

## Karine Megerdooian, Ph.D.

[Personal information removed]

U.S. CITIZEN

---

### PROFILE

Senior scientist with over 10 years of experience in research and development in computational linguistics with focus in the areas of **machine translation, information extraction, and lexical semantics**. Expertise includes in-depth theoretical and hands-on knowledge of linguistic and Natural Language Processing approaches for **less commonly studied languages**, with a strong specialization in Persian. Experience in managing technical and linguistic aspects of NLP projects in industry and academia. Linguistic specialization in the study of **complex predicates and language acquisition of Heritage Speakers** in Persian and Armenian.

---

### EMPLOYMENT HISTORY

- |                               |   |                |
|-------------------------------|---|----------------|
| Oct. 2006<br>to present       | <b>Senior Artificial Intelligence Engineer</b><br>The MITRE Corporation, McLean, VA<br><i>Human Language Technology Group</i><br>Contributing as a computational linguist to projects in Persian and multilingual NLP: event analysis and temporal relations, multilingual information retrieval, and cross-script name-matching. Currently developing projects in computational analysis of Persian language weblogs, conversational Persian dialects, and Tajiki machine translation.   | <b>MITRE</b>   |
| Sept. 2005<br>to Oct. 2006    | <b>Assistant Research Scientist</b><br>Center for Advanced Study of Language, University of Maryland<br>Developed test design and content for cross-linguistic investigation of language acquisition by heritage and L2 speakers. Project manager for the aspects involving linguistic and sociolinguistic analysis of Persian heritage speakers.   | <b>CASL</b>    |
| August 2004<br>to August 2005 | <b>Consultant</b><br>Various projects: <ul style="list-style-type: none"><li>• Inxight Software, Inc.; Sunnyvale, California; 8/04-7/05 and 7/06-10/06<br/>Improved performance of the Inxight Persian text mining system. Developed an automatic transcription system of Persian named entities.</li><li>• Multilingual Named Entity Detection and Transliteration Project; 2/05-8/05<br/>University of Illinois, Urbana-Champaign<br/>Developed Persian and Eastern Armenian entity extraction rules. PI: Prof. Richard Sproat</li><li>• Diagnostic Testing Project (CAST); 2/05-8/05<br/>Language Acquisition Resource Center; San Diego State University<br/>Responsible for designing and developing online proficiency test items for Persian, based on the ACTFL proficiency classification.</li></ul> |                |
| January 2002<br>to Aug. 2004  | <b>Computational Linguist, Technical Lead</b><br>Inxight Software, Inc.; Sunnyvale, California (currently Business Objects)<br>Developed and maintained Persian, Arabic and Slovenian text mining systems as part of the Inxight LinguistX platform, using Xerox finite-state technology. The projects included stemmer, tagger, shallow parsing and information extraction modules, and used a statistical disambiguator.  | <b>Inxight</b> |
| January 1998<br>to April 2001 | <b>Computational Linguist</b><br>Computing Research Laboratory (CRL), New Mexico State University, Las Cruces <ul style="list-style-type: none"><li>• Led the development and acquisition of the linguistic aspects of a large-scale Persian to English machine translation system. Responsible for building the linguistic components and for integration with the computational system (Shiraz project).</li><li>• Acquired ontological knowledge for an English lexicon (MikroLite project).</li></ul>   | <b>CRL</b>     |

- Developed the linguistic components for a Turkish to English machine translation system. Incorporated a pre-existing lexicon and morphological analyzer. Supervised additional lexical acquisition (Expedition project).
- Prepared the linguistic specification for the development of a structural transfer component, used in a machine translation system from low-density languages to English (Boas project).
- Supervised lexicon acquisition for five languages during an upgrade phase for a multilingual machine translation system (Corelli project).

Project managers: Prof. Sergei Nirenburg, Dr. Rémi Zajac, and Dr. Evelyne Viegas

May 1997 **Research Assistant** ISI  
to Aug. 1997 Information Sciences Institute (ISI), Marina del Rey  
Developed a grammar of Persian for a Machine Translation project. PI: Prof. Kevin Knight

---

## EDUCATION

2002 **Ph.D. in Linguistics** University of Southern California, Los Angeles  
Specialization: Syntax-Morphology interface; complex predicates; Armenian and Persian.  
**Dissertation title:** *Beyond Words and Phrases: A Unified Theory of Predicate Composition*  
Dissertation director: Prof. Jean-Roger Vergnaud  
Committee members: Profs. Joseph Aoun, Maria Luisa Zubizarreta, Yen-Hui Audrey Li.

