Búsqueda de Respuestas Bilingüe basada en ILI, el sistema BRILI *

Sergio Ferrández, Antonio Ferrández, Sandra Roger y Pilar López–Moreno
Grupo de Investigación en Procesamiento del Lenguaje Natural
y Sistemas de Información
Departamento de Lenguajes y Sistemas Informáticos
Universidad de Alicante
Carretera San Vicente s/n 03690 Alicante España
{sferrandez,antonio,sroger}@dlsi.ua.es P.Lopez@ua.es

Resumen: Este artículo presenta nuestro sistema de Búsqueda de Respuestas (BR) bilingüe, llamado BRILI. BRILI ha participado en la tarea de BR Inglés–Castellano de la edición del 2006 del "*Cross-Language Evaluation Forum*" (CLEF) quedando en primera posición. BRILI está basado en el uso del módulo ILI ("Inter Lingual Index") de EuroWordNet. Además, se muestra una evaluación sobre el conjunto de preguntas oficiales del CLEF que justifica nuestra estrategia.
Palabras clave: Búsqueda de Respuestas Multilingüe, ILI, EuroWordNet

Abstract: This paper presents our Cross-lingual Question Answering (CL-QA) system, called BRILI. It has participated in the 2006 edition of the Cross-Language Evaluation Forum (CLEF) being ranked first at English–Spanish Question Answering task. BRILI is based on the Inter Lingual Index (ILI) Module of EuroWordNet. Besides, tests on official questions from the CLEF justify our strategy.
Keywords: Cross–Lingual Question Answering, ILI, EuroWordNet

1. Introducción

En estos últimos años, el volumen de información textual en lenguaje natural en diferentes idiomas ha crecido exponencialmente. Este hecho implica la necesidad de gran número de herramientas de Recuperación de Información (RI) que permitan a los usuarios realizar búsquedas de información multilingüe.

Tareas multilingües en RI y Búsqueda de Respuestas (BR) han sido reconocidas como un importante uso en el acceso de información *on-line*, este fenómeno fue corroborado en la pasada edición del "*Cross-Language Evaluation Forum*"(CLEF) 2006 (Magnini et al., 2006).

El objetivo que la BR persigue es encontrar respuestas concisas a preguntas arbitrarias formuladas en lenguaje natural. En entornos multilingües, las preguntas son formulas en lengua diferente a la de los documentos donde se realiza la búsqueda de la respuesta. Este marco de trabajo incrementa la dificultad del proceso de BR.

En este artículo se presenta el sistema BRILI (acrónimo de "Búsqueda de Repuestas usando ILI"), siendo el objetivo principal de este artículo describir nuestra estrategia para afrontar las tareas de BR multilingües (el prototipo de la arquitectura del sistema BRILI fue presentado previamente en (Ferrández y Ferrández, 2006)). Además, se propone una evaluación sobre los conjuntos de preguntas oficiales del CLEF que justifica la bondad de nuestro método.

El resto del artículo está organizado en: la sección 2 detalla el estado de la cuestión estudiando los actuales sistemas de BR multilingüe. En la sección 3 la arquitectura del sistema BRILI es mostrada y analizada, haciendo especial hincapié en la estrategia desarrollada para procesar las tareas multilingües. La sección 4 presenta y discute los resultados obtenidos de la evaluación del sistema usando las preguntas oficiales del CLEF. Y finalmente, la sección 5 redacta nuestras conclusiones y trabajos futuros.

* Este trabajo ha sido llevado a cabo por el consorcio QALL-ME, el Sexto Programa Marco de Investigación de la Unión Europea (UE), referencia: FP6-IST-033860. Los autores agradecen a la UE su apoyo económico, así como a los integrantes del consorcio, por su fructífera colaboración. Para más información sobre el consorcio QALL-ME visite la página web del proyecto, http://qallme.itc.it/.

