

MIT

MIT

MIT

MIT

MIT

MIT

MIT

MIT

MIT

LABORATORY *for*

MIT

COMPUTER

MIT

SCIENCE

MIT

JULY 1998

MIT

SPOKEN

MIT

LANGUAGE

MIT

SYSTEMS

MIT

MIT

MIT

MIT

MIT

MIT

MIT

MIT

MIT Massachusetts
Institute
of Technology



©1998 Massachusetts Institute of Technology

For information or copies of this report, please contact:

Victoria L. Palay
MIT Laboratory for Computer Science
545 Technology Square, NE43-601
Cambridge, MA 02139 USA
palay@mit.edu

Please visit the Spoken Language Systems Group on the
World Wide Web at <http://www.sls.lcs.mit.edu>

SUMMARY *of*

RESEARCH

JULY 1998

Table of Contents

1	Summary of Research	
	Research, Technical, Administrative and Support Staff	viii-xi
	Research Assistants	xii-xv
	Post-Doctoral Associates, Undergraduate Students, Transitions	xv-xvi
	Research Sponsorship	xvii
2	Research Highlights	
	Research Highlights, 1997-1998	
	<i>Victor Zue</i>	3
3	Research Projects	
	JUPITER Data Collection and Analysis	
	<i>Joseph Polifroni, James Glass and Sally Lee</i>	9
	Natural Language Processing in the JUPITER Domain	
	<i>Stephanie Seneff and Joseph Polifroni</i>	12
	Spontaneous Speech Recognition in the JUPITER Domain	
	<i>James Glass</i>	18
	Confidence Scoring for Speech Understanding	
	<i>Christine Pao, Philipp Schmid and James Glass</i>	22
	PEGASUS: Flight Departure/Arrival/Gate Information System	
	<i>Stephanie Seneff, Joseph Polifroni and Philipp Schmid</i>	25
	Using Aggregation to Improve the Performance of Mixture Gaussian Acoustic Models	
	<i>T.J. Hazen and Andrew Halberstadt</i>	27
	BIANCA: A Dialogue Management Engine for PEGASUS	
	<i>Philipp Schmid, Stephanie Seneff and Joseph Polifroni</i>	29
	ANGIE-Based Pronunciation Server	
	<i>Aarati Parmar and Stephanie Seneff</i>	32
4	Thesis Research	
	A Model for Segment-Based Speech Recognition	
	<i>Jane Chang</i>	37
	Hierarchical Duration Modelling for a Speech Recognition System	
	<i>Grace Chung</i>	40
	Discourse Segmentation of Spoken Dialogue: An Empirical Approach	
	<i>Giovanni Flammia</i>	42
	Heterogeneous Acoustic Measurements and Multiple Classifiers for Speech Recognition	
	<i>Andrew Halberstadt</i>	45
	The Use of Speaker Correlation Information for Automatic Speech Recognition	
	<i>T. J. Hazen</i>	47

4

Thesis Research (continued)

The Mole: A Robust Framework for Accessing Information from the World Wide Web <i>HyungJin Kim</i>	50
Sublexical Modelling for Word-Spotting and Speech Recognition Using ANGIE <i>Raymond Lau</i>	52
Probabilistic Segmentation for Segment-Based Speech Recognition <i>Steven Lee</i>	56
A Model for Interactive Computation: Applications to Speech Research <i>Michael McCandless</i>	57
Subword Approaches to Spoken Document Retrieval <i>Kenney Ng</i>	60
A Semi-Automatic System for the Syllabification and Stress Assignment of Large Lexicons <i>Aarati Parmar</i>	62
A Segment-Based Speaker Verification System Using SUMMIT <i>Sridevi Sarma</i>	64
Context Dependent Modelling in a Segment-Based Speech Recognition System <i>Benjamin Serridge</i>	66
Toward the Automatic Transcription of General Audio Data <i>Michelle Spina</i>	67
Porting the GALAXY System to Mandarin Chinese <i>Chao Wang</i>	70
Concatenative Speech Synthesis of Isolated Words Using Sub-Word Units <i>Jon Yi</i>	75