1995 **Master of Arts in Linguistics** University of Southern California, Los Angeles  
**M.A. thesis title:** *Scrambling and Scopal Interactions in a Minimalist Perspective*  
Advisor: Prof. Jean-Roger Vergnaud

1991–1993 **Directed readings in Neural Computation** University of Southern California, Los Angeles  
Supervisor: Prof. William Wagner

1991 **Bachelor of Science in Physics** University of Southern California, Los Angeles  
Coursework included Statistical Mechanics, Quantum Theory, Signal Processing, Integrated Circuits and Electrical Engineering.

---

## INVITED TALKS AND TEACHING

May 2008 **Invited Speaker, USC Student Workshop**  
University of Southern California, Los Angeles  
*Some bare nouns are more equal than others*

January 2008 **Keynote Speaker, Teachers of Persian Workshop**  
University of Maryland, College Park  
*Difficulties for Learning Persian as a Second Language*

October 2007 **Invited Instructor, NATO Advanced Study Institute**  
Advances in Language Engineering for Low- and Middle-Density Languages  
Batumi, Georgia  
*Building Computational Resources and Processors for Persian and Armenian*

June 2006 **Invited Instructor, Linguistics Department**  
University of Tehran, Iran  
Taught a course and tutorial on ‘Finite-State Morphology’ with emphasis on issues of Persian.

June 2006 **Invited Talk, 2<sup>nd</sup> Workshop on Persian Language and Computers**  
June 28-30; University of Tehran, Iran  
*Extending a Persian Morphological Analyzer to Blogs*

- January 2004 to July 2005 **Lecturer, Heritage Language Program**  
Linguistics Department, University of California San Diego  
Responsible for all aspects of instruction, including curriculum development, for Persian and Armenian language analysis and linguistics for heritage speakers.
- May 2005 **Invited Talk, Workshop on Armenian Syntax**  
May 22-25; Pithiviers, France  
*Reflexive Verbs and Voice Morphology in Armenian*
- May 2004 **Invited Instructor, Linguistics Department**  
University of Tehran, Iran  
One-day seminar on 'Introduction to Computational Linguistics'.
- May 2004 **Invited Talks, 1<sup>st</sup> Workshop on Persian Language and Computers**  
May 25-26; University of Tehran, Iran  
*Developing a Persian Part-of-Speech Tagger*  
*A Semantic Template for Light Verb Constructions*
- April 2002 **Invited Talk, USC-GSIL Student Workshop**  
University of Southern California, Los Angeles  
*Parallel Domains and the Projection of Arguments*
- May 2001 **Invited Lecture, New Trends in Linguistics Workshop**  
May 27-29; Allameh Tabatabai University, Tehran, Iran  
*Theories of Verb Formation*
- June 2001 **Invited Talk, Workshop on Issues in Persian Linguistics**  
Islamic Azad University at Khorasgan, Isfahan, Iran  
*Complex Predicates in Persian*
- May 1999 **Guest Instructor, Summer School in Language Engineering**  
New Mexico State University  
*Approaches to Computational Morphology: Applications for Persian*
- Sept. 1998 to Dec. 1998 **Teacher's Assistant, Linguistics Department**  
University of Southern California, Los Angeles  
Taught discussion sessions for the course 'Language and Mind' on acoustic and articulatory phonetics, phonology, speech perception and processing.
- Sept. 1994 to May 1997 **Assistant Lecturer, French and Italian Department**  
University of Southern California  
Full responsibility for all aspects of French language instruction.

---

## LANGUAGES

Native speaker of Eastern Armenian and Persian (Farsi), and near-native speaker of English and French.  
High working knowledge of Western Armenian, Italian, Spanish; basic knowledge of German, Russian and Arabic.

---

## SKILLS

- Expertise in several major areas of Natural Language Processing, especially pertaining to multilingual morphological analysis, information extraction, and machine translation. Specializing in low-density languages with emphasis on Iranian Persian (Farsi), Tajiki Persian, and Armenian.