 © Sociedad Española para el Procesamiento del Lenguaje Natural

2. Estado de la Cuestión

Hoy en día, la mayoría de los actuales sistemas de BR multilingüe (Gillard et al., 2006; Sacaleanu, B. y G. Neumann, 2006; Sutcliffe et al., 2006; Tomás et al., 2006) están basados en el uso de servicios *on-line* de Traducción Automática (TA). Esta tendencia fue confirmada el la edición del 2006 del CLEF (Magnini et al., 2006).

La precisión de los sistemas de BR multilingüe está directamente afectada por su capacidad de analizar y traducir correctamente la pregunta que reciben como entrada al proceso. Imperfectas o ambiguas traducciones de las preguntas ocasionan un efecto negativo en la precisión global del sistema (un estudio sobre el efecto de la TA en la BR multilingüe fue presentado en (Ferrández y Ferrández, 2007)). Otro dato a tener en cuenta y tomando como referencia a (Moldovan et al., 2003), es que la fase de Análisis de la Pregunta es responsable del 36.4 % del total de los errores en la BR de domino abierto.

Esta sección se centra en la tarea de BR bilingüe inglés–castellano, con el objetivo de comparar estrategias y resultados de los actuales sistemas de BR con nuestro sistema (BRILI). En la pasada edición del CLEF 2006, tres diferentes estrategias han sido usadas por los sistemas de BR para resolver dicha tarea.

El sistema (Whittaker et al., 2006) usa una herramienta de TA para traducir la pregunta al lenguaje en el que están escritos los documentos. Esta estrategia es la técnica más simple utilizada hasta el momento. De hecho, cuando se compara el resultado del proceso bilingüe con la ejecución monolingüe en Castellano, la precisión decrece en un 55 %. Este sistema pone de manifiesto la influencia negativa de los servicios de TA en la BR multilingüe.

Por otro lado, el sistema (Bowden et al., 2006) traduce el volumen total de documentos a la lengua en la que la pregunta es formulada. Este sistema utiliza un sistema de TA estadístico que ha sido entrenado usando el "*European Parliament Proceedings Parallel Corpus 1996–2003*" (EUROPARL).

Finalmente, el sistema BRUJA (García-Cumbreres et al., 2006) traduce la pregunta utilizando diferentes servicios *on–line* de TA y heurísticas. Esta técnica consulta diferentes herramientas con el objetivo de obtener un traducción aceptable de la pregunta.

Los tres métodos anteriormente descritos están basados en el uso de herramientas de TA para llevar a cabo la tarea bilingüe de BR. Además, algunos de ellos intentan corregir los errores originados por dichas herramientas por medio de diferentes heurísticas.

Asimismo, se debe tener en cuenta que la baja calidad de los sistemas de TA originan un cúmulo de errores dentro de todos las fases de la BR. Estos hechos causan un decremento importante en la precisión global de los sistemas de BR. Esta situación puede ser corroborada en la pasada edición del CLEF 2006 donde la precisión de los sistemas de BR multilingüe decrece alrededor del 50 % en comparación con las tareas monolingües.

3. El sistema BRILI

En esta sección, se describe la arquitectura y funcionalidad de el sistema BRILI. Siendo el objetivo principal de esta sección detallar nuestra estrategia para afrontar la tarea bilingüe de BR inglés–castellano.

3.1. Arquitectura general

El sistema BRILI esta implementado para localizar respuestas en documentos, donde preguntas y documentos están escritos en lenguas diferentes. El sistema se basa en el uso de patrones sintácticos para analizar las preguntas y localizar las repuestas (para realizar estas tareas se utilizan diferentes herramientas de Procesamiento del Lenguaje Natural (Acebo et al., 1994; Ferrández, Palomar, y Moreno, 1999; Llopis y Vicedo, 2001; Schmid, 1995)). Además, un algoritmo de *"Word Sense Disambiguation (WSD)"* es aplicado con el objetivo de mejorar la precisión (una nueva propuesta de WSD para nombre presentada en (Ferrández et al., 2006)).

La figura 1 muestra la arquitectura global del sistema BRILI.

Para la primera implementación del sistema BRILI, nos hemos apoyado en la fase de indexación de documentos y en el módulo de la extraction de la respuesta de nuestro sistema monolingüe de BR llamado AliQAn (Roger et al., 2005), el cual fue presentado por primera vez en la edición del 2005 del CLEF.