5

Theses, Publications, Presentations and Seminars

Ph.D. and Masters Theses	79
Publications	80
Presentations	82
SLS Seminar Series	83

Research Staff



photo here

VICTOR ZUE

Victor Zue has been associated with MIT since 1970, as a graduate student, teacher and researcher. He is now a Senior Research Scientist, Associate Director of the MIT Laboratory for Computer Science, and head of the SLS group. His main research interest is in the development of conversational systems to facilitate graceful human/computer interactions. He has taught courses at MIT and abroad, written over 150 papers, and delivered numerous talks on this subject. He is a Fellow of the Acoustical Society of America, and currently chairs the Information Science and Technology (ISAT) Study Group for DARPA. In 1994, he was elected Distinguished Lecturer by the IEEE Signal Processing Society.



photo here

JAMES GLASS

James Glass is a Principal Research Scientist and Associate Head of the SLS group. He received his Ph.D. in Electrical Engineering and Computer Science from MIT in 1988. Over the past fifteen years, his research has covered many different areas of the speech communication chain, centered on computer speech recognition and spoken language understanding. In addition to publishing extensively in these areas, he has supervised S.M. and Ph.D. students, and co-taught courses in spectrogram reading and speech recognition. He is one of the original developers of the segment-based SUMMIT speech recognition system.



photo here

T.J. HAZEN

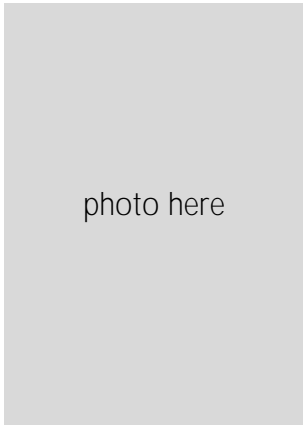
Timothy James (T. J.) Hazen arrived at MIT in 1987 where he received his S.B. degree in 1991, S.M. degree in 1993 and PhD in 1998, all in Electrical Engineering. T.J. joined the SLS group as an undergraduate in 1991 and has been with the group ever since. He is currently working as a research scientist in the group. His primary research interests include acoustic modeling, speaker adaptation, automatic language identification, and phonological modeling.



photo here

LEE HETHERINGTON

Lee Hetherington received his S.B., S.M., and Ph.D. degrees from MIT's Department of Electrical Engineering and Computer Science. He completed his doctoral thesis, "A Characterization of the the Problem of New, Out-of-Vocabulary Words in Continuous-Speech Recognition and Understanding," and joined the SLS group in October 1994. His research interests include many aspects of speech recognition, including search techniques, acoustic measurement discovery, and recently the use of weighted finite-state transduction for context-dependent phonetic models, phonological rules, lexicons, and language models in an integrated search.



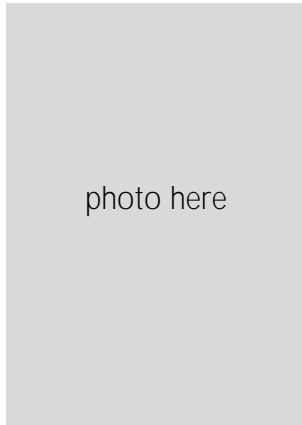
RAYMOND LAU

Raymond Lau received the B.S. in Computer Science and Engineering, the M.S. degree in Electrical Engineering and Computer Science, and the Ph.D. degree in Computer Science, all from the Massachusetts Institute of Technology in 1993, 1994, and 1998, respectively. He was a National Science Foundation fellow and is a member of Eta Kappa Nu. His current research interests are in the area of speech recognition and spoken language systems with a particular focus on subword modelling, search strategies and language modelling.



HELEN MENG

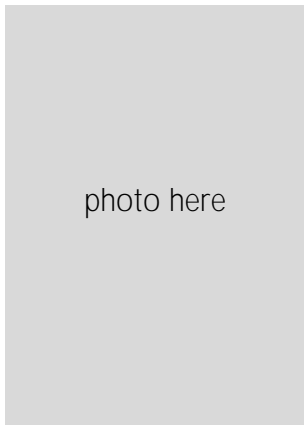
Helen Meng is a Research Scientist in the SLS group. She received her S.B., S.M. and Ph.D. degrees from MIT's Department of Electrical Engineering and Computer Science. Her doctoral thesis, entitled "Phonological Parsing for Bi-directional Letter-to-Sound /Sound-to-Letter Generation," was completed in February 1995, and her master's thesis, entitled "The Use of Distinctive Features for Automatic Speech Recognition," was completed in June 1991. Her research interests include front-end processing for speech recognition, lexical phonology, language modelling, discourse and dialog modelling as well as multi-linguality for conversational systems.



