

Sistema web para la traducción automática interactiva *

Vicent Alabau, Francisco Casacuberta, Luis A. Leiva, Daniel Ortiz-Martínez,
Germán Sanchis-Trilles

Institut Tecnològic d'Informàtica

Universitat Politècnica de València

{valabau,fcn,luileito,dortiz,gsanchis}@iti.upv.es

Resumen

Los sistemas modernos de traducción automática están todavía muy lejos de proporcionar traducciones perfectas. Una alternativa a estos sistemas es la llamada traducción automática interactiva, donde el conocimiento del traductor humano se combina con el sistema de traducción automática para conseguir un resultado de muy alta calidad. En este trabajo presentamos un sistema estadístico de traducción automática interactiva que es capaz de colaborar inteligentemente con el usuario para aumentar la productividad y mejorar la ergonomía del sistema. Además, se presenta un prototipo web mediante el cual es posible probar la utilidad de dicho sistema.

1. Introducción

Los avances en las tecnologías de la información en la sociedad moderna han llevado a la necesidad de métodos más eficientes para la traducción de documentos. Es importante resaltar que los sistemas actuales de traducción

automática (MT, del inglés machine translation) no son capaces de producir textos con calidad suficiente para ser usados directamente [10, 9, 2]. De hecho, los sistemas MT requieren normalmente que un humano revise y postedite los resultados para alcanzar altas cotas de calidad de traducción.

Una forma de sacar ventaja de los sistemas MT es combinarlos con el conocimiento del traductor humano, constituyendo el llamado paradigma de traducción asistida por computador (CAT, del inglés computer-assisted translation).

Una contribución importante a las tecnologías de CAT fue el proyecto TransType (TT) [12, 7, 6, 17]. Este proyecto acarrió un interesante cambio de paradigma en el cual las tecnologías de MT basadas en datos se incrustaban en un entorno de traducción interactivo.

Siguiendo estas ideas, Barrachina et al. [3] propone el uso de sistemas de traducción automática estadística (SMT, del inglés statistical machine translation) para producir hipótesis completas de frases en el idioma destino, o porciones de las mismas, que pueden ser total o parcialmente aceptadas y corregidas por traductores humanos. Cada segmento de texto corregido por el humano es usado por el sistema SMT como información adicional para proporcionar sugerencias mejoradas. En este artículo, también nos centramos en la aproximación interactiva a la traducción automática (IMT, del inglés interactive machine translation).

La Figura 1 ilustra un ejemplo típico de sesión de IMT. Inicialmente, al usuario se le pro-

* Trabajo financiado parcialmente por la EC (FEDER/FSE) y el MEC/MICINN español en el marco del proyecto MIPRCV (CSD2007-00018) bajo el programa "Consolider Ingenio 2010", los proyectos iTransDoc (TIN2006-15694-CO2-01), iTrans2 (TIN2009-14511) y MITTRAL (TIN2009-14633-CO3-01), y la beca AP2005-4023. También financiado por la EC (FEDER) y el MITyC español bajo el proyecto erudito.com (TSI-020110-2009-439) y por la Generalitat Valenciana mediante las subvenciones Prometeo/2009/014 y GV/2010/067, y por el "Vicerrectorado de Investigación de la UPV" mediante la subvención 20091027.

	ORIGEN (x):	Para encender la impresora:
	DESTINO (y):	To power on the printer:
ITER-0	(p) (\hat{s}_h)	() <i>To switch on:</i>
ITER-1	(p) (s_l) (k) (\hat{s}_h)	To <i>switch on:</i> power <i>on the printer:</i>
ITER-2	(p) (s_l) (k) (\hat{s}_h)	To power on the printer: () (#) ()
FINAL	(p \equiv y)	To power on the printer:

Figura 1: Ejemplo de sesión de IMT para traducir una frase de español a inglés. Las hipótesis no validadas se muestran en *cursiva*, mientras que los prefijos aceptados están impresos en fuente normal.

