

# Speech and Natural Language

**Proceedings of a Workshop  
Held at  
Pacific Grove, California  
February 19-22, 1991**

**Sponsored by:  
Defense Advanced Research Projects Agency  
Information Science and Technology Office**

This document contains copies of reports prepared for the DARPA Speech and Natural Language Workshop. Included are reports from DARPA/ISTO sponsored programs and other materials prepared for use at the workshop.

**APPROVED FOR PUBLIC RELEASE  
DISTRIBUTION UNLIMITED**

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the United States Government.

Distributed by  
Morgan Kaufmann Publishers, Inc.  
2929 Campus Drive  
San Mateo, California 94403  
ISBN 1-55860-207-0  
Printed in the United States of America

# TABLE OF CONTENTS

	Page	
<b>Author Index</b> . . . . .	x	
<b>Overview of the Workshop, Patti Price, General Chair</b> . . . . .	3	
 <b>SESSION 1: Speech and Natural Language Efforts in the U.S. and Abroad</b>		
<b>Session Summary, Mark Liberman, Chair; and Patti Price</b> . . . . .	5	
 <b>The ESPRIT Project Polyglot</b>		
L. Boves; Nijmegen University . . . . .	7	
 <b>Some Notes about Research and Development at KTH</b>		
R. Carlson; Royal Institute of Technology. . . . .	12	
 <b>Speech Understanding and Dialogue over the Telephone: An Overview of the ESPRIT SUNDIAL Project</b>		
J. Peckham; Logica. . . . .	14	
 <b>Machine Translation in Europe</b>		
S. Warwick-Armstrong; University of Geneva. . . . .	28	
 <b>ESPRIT, The European Strategic Programme for Research and Development in Information Technology</b> . . . . .		34
 <b>SESSION 2: DARPA Resource Management and ATIS Benchmark Test Poster Session</b>		
<b>Session Summary, David Pallett, Chair</b> . . . . .	49	
 <b>Dragon Systems Resource Management Benchmark Results - February 1991</b> James Baker, Janet Baker, P. Bamberg, L. Gillick, L. Lamel, R. Roth, F. Scattone, D. Sturtevant, O. Ba and R. Benedict; Dragon Systems . . . . .		59
 <b>New Results with the Lincoln Tied-Mixture HMM CSR System</b>		
D. B. Paul; MIT Lincoln Laboratory. . . . .	65	
 <b>Modelling Context Dependency in Acoustic-Phonetic and Lexical Representations</b>		
M. Phillips, J. Glass and V. Zue; MIT Laboratory for Computer Science. . . . .	71	