STEPHANIE SENEFF

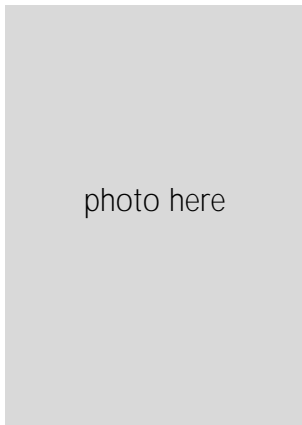
Stephanie Seneff has a B.S. degree in Biophysics and M.S., E.E., and Ph.D. degrees in Electrical Engineering and Computer Science from MIT. Her research interests span a wide spectrum of topics related to conversational systems, including phonological modelling, auditory modelling, computer speech recognition, statistical language modelling, natural language understanding and generation, discourse and dialogue modelling, and prosodic analysis. She has published numerous papers in these areas, and she is currently supervising several students at both master's and doctoral levels.

Technical Staff



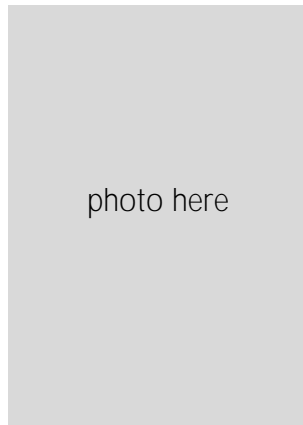
ED HURLEY

Ed Hurley received his B.S. in Physics from MIT in 1985. After working in semiconductor fabrication and parallel processing, he joined the SLS group in October 1994 doing application programming and system administration. His interests are in using the Web as a mechanism for the delivery of spoken language systems, as well as for speech data collection. He is also actively involved in developing telephone based spoken language systems.



CHRISTINE PAO

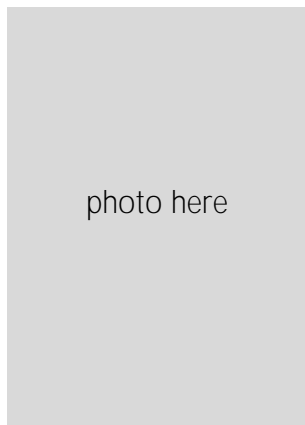
Christine Pao has been a member of the technical research staff since 1992. She is primarily involved in the development and maintenance of the GALAXY system. Her research interests are in discourse and dialog, systems integration with a focus on multilingual systems and language learning, and open microphone issues such as rejection and channel normalization. Christine has a bachelor's degree in Physics from MIT.



JOSEPH POLIFRONI

Joseph Polifroni's interests include language generation, human-computer interaction, and multilingual systems. He has worked on the back-end components of many of the SLS systems, including GALAXY and DINEX in addition to his work on GENESIS, the natural language generation system that is part of the overall GALAXY architecture. He has also contributed to the Spanish and Mandarin Chinese systems. Before joining SLS, Joe worked in the Speech Group at Carnegie Mellon University and was also a consultant for Carnegie Group Inc. in Pittsburgh. In addition, Joe spent two years living in China, teaching English at Shandong University in Jinan.

Administrative & Support Staff



VICTORIA PALAY

Victoria Palay has been a member of the Spoken Language Systems group since 1988. As SLS program administrator, she manages personnel, fiscal, publication and contractual matters as well as space and other group resources. In addition, she supports Victor Zue's duties as LCS Associate Director by coordinating equipment donations made to the Laboratory. Victoria has a B.A. in Government and French Studies from Smith College.



