

Pascal Denis

Equipe MOSTRARE, INRIA NLE – LIFL
Parc scientifique de la Haute Borne
Bât. A Park Plaza 40, avenue Halley
F-59650 Villeneuve d'Ascq (FRANCE)

Phone: +33 (0)3 59 57 79 96
Fax: +33 (0)3 28 77 85 37
Email: pascal.denis@inria.fr
Url: <http://alpage.inria.fr/~pdenis>

EMPLOYMENT

Research Scientist, INRIA, France **2008-Present**

Since February 2012, I am affiliated with the MOSTRARE research group, a joint team between INRIA and University of Lille Computer Science department (LIFL) dedicated to research in information extraction and machine learning. Before that, I was a member of the ALPAGE group, an INRIA – University of Paris 7 team.

Academic Visitor, University of Oxford, UK **Fall 2010**

Temporary member of the Comlab's Computational Linguistics Group.

Natural Language Scientist, Powerset Inc. (now Microsoft), USA **2007-2008**

Member of the Knowledge Resources team, in charge of the development of highly scalable software components for named-entity recognition and tracking, and of their integration in the Powerset Web search engine.

Visiting Research Fellow, IMS, Germany **2005**

Five month research stay at the Institut für Maschinelle Sprachverarbeitung (IMS), Universität Stuttgart, funded by German Academic Exchange Service (DAAD).

Research/Teaching Assistant, The University of Texas at Austin, USA **2000-2007**

Research and teaching in theoretical and computational linguistics.

EDUCATION

Ph.D., Linguistics, The University of Texas at Austin, USA **2007**

Dissertation: *New Learning Models for Robust Reference Resolution*. Advisors: Jason Baldridge and Nicholas Asher. Committee: Laurence Danlos, Andrew Kehler, Jonas Kuhn, Raymond Mooney.

M.A., Artificial Intelligence, IRIT, France **2000**

Thesis: *L'interprétation des fragments phrastiques dans le dialogue : le cas des dialogues temporels*. Advisors: Myriam Bras and Philippe Muller.

M.Sc., Cognitive Science, University of Edinburgh, UK **1999**

Thesis: *Fragments in Dialogue*. Advisor: Dr Alex Lascarides.

B.A., Romance Philology, Université Libre de Bruxelles, Belgium **1998**

Thesis: *Les expressions idiomatiques. De la Syntaxe à la Cognition*. Advisor: Marc Dominicy. *Magna cum laude*.

HONORS AND AWARDS

Best Student Paper Award, NAACL-HLT 2007 **2007**

Award given to the author of the best paper written solely or primarily by a student appearing in the annual NAACL conference proceedings.

Homer Lindsey Bruce Fellowship, The University of Texas at Austin **2005**

Nine-month fellowship awarded in recognition of major achievements since entering Graduate School. Awarded only to the strongest applicants for the University of Texas Continuing University Fellowships.

DAAD Young Researcher Fellowship, German Academic Exchange Service (DAAD) 2005
Five-month scholarship awarded to outstanding graduate and post-doctoral students for conducting study and/or research at a German university.

Nomination for the William S. Livingston Award, The University of Texas at Austin 2005
Distinction awarded to outstanding graduate student employees (research and teaching assistants). Nominated by the Linguistics department.

PARTICIPATION IN RESEARCH PROJECTS

European Research Council (ERC) Advanced Grant, EU 2011-2016
Title: *Strategic Conversation (STAC)*. PI: Nicholas Asher, CNRS Toulouse. Project Member.

Laboratoire d'Excellence (Labex), France 2011-2021
Title: *Empirical Foundations of Linguistics*. Associate Member.

French National Research Agency (ANR), France 2010-2012
Title: *Dynamically Enriching Lexical resources in multilingual and multimodal applications (EdyLex)*. PI: Benoît Sagot, INRIA Alpage. Project Member.

INRIA and Aquitaine Region, France 2010-2012
Title: *Automatic Itinerary Extraction from Travel Logs (ITIPY)*. PI: Renaud Marlet, INRIA Bordeaux. External Collaborator.

