Curriculum Vitae

Andreas Stolcke

Microsoft Research
1065 La Avenida
Mountain View, CA 94043
Phone (650) 693-0884

Email andreas.stolcke@microsoft.com
http://research.microsoft.com/en-us/people/anstolck/

International Computer Science Institute
1947 Center St., Suite 600
Berkeley, CA 94704
Phone (510) 666-2969
Fax (510) 666-2956
Email stolcke@icsi.berkeley.edu

General Interests

All aspects of natural language processing, artificial intelligence, machine learning, formal language theory, speech recognition, speaker modeling and recognition, language and dialect recognition. Strong interests in probabilistic and connectionist language modeling, computational and cognitive linguistics, language acquisition, machine translation, historical linguistics; software designs for language processing.

Education

University of California — **Berkeley, CA**

July 1994 — Ph.D., Computer Science, major in Artificial Intelligence, with minors in Theoretical Computer Science and Cognitive Science. Thesis: *Bayesian Learning of Probabilistic Language Models*. Advisor: Prof. J. A. Feldman.

Technische Universität München — **Munich, Germany**

Nov. 1988 — Diplom in Informatik (Computer Science), with minor in Theoretical Linguistics, *summa cum laude*. Thesis: *Natural Language Generation with Unification-based Grammars— A connectionist approach*. Advisor: Prof. W. Brauer.

Research Experience

Microsoft Research — **Mountain View, CA**

Feb. 2011–present — Principal Researcher in the Conversational Systems Research Center (previously, Conversational Systems Lab). Research on addressee detection, conversational understanding, and language modeling.

International Computer Science Institute — **Berkeley, CA**

Feb. 2011–present — External Fellow, conducting research projects in conversational understanding jointly with Microsoft Research.

SRI International — **Menlo Park, CA**

Oct. 1994–Feb. 2011 — Senior Research Engineer in the Speech Technology and Research Laboratory. Research in statistical language, acoustic, pronunciation, and dialog models for speech recognition, translations, and understanding, as well as speaker and language recognition.
PI and co-PI for research projects funded by NSF, DARPA, DoD, and NASA, including: *Robust Automatic Transcription of Speech (RATS)*, *Global Autonomous Language Exploitation (GALE)*, *Effective Affordable Reusable Speech-to-Text (EARS)*, *TalkPrinting (stylistic speaker recognition)*, *Large Vocabulary Conversational Speech Recognition*, *Information Extraction from Speech*, *Hidden Word-Level Events in Spontaneous Speech*, *Modeling Prosodic Duration Dynamics for Conversational Speech Recognition*, and *Harnessing Speech Prosody for Robust Human-Computer Interaction*.

International Computer Science Institute — **Berkeley, CA**

Apr. 2000–Feb. 2011 — Visiting Senior Researcher with the ICSI Speech Research Group, working on joint projects involving both ICSI and SRI, with a particular focus on processing of multi-party meetings.

CLSP, Johns Hopkins University **Baltimore, MD**

Summer 1997 Member of Team "Discourse Language Modeling" at the Workshop on Innovative Techniques for Large Vocabulary Conversational Speech Recognition.

CLSP, Johns Hopkins University **Baltimore, MD**

Summer 1996 Team Leader, "Dependency Language Modeling" at the Workshop on Innovative Techniques for Large Vocabulary Conversational Speech Recognition.

CLSP, Johns Hopkins University **Baltimore, MD**

Summer 1995 Member of Team "Fast Sparse Data Training/Portability" at the Language Modeling Workshop.

International Computer Science Institute **Berkeley, CA**

Jan. 1989– Sep. 1994 Research Assistant with Prof. J. Feldman. Research on probabilistic and connectionist learning algorithms for formal and natural languages. Design and implementation of a flexible, object-oriented testbed system for stochastic language modeling, estimation and Bayesian induction.