SALLY LEE

Sally Lee joined the Spoken Language Systems group as senior secretary in 1993. She received a B.A. in Studio Art/Art History from Colby College in 1984. She also studied at the Art Institute of Boston and the New York Studio School. In addition to her secretarial duties, Sally has made many of the animated and still icons for SLS programs including GALAXY and JUPITER. She also is responsible for transcribing sentences that are recorded from people calling into the JUPITER system.

Graduate Students



photo here

JANE CHANG

Jane Chang is a doctoral student working on a framework for feature-based speech recognition that better models the inherent variability in human speech. Currently, she is exploring how to use phonological and pronunciation constraints in lexical access. In the past year, she has also worked on other aspects of segmentation, classification and recognition. Jane receives support from an AT&T Fellowship.

Advisors: Victor Zue and James Glass



photo here

GRACE GHUNG

Grace Chung graduated in Electrical Engineering and Mathematics from the University of New South Wales, Sydney, Australia. She earned a Fulbright scholarship to attend MIT and completed her master's degree in June 1997. Her interests are in acoustic modelling and prosodic modelling for speech recognition.

Advisor: Stephanie Seneff



photo here

GIOVANNI FLAMMIA

Giovanni Flammia completed a M.Eng. (Laurea) in Electrical Engineering from the University of Rome in 1988 and an M.S. in Computer Science from McGill University in 1991, funded by a Government of Canada Award. Before joining the SLS group, he did speech processing research at the Center for Personal Communication at the University of Aalborg, Denmark and at CNET France Telecom labs in Lannion. His doctoral research focuses on developing dialogue models and user interfaces for spoken language systems that gather information from the Internet.

Advisor: Victor Zue

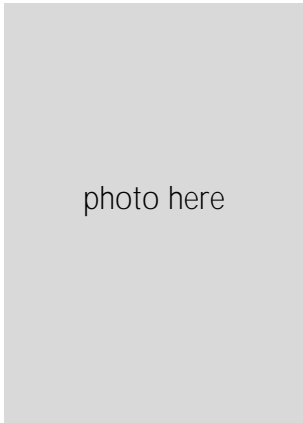


photo here

ANDREW HALBERSTADT

Andrew Halberstadt received the B.S. and M.S. degrees in Electrical Engineering from the University of Rochester in 1992 and 1993, respectively. In addition, he received the Bachelor of Music degree in 1991 from the Eastman School of Music in Rochester, NY. He was the recipient of a Sproul fellowship at the University of Rochester, and is a member of the engineering honor society Tau Beta Pi. His research interests include time-frequency representations, phonetic classification and recognition, speech and audio processing, and pattern recognition.

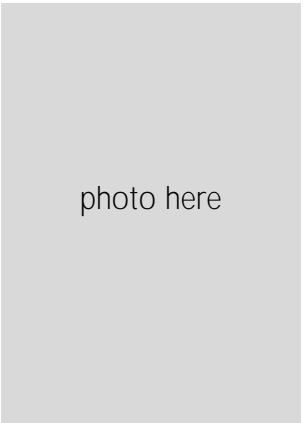
Advisor: James Glass



HYUNG-JIN KIM

Hyung-Jin Kim spent his undergraduate years at MIT and is currently pursuing a Masters of Engineering degree through the SLS group. His research interests include Java, XML, and other web technologies. Currently, he is working a system called the Mole which is a framework for robustly accessing information on HTML pages.

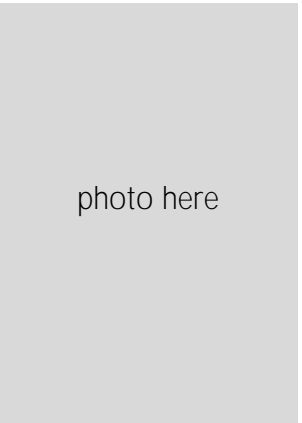
Advisor: Lee Hetherington



STEVEN LEE

Steven Lee is pursuing his master's degree with SLS. He received his S.B. degree from MIT in 1997 and expects to receive his M.Eng. degree in June 1998. He is a member and former president of the Tau Beta Pi Engineering Honor Society, as well as a member and vice-president of the Eta Kappa Nu Honor Society. His current research interest is in probabilistic segmentation.

