

**28th  
Annual Meeting  
of the  
Association for  
Computational Linguistics**

**Proceedings of the Conference**

**6-9 June 1990  
University of Pittsburgh  
Pittsburgh, Pennsylvania, USA**

**Published by the Association for Computational Linguistics**

© 1990, Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Donald E. Walker (ACL)  
Bellcore, MRE 2A379  
445 South Street, Box 1910  
Morristown, NJ 07960-1910, USA

Printed by: **OMNIPRESS** . . . **"The Proceedings Printer"**  
Madison, Wisconsin

## PREFACE

This volume contains the papers presented at the *28th Annual Meeting of the Association for Computational Linguistics*.

Every year it seems that the task of picking the best papers out of hundreds of exciting and new papers becomes more and more difficult. As a community, we should thank all these authors who submitted papers of such high quality, whether their papers were accepted or not, for it is they who make the meeting possible and set its high standards.

Of course, that means that we are indebted more than ever to the members of the Program Committee. They deserve our deepest thanks for the Herculean task of reading and carefully reviewing a near-record number of papers: David Israel, *SRI International*; Karen Jensen, *IBM Corporation*; Aravind Joshi, *University of Pennsylvania*; Richard Larson, *State University of New York, Stony Brook*; Paul Martin, *SRI International*; Kathy McKeown, *Columbia University*; Martha Pollack, *SRI International*; James Pustejovsky, *Brandeis University*; Edward Stabler, *University of California at Los Angeles*; Hans Uszkoreit, *Universität des Saarlandes*; and David Weir, *Northwestern University*.

Kudos go to Rich Thomason of the University of Pittsburgh for acting as the local arrangements coordinator and his staff assistant Stefni Agin; to Dan Flickinger for putting together a most interesting set of tutorials; to Johanna Moore who made demonstrations and exhibits possible; and to my own staff assistant Rhonda Byrne for keeping the program committee's business running smoothly. A special tip-of-the-hat to the managers of MIT's air conditioning, who kept the program committee running on schedule.

Our invited speakers always add a special sparkle to our meetings, so I would like to personally applaud Lila Gleitman from the *University of Pennsylvania* and Fred Jelinek from the *IBM T.J. Watson Research Center*.

Learning is also an important part of our annual meetings, so I also thank the individual instructors for our tutorial sessions: Per-Kristian Halvorsen and John Nerbonne; Terry Langendoen and Mitch Marcus; Irene Heim; and David Rumelhart. Mark Liberman deserves special appreciation for his report on the *ACL Data Collection Initiative*.

Last on this list of thanks, but first in our hearts and minds, is Don Walker. Together with Betty Walker, he is truly the "glue" that makes the *Association for Computational Linguistics* stick together. He does everything from soup to nuts for our Association, and particularly for the conferences, from putting out the Call for Papers, to answering every little question, to publishing the Proceedings. All done selflessly, aside from this meagre reward at the front of each year's Proceedings. We should all think of how much he has done for all of us and the field as a whole.

Robert C. Berwick, *Massachusetts Institute of Technology*  
Chair, Program Committee

# CONFERENCE PROGRAM

## WEDNESDAY, 6 JUNE

- 9:00–12:30 **MORNING TUTORIAL SESSIONS:** David Lawrence Hall
- Room 104 *Unification in the Syntax/Semantics Interface*  
Per-Kristian Halvorsen & John Nerbonne
- Room 105 *Tagging Linguistic Information in a Text Corpus*  
Terry Langendoen & Mitch Marcus
- 2:00–5:30 **AFTERNOON TUTORIAL SESSIONS:** David Lawrence Hall
- Room 104 *Discourse Representation Theory*  
Irene Heim
- Room 105 *Connectionism in Natural Language Processing*  
David Rumelhart
- 7:00–9:00 **Conference Registration and Reception**  
Kurtzman Room, William Pitt Union

