

Srinivas Bangalore

AT&T Labs - Research
Bldg 103, Room E149,
180 Park Avenue,
Florham Park, NJ 07932

E-mail: srini@research.att.com
Work: (973)-360-7041
Home: (973)-984-7420
Fax: (973)-360-8092

EDUCATION

Ph.D., University of Pennsylvania, Philadelphia, PA
Department of Computer and Information Science, May 1997.
Title: Complexity of Lexical Descriptions: Relevance to Partial Parsing.

M.Tech., Indian Institute of Technology, Kanpur, India
Department of Computer Science and Engineering, May 1991.

B.Tech., Osmania University, Hyderabad, India
Department of Computer Science and Engineering, May 1989.

PROFESSIONAL EXPERIENCE

Visiting Associate Professor: Princeton Institute for International and Regional Studies.
Spring 2009: Teaching graduate and senior undergraduate course on *Introduction to Machine Translation*.

Adjunct Associate Professor: Dept. of Computer Science, Columbia University, New York, NY.

Fall 2005: Teaching graduate and senior undergraduate course on Natural Language Processing.
Course Website: <http://www1.cs.columbia.edu/cs4705/>

Principal Technical Staff Member: SIPS, AT&T Research Labs, NJ, April 2002-Present.

Senior Technical Staff Member: SIPS, AT&T Research Labs, NJ, October 1997-April 2002.
My research interests are in combining the flexibility provided by linguistically motivated representations for language with the robustness provided by data driven methods for natural language processing tasks. This theme is reflected in my work on SuperTagging and in the projects I am currently leading.

- Project Anuvaad: Investigating automatic methods for bootstrapping statistical translation models from bilingual texts and bilingual corpora in the context of speech-to-speech translation systems with English as the source language and Spanish, Chinese and Japanese as target languages (<http://www.research.att.com/srini/Projects/Anuvaad/home.html>). Real time broadcast news English-to-Spanish translation at (<http://slshow.research.att.com:8080/ecrf/cgi-bin/bnrtdemo.jsp>).
- Project FERGUS: Concerned with rapid prototyping of components for natural language generation, in particular, a surface realizer, using statistical models trainable from corpora. (<http://www.research.att.com/srini/Projects/Fergus/home.html>)

- **Project MATCH:** Exploring methods for tight integration of speech and gesture information to yield a combined interpretation using finite-state multimodal language understanding techniques in the context of a multimodal application for city help.
(<http://www.research.att.com/info/Projects/MultimodalAccessToCityHelp>)
- **Project WebTalk:** Mining the unstructured information available on the web to automatically build text and spoken dialog applications.
(<http://www.research.att.com/info/Projects/WebTalk>)

Postdoctoral Fellow: CIS Dept., University of Pennsylvania, PA, May 1997-October 1997. Used the Supertagger in a variety of natural language applications such as Information Retrieval, Information Extraction and Machine Translation. Also ported the Supertagger to French.

Team Member: Lexis-Nexis Project, University of Pennsylvania, PA, June 1996-August 1996. Utilized Robust Partial Parser (Supertagger) for the purpose of information extraction from data provided by Lexis-Nexis company. The Supertagger was developed as part of my doctoral thesis.

Team Member: MUC Project, University of Pennsylvania, PA August 1995 - October 1995. Designed and implemented modules for the University of Pennsylvania Coreference Resolution System that stood third at the Sixth Message Understanding Conference.

Research Assistant: XTAG Project, University of Pennsylvania, PA, August 1991 - May 1997.

Designed, implemented and maintained components of XTAG system, a wide-coverage Lexicalized Tree-Adjoining Grammar System, under the direction of Prof. Aravind Joshi. Implemented statistical and linguistic filters to improve the performance of the parser. Performed comparative evaluation of the XTAG system against other wide-coverage grammar systems.

Teaching Assistant: Dept. of CIS, University of Pennsylvania, PA, August 1992 - April 1993. Taught a first-level course on programming for undergraduates. The course involved teaching SCHEME, ML and C++.

M.Tech. Thesis: Indian Institute of Technology, India, January 1990 - May 1991. Implemented a Machine Translation system in the Paninian formalism to translate between Hindi and Telugu. Also developed a Linguist's Workbench in C on Unix using Sunview.

B.Tech. Project: Osmania University, India, January 1989 - May 1989. Designed and developed a Natural Language Interface to a INFORMIX RDBMS based Library Management System, using PROLOG on DOS.