En las siguientes sub–secciones se detallarán los módulos de Identificación del Lenguaje, la fase de Análisis de la Pregunta, el módulo de Referencia Entre Lenguas, el módulo de Selección de Pasajes Relevantes y el módulo de la Extracción de la Respuesta.

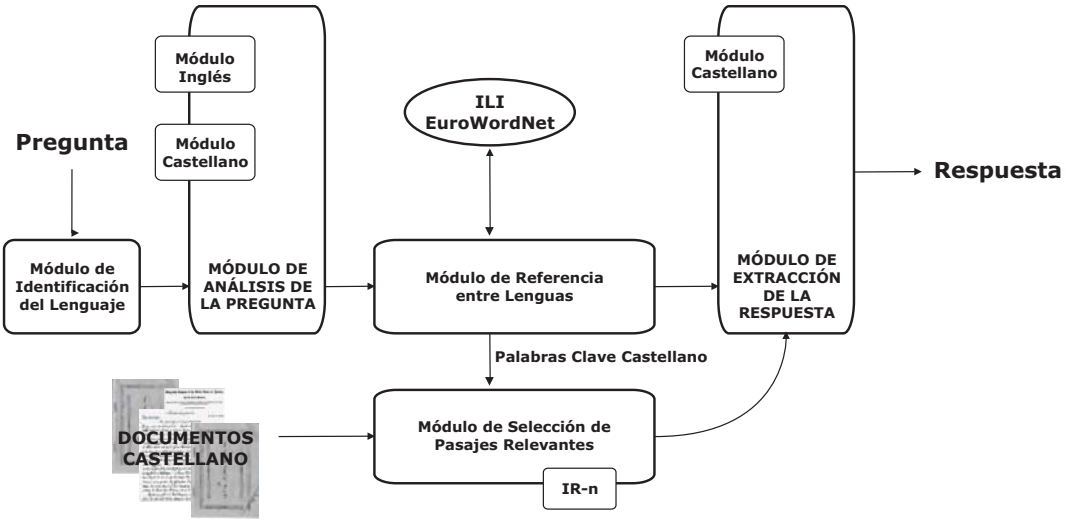

Figura 1: Arquitectura del Sistema BRILI

3.2. Módulo de Identificación del Lenguaje

El módulo de Identificación del Lenguaje ha sido desarrollado con el objetivo de distinguir automáticamente el lenguaje en el que está escrito el texto que recibe como entrada.

Este módulo esté basado en dos técnicas: i) el uso de diccionarios (*"stopwords"* específicas del lenguaje); ii) el uso de partes de las palabras (por ejemplo, *"-ing"* en el caso de inglés). Esta filosofía presenta buenos resultados (Martínez, T., E. Noguera, R. Muñoz y F. Llopis, 2005) para castellano e inglés.

3.3. Módulo de Análisis de la Pregunta

Antes de adentrarnos en la descripción de la fase de Análisis de la Pregunta se deben tener en cuenta algunos aspectos.

Nuestro método se basa fundamentalmente en el análisis sintáctico de la pregunta y los documentos. Para desarrollar esta tarea, el sistema SUPAR (Ferrández, Palomar, y Moreno, 1999) ha sido utilizado. A partir de la salida que SUPAR proporciona, el sistema BRILI extrae los Bloques Sintácticos (BS) de las preguntas y los documentos. Se detectan tres tipos de BS: Sintagma Nominal (SN) simple, Sintagma Verbal (SV) y Sintagma Preposicional (SP). Estos BS forman la unidad sintáctica básica para diseñar e implementar patrones.

El sistema BRILI desarrolla dos tareas principales en la fase de Análisis de la pregunta:

- **Detectar el tipo de respuesta esperada.** El sistema detecta el tipo de información que las posibles respuestas deben satisfacer para llegar a ser respuestas candidatas (por ejemplo: nombres propios, cantidades, fechas, etcétera)
- **Identificar los principales BS de la pregunta.** BRILI extrae los BS del a pregunta que son útiles para encontrar las respuestas en los documentos.