Research with the ICSI speech group headed by Prof. N. Morgan on a spoken language speech understanding system (Postdoctoral Researcher since 7/94). Co-developed a practical system for multiple pronunciation word modeling, based on research with Dr. S. Omohundro. Developed various algorithms for stochastic context-free grammars for use in language modeling and recognition.

Technische Universität München **Munich,Germany**

Spring 87 – Fall 88 Research Assistant with Prof. W. Brauer and Dr. C. Freksa. Research on natural language generation. Developed connectionist implementations of unification and unification-based grammars.

Professional Activities

Fellow, IEEE and IEEE Signal Processing Society

Member, International Speech Communication Association

Co-Editor, *Computer Speech and Language* (2003-2007)

Associate Editor, *IEEE Transactions on Speech and Audio Processing* (2000-2002)

Member of Editorial Board for *Computational Linguistics* (1997-1999), and *Free Speech Journal* (1996-1997), and various conferences.

Co-organized technical meetings: *Workshop on Unsupervised Learning in Natural Language Processing* at ACL-99; *Workshop on Spoken Language Understanding for Conversational Systems and Higher Level Linguistic Information for Speech Processing* at HLT-NAACL 2004; *Special Session on Human Language Technology: Applications and Challenges for Speech Processing* at ICASSP 2005.

Reviewer for *Artificial Intelligence*, *Connection Science*, *Neural Network Review*, *Machine Learning*, *Computer Speech and Language*, *Speech Communication*, *Language and Speech*, *IEEE Transaction on Neural Networks*, and numerous conferences.

Served as invited judge on panel evaluating projects at the Johns Hopkins University CLSP Summer Workshop in 2002, 2004, and 2006.

Teaching Experience

ICSI/University of California **Berkeley, CA**

Spring 2009 & 2012 Guest lectures on language modeling and automatic speech recognition in graduate courses taught by Prof. Nelson Morgan.

International Computer Science Institute **Berkeley, CA**

5/93 – 10/93 Co-organized research seminar on Probabilistic Computational Linguistics, lecturing and leading discussions on select topics.

University of California **Berkeley, CA**

Spring 92/93 Guest lectures in courses on Connectionist Computation, by Prof. Feldman and Dr. Regier.

University of California **Berkeley, CA**

Fall 88 Teaching Assistant with Profs. Rowe and Harvey, introductory Computer Science course on Programming Methodologies, teaching discussion sections, grading exams and consulting during lab sections.

IBM Training Center **Herrenberg, Germany**

11/86 – 5/88 Co-designed and taught introductory and advanced courses on the UNIX operating system.

University of Munich **Germany**

9/86 Teaching Assistant, course on Prolog Programming for Natural Language Processing, Computational Linguistics Summer School.

Publications

most available from http://research.microsoft.com/en-us/people/anstolck/ or http://www.speech.sri.com/people/stolcke/publications.html

Journal articles

"The CALO Meeting Assistant System", in *IEEE Trans. Audio Speech Language Processing* 18(6), 1601-1611, August 2010. With G. Tur et al.

"Improving robustness of MLLR adaptation with speaker-clustered regression class trees", in *Computer Speech and Language* 23(2), 176-199, April 2009. With A. Mandal et al.

"Web resources for language modeling in conversational speech recognition", in *ACM Transactions on Speech and Language Processing* 5(1), December 2007. With I. Bulyko et al.

"Morph-based speech recognition and modeling of out-of-vocabulary words across languages", in *ACM Transactions on Speech and Language Processing* 5(1), December 2007, With M. Creutz et al.

"Speaker Recognition with Session Variability Normalization Based on MLLR Adaptation Transforms", in *IEEE Trans. Audio Speech Language Processing* 15(7), 1987-1998, September 2007, With S. Kajarekar et al.

"Recent Innovations in Speech-to-Text Transcription at SRI-ICSI-UW", *IEEE Trans. Audio, Speech and Language Processing* 14(5), 1729-1744, 2006. With B. Chen et al.

