

## Gestionando Babel: un aula virtual multilingüe en tiempo real

---

Cristina Borrell, investigadora (cborrellc@uoc.edu)  
Antoni Oliver, profesor (aoliverg@uoc.edu)

Universitat Oberta de Catalunya (UOC)  
Internet Interdisciplinary Institute IN3  
Avda. Canal Olímpic S/N  
08860 Castelldefels  
0034646593855

## Resumen

El multilingüismo es una realidad del siglo XXI y el desarrollo de las nuevas tecnologías en las últimas décadas se revela como un método eficaz para hacerle frente. Actualmente, las comunidades son cada vez más multiculturales y la sociedad necesita abastecerse de herramientas capaces de gestionar el multilingüismo derivado de esta condición. La Universitat Oberta de Catalunya, en el marco de la Cátedra de Multilingüismo Linguamón-UOC, ha integrado un entorno virtual de aprendizaje con un sistema de traducción automática. De este modo se consigue un espacio de aprendizaje virtual capaz de acoger alumnos y profesores de diferentes comunidades de habla, en el cual el usuario podrá expresarse y obtener la información en la lengua que desee.

Las aplicaciones con las cuales se ha llevado a cabo el proyecto, Moodle y Apertium, son de código abierto. De este modo se consigue llegar al máximo de usuarios posible sin ningún tipo de coste ni restricción, y con la libertad de saber que el conjunto se puede modificar en función de las propias necesidades. Este entorno integrado incluye un detector de lengua y un módulo de preedición del texto de entrada, para conseguir una traducción de mejor calidad sin ningún tipo de intervención por parte del usuario. Además, su disposición totalmente modular permite adaptar fácilmente la plataforma a las características de los usuarios. Los autores lo han programado en función de la tipología textual observada en un estudio lingüístico previo, tomando como referencia los corpus académicos de la UOC.

El traductor integrado funciona para los mensajes en catalán, español, inglés y francés, aunque se pueden incorporar otros pares de lenguas. En cualquier momento, el usuario tiene la opción de ver los mensajes en la lengua que elija. Asimismo, siempre será consciente de si está leyendo un mensaje en lengua original o una traducción, porque queda explícito en la interfaz. Como el proceso no requiere ninguna intervención por parte del usuario, es posible que alguna de las traducciones presente errores. En caso de que la traducción no fuera del todo adecuada, y gracias a un sistema de asignación de roles, existe la posibilidad de modificar y corregir los mensajes salientes.

Todo el sistema se ha sometido a una encuesta repartida entre los estudiantes, con el objetivo de conocer el grado de satisfacción de los usuarios. En general, el aula multilingüe se ha visto con buenos ojos, se considera una riqueza y se ha calificado la experiencia de positiva, sobre todo por los hablantes del catalán que a menudo deben renunciar al uso de su lengua a favor del español.

## Introducción

La sociedad actual, cada vez más globalizada, interrelacionada y en continuo movimiento, ha hecho evolucionar las necesidades del ser humano. La movilidad es una de las facetas más visibles de la mundialización, con repercusiones en todos los ámbitos. En las aulas, por ejemplo, ahora es perfectamente normal agrupar alumnos de cualquier nacionalidad. La enseñanza en línea no ha sido una excepción. En este sentido, Internet se ha convertido en un instrumento esencial para el desarrollo de las personas, y la sociedad del conocimiento se ha adaptado a esta nueva tendencia. Gracias a las TIC, pues, el acceso a la información y a la formación es posible para todo el mundo y en cualquier contexto. La Universitat Oberta de Catalunya quiere contribuir dentro de lo posible al desarrollo y progreso de todas las personas a quienes pueda llegar, facilitando al máximo la formación continua en función de sus necesidades a lo largo de la vida.

Así pues, la UOC, en el marco de la cátedra de multilingüismo Linguamón-UOC, ha querido incidir en este fenómeno y aportar mejoras para las aulas multiculturales. En este sentido, se ha desarrollado un espacio virtual de aprendizaje multilingüe en tiempo real. Esta aplicación se ha construido integrando un sistema de traducción automática (Apertium) en un entorno virtual d'aprendizaje (Moodle). Como estas aplicaciones se distribuyen bajo una licencia de software libre, se prevé liberar todo el proyecto, hecho muy importante si se tiene en cuenta que el producto final podrá ser utilizado por cualquier persona o institución de manera totalmente libre y gratuita.

