Ariya Rastrow

CONTACT INFORMATION	CSEB 322	*Voice:* (703) 300-6728
	Center for Language and Speech Processing	*Fax:* (410) 516-5050
	Johns Hopkins University	*E-mail:* ariya@jhu.edu
	Baltimore, MD 21218, USA	*Web:* www.clsp.jhu.edu/∼arastrow

RESEARCH INTERESTS

Automatic Speech Recognition (ASR), Human Language Technology, Language Modeling, Statistical Modeling, Machine Learning, Information Retrieval, Data Mining,

EDUCATION

Ph.D. Candidate in Electrical and Computer Enginnering
Johns Hopkins University, Baltimore, MD   **Fall 2006 - Present**
Advisor: Professor Frederick Jelinek (deceased)
Expected graduation date: April 2012
Relevant Courses: Statistical Pattern Recognition, Machine Learning, Random Signal Analysis, Information Theory, Information Extraction, Graph Theory, Matrix Analysis and Linear Algebra, Natural Language Processing, Information Retrieve/Web Agents

M.S.E. in Electrical and Computer Engineering
Johns Hopkins University, Baltimore, MD   **May 2007**

Bachelor of Science in Electrical Enginnering
Sharif University of Technology, Tehran, Iran   **Aug. 2006**

EMPLOYMENT AND RESEARCH EXPERIENCE

Johns Hopkins University, Baltimore, MD   **September 2006 - Present**
Center for Language and Speech Processing and Human Language Technology COE
Research advisor: Frederick Jelinek 2006-2010 (deceased).
Current advisors: Sanjeev Khudanpur (ECE) and Mark Dredze (CS).

IBM T.J. Watson Research Center, Yorktown Heights, NY
Extended Summer Research Intern   **September - December 2008**
Speech Group (Manager: Bhuvana Ramabhadran)

- Proposed and implemented a new framework for discriminative training of n-gram language models and discriminative adaptation of language models to a new domain.

IBM T.J. Watson Research Center, Yorktown Heights, NY
Research Intern   **Summer 2008**
Speech Group (Manager: Bhuvana Ramabhadran)

- Proposed and implemented a new method for detecting out-of-vocabulary (OOV) words in the output of a large vocabulary continuous speech recognition (LVCSR) system. The proposed method uses a hybrid system combining words and variable-length data-driven sub word units.

CLSP Summer Research Workshop, Baltimore, MD
Researcher (Team Leader: Hynek Hermansky)   **Summer 2007**

- Participated in the JHU WS07 group called "Recovering from Model Inconsistency in Multi-lingual Speech Recognition"
- The group proposed novel methods based on inconsistency between unconstrained and constrained speech recognizers for detection of out-of-vocabulary (OOV) segments and erroneous regions in the output of large vocabulary continuous speech recognition (LVSCR) systems.

Optical and Wireless Networking Lab, Tehran, Iran
Research Assistant (Advisor: Prof. J. Salehi)   **June 2005 - May 2006**
Sharif University of Technology

- Implemented Spread-Time Code Division Multiple Access (ST-CDMA) technique for band limited access to the communication channel.

- The Frederick Jelinek Fellowship, Johns Hopkins University, Fall 2011.
- Dean's Graduate Tuition Fellowship, Johns Hopkins University, Fall 2006 - Present
- Human Language Technology Center of Excellence RA 2007 - Present
- Ranked 61st in nationwide university entrance exam (500,000 participants), Iran, Summer 2001
- Silver Medal of National Mathematic Olympia, Iran, Summer 2000

TEACHING
EXPERIENCE

Teaching Assistant, Signal and Systems
Johns Hopkins University   **Spring 2009**
Teaching Assistant, Introduction to VLSI
Johns Hopkins University   **Spring 2007**
Teaching Assistant, ECE laboratory
Johns Hopkins University   **Fall 2006**
Teaching Assistant, Digital Signal Processing
Sharif University of Technology   **Fall 2006**

PUBLICATIONS

A. Rastrow, M. Dredze and S. Khudanpur "Adapting n-gram Maximum Entropy Language Models with Conditional Entropy Regularization", IEEE Automatic Speech Recognition and Understanding (ASRU) Workshop, 2011.

A. Rastrow, M. Dredze and S. Khudanpur "Efficient Discriminative Training of Long-span Language Models", IEEE Automatic Speech Recognition and Understanding (ASRU) Workshop, 2011.