- Development expertise in unification-based systems, Xerox finite-state tools, morphological and syntactic analysis and generation, lexicography, ontological knowledge representation. Experience in corpus linguistics.
- Linguistic specialization in Distributed Morphology, Minimalist Program, Principles and Parameters Theory. Knowledge of HPSG.
- Familiar with Perl and scripting languages; basic knowledge of Java.
- Curriculum development for teaching languages and linguistics to heritage speakers

---

## PROFESSIONAL ACTIVITIES AND MEMBERSHIPS

- Journal and conference reviews for *Word Structure*, *Language Sciences*, *Natural Language and Linguistic Theory*, *Computers and the Humanities*, *Journal of Linguistics*, *Linguistica Atlantica*, and SALT conference.
- Proposal review for the NSF Linguistics Program.
- Organizer for the **International Conference on Complex Predicates in Iranian Languages**. Paris, France. July 5-6, 2008. *URL*: [www.iranianlinguistics.org/complexpredicates](http://www.iranianlinguistics.org/complexpredicates).
- Organizer for **Computational Approaches to Arabic Script-based Languages (CAASL)**. Stanford, California; July 21-22, 2007 and Geneva, Switzerland; August 28, 2004. *URL*: [arabicscript.org](http://arabicscript.org)
- Launched the **Society for Iranian Linguistics**. *URL*: [iranianlinguistics.org](http://iranianlinguistics.org).
- Professional memberships: Linguistic Society of America (LSA), Association for Computational Linguistics (ACL), International Society for Iranian Studies, Society for Armenian Studies.

---

## ACADEMIC HONORS AND AWARDS

- Member of the Technical Program Committee for the Conference on Language and Technology (CLT07), held in University of Peshawar, Pakistan. August 7-11, 2007.
- Member of the editorial board of *The Linguistics Journal*.
- Recipient of Ahmanson Foundation Dissertation Fellowship; University of Southern California, LA, 2000-2001.

---

## PUBLICATIONS

### Book Contributions

- Megerdooonian, Karine. To appear. "Low-Density Language Strategies for Persian and Armenian". In *Advances in Language Engineering for Low- and Middle-Density Languages*; edited by Sergei Nirenburg and Oleg Kapanadze. IOS Press of Amsterdam.
- Megerdooonian, Karine. 2008. "Parallel Nominal and Verbal Projections". In *Foundational Issues for Linguistic Theory*; edited by Robert Freidin, Carlos P. Otero and Maria Luisa Zubizarreta. MIT Press.
- Megerdooonian, Karine, 2003. "Text Mining, Corpus Building and Testing". In *A Handbook for Language Engineers*; edited by Ali Farghaly. CSLI Publications: Stanford, CA.

### Book and Proceedings Editing

- Megerdooonian, Karine and Nick Pendar (eds.). To appear. *Foundational Issues in the Computational Processing of Persian*.
- Megerdooonian, Karine and Ali Farghaly (eds.). 2004. *Proceedings of the 1<sup>st</sup> Workshop on Computational Approaches to Arabic Script-Based Languages*. COLING 2004.
- Megerdooonian, Karine and Leora Ann Bar-el (eds.). 2001. *Proceedings of the 20th West Coast Conference on Formal Linguistics*. Cascadilla Press.

### Journal Publications

- Megerdooonian, Karine. To appear. "The Status of the Nominal in Persian Complex Predicates". *Natural Language and Linguistic Theory*, Springer Netherlands.

- Megerdooimian, Karine. 2005. "Transitivity Alternations and Causative Constructions in Eastern Armenian". In *Annual of Armenian Linguistics* 21.
- Megerdooimian, Karine. 2001. "Event Structure and Complex Predicates in Persian". In *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 46(1/2): 97-125.

### Peer-Reviewed Conference Proceedings

- Megerdooimian, Karine and Dan Parvaz, 2008. "Low-Density Language Bootstrapping: The Case of Tajiki Persian". In *Proceedings of Language Resources and Evaluation Conference (LREC 2008)*. Marrakech, Morocco.
- Megerdooimian, Karine, 2004. "Finite-State Morphological Analysis of Persian". In *Proceedings of the Workshop on Computational Approaches to Arabic Script-Based Languages, COLING 2004*. University of Geneva, Switzerland.
- Amtrup, Jan W., Karine Megerdooimian and Remi Zajac, 2000. "Rapid Development of Translation Tools: Application to Persian and Turkish". In *Proceedings of COLING 2000*, Saarbrücken, Germany.
- Megerdooimian, Karine, 2000. "Unification-Based Persian Morphology". In *Proceedings of CICLing 2000*. Alexander Gelbukh (ed.). Centro de Investigacion en Computacion-IPN, Mexico.
- Amtrup, Jan W., Karine Megerdooimian and Remi Zajac, 1999. "Rapid Development of Translation Tools". In *Proceedings of Machine Translation Summit VII*, Singapore, pp. 385-389, September 1999.

