Speech and Natural Language

Proceedings of a Workshop
Held at
Philadelphia, Pennsylvania
February 21-23, 1989

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 1-55860-073-6

Printed in the United States of America

TABLE OF CONTENTS

Page

FOREWORD vii

CONFERENCE OVERVIEW 1

Overview of the Workshop 1

Workshop Agenda 3

Reports From Session Chairs 5

Tutorials 37

TECHNICAL PAPERS

AT&T 75

"Speaker Independent Phonetic Transcription of Fluent Speech for Large Vocabulary Speech Recognition", S.E. Levinson, M.Y. Liberman, A. Ljolje, and L.G. Miller

BBN 81

"Rapid Porting of the Parlance™ Natural Language Interface", Madeleine Bates

"A Common Facts Data Base", William Crowther

"The BBN BYBLOS Continuous Speech Recognition System", Richard Schwartz, Chris Barry, Yen-Lu Chow, Alan Derr, Ming-Whei Feng, Owen Kimball, Francis Kubala, John Makhoul, and Jeffrey Vandegrift

"Speaker Adaptation from Limited Training in the BBN BYBLOS Speech Recognition System", Francis Kubala, Ming-Whei Feng, John Makhoul, and Richard Schwartz

"The BBN Spoken Language System", Sean Boisen, Yen-Lu Chow, Andrew Haus, Robert Ingria, Salim Roukas, and David Stallard

TABLE OF CONTENTS (Cont'd)

TECHNICAL PAPERS (Cont'd)

Page

"Portability in the Janus Natural Language Interface", Ralph M. Weischedel, Robert J. Bobrow, Damaris Ayuso, and Lance Ramshaw

CMU 119

"The Design of Voice-Driven Interfaces", Alexander I. Rudnicky

"Recent Progress in the SPHINX Speech Recognition System", Kai-Fu Lee, Hsiao-Wuen Hon, and Mei-Yuh Hwang

"The MINDS System: Using Context and Dialog to Enhance Speech Recognition", Sheryl R. Young

"Understanding Spontaneous Speech", Wayne Ward

DRAGON 143

IBM 145

"Goats to Sheep: Can Recognition Rate be Improved for Poor Tangora Speakers?", Catalina M. Danis

ISI 151

"A Flexible Interface for Linking Applications to Penman's Sentence Generator", Robert T. Kasper

LINCOLN 159

"The Lincoln Continuous Speech Recognition System: Recent Developments and Results", Douglas B. Paul

MIT 167

"TINA: A Probabilistic Syntactic Parser for Speech Understanding Systems", Stephanie Seneff

TECHNICAL PAPERS (Cont'd)

Page

"The MIT SUMMIT Speech Recognition System: A Progress Report", Victor Zue, James Glass, Michael Phillips, and Stephanie Seneff

NIST 191

NMSU 193

NOSC 195

"Plans for a Task-Oriented Evaluation of Natural Language Understanding Systems", Beth M. Sundheim

NYU 203

"Analyzing Telegraphic Messages", Ralph Grishman and John Sterling

Penn 209

"Lexicalized TAGs, Parsing and Lexicons", Anne Abeille, Kathleen Bishop, Sharon Cote, Aravind K. Joshi, and Yves Schabes

"Elements of a Computational Model of Cooperative Response Generation", Brant A. Cheikes and Bonnie L. Webber

"Intonation and Syntax in Spoken Language Systems", Mark Steedman

Signition 229

"Auditory Speech Preprocessors", George Zweig

SRI 237

"SRI's DECIPHER System", Hy Murveit, Michael Cohen, Patti Price, Gay Baldwin, Mitch Weintraub, and Jared Bernstein

"Integrating Speech and Natural-Language Processing", Robert Moore, Fernando Pereira, and Hy Murveit

TECHNICAL PAPERS (Cont'd)

Page

TI 249

"Chart Parsing of Stochastic Spoken Language Models", Charles Hemphill and Joseph Picone

"Initial Draft Guidelines for the Development of the Next-Generation Spoken Language Systems Speech Research Database", George R. Doddington

UCB 263

Unisys 265

"Analyzing Explicitly-Structured Discourse in a Limited Domain: Trouble and Failure Reports", Catherine N. Ball

"Reducing Search by Partitioning the Word Network", John Dowding

"Porting PUNDIT to the Resource Management Domain", Lynette Hirschman, Francois-Michel Lang, John Dowding and Carl Weir

Xerox 283

"Automatic Discovery of Contextual Factors Describing Phonological Variation", Francine Chen and Jeff Shrager

"Probabilistic Models of Short and Long Distance Word Dependencies in Running Text", Julien Kupiec

Foreword

The first DARPA-sponsored Speech and Natural Language Workshop took place in Philadelphia, Pennsylvania, on 21-23 February 1989. The workshop dealt with all aspects of the DARPA speech and natural language programs, placing special emphasis on spoken language understanding. It was the first time that representatives of all the DARPA-sponsored groups met together, and a special effort was made to cross-educate the two communities.

The workshop was quite successful along a number of dimensions. Over 120 people participated. In addition to the DARPA contractors, representatives of AT&T and IBM attended, making important contributions. Much progress was reported, many new ideas were surfaced, and a great deal of cross-fertilization occurred both in and outside the sessions. It was especially encouraging to see the speech and natural language communities talking to one another.

This Proceedings begins with an overview of the workshop, a copy of the agenda, summaries of the various sessions by their respective chairs, and two tutorial papers. The bulk of the Proceedings consists of many of the technical papers presented during the workshop. (Papers being published elsewhere are not included.) A brief summary of each contractor's program appears before the papers from that contractor.

Many people contributed to the success of the workshop. Special recognition should go to Lynette Hirschman of Unisys, who served as General Chair, and Mary Sullivan of SAIC, who worked tirelessly to make it all run smoothly.

Charles L. Wayne
Program Manager