Programmer: MCEME, Hyderabad, India, May 1987 - August 1988. Designed and developed a vehicle simulator on IBM-386 in Pascal 4.0, coupled with graphical displays for scene management. Used as a training program at Military College of Electrical and Mechanical Engineers (MCEME), Hyderabad, India.

MENTORING and TEACHING EXPERIENCE

Organized, developed course material and co-taught a two week project-oriented course on NLP at IIIT, Hyderabad, India, December 2007.

Member of the AT&T Labs Fellowship committee chartered to provide financial support and mentoring advice to PhD students.

AT&T Labs Fellowship Mentor for Vanessa Murdock, PhD student, University of Massachusetts.

PhD Committee Member for John Chen, University of Delaware, 2001.

Organized, developed course material and co-taught a three week course on NLP at Anna University, Chennai, India, December 2001.

Supervisor for summer interns since 1999.

RESEARCH INTERESTS

Parsing Methodologies, Formal Grammars, Statistical Language Processing, Conversational Systems, Machine Learning and Data Mining, Information Retrieval, Machine Translation, Information Theory.

HONORS and AWARDS

AT&T Outstanding New Mentor Award, in recognition of outstanding support and dedication to AT&T Labs Mentoring Program, 2001.

Morris and Dorothy Rubinoff award for “outstanding dissertation that has resulted in or could lead to innovative applications of computer technology”, University of Pennsylvania, 1998.

Institute for Research in Cognitive Science Fellowship, University of Pennsylvania, 1994-1996.

Dean’s Fellowship, University of Pennsylvania, PA, 1991 - 1992.

Gold Medal for standing University First in B.Tech, Osmania University, India, 1989.

Gold Medal for standing First in B.Tech in the state of Andhra Pradesh, India, 1989.

National Merit Scholarship, India, 1985 - 1989.

SERVICE and PROFESSIONAL ACTIVITIES

Special Issue Editor on Spoken Language Understanding, Speech Communication Journal, 2005.

Co-organizer, Spoken Language Understanding Workshop, AAAI 2005.

General Workshop Chair, ACL 2004.

Co-organizer, Spoken Language Understanding Workshop, NAACL 2004.

NSF Proposal Reviewer, 2003, 2004.

Editorial Board Member, Journal of Computational Linguistics, 2001-2003.

Program Committee Member, ASRU 2005, NAACL 2004, NAACL 2003, ACL 97, ICON Conferences, TAG+ workshops.

Reviewer, Journal of Computational Linguistics, Journal of Natural Language Engineering, IEEE Transactions on Speech and Audio Processing.

Reviewer for Association of Computational Linguistics (ACL), EACL, COLING and AAAI conferences since 1996.

Member, Association for Computational Linguistics

Member, Association for Computing Machinery

Chairman of the Student Chapter of Computer Society of India at Hyderabad, 1988-1989, during which the branch was awarded the Best Student Branch Award.

Secretary of the Student Chapter of Computer Society of India at Hyderabad, 1987-1988.

PATENTS

Granted

US Patent# 6317707: Automatic Clustering Of Tokens From A Corpus For Grammar Acquisition

US Patent# 6415248: Method For Building Linguistic Models From A Corpus

US Patent# 6751584: Automatic Clustering Of Tokens From A Corpus For Grammar Acquisition

US Patent# 6868383: Systems and methods for extracting meaning from multimodal inputs using finite-state devices

US Patent# 6947885: Probabilistic model for natural language generation

US Patent# 7069215: Systems and methods for extracting meaning from multimodal inputs using finite-state devices

US Patent# 7113903: Method and apparatus for providing stochastic finite-state machine translation

US Patent# 7225131: System and method for accessing and annotating electronic medical records from multi-modal and unimodal inputs

US Patent# 7231341: System and method for natural language generation

US Patent# 7257575: Systems and methods for generating markup-language based expressions from multi-modal and unimodal inputs

US Patent# 7295975: Systems and methods for extracting meaning from multimodal inputs using finite-state devices

BIBLIOGRAPHY

Tutorials

“Multimodality in Human-Computer Interfaces”, with M. Johnston, ICASSP, Honolulu, Hawaii, 2007.

“Multimodal Language Processing”, with M. Johnston, ACL-Coling, Sydney, Australia, 2006.