"Enriching Speech Recognition with Automatic Detection of Sentence Boundaries and Disfluencies", *IEEE Trans. Audio, Speech and Language Processing* 14(5), 1526-1540, 2006. With Y. Liu, E. Shriberg, D. Hillard, M. Ostendorf, and M. Harper.

"Morphology-Based Language Modeling for Conversational Arabic Speech Recognition", *Computer Speech and Language* 20(4), 589-608, 2006 With K. Kirchhoff, D. Vergyri, J. Bilmes, and K. Duh.

"A Study in Machine Learning from Imbalanced Data for Sentence Boundary Detection in Speech", *Computer Speech and Language* 20(4), 468-494, 2006. With Y. Liu, N. V. Chawla, M. P. Harper, and E. Shriberg.

"Modeling prosodic feature sequences for speaker recognition", *Speech Communication* 46(3-4), 455-472, 2005. With E. Shriberg, L. Ferrer, S. Kajarekar, and A. Venkataraman.

"Modeling Word-level Rate-of-Speech Variation in Large Vocabulary Conversational Speech Recognition", *Speech Communication* 41, 273-285, 2003. With J. Zheng and H. Franco.

"Improved modeling and efficiency for automatic transcription of Broadcast News", *Speech Communication* 37(1-2), 2002. With A. Sankar et al.

"Integrating Prosodic and Lexical Cues for Automatic Topic Segmentation", *Computational Linguistics* 27(1), 2001. With G. Tür et al.

"Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech", *Computational Linguistics* 26(3), 2000. With K. Ries et al.

"Finding consensus in speech recognition: word error minimization and other applications of confusion networks", *Computer Speech and Language* 14(4), 2000. With L. Mangu and E. Brill.

"Prosody-Based Automatic Segmentation of Speech into Sentences and Topics", *Speech Communication* 32(1-2), 2000. With E. Shriberg et al.

"MAESTRO: Conductor of Multimedia Analysis Technologies", *Communications of the ACM* 43(2), 2000. With Z. Rivlin et al.

"Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech?", *Language and Speech* 41(3-4), 1998. With E. Shriberg et al.

"L0—The First Five Years of an Automated Language Acquisition Project", *Artificial Intelligence Review* 10(1-2), 1996. With J. Feldman et al.

"An Efficient Probabilistic Context-Free Parsing Algorithm that Computes Prefix Probabilities", *Computational Linguistics* 21(2), 1995.

"Unification as Constraint Satisfaction in Structured Connectionist Networks", *Neural Computation* 1, 1989.

Refereed conference papers

Published over 150 papers in conference proceedings. A selection of recent and representative papers appears below. For a complete list see http://research.microsoft.com/en-us/people/anstolck/ and http://www.speech.sri.com/people/stolcke/publications.html.

"Learning When to Listen: Detecting System-Addressed Speech in Human-Human-Computer Dialog", Proc. Interspeech, 2012. With E. Shriberg et al.

"Language Modeling of Nonverbal Vocalizations in Spontaneous Speech", *Text, Speech and Dialogue, 15th International Conference*, Springer, 2012. With D. Prylipko et al.

"Effects of audio and ASR quality on cepstral and high-level speaker verification systems", Proc. Odyssey Speaker and Language Recognition Workshop, 2012. With M. Graciarena et al.

"Speaker recognition with region-constrained MLLR transforms", Proc. IEEE ICASSP, 2012. With A. Mandal and E. Shriberg.

"Effective Arabic Dialect Classification Using Diverse Phonotactic Models", Proc. Interspeech, 2011. With M. Akbacak et al.

"Language-independent constrained cepstral features for speaker recognition", Proc. IEEE ICASSP, 2011. With E. Shriberg

"Making the Most from Multiple Microphones in Meeting Recognition", Proc. IEEE ICASSP, 2011.