porciona una frase de entrada \mathbf{x} para que la traduzca. La referencia \mathbf{y} es la traducción que le gustaría alcanzar al usuario al final de la sesión de IMT. En la iteración 0, el usuario no ha proporcionado todavía ninguna corrección al sistema. Por este motivo, se muestra vacío. Por lo tanto, el sistema de IMT debe proporcionar una traducción inicial completa \mathbf{s}_h , como si fuera un sistema de SMT convencional. En la siguiente iteración, el usuario valida un prefijo \mathbf{p} como correcto posicionando el cursor en una posición concreta de \mathbf{s}_h . En este caso, después de las palabras “*To print a*”. El usuario está indicando implícitamente que el resto de la frase \mathbf{s}_l es potencialmente incorrecta. Después, el usuario introduce una nueva palabra k , la cual se asume que es diferente de la primera palabra s_{l_1} del sufijo no validado \mathbf{s}_l , $k \neq s_{l_1}$. Una vez hecho esto, el sistema sugiere una nueva hipótesis $\hat{\mathbf{s}}_h$, sujeta a $\hat{s}_{h_1} = k$. De nuevo, el usuario valida un nuevo prefijo e introduce una palabra. El proceso continúa hasta que la frase completa es correcta. Las frases corregidas se validan introduciendo la palabra especial “#”.

Más tarde, el paradigma de IMT se extendió en [21], el cual añadió la posibilidad de *rechar* un prefijo dado.

Por otra parte, internet, que proporciona acceso universal a la información y comunicación instantánea entre usuarios, ha creado una li-

bertad física y geográfica que era inconcebible en el pasado [5]. El paradigma de IMT mencionado en esta introducción ha demostrado funcionar bien en una arquitectura basada en la web. Otros sistemas web similares para el procesamiento de lenguaje natural [20] han mostrado que la retroalimentación del usuario puede mejorar la precisión del sistema, incrementar su ergonomía y la aceptación por parte del usuario de los sistemas MT. Además, dado que los usuarios operan a través de un navegador web, el sistema también proporciona una plataforma portable, independiente del sistema, y no requiere potencia computacional ni grandes cantidades de espacio en disco en las máquinas cliente.

2. Paradigma

2.1. Traducción interactiva

En esta sección describiremos brevemente el marco estadístico para el paradigma de IMT. El paradigma de IMT puede verse como una evolución del paradigma de SMT, el cual ha demostrado ser un marco eficiente para construir con poco esfuerzo sistemas de traducción, siempre que se tengan disponibles corpus adecuados [8]. La ecuación fundamental de la traducción estadística se define del siguiente

modo:

$$\hat{\mathbf{y}} = \operatorname{argmax}_{\mathbf{y}} Pr(\mathbf{y} | \mathbf{x}) \quad (1)$$

$$= \operatorname{argmax}_{\mathbf{y}} Pr(\mathbf{x} | \mathbf{y}) Pr(\mathbf{y}) \quad (2)$$

donde $Pr(\mathbf{x} | \mathbf{y})$ es el *modelo de traducción* que da cuenta de la correlación entre las frases origen y destino y $Pr(\mathbf{y})$ es el *modelo de lenguaje* que mide la corrección de \mathbf{y} como frase de la lengua destino.

En la práctica, se adopta comúnmente el modelado de la distribución de probabilidad a posteriori $Pr(\mathbf{y} | \mathbf{x})$. Para ello, distintos autores [19, 14] proponen el uso de los llamados modelos log-lineales, en los que la regla de decisión viene dada por la expresión:

$$\hat{\mathbf{y}} = \operatorname{argmax}_{\mathbf{y}} \sum_{m=1}^M \lambda_m h_m(\mathbf{x}, \mathbf{y}) \quad (3)$$

donde $h_m(\mathbf{x}, \mathbf{y})$ es una función de puntuación relevante para traducir de \mathbf{x} a \mathbf{y} , M es el número de modelos (o características) y λ_m son los pesos de la combinación log-lineal.

Las instancias más habituales de los modelos log-lineales incluyen modelos de segmentos [23, 11] (del inglés, modelos *phrase-based* o modelos PB). Los modelos PB permiten capturar información contextual al trabajar con segmentos de frases en lugar de con palabras. La idea básica de estos modelos consiste en dividir la frase origen en segmentos (conjuntos de palabras consecutivas de la frase origen o destino), después elegir la traducción de cada segmento y por último reordenar dichas traducciones para componer la frase destino. Los modelos PB se han empleado a lo largo de todo este trabajo.

En los modelos log-lineales, el problema de maximización formulado en la Ec. 3 se resuelve por medio de un algoritmo de búsqueda en haz como el introducido en [13] para su uso en el campo de reconocimiento de voz. El algoritmo de búsqueda en haz genera soluciones parciales, denominadas *hipótesis*, hasta que se obtiene una traducción completa; estas hipótesis son almacenadas en una pila y ordenadas por su puntuación. Dicha puntuación viene da-

da por una combinación log-lineal de características.

