsulado al comienzo del proceso y la sitúa en el TM.

4. Análisis del corpus

4.1 El modelo cero

El modelo cero es el que usaremos para el análisis del corpus. Se trata de un modelo que sirve para detectar qué es necesario hacer (qué operaciones se han de llevar a cabo),

además de para obtener un equivalente fijo para cada palabra. En él, el sistema lee la frase de izquierda a derecha, palabra por palabra, y ofrece en la LM una frase en la que las palabras originales se sustituyen, en el mismo orden que el original, por un equivalente que se encuentra recogido en el diccionario bilingüe y que puede estar formado por una o más palabras, o por ninguna. Así, para una frase en LO con n palabras, el sistema ofrecería su traducción sin tener en cuenta para ello las relaciones existentes ellas.

LO	P ₁	P ₂	P ₃	P ₄	...	P _n
LM	T(P ₁)	T(P ₂)	T(P ₃)	T(P ₄)	...	T(P _n)

Con el siguiente ejemplo ilustramos mejor cuál es el procedimiento:

Die Politiker haben eine Entscheidung über die Krise der Flüchtlinge getroffen (1)
La políticos han una decisión sobre la crisis el refugiados acertado
Los políticos han tomado una decisión sobre la crisis de los refugiados

El traductor busca las formas superficiales de la frase original en el diccionario bilingüe y ofrece para cada palabra una traducción, tras lo cual sustituye la oración alemana por la española:

<i>Die</i>	<i>Politiker</i>	<i>haben</i>	<i>eine</i>	<i>Entscheidung</i>	<i>über</i>	<i>die</i>	<i>Krise</i>	<i>der</i>	<i>Flüchtlinge</i>	<i>getroffen</i>
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
La	político	tienen	una	decisión	sobre	la	crisis	el	refugiados	acertado

Como se puede observar, al traducir palabra por palabra aparecen varios problemas, como por ejemplo la falta de concordancia entre determinantes y sustantivos y errores en la traducción del verbo debido a la separación de los elementos que lo constituyen, ya que las palabras *haben* y *getroffen* se corresponden en alemán con el auxiliar y el participio del *Perfekt* (pretérito perfecto compuesto), por lo que la traducción correcta no es *tener* y *acertado*, de forma aislada, sino *han acertado*. Además, la expresión *eine Entscheidung treffen* (tomar una decisión) es una colocación, por lo que el verbo *treffen*, el infinitivo del participio *getroffen*, debe traducirse por *tomar*, y no por *acertar*.

4.2 Método de análisis y evaluación

Con el modelo cero se traducirá un texto de carácter general de 730 palabras. Para asegurar la calidad del texto original, éste se ha elegido de la edición digital de uno de los periódicos de mayor tirada nacional en Alemania, *Die Zeit*. De él se realizarán dos traducciones, una mediante el modelo cero y otra mediante un traductor humano y se compararán para ver qué errores de traducción se cometen con el modelo cero y qué particularidades del TO son los que los causan. En el anexo se encuentra el texto original y las traducciones.

No obstante, la tarea del traductor no será la de realizar una traducción profesional, sino el tipo de traducción que cabría esperar de un sistema de TA, como por ejemplo Apertium, una vez que se hubieran introducido los diccionarios y reglas necesarias en los archivos correspondientes. De realizarse una traducción profesional, el traductor haría uso de una serie de técnicas de traducción (adición, omisión, paráfrasis, etc.) que son necesarias para conseguir un texto idiomático, pero que dificultarían demasiado la comparación de ambas traducciones. Además, sería imposible conseguir que un sistema de TABR empleara estas técnicas, por lo que no sería el objetivo de la TA con Apertium obtener un texto de esas características.

Una vez realizada la comparación de ambas traducciones, se categorizarán los errores que se han encontrado y se cuantificarán su magnitud. De esta manera podremos establecer la recurrencia de cada tipo de error dentro del conjunto del texto, y, a su vez, se podrá establecer una prioridad, de manera que la calidad de los textos traducidos automáticamente mejorará conforme se vayan añadiendo al paquete de datos es-de reglas que permitan solucionar las dificultades que son más recurrentes para la traducción del alemán al español.

4.3 Resultados del análisis

Los tipos de problemas que se han identificado en la traducción del texto original con el modelo cero quedan recogidos en la siguiente tabla.

AA	Artículo ausente: <i>Aber Fleischliebhaber können...</i> Pero Amantes de la carne pueden... Pero los amantes de la carne pueden...
----	--

CDSA	Concordancia entre determinante, sustantivo y adjetivo: <i>Der tägliche Verbrauch</i> El diarios uso El diario uso
SCO ACO	Sustantivos y adjetivos compuestos: <i>Wasserverbrauch</i> <i>wasserregulierende Duschkopf</i> <i>Wasser – verbrauch</i> <i>wasser – regulierende Dusch–kopf</i> agua – uso agua – que regula ducha–cabeza uso de agua alcachofa de ducha reguladora de agua
CS	Complemento suplemento. Verbos que rigen una preposición concreta, y los sustantivos derivados: <i>Verzichten auf</i> → <i>Verzicht auf</i> Renunciar sobre → Renuncia sobre Renunciar a → Renuncia a
ULMP	Unidades léxicas multi-palabra. Conjuntos de palabras que tienen una traducción fija: <i>wie viel</i> <i>am Meisten</i> como mucho en el mayoría cuánto más, la mayor parte
EFV	Elección incorrecta de la forma verbal debido a la ambigüedad. <i>steckt</i> (<i>stecken</i> : poner) puede ser: 3ª persona singular, presente de indicativo: pone; 2ª persona plural, imperativo: poned.
M	Mayúsculas. En alemán, todos los sustantivos se escriben con mayúsculas: <i>Der Wasserverbrauch</i> El Uso de agua El uso de agua
OSA	Orden sustantivo-adjetivo: <i>In einem wasserreichen Land</i> En un rico en agua país En un país rico en agua
PPE	Posición del pronombre enclítico: <i>Der Wasserverbrauch bemisst sich</i> nicht nur... El uso de agua mide se no solo... El uso de agua no se mide solo...
PR	Pronombre relativo. En alemán tiene la misma forma que el artículo que acompaña al sustantivo al que se refiere la oración de relativo. No solo se trata de traducir correctamente el pronombre, sino también de eliminar la coma obligatoria en las subordinadas alemanas: <i>Water Footprint Network bietet einen Kalkulator an, der...</i> Water Footprint Network ofrece una calculadora en, el... Water Footprint Network ofrece una calculadora que...
VS	Verbo separable: (<i>ausmachen</i> : constituir) <i>Der direkte Verzehr der Bohne macht nur einen Anteil des weltweiten Anbaus aus</i> El directo consumo del haba hace solo una parte del mundial cultivo de El consumo directo del haba constituye solo una parte del cultivo mundial
VSE	Verbo <i>ser</i> o <i>estar</i> . En alemán existe solo el verbo <i>sein</i> , que en ocasiones se traduce también por el verbo <i>tener</i> . Además, cuando el verbo <i>sein</i> es parte de la forma compuesta de un verbo, se traduce por el verbo <i>haber</i> .

