

# Cross-Lingual Question Answering Using Inter Lingual Index Module of EuroWordNet\*

Sergio Ferrández and Antonio Ferrández

Natural Language Processing and Information Systems Group  
Department of Software and Computing Systems  
University of Alicante, Spain  
{sferrandez,antonio}@dlsi.ua.es

**Abstract.** This paper outlines the BRILI cross-lingual English-Spanish-Catalan Question Answering (QA) system. The BRILI is being designed at University of Alicante and will be capable to answer English, Spanish and Catalan questions from English, Spanish and Catalan documents. The starting point is our monolingual Spanish QA system [11] which was presented at the 2005 edition of the Cross-Language Evaluation Forum (CLEF). We describe the extensions to our monolingual QA system that are required, especially the language identification module and the strategy used for the question processing module. The Inter Lingual Index (ILI) Module of EuroWordNet (EWN) is used by the question processing modules. The aim of this is to reduce the negative effect of question translation on the overall accuracy of QA systems.

## 1 Introduction

The aim of a Question Answering (QA) system is to localize the correct answer to a question in natural language in a non-structured collection of documents, also the situations where the system is not able to provide an answer should be detected. In the case of a Cross-Lingual QA (CL-QA) system, the question is formulated in a language different from the one of the documents, which increases the difficulty. Nowadays, multilingual QA systems have been recognized as an important issue for the future of Information Retrieval (IR).

In this paper we present BRILI (Spanish acronym for "Question Answering using Inter Lingual Index Module"). It is a CL-QA system for Spanish, English and Catalan. It is designed to localize answers from documents, where both answers and documents are written in the three languages. The system is based on complex pattern matching using NLP tools [1, 4, 7, 12]. Beside, Word Sense Disambiguation (WSD) is applied to improve the system (a new proposal of WSD for nouns based on [2]).

BRILI is fully automatic, including the modules of language identification and question processing. The main goal of this paper is to describe these modules

---

\* This research has been partially funded by the Spanish Government under project CICYT number TIC2003-07158-C04-01 and by the Valencia Government under project number GV04B-268.