Sin embargo, la Ec. 1 debe ser modificada de acuerdo con el paradigma de IMT con objeto de tener en cuenta la parte de la frase destino que ya ha sido traducida, en concreto \mathbf{p} y k :

$$\hat{\mathbf{s}}_h = \operatorname{argmax}_{\mathbf{s}_h} Pr(\mathbf{s}_h | \mathbf{x}, \mathbf{p}, k) \quad (4)$$

Donde el problema de maximización se define sobre el sufijo \mathbf{s}_h . Esto permite reescribir la Ec. 4 descomponiendo adecuadamente el lado derecho y eliminando términos constantes, alcanzando el criterio equivalente:

$$\hat{\mathbf{s}}_h = \operatorname{argmax}_{\mathbf{s}_h} Pr(\mathbf{p}, k, \mathbf{s}_h | \mathbf{x}). \quad (5)$$

Un ejemplo del sentido de estas variables puede apreciarse en la Figura 1.

Dado que $(\mathbf{p} k \mathbf{s}_h) = \mathbf{y}$, la Ec. 5 es muy similar a la Ec. 1. La principal diferencia reside en que la maximización se lleva a cabo sobre el conjunto de sufijos \mathbf{s}_h que completan $(\mathbf{p} k)$ en vez de sobre frases completas (\mathbf{y} en la Ec. 1). Esto implica que podemos usar los mismos modelos si los algoritmos de búsqueda son adecuadamente modificados [3].

2.2. Traducción interactiva basada en modelos PB

Los modelos PB pueden adaptarse fácilmente para su uso en un escenario de IMT. La modificación más importante consiste en introducir el uso de grafos de palabras, o *word graphs* en inglés. El uso de word graphs en IMT ha sido estudiado en [3], en combinación con dos técnicas diferentes de traducción, a saber, la técnica de *Alignment Templates* [15, 16], y la técnica de transductores de estados finitos [4].

2.3. Generación de word graphs

Un word graph es un grafo acíclico ponderado, en el que cada nodo representa una traducción parcial y cada arco está etiquetado con una palabra de la frase destino y ponderado de acuerdo con las puntuaciones dada por el modelo de traducción (para más detalles véase [22]). En [17], se propone el uso de word

graphs como interfaz entre un sistema de traducción basado en alignment templates y el motor de traducción interactiva. Análogamente, en este trabajo se usan word graphs contruídos durante el proceso de búsqueda de un sistema de traducción basado en modelos PB.

Durante el proceso de búsqueda que lleva a cabo el motor de traducción, es posible crear un grafo de segmentos o *segment graph* en inglés. En dicho segment graph, cada nodo representa un estado del modelo de traducción, y cada arco una transición ponderada con un segmento de frase asociado. Cada vez que se extiende una hipótesis, se añade un nuevo arco conectando el estado de dicha hipótesis con el estado de la hipótesis extendida. El nuevo arco se etiqueta con el segmento de palabras destino que ha sido incorporado a la hipótesis extendida, y se le asocia un peso determinado con la puntuación del modelo de traducción.

Una vez que se ha generado el segment graph, puede convertirse fácilmente en un word graph mediante la introducción de estados artificiales para las palabras de los segmentos asociados a los arcos.

2.4. Traducción interactiva basada en word graphs

Durante el proceso de traducción interactiva para una frase origen, el sistema hace uso del word graph generado para dicha frase con objeto de completar los prefijos aceptados por el usuario. Específicamente, el sistema encuentra el mejor camino en el word graph asociado a un prefijo dado, de forma que pueda completarse la frase destino. Esto permite proveer al usuario de numerosas sugerencias para completar cada prefijo.

Un problema común en traducción interactiva aparece cuando el usuario establece un prefijo que no se encuentra en el word graph, ya que en este caso el sistema no puede encontrar un camino en el word graph para devolver un sufijo apropiado. El procedimiento habitual para resolver este problema consiste en llevar a cabo una búsqueda *tolerante* en el word graph. Dicha búsqueda tolerante utiliza el concepto bien conocido de distancia de Levenshtein con objeto de obtener la cadena más

similar al prefijo dado por el usuario (para más detalles véase [17]).

