

El proyecto *EX-TermGram*: extracción automática de términos y definiciones utilizados en el dominio de las gramáticas de español

Giuseppe Forte (FAO) - María Lozano Zahonero (Universidad de Roma "Tor Vergata")

"The more that is known about a language, the more we can find out about it", W. Labov (1972). *Some Principles of Linguistic Methodology*. Filadelfia: University of Pennsylvania Press, p. 202.

Uno de los principales problemas de cualquier esfera del conocimiento es organizar y definir los términos y conceptos que utiliza. Como es sabido, la gramática del español, y más concretamente la gramática de español para italianos, son disciplinas que no disponen de amplios recursos metalingüísticos ni terminográficos. Partiendo de estas dos premisas y dentro del ámbito del proyecto *Contrastiva*, se ha elaborado *EX-TermGram*, un proyecto de extracción automática de términos y definiciones utilizados en el dominio de las gramáticas de español, con especial atención a las gramáticas de español para italianos. El objetivo último es la elaboración de un banco de datos y/o banco de conocimiento que organice este dominio específico y pueda constituir un útil instrumento para la realización de trabajos de investigación y obras gramaticales y metagramaticales de carácter tanto sincrónico como diacrónico.

En nuestra ponencia se presentarán la justificación y alcance del proyecto y de sus distintas fases, los fundamentos para la elaboración del método de extracción automática y los experimentos y tareas realizados con tal fin, los problemas encontrados y sus posibles soluciones, así como los resultados obtenidos y el estado actual del proyecto.