Para la primera tarea, se ha diseñado una taxonomía basada en *"WordNet Based–Types"* y *"EuroWordNet Top–Concepts"*. Nuestra clasificación se compone de las siguientes categorías: persona, profesión, grupo, objeto, lugar–país, lugar–ciudad, lugar-capital, lugar, abreviatura, evento, numérico–económico, numérico–edad, numérico–medida, numérico–periodo, numérico–porcentaje, numérico–cantidad, temporal–año, temporal-mes, temporal–fecha, efeméride y definición.

El tipo de respuesta esperada es detectado por medio de un conjunto de patrones sintácticos. El sistema posee alrededor de 200 patrones para determinar la categoría de la pregunta dentro de nuestra taxonomía. El procedimiento es el siguiente: BRILI compara los BS de los patrones con los BS de la pregunta, el resultado de estas comparaciones determinan el tipo.

A continuación se muestra un ejemplo que detalla el comportamiento de esta tarea:

- **Pregunta 06 del CLEF 2006:** *Which country did Iraq invade in 1990?*

- **Patrón:** [WHICH] [sinónimo de COUNTRY]
- **Tipo de Respuesta Esperada** lugar–país

El objetivo que la segunda tarea persigue es determinar qué BS de la pregunta son útiles para localizar las respuestas en los documentos. Además, en estos BS se encuentran las palabras clave que deben ser traducidas por el Módulo de Referencia Entre Lenguas. El ejemplo siguiente muestra el proceso de esta tarea en una pregunta de tipo temporal–año.

- **Pregunta 75 del CLEF 2006:** *In which year was the Football World Cup celebrated in the United States?*
- **SBs:** [SN Football World Cup] + [SV to celebrate] + [SP in [SN United States]]
- **Palabras Clave:** Football World Cup celebrate United States

Además, BRILI, con el objetivo de mejorar la búsqueda de las respuestas, descarta algunas palabras de las preguntas. Por ejemplo, en la pregunta siguiente las palabras *"can"* y *"be"* no son utilizadas en la localización de las posibles respuestas.

- **Pregunta 33 del CLEF 2004:** *How can an allergy be treated?*
- **SBs:** [SV to treat] + [SN allergy]
- **Palabras Clave:** treat allergy

3.4. Módulo de Referencia Entre Lenguas

En está sub–sección se describe el módulo de Referencia Entre Lenguas (REL) el cual se encarga de traducir las palabras clave de las pregunta a la lengua en la que se va a realizar la búsqueda de posibles respuestas.

El módulo REL estable enlaces de referencia entre las lenguas que forman la BR bilingüe. Dicho módulo recibe como entrada las palabras clave detectadas en la fase de Análisis de la Pregunta. Con el objetivo de ilustrar el proceso, veamos el siguiente ejemplo (Figura 2) en donde se muestran las referencias establecidas por el módulo REL para la palabra clave "*president*".

Cómo se aprecia en el ejemplo, el módulo LER encuentra más de una referencia en Castellano para la palabra inglesa "*president*".

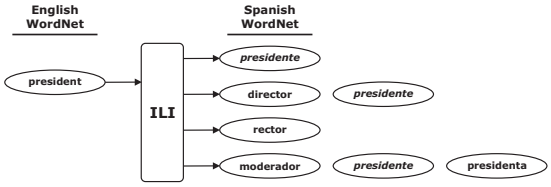

Figura 2: Referencias para la palabra "president"

La estrategia empleada para poder realizar la BR haciendo uso de toda la información que ILI proporciona al sistema, es valorar y ordenar todas las referencias establecidas por frecuencia. En el ejemplo anterior, la palabra en Castellano más valorada sería "*presidente*".

Por otro lado, se debe tener en cuenta que en algunos casos el sistema se encuentra ante palabras inglesas que no están referenciadas al Castellano por el módulo ILI. En estos casos el sistema utiliza un diccionario *on–line*[1] para poder determinar las traducciones. Además, el sistema BRILI se apoya en el uso de colecciones bilingües de organizaciones y lugares para traducir este tipo de entidades que en algunos casos no están referenciadas por ILI.