Y los resultados del análisis son los siguientes:

Palabras totales del texto: 730		
Tipo	Nº de apariciones	porcentaje sobre el total
M	155	27,48 %
CDSA	138	24,47 %
OSA	56	9,93 %
SCO	53	9,40 %
EFV	41	7,27 %
PV	37	6,56 %
AA	29	5,14 %
ULMP	17	3,01 %
VSE	10	1,77 %
CS	9	1,60 %
PR	8	1,42 %
VS	6	1,06 %
ACO	4	0,71 %
PPE	1	0,18 %
Total	564	100 %

Tabla 1. Resultados del análisis

En la columna central se recoge el número de veces que aparece cada elemento en el texto, haya supuesto o no un error de traducción mediante el modelo cero. La razón para incluir en el recuento los casos que no han generado error es la siguiente. Tomemos como ejemplo la frase 5 del texto original

*Wassersparen ist eine **Lieblingsdisziplin** der Deutschen*

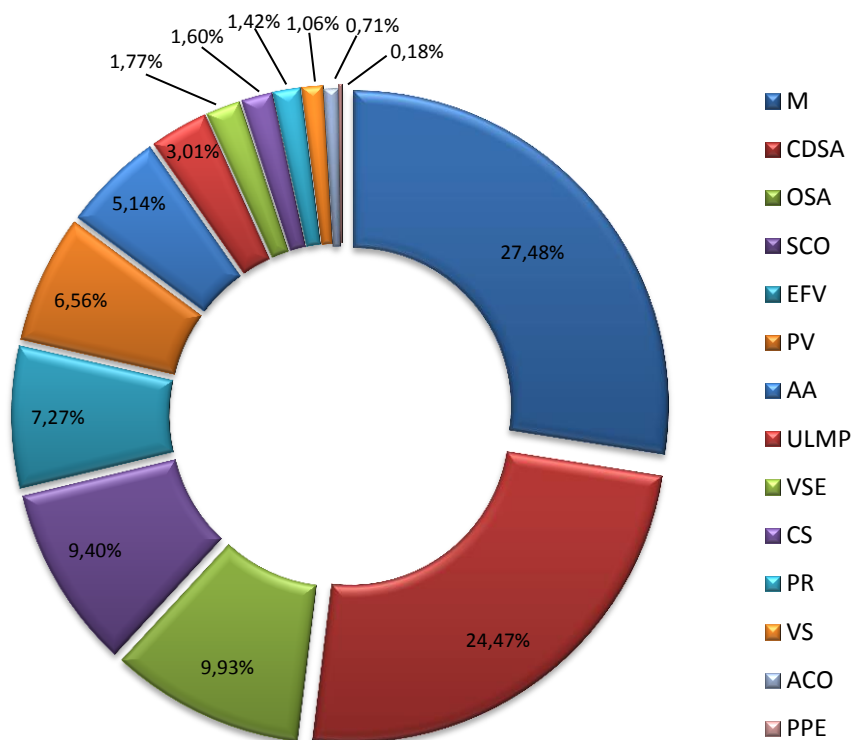
y en concreto la parte resaltada con negrita, que el modelo cero ha traducido como

Ahorro de agua **es una Disciplina favorita** el Alemanes.

Aunque en muchas otras frases no se ha establecido correctamente la concordancia entre el determinante y el sustantivo, ni se ha traducido correctamente el verbo *sein* por *ser* o por *estar*, en la parte resaltada en negrita de esta oración sí que se ha logrado. Sin embargo, el hecho de que se haya traducido correctamente no es producto de contar con un sistema que haya sido capaz de realizar las operaciones necesarias para analizar el

TO y ofrecer la traducción correcta. Es simplemente una consecuencia de que en el diccionario bilingüe el modelo cero contara con equivalentes para *ist* y *eine* que casualmente son correctos en esta ocasión, además de que el equivalente de *Lieblingsdisziplin*, que en alemán es un sustantivo de género femenino, sea en español también de género femenino. En caso de que el género de *disciplina* no fuera femenino sino masculino, la concordancia de género femenino entre el artículo y el sustantivo en la frase original no se mantendría en la traducción del modelo cero. Igualmente, si el equivalente correcto para *ist* hubiera sido *está*, la traducción del modelo cero también hubiera sido errónea. Por eso, si no incluyéramos en el recuento estos casos correctamente traducidos estaríamos ofreciendo un recuento de los casos de recurrencia más bajo del que sería realmente.

La tercera columna de la tabla 1 presenta los mismos datos de recurrencia, pero expresados como porcentaje sobre el total de problemas encontrados.



Tal y como se aprecia de manera visual en el gráfico, los dos elementos que más problemas dan son los que aparecen con mayor frecuencia en el texto: las mayúsculas

de los sustantivos (27,48 %) y las concordancias de género y número entre los determinantes, los sustantivos y los adjetivos (24,47 %). Además, tal y como se deduce del tercer elemento en orden descendente, el número de sustantivos que son a su vez formas compuestas es de un 9,93 %, lo que indica que más de un tercio de los sustantivos del texto son compuestos. Esto nos da una idea de la importancia que tiene para traducir del alemán al español que Apertium cuente con las reglas necesarias para procesarlos. Como explicaremos en el siguiente apartado, hay determinados tipos de compuestos cuya traducción es sencilla, puesto que para ello basta con que se introduzcan tal cual en los diccionarios. Este tipo de compuestos y la posición relativa del sustantivo y el adjetivo (9,40 %), cuya severidad con respecto a la comprensión del texto es baja, son probablemente dos de los elementos que serían menos complejos de tratar.

Probablemente sean dos de los problemas relacionados con los verbos –la posición del verbo (6,56 %) y los verbos separables (1,06 %)– los que sí que afectan en mayor medida a la comprensibilidad. Pese a tener una menor presencia en el texto, situar el verbo en una posición muy alejada de la que tendría en español o traducir un verbo separable alemán al español como una forma verbal más una partícula aislada cuando los dos elementos del verbo separable se encuentran en puntos opuestos de la oración alemana pueden dificultar mucho la comprensión del texto. Su severidad es, por lo tanto, mayor.

El resto de elementos incluyen cuestiones que afectan poco a la comprensibilidad, como por ejemplo la ausencia de artículo en determinados casos (5,14 %), la inversión en la posición del pronombre átono con respecto al verbo (0,18 %) o la aparición de unidades multi-palabra (3,01 %), cuya solución probablemente sea exclusivamente introducirlas tal cual en los diccionarios.

5. Propuestas de solución en Apertium

En este punto ofrecemos posibles soluciones para algunos de los problemas encontrados.

- **Tratamiento de las declinaciones.** La declinación no se ha analizado específicamente en el texto debido a que no existe un correlato en español para la declinación alemana. No obstante, se trata de un problema a resolver si se quiere traducir del alemán al español, aunque a priori se trata de un operación más sencilla que la de generar la declinación al traducir de español al alemán.

El principal problema de la declinación tiene que ver con las terminaciones de los artículos y adjetivos –en menor medida, también de los sustantivos– una vez que se declinan. Intentaremos simplificar la explicación mediante un ejemplo extraído de la frase 2 del texto:

[...] *an der Dauer des morgendlichen Duschens*
[...] durante la duración de la matutina ducha
[...] durante la duración de la ducha matutina

Los elementos que encontramos son: la preposición *an* (cuya traducción en este caso sería *durante*), que en este ejemplo rige el caso dativo (aunque en otros puede regir también acusativo), el determinante *der* (el), el sustantivo *Dauer* (duración), el determinante de genitivo *des* (del), el adjetivo *morgendlich* (matutino) declinado en genitivo y el sustantivo *Duschen* (nominalización del verbo *duschen*, duchar) declinado también en genitivo.

El problema reside, por ejemplo, en que los determinantes son elementos ambiguos con varios posibles análisis morfológicos:

der: masculino nominativo (*el*), femenino dativo (*a la*), femenino genitivo (*de la*), o plural genitivo (*de los / de las*)

des: masculino genitivo (*del*) o neutro genitivo (*del*)

En este caso, *Dauer* es un sustantivo femenino en alemán, por lo que le correspondería el artículo *die*, pero al estar declinado en dativo por la presencia de la preposición *an*, toma la forma correspondiente al dativo, *der*. La ventaja al traducir al español es que la única variación que experimentan los determinantes viene dada por la concordancia de género y número del sustantivo al que acompañan. En español, el caso queda recogido en realidad en la preposición *de*

que se añade al artículo en esta ocasión. Por lo tanto, en cierta medida, la presencia de la preposición *an* es un factor que se puede aprovechar para facilitar la fase de transferencia, ya que restringe los posibles análisis de *der* a uno solo, el dativo femenino. Este tipo de análisis se podría llevar a cabo en el desambiguador léxico categorial. No obstante, hay muchas otras ocasiones en las que no hay preposición alguna y la declinación viene dada por la función sintáctica del sintagma en cuestión, lo cual podría complicar el análisis.