Ahora el alumno, sea de la cultura que sea y hable la lengua que hable, podrá acceder a la información, leer los mensajes del profesor y dirigirse a sus compañeros sin necesidad de recurrir a una lengua franca. Todos podrán expresarse en su lengua materna, en total libertad, y todos recibirán la formación en su misma lengua, sin ningún tipo de retraso. Esto es lo que permite el aula virtual multilingüe.

El artículo se organiza de la siguiente manera: en primer lugar explicaremos brevemente qué son las aplicaciones Moodle y Apertium. Después explicaremos los aspectos técnicos que se han desarrollado en este proyecto para integrar estas dos aplicaciones, en concreto la creación de un corrector ortográfico automático. A continuación explicaremos los aspectos lingüísticos que se han estudiado para alimentar los correctores ortográficos. Por último presentaremos el prototipo y las conclusiones.

## ¿Qué es Moodle?

Moodle (<http://moodle.org>) es un sistema de gestión de cursos en línea (CMS - *course management system*) que se distribuye gratuitamente como software libre bajo la licencia pública GNU. La palabra Moodle era al principio un acrónimo de *Modular Object-Oriented Dynamic Learning Environment* (Entorno de Aprendizaje Dinámico Orientado a Objetos y Modular). Este sistema permite crear de manera sencilla cursos que se pueden realizar 100 % en línea. Como el sistema es muy fácil de instalar y mantener, no tiene demasiados requisitos de hardware ni de software y la creación de cursos es muy sencilla, Moodle se ha convertido en un sistema CMS realmente popular.

## ¿Qué es Apertium?

Apertium es una plataforma de traducción automática de software libre que se distribuye bajo licencia GPL. Mediante las herramientas que ofrece esta plataforma se pueden desarrollar nuevos pares de lenguas o bien mejorar los pares de lenguas ya existentes. Es decir, además de distribuir todo lo necesario para poder usar el sistema de traducción automática también se distribuyen todas las herramientas necesarias para poder crear nuevos sistemas:

- Un motor de traducción común para todos los pares de lenguas disponibles
- Datos lingüísticos para varios pares de lenguas
- Herramientas para gestionar los datos lingüísticos necesarios para construir un sistema de traducción automática para un determinado par de lenguas.

Apertium funciona mediante transferencia superficial. Esta metodología consiste en analizar sintácticamente la oración en la lengua de partida y transferir la información a la lengua de llegada para poder generar la oración traducida. Este análisis sintáctico no es profundo, sino superficial, por lo tanto, el árbol sintáctico que se genera no es completo y pueden faltar relaciones entre algunos constituyentes.

Esta metodología de traducción automática está inicialmente concebida para pares de lenguas emparentadas (como catalán-castellano, catalán-francés, castellano-portugués, etc.). De todos modos, se ha expandido la metodología para poder tratar pares de lenguas más divergentes (catalán-inglés, por ejemplo).

## ¿Cómo funciona Apertium?

Etapas del procesamiento del texto de entrada de Apertium:

- desformateo: se tratan varios formatos de entrada (texto, RTF, html, etc.) y se separa el texto a traducir de la información de formato;
- análisis morfológico: se asignan todas las informaciones morfológicas posibles a cada palabra del texto de entrada;
- desambiguación categorial: se desambigua la información morfológica asignada en la etapa anterior, de forma que a cada palabra sólo le quede una posible lectura;
- análisis sintáctico superficial: se construye un árbol de análisis superficial de la frase de partida;
- transferencia sintáctica superficial: se transfiere el árbol de análisis superficial de la frase de partida a un árbol de análisis superficial equivalente en la lengua de llegada;
- transferencia léxica: se busca y se selecciona la traducción más adecuada en la lengua de llegada de las palabras de la frase de partida;
- generación morfológica: se generan las formas adecuadas de las palabras traducidas (sea plural o singular, el tiempo y la persona para los verbos, etc.);
- reformateo: se recupera la información de formato para recuperar el formato del documento original.