### Other Conference Proceedings

- Megerdooimian, Karine, 2006. "Extending a Persian Morphological Analyzer to Blogs". In *Proceedings of the Second Workshop on Persian Language and Computers*. Tehran University, Iran. June 28-30.
- Megerdooimian, Karine, 2004. "Developing a Persian Part-of-Speech Tagger". In *Proceedings of the First Workshop on Persian Language and Computers*. Tehran University, Iran. May 25-26.
- Megerdooimian, Karine, 2004. "A Semantic Template for Light Verb Constructions". In *Proceedings of the First Workshop on Persian Language and Computers*. Tehran University, Iran. May 25-26.
- Megerdooimian, Karine and Shadi Ganjavi, 2001. "Against Optional *Wh*-Movement". In *Proceedings of WECOL 2000*, Volume 12. California State University, Fresno.
- Megerdooimian, Karine, 2001. "Aspect and Partitive Objects in Finnish". In *Proceedings of WCCFL 2000*. Cascadilla Press.
- Megerdooimian, Karine, 2000. "Projection of Direct Objects". In *Proceedings of WECOL 1999*, Volume 11. California State University, Fresno.

### Technical Reports & Unpublished Manuscripts

- Megerdooimian, Karine, 2004. "Developing a Persian Text Mining System". Ms., Inxight Software, Sunnyvale.
- Megerdooimian, Karine, 2004. "Dichotomy in the Structure of the Lexicon: Transitivity Splits in Armenian". Ms., University of California, San Diego.
- Amtrup, Jan W., Hamid Mansouri Rad, Karine Megerdooimian and Remi Zajac, 2000. "Persian-English Machine Translation: An Overview of the Shiraz Project". NMSU, CRL, Memoranda in Computer and Cognitive Science (MCCS-00-319).
- Megerdooimian, Karine, 2000. "Persian Computational Morphology: A Unification-Based Approach". NMSU, CRL, Memoranda in Computer and Cognitive Science (MCCS-00-320).
- Megerdooimian, Karine, 2000. "A Computational Analysis of the Persian Noun Phrase". NMSU, CRL, Memoranda in Computer and Cognitive Science (MCCS-00-321).
- Megerdooimian, Karine and Remi Zajac, 2000. "Processing Persian Text: Tokenization in the Shiraz Project". NMSU, CRL, Memoranda in Computer and Cognitive Science (MCCS-00-322).

---

### RECENT CONFERENCE PRESENTATIONS (UNPUBLISHED)

- Preverbal Nominals and Telicity in Persian Complex Predicates. Talk presented at the *International Conference on Complex Predicates in Iranian Languages*, Paris, France. July 2008.
- Second Position Clitics in the *vP* Phase: The Case of the Armenian Auxiliary (with Arsalan Kahnemuyipour). Talk presented at the *LSA 2008 Meeting*, Chicago. January 2008.

- The Auxiliary Clitic and Sentential Stress in Eastern Armenian. Talk presented at the *GURT 2007: Little words – their history, phonology, syntax, semantics, pragmatics, and acquisition*, Georgetown University, Washington, DC. March 2007.
- Testing the Linguistic Competence of Persian Heritage Speakers (with Ilhan Cagri and Scott Jackson). Talk presented at the *Second Language Research Forum: Colloquium on understanding the interlanguage systems of heritage language learners*, Urbana-Champaign, Illinois. October 2007.
- On Incorporation in Persian Verbal Predicates. Talk presented at the *Second International Conference on Aspects of Iranian Linguistics*, Hamburg, Germany. June 2007.
- Machine Translation of Persian Complex Predicates (with Jan W. Amtrup). Talk presented at the *Second International Conference on Aspects of Iranian Linguistics*, Hamburg, Germany. June 2007.
- A Typology of Persian Complex Predicates. Talk presented at the *First International Conference on Aspects of Iranian Linguistics*, Leipzig, Germany. June 2005.
- Asymmetries in Form and Meaning: Surface Realization and Interface Conditions. Talk presented at *Approaching Asymmetry at the Interfaces* workshop, held in Université de Québec à Montréal, Canada. Oct. 24-25, 2003.
- Aspect in Complex Predicates. Talk presented at the *Workshop on Complex Predicates, Particles and Subevents*, University of Konstanz, Germany. Sept. 30-Oct. 2, 2002.
- The Derivation/Inflection Distinction and Post-Syntactic Merge (with Arsalan Kahnemuyipour). Talk presented at the *Canadian Linguistic Association / Association Canadienne de Linguistique*. University of Toronto, Canada. May 25-28, 2002.