Sin embargo, este es solo uno de los problemas que pueden darse, ya que existen otros factores a tener en cuenta, como por ejemplo el hecho de que en el caso de los genitivos debe generarse la preposición *de*. Por ello, se hace imprescindible un análisis en profundidad de los problemas derivados de la declinación alemana para poder realizar las propuestas de solución necesarias.

- **Mayúsculas.** Apertium ya cuenta con una manera de trabajar con aquellas palabras que contienen mayúsculas, de manera que puede discernir, por ejemplo, entre nombres comunes y nombres propios (ver http://wiki.apertium.org/wiki/Letter_case_handling).

En el caso del alemán, la mayúscula inicial de los sustantivos se procesaría en el diccionario, dentro del módulo de transferencia léxica. En este módulo se lee la palabra de entrada con mayúscula y al encontrar la coincidencia en el diccionario realiza automáticamente la sustitución con el sustantivo español en minúscula, a no ser que se trate de un inicio de oración, en cuyo caso se aplicarían las reglas generales de gestión de mayúsculas. En cualquier caso, un requisito indispensable es el de introducir previamente en los diccionarios la mayor cantidad posible de sustantivos.

- **Concordancia del determinante, adjetivo y sustantivo.** Como ya se ha comentado en el apartado 3, el problema de las concordancias se resuelve en el módulo de transferencia estructural, en concreto dentro del *chunker*. Para cada nueva secuencia de formas léxicas que se le introduce, el módulo de transferencia busca una coincidencia dentro del catálogo de patrones que previamente se han introducido. Al detectarlo, aplica la regla más larga que lleve asociada ese patrón en

concreto. En este caso se introduciría una regla que especificara que el determinante y el adjetivo deben tomar el género y el número del sustantivo

- **Orden sustantivo–adjetivo.** Al igual que el caso de la concordancia, el orden alemán adjetivo–sustantivo cambiaría a sustantivo–adjetivo en el módulo de transferencia estructural mediante el patrón y la regla de transferencia correspondientes.

Tanto el problema de la concordancia como del orden del sustantivo y el adjetivo son problemas que ya se han resuelto en pares como en→es o en→ca, por lo que se podrían utilizar la mismas reglas para el alemán.

- **Compuestos.** La composición es una de las dos formas mediante las que el alemán forma el léxico. En este proceso se produce una yuxtaposición de dos o más sustantivos, verbos, adjetivos, adverbios o preposiciones. Sólo la palabra que queda en la parte derecha del compuesto es la que se flexiona.

En el texto que se ha analizado se han encontrado tanto sustantivos compuestos como adjetivos compuestos. La particularidad de estas palabras, sobre todo en el caso de los sustantivos, es que en algunas ocasiones la traducción del compuesto no se deriva de la traducción de sus partes (*Haus*: casa + *Meister*: maestro→*Hausmeister*: conserje), mientras que en otras, sí (*Aktien*: acciones + *Markt*: mercado→*Aktienmarkt*: mercado de acciones).

La manera de tratar en Apertium con el primer tipo de compuestos, en el que su significado difiere del de los constituyentes, es incluirlos siempre en los diccionarios, puesto que es la única forma de traducirlos. Sin embargo, para evitar engrosar el diccionario con el segundo tipo de compuestos, convendría encontrar la forma de que el traductor consiguiera separar los elementos del compuesto y los tradujera y ordenara para formar la traducción correcta.

Otro problema al respecto es que, mientras que muchos compuestos se forman con una mera yuxtaposición de las palabras (*Gast*: huésped + *Haus*: casa → *Gasthaus*: casa de huéspedes), otros se caracterizan por incluir entre las palabras que forman el compuesto un elemento epentético, carente de significado y sin función dentro de

la flexión. En algunos casos el elemento epentético es la terminación de genitivo (*Bund*: federación + *Bank*: banco → *Bundesbank*: banco federal) y en otros la terminación del plural (*Kart*: carta + *Haus*: casa → *Kartenhaus*: castillo de naipes), con los posibles cambios vocálicos que puedan producirse (*Gans*: ganso + *Braten*: asado → *Gänsebraten*: ganso asado). Además, también hay casos en los que se suprime la última letra de la primera palabra (*Krone*: corona + *Prinz*: príncipe → *Kronprinz*: príncipe heredero). Todo esto representa un problema añadido a la traducción de los compuestos, ya que no se trata solo de poder separar el compuesto en sus constituyentes, sino identificar la extensión de los constituyentes, especialmente en el caso de supresión vocálica.

La aparición de este tipo de elementos epentético, característicos de las lenguas germánicas, ya ha sido abordada en Apertium para el noruego, aunque no ha sido resuelta (ver <http://wiki.apertium.org/wiki/Compounds>). En este caso se apunta a que una posible solución sería la de añadir etiquetas a las palabras, de manera que se pueda indicar cuáles podrían formar parte de un compuesto o aparecer como una palabra aislada, y cuáles son solo posibles dentro del compuesto, bien por incluir el conector epentético o por verse acortadas debido al proceso de composición. Así, palabras como *Krone* (corona) o *Geschichte* (historia), que constan únicamente de las formas *Krone* y *Geschichte* para el singular y *Kronen* y *Geschichten* para el plural, forman todos sus compuestos con las formas *Kron* y *Geschicht*, que no existen de forma aislada y, por lo tanto, son las que deberían etiquetarse como formas que tan solo son posibles dentro de un compuesto. Quizás debiera realizarse un análisis exhaustivo del mayor número posible de sustantivos compuestos y realizar una lista de qué palabras son las que realizan este tipo de supresión, de manera que baste una etiqueta por palabra para que Apertium pueda tratar todos los sustantivos derivados de la misma sin tener que incluirlos en su totalidad en los diccionarios.

- **Elección de la forma verbal.** Este problema es similar en muchos casos al de la concordancia. En realidad, se trata de un caso de ambigüedad, ya que el sistema tiene que decidir cuál de las posibles traducciones de una forma verbal ambigua es la correcta. Por lo tanto Apertium, tal y como ya se hace para el inglés, debería

poder obtener la traducción correcta con ayuda del desambiguador léxico categorial.

- **Posición del verbo.** En alemán, la posición del verbo en la oración varía según varios factores. Así, como vemos en los siguientes dos ejemplos, la transformación de una frase simple en una oración subordinada dentro de una oración compleja mayor implica que el verbo se desplaza al final de la oración.

*Ich **gehe** heute zur Arbeit* (2)

Yo voy hoy al trabajo

Hoy **voy** al trabajo

[*Ich bin sicher, dass*] *ich heute zur Arbeit **gehe*** (3)

[Yo estoy seguro de que] yo hoy al trabajo **voy**

[Estoy seguro de que] hoy voy al trabajo

Si la forma verbal elegida es compuesta, como por ejemplo en el caso del *Perfekt* (pretérito perfecto simple), el verbo auxiliar ocupará la posición que ocuparía la forma verbal simple, mientras que el participio se desplaza al final de la oración.

*Ich **lese** ein Buch* (4)

Yo **leo** un libro

Leo un libro

*Ich **habe** ein Buch **gelesen*** (5)

Yo **he** un libro **leído**

He leído un libro

Si la forma compuesta, además, está incluida dentro de una oración subordinada, toda ella se desplazará al final de la oración, pero con una inversión del verbo auxiliar.

*Ich **habe** einen Apfel **gegessen*** (6)

Yo **he** una manzana **comido**

Me **he comido** una manzana

*Carlos sagt, dass ich einen Apfel **gegessen habe*** (7)

Carlos dice que yo una manzana **comido he**

Carlos dice que me **he comido** una manzana

Esta circunstancia hace que el sistema deba ser capaz tanto de reordenar el verbo dentro de la oración para adecuar el texto al orden sintáctico español como de identificar los elementos que lo componen, a pesar de poder estar separados dentro de la oración por varios complementos.