## THURSDAY MORNING, 7 JUNE

- 9:00–9:15 **Opening remarks and announcements**
- 9:15–9:40 *Polynomial Time Parsing of Combinatory Categorical Grammars*  
K. Vijay-Shanker & David J. Weir
- 9:40–10:05 *Intonation and Structure in Spoken Language Understanding*  
Mark Steedman
- 10:25–10:50 *Prosody, Syntax, and Parsing*  
John Bear & Patti Price
- 10:50–11:15 *Empirical Study of Predictive Powers of Simple Attachment Schemes for Post-Modifier Prepositional Phrases*  
Greg Whittemore, Kathleen Ferrara, & Hans Brunner
- 11:15–11:40 *Structural Disambiguation with Constraint Propagation*  
Hiroshi Maruyama
- 11:40–12:05 *Memory Capacity and Sentence Processing*  
Edward Gibson

## THURSDAY AFTERNOON, 7 JUNE

- 1:30–1:55 *Transforming Syntactic Graphs Into Semantic Graphs*  
Hae-Chang Rim, Jungyun Seo, & Robert F. Simmons
- 1:55–2:20 *A Compositional Semantics for Focusing Subjuncts*  
Daniel Lyons & Graeme Hirst
- 2:20–2:45 *Designer Definites in Logical Form*  
Mary P. Harper
- 3:05–3:30 *Mixed Initiative in Dialogue: An Investigation Into Discourse Segmentation*  
Marilyn Walker & Steve Whittaker

- 3:30–3:55 *Performatives in a Rationally-Based Speech Act Theory*  
Philip R. Cohen & Hector J. Levesque
- 3:55–&4:20 *Normal State Implicature*  
Nancy L. Green
- 4:40–5:05 *The Computational Complexity of Avoiding Conversational Implicatures*  
Ehud Reiter
- 5:05–5:30 *Free Indexation: Combinatorial Analysis and a Compositional Algorithm*  
Sandiway Fong
- 5:30–5:55 *Licensing and Tree Adjoining Grammar in Government Binding Parsing*  
Robert Frank

### FRIDAY MORNING, 8 JUNE

- 9:00-9:25 *A Simplified Theory of Tense Representations and Constraints on Their Composition*  
Michael R. Brent
- 9:25-9:50 *Solving Thematic Divergences in Machine Translation*  
Bonnie Dorr
- 9:50–10:15 *A Syntactic Filter on Pronominal Anaphora for Slot Grammar*  
Shalom Lappin & Michael McCord
- 10:35–11:00 *Acquiring Core Meanings of Words, Represented as Jackendoff-Style Conceptual Structures, from Correlated Streams of Linguistic and Non-Linguistic Input*  
Jeffrey Mark Siskind
- 11:00–12:00 **Structural Sources of Verb Meaning**  
*Lila Gleitman, INVITED SPEAKER*

### FRIDAY AFTERNOON, 8 JUNE

- 1:30–1:55 *Types in Functional Unification Grammars*  
Michael Elhadad
- 1:55–2:20 *Defaults in Unification Grammar*  
Gosse Bouma
- 2:20–2:45 *Expressing Disjunctive and Negative Feature Constraints with Classical First-Order Logic*  
Mark Johnson
- 3:05–3:30 *Lazy Unification*  
Kurt Godden
- 3:30–3:55 *Zero Morphemes in Unification-Based Combinatory Categorical Grammar*  
Chinatsu Aone & Kent Wittenburg
- 3:55–4:20 *The Limits of Unification*  
Robert J.P. Ingria
- 4:40–5:05 *Asymmetry in Parsing and Generating With Unification Grammars: Case Studies from ELU*  
Graham Russell, Susan Warwick, & John Carroll
- 5:05–5:30 *Automated Inversion of Logic Grammars for Generation*  
Tomasz Strzalkowski & Ping Peng
- 5:30–5:55 *Algorithms for Generation in Lambek Theorem Proving*  
Erik-Jan van der Linden & Guido Minnen

- 6:30--7:30     **RECEPTION**  
 Galleria, Forbes Quadrangle
- 7:30--10:00   **BANQUET**  
 Assembly Room, William Pitt Union  
*Presidential Address:* Jerry R. Hobbs