BOOKS

“Complexity of Lexical Descriptions and its Relevance to Natural Language Processing: A SuperTagging Approach”, with Aravind Joshi (eds), MIT Press, expected Spring 2009.

BOOK CHAPTERS

“Statistical Machine Translation through Global Lexical Selection and Sentence Reconstruction”, with S. Kanthak and P. Haffner. In *Learning Machine Translation*, C. Goutte, N. Cancedda, M. Dymetman and G. Foster, MIT Press, 2008.

“Dependency and Valency in other theories: Tree Adjoining Grammar”, with O. Rambow and A.K. Joshi. In *Dependency and Valency: An International Handbook of Contemporary Research*, V. Agel, L.M. Eichinger, H.W. Eroms et. al.(eds.), Portico Publications, 2003.

“Localizing Dependencies and SuperTagging”, In *Data-Oriented Parsing*, Rens Bod, Remko Scha and Khalil Sima'an (eds.), CSLI Publications, 2003.

“A computational model of the grammatical aspects of word recognition as supertagging”, with A.E. Kim and John Trueswell. In *The Lexical Basis of Sentence Processing: Formal, computational and experimental issues*, Paola Merlo and Suzanne Stevenson (eds.), John Benjamins, 2002.

“Glean: Using Syntactic Information in Document Filtering”, with R. Chandrasekar, *Encyclopedia of Library and Information Science, Volume 71*, A. Kent (ed). Marcel Dekker, 2002.

“Glean: Using Syntactic Information in Document Filtering”, with R. Chandrasekar, *Encyclopedia of Microcomputers, Volume 28*, A. Kent and J.G. Williams (eds). Marcel Dekker, 2002.

“Performance Evaluation of Supertagging for Partial Parsing”, *Advances in Probabilistic and Other Parsing Technologies*, Harry Bunt and Anton Nijholt (eds.), Kluwer, 2000.

“Developing a Wide-Coverage CCG System” with C. Doran, *CSLI Lecture Notes*, A. Abeille and O. Rambow (eds.), 2000.

“Status of the XTAG System” with C. Doran, B.A. Hockey, A. Sarkar and F. Xia *CSLI Lecture Notes*, A. Abeille and O. Rambow (eds.), 2000.

“Explanation-based Learning and Finite State Transducers: Application for Parsing Lexicalized Tree-Adjoining Grammars”, In *Extended finite state models of language*, Andrs Kornai, editor, Cambridge University Press, 1999.

“Using Parsed Corpora for Circumventing Parsing”, with Aravind Joshi, in *Symbolic, Connectionist and Statistical Approaches to Learning for Natural Language Processing*, S. Wermter, E. Riloff and G. Scheler (Eds), Springer Verlag, 1996.

JOURNAL PUBLICATIONS

“Robust Multimodal Language Processing”, with M. Johnston, accepted *Computational Linguistics*, 2009.

“Discriminative Machine Translation using Global Lexical Selection”, with S. Venkatapathy, *ACM TALIP*, accepted for publication, 2009.

“Combining lexical, syntactic and prosodic cues for improved online dialog act tagging”, with V. Rangarajan and S. Narayanan, *Computer Speech and Language*, accepted for publication, 2008.

“Learning the Structure of Task-driven Human-Human Dialogs”, with A. Stent and G. diFabrizio, *IEEE Transactions on Audio, Speech and Language*, 16(7), 2008.

“Exploiting acoustic and syntactic features for automatic prosody labeling in a maximum entropy framework”, with V. Rangarajan and S. Narayanan, *IEEE Transactions on Audio, Speech and Language*, 16(4), 2008.

“Introduction to the Special Issue on Spoken Language Understanding in Conversational Systems”, with D. Hakkani-Tur and G. Tur, In *Journal of Speech Communication Special Issue on Spoken Language Understanding*, Vol. 48, No. 3-4, pp. 233-238, 2006.

“Automated Extraction of Tree-Adjoining Grammars from Treebanks”, with J. Chen and K. Vijayshanker, In *Journal of Natural Language Engineering*, 2005.

“The AT&T Spoken Language Understanding System”, with N.Gupta, D.Tur, G.Tur, G.Riccardi, M.Rahim, In *IEEE Transactions on Speech and Audio Processing*, 2005.

“Finite-state Multimodal Integration and Understanding”, with Michael Johnston, In *Journal of Natural Language Engineering*, Cambridge University Press, 2005.