*Ich **bin** heute vormittags mit deinem Vater im Einkaufszentrum **begegnet**.* (8)

Yo **he** hoy por la mañana con tu padre en el centro comercial **encontrado**

Esta mañana **me he encontrado** con tu padre en el centro comercial.

Este tipo de elementos deberían solucionarse en Apertium en el submódulo *interchunk*, dentro del módulo de transferencia estructural, en el que se podría definir los siguientes patrones y reglas dentro del archivo .t2x:

- Para las oraciones subordinadas, como en (7), un patrón que identifique la estructura <sujeeto> <sintagma₁> <sintagma₂> ... <sintagma_n> <participio> <verbo auxiliar> y una regla que la reordene como <sujeeto> <verbo auxiliar> <participio> <sintagma₁> <sintagma₂> ... <sintagma_n>

*Carlos sagt, dass ich einen Apfel **gegessen habe***

Carlos dice que yo una manzana **comido he**

*Carlos sagt, dass ich **habe gegessen** einen Apfel*

Carlos dice que yo **he comido** una manzana

Para las formas verbales no compuestas se introduciría el mismo patrón, pero sin el verbo auxiliar

- Para oraciones no subordinadas, como en (6), otro que identifique la estructura formada por <sujeeto> <verbo auxiliar> <sintagma₁> <sintagma₂> ... <sintagma_n> <participio> y la correspondiente regla que la reordene como <sujeeto> <verbo auxiliar> <participio> <sintagma₁> <sintagma₂> ... <sintagma_n>

*Ich **habe** einen Apfel **gegessen***

Yo **he** una manzana **comido**

Ich habe gegessen einen Apfel
Yo **he comido** una manzana

El problema es que deben establecerse distintas reglas en función del número de sintagmas presentes en la oración, por lo que si se quiere cubrir la posibilidad de que el módulo de transferencia estructural reordene el verbo dentro de una oración con cuatro sintagmas, el número de reglas que habrá que introducir en el *interchunk* será también cuatro: una regla para un sintagma, otra para dos, una tercera para tres sintagmas y una última para los cuatro.

- **Ausencia de artículo.** En general el alemán y el español emplean los artículos determinado e indeterminado de manera similar (Castell, 2011:253). No obstante hay algunos casos en los que en el texto original no se emplea artículo alguno, cuando en español sí que se requiere uno. Uno de ellos es el que aparece en las frases 3, 20, 26 y 34 del texto analizado. Se trata de frases que comienzan por un sustantivo que no precisa el artículo en alemán, pero sí en español. Sin embargo, aunque en estos casos en particular sea fácil establecer una regla para añadir el artículo, a priori no parece posible establecer de manera general una pauta que describa cuándo la ausencia de artículo en alemán requiere o no la introducción de uno en el texto en español, por lo que este problema necesita de un análisis en mayor profundidad, probablemente en base a un corpus textual amplio suficientemente representativo del que pueda extraerse una casuística.
- **Verbos separables.** Son verbos formados por la unión de una base verbal y un prefijo separable. La unión del prefijo matiza e incluso cambia el significado de la base verbal, pero ese no es el problema. La dificultad de este tipo de verbos es doble. Por un lado, al conjugarlos en determinados tiempos verbales —siempre y cuando no se trate de una oración subordinada—, la base verbal conjugada permanece en su posición, mientras que el prefijo pasa a ocupar el último lugar en la oración y permanece invariable. Este sería el caso, por ejemplo, del presente de indicativo.

Ich stelle einen Laptop auf den Tisch (stellen: poner) (9)
Yo **pongo** un portátil sobre la mesa

Ich stelle einen Laptop her (*herstellen*: fabricar) (10)
Yo **fabrico** un portátil

En el caso de que la frase sea subordinada, el verbo se mantiene sin separar al final de la oración.

Kennst du den neuen Laptop, der Toshiba jetzt herstellt? (11)
¿Conoces tú el nuevo portátil que Toshiba ahora fabrica?
¿Conoces el nuevo portátil **que fabrica** ahora Toshiba?

Por otro lado, el participio (*Partizip II*) de este tipo de verbos difiere del de los verbos normales. En el caso de los verbos separables, el participio se forma insertando la partícula *-ge-* entre el prefijo y la base verbal, en vez de añadirse en forma de prefijo, como sucede con los verbos no separables.

Ich habe einen Laptop hergestellt (12)
Yo **he** un portátil **fabricado**
He fabricado un portátil

Lo mismo sucede con las oraciones de infinitivo, en las que la preposición *zu* se inserta entre el prefijo y la base verbal.

Ich habe es versucht, einen Laptop herzustellen (13)
Yo **he** lo intentado, un portátil **fabricar**
He intentado fabricar un portátil

Para conseguir traducir el verdadero significado del verbo separable y no únicamente el de su base verbal, debe poder identificarse su prefijo como parte semánticamente indivisible del verbo separable.

En Apertium ya se han producido algunos intentos de lidiar con estos problemas, tanto en inglés con los *phrasal verbs* (ver http://wiki.apertium.org/wiki/M%C3%B3dulo_de_procesamiento_de_expresiones_separables) como para lenguas germánicas con los verbos separables, en concreto para el par afrikàans-inglés (ver http://wiki.apertium.org/wiki/Separable_verbs). Para este par de lenguas se ofrecen tres posibles soluciones, si bien hasta el momento todas comportan problemas sin solución. No obstante, una de ellas la podríamos utilizar en el caso del alemán para las formas compuestas de los verbos.

Consistiría en declarar todas las partículas que forman verbos separables, ya que estas son limitadas, y etiquetarlas como formas que solo pueden aparecer formando un verbo separable. Esta regla tan solo tiene alguna excepción, por lo que funcionaría en la mayoría de casos. Una vez declaradas las partículas separables, habría que elaborar una regla que especificara que si el sistema encuentra cualquiera de ellas seguida de *-ge-* o *-zu-* trate el verbo separable de manera unitaria y lo flexione como un participio si aparece la partícula *-ge-* o como infinitivo si es *-zu-*.

En cambio, si la forma superficial del verbo es una forma simple, la partícula separable se encontrará siempre aislada al final de la oración. Una posible solución para reunir el prefijo y la base verbal podría ser la de hacer que la transferencia léxica de los verbos no se realizara completamente hasta después del módulo de transferencia estructural. Como ya hemos comentado, sería necesario declarar todas las partículas que pueden formar parte de un verbo separable como paradigmas en los ficheros e introducir un patrón para el módulo de transferencia en el que se indique que la partícula es parte semánticamente indivisible del verbo separable y debe traducirse como tal si esta aparece antes de un signo ortográfico (coma, punto, dos puntos, signo de interrogación, etc.) o precediendo a las conjunciones *und* (y) y *oder* (o). Estos son los únicos casos en los que una partícula puede aparecer en esa posición en la oración. Si al llegar al final de la oración el sistema no ha encontrado ninguna partícula, procedería entonces a traducir la forma verbal tal cual ha aparecido anteriormente en la oración.

- **Traducción del verbo *sein* por *ser* o *estar*.** El verbo *sein*, al igual que el verbo inglés *to be*, representa un problema debido a que hay multitud de excepciones y giros en los que los verbos *ser* y *estar* no participan de la regla general, e incluso en algunas ocasiones ni siquiera se traduce por ninguno de estos dos verbos.

Un ejemplo serían las siguientes frases:

Ich bin David (14)
Yo **soy** David
Soy David

Ich bin nervös (15)

Yo **Estoy** nervioso
Estoy nervioso

Ich bin alt (16)
Yo **soy** mayor
Soy rápido

Ich bin 35 Jahre alt (17)
Yo **soy** 35 años viejo
Tengo 35 años

En el par inglés-español ya hay solucionados algunos de los problemas relacionados con los verbos ser y estar. Sin embargo, en alemán hay un problema añadido ya que las formas verbales compuestas de pasado se forman en alemán usando como auxiliar el verbo *sein* (ser) para los verbos que impliquen movimiento o cambio de estado y el verbo *haben* (tener) para el resto, por lo que según el contenido de la oración el módulo de transferencia estructural debería analizar el resto de elementos para tomar una decisión al respecto.