**SATURDAY MORNING, 9 JUNE**

- 9:00-9:25     *Multiple Underlying Systems: Translating User Requests Into Programs to Produce Answers*  
 Robert J. Bobrow, Philip Resnick, Ralph M. Weischedel
- 9:25-9:50     *Computational Structure of Generative Phonology and Its Relation to Language Comprehension*  
 Eric Sven Ristad
- 10:10--11:10   **Stochastic Methods for Context-Free Grammars**  
*Fred Jelinek, INVITED SPEAKER*
- 11:10--12:15   **ACL REPORT, BUSINESS MEETING, & ELECTIONS**
- 11:10--11:30   *A Report on the ACL Data Collection Initiative*  
 Mark Y. Liberman
- 11:30--12:15   **Business Meeting & Elections**  
**Nominations for ACL Offices for 1991**  
*President:* Ralph Grishman, *New York University*  
*Vice President:* Kathy McKeown, *Columbia University*  
*Secretary-Treasurer:* Don Walker, *Bellcore*  
*Executive Committee (1991-1992):* Martha Palmer, *Unisys*  
*Executive Committee (1991-1993):* Fernando Pereira, *AT&T Bell Labs*  
*Nominating Committee (1991-1993):* Jerry Hobbs, *SRI International*

**SATURDAY AFTERNOON, 9 JUNE**

- 1:30--1:55     *Parsing the LOB Corpus*  
 Carl G. de Marcken
- 1:55--2:20     *Automatically Extracting and Representing Collocations for Language Generation*  
 Frank A. Smadja & Kathleen R. McKeown
- 2:20--2:45     *Disambiguating and Interpreting Verb Definitions*  
 Yael Ravin
- 3:05--3:30     *Noun Classification from Predicate-Argument Structures*  
 Donald Hindle
- 3:30--3:55     *Deterministic Left to Right Parsing of Tree Adjoining Languages*  
 Yves Schabes & K. Vijay-Shanker
- 3:55--4:20     *An Efficient Parsing Algorithm for Tree Adjoining Grammars*  
 Karin Harbusch
- 4:40--5:05     *Lexical and Syntactic Rules in a Tree Adjoining Grammar*  
 Anne Abeillé
- 5:05--5:30     *Bottom-Up Parsing Extending Context-Freeness in a Process Grammar Processor*  
 Massimo Marino
- 5:30--5:55     *A Hardware Algorithm for High Speed Morpheme Extraction and Its Implementation*  
 Toshikazu Fukushima, Yutaka Ohyama, & Hitoshi Miyai

# TUTORIALS

## **Unification in the Syntax/Semantics Interface**

Per-Kristian Halvorsen, *Xerox PARC & CSLI*, and John Nerbonne, *Hewlett-Packard & CSLI*

The introduction of unification-based formalisms for linguistic description has had a major impact on syntactic processing techniques and even on syntactic theory, especially in the development of feature-based grammars. More recently the use of unification has found success in semantic processing, raising anew some central issues concerning the syntax/semantics interface (e.g. compositionality and the virtues of "direct" interpretation). The tutorial will present the basic techniques underlying the use of unification in semantic processing, and demonstrate how these are profitably extended. Major topics are (i) Compositionality and systematicity in semantic interpretation revisited from the perspective of unification grammars; (ii) Structure-sharing, the ability to cater to the information shared in linguistic signs, employed not only in logic translations, but also in ancillary information associated with interrogative, relative, anaphoric and reflexive pronouns; (iii) Exploitation of the ability to underspecify information as an attractive option to multiplying semantics analyses in treating ambiguities; (iv) New ways of integrating semantic and syntactic constraints provided by a unitary basis for processing both semantic and syntactic information.

## **Tagging Linguistic Information in a Text Corpus**

Terry Langendoen, *University of Arizona*, and Mitch Marcus, *University of Pennsylvania*

Within the next few months, the widespread availability of extremely large corpora of computer readable texts through the ACL Data Collection Initiative will provide a new research tool for both linguistics and computational linguistics. Work is also beginning on explicit annotation of large corpora with several aspects of linguistic structure. This tutorial is intended for those who wish either to engage in text annotation projects of their own, or to exploit the availability of new large annotated text and transcribed speech corpora such as the Penn Treebank. We will discuss at length a proposed set of draft standards developed by the internationally sponsored Text Encoding Initiative for the encoding and interchange of phonological, morphological and syntactic information, focussing both on the issues involved in developing such standards and on the particulars of the present proposed standard. We will outline a set of general design issues for large corpus annotation efforts, with focus on achieving the productivity rates necessary to annotate millions of words of text. The particular schemes used within the Penn Treebank project will be presented. We will also briefly survey a range of recent results deriving from the use of both annotated and unannotated corpora as research tools, using statistical, information-theoretic, and classical-AI symbolic methodologies.