“Extracting Clauses for Spoken Language Understanding in Conversational Systems”, with Narendra Gupta, In *TAL Journal, Special Issue on Spoken corpus processing*, Hermes Science, 2004.

“Stochastic Finite-State Models for Spoken Language Machine Translation”, with Giuseppe Riccardi, In *Machine Translation*, Kluwer Publications, Vol. 17, No. 3, 2002.

“A Lightweight Dependency Analyzer for Partial Parsing”, In *Natural Language Engineering*, Cambridge University Press, Vol 6:113-138, 2000.

“Head Transducer Models for Speech Translation and their Automatic Acquisition from Bilingual Data”, with H. Alshawi and S. Douglas, In *Machine Translation*, 2000.

“Learning Dependency Translation Models as Collections of Finite State Head Transducers”, with H. Alshawi and S. Douglas, In *Computational Linguistics*, Vol 26, No. 1, March 2000, MIT Press.

“Supertagging: An Approach to Almost Parsing”, with A.K. Joshi, In *Computational Linguistics*, Vol 25, No. 2, June 1999, MIT Press.

“Glean: Using Syntactic Information in Document Filtering”, with R. Chandrasekar *Information Processing and Management*, Vol. 34, No. 5, (1998), Elsevier Science.

“Automatic Induction of Rules for Text Simplification”, with R. Chandrasekar *Knowledge Based Systems*, Vol. 10 (1997), Elsevier Science.

“Explanation-based Learning and Finite State Transducers : Applications to Parsing Lexicalized Tree Adjoining Grammars”, *Journal of Natural Language Engineering*, Vol.2, No. 4, Cambridge University Press.

CONFERENCE PUBLICATIONS

“Role of Natural Language Understanding in Voice Local Search”, with J. Feng and M. Gilbert, *Interspeech 2009*, Brighton, UK, 2009.

“MICA: A Probabilistic Dependency Parser Based on Tree Insertion Grammars”, with P. Boullier, A. Nasr, O. Rambow and B. Sagot, *NAACL-HLT 2009*, Boulder, Colorado, 2009.

“Tightly coupling Speech Recognition and Search”, with T. Mishra, *NAACL-HLT 2009*, Boulder, Colorado, 2009.

“Incremental Parsing Models for Dialog Task Structure”, with A. Stent, *European ACL 2009*, Athens, Greece, 2009.

“Effects of Word Confusion Networks on Voice Search”, with J. Feng, *European ACL 2009*, Athens, Greece, 2009.

“Query Parsing for Voice-enabled Mobile Local Search”, with J. Feng, *ICASSP*, Taipei, Taiwan, 2009.

“Incorporating discourse context in spoken language translation through dialog acts”, with S. Rangarajan and S. Narayanan, *IEEE SLT 2008*, Goa, India, 2008.

“Statistical Shared Plan-Based Dialog Management”, with A. Stent, *Interspeech 2008*, Brisbane, Australia, 2008.

“Factored translation models for enriching spoken language translation with prosody”, *Interspeech 2008*, Brisbane, Australia, 2008.

“Enriching spoken language translation with dialog acts”, with V. Rangarajan and S. Narayanan, *ACL*, Columbus, Ohio, 2008.

“Referring Expression Generation Using Speaker-based Attribute Selection and Trainable Realization”, with A. Stent and G. diFabrizio, *INLG*, Salt Fork, OH, 2008.

“Modeling the Intonation of Discourse Segments for Improved online Dialog Act Tagging”, with V. Rangarajan and S. Narayanan, *ICASSP*, Las Vegas, 2008.

“Where Do The Words Come From? Learning Models for Lexical Selection and Word Ordering from Spoken Dialog Corpora”, with A. Stent and G. diFabrizio, *ICASSP*, Las Vegas, 2008.

“Exploiting prosodic features for dialog act tagging in a discriminative modeling framework”, with V. Rangarajan and S. Narayanan, *Interspeech*, Antwerp, 2007.

“Statistical Machine Translation through Global Lexical Selection and Sentence Reconstruction”, with S. Kanthak and P. Haffner, *ACL*, Prague, Czech Republic, 2007.

“Exploiting acoustic and syntactic features for prosody labeling in a maximum entropy framework”, with V. Rangarajan and S. Narayanan, *HLT-NAACL*, Rochester, 2007.