David ist in der Schule (18)
David **está** en el colegio

vs.

David ist in die Schule gegangen (19)
David **está** en el colegio **ido**
David **ha ido** al colegio

En cualquier caso, convendría un análisis exhaustivo de este problema para ver qué regularidades pueden extraerse con el fin de poder establecer patrones que puedan ser reconocidos por el módulo de transferencia estructural, de manera que mediante dichos patrones puedan reordenarse los elementos verbales en el módulo *interchunk*, especialmente en frases cortas.

- **Pronombres de relativo.** El problema con los pronombres de relativo alemanes es que, cuando estos toman la forma del artículo determinado (*der, die, das*), aparecen declinados según la función sintáctica de la oración de relativo, en cuyo caso la traducción sería *que* o *cual*:

Das ist das Haus, das ich gekauft habe (20)

Esto es el casa, el yo comprado he

Esta es la casa que yo he comprado

Dieser ist der Mann, dem ich das Haus gekauft habe (21)

Este es el hombre, al yo el casa comprado he

Este es el hombre al que/cual he comprado la casa

La dificultad de la traducción viene dada, por un lado, por la problemática de tratar con el sistema de declinación alemán, que ya hemos comentado en el primer apartado. Ante la ambigüedad, el sistema debe ser capaz de decidir qué forma morfológica se está empleando, si se trata realmente de un artículo o de un pronombre. El desambiguador léxico categorial sería en este caso el encargado de tomar esta decisión en Apertium.

Por otro lado, hay que tener también en cuenta que el verbo que se emplea en la oración de relativo puede requerir un complemento preposicional, por lo que el pronombre de relativo puede aparecer acompañado de la preposición que necesite el verbo, en vez de presentarse solo.

Der Bus, **auf den** ich warte, kommt in 10 Minuten. (22)

El autobús, **sobre el** yo espero, viene en 10 minutos

El autobús **al que** espero viene en 10 minutos

En estos casos es necesario que el sistema pueda identificar la relación semántica entre la preposición y el verbo para ofrecer el equivalente adecuado, ya se trate de un verbo con o sin complemento preposicional en español. Este problema debería poder solucionarse mediante una regla introducida en el módulo de transferencia estructural. Para que el sistema reconozca la preposición se puede aprovechar la propia estructura de la oración de relativo en estos casos, que tal y como se puede ver en el ejemplo (22) sigue el patrón <coma><preposición alemana>...<verbo alemán><coma>, de manera que se sustituyera por otra del tipo <que>....<verbo>, en la que se eliminan comas y se sustituye el verbo alemán por el verbo español, en este caso con complemento preposicional.

- **Posición del pronombre enclítico.** En alemán, hay ocasiones en que los pronombres átonos siguen al verbo, de manera que aparecen como enclíticos. Es el caso, por ejemplo, de los verbos reflexivos, que se forman con los pronombres reflexivos *mich, dich, sich, uns, euch, sich* (*me, te, se, nos, os, se*) más el verbo en cuestión. El problema radica en que, mientras que en español generalmente las posiciones del pronombre y el verbo no cambian en los verbos reflexivos, en alemán la posición depende del tiempo verbal empleado (simple o compuesto) y del tipo de oración (principal o subordinada).

Por ejemplo, en el caso de oraciones con tiempos simples, el pronombre reflexivo se sitúa detrás del verbo.

*Ich duche **mich** am Morgen* (23)
 Yo ducho **me** por la mañana
 Yo **me** ducho por la mañana

Para este tipo de casos ya se ha encontrado solución en el par inglés-español, de manera que frases como *I see **myself** in the mirror* se traducen correctamente como *Yo **me** veo en el espejo.*

Sin embargo, en otros casos como por ejemplo en oraciones subordinadas alemanas, el pronombre aparece directamente tras el sujeto y fuera del verbo compuesto.

*Ich erinnere **mich** nicht, ob ich **mich** am Morgen geduscht **habe*** (24)
 Yo acuerdo de **me** no si yo **me** por la mañana duchado **he**
 No **me** acuerdo de si **me** he duchado por la mañana

Nuevamente, para solucionar este problema en Apertium se pueden definir un patrón específico para este tipo de casos y una regla de transferencia para el mismo. De hecho, puesto que la posición normal del pronombre reflexivo en español es tras el sujeto, una de las posibilidades que se podría incluir en el apartado de la posición del verbo es la de escribir una regla para las oraciones subordinadas que especifique que, si tras el sujeto hay un pronombre reflexivo, el verbo pase del final de la oración a la posición inmediatamente posterior al pronombre reflexivo.

6. Conclusión

Tal y como se ha visto a lo largo del trabajo, hay multitud de problemas que deben resolverse para conseguir que Apertium realice una traducción aceptable de un texto alemán al español. Algunos de estos problemas tienen fácil solución, como por ejemplo la concordancia entre determinante, sustantivo y adjetivo, o la inversión del orden sustantivo-adjetivo, máxime teniendo en cuenta que se trata de cuestiones que ya se han abordado para otros pares de lenguas, principalmente el par inglés-español.

Sin embargo, en otros casos la solución no está tan clara. Se trata de problemas muy ligados a la lengua alemana en particular y a las lenguas germánicas en general. Así, el tratamiento de las formas compuestas, los verbos separables y la multitud de posiciones que el verbo puede ocupar dependiendo del tipo de oración que se trate son tres elementos para los que en el presente trabajo se ha apuntado a posibles soluciones, si bien estas soluciones son parciales. Además, queda por ver que las estrategias presentadas puedan realmente acabar implementándose en la plataforma Apertium y sirvan para solventar los problemas para los que han sido propuestas.

Hay que tener en cuenta también que, a día de hoy, los archivos de trabajo del par alemán-español no cuentan con la información suficiente para poder trabajar con la herramienta, por lo que la mera introducción de reglas no basta. Por ello, es necesario empezar por definir categorías e introducir léxico en los diccionarios para que pueda realmente comprobarse si las reglas funcionan de manera sistemática o no. Así, el presente trabajo pretende también servir de base para otros futuros, de enfoque quizás más práctico y orientados, por ejemplo, a la mejora léxica del par alemán-español, al análisis en mayor profundidad de algunas de las cuestiones concretas que ya se han abordado en este o a la introducción de las reglas en los archivos del par para comprobar que realmente funcionan.

