

AVI.cat: Asistente virtual para la mejora de la redacción en catalán

AVI.cat: a virtual assistant for the improvement of writing skills in Catalan

Antoni Oliver, Salvador Climent, Marta Coll-Florit

Universitat Oberta de Catalunya
Avda. Tibidabo 39-43 08035 Barcelona
aoliverg,scliment,mcollfl@uoc.edu

Resumen: En esta demostración presentamos un primer prototipo de asistente para la mejora de la redacción en catalán. El sistema va más allá de un simple corrector gramatical, ya que propone enlaces a gramáticas y ejercicios que permiten al usuario practicar los aspectos donde presenta más carencias. El sistema funciona también como evaluador de nivel y permite realizar un seguimiento de las mejoras del usuario.

Palabras clave: corrección gramatical

Abstract: In this demo we present a first prototype of an assistant for the improvement of writing skills in Catalan. The system is more than a grammatical checker as it proposes links to grammatical explanations and exercises, allowing the user to practice specific aspects. The program also works as a level evaluator and allows to track the user's improvements.

Keywords: grammar checking

1. *Introducción*

Este proyecto está desarrollando un asistente virtual para la mejora de la redacción en catalán. El proyecto se está desarrollando a raíz de una convocatoria de proyectos internos de innovación docente APLICA de la Universitat Oberta de Catalunya (UOC). Esta aplicación se integra en el procesador de textos OpenOffice.org Writer y tiene dos modos de funcionamiento: (1) evaluador de nivel y (2) corrector gramatical y asesor lingüístico.

Los procesadores de textos habituales (Microsoft Word y OpenOffice.org Writer) ofrecen correctores ortográficos para el catalán, pero no correctores gramaticales. La diferencia entre ambos es que mientras los correctores ortográficos sólo dicen si una palabra es o no posible en una lengua, los correctores gramaticales analizan el contexto donde se inserta y dicen si la palabra es correcta en ese contexto.

El asesor virtual AVI.cat se integra en Open Office.org Writer (con posibilidad de futura integración en otros procesadores) y funciona como corrector gramatical, pero no sólo eso, sino que, además, es una herramienta de evaluación de la calidad de redacción

-haciendo las funciones de prueba de nivel y un asesor para la mejora de la calidad de escritura del usuario.

2. *Modos de funcionamiento*

Cuando AVI.cat funciona en modo corrector gramatical, detecta el error, emite el correspondiente mensaje con la propuesta de corrección y, además, recomienda los enlaces apropiados a piezas de conocimiento y ejercicios gramaticales donde el usuario, si lo desea, podrá trabajar el error cometido y así mejorar sus competencias en ese aspecto concreto. En el segundo modo de funcionamiento (como herramienta de evaluación), a partir del documento del usuario, AVI.cat emite una valoración cuantitativa y cualitativa del nivel del texto y también, como en el caso anterior, ofrece unos enlaces para trabajar los errores específicos que el usuario ha cometido y así mejorar su nivel de redacción en aquellos aspectos concretos donde se concentran sus carencias.

3. *Recursos utilizados*

Para implementar el asesor virtual se parte de los recursos libres que se relacionan a

continuación, que se integrarán, mejorarán y serán dotados de nuevas funciones.

Por un lado, existe un buen corrector gramatical del catalán, *El Corrector*¹ (Badia et al., 2004), desarrollado por la Universidad Pompeu Fabra a raíz de un concurso público de la Generalitat de Cataluña. Este corrector destaca por su exhaustivo estudio empírico de los errores gramaticales más frecuentes y por ser una herramienta totalmente libre. En cambio, su punto más débil es la falta de una integración adecuada con los procesadores de textos. En este proyecto aprovechamos especialmente el documento *Especificacions d'errors d'El Corrector* (Domingo et al., 2008), una recopilación y clasificación de especificaciones de errores gramaticales para el catalán. Nuestras reglas se desarrollarán siguiendo esta guía y manteniendo los códigos asignados. Para cada tipo de error se ofrece una información que consta de una descripción, la especificación del elemento a marcar, la especificación de la propuesta de corrección, el mensaje de error, ejemplos así como una serie de observaciones.

Por otra parte, la herramienta libre *LanguageTool*² (Naber, 2003) proporciona un corrector gramatical genérico para OpenOffice.org Writer. LanguageTool permite crear correctores gramaticales para cualquier lengua y ofrece todas las herramientas de integración con el procesador de textos. Las reglas se especifican en lenguaje XML y actualmente incorpora un conjunto de reglas muy reducido para el catalán. Durante el proyecto desarrollaremos inicialmente las reglas del documento de especificaciones de *El Corrector* en el formato XML propio de *LanguageTool*.

En cuanto a los enlaces que ofrece el asistente en sus funciones de asesor lingüístico, se están seleccionando entre los recursos libres y gratuitos disponibles en Internet, partiendo en primer lugar y como fuente principal de los apartados Criterios lingüísticos y Dudas más frecuentes de la web de los Servicios Lingüísticos de la UOC³. Además de esta fuente principal, se ha llevado a cabo una importante tarea de búsqueda de enlaces a información gramatical y ejercicios y se ha clasificado toda esta información de acuerdo a los códigos de errores de las especificacio-

nes de *El Corrector*. Esto permitirá relacionar directamente el error detectado con la información gramatical y ejercicios asociados.

4. Conclusiones

El proyecto Avi.cat se está desarrollando este año 2011 y tiene prevista su finalización en diciembre. Durante la demostración el sistema contará con un número suficiente de reglas y enlaces gramaticales y a ejercicios. Una vez finalizado su desarrollo se distribuirá con una licencia de software libre. Los desarrolladores tienen la esperanza que la herramienta sea útil para la mejora del nivel de redacción en catalán de los usuarios y no descartan el desarrollo de herramientas similares para otras lenguas. En este sentido, creen que este tipo de herramientas puede ser útil para la mejora del nivel de redacción en lenguas extranjeras si se crean conjuntos de reglas específicas para esta necesidad, que recojan los errores más frecuentes de redactores no nativos.

Bibliografía

- Badia, T., A. Gil, M. Quixal, y O. Valentín. 2004. NLP-enhanced error checking for catalan unrestricted text. En *Proceedings of the fourth international conference on Language Resources and Evaluation, LREC*, página 1919–1922.
- Domingo, Judith, Guillem Massó, Òscar Puente, Bernat Grau, Toni Badia, y Martí Quixal. 2008. Especificacions d'errors d'El corrector. Informe técnico, GLiCom, Grup de Lingüística Computacional. Universitat Pompeu Fabra.
- Naber, D. 2003. *A rule-based style and grammar checker*. Master's thesis, University of Bielefeld.

¹www.elcorrector.cat

²www.languagetool.org

³<http://www.uoc.edu/serveilinguistic/home/index.html>