French National Research Agency (ANR), France 2009-2011
Title: *Large coverage probabilistic syntactic parsing of French (SEQUOIA)*. Grant ANR-08-EMER-013. PI: Alexis Nasr, Université de la Méditerranée. Project Member.

Pôle de compétitivité System@tic, France 2008-2010
Title: *Semi-automatic and Collaborative Retrieval of Information Based on Ontologies (SCRIBO)*. <http://scribo.ws>. Project Member.

Information and Intelligent Systems, National Science Foundation (NSF), USA 2006-2008
Title: *Extracting and Using Discourse Structure to Resolve Anaphoric Dependencies: Combining Logico-Semantic and Statistical Approaches*. Grant IIS-0535154. PI: N. Asher, Co-PIs: J. Baldridge; J. Kuhn, The University of Texas at Austin. <http://comp.ling.utexas.edu/discor/>. Research Assistant.

Morris Memorial Grant, New York Community Trust, USA 2002-2003
Title: *Linguistic Research on Text Interpretation*. PI: Carlota Smith, The University of Texas at Austin. Research Assistant.

PUBLICATIONS

Refereed Journals

P. Denis and J. Baldridge. 2009. Global joint models for coreference resolution and named entity classification. In *Procesamiento del Lenguaje Natural*, 42, Barcelona: Sociedad Española para el Procesamiento del Lenguaje Natural, 87–96.

P. Amsili, P. Denis and L. Roussarie. 2005. Anaphores abstraites en français: représentation formelle. In *Traitement automatique des langues (TAL)*, 46(1), Paris: Hermès, 15–39.

L. Prévot, P. Muller, P. Denis and L. Vieu. 2002. Une approche sémantique et rhétorique du dialogue. Un cas d'étude: l'explication d'itinéraires. In *Traitement automatique des langues (TAL)*, 43(2), Paris: Hermès, 43–70.

J. Busquets and P. Denis. 2001. L'ellipse modale en français: le cas de *pouvoir* et *devoir*. In *Les Cahiers de Grammaire*, 26, Toulouse, France, 55–74.

Refereed Conferences

- P. Denis and P. Muller. 2011. Predicting globally-coherent temporal structures via endpoint inference and graph decomposition. To appear in *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, Barcelona, Spain.
- A. Bittar, P. Amsili, P. Denis, and L. Danlos. 2011. French TimeBank: An ISO-TimeML Annotated Reference Corpus. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, Short paper, Portland, Oregon.
- A. Bittar, P. Amsili, and P. Denis. 2011. French TimeBank : un corpus de référence sur la temporalité en français. To appear in *Actes de Traitement Automatique des Langues Naturelles (TALN 2011)*, Montpellier, France.
- E. Lassalle and P. Denis. 2011. Leveraging different meronym discovery methods for bridging resolution in French. To appear in *Anaphora Processing and Applications, Proceedings of the 8th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2011), Selected papers*. LNAI, Springer.
- E. de la Clergerie, B. Sagot, R. Stern, P. Denis, G. Recourcé, and V. Mignot. 2011. Extracting and Visualizing Quotations from News Wires. In Z. Vetulani (Ed.), *Human Language Technology. Challenges for Computer Science and Linguistics - 4th Language and Technology Conference, LTC 2009, Revised Selected Papers*, LNAI, Springer, 522-532.
- P. Denis and P. Muller. 2010. Comparison of different algebras for inducing the temporal structure of texts. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, Beijing, China.
- M. Candito, J. Nivre, P. Denis, and E. Henestroza-Anguiano. 2010. Benchmarking of statistical dependency parsers for French. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, Short paper, Beijing, China.
- P. Denis and B. Sagot. 2010. Exploitation d'une ressource lexicale pour la construction d'un étiqueteur morpho-syntaxique état-de-l'art du français. In *Actes de Traitement Automatique des Langues Naturelles (TALN 2010)*, Montreal, Canada.
- S. Afantenos, P. Denis, P. Muller, and L. Danlos. 2010. Learning recursive segments for discourse parsing. In *Proceedings of the 7th international conference on Language Resources and Evaluation (LREC 2010)*, Valletta, Malta.
- M. Candito, B. Crabbé, and P. Denis. 2010. Statistical French dependency parsing: treebank conversion and first results. In *Proceedings of the 7th international conference on Language Resources and Evaluation (LREC 2010)*, Valletta, Malta.
- P. Denis and B. Sagot. 2009. Coupling an annotated corpus and a morphosyntactic lexicon for state-of-the-art POS tagging with less human effort. In *Proceedings of the 23rd Pacific Asia Conference on Language, Information, and Computation (PACLIC 23)*, Hong Kong, China.
- E. de la Clergerie, B. Sagot, R. Stern, P. Denis, G. Recourcé, and V. Mignot. 2009. Extracting and visualizing quotations from news wires. In *Proceedings of The Language and Technology Conference (LTC 09)*, Poznań, Poland.
- M. Candito, B. Crabbé, P. Denis, and F. Guérin. 2009. Analyse syntaxique du français : des constituants aux dépendances. In *Actes de Traitement Automatique des Langues Naturelles (TALN 2009)*, Senlis, France.
- P. Denis and J. Baldridge. 2008. Specialized models and ranking for coreference resolution. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2008)*, Honolulu, Hawaii.
- P. Denis and J. Baldridge. 2007. Joint determination of anaphoricity and coreference resolution using integer programming. In *Proceedings of the Annual Meeting of the North American Chapter of the Association for Computational Linguistics - Human Language Technology Conference (NAACL/HLT 2007)*, Rochester, New York, 236–243.