Anexo

Tabla de traducciones y clasificación

Fuente		http://www.zeit.de/wirtschaft/2016-03/wasserverbrauch-weltwassertag-wasser-fussabdruck/komplettansicht		
ALEMÁN (ORIGINAL)		ESPAÑOL (TA)	ESPAÑOL (TH)	TIPO DE PROBLEMA
1	Wie viel Wasser steckt im Burger?	Cómo mucho Agua poned en el Hamburguesa?	¿Cuánta agua contiene una hamburguesa?	ULMP, M, EFV, CDSA, M
2	Der Wasserverbrauch eines jeden bemisst sich nicht nur an der Dauer des morgendlichen Duschens.	El Consumo de agua de uno cada uno mide se no sólo en el Duración del matutino Ducha.	El consumo de agua de cada uno no se mide solo durante la duración de la ducha matutina.	CDSA, M, CO, ULMP, PPR, CS, CDSA, M, CDSA, OSA, M
3	Essgewohnheiten beeinträchtigen viel mehr den eigenen Wasserfußabdruck.	Hábitos alimenticios estorban mucho más el propio Huella de agua.	Los hábitos alimenticios influyen mucho más la huella de agua propia.	AA, CO, EFV, CDSA, OSA, M, CO
4	Von Judith Colling	De Judith Colling	De Judith Colling	M, M
5	Wassersparen ist eine Lieblingsdisziplin der Deutschen.	Ahorrar de agua es una Disciplina favorita el Alemanes.	Ahorrar agua es una disciplina favorita de los alemanes	CO, VSE, CDSA, M, CO, CDSA, M
6	Trotzdem machen der wasserregulierende Duschkopf, eine Waschmaschine mit automatischem Wasserstopp oder der Verzicht auf einen grünen Rasen im Sommer die Deutschen noch nicht zu echten Wassersparern.	Sin embargo hacen el reguladores de agua Alcachofa de ducha, una Lavadora con automático Parada de agua o el Renuncia en uno florecer Césped en el Verano la Alemán todavía no demasiado auténtico A ahorradores de agua.	Sin embargo, la alcachofa de ducha reguladora de agua, una lavadora con parada automática de agua, o la renuncia a un césped verde en verano no hacen a los alemanes unos verdaderos ahorradores de agua.	PV, CDSA, OSA, CO, M, CO, CDSA, CO, CDSA, OSA, M, CDSA, M, CO, CS, M, CDSA, OSA, M, OSA, M, CO
7	Die Umweltorganisation WWF erinnert zum Weltwassertag an diesem Dienstag daran, dass der direkte Verbrauch nur einen kleinen Anteil des gesamten Wasserkonsums ausmacht.	La Organización ambiental WWF recordad al Día mundial del agua en a esto Martes en eso, que el directos Consumo sólo uno pequeño Parte del completo consumo de agua apaga.	La organización ecológica WWF recuerda este martes en el Día mundial del agua que el uso directo solo constituye una pequeña parte del consumo total de agua.	CDSA, CO, M, M, PV, EFV, CDSA, M, CO, CDSA, M, CDSA, OSA, M, PV, EFV, CDSA, OSA, M, CDSA, OSA, M, CO
8	Zwar liegt der tägliche Verbrauch pro Kopf in Deutschland bei nur noch etwa 120 Litern.	En verdad estad el diarios Consumo por Cabeza en Alemania en sólo todavía por ejemplo 120 A litros.	En realidad, el consumo diario de agua per cápita en Alemania se sitúa aún en unos 120 litros.	PV, EFV, CDSA, OSA, M, ULMP, CS, ULMP, M
9	Dafür umfasst der Wasserfußabdruck eines	Para eso abarcado el Huella de agua de uno De ciudadano de	Para la cifra, la huella de agua de un ciudadano	AC, PV, EFV, CDSA, CO,

	Durchschnittsbürgers nach Berechnungen des WWF 5.288 Liter am Tag, was etwa 25 vollen Badewannen entspricht.	promedio a Cálculos del WWF 5.288 Litro en el Día, qué por ejemplo 25 lleno Bañeras corresponde.	medio asciende según cálculos de WWF a 5.288 litros al día, lo que corresponde a unas 25 bañeras llenas.	M, CDSA, M, P, M, CO, M, CDSA, OSA, P, M, CO
10	Der Wasserfußabdruck beinhaltet neben dem direkten Verbrauch auch das sogenannte virtuelle Wasser.	El Huella de agua contenido al lado de al directo Consumo también eso llamados virtuales Agua.	La huella de agua incluye también además del uso directo la llamada agua virtual.	CDSA, M, CO, EFV, P, CDSA, OSA, M, CDSA, OSA, M
11	Darunter versteht die Organisation Water Footprint Network alle Wassermengen, die bei der Herstellung von Lebensmitteln und Industriegütern verbraucht werden, also in den Endprodukten, die wir konsumieren, enthalten sind.	Debajo de eso comprended la Organización Water Footprint Red todos Cantidades de agua, la en el Fabricación de A alimentos y A propiedades industriales consumido son, pues en el Productos finales, la nosotros consumen, contenido son.	Por ese nombre entiende la organización Water Footprint Network todas las cantidades de agua que se emplean durante la fabricación de alimentos y bienes industriales, es decir los productos finales, que consumimos.	AC, PV, EFV, CDSA, M, CDSA, M, CO PR, CDSA, M, CO, CDSA, M, CO, EFV, CDSA, M, CO, PR, EFV, EFV
12	Die Berechnung exakter Zahlen ist kompliziert und kann durchaus zu unterschiedlichen Ergebnissen führen.	La Cálculo más exacto Números es complicado y puedo absolutamente demasiado diferente resultados estar en cabeza.	El cálculo de las cifras exactas es complicado y puede llevar a resultados completamente distintos.	CDSA, M, AA, CDSA, OSA, M, VSE, PV, EFV CDSA, OSA, AA, M,
13	Je nachdem wo und unter welchen Bedingungen das Produkt gefertigt wird, kann der Wasserverbrauch stark variieren.	Respectivamente después de que dónde y debajo de cuál Condiciones eso Producto elaborado es, puede el Uso de agua fuerte variar.	Dependiendo de dónde y bajo qué condiciones se elabore el producto, el uso de agua puede variar enormemente.	ULMP, CDSA, M, PV, EFV, CDSA, M, PV, EFV, CDSA, M, CO
14	Das erklärt, weshalb das Umweltbundesamt lediglich 3.900 Liter als Wasserfußabdruck des durchschnittlichen Deutschen angibt, etwa 1.000 Liter weniger als der WWF.	Eso explicado, por qué el Oficina alemana de Medio Ambiente solo 3.900 Litro como Huella de agua del medio Alemanes señala, aproximadamente 1.000 Litro menos como el WWF.	Eso explica por qué la Oficina alemana de Medio Ambiente solo calcula una huella de agua del alemán medio de 3.900 litros, unos 1.000 litros menos que la WWF.	EFV, CDSA, M, CO, PV, EFV, CDSA, M, AA, Mn CO, CDSA, OSA, M, CDSA, M, ULMP, CDSA, M
15	Dieser berechnet seine Zahlen aufgrund der weltweiten Durchschnittswerte, die aus den Studien des Water Footprint Networks hervorgehen.	Este calculado sus Números a causa de el universal Valores medios, la el Estudios del Water Footprint Network resultar.	Esta calcula sus cifras teniendo en cuenta los valores medios mundiales recogidos en los estudios de Water Footprint Network.	EFV, CDSA, M, CDSA, OSA, M, CO, PR, CS, CDSA, M, CDSA
16	Diese Grunddaten nutzt der WWF, um den virtuellen Wassergehalt weit verbreiteter Produkte darzustellen.	Estos Datos básicos servid el WWF, sobre el virtual contenido en agua lejos más difundido Productos representar.	La WWF usa estos datos básicos para representar el contenido de agua virtual de productos muy difundidos.	CDSA, M, CO, M, EFV, PV, VS, CDSA, OSA, CDSA, M,
17	Zum Beispiel den Wassergehalt eines Hamburgers mit den typischen Zutaten Fleisch, Salat,	Al Ejemplo el Contenido de agua de uno De hamburguesa con el típico Ingrediente Carne,	Por ejemplo, del contenido de agua de una hamburguesa con los	ULMP, CDSA, M, CO, CDSA,