**BYBLOS Speech Recognition Benchmark Results**

F. Kubala, S. Austin, C. Barry, J. Makhoul, P. Placeway and R. Schwartz;

BBN Systems &amp; Technologies . . . . . 77

**Integration of Diverse Recognition Methodologies through Reevaluation of N-Best Sentence Hypotheses**

M. Ostendorf, A. Kannan, Boston University; S. Austin, BBN Systems &amp; Technologies;

O. Kimball, Boston University; R. Schwartz,

and J. R. Rohlicek, BBN Systems &amp; Technologies . . . . . 83

**Development and Preliminary Evaluation of the MIT ATIS System**

S. Seneff, J. Glass, D. Goddeau, D. Goodine, L. Hirschman, H. Leung, M. Phillips,

J. Polifroni and V. Zue; MIT Laboratory for Computer Science . . . . . 88

**Speech Recognition in SRI's Resource Management and ATIS Systems**

H. Murveit, J. Butzberger and M. Weintraub; SRI International . . . . . 94

**Evaluation of the CMU ATIS System**

W. Ward; Carnegie Mellon University . . . . . 101

**Using Semantics to Correct Parser Output for ATIS Utterances**

S. Young; Carnegie Mellon University . . . . . 106

**BBN HARC and DELPHI Results on the ATIS Benchmarks - February 1991**

S. Austin, D. Ayuso, M. Bates, R. Bobrow, R. Ingria, J. Makhoul, P. Placeway, R. Schwartz

and D. Stallard; BBN Systems &amp; Technologies . . . . . 112

**A Textual Processor to Handle ATIS Queries**

D. O'Shaughnessy; INRS-Telecommunications . . . . . 116

**Stochastic Representation of Conceptual Structure in the ATIS Task**

R. Pieraccini, E. Levin, and C-H. Lee; AT&amp;T Bell Laboratories . . . . . 121

**Augmented Role Filling Capabilities for Semantic Interpretation of Spoken Language**

L. Norton, M. Linebarger, D. Dahl and N. Nguyen; Unisys . . . . . 125

**The Use of a Commerical Natural Language Interface in the ATIS Task**

E. Tzoukermann; AT&amp;T Bell Laboratories . . . . . 134

**SESSION 3: Machine Translation****Session Summary, Jaime Carbonell, Chair . . . . . 139****Machine Translation Using Abductive Inference**

J. R. Hobbs and M. Kameyama; SRI International . . . . . 141

<b>A Statistical Approach to Sense Disambiguation in Machine Translation</b> P. F. Brown, S. A. Della Pietra, V. J. Della Pietra and R. L. Mercer; IBM TJ Watson Research Center . . . . .	146
<b>Identifying Word Correspondences in Parallel Texts</b> W. A. Gale and K. W. Church; AT&T Bell Laboratories . . . . .	152
<b>SESSION 4: Speech I</b>	
<b>Session Summary, Richard Lyon, Chair . . . . .</b>	159
<b>Field Test Evaluations and Optimization of Speaker Independent Speech Recognition for Telephone Applications</b> C. Gagnoulet and C. Sorin; CNET. . . . .	160
<b>Collection and Analysis of Data from Real Users: Implications for Speech Recognition/Understanding Systems</b> J. Spitz and the Artificial Intelligence Speech Technology Group; NYNEX Science and Technology. . . . .	164
<b>Autodirective Microphone Systems for Natural Communication with Speech Recognizers</b> J. L. Flanagan and R. Mammone; CAIP Center, Rutgers University G. W. Elko; AT&T Bell Laboratories . . . . .	170
<b>Signal Representation, Attribute Extraction and, the Use of Distinctive Features for Phonetic Classification</b> H. M. Meng, V. W. Zue and H. C. Leung; MIT Laboratory for Computer Science. . . . .	176
<b>SESSION 5: Natural Language I</b>	
<b>Session Summary, James Allen, Chair . . . . .</b>	183
<b>The Mapping Unit Approach to Subcategorization</b> R. Bobrow, R. Ingria and D. Stallard; BBN Systems & Technologies. . . . .	185
<b>A Template Matcher for Robust NL Interpretation</b> E. Jackson, D. Appelt, J. Bear, R. Moore and A. Podlozny; SRI International . . . . .	190
<b>Fixed and Flexible Phrase Structure: Coordination in Tree Adjoining Grammars</b> A. K. Joshi and Y. Schabes; University of Pennsylvania . . . . .	195
<b>Efficient Bottom-Up Parsing</b> R. Moore and J. Dowding; SRI International. . . . .	200

	<b>Page</b>
<b>Partial Parsing: A Report on Work in Progress</b> R. Weischedel, D. Ayuso, R. Bobrow, S. Boisen, R. Ingria and J. Palmucci; BBN Systems & Technologies . . . . .	204
<b>SESSION 6: Demonstrations and Videotapes of Speech and Natural Language Technologies</b>	
<b>Session Summary, Mari Ostendorf, Chair . . . . .</b>	<b>211</b>
<b>Analog Implementations of Auditory Models</b> R. F. Lyon; Apple Computer . . . . .	212
<b>Using Spoken Language to Facilitate Military Transportation Planning</b> M. Bates, D. Ellard, P. Peterson and V. Shaked; BBN Systems & Technologies . . . . .	217
<b>SESSION 7: Natural Language II</b>	
<b>Session Summary, Salim Roukos, Chair. . . . .</b>	<b>221</b>
<b>Statistical Agenda Parsing</b> R. J. Bobrow; BBN Systems & Technologies . . . . .	222
<b>Some Results on Stochastic Language Modelling</b> R. De Mori and R. Kuhn; McGill Universit . . . . .	225
<b>Parsing the Voyager Domain Using Pearl</b> D. M. Magerman and M. P. Marcus; University of Pennsylvania . . . . .	231
<b>Calculating the Probability of a Partial Parse of a Sentence</b> F. Kochman and J. Kupin; Institute for Defense Anlayses . . . . .	237
<b>A Trellis-Based Algorithm for Estimating the Parameters of a Hidden Stochastic Context-free Grammar</b> J. Kupiec; Xerox Palo Alto Research Center. . . . .	241
<b>SESSION 8: Speech II</b>	
<b>Session Summary, Kai-Fu Lee, Chair. . . . .</b>	<b>247</b>
<b>Continuous Speech Recognition Using Segmental Neural Nets</b> S. Austin, J. Makhoul and R. Schwartz; BBN Systems & Technologies G. Zavaliagkos; Northeastern University. . . . .	249