P. Denis and J. Baldrige. 2007. A ranking approach to pronoun resolution. In *Proceedings of the International Joint Conference in Artificial Intelligence (IJCAI 2007)*, Hyderabad, India, 1588–1593.

P. Denis and J. Kuhn. 2006. Applying an LFG parser in coreference resolution. Experiments and analysis. In *Proceedings of the International Lexical Functional Grammar Conference (LFG 2006)*, Konstanz, Germany, 166–183.

N. Asher and P. Denis. 2004. Dynamic typing for lexical semantics. A case study: the genitive construction. In A. Varzi and L. Vieu, editors, *Formal Ontology in Information Systems. Proceedings of the Third International Conference (FOIS 2004)*, IOS Press, Amsterdam, 165–176.

P. Denis and P. Muller. 2004. A semantics for temporally dependent expressions. In O. Bonami and P. Cabredo Hofherr, editors, *Empirical Issues in Syntax and Semantics 5, Papers from CSSP 2003*, Paris, 45–62. <http://www.cssp.cnrs.fr/eiss5/>.

P. Denis. 2003. A fresh look at goal motion constructions. In *Papers from the 39th Regional Meeting of the Chicago Linguistics Society (CLS 39)*, Chicago, IL.

Refereed Workshops

N. Asher, P. Denis, and B. Reese. 2006. Names and pops and discourse structure. In *Proceedings of the 2006 Workshop on Constraints in Discourse Structure (CID 2006)*. Maynooth, Ireland.

N. Asher, J. Hunter, P. Denis, and B. Reese. 2006. Evidentiality and intensionality: Two uses of reportative constructions in discourse. In *Proceedings of the 2006 Workshop on Constraints in Discourse Structure (CID 2006)*, Maynooth, Ireland.

N. Asher and P. Denis. 2005. Lexical ambiguity as type disjunction. In P. Bouillon and Kyoko Kanzaki, editors, *Proceedings of the Third International Workshop on Generative Approaches to the Lexicon (GL2005)*, Geneva, Switzerland. 10–17.

P. Denis, J. Kuhn and S. Wechsler. 2003. V-PP goal motion complexes in English: an HPSG account. In *Proceedings of the ACL-SIGSEM workshop: The Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, Toulouse, France, 121–132.