A su vez, con el objetivo de reducir el efecto negativo de traducciones incorrectas de algunos nombre propios, la fase de Extracción de la Respuesta utiliza las palabras clave traducidas por el módulo REL y dichas palabras en el lenguaje original de la pregunta (teniendo en cuenta que las soluciones obtenidas a partir de dichas palabras son valoradas un 20 % menos).

La estrategia establecida en el módulo REL de no realizar una TA de la pregunta, que por otro lado es el objetivo real que persigue el sistema BRILI, se lleva a cabo gracias al análisis de la pregunta en su lenguaje original y el uso del módulo ILI de EuroWordNet (Vossen, 1998).

A continuación se muestra un ejemplo del proceso realizado por el módulo REL:

- **Pregunta 101 del CLEF 2004:** *What army occupied Haiti?*
- **BS:** [SN army] + [SV to occupy] + [SN Haiti]
- **Tipo de Respuesta Esperada** grupo
- **Palabras Clave:** army occupy Haiti

[1] http://www.wordreference.com

- **Referencias :**
  - army ↦ ejército
  - occupy ↦ absorber ocupar atraer residir vivir colmar rellenar ocupar llenar
  - Haiti ↦ Haití
- **Principales BS utilizados en la BR:** [SN ejército] + [SV ocupar (verbo más valorado)] + [SN Haití]

3.5. Módulo de Selección de pasajes Relevantes

El módulo de Selección de pasajes Relevantes se basa en el uso de la herramienta de RI, IR–n (Llopis y Vicedo, 2001).

IR–n recibe como entrada las palabras clave traducidas por el módulo REL, por ejemplo, en la pregunta 92 del CLEF 2006, "*Where is Interpol's headquarters?*", IR-n recibe las palabras *"Interpol"* y *"comandancia"*.

El sistema IR–n proporciona una lista de pasajes relevantes en donde el sistema aplicará la búsqueda de respuestas candidatas.

3.6. Módulo de Extracción de la Respuesta

La fase final del sistema BRILI se realiza en el módulo Extracción de la Respuesta. Dicho módulo usa los BS de la pregunta y diferentes conjuntos de patrones sintácticos (según el tipo de respuesta espera) con información léxica, sintáctica y semántica para localizar respuestas candidatas.

En el ejemplo siguiente se muestra el proceso completo de BR donde la aplicación de un patron de localización de respuestas es detallada.

- **Pregunta 70 del CLEF 2005:** *Which French president inaugurated the Eurotunnel?*
- **SB:** [SN French president] + [SV to inaugurate]+ [SN Eurotunnel]
- **Tipo de Respuesta Esperada** persona
- **Principales BS utilizados en la BR:** [SN presidente francés] + [SV inaugurar]+ [SN Eurotunel]
- **Frase de los documentos:** ... el Presidente Francés, *Jacques Chirac*, inauguró el Eurotunel ...
- **Patrón sintáctico**: [SN (*"Presidente Francés"*), aposición [SN (SOLUCIÓN)]] + [SV (*"inaugurar"*)] + [SN *"Eurotunel"*)]
- **Respuesta Candidata**: Jacques Chirac

4. Resultados Experimentales

4.1. Entorno de Evaluación

Para llevar a cabo la evaluación del sistema que en este artículo se presenta, se han utilizado los conjuntos de 200 preguntas de las tareas de BR de las ediciones del CLEF 2004, 2005 y 2006 (en inglés y en castellano) y el conjunto de documentos en Castellano ("*EFE 1994–1995 Spanish corpora*") propuesto por el CLEF para realizar las búsquedas de las soluciones. El corpus propuesto proporciona un marco de trabajo adecuado para la evaluación de los sistemas de BR.

El conjunto de preguntas está compuesto principalmente por *"factoid questions"* y "*definition questions*". Las "*factoid questions*"son preguntas basadas en hechos, solicitan el nombre de una persona, una localidad, el día que ocurrió algo, etcétera.