“Acoustic-Syntactic Maximum Entropy Model for Automatic Prosody Labeling”, with V. Rangarajan and S. Narayanan, *IEEE-ACL Conference on Spoken Language Technology*, Aruba, 2006.

“Sequence Classification for Machine Translation”, with P. Haffner and S. Kanthak, *ICSLP*, Pittsburgh, USA, 2006.

“Learning the Structure of Task-driven Human-Human Dialogs”, with G. Di Fabbrizio and A. Stent, *ACL-Coling*, Sydney, Australia, 2006.

“Edit Machines for Robust Multimodal Language Processing”, with M. Johnston, *EACL*, Trento, Italy, 2006.

“Learning Edit Machines for Robust Multimodal Understanding”, with M. Johnston, *ICASSP*, Toulouse, France, 2006.

“Towards Learning to Converse: Structuring Task-Oriented Human-Human Dialogs”, with G. di Fabbrizio and A. Stent, *ICASSP*, Toulouse, France, 2006.

“Automatic Diacritization of Arabic Transcripts for Automatic Speech Recognition”, with S. Ananthakrishnan and S. Narayanan, *International Conference on Natural Language Processing*, Kanpur, India, 2005.

“Classification of Structured Descriptions”, with O. Rambow, *ICASSP*, Philadelphia, PA, 2005.

“Supertagging Nodes of a Dependency Tree”, with R. Bhimavarapu and O. Rambow, *International Conference on Natural Language Processing*, Hyderabad, India, 2004.

“Question-Answering in WebTalk: An Evaluation Study”, with J. Feng and M. Rahim, *ICSLP 2004*, Korea, 2004.

“Compiling Boostexter Rules into a Finite-state Transducer”, *Association of Computational Linguistics (ACL-2004)*, Barcelona, Spain, 2004.

“MATCHkiosk: A Multimodal Interactive City Guide”, with M. Johnston, *Association of Computational Linguistics (ACL2004)*, Barcelona, Spain, 2004.

“Question-answering in webtalk: an evaluation study”, with J. Feng and M. Rahim, *In Proceedings of Interspeech*, 2004.

“Balancing Data-driven and Rule-based Approaches in the Context of a Multimodal Conversational System”, with M. Johnston, *NAACL 2004*, Boston, 2004.

“Robust Multimodal Understanding”, with M. Johnston, *ICASSP 2004*, Montreal, Canada, 2004.

“Creating a Finite-State Parser with Application Semantics”, with O. Rambow, T. Butt, A. Nasr and R.Sproat, *International Conference on Computational Linguistics (COLING 2002)*, Taipei, Taiwan, 2002.

“Towards Automatic Generation of Natural Language Generation Systems”, with J. Chen, O. Rambow and M. Walker, *International Conference on Computational Linguistics (COLING 2002)*, Taipei, Taiwan, 2002.

“Bootstrapping Bilingual Data using Consensus Translation for a Multilingual Instant Messaging System”, with V. Murdock and G. Riccardi, *International Conference on Computational Linguistics (COLING 2002)*, Taipei, Taiwan, 2002.

“MATCH: An Architecture for Multimodal Dialogue System”, with M. Johnston, et.al, *Association of Computational Linguistics (ACL 2002)*, Philadelphia, 2002.

“Extracting Clauses for Spoken Language Understanding in Conversational Systems”, with N. Gupta and M. Rahim, *International Conference on Speech and Language Processing 2002 (ICSLP-2002)*, Colorado, 2002.

“A Finite-State Approach to Machine Translation”, with G. Riccardi, *North American Association of Computational Linguistics (NAACL 2001)*, Pittsburgh, USA, May 2001.

“Impact of Quality and Quantity of Corpora on Stochastic Generation”, with J. Chen and O. Rambow, *Empirical Methods in Natural Language Processing*, Pittsburgh, May 2001.

“Finite-State Models for Lexical Reordering in Spoken Language Translation”, with G. Riccardi, *International Conference on Speech and Language Processing 2000 (ICSLP-2000)*, Beijing, China, October 2000.

“Tight-coupling of Multimodal Language Processing with Speech Recognition”, with M. Johnston, *International Conference on Speech and Language Processing 2000 (ICSLP-2000)*, Beijing, China, October 2000.

“Corpus-Based Lexical Choice in Natural Language Generation”, with O. Rambow, *Association of Computational Linguistics (ACL 2000)*, Hongkong, China, October 2000.