Editorial Work

P. Denis, E. McCready, A. Palmer, and B. Reese (Eds.). 2006 *Proceedings of the 2004 Texas Linguistics Society Conference: Issues at the Semantics-Pragmatics Interface*, Cascadilla Proceedings Project, Somerville, MA, 2006. <http://www.lingref.com/cpp/tls/2004/index.html>.

Technical Reports

B. Reese, P. Denis, N. Asher, J. Baldrige, and J. Hunter. 2007. *Reference Manual for the Analysis and Annotation of Rhetorical Structure*, Department of Linguistics, The University of Texas at Austin. Available from <http://comp.ling.utexas.edu/discor>.

Invited Talks

Predicting globally-coherent temporal structures via endpoint inference and graph decomposition. CLG seminar, University of Oxford, August 2011.

Specialized rankers and global models for coreference. Computational Linguistics Colloquium, Universität Heidelberg, July 2009.

Balancing multiple models for coreference resolution using integer programming. Séminaire Talaris, Loria, Nancy, May 2009.

Specialized models and ranking for coreference resolution. Séminaire Alpage, Université Paris 7, Paris, October 2008.

Integrating anaphoricity, named entity classification, and coreference resolution using Integer Linear Programming. Powerset Inc., San Francisco, CA, March 2007. (Job talk)

Global models and rich features for automated coreference resolution. Department of Linguistics, Indiana University, Bloomington, IN, February 2007. (Job talk)

Modéliser l'ambiguïté lexicale avec des types disjoints. Lattice, Université Paris 7, Paris, June 2005.

Learning Noun Phrase coreference using different syntactic views. IMS Institutversammlung, Universität Stuttgart, Germany, January 2006; and Département de Linguistique, Université d'Ottawa, Canada, February 2006. (Job talk)

SERVICE

Journal reviewing

ACM Transactions on Speech and Language Processing; ACM Computing Surveys; Journal of Linguistics; Traitement Automatique des Langues (TAL)

Conference program committee/reviewing

ACL 2007, 2009, 2011, 2012; IJCAI 2011; AAAI 2012, EMNLP 2009, 2010, 2011; COLING 2010; EACL 2012; TALN 2011, 2012; CICLING 2008; IEEE-ICSC 2009; LSRL 2005; SuB 2008; TLS XII; NAACL/HLT 2009 Workshop on Integer Linear Programming for NLP; ATALA Workshop on CRFs 2010; CID 2011; ATALA Workshop on Corpus Annotation 2010; ATALA Workshop on Discourse et TAL 2012

Conference/Event organizing

Co-chair of *Texas Linguistic Society 8th Annual Conference (TLS VIII)*, May 2004, Austin, TX, USA
Co-organizer of *TempEval-2: Evaluating Events, Time Expressions, and Temporal Relations*, July 2010, Uppsala, Sweden

Co-organizer of *Constraints in Discourse (CID 2011)*, Agay, France.

Member of professional organizations

The Association for Computational Linguistics (ACL); ACL's Special Interest Group on the Lexicon (SIGLEX), on Semantics (SIGSEM), and on Natural Language Learning (SIGNLL); The French Association for Natural Language Processing (ATALA)

OPEN SOURCE SOFTWARE DEVELOPMENT

Toolkit for Advanced Discriminative Modeling (TADM) <http://tadm.sourceforge.net/>
Developer. High performance C++ and Python toolkit for training maximum entropy and perceptron models. LGPL.

Coreference Resolution Toolkit (CoRTex) <http://comp.ling.utexas.edu/software/cortex>
Original author and main developer. Python toolkit for coreference resolution. LGPL.

Maximum-Entropy Lexicon-Enriched Tagger (MELT) <http://lingwb.gforge.inria.fr/>
Co-developed with B. Sagot. High precision probabilistic, lexically-enriched tagger for French. Currently being integrated within the Alpage linguistic pipeline.

Newswire Quotation Extractor & Navigator (SAPIENS) <http://passage.inria.fr/SAPIENS/>
Developer of the anaphora resolution module. LGPL.