2.5. Acciones de ratón

En el paradigma de IMT tradicional, la única retroalimentación que recibe el sistema se produce cuando el usuario introduce una palabra nueva. Sin embargo, en este artículo mostramos cómo enriquecer la interacción hombre-máquina mediante la introducción de Acciones de Ratón (AR) como fuente adicional de información para el sistema. Por ello, consideramos dos tipos de AR: AR *no explícitas de posicionamiento* y AR *explícitas de interacción*.

AR no explícitas de posicionamiento

Al corregir una hipótesis y antes de introducir una palabra nueva, el usuario necesita posicionar el cursor en el lugar donde quiere introducir dicha palabra. En este trabajo, vamos a asumir que dicha acción se lleva a cabo mediante una AR, a pesar de que la misma idea podría ser aplicada en el caso de que dicha acción fuera llevada a cabo de otro modo. Es importante recalcar que, al hacer esto, el usuario ya está proporcionando una información valiosa al sistema: está validando el prefijo hasta la posición donde ha situado el cursor y, además, está indicando que cualquier palabra que se sitúe a continuación debe ser considerada como incorrecta. Por ello, el sistema ya puede capturar este hecho y proporcionar una nueva hipótesis, cuyo sufijo será aquel de probabilidad máxima teniendo en cuenta la restricción de que la primera palabra no puede ser la primera palabra del sufijo invalidado. Somos conscientes de que esto no implica de manera directa que el nuevo sufijo sea correcto, ni tan siquiera que lo sea su primera palabra. Sin embargo, lo peor que puede pasar es que el usuario tenga que introducir la palabra de todos modos, como iba a hacer en un principio. Por otro lado, si el sufijo es correcto, el usuario ya no necesitará corregir esa palabra.

Se puede ver un ejemplo de dicho comportamiento en la Figura 2. En este ejemplo, el sistema SMT proporciona una traducción que no satisface al usuario. Por ello, posiciona el

ORIGEN (x):		Para encender la impresora:
DESTINO (y):		To power on the printer:
ITER-0	(p) (ŝ _h)	() To switch on:
ITER-1	(p) (s _l) (ŝ _h)	To switch on: power on the printer:
ITER-2	(p) (s _l) (k) (ŝ _h)	To power on the printer: () (#) ()
FINAL	(p ≡ y)	To power on the printer:

Figura 2: Ejemplo de acción de ratón no explícita de posicionamiento que resuelve un error cambiando una palabra. Al validar el prefijo \mathbf{p} , el usuario está indicando implícitamente que quiere que el sufijo sea modificado, con lo cual el sistema produce inmediatamente un nuevo sufijo \mathbf{s}_h , que en este caso es correcto. En **ITER-1**, el carácter | indica la posición donde se llevó a cabo la AR, \mathbf{s}_l es el sufijo rechazado por esa AR, y $\hat{\mathbf{s}}_h$ es el sufijo que el sistema sugiere después de observar que \mathbf{s}_l debe ser considerado incorrecto. El carácter # es un carácter especial que el usuario introduce para indicar que la hipótesis debe ser aceptada.

cursor antes de la palabra “switch”, con el propósito de introducir la palabra “power”. Al hacer esto, está validando el prefijo “To”, y señalando implícitamente que quiere cambiar la palabra “switch”. Antes de que teclee nada, el sistema se da cuenta de que va a cambiar la palabra situada tras el cursor, y reemplaza el sufijo por otro. Puesto que el nuevo sufijo propuesto es el que el usuario tenía en mente, el usuario únicamente necesita aceptar la hipótesis.

Llamamos a este tipo de hipótesis AR *no explícita de posicionamiento* porque no necesita ninguna acción adicional por parte del usuario: ya ha llevado a cabo una AR para posicionar el cursor en el sitio deseado, y el sistema se aprovecha de esta acción para sugerir una nueva hipótesis.

Es importante destacar que cualquier mejora en el rendimiento obtenida gracias a la introducción de AR no explícitas no requiere ningún esfuerzo adicional por parte del usuario, puesto que el usuario necesita posicionar el cursor antes de introducir una nueva palabra, y por lo tanto no tienen ningún coste.