<b>A Dynamical System Approach to Continuous Speech Recognition</b> V. Digalakis, Boston University; J. R. Rohlicek, BBN Systems & Technologies; and M. Ostendorf, Boston University. . . . .	253
<b>Recent Progress in Robust Vocabulary-Independent Speech Recognition</b> H-W. Hon and K-F. Lee; Carnegie Mellon University . . . . .	258
<b>Context Dependent Modeling of Phones in Continuous Speech Using Decision Trees</b> L. R. Bahl, P. V. de Souza, P. S. Gopalakrishnan, D. Nahamoo and M. A. Picheny; IBM TJ Watson Research Center . . . . .	264
<b>SESSION 9: Speech III</b>	
<b>Session Summary, Francis Kubala, Chair . . . . .</b>	271
<b>Bayesian Learning of Gaussian Mixture Densities for Hidden Markov Models</b> J-L. Gauvain and C-H. Lee; AT&T Bell Laboratories . . . . .	272
<b>A Study on Speaker-Adaptive Speech Recognition</b> X. D. Huang; Carnegie Mellon University. . . . .	278
<b>Experience with a Stack Decoder-Based HMM CSR and Back-Off N-Gram Language Models</b> D. B. Paul; MIT Lincoln Laboratory. . . . .	284
<b>Lexical Access with a Statistically-Derived Phonetic Network</b> M. D. Riley and A. Ljolje; AT&T Bell Laboratories . . . . .	289
<b>A Dynamic Language Model for Speech Recognition</b> F. Jelinek, B. Meriardo, S. Roukos and M. Strauss; IBM TJ Watson Research Center. . . . .	293
<b>SESSION 10: Corpora and Evaluation</b>	
<b>Session Summary, Clifford Wienstein, Chair. . . . .</b>	297
<b>Third Message Understanding Evaluation and Conference (MUC-3): Phase 1 Status Report</b> B. M. Sundheim; Naval Ocean Systems Center . . . . .	301

<b>A Procedure for Quantitatively Comparing the Syntactic Coverage of English Grammars</b>	
E. Black, S. Abney, D. Flickenger, C. Gdaniec, R. Grishman, P. Harrison, D. Hindle, R. Ingria, F. Jelinek, J. Klavans, M. Liberman, M. Marcus, S. Roukos, B. Santorini and T. Strzalkowski; IBM TJ Watson Research Center . . . . .	306
<b>Evaluating Text Categorization</b>	
D. D. Lewis; University of Massachusetts . . . . .	312
<b>A Proposal for Incremental Dialogue Evaluation</b>	
M. Bates and D. Ayuso; BBN Systems & Technologies . . . . .	319
<b>SESSION 11: Natural Language III</b>	
<b>Session Summary, Mitch Marcus, Chair . . . . .</b>	<b>323</b>
<b>Discourse Structure in the TRAINS Project</b>	
J. F. Allen; University of Rochester . . . . .	325
<b>Studies in Part of Speech Labelling</b>	
M. Meteer, R. Schwartz and R. Weischedel; BBN Systems & Technologies . . . . .	331
<b>Lexico-Semantic Pattern Matching as a Companion to Parsing in Text Understanding</b>	
P. S. Jacobs, G. R. Krupka and L. F. Rau; General Electric Corporate Research and Development Center . . . . .	337
<b>Automatic Acquisition of Subcategorization Frames from Tagged Text</b>	
M. R. Brent and R. C. Berwick; MIT Artificial Intelligence Laboratory . . . . .	342
<b>Fast Text Processing for Information Retrieval</b>	
T. Strzalkowski and B. Vauthey; New York University . . . . .	346
<b>SESSION 12: SLS and Prosody</b>	
<b>Session Summary, Edward Neuburg, Chair . . . . .</b>	<b>353</b>
<b>Interactive Problem Solving and Dialogue in the ATIS Domain</b>	
S. Seneff, L. Hirschman and V. W. Zue; MIT Laboratory for Computer Science . . . . .	354
<b>Collection of Spontaneous Speech for the ATIS Domain and Comparative Analyses of Data Collected at MIT and TI</b>	
J. Polifroni, S. Seneff and V. W. Zue; MIT Laboratory for Computer Science . . . . .	360