## **Discourse Representation Theory**

Irene Heim, *Massachusetts Institute of Technology*

Discourse Representation Theory (DRT) refers to a semantic analysis of indefinites and anaphora advanced by Kamp (1981) and Heim (1982) and to modifications and extensions thereof in subsequent work by various authors. Like any concrete piece of linguistic analysis, DRT is a package deal and eclectically incorporates several ideas and assumptions that have been around in the philosophy and linguistics literature for a while and will probably continue to be explored long after DRT is obsolete. The main goal of this tutorial is not just to show how the DRT analysis works, but to isolate the distinct items in the 'package' and clarify the role of each within DRT as well as in the broader context of semantic theory. We will specifically attend to aspects of the following general issues: the syntax-semantics interface; context-dependency and presupposition; and the syntax and semantics of natural language quantification structures.

## **Connectionist Approaches to Language**

David Rumelhart, *Stanford University*

This tutorial will provide an overview and introduction to connectionist techniques and their applications to linguistic information processing. I will discuss issues of learning and computation via constraint satisfaction, and will illustrate the applications of these to language processing problems.

## TABLE OF CONTENTS

<i>Polynomial Time Parsing of Combinatory Categorical Grammars</i> K. Vijay-Shanker & David J. Weir .....	1
<i>Intonation and Structure in Spoken Language Understanding</i> Mark Steedman .....	9
<i>Prosody, Syntax, and Parsing</i> John Bear & Patti Price .....	17
<i>Empirical Study of Predictive Powers of Simple Attachment Schemes for Post-modifier Prepositional Phrases</i> Greg Whittemore, Kathleen Ferrara, & Hans Brunner .....	23
<i>Structural Disambiguation with Constraint Propagation</i> Hiroshi Maruyama .....	31
<i>Memory Capacity and Sentence Processing</i> Edward Gibson .....	39
<i>Transforming Syntactic Graphs Into Semantic Graphs</i> Hae-Chang Rim, Jungyun Seo, & Robert F. Simmons .....	47
<i>A Compositional Semantics for Focusing Subjuncts</i> Daniel Lyons & Graeme Hirst .....	54
<i>Designer Definites in Logical Form</i> Mary P. Harper .....	62
<i>Mixed Initiative in Dialogue: An Investigation Into Discourse Segmentation</i> Marilyn Walker & Steve Whittaker .....	70
<i>Performatives in a Rationally-Based Speech Act Theory</i> Philip R. Cohen & Hector J. Levesque .....	79
<i>Normal State Implicature</i> Nancy L. Green .....	89
<i>The Computational Complexity of Avoiding Conversational Implicatures</i> Ehud Reiter .....	97
<i>Free Indexation: Combinatorial Analysis and a Compositional Algorithm</i> Sandiway Fong .....	105
<i>Licensing and Tree Adjoining Grammar in Government Binding Parsing</i> Robert Frank .....	111
<i>A Simplified Theory of Tense Representations and Constraints on Their Composition</i> Michael R. Brent .....	119
<i>Solving Thematic Divergences in Machine Translation</i> Bonnie Dorr .....	127
<i>A Syntactic Filter on Pronominal Anaphora for Slot Grammar</i> Shalom Lappin & Michael McCord .....	135
<i>Acquiring Core Meanings of Words, Represented as Jackendoff-Style Conceptual Structures, from Correlated Streams of Linguistic and Non-Linguistic Input</i> Jeffrey Mark Siskind .....	143