Web de Apertium: <http://xixona.dlsi.ua.es/apertium-www/>. En esta web se puede consultar información sobre el proyecto. Las descargas se pueden hacer desde esta página o desde Source Forge: <http://sourceforge.net/projects/apertium/>

## Aspectos técnicos: funcionamiento de la plataforma integrada

La idea del proyecto consiste en traducir automáticamente los mensajes de Moodle a la lengua que seleccione el usuario, de modo que todo el mundo pueda acceder a la información en la lengua que desee. Con todo, la conexión de Moodle y Apertium no se hace directamente, es decir, no se traducen directamente los mensajes que aparecen a Moodle con Apertium. En este sentido, y gracias a proyectos anteriores (Climent-2005 y Moré-2005), se ha observado que, en general, se produce una cantidad elevada de faltas de ortografía en los mensajes enviados a foros académicos. Además, presentan algunas características propias (como abreviaturas, uso de emoticonos, etc.) que hacen que el rendimiento de los sistemas de traducción automática para este tipo de mensajes no sea demasiado satisfactorio.

El primer paso del proceso de traducción de los mensajes es la detección automática de la lengua en la que está escrito el mensaje original. En nuestro sistema se ha restringido la lengua de los mensajes al catalán, castellano, inglés o francés. Para evitar la intervención del usuario en todo el proceso, esta información no se pide sino que se determina mediante un algoritmo de detección automática de lengua, a saber el TextCat de Gertjan van Noord (<http://www.let.rug.nl/~vannoord/TextCat/>) basado en Cavnar-94.

Una vez determinada la lengua del mensaje, y como paso previo a la traducción, los mensajes pasan por un corrector ortográfico. El corrector ortográfico se ha diseñado para que sea totalmente automático y no requiera ninguna intervención por parte del usuario. Por esta razón, sólo corrige aquellas faltas que son claramente errores. Una opción alternativa sería hacer uso de un corrector ortográfico normal que requiera la intervención del usuario, idea que se descartó porque puede provocar el rechazo del usuario, que se encontraría con la necesidad de hacer un paso más. En todo caso, el usuario puede hacer uso de otros correctores ortográficos, como por ejemplo el que ofrece el navegador Mozilla.

El módulo de corrección ortográfica que hemos desarrollado funciona en tres etapas:

- Una primera etapa realiza una serie de cambios tipográficos en el texto. Este módulo funciona con expresiones regulares del estilo `s/ /' /g`. Otra transformación que realiza, por ejemplo, es el cambio del punto de las *eles* geminadas del catalán por el punto medio.
- Una segunda etapa se encarga de hacer una serie de cambios *ad-hoc* que vienen especificados por un fichero que se puede adaptar a las propias necesidades. Estos cambios *ad-hoc* sirven para expresar errores típicos y sus soluciones, para resolver aquellos casos en que la elección de la opción correcta por parte de la tercera etapa sea difícil o imposible. Por ejemplo, en catalán podría expresar cambiar `"*susto"` por `"ensurt"`.
- La tercera etapa de la corrección ortográfica funciona mediante un formulario y un diccionario (o más) dado por el usuario. Un diccionario de formas o formulario es una lista de palabras (cuanto más extensa mejor) flexionadas en todas sus formas (singular, plural, los diferentes tiempos verbales y personas y números, etc.). Este formulario se puede completar con diccionarios personalizados. El corrector verifica si todas las palabras del texto a corregir están en esta lista y

si no están intenta encontrar la palabra de la lista que tenga más posibilidades de ser la palabra correcta.