G. Kurata, A. Sethy, B. Ramabhadran, **A. Rastrow**, N. Itoh and M. Nishimura "Acoustically discriminative language model training with pseudo-hypothesis", Speech Communication, Volume 5, Issue 2, pages 219-228, 2011

A. Rastrow, M. Dreyer, A. Sethy, S. Khudanpur, B. Ramabhadran and M. Dredze Hill Climbing on Speech Lattices: A New Rescoring Framework, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2011

C. Parada, M. Dredze, A. Sethy, and **A. Rastrow** "Learning Sub-Word Units for Open Vocabulary Speech Recognition", The Association for Computational Linguistics (ACL), 2011.

A. Rastrow, F. Jelinek, A. Sethy, B. Ramabhadran "Unsupervised Model Adaptation using Information-Theoretic Criterion", North American Chapter of the Association for Computational Linguitics (NAACL), June 2010.

D. Povey, L. Burget, et al. "Subspace Gaussian Mixture Models - A Structured Model for Speech Recognition", Computer Speech and Language.

A. Goshal, D. Povey, et al. "A Novel Estimation of Feature-space MLLR for Full-Covariance Models", IEEE International Conference on Acoustics, Speech and Signal Processing, March 2010.

L. Burget, P. Schwarz, et. al. "Multilingual Acoustic Modeling for Speech Recognition based on Subspace Gaussian Mixture Models", IEEE International Conference on Acoustics, Speech and Signal Processing, March 2010.

D. Povey, L. Burget, et al. "Subspace Gaussian Mixture Models for Speech Recognition", IEEE International Conference on Acoustics, Speech and Signal Processing, March 2010.

N. Goel, S. Thomas, et al. "Approaches to Automatic Lexicon Learning with Limited Training Examples", IEEE International Conference on Acoustics, Speech and Signal Processing, March 2010.

A. Rastrow, A. Sethy, B. Ramabhadran "Constrained Discriminative Training of N-gram Language Model", IEEE Automatic Speech Recognition and Understanding (ASRU) Workshop, December 2009.

A. Rastrow, A. Sethy, B. Ramabhadran, F. Jelinek "Towards Using Hybrid Word and Fragment Units for Vocabulary Independent LVCSR Systems", In Proceedings of Interspeech, September 2009.

C.M. White, **A. Rastrow**, S. Khudanpur, F. Jelinek "Unsupervised Estimation of the Language Model Scaling Factor", In Proceedings of Interspeech, September 2009.

A. Rastrow, A. Sethy, B. Ramabhadran, "A New Method For OOV Detection Using Hybrid Word/Fragment System", IEEE International Conference on Acoustics, Speech and Signal Processing, April 2009.

L. Burget, P. Schwarz, P. Matejka, M. Hannemann, **A. Rastrow**, C. White, S. Khudanpur, H. Hermansky, J. Cernocky, "Combination of Strongly and Weakly Constrained Recognizers for Reliable Detection of OOVs", IEEE International Conference on Acoustics, Speech and Signal Processing, April 2008.

PRESENTATIONS AND INVITED TALKS	- ISI USC, Los Angles, CA, November 2011 - SDL Language Weaver, Los Angles, CA, November 2011 - Center for Language and Speech Processing, Baltimore, MD, November 2011 - NAACL, Los Angeles, CA, June 2010 - Center for Language and Speech Processing, Baltimore, MD, April 2010 - Statistical Machine Learning and Applications Symposium (SMiLe), IBM T.J. Watson Research Center, Yorktown Heights, NY Oct 2009 - Human Language Technology Center of Excellence, Baltimore, MD, February 2009 - IBM T.J. Watson Research Lab, Yorktown Heights, NY, January 2009 - Center for Language and Speech Processing, Baltimore, MD, December 2008 - JHU WS07 Workshop Closing, Baltimore, MD, August 2007

SKILLS

Programming:	C, C++, Java, Perl, Python, UNIX shell scripting
Libraries and Tools:	HTK, Attila (IBM recognizer), MATLAB, SRILM toolkit, Google OpenFST, AT&T FSM Library, LaTeX
Operating Systems:	Unix/Linux, Mac OS X, Windows
Languages:	Fluent in Farsi (Persian) and English, Familiar with Arabic

PROFESSIONAL ACTIVITIES AND SERVICE

- Reviewer for NAACL 2010, ICASSP 2011, EMNLP 2011, INTERSPEECH 2011, ASRU 2011
- Member, Institute of Electrical and Electronics Engineers, Inc.
- Member, Association for Computational Linguistics (ACL)
- Public Relations Chair, Iranian Cultural Society (ICS) of Johns Hopkins University

REFERENCES   Available upon request