"Improving Language Recognition with Multilingual Phone Recognition and Speaker Adaptation Transforms", Proc. Odyssey Speaker and Language Recognition Workshop, 2010. With M. Akbacak et al.

"Acoustic Front-End Optimization for Bird Species Recognition", Proc. IEEE ICASSP, 2010. With M. Graciarenaet et al.

"Leveraging Speaker Diarization for Meeting Recognition from Distant Microphones", Proc. IEEE ICASSP, 2010. With G. Friedland et al.

"Exploiting User Feedback for Language Model Adaptation in Meeting Recognition", Proc. IEEE ICASSP, 2009. With D. Vergyri et al.

"The Case for Automatic Higher-Level Features in Forensic Speaker Recognition", Proc. Interpeech, 2008. With E. Shriberg.

"Open-Vocabulary Spoken Term Detection Using Graphone-Based Hybrid Recognition Systems", Proc. IEEE ICASSP, 2008. With M. Akbacak et al.

"Recognizing Arabic Speakers with English Phones", Proc. Odyssey Speaker and Language Recognition Workshop, 2008. With S. Kajarekar.

"Detecting Nonnative Speech Using Speaker Recognition Approaches", Proc. Odyssey Speaker and Language Recognition Workshop, 2008. With E. Shriberg et al.

"Reranking Machine Translation Hypotheses With Structured and Web-based Language Models", Proc. IEEE ASRU Workshop, 2008. With W. Wang et al.

"Detecting Deception Using Critical Segments", Proc. Eurospeech, 2007. With F. Enos et al.

"fMPE-MAP: Improved Discriminative Adaptation for Modeling New Domains", Proc. Eurospeech, 2007. With J. Zheng

"NAP and WCCN: Comparison of Approaches Using MLLR-SVM Speaker Verification System", Proc. IEEE ICASSP, 2007. With S. Kajarekar.

"Unsupervised Language Model Adaptation for Meeting Recognition", Proc. IEEE ICASSP, 2007. With G. Tur.

"Improved Speech Activity Detection Using Cross-Channel Features for Recognition of Multiparty Meetings", Proc. ICSLP, 2006. With K. Boakye.

"Within-Class Covariance Normalization for SVM-based Speaker Recognition", Proc. ICSLP, 2006. With A. O. Hatch and S. Kajarekar.

"Cross-domain and Cross-language Portability of Acoustic Features Estimated by Multilayer Perceptrons", Proc. IEEE ICASSP, 2006. With F. Grezl et al.

"MLLR Transforms as Features in Speaker Recognition", Proc. Eurospeech, 2005. with L. Ferrer et al.

"Improved Phonetic Speaker Recognition Using Lattice Decoding", Proc. IEEE ICASSP, 2005. With A. O. Hatch and B. Peskin.

"Progress in Meeting Recognition: The ICSI-SRI-UW Spring 2004 Evaluation System", Proc. NIST ICASSP Meeting Recognition Workshop, 2004. With C. Wooters et al.

"The Use of a Linguistically Motivated Language Model in Conversational Speech Recognition", Proc. ICASSP, 2004. With W. Wang and M. P. Harper.

"TRAPping Conversational Speech: Extending TRAP/Tandem approaches to conversational telephone speech recognition", Proc. ICASSP, 2004. With N. Morgan et al.

"Getting More Mileage from Web Text Sources for Conversational Speech Language Modeling using Class-Dependent Mixtures", Proc. HLT-NAACL, 2003. With I. Buyko and M. Ostendorf.

"A prosody-based approach to end-of-utterance detection that does not require speech recognition", Proc. ICASSP, 2003. With L. Ferrer et al.

"Prosodic Knowledge Sources for Automatic Speech Recognition", Proc. ICASSP, 2003. With D. Vergyri et al.

"Speaker Recognition using Prosodic and Lexical Features", Proc. IEEE Speech Recognition and Understanding Workshop, 2003. With S. Kajarekar et al.

