

PROGRAMA JUAN DE LA CIERVA

Curriculum vitae Impreso normalizado

Número de hojas que contiene: 10

Nombre: Gemma Boleda Torrent

Fecha: 13/03/07

- Este curriculum no excluye que durante el proceso de evaluación se le requiera para ampliar y justificar la información aquí contenida.

DATOS PERSONALES

Apellidos: Boleda Torrent

Nombre: Gemma

DN/Pasaporte: 44016678 Fecha de nacimiento: 29/03/76

Sexo: femenino

Nacionalidad: Española

SITUACIÓN PROFESIONAL ACTUAL

Organismo: Universidad Pompeu Fabra

Facultad, Escuela o Instituto: Facultad de Traducción e Interpretación

Depto./Unidad: Departamento de Traducción y Filología

Dirección postal: La Rambla 30-32, 08002 Barcelona

País: España

Teléfono (indicar prefijo, número y extensión): (93) 542-29-40

Fax: (93) 542-16-17

Correo electrónico: gemma.boleda@upf.edu

Especialización (Códigos UNESCO): 570104

Categoría profesional: Becaria de investigación

Fecha de inicio: 01/01/2007

Situación administrativa

☐ Plantilla

☐ Contratado

☐ Interino

☐ Becario

☐ Otras situaciones especificar:

Dedicación A tiempo completo ☐

A tiempo parcial ☐

LÍNEAS DE INVESTIGACIÓN

Breve descripción, por medio de palabras claves, de la especialización y líneas de investigación actuales.

Lingüística computacional, procesamiento del lenguaje natural, aprendizaje automático, semántica léxica, lingüística sincrónica

FORMACIÓN ACADÉMICA

Titulación Superior	Centro	Fecha
Filología Española	Universidad Autónoma de Barcelona	2000

Doctorado	Centro	Director/a tesis	Fecha

ACTIVIDADES ANTERIORES DE CARÁCTER CIENTÍFICO (*)

<i>Puesto</i>	<i>Institución</i>	<i>Fechas</i>
Contrato de investigación (proyectos e-Title y METIS-II)	Fundación Barcelona Media	Ene 2005 - ene 2006
Asistente en proyecto de Innovación docente "Creación de una base de datos lingüística para la docencia"	Depto. Traducción y Filología, Universidad Pompeu Fabra	1999-2000
Asistente de investigación	Depto. Procesamiento de Información Lingüística, Universidad Colonia, Alemania	1997-1998

(*) La información contenida en el cuadro anterior se utilizará para acreditar el cumplimiento de los requisitos establecidos en el apartado vigésimo cuarto.3 de la Resolución de convocatoria. El órgano competente para la instrucción puede solicitar al candidato la verificación documental de lo declarado con anterioridad en cualquier momento de la tramitación de su expediente. (**) Si el Organismo es un centro mixto deberá indicarse tal situación con mención expresa de todos los centros que participan en su gestión.

IDIOMAS (R = REGULAR, B = BIEN, C = CORRECTAMENTE)

<i>Idioma</i>	<i>Habla</i>	<i>Lee</i>	<i>Escribe</i>
inglés	C	C	C
castellano	C	C	C
catalán	C	C	C
alemán	C	C	C
francés	B	C	B

PARTICIPACIÓN EN PROYECTOS DE INVESTIGACIÓN

Título del proyecto: Acción Integrada trilateral "Natural language ontology for reference to facts and eventualities"

Entidad financiadora: MEC, HA2005-0100 y HF2005-0177

Entidades participantes: Universidad Pompeu Fabra, Universidad Lille 3, Universidad de Osnabrück

Duración, desde: 2005 **hasta:** 2006

Cuantía de la subvención: 10831 y 11080 € respectivamente

Investigador responsable: L. McNally, G. Katz, P. Miller

Número de investigadores participantes: 19

Título del proyecto: METIS-II: Statistical Machine Translation using Monolingual Corpora: from Concept to Implementation (IST – FP6-003768)