4.2. Análisis de Resultados

El objetivo de estos experimentos es evaluar y discutir la bondad de la estrategia de BR bilingüe diseñada e implementada en BRILI. Para realizar este análisis se proporcionan la precisión del sistema en su ejecución bilingüe inglés–castellano y monolingüe castellano. Más delante, se comparan los resultados con los actuales sistemas de BR bilingüe presentados de la pasada edición del CLEF 2006.

En el cuadro 4 se muestra la precisión[2] del sistema BRILI (ver filas 2,4 y 6). A su vez, estos resultados se comparan con la ejecución monolingüe en castellano (ver filas 1, 3 y 5) y con las presiciones obtenidas por los sistemas que en la actualidad forman el estado de la cuestión de la BR bilingüe (ver filas 7, 8 y 9).

Cuando se compara la precisión obtenida en las tareas de BR bilingüe con las ejecuciones monolingües en Castellano, se aprecia que el decremento sufrido en la precisión global del sistema BRILI está al rededor

[2]Para calcular la precisión de nuestros sistemas se han considerado las respuestas correctas y las respuestas inexactas cuando éstas contienen más información que la requerida por la pregunta.

	Sistema	Conjunto de Datos	Precisión (%)
		CLEF 2004	
1	AliQAn	200 Preguntas castellano	41.5
2	BRILI	200 Preguntas inglés	33.5
		CLEF 2005	
3	AliQAn	200 Preguntas castellano	51.5
4	BRILI	200 Preguntas inglés	42.5
		CLEF 2006	
5	AliQAn	200 Preguntas castellano	50.5
6	BRILI	200 Preguntas inglés	44
	Participantes en	**CLEF 2006** (Magnini et al., 2006)	
7	(Whittaker et al., 2006)	200 Preguntas inglés	6
8	(Bowden et al., 2006)	200 Preguntas inglés	19
9	(García-Cumbreres et al., 2006)	200 Preguntas inglés	19.5

Cuadro 1: Evaluación

del 17 %. Este dato revela y justifica la calidad del método implementado en BRILI, ya que, como se mostró en la pasada edición del CLEF (Magnini et al., 2006), la precisión de las tareas bilingües inglés–castellano se establecido en alrededor de un 50 % menos que la precisión obtenida por los sistemas de BR monolingüe Castellana.

Además, comparando con el resto de sistemas que en la actualidad desarrollan la tarea bilingüe que en esta sección se discute, y que han participado en la pasada edición del CLEF 2006, el sistema BRILI obtiene mejores resultados (BRILI ha quedado en la primera posición en la tarea de BR bilingüe inglés–castellano en el CLEF 2006).

5. Conclusiones y Trabajos Futuros

En este articulo, la estrategia de BR bilingüe inglés–castellano diseñada e implementada dentro del sistema BRILI ha sido detallada y evaluada.

El sistema BRILI reduce el uso de la TA con el objetivo de evitar el efecto negativo que causan este tipo de estrategias en la precisión global de los sistemas de BR. En el sistema BRILI se introducen dos mejoras que intentan alcanzar dicho objetivo: i) El sistema considera más de una traducción por palabra gracias a los diferentes synsets de cada palabra que están referenciados por el módulo ILI de EuroWordNet; ii) Al contrario que los actuales sistemas de BR, el análisis de la pregunta se realiza en su lenguaje original, sin ningún tipo de traducción.

Los experimentos realizados sobre los conjuntos de preguntas oficiales de las ediciones del CLEF 2004, 2005 y 2006 justifican nuestra estrategia.

En la actualidad se está diseñando e implementando la integración de un reconocedor de entidades con el objetivo de detectar y clasificar entidades para así poder establecer las necesidades de traducción de las mismas. Por ejemplo, en la pregunta 059 del CLEF 2006, *What is Deep Blue?*, la entidad "*Deep Blue*" no debería ser traducida.

A su vez, se está incorporando conocimiento multilingüe de Wikipedia[3] con el objetivo de traducir entidades nombradas y ampliar las colecciones bilingües de organizaciones y lugares.