“Exploiting a Probabilistic Hierarchical Model for Generation”, with O. Rambow, *International Conference on Computational Linguistics (COLING 2000)*, Saarbrucken, Germany, August 2000.

“Finite-state Multimodal Parsing and Understanding”, with M. Johnston, *International Conference on Computational Linguistics (COLING 2000)*, Saarbrucken, Germany, August 2000.

“Evaluation Metrics for Generation”, with O. Rambow and S. Whittaker, *International Conference on Natural Language Generation (INLG 2000)*, Mitzpe Ramon, Israel, July 2000.

“New Models for Improving Supertag Disambiguation”, with J. Chen and K. Vijay-Shanker, *European Association of Computational Linguistics (EACL 99)*, Bergen, Norway, June 1999.

“Learning Phrase-based Head Transduction Models for Translation of Spoken Utterances”, with H. Alshawi and S. Douglas, *International Conference on Speech and Language Processing 98 (ICSLP-98)*, Sydney, Australia, October 1998 .

“Automatic Acquisition of Hierarchical Transduction Models for Machine Translation”, with H. Alshawi and S. Douglas, *Proceedings of the 36th Conference of Association of Computational Linguistics (ACL 98)*, Montreal, Canada, August 1998.

“Using Supertags in Document Filtering: The Effect of Increased Context on Information Retrieval Effectiveness.”, with R. Chandrasekar, *Recent Advances in NLP’97 (RANLP’97)*, Tzigov Chark, Bulgaria, September 1997.

“Incremental Processing Using Lexicalized Tree-Adjoining Grammar: Symbolic and Connectionist Approaches”, with Albert Kim and John Trueswell, *Conference on Computational Psycholinguistics*, Berkeley, California, August 1997.

“EAGLE: An Extensible Architecture for General Linguistic Engineering”, with B. Baldwin, C. Doran, J.C. Reynar, M. Niv and M. Wasson, *RIAO’97*, Montreal, Quebec, June 1997.

“Using Syntactic Information in Document Filtering: A Comparative Study of Part-of-speech Tagging and Supertagging”, with R. Chandrasekar, *RIAO’97*, Montreal, Quebec, June 1997.

“Automatic Induction of Rules for Text Simplification”, with R. Chandrasekar, *International Conference on Knowledge-based Computer Systems*, Mumbai, India, December 1996.

“Almost Parsing Technique for Language Modeling”, *International Conference on Speech and Language Processing 96 (ICSLP-96)*, Philadelphia, PA, USA, October 1996.

“Motivations and Methods for Text Simplification”, with R. Chandrasekar and C. Doran, *International Conference on Computational Linguistics (COLING 96)*, Copenhagen, August 1996.

“Exploiting Supertag Representation for Fast Coreference Resolution”, with B. Baldwin, *International Conference on NLP+IA/TAL+AI 96*, Moncton, NB, Canada, June 1996.

“Novel Applications of Explanation-based Learning for Parsing Lexicalized Tree-Adjoining Grammars”, with Aravind Joshi, *Proceedings of the 33rd Conference of Association of Computational Linguistics (ACL 95)*, MIT, Boston, June 1995.

“Lexicalization and Grammar Development”, with C. Doran, D. Egedi and T. Becker, *KONVENS 94*, University of Vienna, Vienna, Austria, September 1994.

“Integration of Structural and Statistical Information: Role of Complexity of Description of Primitives”, with Aravind Joshi, *New Methods in Language Processing*, UMIST, Manchester, United Kingdom, September 1994.

“XTAG System – A Wide Coverage Grammar for English”, with C. Doran, D. Egedi, B.A. Hockey and M. Zaidel, *International Conference on Computational Linguistics (COLING 94)*, Kyoto University, Kyoto, Japan, August 1994.

“Disambiguation of Super Parts of Speech (or Supertags): Almost Parsing”, with Aravind Joshi, *International Conference on Computational Linguistics (COLING 94)*, Kyoto University, Japan, August 1994.

“Feature-based TAG in place of Multi-Component Adjunction: Computational Implication”, with B.A. Hockey, *Natural Language Processing Pacific Rim Symposium*, Fukuoka Institute of Technology, Fukuoka, Japan, December 1993.

WORKSHOP PUBLICATIONS

“Three models for discriminative machine translation using Global Lexical Selection and Sentence Reconstruction”, with S. Venkatapathy, *HLT-NAACL workshop*, Rochester, USA, 2007.