Entidad financiadora: Unión Europea - programa IST

Duración desde: 2004 **hasta:** 2007

Investigadores principales en la UPF: Toni Badia

Coordinador General: ILSP (Atenas, Grecia)

Otros participantes: IAI (Saarbrücken, Alemania), Leuven University (Bélgica)

Título del proyecto: Arquitectura integrada para el tratamiento avanzado de textos (ARQUITEXT)

Entidad financiadora: MEC, HUM2004-05321-C02-02

Duración desde: 2005 **hasta:** 2008

Cuantía de la subvención: 25.920€

Investigador responsable: Toni Badia Cardús

Número de investigadores participantes: 10

Título del proyecto: Construcciones no canónicas en el discurso oral: estudio transversal y comparativo (NOCANDO)

Entidad financiadora: MEC, HUM2004-04463

Duración desde: 2005 **hasta:** 2008

Cuantía de la subvención: 21.800€

Investigador principal: Enric Vallduví Botet

Número de investigadores participantes: 10

Título del proyecto: Sistema de preparación automatizada de documentos (PrADo)

Entidad financiadora: MCYT-DGI, TIC2000-1681-C02-01

Duración desde: 2001 **hasta:** 2004

Cuantía de la subvención: 36.000€

Investigador principal: Toni Badia, UPF; Sergio Balari Ravera, subproyecto en la UAB

Número de investigadores participantes: 12

Título del proyecto: Integración de técnicas y herramientas de tagging, parsing y unificación (PB98-1151)

Entidad financiadora: DGES - MEC

Duración desde: 1999 **hasta:** 2002

Investigador principal: Gabriel Amores (Universidad de Sevilla)

PUBLICACIONES

Indicar volumen, páginas inicial y final (año) y clave.

CLAVE: L= libro completo, CL.= capítulo de libro, A= artículo, R= revisión/"review", E= editor/a

(*) En el caso de aquellas publicaciones que estén en tramitación y aún no hayan sido publicadas, indicar únicamente la situación en la que se encuentra la publicación. (**) Con carácter opcional, se podrán indicar los aspectos que considere más destacados de cada publicación para evaluar su calidad (p.ej. el índice de impacto de la revista, posición de la revista en los listados de los campos correspondientes, citas recibidas u otros indicadores de repercusión).

Capítulos de libro

Louise McNally y Gemma Boleda. 2004. Relational nouns as properties of kinds. En *Empirical Issues in Formal Syntax and Semantics 5*, ed. O. Bonami y P. Cabredo Hofherr, publicado electrónicamente en <http://www.cssp.cnrs.fr/eiss5>, 179-196.