Por ello, ahora estamos considerando dos situaciones diferentes: la primera, la del paradigma IMT tradicional, en la cual el sistema necesita encontrar un sufijo apropiado según

la ecuación 5, y una nueva, en la cual el sistema necesita encontrar un sufijo cuya primera palabra no debe ser una palabra k dada, sino que debe ser *diferente* a la palabra s_{l_1} . Esta restricción se puede expresar mediante la ecuación siguiente:

$$\hat{\mathbf{s}}_h = \operatorname{argmax}_{\mathbf{s}_h: s_{h_1} \neq s_{l_1}} Pr(\mathbf{p}, \mathbf{s}_h | \mathbf{x}, \mathbf{s}_l) \quad (6)$$

donde \mathbf{s}_l es el sufijo generado en la iteración anterior, y que ya ha sido descartado por el usuario, y s_{l_1} es la primera palabra de \mathbf{s}_l . En esta fórmula, omitimos k porque el usuario no ha introducido ninguna palabra.

AR explícitas de interacción

En caso de que el sistema sea eficiente y proporcione sugerencias que son de una calidad aceptable, es fácil imaginar una situación en la cual el usuario requiera al sistema que modifique un sufijo determinado, sin llegar a introducir ninguna palabra. En este caso, modelamos este hecho como una clase de AR diferente, *explícita de interacción*, puesto que el usuario necesita indicar *explícitamente* que quiere que se modifique el sufijo. Esto contrasta con las AR de la sección anterior, en las cuales el usuario no realizaba ninguna acción de ma-

nera explícita. Sin embargo, si el sistema de traducción automática subyacente proporciona sufijos lo suficientemente buenos, el usuario se dará cuenta rápidamente de que realizar una AR es menos costoso que introducir una palabra entera, y por lo tanto se aprovechará de este hecho pulsando con el ratón sistemáticamente antes de introducir una palabra nueva. Por ello, en este caso asumiremos también que el usuario pulsa el ratón siempre antes de corregir una palabra incorrecta, requiriendo al sistema que modifique el sufijo. Al hacer esta asunción, estamos también asumiendo un perfil de usuario diferente, en el cual el usuario adquiere una postura más participativa e interactiva, al pulsar el ratón en lugares donde no es estrictamente necesario.

En este caso, por tanto, sí que hay un coste asociado a esta AR, puesto que el usuario sí que debe llevar a cabo acciones adicionales, que pueden o no ser beneficiosas. Es posible, por tanto, que el usuario tenga que introducir la palabra que tenía en mente incluso después de pedir repetidamente que se cambie el sufijo, con lo cual las AR adicionales que hubiera llevado a cabo habrían resultado inútiles.

3. Prototipo

Con objeto de ilustrar el marco teórico de IMT, hemos creado un prototipo que implementa este concepto de traducción automática interactiva¹. A continuación mostramos sus principales características, el diseño y la arquitectura del sistema, así como el protocolo de interacción usuario-máquina.

3.1. Diseño

En primer lugar, cabe comentar que este prototipo no está pensado principalmente para ser un producto comercial. En su lugar, provee una interfaz intuitiva que ayuda a mostrar las funcionalidades de un sistema IMT. Así, su arquitectura ha sido concebida pensando en dos aspectos: accesibilidad y flexibilidad. El primero es necesario para llegar a un gran número de usuarios potenciales. El segundo permite a

los investigadores probar diferentes técnicas y protocolos de interacción, reduciendo el esfuerzo en la implementación.

Por esta razón, se desarrolló una interfaz de programación de aplicaciones (API) para herramientas CAT [1]. Este API permite una separación nítida entre la interfaz del cliente y el sistema de traducción, usando un protocolo de comunicación en red y proporcionando un conjunto bien definido de funciones. Además permite la creación, implementación y personalización de herramientas profesionales para usar sistemas de IMT con un mínimo esfuerzo.

El diagrama de la arquitectura del sistema se muestra en la Figura 3. Por un lado, el cliente IMT proporciona una Interfaz de Usuario (UI) la cual utiliza el API de comunicación para conectarse con servidores IMT a través de la Web. Los requerimientos de hardware en la parte del cliente son mínimos, ya que todo el proceso de traducción se lleva a cabo en el servidor. Así, virtualmente cualquier ordenador (incluyendo netbooks, tablets, o incluso teléfonos móviles) puede hacer uso de este sistema. Por otro lado, el servidor, que no se da cuenta de los detalles de la implementación del cliente, emplea modelos estadísticos para realizar las traducciones.