Que una palabra no esté en la lista no quiere decir necesariamente que sea una falta de ortografía. Un caso típico de esta situación es los nombres propios, especialmente los extranjeros, que no estarán incluidos en el formario. Para evitar este hecho el corrector que hemos diseñado sólo corrige las palabras escritas en minúsculas. Esto puede considerarse una limitación del sistema, pero cabe recordar que nuestro corrector no tiene ninguna intervención del usuario y que sólo puede corregir aquello que sea muy seguro. Una vez tiene una palabra escrita en minúscula no contenida en el formario ni en los diccionarios del usuario el algoritmo debe ser capaz de encontrar las palabras más parecidas a ésta y decidir cuál tiene más posibilidades de ser la corrección. Esta búsqueda, además, tiene que ser muy rápida y sobre una lista de formas muy extensa. Para hacer esto realiza los siguientes pasos:

- Elimina un carácter en cada una de las posiciones posibles y verifica si el resultado está en el formario. Esto funcionaría para un error del estilo "*\*dessig*" escrito en catalán en vez de "*desig*", en este caso el sistema probaría *\*essig*, *\*dssig*, *desig*, *desig*, *\*dessg*, *\*dessi* y encontraría la opción correcta *desig*.

- Añade un carácter a todas las posiciones de la palabra y verifica si el resultado está en el formario. Lo intenta con todos los caracteres posibles de la lengua. Si continuamos con el ejemplo de "*\*dessig*" el sistema haría *\*adessig*, *\*bdessig*, *\*cdessig*,..., *\*daessig*, *\*dbessig*, etc. En este caso no encontraría ninguna palabra.

- Cambia cada carácter de la palabra por todos los caracteres posibles de la lengua y verifica si el resultado está en el formario. Continuando con el mismo ejemplo tendríamos *\*aessig*, *\*bessig*, *\*cessig*, ..., *\*dassig*, *\*dbssig*, *\*dcssig*, etc. En este caso encontraría dos posibles palabras de la lengua: *dessis* y *dessin*, las dos formas del verbo *dar*.

El corrector ahora debe decidir cual de las tres posibles formas correctas (*desig*, *dessis* y *dessin*) es la que tiene más posibilidades de ser la corrección de *\*dessig*. Para hacer esto se ha desarrollado una estadística de errores más frecuentes, que nos indicará que el cambio *\*ss:s* (de *\*dessig* a *desig*) es más frecuente que el cambio *g:s* (de *\*dessig* a *dessis*) y que el cambio *g:n* (de *\*dessig* a *dessin*).

Una vez corregido el mensaje ya se puede traducir automáticamente y presentar la versión original y sus traducciones. La versión original que se presenta es la versión sin corregir ortográficamente, aun cuando las traducciones que se presentan son generadas a partir de la versión corregida.

## Aspectos lingüísticos

Como ya se ha dicho, el corrector ortográfico automático se ha adaptado a una tipología textual muy concreta, por tal de sacarle el máximo de rendimiento posible. El corrector necesitará los siguientes recursos:

- Formarios para cada una de las lenguas
- Listas de *ad-hocs*
- Estadísticas de errores más frecuentes

## ***Formarios***

Un formario es una lista de palabras de una determinada lengua flexionadas en todas sus formas. Estas listas pueden llegar a tener medidas considerables, especialmente para lenguas mayoritariamente flexivas, como el catalán, el castellano o el francés. En este apartado explicaremos de dónde hemos obtenido los formarios para las lenguas de trabajo y la medida de estos formarios.

- Catalán: se ha usado el diccionario morfológico que ha sido desarrollado por el grupo Glicom de la Universitat Pompeu Fabra para la creación de "El corrector", un corrector ortográfico y gramatical del catalán que se liberará como software libre. De este diccionario se han aprovechado sólo las formas y se ha desestimado el resto de información. El formario del catalán cuenta con un total de 634.000 entradas.
- Castellano: se ha usado el diccionario del corrector Ispell y cuenta con 653.000 entradas
- Los formarios del inglés y del francés todavía no se han confeccionado, pero muy probablemente se harán también a partir de los diccionarios del corrector Ispell.

Para poder hacer búsquedas muy rápidas en estas listas de medida considerable, la información se ha convertido en bases de datos Berkeley. Las bases de datos Berkeley están organizadas como pares clave-valor y permiten una búsqueda muy rápida. Así, usan la clave para almacenar la forma y el valor simplemente se instancia (adquiere un valor cualquiera, en este caso "1"), de forma que para verificar si una forma aparece en el formario sólo hace falta comprobar si existe una entrada con esta clave.