Artículos

- G. Boleda, S. Schulte im Walde, T. Badia. Modelling Polysemy with Multi-label Classification. En proceso de revisión para *ACL 2007*.
- G. Boleda, S. Schulte im Walde, T. Badia. Analysis of Agreement about the Semantic Class of Adjectives. Aceptado para su publicación en *Research on Language and Computation: special issue on ambiguity and semantic judgments*, editado por Massimo Poesio and Ron Artstein.
- T. Badia, G. Boleda, M. Melero, A. Oliver (en prensa) El proyecto METIS-II. *Revista de Procesamiento del Lenguaje Natural*. ISSN: 1135-5948.
- A. Oliver, T. Badia, G. Boleda, M. Melero (en prensa) Traducción automática estadística basada en n-gramas. *Revista de Procesamiento del Lenguaje Natural*. ISSN: 1135-5948.
- G. Boleda, S. Bott, C. Castillo, R. Meza, T. Badia, V. López. 2006. CUCWeb: a Catalan corpus built from the Web. En *Proceedings of the Second Workshop on the Web as a Corpus at EACL'06*. Trento, Italia.
- T. Badia, G. Boleda, M. Melero, A. Oliver. 2005. An n-gram approach to exploiting a monolingual corpus for Machine Translation. En *Proceedings of the Second Workshop on Example-based Machine Translation*, MT Summit X, Phuket, Tailandia.
- Boleda, G., T. Badia, S. Schulte im Walde. 2005. Morphology vs. Syntax in Adjective Class Acquisition. En *Proceedings of ACL-SIGLEX 2005 Workshop on Deep Lexical Acquisition*, Ann Arbor, EEUU.
- Mayol, L., G. Boleda, T. Badia. 2005. Automatic acquisition of syntactic verb classes with basic resources. *Language Resources and Evaluation*, 39(4):295-312.
- Boleda, Gemma, Toni Badia y Eloi Batlle. 2004. Acquisition of Semantic Classes for Adjectives from Distributional Evidence. En *Proceedings of 20th International Conference on Computational Linguistics (COLING 2004)*, 1119-1125, Ginebra, Suiza. ISBN:1-932432-48-5
- Padó, Sebastian y Gemma Boleda. 2004. The Influence of Argument Structure on Semantic Role Assignment. En *Proceedings of 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-04)*, Barcelona.
- Boleda, Gemma y Laura Alonso. 2003. Clustering Adjectives for Class Acquisition. En *Proceedings of the 10th Conference of The European Chapter of the Association for Computational Linguistics (EACL 2003) Student Research Workshop*, pages 9-16, Budapest, Hungría. ISBN: 1-932432-01-9
- Alsina, Àlex, Toni Badia, Gemma Boleda, Stefan Bott, Àngel Gil, Martí Quixal, Oriol Valentín. 2002. CATCG: a general purpose parsing tool applied. En *Proceedings of Third International Conference on Language Resources and Evaluation (LREC 2002)*, Las Palmas, Vol. III, 1130-1135. ISBN: 2-9517408-0-8
- Alsina, Àlex, Toni Badia, Gemma Boleda, Stefan Bott, Àngel Gil, Martí Quixal, Oriol Valentín. 2002. CATCG: un sistema de anàlisi morfosintàctic per al català. *Revista de Procesamiento del Lenguaje Natural*, 29, Sept. 2002, 309-310. ISSN: 1135-5948

- Badia, Toni, Gemma Boleda, Jenny Brumme, Carme Colominas, Mireia Garmendia, Martí Quixal. 2002. BancTrad: un banco de corpus anotados con interfície web. *Revista de Procesamiento del Lenguaje Natural*, 29, Sept. 2002, 293-294. ISSN: 1135-5948
- Badia, Toni, Gemma Boleda, Carme Colominas, Mireia Garmendia, Agnès González, Martí Quixal. 2002. BancTrad: a web interface for integrated access to parallel annotated corpora. En *Proceedings of the First International Workshop On Language Resources For Translation Work And Research, 3rd LREC Conference (LREC 2002)*, Las Palmas.
- Badia, Toni, Gemma Boleda, Carme Colominas, M. Garmendia, Agnès González, Martí Quixal. 2002. Eines de lingüística computacional per a la traducció: corpus paral·lels anotats. En *Proceedings of the 2nd International Conference on Specialized Translation*, Barcelona, Spain, 129-137. ISBN: 84-477-0820-9
- Badia, Toni, Gemma Boleda, Martí Quixal. 2001. Curso sobre Tecnologías de la lengua (segunda edición). *QUARK, Ciencia, Medicina, Comunicación y Cultura*, 21, Jul - Sept. 2001, 14-16. ISSN: 1135-8521
- Badia, Toni, Gemma Boleda, Martí Quixal, Eva Bofias. 2001. A modular architecture for the processing of free text. En *Proceedings of the Workshop on Modular Programming applied to Natural Language Processing at EUROLAN 2001*, Iasi, Romania, August 2001, 11-18

Informes tècnics

- Aguilar, Lourdes, Àlex Alsina, Anna-Belén Avilés, Toni Badia, Sergio Balari, Gemma Boleda, Stefan Bott, Jenny Brumme, Carme Colominas, Anna Espunya, Josep Fontana, Jordi Fontseca, Àngel Gil, Carmen Hernández, Laia Mayol, Louise McNally, Carme de la Mota, Martí Quixal, Yolanda Rodríguez, Oriol Valentín, Enric Vallduví, Teresa Vallverdú. 2004. PrADo: Preparación Automatizada de Documentos (TIC 2000-1681). Informe técnico.

