

# Ariadna Font Llitjós

Language Technologies Institute ~ School of Computer Science ~ Carnegie Mellon University

5000 Forbes Ave. Pittsburgh, PA 15213 Tel: 412-268-3653 (office) 412-654-4861 (cell)

[aria+cv@cs.cmu.edu](mailto:aria+cv@cs.cmu.edu)

<http://www.cs.cmu.edu/~aria>

**RESEACRH INTERESTS:** Machine Translation, Machine Learning, Natural Language Processing and Understanding, User Interfaces, Dialog Systems and Speech Synthesis.

## EDUCATION

- Aug 01 – July 2007 **PhD candidate**, *Language Technologies Institute, Carnegie Mellon University*.  
(*expected*) Thesis topic: **Improving Rule-Based Machine Translation Systems with an Automatic Rule Refiner**. Advisor: Dr. Jaime Carbonell
- Aug 99 – May 01 **Masters in Science**, *Language Technologies Institute, Carnegie Mellon*.  
Masters thesis: **Knowledge of Language Origin Improves Pronunciation Accuracy of Proper Names**. Applied Machine Learning techniques to augment basic pronunciation models. Advisor: Dr. Alan W Black.
- Sep 97 – Jun 99 **Masters in Cognitive Science and Language**, *Universitat Pompeu Fabra, Universitat Politècnica de Catalunya, Universitat Autònoma de Barcelona, Universitat de Barcelona*. Advisor: Dr. Toni Badia
- Sep 92 – Jul 96 **BA with Honors in Translation and Interpreting (English, German)**, *Universitat Pompeu Fabra*. Senior thesis: **A bottom-up chart parser for PATR unification grammars in prolog**. Advisor: Dr. Toni Badia.

## WORK EXPERIENCE

- May 2002-Present **Research Assistant**, *Language Technologies Institute, Carnegie Mellon University*. Project AVENUE: focuses on reducing development time and cost of MT, with special emphasis to languages with limited resources. (<http://www.cs.cmu.edu/~avenue>). I worked with **Mapudungun** and **Quechua**.
- April and Nov 2002 **Lead Technical Consultant** for the AVENUE project (CMU) in Temuco, Chile. Supervised development of resources and NLP tools for Mapudungun–Spanish (transcribed spoken corpus, dictionary, morphological analyzer) and provided technical assistance setting up and using software developed at CMU. Instituto de Estudios Indígenas at Universidad de la Frontera.
- May – Jul 2000 **Summer Intern at Microsoft Research** (Redmond). Worked with the NLP team in building the Spanish Generation Module for the general purpose, multilingual Machine Translation System (NLPWIN).  
Mentors: Steve Richardson and Maite Melero. Manager: Karen Jensen.
- Sep 97 - Jul 99 **Research Assistant**, *Institute for Applied Linguistics (IULA), Universitat Pompeu Fabra*. Worked on implementation of a Catalan Constraint Grammar. Directed by Dr.Toni Badia.

## Ariadna Font Llitjós

- Sep 96 - Mar 97 Design and implementation of a Catalan Analysis Grammar for a transfer, LFG-based Machine Translation system (LEKTA).
- Apr - Aug 1997 **Research Assistant**, *Institut für Maschinelle Sprachverarbeitung, University of Stuttgart*. Extraction of subcategorization patterns of English verbs from the British National Corpus with Xkwic tool. Directed by Dr. Ulrich Heid.

## TEACHING EXPRIENCE

- Sept – Dec 2004 **Teaching Assistant**, *School of Computer Science, CMU*.  
[Technology Consulting in the Community](#) (15-391).
- Sept – Dec 2001 Language Technologies (11-682), which evolved into the current [Introduction to IR, NLP, MT, and Speech \("Words for Nerds"\)](#).
- Feb - Jul 1999 **Co-taught two courses** in the Graduate Program: Text Management for Professionals: Translation and Processing. *Universitat Pompeu Fabra*.
- Aug - Sep 1995 Taught “Introduction to NLP and Grammars” at the *Information Systems Department of the Universidad Tecnológica Nacional, Buenos Aires*.

## SKILLS

**Natural Languages:** Catalan and Spanish (native), English (proficient) German, French (advanced), Portuguese, Italian (written and spoken understanding) and Quechua (beginner).