<b>Integrating Syntax and Semantics into Spoken Language Understanding</b> L. Hirschman, S. Seneff, D. Goodine and M. Phillips; MIT Laboratory for Computer Science . . . . .	366
<b>The Use of Prosody in Syntactic Disambiguation</b> P. Price, SRI International; M. Ostendorf, Boston University; S. Shattuck-Hufnagel, MIT; and C. Fong, Boston University . . . . .	372
<b>Predicting Intonational Boundaries Automatically from Text: The ATIS Domain</b> M. Q. Wang; Cambridge University and J. Hirschberg; AT&T Bell Laboratories . . . . .	378
<b>Use of Prosody in Syntactic Disambiguation: An Analysis-By-Synthesis Approach</b> C. W. Wightman, N. M. Veilleux and M. Ostendorf; Boston University . . . . .	384
 <b>ADDITIONAL PAPERS</b>	
<b>The Consortium for Lexical Research</b> New Mexico State University, Computing Research Laboratory. . . . .	393
<b>A Proposal for Lexical Disambiguation</b> G. A. Miller and D. A. Teibel; Princeton University, Cognitive Science Laboratory . . . . .	395
 <b>SITE REPORTS</b> . . . . .	 401

## AUTHOR INDEX

Abney, S. . . . .	306	Ellard, D. . . . .	217
Allen, J. F.. . . . .	183, 325, 431	Feiner, S. . . . .	413
Appelt, D. . . . .	190	Flanagan, J. L. . . . .	170
Austin, S. . . . .	77, 83, 112, 249	Flickenger, D.. . . .	306
Ayuso, D. . . . .	112, 204, 319	Fong, C. . . . .	372
Ba, O. . . . .	59	Gagnoulet, C.. . . .	160
Bahl, L. R. . . . .	264	Gale, W. A. . . . .	152
Baker, J. K.. . . .	59	Gauvain, J.-L.. . . .	272
Baker, J. M. . . . .	59, 414	Gdaniec, C. . . . .	306
Bamberg, P. . . . .	59	Gillick, L. S. . . . .	59, 414
Barry, C. . . . .	71	Glass, J.. . . .	71, 88
Bates, M. . . . .	112, 217, 319	Goddeau, D.. . . .	88
Bear, J. . . . .	190	Goodine, D.. . . .	88, 366
Benedict, R. . . . .	59	Gopalakrishnan, P. S.. . . .	264
Bernstein, J. . . . .	423	Grishman, R. . . . .	306, 421
Berwick, R. C. . . . .	342	Harrison, P. . . . .	306
Bisiani, R. . . . .	412	Hindle, D.. . . .	306
Black, E. . . . .	306	Hirschberg, J. . . . .	378
Bobrow, R.. . . .	112, 185, 204, 222	Hirschman, L.. . . .	88, 354, 366, 416
Boisen, S. . . . .	204	Hobbs, J. R. . . . .	141, 426
Boves, L.. . . .	7	Hon, H.-W. . . . .	258
Brent, M. R. . . . .	342	Hovy, E. H. . . . .	432
Brown, P. F. . . . .	146	Huang, X. D. . . . .	278
Butzberger, J.. . . .	94	Ingria, R. . . . .	112, 185, 204, 306
Carbonell, J. . . . .	139	Jackson, E. . . . .	190
Carlson, R.. . . .	12	Jacobs, P. S. . . . .	337
Church, K. W. . . . .	152	Jelinek, F.. . . .	293, 306
Dahl, D.A. . . . .	125, 427	Joshi, A. K.. . . .	195, 429
Della Pietra, S. A. . . . .	146	Kameyama, M. . . . .	141
Della Pietra, V. J. . . . .	146	Kannan, A. . . . .	83
De Mori, R. . . . .	225	Kimball, O. . . . .	83
de Souza, P. V.. . . .	264	Klavans, J. . . . .	306
Digalakis, V.. . . .	253	Kochman, F. . . . .	237
Dowding, J. . . . .	200	Krupka, G. R.. . . .	337
Elko, G. W. . . . .	170	Kubala, F.. . . .	77, 271