“Finite-State Transducer-based Statistical Machine Translation using Joint Probabilities”, with P. Haffner and S. Kanthak, *IWSLT*, Kyoto, Japan, 2006.

“Combining Stochastic and Grammar-based Methods with Finite-state Edit Machines”, with M. Johnston, *ASRU-2005*, San Juan, Puerto Rico, 2005.

“WebTalk: Mining Websites for Automatically Building Dialog Systems”, with J. Feng and M. Rahim, *ASRU*, December 2003.

“Segmenting Spoken Language Utterances into Clauses for Semantic Classification”, with N. Gupta, *ASRU*, December 2003.

“Balancing data-driven and rule-based approaches in the context of a Multimodal Conversational System”, with M. Johnston, *ASRU*, December 2003.

“Extracting Clauses for Spoken Language Understanding in Conversational Systems”, with N. Gupta, *Empirical Methods in Natural Language Processing 2002 (EMNLP-2002)*, Philadelphia, June 2002.

“Reranking an N-gram SuperTagger”, with J.Chen, M. Collins, O.Rambow, *Sixth International Workshop on Tree-Adjoining Grammars (TAG+6)*, Italy, May 2002.

“Context-free parsing of a Tree-Adjoining Grammar using Finite-State Machines, with A. Nasr, O. Rambow, and J.Chen, *Sixth International Workshop on Tree-Adjoining Grammars (TAG+6)*, Italy, May 2002.

“MATCH: Multimodal Access To City Help”, with M. Johnston, *Workshop on Automatic Speech Recognition and Understanding*, Italy, December 2001.

“Computing Consensus Translation from Multiple Machine Translation Systems”, with G. Riccardi and G. Bordel, *Workshop on Automatic Speech Recognition and Understanding*, Italy, December 2001.

“Finite-state Methods for Multimodal Parsing and Integration”, with M. Johnston, In *Finite State Methods in Natural Language Processing*, Helsinki, August 2001.

“Natural Language Generation in Dialog Systems”, with O. Rambow and M. Walker, In *Workshop on Human Language Technology (HLT)*, San Diego, March 2001.

“Using TAG, a Tree Model, and a Language Model for Generation”, with O. Rambow, *Fifth International Workshop on Tree-Adjoining Grammars (TAG+)*, TALANA, Paris, May 2000.

“Stochastic Finite-State models for Spoken Language Machine Translation”, with G. Riccardi, *Workshop on Embedded Machine Translation Systems*, Seattle, April 2000.

“Learning Head-dependency Relations From Unannotated Corpora”, with G. Riccardi and P. Sarin, *Automatic Speech Recognition and Understanding Workshop*, Keystone, Colorado, December 1999.

“Automatic Acquisition of Phrase Grammars for Stochastic Language Modeling”, with G. Riccardi, *Sixth Workshop on Very Large Corpora*, Montreal, Canada, August 1998.

“Transplanting Supertags from English to Spanish”, *Workshop on Tree-Adjoining Grammars*, Philadelphia, August 1998.

“Grammar & Parser Evaluation in the XTAG Project”, with C. Doran, B.A. Hockey and A. Sarkar, *Workshop on the Evaluation of Parsing Systems*, Granada, Spain, May 1998.

“Catering to the needs of Web users: Integrating Retrieval and Browsing”, with Z. Lacroix, A. Sahuguet and R. Chandrasekar, *Workshop on Conceptual Modelling in Multimedia Information Seeking*, ER97, Los Angeles, California, November 1997.

“Performance Evaluation of SuperTagging for Partial Parsing”, *Fifth International Workshop on Parsing Technologies*, Boston, September 1997.

“Catering to the Web User’s Needs: Integrating Retrieval and Browsing”, with Z. Lacroix and A. Sahuguet, *Third International Workshop on Multimedia Information Systems*, Como, Italy, September 1997.

“Mother of Perl: A Multi-tier Pattern Description Language”, with B. Baldwin, C. Doran, J.C. Reynar and M. Niv, *International Workshop on Lexically Driven Information Extraction*, Frascati, Italy, July 1997.

“Maintaining the forest and burning out the underbrush in XTAG”, with C. Doran, B. Hockey, P. Hopely, J. Rosenzweig, A. Sarkar, F. Xia, A. Nasr and O. Rambow, *Workshop on Computational Environments for Practical Grammar Development (ENVGRAM ’97)*, Madrid, July 1997.