ESTANCIAS EN CENTROS DE INVESTIGACIÓN

CLAVE: D=doctorado, P=postdoctoral. Y= invitado/a, C=contratado/a, O=otras (especificar)

Centro: Depto. Procesamiento de Información Lingüística (*Sprachliche Informationsverarbeitung*),
Universidad de Colonia

Localidad: Colonia **País:** Alemania **Fecha:** 1997-1998

Duración (semanas): 35

Tema: Traducción automática

Clave: O (estudios de licenciatura y contrato como asistente de investigación)

Centro: Instituto de Fonética y Lingüística Computacional (CoLi), Universidad del Sarre.

Localidad: Saarbrücken **País:** Alemania **Fecha:** 2003

Duración (semanas): 12

Tema: Semántica Léxica Computacional

Clave: Y

Centro: Instituto de Fonética y Lingüística Computacional (CoLi), Universidad del Sarre.

Localidad: Saarbrücken **País:** Alemania **Fecha:** 2004

Duración (semanas): 7

Tema: Semántica Léxica Computacional

Clave: Y

CONGRESOS

- Boleda, G., S. Bott, R. Meza, C. Castillo, T. Badia, V. López. 2005. Usant la web per estudiar el català. *III Jornades sobre el català a les noves tecnologies*, 14-16 Abril, Barcelona.
- Mayol, L., G. Boleda, T. Badia. 2005. Automatic learning of syntactic verb classes. *Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes*, 28 feb-1 marzo, Saarbrücken, Alemania.
- Boleda, G., S. Bott, B. Poblete, C. Castillo, M.E. Fuenmayor, T. Badia, V. López. 2004. CuCWeb: un corpus del català construït a partir de la web. *II Congrés Online de l'Observatori per a la Cibersocietat*, Barcelona.
- Colominas, Carme y Gemma Boleda. 2004. The extraction of translationally relevant information from small ad-hoc corpora. *Third International Conference on Corpus Use and Learning to Translate (CULT-BCN)*, Barcelona.
- McNally, Louise y Gemma Boleda. 2003. Relational Adjectives as Properties of Kinds. *CSSP 2003 (The Fifth Syntax and Semantics Conference in Paris)*, 2-4 Octubre, París, Francia.
- Padó, Sebastian y Gemma Boleda. 2003. Towards Better Understanding of Automatic Semantic Role Assignment. *Prospects and Advances in the Syntax/Semantics Interface*, Nancy, Francia.
- Alonso, Laura and Gemma Boleda. 2002. An Approach to Catalan Adjective Lexical Classes by Clustering. *Workshop on Quantitative Investigations for Theoretical Linguistics*, Osnabrück, Alemania, 3-5 Octubre 2002

INVESTIGADORES DE REFERENCIA

(Se podrán aportar hasta 3 nombres de investigadores que puedan dar referencias del solicitante.
Este apartado no es de obligado cumplimiento para el solicitante)

APELLIDOS, NOMBRE: Badia Cardús, Toni

RELACIÓN PROFESIONAL CON EL SOLICITANTE: Director de investigación
(Profesor-tutor, Director de investigación, Empleador-jefe, Colaborador, Otros)

DURACIÓN DE LA RELACIÓN (indicar fechas): 2001-2007

ORGANISMO AL QUE PERTENECE EN LA ACTUALIDAD: Universidad Pompeu Fabra, España
(Institución y país)