Bibliografía

Acebo, S., A. Ageno, S. Climent, J. Farreres, L. Padró, R. Placer, H. Rodriguez, M. Taulé, y J. Turno. 1994. MACO: Morphological Analyzer Corpus-Oriented. *ESPRIT BRA-7315 Aquilex II, Working Paper 31.*

Bowden, M., M. Olteanu, P. Suriyentrakorn, J. Clark, y D. Moldovan. 2006. LCC's PowerAnswer at QA@CLEF 2006. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Ferrández, A., M. Palomar, y L. Moreno. 1999. An Empirical Approach to Spanish

[3] http://www.wikipedia.org/

Anaphora Resolution. *Machine Translation. Special Issue on Anaphora Resolution In Machine Translation*, 14(3/4):191–216.

Ferrández, S. y A. Ferrández. 2006. Cross-lingual question answering using inter lingual index module of eurowordnet. *Advances in Natural Language Processing. Research in Computing Science. ISSN: 1665-9899*, 18:177–182, February.

Ferrández, S. y A. Ferrández. 2007. The negative effect of machine translation on cross–lingual question answering. *Proceedings of the CICLing 2007, LNCS 4394*, páginas 494–505, February.

Ferrández, S., S. Roger, A. Ferrández, A. Aguilar, y P. López-Moreno. 2006. A new proposal of Word Sense Disambiguation for nouns on a Question Answering System. *Advances in Natural Language Processing. Research in Computing Science. ISSN: 1665-9899*, 18:83–92, February.

García-Cumbreres, M.A., L.A. Ureña-López, F. Martínez-Santiago, y J.M. Perea-Ortega. 2006. BRUJA System. The University of Jaén at the Spanish task of CLEFQA 2006. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Gillard, L., L. Sitbon, E. Blaudez, P. Bellot, y M. El-Béze. 2006. The LIA at QA@CLEF-2006. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Llopis, F. y J.L. Vicedo. 2001. Ir-n, a passage retrieval system. *In Workshop of Cross-Language Evaluation Forum (CLEF).*

Magnini, B., D. Giampiccolo, P. Forner, C. Ayache, V. Jijkoun, P. Osevona, A. Peñas, , P. Rocha, B. Sacaleanu, y R. Sutcliffe. 2006. Overview of the CLEF 2006 Multilingual Question Answering Track. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Martínez, T., E. Noguera, R. Muñoz y F. Llopis. 2005. Web track for CLEF2005 at ALICANTE UNIVERSITY. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Moldovan, D.I., M. Pasca, S.M. Harabagiu, y M. Surdeanu. 2003. Performance issues and error analysis in an open-domain question answering system. *ACM Trans. Inf. Syst*, 21:133–154.

Roger, S., S. Ferrández, A. Ferrández, J. Peral, F. Llopis, A. Aguilar, y D. Tomás. 2005. AliQAn, Spanish QA System at CLEF-2005. *In Workshop of Cross-Language Evaluation Forum (CLEF).*

Sacaleanu, B. y G. Neumann. 2006. Dfki-lt at the clef 2006 multiple language question answering track. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Schmid, H. 1995. TreeTagger — a language independent part-of-speech tagger. *Institut fur Maschinelle Sprachverarbeitung, Universitat Stuttgart.*

Sutcliffe, R.F.E., K. White, D. Slattery, I. Gabbay, y M. Mulcanhy. 2006. Cross-language French-English Question Answering using the DLT System at CLEF 2006. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Tomás, D., J.L Vicedo, E. Bisbal, y L. Moreno. 2006. Experiments with LSA for Passage Re-Ranking in Quesion Answering. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

Vossen, P. 1998. Introduction to eurowordnet. *Computers and the Humanities*, 32:73–89.

Whittaker, E.W.D., J.R. Novak, P. Chatain, P.R. Dixon, M.H. Heie, y S. Furui. 2006. CLEF2005 Question Answering Experiments at Tokyo Institute of Technology. *In Workshop of Cross-Language Evaluation Forum (CLEF)*, September.

.