	Käse und verschiedene Soßen (siehe Grafik).	Ensalada, Queso y diferentes Salsas (véase Gráfico).	ingredientes típicos: carne, lechuga, queso y distintas salsas (véase el gráfico).	M, CDSA, OSA, M, M, M, M, CDSA, AA, M
18	Dabei ist zu beachten, dass der Verbrauch eines Burgers, der nur mit regionalen Produkten zubereitet wird, stark von den 2.453 Litern des WWF-Durchschnittsburgers abweichen kann.	En eso es demasiado tienen en cuenta, que el Uso de uno hamburguesa, el sólo con regional A productos preparado es, fuerte de el 2.453 A litros del WWF-Durchschnittsburgers desviarse puede.	A este respecto, hay que tener en cuenta que el uso de una hamburguesa que solo se prepara con productos regionales puede diferir enormemente de los 2.453 litros calculados por la WWF para una hamburguesa media.	CA, ULMP, CDSA, M, CDSA, M, PR, PV, EFV, CDSA, OSA, M, PV, CDSA, M, CDSA, M, CDSA, M, CO,
19	Wie kommt das Wasser in den Burger?	Cómo venid eso Agua en el hamburguesa?	¿Cómo llega el agua a la hamburguesa?	EFV, CDSA, M, CDSA, M
20	Rindfleisch, die Hauptkomponente eines jeden Burgers, ist dabei ein besonders verschwenderisches Nahrungsmittel.	Carne de vaca, la Componente principal de uno cada uno De hamburgués, está en eso un especialmente pródigo Alimento.	La carne de vacuno, el componente principal de cada hamburguesa, es al mismo tiempo un alimento especialmente derrochador.	AA, CDSA, M, CO, CDSA, M, VSE, CA, CDSA, OSA, CO, M
21	Im weltweiten Durchschnitt benötigt die Produktion eines Kilos 15.415 Liter Wasser.	En el universal Promedio necesitado la Producción de uno Kilos 15.415 Litro Agua.	De media mundial la producción de un kilo necesita 15.415 litros de agua.	CDSA, OSA, M, PV, EFV, CDSA, M, CDSA, M, CDSA, M, M
22	Am meisten davon entfällt auf den Anbau von Futtermitteln wie Soja und Mais, nämlich 15.300 Liter.	En el meisten de eso se suprime en el Cultivo de A piensos cómo Soja y Maíz, mismo 15.300 Litro.	La mayoría de ella se corresponde con el cultivo de piensos como la soja y el maíz, es decir, 15.300 litros.	ULMP, CA, PV, CS, EFV, CDSA, M, AA, M, CO, AA, M, AA, M, CDSA, M
23	Bei nur einer Mahlzeit verspeist der Burger-Fan so fast die Hälfte seines durchschnittlichen Tagesverbrauchs.	En sólo a una Comida comido el Hamburguesa-Fan así casi la Mitad del suyo medio De uso diario.	En una sola comida, el fan de las hamburguesas consume de esta manera casi la mitad de su uso diario medio.	CDSA, M, CO, PV, EFV, CDSA, M, M, CDSA, M, CDSA, OSA, M, CO
24	Aber Fleischliebhaber können den eigenen Fußabdruck auch ohne Verzicht auf den Burger-Genuss reduzieren.	Sin embargo Amante carnal pueden el propio Huella también sin Renuncia en el Hamburguesa-Placer reducen.	Sin embargo, los amantes de la carne también pueden reducir su huella sin renunciar al placer de una hamburguesa.	AA, M, CO, PV, EFV, CDSA, OSA, M, CO, CS, M, CDSA, CO, M, M
25	Während nämlich ein Kilo Rindfleisch aus Brasilien einen Abdruck von 19.488 Litern hinterlässt, fällt ein Kilo aus Österreich nur mit 8.300 Litern ins Gewicht.	Durante mismo un Kilo Carne de vaca de Brasil uno Copia de 19.488 A litros deja, cae un Kilo de Austria sólo con 8.300 A litros en lo Peso.	Mientras que un kilo de carne de vacuno de Brasil deja tras de sí una huella de agua de 19.488 litros, un kilo procedente de Austria solo supondría 8.300 litros.	CDSA, M, M, PV, EFV, CO, CDSA, M, CDSA, M, CDSA, M, ULMP, CDSA, M, CDSA, M
26	Vegetarier sparen Wasser	Vegetariano ahorrar Agua	Los vegetarianos ahorran	AA, EFV, M

			agua	
27	Genau zu dieser Differenzierung und Achtsamkeit ruft der WWF die Verbraucher anlässlich des Weltwassertags auf.	Exactamente demasiado este Diferenciación y Cuidado llamado el WWF la Consumidor con ocasión de del Weltwassertags en.	Precisamente sobre esta diferenciación y sobre la concienciación, la WWF llama la atención de los consumidores con ocasión del Día mundial del agua.	CDSA, M, M, VS, CDSA, M, CDSA, M, CO
28	Aber eine auf den ersten Blick niedrige Wasserverbrauchszahl allein identifiziert das Produkt noch nicht als wirklich umweltschonend.	Sin embargo una en el primero Mirada bajos Número de uso de agua solo identificado eso Producto todavía no como real protegiendo medio ambiente.	Sin embargo, una cifra de uso de agua a primera vista baja no identifica por sí misma un producto como respetuoso con el medio ambiente.	CDSA, ULMP, OSA, M, CO, EFV, CDSA, M,
29	Es macht einen Unterschied, ob diese Wassermenge in einem wasserreichen Land wie Deutschland oder einem sehr trockenen Land wie zum Beispiel Spanien verbraucht wird.	Ello hace un Diferencia, si estos Cantidad de agua en a uno rico en agua País cómo Alemania o a uno muy seco País cómo al Ejemplo España usado es.	Hay una gran diferencia entre que esa cantidad de agua se use en un país con abundancia de agua como Alemania o en uno muy seco como España.	CS, CDSA, M, CDSA, M, CO, PV, EFV, CDSA, OSA, CO, M, CDSA, OSA, M, ULMP,
30	Deutschland importiert etwa 50 Prozent des virtuellen Wasserverbrauchs	Alemania importado por ejemplo 50 Por ciento del virtual De uso de agua.	Alemania importa aproximadamente un 50 por ciento del uso de agua virtual.	EFV, AA, M, CDSA, OSA, M, CO
31	Viele sehr wasserintensive Importgüter wie Kaffee und Baumwolle stammen dabei aus Ländern, in denen Wasserknappheit herrscht oder die mit landwirtschaftlichen Problemen kämpfen.	Muchos muy intensivos de agua Propiedades de importación cómo Café y Algodón proceden en eso de A países, en a esos Escasez de agua gobernada o la con agropecuario A problemas luchar.	Muchos de los bienes de importación que requieren mucha agua, como el café y el algodón, provienen de países en los que reina la escasez de agua o en los que luchan con problemas agrícolas.	CDSA, OSA, M, CO, M, M, CS, AA, M, PR, PV, EFV, AA, PR, PV, EFV, CDSA, OSA, CO, M
32	So ist Brasilien zwar eigentlich ein wasserreiches Land, doch aufgrund unkontrollierter Wasserverschmutzung durch die Landwirtschaft und dem einseitigen Anbau von Futterpflanzen wie Mais und Soja ist Wasser dort ein kostbares Gut.	Así está Brasil en verdad en realidad un rico en agua País, no obstante a causa de más incontrolado Contaminación del agua por la Agricultura y a eso unilateral Cultivo de Plantas forrajeras cómo Maíz y Soja es Agua allí un valioso Bien.	Así, aunque Brasil es en realidad un país rico en agua, el agua allí es un bien preciado debido a la contaminación incontrolada del agua a causa de la agricultura y de los monocultivos de plantas forrajeras como el maíz y la soja.	CDSA, OSA, CO, M, CDSA, AA, CDSA, OSA, M, CO, M, CO, CDSA, OSA, M, M, CO, AA, M, AA, M, VSE, AA, M, CDSA, OSA, M
33	Brasilien ist der weltweite größte Sojaexporteur.	Brasil es el universales mayores Exportador de sojas.	Brasil es el mayor exportador de soja del mundo	VSE, CDSA, OSA, M, CO
34	80 Prozent der Sojaernte werden zu Mehl verarbeitet und landen als Futter in Tiertrögen.	80 Por ciento el Cosecha de sojas son demasiado Harina transformado y aterrizar como Forraje en Artesas de animales.	El 80 por ciento de la cosecha de soja se transforma en harina y acaban como forraje en las artesas de los animales.	AA, CDSA, M, CO, PV, M, M, M, CO
35	Auch bei Rindfleisch aus europäischen Ländern ist es	También en Carne de vaca de europeo a países está ello a	A causa de eso, es también muy probable que en el	AA, M, CO, CDSA, OSA,