Advisor: James Glass



KAREN LIVESCU

Karen Livescu received her B.A. in Physics at Princeton University in 1996. She spent the following year at the Technion in Haifa, Israel, as a visiting student in the Electrical Engineering department. Karen started graduate study in the SLS group in September 1997. She is a National Science Foundation fellow and plans to pursue research in speech recognition.

Advisor: James Glass



MICHAEL McCANDLESS

Michael McCandless is working towards a doctoral degree in the area of speech recognition. He is constructing a novel framework which will unify all stages in the speech recognition process. The framework is cast within an interactive tool, based on the Python language, which enables rapid prototyping of new recognition domains and exploration of design tradeoffs. His master's thesis was in the area of automatic learning of language structure for improving speech recognition. Michael is co-author of the IEEE Expert Internet Services Department and a member of the American Association for the Advancement of Science.

Advisor: James Glass

Graduate Students

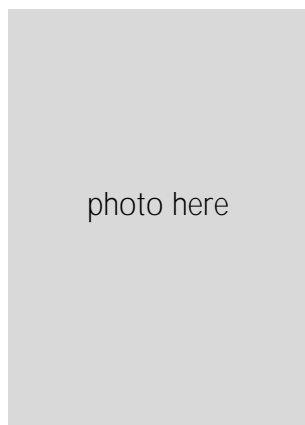


photo here

KENNEY NG

Kenney Ng's current research interest is in the area of information retrieval of spoken documents, which is the task of identifying those speech messages stored in a large collection that are relevant to a query provided by a user. Prior to his return to MIT in 1995, Kenney was a member of the Speech and Language Department at BBN Systems and Technologies where he did research on large vocabulary recognition of conversational speech, word spotting, topic spotting, probabilistic segmental speech models, and noise compensation. He received his B.S. and M.S. degrees in EECS from MIT in 1990.

Advisor: Victor Zue

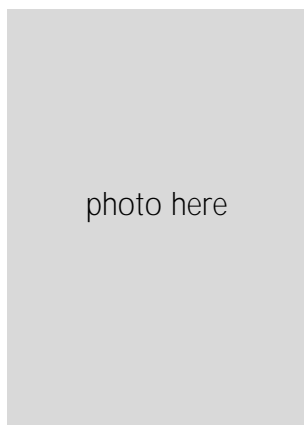


photo here

SRIDEVI SARMA

Sridevi Sarma received her bachelor's degree in Electrical Engineering from Cornell University in 1994. She completed her master's thesis, which investigates speaker verification using a segment-based approach in June 1997. Sridevi is a National Science Foundation fellow.

Advisor: Victor Zue

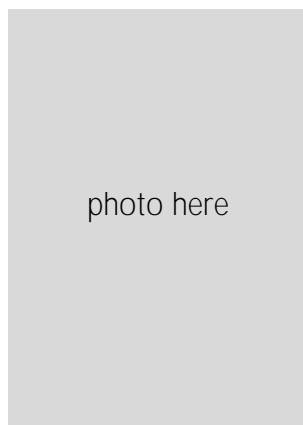


photo here

MICHELLE SPINA

Michelle Spina received the B.S. in Electrical Engineering from the Rochester Institute of Technology in 1991, and the S.M. in Electrical Engineering from MIT in 1994. She is currently pursuing a Ph.D. degree in the SLS group. Michelle's research interests include automatic indexing of audio content, speech recognition and understanding, and biomedical issues of speech processing as they relate to automatic speech recognition. Her current research involves general sound understanding, and orthographic analysis of general audio data. Michelle was a 1995 Intel Foundation Graduate Fellow, and is a member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi.