"Prosody-Based Automatic Detection of Annoyance and Frustration in Human-Computer Dialog", Proc. ICSLP, 2002. With J. Ang et al.

"Building an ASR System for Noisy Environments: SRI's 2001 SPINE Evaluation System", Proc. ICSLP, 2002. With V. R. R. Gadde et al.

"Improved Maximum Mutual Information Estimation Training of Continuous Density HMMs", Proc. EUROSPEECH, 2001. With J. Zheng et al.

"The SRI March 2000 Hub-5 Conversational Speech Transcription System, Proc. NIST Speech Transcription Workshop, College Park, MD, 2000". With H. Bratt et al.

"Modeling the Prosody of Hidden Events for Improved Word Recognition", Proc. EUROSPEECH, 1999. With E. Shriberg et al.

"Combining Words and Prosody for Information Extraction from Speech", Proc. EUROSPEECH, 1999. With D. Hakkani-Tür et al.

"How far do speakers back up in their repairs? A quantitative model", Proc. ICLSP, 1998. With E. Shriberg.

"Efficient Lattice Representation and Generation", Proc. ICSLP, 1998. With F. Weng and A. Sankar.

"Entropy-based Pruning of Backoff Language Models", Proc. DARPA Broadcast News Transcription and Understanding Workshop, Lansdowne, VA, 1998.

"Explicit Word Error Minimization in N-best List Rescoring", Proc. EUROSPEECH, 1997. With Y. Konig and M. Weintraub.

"A Prosody-Only Decision-Tree Model for Disfluency Detection", Proc. EUROSPEECH, 1997. With E. Shriberg and R. Bates.

"A Study on Multilingual Speech Recognition", Proc. EUROSPEECH, 1997. With F. Weng et al.

"Neural-Network Based Measures of Confidence for Word Recognition", Proc. ICASSP, 1997. With M. Weintraub et al.

"Acoustic Modeling for the SRI Hub4 Partitioned Evaluation Continuous Speech Recognition System", Proc. DARPA Speech Recognition Workshop, Chantilly, VA, 1997 With A. Sankar et al.

"Hub4 Language Modeling Using Domain Interpolation and Data Clustering", Proc. DARPA Speech Recognition Workshop, Chantilly, VA, 1997. With F. Weng and A. Sankar.

"Statistical language modeling for speech disfluencies", Proc. ICASSP-96. With E. Shriberg.

"Automatic linguistic segmentation of conversational speech", Proc. ICLSP-96. With E. Shriberg.

"Word predictability after filled pauses: A corpus-based study", Proc. ICLSP-96. With E. Shriberg.

"Using a Stochastic Context-Free Grammar as a Language Model for Speech Recognition", Proc. ICASSP-95. With D. Jurafsky et al.

"Inducing Probabilistic Grammars by Bayesian Model Merging", in *Grammatical Inference and Applications*, R. C. Carrasco and J. Oncina, eds., Springer, 1994. With S. Omohundro.

"Multiple-pronunciation Lexical Modeling in a Speaker-independent Speech Understanding System", Proc. ICSLP, 1994. With C. Wooters.

"Precise n-gram Probabilities from Stochastic Context-free Grammars", Proc. ACL-94. With J. Segal.

"Hidden Markov Model Induction by Bayesian Model Merging", Advances in Neural Information Processing Systems 5, S. J. Hanson, J. D. Cowan and C. L. Giles (eds.), Morgan Kaufman, 1993. With S. Omohundro.

"Gapping and Frame Semantics: A fresh look from a cognitive perspective", Proc. 13th Intl. Conf. Computational Linguistics (COLING), Helsinki, August 1990.

"Miniature Language Acquisition: A touchstone for cognitive science", in Proc. 12th Annual Conf. Cognitive Science Soc., Cambridge, MA, July 1990. With J. Feldman, G. Lakoff and S. H. Weber.