“Gleaning information from the Web: Using Syntax to Filter out Irrelevant Information”, with R. Chandrasekar, *Workshop on Natural Language Processing for the World Wide Web*, AAAI Spring Symposium, Stanford University, California, March 1997.

“An Approach to Robust Partial Parsing and Evaluation Metrics”, with C. Doran, B.A. Hockey, and A.K. Joshi, *Workshop on Robust Parsing, European Summer School in Logic, Language and Information (ESSLLI)*, August 1996.

“Explanation-Based Learning and Finite State Transducers: Applications to Parsing Lexicalized Tree-Adjoining Grammars”, *Workshop on Extended Finite State Models of language*, Budapest, Hungary, August 1996.

“University of Pennsylvania: Description of the University of Pennsylvania System used for MUC-6”, with B. Baldwin, J.C. Reynar, M. Collins, J. Eisner, A. Ratnaparkhi, J. Rosenzweig and A. Sarkar, *Sixth Message Understanding Conference (MUC-6)*, Maryland, October 1995.

“Parse Ranking and Heuristics”, with C. Doran and S. Kulick, *Fourth International Workshop on Parsing Technology (IWPT 95)*, Prague, September 1995.

“Using Parsed Corpora for Circumventing Parsing”, with Aravind Joshi, *International Joint Conference on Artificial Intelligence workshop on New Approaches to Learning for Natural Language Processing*, Montreal, August 1995.

“Status of the XTAG System”, with C. Doran, D. Egedi and B.A. Hockey, *Third International Workshop on Tree-Adjoining Grammars (TAG+)*, TALANA, Paris, September 1994.

“Bootstrapping a Wide-coverage CCG from FB-LTAG”, with C. Doran, *Third International Workshop on Tree-Adjoining Grammars (TAG+)*, TALANA, Paris, September 1994.

“Complexity of Description of Primitives: Relevance to Local Statistical Computations”, with Aravind Joshi, *Balancing Act Workshop*, Las Cruces, New Mexico, June 1994.

INVITED TALKS

“Machine Translation in a Mobile World: From Models to Monetization”, Keynote Lecture at ICON 2008, Pune, India, December, 2008.

“Machine Translation”, Princeton University, November, 2008

“Lexicalized Representations and Language Generation”, at University of Washington, July, 2008.

“Spoken Dialog Systems”, at IIIT, Hyderabad, India, December, 2007.

“Multimodal and Multilingual Interfaces”, at Microsoft Research, India, February 2007.

“Localizing Dependencies: Consequences for NLP”, at SUNY, Stony Brook, November 2004.

“Balancing Data-driven and Rule-based Approaches in the Context of a Multimodal Conversational System”, at MITRE, November 2004.

“Corpus-based Natural Language Generation”, at Microsoft Research, March 2002.

“Evaluation Metrics for Natural Language Generation”, at Columbia University, October 2001.

“Stochastic Finite-State Transducers for Machine Translation”, at University of Maryland, June 2001.

“A Glimpse at NLP Techniques”, at Johns Hopkins University, Baltimore, June 2001.

“Complexity of Descriptions”, at the Workshop on Language Modeling for Biology, University of Pennsylvania, February 2001.

“Roles for Annotated and Unannotated Corpora in Natural Language Applications”, at IIIT, Hyderabad, India, January 2001.

“Exploiting a Probabilistic Hierarchical Model for Generation”, at IMPACTS 2000, workshop on Natural Language Generation, Dagstuhl, Germany, July 2000.

“Issues in Large-scale, Portable Machine Translation Systems”, at National Centre for Software Technology, Mumbai, India, August 1997.

“Supertags and Synchronous Lexicalized Tree-Adjoining Grammars”, at Center for Development of Advanced Computing (CDAC), Pune, India, February 1996.

“Supertagging and Almost Parsing”, at University of Paris VII, Paris, October 1995.

“Circumventing Parsing using Parsed Corpora”, at Bell Northern Research, Canada, August 1995.

“Almost-Parsing with Lexicalized Tree-Adjoining Grammars”, at Speech Research Symposium (SRS 95), Johns Hopkins University, June 1995.

“The XTAG System, Supertagging and EBL”, at SRI, Cambridge, September 1994.

“Disambiguating Supertags”, at ATR, Kyoto, Japan, December 1993.