	deshalb sehr gut möglich, dass das Futter für die Kuh aus Brasilien stammt.	causa de eso muy bueno posible, que eso Forraje para la Vaca de Brasil proceded.	caso de la carne de vacuno de los países europeos el forraje para las vacas provenga de Brasil.	M, VSE, PV, ULMP, CDSA, M, CDSA, M, CS, M
36	So kann kostbares Wasser aus Brasilien indirekt in einem Burgerbratling aus Österreich landen, der auf den ersten Blick Wasserressourcen schont.	Así puede valioso Agua de Brasil indirecto en a uno hamburguesa vegetariana de Austria aterrizar, el en el primero Mirada Recursos de agua proteged.	Así, la preciada agua de Brasil puede acabar indirectamente en una hamburguesa vegetal de Austria, que en principio es respetuosa con los recursos hídricos.	CDSA, OSA, M, M, PV, EFV, CDSA, M, CO, M, ULMP, PV EFV, M, CO
37	Fleischesser hinterlassen generell einen größeren Abdruck als Vegetarier.	Comedor carnal dejado general uno mayor Copia como Vegetariano.	Los comedores de carne dejan atrás por lo general una mayor huella que los vegetarianos.	AA, EFV, VS, CDSA, OSA, M, AA, M
38	Obwohl Soja eine wichtige Grundlage vegetarischer und veganer Gerichte ist, macht der direkte Verzehr der Bohne nur einen geringen Anteil des weltweiten Anbaus aus.	Aunque Soja una importantes Base más vegetariano y vegano Platos es, haced el directos Consumo el Haba sólo uno pequeño Parte del universal cultivo de.	Aunque la soja es una base importante para los platos vegetarianos y veganos, el consumo directo del haba representa solo una reducida parte del cultivo mundial.	AA, PV, EFV, VSE, CDSA, OSA, M, AA, AA, CDSA, M, PV, VS, EFV, CDSA, OSA, M, CDSA, M, CDSA, OSA, M, CDSA, M
39	Durch Anbau von Futtermitteln und zusätzlich die Haltung der Tiere wird zur Herstellung eines Kilos Fleisch deutlich mehr virtuelles Wasser benötigt, als wenn der Mensch dieselbe Menge einer pflanzlichen Alternative selbst verspeist.	Por Cultivo de A piensos y adicional la Postura el Animales la Fabricación de uno Kilos Carne claro más virtual Agua necesitado, como cuando el Persona la misma Cantidad Multitud a una vegetal Alternativa de forma misma comido.	Mediante el cultivo de piensos y la cría de ganado, se necesita para la producción de un kilo de carne bastante más agua virtual que si las personas comieran la misma cantidad de una alternativa vegetal.	AA, M, M, CDSA, M, CDSA, M, M, CDSA, M PV, EFV, CDSA, OSA, M, PV, EFV, CDSA, M, CDSA, M, CDSA, OSA, M
40	Der eigene Wasserfußabdruck	El propios Huella de agua	La huella de agua personal.	CDSA, OSA, M, CO
41	Water Footprint Network bietet einen Kalkulator an, der den ganz persönlichen Wasserfußabdruck bestimmen soll.	Water Footprint Red ofrece uno Calculador en, el el completamente personal Huella de agua disponen debe.	Water Footprint network ofrece una calculadora que permite determinar la huella de agua personal.	VS, EFV, CDSA, M, PR, PV, EFV, CDSA, OSA, M, CO,
42	In Abhängigkeit von Nationalität, Geschlecht, Fleischverzehr und verfügbarem Jahresgehalt wird der eigene Verbrauch geschätzt.	En Dependencia de Nacionalidad, Sexo, Consumo carnal y disponible Sueldo anual es el propios Uso estimado.	El uso individual se calcula con arreglo a la nacionalidad, el sexo, el consumo de carne y el salario disponible.	ULMP, AA, M, AA, AA, M,CO, CDSA, OSA, M, CO, PV, EFV, CDSA, OSA, M
43	Auch hier sollten die Ergebnisse nur als grobe Richtwerte verstanden werden.	También aquí deberían la Resultados sólo como gruesos Valores de tipo comprendido son.	También en este caso, los resultados deben entenderse como valores indicativos aproximados.	PV, EFV, CDSA, M, CDSA, OSA, M, CO,

44	In der Realität ist der Verbrauch natürlich viel stärker von der persönlichen Lebensführung bestimmt.	En el Realidad es el Uso natural mucho más fuerte de el personal estilo de vida específico.	En realidad, el uso está por supuesto mucho más determinado por el estilo de vida personal.	VSE, PV, EFV, CDSA, M, ULMP, CDSA, OSA, M, CO,
45	Nimmt man fiktiv an, dass ein deutscher Vorstandschef sein gesamtes Nettoeinkommen konsumiert und ein großer Fleischliebhaber ist, hinterlässt er nach der Berechnung von Water Footprint Network einen Wasserfußabdruck von 14.463 Liter am Tag.	Coge uno ficticio en, que un más alemán Jefe de junta directiva ser completo Sueldo neto de forma misma consumido y un grande Amante carnal es, deja él a el Cálculo de Water Footprint Red uno Huella de agua de 14.463 Litro en el Día.	Pongamos por ejemplo que un jefe de una junta directiva alemán que gasta todo su sueldo neto y es una gran amante de la carne. Según los cálculos de Water Footprint Network, deja tras de sí una huella de agua de 14.463 litros al día.	ULMP, VS, EFV, CDSA, OSA, M, CO, PV, EV, CDSA, OSAM, CO, PV, VSE, EFV, PV, EFV, CDSA, M, CDSA, M, CO, CDSA, M, EF
46	Der Abdruck einer vegetarisch lebenden Studentin, die den Bafög-Höchstsatz bezieht, ist mit 1.633 Litern deutlich unscheinbarer.	El Copia a una vegetariano viviendo Estudiante, la el Tarifa máxima de Bafög recibid, es con 1.633 A litros claro más insignificante.	La huella de una estudiante vegetariana que recibe la cuantía máxima del crédito para los estudios es con 1.633 litros claramente incomparable.	CDSA, M, CDSA, OSA, M, PR, PV, EFV, CDSA, CO, M CDSA, M

7. Bibliografía

Arnold, D.J., Balkan, L., Meijer, S., Humphreys, R.L. and Sadler, L. (1994) *Machine Translation: an Introductory Guide*. Londres: Blackwells-NCC.

Castell, Andreu (2011) *Gramática de la lengua alemana*. Madrid: Editorial idiomas

Corcoll, Brigitte y Roberto. (2006) *Programm. Alemán para hispanohablantes*. Barcelona: Herder.

Fernández Guerra, Ana (2000) *Machine Translation. Capabilities and Limitations*. Valencia: Universitat de València, Lengua Inglesa, D. L.

Forcada, Mikel L. (2010) "Machine Translation Today", en Yves Gambier, Luc Van Doorslaer, ed., *Handbook of Translation Studies*, vol. 1, 215-223. Amsterdam ; Philadelphia : John Benjamins Publishing

Forcada, Mikel L.; Ivanov Bonev, B.; Ortiz Rojas, S.; Pérez Ortiz, Juan A.; Sánchez Martínez, F.; Ramírez Sánchez, G.; Montava, Marco A. y Tyres, Francis M. (2010) *Documentation of the Open-Source Shallow-Transfer Machine Translation Platform Apertium*. Grup Transducens, Universitat d'Alacant. Versió electrònica: <http://xixona.dlsi.ua.es/~fran/apertium2-documentation.pdf>

Forcada, Mikel L., Mireia Ginestí-Rosell, Jacob Nordfalk, Jim O'Regan, Sergio Ortiz-Rojas, Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, Francis M. Tyres, (2011) "Apertium: a free/open-source platform for rule-based machine translation", en *Machine Translation*, (Special Issue on Free/Open-Source Machine Translation) 25:2, 127-144

Hutchins, John W.; Somers, Harold (1995) *Introducción a la traducción automática*. Madrid: Visor Distribuciones, S.A.

Hutchins, John (2003) "ALPAC: The (IN)Famous Report", en Sergei Nirenburg, Harold Somers y Yorick Wilks,ed., *Readings in Machine Translation*, 131-135. Cambridge London: The MIT Press

Sager, J.C. (1994) *Language engineering and translation: Consequences of automation*. Amsterdam: Benjamins.

Trujillo, Arturo (1999) *Translation Engines: Techniques for machine translation*. Londres: Springer