

## Practico 1 - Traducción automática y evaluación

En la clase del viernes 19-03-2010 dictada por Laura, vieron algunos problemas que habitualmente encontramos al tratar o procesar un lenguaje natural y la idea de este práctico es ver cómo algunos sistemas de TA los resuelven (o no). Las principales divergencias entre lenguajes se originan en la ambigüedad a distintos niveles:

- A nivel **léxico**: una misma palabra puede tener varios significados; por ejemplo *banco* (entidad financiera, asiento en una plaza, verbo “banca”?).
- A nivel **estructural**: puede originar distintos significados de una frase de acuerdo a la estructura de constituyentes que tengan o a cómo se la analiza; por ejemplo *vi a un hombre con un telescopio* (tengo un telescopio o lo tiene el hombre?)
- Construcciones **idiomáticas**: son expresiones que no tienen una traducción directa, siendo difícil expresar en el idioma destino (TL)<sup>1</sup> el significado de una sola palabra en el idioma origen (SL); por ejemplo “tirar la toalla” o “estirar la pata”.
- A nivel **referencial**, la resolución de anáforas y catáforas implica determinar la entidad lingüística previa o posterior a que hacen referencia; ejemplo “se *lo* advertí” (que cosa le advertí?).
- A nivel **pragmático**: una oración no siempre significa lo que se está diciendo, un caso muy particular es la **ironía**, que puede cambiar por completo el mensaje.

Para cada uno de estos tipos de problemas (o al menos para un par!):

1. Elijan un par de idiomas SL/TL distintos para trabajar y encuentren un ejemplo que ilustre el problema; ej. si eligen Español/Inglés la palabra “pata” podría dar lugar una ambigüedad léxica
2. Para el ejemplo elegido, armar una frase (SL) y dar su traducción correcta en TL
3. Traducir la misma frase de ejemplo con algunos de estos traductores automáticos:  
LexiCool ofrece traducciones con Google – Systran – Bing – Reverso – SDL  
(<http://www.lexicool.com/translate.asp?IL=3>),  
IM Translator <http://translation2.paralink.com/>,  
ProMT (<http://www.online-translator.com/>),  
WordLingo ([http://www.worldlingo.com/en/products\\_services/worldlingo\\_translator.html](http://www.worldlingo.com/en/products_services/worldlingo_translator.html))
4. Analizar las traducciones automáticas y decidir si son correctas. En caso de que no lo sean, por qué les parece que falló? Es una consecuencia directa del tipo de divergencias que indicaron?
5. De ser posible encuentren ejemplos en textos reales, por ejemplo leyendo noticias o blogs.

Para esta actividad se necesita conocimiento de otro idioma distinto del Español, como Inglés, Ruso, Griego, etc pero en caso de no poseerlo, o también por simple curiosidad, pueden hacer el práctico usando los textos A y B (links más abajo) de la siguiente manera: con algunos de los traductores mencionados arriba, traducir A del inglés al español y comparar el resultado con B. Para cada error encontrado en la traducción automática, tratar de identificar el problema que lo originó (si podría ser una ambigüedad no resuelta, falta de vocabulario, etc)

A) Artículo “On the cruelty of really teaching computing science” de E. Dijkstra

[http://www.smaldone.com.ar/documentos/ewd/sobre\\_la\\_crueldad.html](http://www.smaldone.com.ar/documentos/ewd/sobre_la_crueldad.html)

B) Traducción del mismo artículo al español

<http://userweb.cs.utexas.edu/users/EWD/transcriptions/EWD10xx/EWD1036.html>

La clase que viene (viernes 26-03-2010, aula 14) podemos ver algunos de los casos estudiados y discutir los resultados. Dudas/consultas/comentarios se pueden mandar a [pestrella@gmail.com](mailto:pestrella@gmail.com)

---

<sup>1</sup> TL = *Target Language* y SL = *Source Language* y cuando traducimos pasamos de un SL a un TL.