<i>Types in Functional Unification Grammars</i>	
Michael Elhadad .....	157
<i>Defaults in Unification Grammar</i>	
Gosse Bouma .....	165
<i>Expressing Disjunctive and Negative Feature Constraints with Classical First-Order Logic</i>	
Mark Johnson .....	173
<i>Lazy Unification</i>	
Kurt Godden .....	180
<i>Zero Morphemes in Unification-Based Combinatory Categorical Grammar</i>	
Chinatsu Aone & Kent Wittenburg .....	188
<i>The Limits of Unification</i>	
Robert J.P. Ingria .....	194
<i>Asymmetry in Parsing and Generating With Unification Grammars: Case Studies from ELU</i>	
Graham Russell, Susan Warwick, & John Carroll .....	205
<i>Automated Inversion of Logic Grammars for Generation</i>	
Tomasz Strzalkowski & Ping Peng .....	212
<i>Algorithms for Generation in Lambek Theorem Proving</i>	
Erik-Jan van der Linden & Guido Minnen .....	220
<i>Multiple Underlying Systems: Translating User Requests Into Programs to Produce Answers</i>	
Robert J. Bobrow, Philip Resnick, Ralph M. Weischedel .....	227
<i>Computational Structure of Generative Phonology and Its Relation to Language Comprehension</i>	
Eric Sven Ristad .....	235
<i>Parsing the LOB Corpus</i>	
Carl G. de Marcken .....	243
<i>Automatically Extracting and Representing Collocations for Language Generation</i>	
Frank A. Smadja & Kathleen R. McKeown .....	252
<i>Disambiguating and Interpreting Verb Definitions</i>	
Yael Ravin .....	260
<i>Noun Classification from Predicate-Argument Structures</i>	
Donald Hindle .....	268
<i>Deterministic Left to Right Parsing of Tree Adjoining Languages</i>	
Yves Schabes & K. Vijay-Shanker .....	276
<i>An Efficient Parsing Algorithm for Tree Adjoining Grammars</i>	
Karin Harbusch .....	284
<i>Lexical and Syntactic Rules in a Tree Adjoining Grammar</i>	
Anne Abeillé .....	292
<i>Bottom-Up Parsing Extending Context-Freeness in a Process Grammar Processor</i>	
Massimo Marino .....	299
<i>A Hardware Algorithm for High Speed Morpheme Extraction and Its Implementation</i>	
Toshikazu Fukushima, Yutaka Ohyama, & Hitoshi Miyai .....	307

## AUTHOR INDEX

Anne Abeillé .....	292	Michael McCord .....	135
Chinatsu Aone .....	188	Kathleen R. McKeown .....	252
John Bear .....	17	Guido Minnen .....	220
Robert J. Bobrow .....	227	Hitoshi Miyai .....	307
Gosse Bouma .....	165	Yutaka Ohyama .....	307
Michael R. Brent .....	119	Ping Peng .....	212
Hans Brunner .....	23	Patti Price .....	17
John Carroll .....	205	Yael Ravin .....	260
Philip R. Cohen .....	79	Ehud Reiter .....	97
Carl G. de Marcken .....	243	Philip Resnick .....	227
Bonnie Dorr .....	127	Hae-Chang Rim .....	47
Michael Elhadad .....	157	Eric Sven Ristad .....	235
Kathleen Ferrara .....	23	Graham Russell .....	205
Sandiway Fong .....	105	Yves Schabes .....	276
Robert Frank .....	111	Jungyun Seo .....	47
Toshikazu Fukushima .....	307	Robert F. Simmons .....	47
Edward Gibson .....	39	Jeffrey Mark Siskind .....	143
Kurt Godden .....	180	Frank A. Smadja .....	252
Nancy L. Green .....	89	Mark Steedman .....	9
Mary P. Harper .....	62	Tomasz Strzalkowski .....	212
Karin Harbusch .....	284	Erik-Jan van der Linden .....	220
Donald Hindle .....	268	K. Vijay-Shanker .....	1, 276
Graeme Hirst .....	54	Marilyn Walker .....	70
Robert J.P. Ingria .....	194	Susan Warwick .....	205
Mark Johnson .....	173	David J. Weir .....	1
Shalom Lappin .....	135	Ralph M. Weischedel .....	227
Hector J. Levesque .....	79	Steve Whittaker .....	70
Daniel Lyons .....	54	Greg Whitemore .....	23
Massimo Marino .....	299	Kent Wittenburg .....	188
Hiroshi Maruyama .....	31		