Other publications

"The SRI-ICSI Spring 2007 Meeting and Lecture Recognition System", in R. Stiefelhagen et al. (eds.), *Multimodal Technologies for Perception of Humans. International Evaluation Workshops CLEAR 2007 and RT 2007*, Lecture Notes in Computer Science 4625, Springer, 2008. With K. Boakye et al.

"Prosody Modeling for Automatic Speech Recognition and Understanding", in M. Johnson, S. Khudanpur, M. Ostendorf, and R. Rosenfeld (eds.), *Mathematical Foundations of Speech and Language Processing*, IMA Volumes in Mathematics and its applications, Springer, 2004. With E. Shriberg.

"Language Modeling for Multilingual Speech Translation", in M. Rayner et al. (eds.), *The Spoken Language Translator*, pp. 250-264, Cambridge University Press, 2000. With F. Weng and M. Cohen.

"Linguistic Knowledge and Empirical Methods in Speech Recognition", *AI Magazine* 18(4), 1997.

Invited Talks and Lectures

Markovian Combination of Language and Prosodic Models for Better Speech Understanding and Recognition, at the IEEE Automatic Speech Recognition and Understanding Workshop, Madonna di Campiglio, Italy, December 2001.

Invited talk and paper on *Prosody Modeling for Automatic Speech Understanding: An Overview of Recent Research at SRI*, at the ISCA Tutorial and Research Workshop on Prosody in Speech Recognition and Understanding, Red Bank, NJ, October 2001.

Hidden Events for Segmentation and Recognition, IBM Watson Research Center, Yorktown Heights, NY, January 2000.

Hidden Events for Segmentation and Recognition, NASA/RIACS, Moffet Field, CA, November 1999.

Two invited talks on *Fast Training and Portability* and on *Dependency Language Modeling*, at Intl. Conf. Spoken Language Processing, Philadelphia, October 1996.

Probabilistic Models in Speech Recognition, guest lecture for CS354: Probabilistic Reasoning in Computing, Stanford University, November 1995.

Bayesian Model Merging for Probabilistic Grammar Induction, Computational Learning Seminar, CSLI, Stanford University, May 1995.

Learning Probabilistic Grammars by Model Merging, Linguistics Department, U.C. Berkeley, April 1994.

Earley-parsing with stochastic context-free grammars, Probability in Computational Linguistics Seminar, ICSI, Berkeley, CA, August 1993.

Connectionist approaches to language processing, guest lecture, Computer Science Division, UC Berkeley, March 1993.

Vector Space Grammars for Learning Syntactic Categories, invited talk, Department of Psychology, UC Davis, August 1991.

Learning Feature Semantics in Simple Recurrent Networks, ICSI, Berkeley, CA, September 1990.

Patents

"Method and apparatus for tailoring the output of an intelligent automated assistant to a user", G. Tur et al., US 2012/0166365 A1.

"Method and apparatus for adding new vocabulary to interactive translation and Dialogue Systems", K. Precoda et al., US 2012/0029904 A1.

"Method and apparatus for speaker recognition", Harry Bratt et al., US 2008/0010065 A1.

"Method and apparatus for fusion of recognition results from multiple types of data sources", G. K. Myers et al., US 2005/0125224 A1.

Software

Author of SRILM, a widely used open-source toolkit for statistical language modeling. http://www.speech.sri.com/projects/srilm/ . For more information see:

"SRILM at Sixteen: Update and Outlook", Proc. IEEE ASRU Workshop, 2011. With J. Zheng et al.

"SRILM—An Extensible Language Modeling Toolkit", Proc. ICSLP, 2002.

Awards

IBM Graduate Fellowship, University of California, 1989-1992.

Computer Speech and Language 2003 Paper Award, for "Finding consensus in speech recognition: word error minimization and other applications of confusion networks". With L. Mangu and E. Brill.

ICASSP 2005 Student paper award for "Improved Phonetic Speaker Recognition Using Lattice Decoding". With A. O. Hatch and B. Peskin.