## ***Confección del corpus***

Para adaptar al máximo el corrector a la tipología textual prevista para los mensajes de Moodle, hemos llevado a cabo un estudio tomando una muestra de los mensajes de los foros del Campus Virtual de la UOC. La intención era obtener dos corpus: uno para el catalán con los mensajes que hemos encontrado en catalán y otro en castellano para los mensajes en esta lengua, para poder estudiar qué tipo de lenguaje se emplea en estos casos. En el Campus virtual, dentro del apartado Comunidad > Foros, hemos recopilado los mensajes clasificados como "Académicos". Esto ha sido posible para el catalán y para el castellano, pero no para el inglés ni el francés porque no disponíamos de los datos necesarios (en los foros no hay mensajes en estas lenguas).

Hemos intentado escoger los foros por su diversidad, que comprendieran todas las temáticas y usuarios posibles, para conseguir un léxico lo más extenso posible. Nos hemos quedado con los 25 mensajes más recientes de cada foro en el momento de la consulta (principios de octubre de 2007). En total hemos analizado 290 mensajes de 17 foros diferentes, con usuarios de edades comprendidas desde los 20 hasta mayores de 45 años, con participación tanto de alumnos como de profesores.

## ***Listas de ad-hocs***

A partir de los corpus constituidos, hemos analizado los errores y características de esta tipología

textual, para sistematizar en la medida de lo posible la corrección automática de errores. Este proceso consiste en identificar cada falta, clasificarla según su origen y buscar técnicas de corrección adecuadas y específicas a cada tipo de error. Esto nos ha permitido elaborar las listas de *ad-hocs*: relación de palabras que normalmente son susceptibles de ser mal escritas (y que constan como errores en el corpus), junto con su corrección. Para este corrector automático en concreto, las listas de *ad-hocs* se han confeccionado a partir de dos fuentes de información: principalmente a partir de los errores encontrados en el corpus elaborado con este fin, pero también de listas de corrección de los servicios lingüísticos de las universidades catalanas.

Así pues, hemos confeccionado dos listas de *ad-hocs*, una para el catalán y otra para el castellano. Cada lista tiene la palabra donde hemos encontrado el error, su forma en femenino (o masculino si procede) y las formas en plural, siguiendo el patrón del error y corregidas en cada caso:

incluïdo	incluido
incluïdos	incluidos
incluïda	incluida
incluïdas	incluidas

En los casos en qué la palabra mal escrita era un verbo, hemos incluido en la lista las formas siguientes: el infinitivo, la primera persona del presente de indicativo, la primera persona del pretérito imperfecto y, en los casos pertinentes, la forma que aparecía mal escrita en el corpus (fuera el tiempo verbal que fuera):

obserbár	observar
obserbaba	observaba
obserbo	observo

La lista de *ad-hocs* también contiene una relación estadística de los cambios más frecuentes, a saber, por ejemplo, en cuántas ocasiones he utilizado *b* cuando en realidad tenía que utilizar *v*. De este modo la estadística nos servirá para decantarnos por una forma u otra en los casos en que haya ambigüedad. Por ejemplo, cada vez que nos indique que hay un error en la palabra *\*aixo*, el corrector sabrá que debe sustituirla por *això*, porque el cambio *o:ò* es uno de los más frecuentes del corpus.



## ***Estadísticas de errores más frecuentes***

Hemos contabilizado y clasificado todos los errores del corpus según los criterios establecidos anteriormente, y como dato más significativo se puede remarcar que la mayoría de errores son de acentuación, tanto en catalán como en castellano. No es fácil decir si se trata de un error intencional o no, porque si bien es cierto que el hecho de no acentuar un texto forma parte del nuevo registro usado en los correos electrónicos, no se mantiene la coherencia. Con esto queremos decir que en un mensaje tanto se acentúa como no se acentúa, por lo cual no podemos considerar que sea un error intencional, sino más bien ortográfico. En catalán, por ejemplo, nos hemos encontrado casos como este:

*\*informatica de sistemes o de gestió.*

Además, en el caso concreto del catalán, si bien la mayoría de errores de acentuación son causados por la falta de acentuación (acentuación negativa), muchos de los errores de acentuación positiva (acentuar palabras incorrectamente) denotan una clara influencia del castellano (*\*podría*).

