Regular Sound Changes for Cross-Language Information Retrieval
Michael P. Oakes (University of Sunderland, UK)

Cognate words are vocabulary items which occur in two or more historically related languages, such that they have similar meanings, and one can be transformed into the other by a predictable series of phonological changes. The discovery of such sound changes is potentially relevant to Cross-Language Information Retrieval: the automatic conversion of query terms in one language into their equivalents in a second, historically related, language will enable documents in that second language can be retrieved. We will describe one computer program called JAKARTA, that given a sample word list from two related languages, extracts the probable rules for predicting any word of one language from that of the other. As a case study we will examine vocabulary lists for Galician and Spanish.

Martes, 14 de marzo de 2006, 11:30, Aula de Grados

Organizadores: Miguel A. Alonso Pardo (alonso@udc.es) y Jesús Vilares Ferro (jvilares@udc.es)