CARGO: Profesor titular, director del grupo de investigación GLiCom

TELÉFONO: 93-542-24-14 CORREO ELECTRÓNICO: toni.badia@upf.edu

APELLIDOS, NOMBRE: Schulte im Walde, Sabine

RELACIÓN PROFESIONAL CON EL SOLICITANTE: Directora de investigación
(Profesor-tutor, Director de investigación, Empleador-jefe, Colaborador, Otros)

DURACIÓN DE LA RELACIÓN (indicar fechas): 2004-2007

ORGANISMO AL QUE PERTENECE EN LA ACTUALIDAD: Universidad de Stuttgart, Alemania
(Institución y país)

CARGO: Investigadora

TELÉFONO: +49-711-685-84584 CORREO ELECTRÓNICO: schulte@ims.uni-stuttgart.de

APELLIDOS, NOMBRE: McNally Seifert, Louise

RELACIÓN PROFESIONAL CON EL SOLICITANTE: Colaborador
(Profesor-tutor, Director de investigación, Empleador-jefe, Colaborador, Otros)

DURACIÓN DE LA RELACIÓN (indicar fechas): 2001-2007

ORGANISMO AL QUE PERTENECE EN LA ACTUALIDAD: Universidad Pompeu Fabra, España
(Institución y país)

CARGO: Profesora titular

TELÉFONO: 93-542-22-48 CORREO ELECTRÓNICO: louise.mcnally@upf.edu

OTROS MÉRITOS O ACLARACIONES QUE SE DESEE HACER CONSTAR
(utilice únicamente el espacio de una página DIN A4)

Premios y Becas

- 2005-2006: Beca pre-doctoral de la Fundación Caja Madrid (2005-2006)
- 2001-2004: Beca FPI (*Formació de Personal Investigador*; ref. 2001FPI 00582) de la Generalitat de Catalunya
- 2001: Mención especial en los Premios Nacionales de Fin de Carrera 1999-2001 otorgados por el Ministerio de Educación, Cultura y Deportes
- 2001: Premio extraordinario de licenciatura de Filología Hispánica de la Universitat Autònoma de Barcelona del curso 1999-2000
- Sept-dic 2000: Beca de Introducción a la Investigación del Consejo Superior de Investigaciones Científicas. *Institut d'Investigació en Intel·ligència Artificial* (IIIA, CSIC).
- 1997-1998: Beca del programa Sócrates-Erasmus para la estancia de un curso académico en la Universidad de Colonia (Alemania).
- 1997: Beca del *Deutscher Akademischer Austauschdienst* (Servicio Alemán de Intercambios Académicos) para un curso de lengua en la *Universität Gesamthochschule Essen*.
- 1994: Beca del Ministerio de Educación, Cultura y Deportes para la matrícula del primer curso de licenciatura por Matrícula de Honor de C.O.U.

Actuaciones de referee

Referee para los congresos *Student Research Workshop at EACL 2006* (Trento, Italia), *Penn Linguistics Colloquium* (2005 y 2006), *Student Research Workshop at ACL 2007* (Praga, República Checa).

Conferencias invitadas

- 11 Noviembre 2004: Acquisition of Semantic Classes for Adjectives. Conferencia en el *Colloquium of the International Post-Graduate College on Language Technology and Cognitive Systems*, Universität des Saarlandes, Saarbrücken, Alemania.
- 25 Noviembre 2004: A Quantitative Approach to the Lexical Semantics of Adjectives. *Computational Linguistics Colloquium*, Universität des Saarlandes, Saarbrücken, Alemania.
- 10 Diciembre 2004: Acquiring Semantics Classes for Adjectives through Clustering. *Computational Linguistics Seminar*, King's College, Londres, Gran Bretaña.
- 28 Junio 2005: Adquisició de classes semàntiques adjectivals. *III Workshop del Programa de Doctorado en Ciencia Cognitiva y lenguaje: "Acquisition"*, Barcelona.

Filiaciones profesionales

Association for Computational Linguistics