Podemos apreciar que la cantidad de acentos “negativos” es muy importante, y que por lo tanto es probable que comporten un grado de errores muy elevado. Llegados a este punto, es necesario diferenciar entre los errores y los “no errores”, errores que son palabras incorrectas en el contexto en que son utilizadas. En un traductor automático, este tipo de errores conlleva mucha controversia. En este sentido, podemos considerar el caso en catalán de *angles* como buen ejemplo para ilustrar los posibles problemas para el traductor. Si no hay ningún tipo de edición previa a la corrección automática, el traductor, por ejemplo, nos convertirá esta palabra en *\*ángulos* cuando lo traducimos hacia el castellano, si bien en ningún caso no se hace referencia a la figura geométrica, si no a la lengua, el *inglés*.

Como resumen, la mayoría de los ítemes contemplados en la lista de errores son causados por una acentuación incorrecta, que en el caso del catalán se debe a una fuerte influencia del español y a la inseguridad de los usuarios a la hora de escribir en lengua catalana, y que en general se justifica por la falta de atención propia del registro escrito de los mensajes de correo electrónico. Tras el estudio, podemos afirmar, que se comenten más faltas en catalán que en castellano.

## **Prototipo**

Uno de los aspectos fundamentales de esta integración es la forma de presentar las diferentes versiones de un mismo mensaje. La primera idea que surge es la de presentar los mensajes en todas las lenguas (el original y las diversas traducciones) en una misma pantalla. Esta opción puede ser válida para entornos bilingües, pero si el entorno es multilingüe no es una solución adecuada, puesto que enseguida se verán las pantallas llenas de texto y será difícil de encontrar el mensaje en la lengua deseada.

Así pues, se ha optado por otra presentación, en la cual cada usuario puede definir su lengua por defecto. El usuario verá en todo momento todos los mensajes en la lengua deseada, sea ésta la lengua original o bien una traducción, y teniendo constancia de esta información. De este modo, podemos saber si leemos un original o bien una traducción, y además, en caso de leer una

traducción sabremos cuál es la lengua original y tendremos un acceso inmediato a la versión original.



Figura 1. Página de inicio del sistema al escoger como lengua por defecto el inglés.



Figura 2. Página de inicio del sistema al escoger como lengua por defecto el catalán.

En la figura 1 podemos ver la página de inicio al sistema si escogemos como lengua por defecto el inglés. Fijémonos que el texto de la derecha "*This site is for implementing...*" nos aparece en inglés y nos indica que es el texto original. En cambio, en la figura 2 observamos la misma pantalla si escogemos como lengua por defecto el catalán. Como podemos ver, el texto de la derecha aparece en catalán, pero sabemos que es una traducción. Mediante los enlaces de debajo del texto podemos visualizar el texto original (figura 3) o todas las traducciones (figura 4).

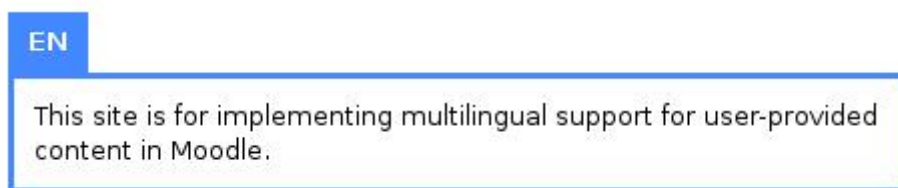


Figura 3. Pantalla que muestra el texto original.

