

Vieiros

Desenvolvido polo Centro Ramón Piñeiro

O corpus da lingua actual amplíase até os vinte millóns de formas ortográficas

A base de datos está a disposición de todas as persoas que o desexen con tan só solicitar unha clave de acceso.

Redacción - 18:20 28/03/2007

Tags: actual centro ramón piñeiro corpus lingua política lingüística

A nova versión do Corpus de Referencia do Galego Actual (Corga) engade seis novos millóns de formas ortográficas á súa base de datos, co que o total ascende a 19,5 millóns. O Corga recolle textos na nosa lingua escritos en diferentes publicacións e obras literarias dende o ano 1975 e está desenvolvido polo Centro Ramón Piñeiro para a investigación en Humanidades.

Toda esa información está codificada empregando o estándar XML, e permite facer buscas de palabras ou expresións e seleccionar parámetros como tipos de texto, épocas, temas ou unha combinación dos anteriores. O recurso está destinado principalmente aos investigadores da lingua actual, aínda que todas as persoas que desexen consultalo poden solicitar unha clave de acceso.

O Centro Ramón Piñeiro segue a traballar na ampliación do corpus e en dotalo de novas funcións, como a incorporación dunha etiquetaxe morfosintáctica que aumentará a flexibilidade das consultas. O proxecto está dirixido por Guillermo Rojo, e conta con Eva Domínguez como coordinadora lingüística e con Mario Barcala como responsábel da parte informática.

Aínda que xa está dispoñíbel dende o 12 de marzo, o servizo presentouse oficialmente o mércores, coa presenza da secretaria xeral de Política Lingüística Marisol López, quen valorou o feito de que as ferramentas postas a disposición da sociedade polo Centro Ramón Piñeiro están a conseguir un mellor coñecemento da lingua.