TEACHING EXPERIENCE

Main Instructor, Master Linguistique Informatique (LI), Université Paris 7 2009-2012
Sémantique Automatique du Langage Naturel (Graduate level, Fall 2009, Fall 2010, Fall 2012)

Invited Speaker, Journées de Sémantique et de Modélisation (JSM), Loria, Nancy
Automatic Reference Resolution. A Quick Overview. Outreach lecture.

Teaching Assistant, Linguistics Department, The University of Texas at Austin 2000-2003*Computational Linguistics I* (Graduate level, Spring 2003, Instructor: Jonas Kuhn)*Introduction to Cognitive Science* (Undergraduate level, Spring 2003, Instructor: Carlota Smith)*Introduction to Computational Linguistics* (Undergraduate level, Fall 2002, Instructor: Bernard Schwarz)*Automata Theory* (Undergraduate level, Spring 2001, Instructor: Rajesh Bhatt)*Language and the Internet* (Undergraduate level, Fall 2000, Instructor: Rajesh Bhatt)**Language Instructor, Université Libre de Bruxelles 2001-2003**Introductory and advanced French language classes at the *Summer School in French Language and Literature*.**ADVISING****Ph.D. co-supervision**André Bittar, Université Paris 7, *Building a TimeBank for French: A Reference Corpus Annotated According to the ISO-TimeML Standard*, defended on November 19th, 2010 (main advisor: Laurence Danlos)

Emmanuel Lassalle, Université Paris 7 (main advisor: Laurence Danlos)

Chloé Braud, Université Paris 7 (main advisor: Laurence Danlos)

M.A. thesis supervision

Emmanuel Lassalle, Ecole Normale Supérieure, defended in September 2010

Chloé Braud, Université Paris 7, defended in June 2011

Internships

Altaf Rahman, University of Texas at Dallas, USA, June-August 2011

Alexis Vanacker, Ecole des Mines de Nancy, June-August 2009

PHD COMMITTEES**Computer Science Department, Universitat Politècnica De Catalunya May 2012**

External member on Emili Sapena's dissertation committee

DEPARTMENTAL COMMITTEES**Computer Science Department, Université de Marseille, France May 2012**

External member on Computer Science Maître de Conférences Search Committee

Computer Science Department, Université de Marseille, France May 2011

External member on Computer Science Maître de Conférences Search Committee

Linguistics Department, The University of Texas at Austin 2002-2007

Co-chair of the Associated Students of Linguistics (ASOL) executive committee (2006-2007)

Graduate Student Representative to the Graduate Student Assembly (GSA) (2005-2006)

Student member of the Computational Linguistics Search Committee (2002-2003, 2004-2005)

REFERENCES

Nicholas Asher

Professor
Philosophy Department
University of Texas at Austin
1 University Station, C3500
Austin, TX 78712
USA
+1 512 471-4857
nasher@mail.utexas.edu

Jason Baldrige

Assistant Professor
Linguistics Department
University of Texas at Austin
1 University Station, B5100
Austin, TX 78712-0198
USA
+1 512 232-7682
jbaldrid@mail.utexas.edu

Andrew Kehler

Associate Professor
Linguistics Department
University of California San Diego
9500 Gilman Drive, 0108
La Jolla, CA 92093-0108
USA
+1 858 534-6239
keehler@ling.ucsd.edu

Jonas Kuhn

Professor
Institut für Linguistik
Universität Potsdam
Karl-Liebknecht-Str. 24/25
D-14476 Golm
Germany
+49 331 977-2444
kuhn@ling.uni-potsdam.de

Raymond Mooney

Professor
Computer Science Department
University of Texas at Austin
1 University Station, C0500
Austin, TX 78712-0233
USA
+1 512 471-9558
mooney@cs.utexas.edu

James Pustejovsky

Professor
Computer Science Department
Brandeis University
415 South Street
Waltham, MA 02454-9110
USA
+1 781 736-2709
jamesp@cs.brandeis.edu