Advisor: Victor Zue

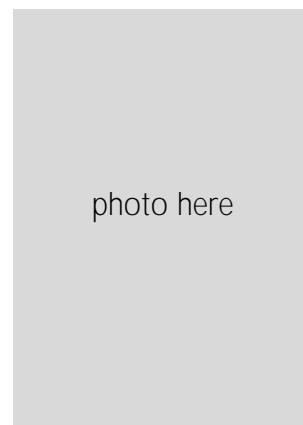


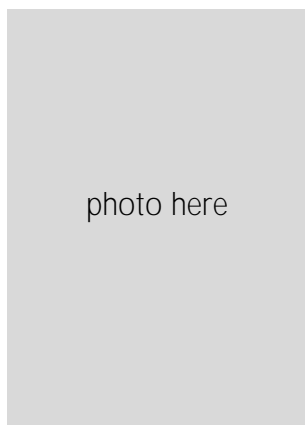
photo here

CHAO WANG

Chao Wang received her bachelor's degree in Biomedical Engineering, with a minor in Computer Science from Tsinghua University, Beijing, China in 1994. She started her graduate study in MIT in September 1995 and joined the SLS group in April 1996. Chao's master's degree, completed in June 1997, worked on porting the GALAXY system to Mandarin Chinese.

Advisor: Stephanie Seneff

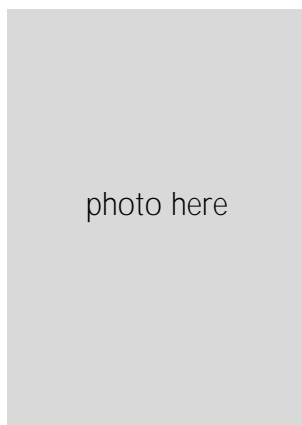
Post-Doctoral Associates



JON YI

Jon Yi received the S.B. and the M.Eng. degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 1997 and 1998, respectively. He also graduated in 1997 with a minor in Music. At SLS he has worked on developing a Mandarin Chinese concatenative speech synthesizer and a UNICODE/Java World Wide Web interface for the GALAXY system. His research interests include speech synthesis, communications systems, and multilingual speech understanding systems.

Advisor: James Glass



PHILIPP SCHMID

Philipp Schmid received his Ph.D. in Computer Science and Engineering from the Oregon Graduate Institute of Science and Technology in December 1996. In his Ph.D. thesis he investigated the use of "Explicit N-Best Formant Features for Segment-Based Speech Recognition". He joined SLS in January 1997 as a Postdoctoral Associate. He is interested in building conversational systems for real users. His main focus has been in the area of dialogue management, information retrieval from web-based knowledge sources, and flow of control issues. He has been working on PEGASUS, the flight arrival and departure information system, where he helped develop a new dialogue control mechanism.



NIKKO STRÖM

Nikko Ström received the Master of Science, (Engineering Physics) degree in 1991, and the Ph.D. degree in Electrical Engineering (Department of Speech, Music, and Hearing) in 1997 at the Royal Institute of Technology (KTH), Stockholm, Sweden. He joined SLS in May '98 as a Postdoctoral Associate. His main areas of interest are human/machine dialogue, lexical search in automatic speech recognition, and acoustic/phonetic modeling.

Undergraduate Students

Ivan Gonzalez-Gallo
Simo Kamppari
Fernando Perez
Rafael Schloming
Archit Shah
Aleem Siddiqui
Samuel Wong
James Wood
Minnan Xu

Transitions

Jane Chang, Ph.D. June 1998
Giovanni Flammia, Ph.D., June 1998
T.J. Hazen, Ph.D., January 1998/SLS Research Scientist
Jim Hugunin
Raymond Lau, Ph.D., March 1998/SLS Research Scientist
Steve Lee, M.Eng, June 1998
Michael McCandless, Ph.D., June 1998
Helen Meng, April 1998 joined Chinese University of Hong Kong
Aarati Parmar, M.Eng., June 1997
Sridevi Sarma, S.M., June 1997
Benjamin Serridge, M.Eng., August 1997

Research Sponsorship

Defense Advanced Research Projects Agency¹
Bell Atlantic Corporation
BellSouth Intelliventures
National Science Foundation²

In addition, discretionary funds for research are provided by ATR Interpreting Telecommunications Research Laboratories and Intel Corporation.

1. Contract No. N66001-96-C-8526, from the Information Technology Office, monitored by the Naval Command Control, and Ocean Surveillance Center and contract no. DAAN02-98-K0003, monitored through US Army Soldier Systems Command.

2. This material is based upon work supported by NSF grant no. IRI-9618731.