3.2. Protocolo de Interacción del Usuario

El protocolo que rige el proceso de interacción con un sistema IMT tiene los siguientes pasos:

1. El sistema propone una traducción completa del segmento de texto seleccionado.
2. El usuario valida el prefijo más largo de la traducción que está libre de errores y/o corrige el primer error que encuentra en el sufijo. Las correcciones se introducen mediante teclado o clicks del ratón.
3. De esta manera, se extiende y consolida un nuevo prefijo, basado en el prefijo anterior (validado) y el feedback de la interacción del usuario. Usando este nuevo prefijo, el sistema sugiere una adecuada continuación del mismo.

¹ disponible en <http://cat.it.i.upv.es/imt>

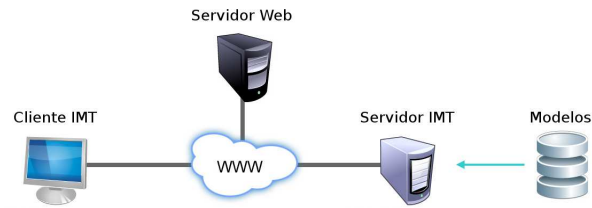


Figura 3: Arquitectura del sistema IMT.

- Los pasos 2 y 3 se repiten hasta que se alcanza la traducción que buscaba el usuario.

3.3. Funcionalidades

La siguiente es una lista de las características que soporta el prototipo:

- Cuando el usuario corrige la solución propuesta por el sistema, se propone un nuevo (y mejorado) sufijo.
- El usuario puede realizar *acciones* mediante atajos de teclado o gestos de ratón. Las posibles acciones a realizar sobre el sufijo propuesto son:

Sustitución Sustituye la primera palabra o carácter erróneos del sufijo.

Borrado Elimina la primera palabra o carácter erróneos del sufijo.

Inserción Inserta una palabra antes del sufijo.

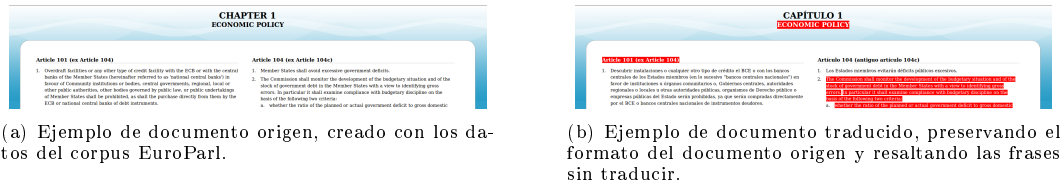
Rechazo La palabra rechazada se marca como incorrecta, y no volverá a aparecer en las siguientes hipótesis.

- En cualquier momento, el usuario es capaz de visualizar el documento original (Figura 4(a)), así como un borrador de la traducción actual formateada adecuadamente (Figura 4(b)).
- Se presenta una lista de los documentos disponibles para traducir al usuario (Figura 5) de forma que pueda probar el prototipo bajo diferentes condiciones, ej. corpus de documentos y pares de lenguaje origen-destino.

3.4. Implementación

Este prototipo aprovecha Internet para permitir la conexión simultánea y global de usuarios desde cualquier parte del mundo, coordinando tecnologías cliente-servidor — ej. la interfaz del cliente se ha construido mediante estándares web tales como HTML y JavaScript, mientras que el motor de traducción fue escrito en C++ usando el toolkit Thot [18].

Para empezar, la Interfaz Web (WUI, del inglés Web User Interface) carga un índice de todos los corpus de traducción disponibles (Figura 5). Además, los usuarios pueden subir sus propios documentos y contribuir así a engrosar el conjunto de corpus previamente mencionado. El usuario después escoge un corpus y navega a la página con la aplicación de traducción (Figura 6), donde empieza a traducir uno a uno los segmentos de texto. La extracción de dichos segmentos es posible gracias a que los documentos se convierten a un formato intermedio (TMX, un estándar abierto para el intercambio de memorias de traducción). Así, el formato del documento original puede ser preservado fácilmente y aplicado una vez traducido (parte de o todo) el mismo. Durante el proceso de traducción, el feedback del usuario es procesado por el servidor IMT. De este modo, tanto el usuario como el sistema se guían y ayudan mutuamente, optimizando rendimiento y usabilidad. Todas las correcciones se guardan en logs de texto plano en el servidor, de forma que el usuario puede retomar los documentos a traducir desde el punto en el que los dejó, permitiendo alternativamente que otros usuarios puedan ayudar a traducir todo el corpus de documentos.



(a) Ejemplo de documento origen, creado con los datos del corpus EuroParl.