**Computer languages:** C++, Perl, HTML, CGI, VoiceXML.

**Computer Environments:** Linux/UNIX, Windows XP, OSX.

## HONORS

**2005 Google Anita Borg Memorial Scholarship Winner**, Google.

**Graduate Student Citizenship Award 2004**, SCS at CMU.

**“La Caixa” Fellowship** to pursue graduate studies in the US 1999-2001.

**Grace Hopper Celebration Scholarship 2002 and 2004**, USA.

## ACTIVITIES

- June-August 05 **Built a Quechua-to-Spanish MT prototype system** (Cusco, Peru) as part of a graduate course designed to grow a vision (V-Unit), part of CMU’s [TechBridgeWorld](#) initiative.
- April 2003 **Member of the ACL-2003 Student Research Workshop Program Committee.**
- Feb 2003 **SCS DAY creator and main coordinator** [www.cs.cmu.edu/~scsday](http://www.cs.cmu.edu/~scsday)  
SCS Day is a celebration of the diversity in the School of Computer Science at CMU, including workshops, exhibits, games, and a talent show, where all the members of the SCS community get together to display their talents and share their skills.
- 2001 - Present **[Women@SCS](#) graduate committee member.** Committee head (Jan-Aug 03), organize and participate in events such as Outreach Roadshows, inviting speakers, Expanding Your Horizons Workshop, conferences, panels, talks.  
[www.women.cs.cmu.edu](http://www.women.cs.cmu.edu).
- 2001 – 2002 **Dec5 member** support organization of CMU SCS grad students.
- 1999 – 2001 **Language Technologies Institute student representatives.**

## SELECTED PUBLICATIONS

- [A Walk on the Other Side: Adding Statistical Components to a Transfer-Based Translation System](#) with Stephan Vogel. To appear in Syntax and Structure in Statistical Translation (SSST) Workshop at HLT-NAACL, 26 April 2007, Rochester, New York, USA.
- [The Inner Works of an Automatic Rule Refiner for Machine Translation](#) with William Ridmann. METIS-II Workshop, January 11, 2007, Leuven, Belgium.
- [Automating Post-Editing to Improve MT Systems](#) with Jaime Carbonell. Automated Post-Editing Workshop at AMTA, August 12, 2006, Boston, USA.
- [Giving the Power to Bilingual Speakers](#). Position Paper for the Automated Post-Editing Workshop at AMTA, August 12, 2006, Boston, USA.
- [Can the Internet help improve Machine Translation](#). Doctoral Consortium at HLT-NAACL, June 4, 2006, New York, USA.
- [Building Machine translation systems for indigenous languages](#), with Roberto Aranovich and Lori Levin. Second Conference on the Indigenous Languages of Latin America (CILLA II), 27-29 October 2005, Texas, USA.
- [A Framework for Interactive and Automatic Refinement of Transfer-based Machine Translation](#), with Jaime Carbonell and Alon Lavie. EAMT 10th Annual Conference, 30-31 May 2005, Budapest, Hungary.
- [Error Analysis of Two Types of Grammar for the Purpose of Automatic Rule Refinement](#), with Katharina Probst and Jaime Carbonell. AMTA, 2004, Washington DC.
- [The Translation Correction Tool: English-Spanish user studies](#), with Jaime Carbonell. LREC, 2004. Lisbon, Portugal.
- Lavie, A., S. Vogel, L. Levin, E. Peterson, K. Probst, **A. Font Llitjós**, R. Reynolds, J. Carbonell, and R. Cohen, [Experiments with a Hindi-to-English Transfer-based MT System under a Miserly Data Scenario](#). ACM Transactions on Asian Language Information Processing (TALIP), 2(2). June 2003. Pages 143-163.
- Panel member "From Bits to Bots: Women Everywhere, Leading the Way": [AVENUE, Automatic Machine Translation for low-density languages](#). [Grace Hopper Celebration](#), 2002 Vancouver, Canada.
- [Building Voicexml-based applications](#), with Christina Bennett, Stefanie Shriver, Alexander Rudnicky and Alan W Black. *ICSLP 2002*. Denver, USA.
- [Unit Selection without a Phoneme Set](#), with Alan W Black. *IEEE Synthesis Workshop 2002*. Santa Monica, USA.
- [Automatic Pronunciation of Proper Names using Language Origin Classes and Unsupervised Clustering](#). *ACL Student Session 2002*, Philadelphia, USA.
- [Evaluation and collection of proper name pronunciations online](#), with Alan Black. *LREC 2002*, Las Palmas, Spain.
- [Improving Pronunciation Accuracy of Proper Names with Language Origin Classes](#). Masters Thesis, August 2001. CMU-LTI-01-169. Carnegie Mellon University.
- [Knowledge of Language Origin Improves Pronunciation Accuracy of Proper Names](#), with Alan Black. *Eurospeech 2001*, Aalborg, Denmark