

## Transtype 2. Un sistema de ayuda a la traducción

Antonio L. Lagarda, Luis Rodríguez, Elsa Cubel, Enrique Vidal  
and Francisco Casacuberta

Instituto Tecnológico de Informática  
Universiad Politécnica de Valencia  
E-46071 Valencia, Spain  
{alagarda,lrodrig,ecubel,evidal,fcn}@iti.uv.es

**Resumen:** TransType2 es un proyecto enmarcado en el ámbito de la traducción asistida. Su objetivo consiste en proporcionar traducciones de alta calidad combinando la experiencia de traductores humanos con sistemas de traducción automática.

**Palabras clave:** Traducción asistida por computador, traducción automática.

**Abstract:** TransType2 is a project framed in assisted translation field. Its purpose is to provide high-quality translations by combining human translators' knowledge with machine translation systems.

**Keywords:** Computer-Assisted Translation, Machine Translation

### 1. Introducción

TransType2 (TT2) (SchlumbergerSema S.A. et al., 2001) es un proyecto financiado por la Comisión Europea dentro del *IST Programme (IST-2001-32091)*. El objetivo de TT2 es el desarrollo de un sistema de traducción asistida por computador (CAT, *Computer-Assisted Traslation*) que contribuya a solventar la creciente demanda de traducciones de alta calidad existente actualmente.

Para conseguir este objetivo, y con el fin de asegurar un mejor resultado en las traducciones obtenidas, TT2 combina los siguientes paradigmas. Por un lado, se utiliza un sistema de traducción automática (MT, *Machine Translation*) que permite reducir el tiempo necesario para realizar el proceso de traducción. Por otra parte, se incorpora el conocimiento de un traductor humano con el fin de asegurar una traducción de calidad.

### 2. Sistema de traducción

En este proyecto, se ha optado por incluir un motor de traducción basado en modelos estadísticos. Dicho motor de traducción está basado en Transductores de estados finitos estocásticos (SFSTs)(Casacuberta, 2000), los cuales ya habían demostrado anteriormente su buen comportamiento y eficiencia en tareas pertenecientes a dominios restringidos (Vidal, 1997).

El proceso seguido con estos modelos consiste, en primer lugar, en la inferencia de un SFST partiendo de ejemplos compuestos de posibles frases a traducir y sus correspondientes traducciones. Para realizar este proceso de inferencia, se utiliza la técnica denominada GIATI (Casacuberta, 2000). Una vez obtenido el SFST, es posible llevar a cabo la traducción de una frase concreta realizando un proceso de búsqueda (basado en el algoritmo de Viterbi (Viterbi, 1967)) sobre el SFST.

Las traducciones obtenidas mediante el sistema de traducción no garantizan un nivel óptimo de calidad. En este punto, interviene el traductor humano con el objetivo de evaluar y corregir, si es necesario, dichas traducciones. El proceso consiste en que una vez que el sistema de traducción proporciona una traducción, el traductor humano verifica la corrección de ésta pudiendo aceptarla en su totalidad o sugerir algunos cambios. Concretamente, el traductor puede aceptar parte de la traducción y/o añadirle nuevos elementos (palabras, letras, ...) generando un prefijo de lo que considera debe ser la traducción correcta. Una vez realizado este proceso, el sistema de traducción procede a completar este prefijo, de manera que la traducción resultante está compuesta por el prefijo proporcionado por el traductor seguido de la sugerencia realizada por el sistema de traduc-