(b) Ejemplo de documento traducido, preservando el formato del documento origen y resaltando las frases sin traducir.

Figura 4: Traduciendo documentos con el sistema IMT propuesto.



Figura 5: Página principal, mostrando los corpus disponibles.

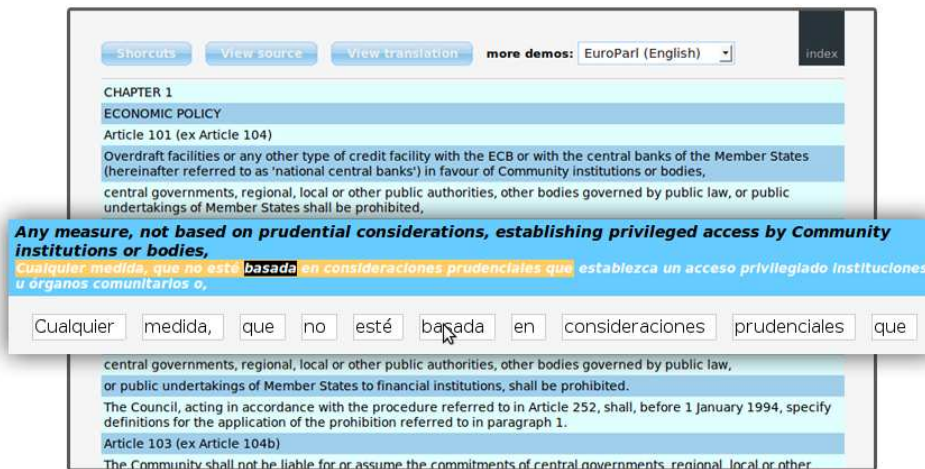


Figura 6: Interfaz del prototipo. El texto del documento origen es extraído en segmentos automáticamente.

Puesto que los usuarios trabajan dentro de un navegador Web, el sistema también proporciona compatibilidad multiplataforma. Ade-

más, la máquina del cliente no requiere ni potencia de cálculo ni espacio en disco. La comunicación cliente-servidor se basa en conexiones

HTTP asíncronas, proporcionando así una experiencia interactiva más rica — no actualiza la página totalmente, sólo en caso de cambiar a otros corpus de traducción. Por otra parte, el servidor Web se comunica con el motor IMT a través de sockets TCP binarios. Por lo tanto, los tiempos de latencia son muy reducidos — un requisito necesario para el consuelo del usuario. Además, es posible la comunicación entre diferentes dominios Web. De esta manera, es posible cambiar entre diferentes motores de IMT desde la misma WUI.

3.5. Evaluación

Antes de implementar el prototipo se realizaron pruebas experimentales para evaluar la factibilidad del marco de trabajo. Dichos experimentos simulan la interacción del usuario con el objetivo de medir el esfuerzo requerido para generar traducciones libre de errores. En [?, 3] se pueden ver los detalles de las pruebas y los resultados en corpus reales. Dichos resultados sugieren que nuestro enfoque interactivo puede reducir sustancialmente el esfuerzo de escritura respecto a tener que post-editar toda la salida que devuelve automática un sistema de traducción. Específicamente, la mejor reducción de esfuerzo se obtuvo con la tarea Español-Inglés, correspondiente a un ahorro del 80%.

4. Conclusiones y trabajo futuro

En este artículo hemos presentado una posible implementación de sistema de IMT, tanto desde el punto de vista teórico como desde el punto de vista de su implementación práctica mediante una interfaz de interacción via web. Para ello, hemos tenido en cuenta tanto el planteamiento clásico de IMT, en el cual la interacción con el usuario sucede únicamente cuando el usuario introduce una palabra. Como añadidura, hemos introducido también un nuevo protocolo de interacción con el usuario, mediante la consideración de las acciones de ratón del usuario como retroalimentación importante para el sistema.

Además, hemos presentado un prototipo de aplicación de este marco teórico, en el cual se

puede observar el funcionamiento del sistema implementado. Este prototipo, basado en una interfaz web, proporciona las sugerencias de traducción al usuario mediante un protocolo de comunicación en red. Es importante remarcar que las sugerencias presentadas funcionan de manera predictiva, esto es, al escribir parcialmente una palabra se propone completarla junto con el mejor sufijo según el feedback del usuario.