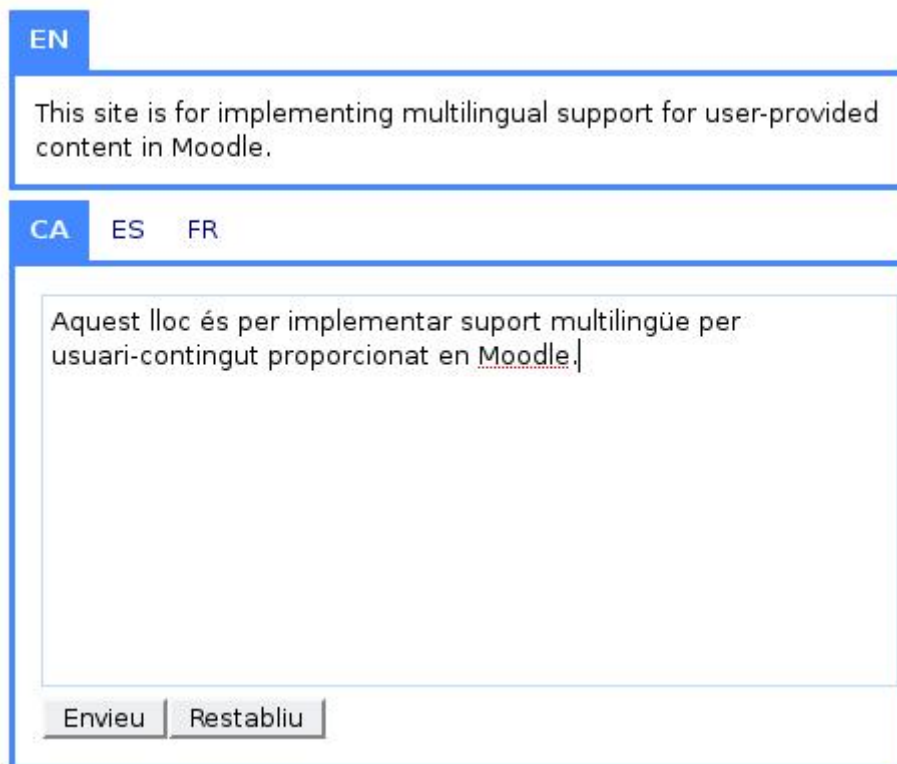


Figura 4. Pantalla que muestra el texto original y las diferentes traducciones (que se pueden seleccionar con las pestañas).

Una característica interesante de la visualización de las traducciones es que se puede establecer una serie de permisos que permitan a todos los usuarios o a un grupo de ellos modificar las traducciones realizadas automáticamente. De este modo un usuario con suficientes permisos puede corregir los posibles errores de traducción cometidos por el traductor automático y mejorar así la calidad de la traducción.

## Conclusiones

El entorno de aprendizaje que hemos presentado en este artículo está listo para acoger cursos y alumnos multilingües, mostrar los mensajes en catalán, castellano, inglés o francés o traducir los mensajes entre los usuarios. Para saber si el funcionamiento general es satisfactorio o no, si el usuario encuentra la interfaz agradable y fácil de usar y el traductor traduce correctamente y en tiempo real, se sometió la aplicación a un periodo de pruebas en un contexto de aprendizaje virtual real.

Posteriormente, se diseñó una encuesta para descubrir el grado de satisfacción de los alumnos y de los profesores. Los resultados obtenidos muestran que el grado de satisfacción de los usuarios depende en última instancia de sus características: si los usuarios están acostumbrados a comunicarse en más de una lengua apreciarán el potencial multilingüe de la herramienta. Si, por el contrario, son usuarios monolingües, sólo les interesará poder consultar la información en su lengua, porque no están sensibilizados con el resto de lenguas, ni siquiera las consultan. Sin embargo, la impresión general es positiva y demuestra que la mayoría de usuarios ven con buenos ojos el entorno virtual de

aprendizaje multilingüe.

### **Bibliografía**

[Cavnar-94] Cavnar, W. B. and J. M. Trenkle, N-Gram-Based Text Categorization In Proceedings of Third Annual Symposium on Document Analysis and Information Retrieval, Las Vegas, NV, UNLV Publications/Reprographics, pp. 161-175, 11-13 April 1994

[Climent-2005] Climent S, J. Moré, A. Oliver, M. Salvatierra, I. Sànchez y M. Vázquez. Technologies de la traducció per a la gestió de la doble oferta docent en català i castellà a la UOC. Zeitschrift für Katalanistik/Revista d'Estudis Catalans 18. p. 31-57 (ISSN 0932-2221

[More-2005]

[Moré-2005] J. Moré, S. Climent, A. Oliver, M. Taulé. Análisis de los fenómenos lingüísticos de los mensajes de correo electrónico en catalán desde la perspectiva de la traducción automática. XXI Congreso de la SEPLN. Granada. Septiembre 2005.