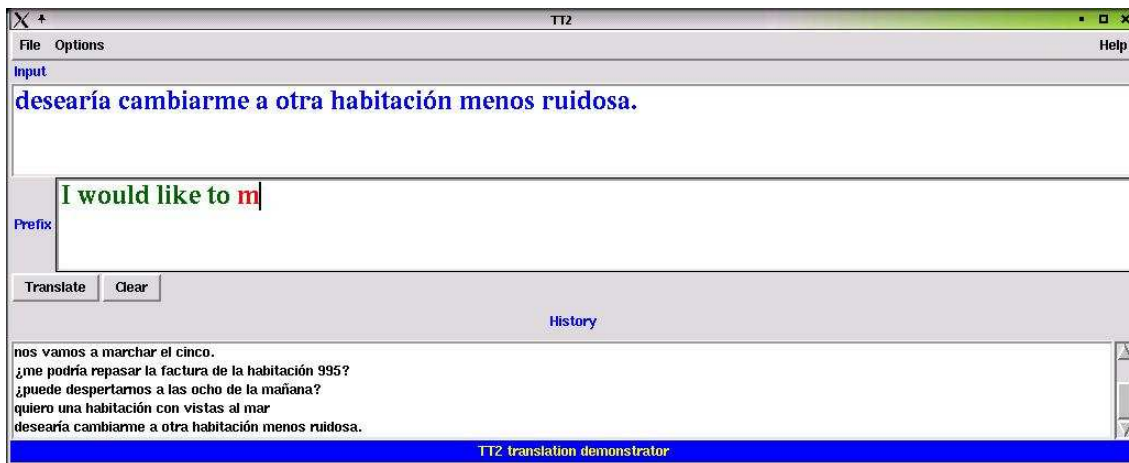


Figura 1: Captura de pantalla del prototipo

ción, obteniéndose así una traducción completa.

### 3. Descripción del sistema

Todo lo comentado en el apartado anterior se plasma en un prototipo (figura 1) que, sin constituir la versión definitiva del proyecto, sí puede considerarse como una primera aproximación a lo que constituirá el resultado definitivo. Este prototipo se compone del motor de traducción asistida empujado en un interfaz TCL/TK desarrollado todo ello en el Instituto Tecnológico de Informática de Valencia.

El funcionamiento de este prototipo es muy simple. En primer lugar, se debe seleccionar el sistema de traducción a utilizar. Seguidamente, habrá que seleccionar un fichero que contenga el conjunto de frases que se pretende traducir. En este momento el sistema está listo para ser utilizado. El proceso básico a seguir consiste en seleccionar una de las frases de entrada disponibles y solicitar una traducción pulsando en el botón *Translate*. Una vez que el sistema nos ha proporcionado una traducción, es posible refinarla de la siguiente manera. En primer lugar debemos situar el cursor (dentro del panel *Prefix*) al final del fragmento de traducción que consideremos correcto. Seguidamente podemos completar dicho fragmento con un conjunto de caracteres de manera que la secuencia de palabras situada a la izquierda del cursor de forma que el sistema complete la traducción desde ese punto sin modificar ese prefijo.

Este prototipo permite al usuario interactuar con el sistema con métodos “clásicos”, es decir, utilizando periféricos co-

mo el teclado y el ratón. Además, incorpora la posibilidad de que el usuario se comunique con el sistema a través de la voz.

### Bibliografía

- Casacuberta, Francisco. 2000. Inference of finite-state transducers by using regular grammars and morphisms. En A.L. Oliveira, editor, *Grammatical Inference: Algorithms and Applications*, volumen 1891 de *Lecture Notes in Computer Science*. Springer-Verlag, páginas 1–14. 5th International Colloquium Grammatical Inference -ICGI2000-. Lisboa. Portugal.
- SchlumbergerSema S.A., Instituto Tecnológico de Informática, Rheinisch Westfälische Technische Hochschule Aachen Lehrstuhl für Informatik VI, Recherche Appliquée en Linguistique Informatique Laboratory University of Montreal, Celer Soluciones, Société Gamma, y Xerox Research Centre Europe. 2001. TT2. TransType2 - computer assisted translation. Project technical annex.
- Vidal, Enrique. 1997. Finite-state speech-to-speech translation. En *Int. Conf. on Acoustics Speech and Signal Processing (ICASSP-97), proc., Vol.1*, páginas 111–114, Munich.
- Viterbi, Andrew. 1967. Error bounds for convolutional codes and a asymptotically optimal decoding algorithm. *IEEE Transactions on Information Theory*, 13:260–269.