Actualmente estamos implementando multimodalidad en el prototipo, para poder usar otros dispositivos de entrada tales como lápiz electrónico o pantallas táctiles. Esta aproximación ya ha sido probada con éxito en trabajos anteriores [20, ?]. Como trabajo futuro cabe citar la evaluación del sistema con usuarios reales, para comparar con los resultados preliminares citados en la sección 3.5.

Referencias

- [1] V. Alabau, D. Ortiz, V. Romero, and J. Ocampo. A multimodal predictive-interactive application for computer assisted transcription and translation. In *ICMI-MLMI '09: Proceedings of the 2009 international conference on Multimodal interfaces*, pages 227–228, New York, NY, USA, 2009. ACM.
- [2] D.J. Arnold. *Computers and Translation: A translator's guide*, chapter 8, pages 119–142. 2003.
- [3] S. Barrachina, O. Bender, F. Casacuberta, J. Civera, E. Cubel, S. Khadivi, A.L. Lagarda, H. Ney, J. Tomás, E. Vidal, and J.M. Vilar. Statistical approaches to computer-assisted translation. *Computational Linguistics*, 35(1):3–28, 2009.
- [4] F. Casacuberta and E. Vidal. Learning finite-state models for machine translation. *Machine Learning*, 66(1):69–91, 2007.
- [5] O. Craciunescu, C. Gerding-Salas, and S. Stringer-O'Keefe. Machine translation and computer-assisted translation: a new

- way of translating? *Translation Journal*, 8(3), 2004.
- [6] G. Foster. *Text Prediction for Translators*. PhD thesis, Université de Montréal, 2002.
- [7] G. Foster, P. Langlais, and G. Lapalme. User-friendly text prediction for translators. In *Proc. of EMNLP'02*, pages 148–155, 2002.
- [8] J. Hutchings and H. Somers. An introduction to machine translation. In *Ed. Academic Press.*, 1992.
- [9] J. Hutchins. Retrospect and prospect in computer-based translation. In *Proc. of MT Summit VII*, pages 30–44, 1999.
- [10] M. Kay. It's still the proper place. *Machine Translation*, 12(1-2):35–38, 1997.
- [11] P. Koehn, F.J. Och, and D. Marcu. Statistical phrase-based translation. In *Proc. HLT/NAACL'03*, pages 48–54, 2003.
- [12] P. Langlais, G. Lapalme, and M. Loranger. Transtype: Development-evaluation cycles to boost translator's productivity. *Machine Translation*, 15(4):77–98, 2002.
- [13] B.T. Lowerre. *The harpy speech recognition system*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 1976.
- [14] F. Och and H. Ney. Discriminative training and maximum entropy models for statistical machine translation. In *Proc. of the ACL'02*, pages 295–302, 2002.
- [15] F. Och, C. Tillmann, and H. Ney. Improved alignment models for statistical machine translation. In *Proc. of EMNLP/WVLC'99*, pages 20–28, 1999.
- [16] F.J. Och and H. Ney. The alignment template approach to statistical machine translation. *Comput. Linguist.*, 30(4):417–449, 2004.
- [17] F.J. Och, R. Zens, and H. Ney. Efficient search for interactive statistical machine translation. In *Proc. of EACL'03*, pages 387–393, 2003.
- [18] D. Ortiz, I. García-Varea, and F. Casacuberta. Thot: a toolkit to train phrase-based statistical translation models. In *Proc. of the Machine Translation Summit X*, pages 141–148, Phuket, Thailand, September 2005.
- [19] K. Papineni, S. Roukos, and T. Ward. Maximum likelihood and discriminative training of direct translation models. In *Proc. of ICASSP'98*, pages 189–192, 1998.
- [20] V. Romero, L.A. Leiva, A.H. Toselli, and E. Vidal. Interactive multimodal transcription of text images using a web-based demo system. In *Proceedings of IUI*, 2009.
- [21] G. Sanchis-Trilles, D. Ortiz-Martínez, J. Civera, F. Casacuberta, E. Vidal, and H. Hoang. Improving interactive machine translation via mouse actions. In *EMNLP 2008: conference on Empirical Methods in Natural Language Processing*, 2008.
- [22] N. Ueffing, F. Och, and H. Ney. Generation of word graphs in statistical machine translation. In *Proc. of EMNLP'02*, pages 156–163, 2002.
- [23] R. Zens, F.J. Och, and H. Ney. Phrase-based statistical machine translation. In *Proc. of KI'02*, pages 18–32, 2002.