Kuhn, R. . . . .	225	Picheny, M. A. . . . .	264
Kupiec, J. . . . .	241	Pieraccini, R. . . . .	121
Kupin, J. . . . .	237	Placeway, P. . . . .	77, 112
Lamel, L. . . . .	59	Podlozny, A. . . . .	190
Lee, C.-H. . . . .	121, 272	Polifroni, J. . . . .	88, 360
Lee, K.-F. . . . .	247, 258	Price, P. . . . .	3, 5, 372, 408, 425
Leung, H. C. . . . .	88, 176	Rau, L. F. . . . .	337
Levin, E. . . . .	121	Reddy, R. . . . .	411
Lewis, D. D. . . . .	312	Riley, M. D. . . . .	289
Liberman, M. Y. . . . .	5, 306	Rohlicek, J. R. . . . .	83, 253, 409
Linebarger, M. . . . .	125	Roth, R. . . . .	59
Ljolje, A. . . . .	289	Roukos, S. . . . .	221, 293, 306
Lyon, R. F. . . . .	159, 212	Santorini, B. . . . .	306
Magerman, D. M. . . . .	231	Scattono, F. . . . .	59
Makhoul, J. . . . .	77, 112, 249, 406, 407	Schabes, Y. . . . .	195
Mammone, R. . . . .	170	Schubert, L. K. . . . .	431
Marcus, M. P. . . . .	231, 306, 323, 429, 430	Schwartz, R. . . . .	77, 83, 112, 249, 331, 406
McKeown, K. . . . .	413	Seneff, S. . . . .	88, 354, 360, 366
Meng, H. M. . . . .	176	Shaked, V. . . . .	217
Mercer, R. L. . . . .	146	Shattuck-Hufnagel, S. . . . .	372
Merialdo, B. . . . .	293	Shore, J. . . . .	415
Meteer, M. . . . .	331	Silverman, H. F. . . . .	410
Miller, G. A. . . . .	395	Sorin, C. . . . .	160
Moore, R. . . . .	190, 200, 425	Spitz, J. . . . .	164
Murveit, H. . . . .	94, 424	Stallard, D. . . . .	112, 185
Nahamoo, D. . . . .	264	Steedman, M. . . . .	429
Neuburg, E. P. . . . .	353	Strauss, M. . . . .	293
Nguyen, N. . . . .	125	Strzalkowski, T. . . . .	306, 346
Norton, L. . . . .	125	Sturtevant, D. . . . .	59
O'Shaughnessy, D. . . . .	116	Sundheim, B. M. . . . .	301, 419
Ostendorf, M. . . . .	83, 211, 253, 372	Teibel, D. A. . . . .	395
. . . . .	384, 408, 409	Tzoukermann, E. . . . .	134
Pallett, D. S. . . . .	49, 418	Vauthey, B. . . . .	346
Palmucci, J. . . . .	204	Veilleux, N. M. . . . .	384
Paul, D. B. . . . .	65, 284, 417	Wang, M. Q. . . . .	378
Peckham, J. . . . .	14	Ward, W. . . . .	101
Peterson, P. . . . .	217	Warwick-Armstrong, S. . . . .	28
Phillips, M. . . . .	71, 88, 366	Wayne, C. L. . . . .	403

Webber, B. . . . .	429
Weinstein, C. J. . . . .	297, 417
Weintraub, M. . . . .	94, 424
Weischedel, R. . . . .	204, 331, 405
Wightman, C. W. . . . .	384
Wilensky, R. . . . .	428
Wilks, Y. . . . .	420
Young, S. . . . .	106
Zavaliagos, G. . . . .	249
Zue, V. . . . .	71, 88, 176, 354, 360, 416
Zweig, G. . . . .	422