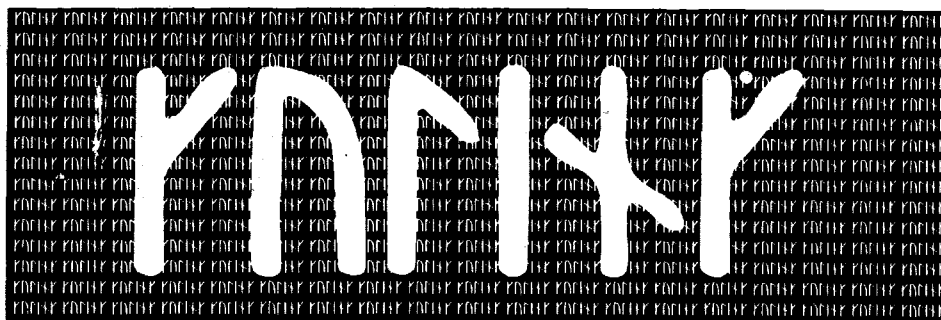


# COLING-96

The 16th International Conference on Computational Linguistics



## PROCEEDINGS

VOL. 1

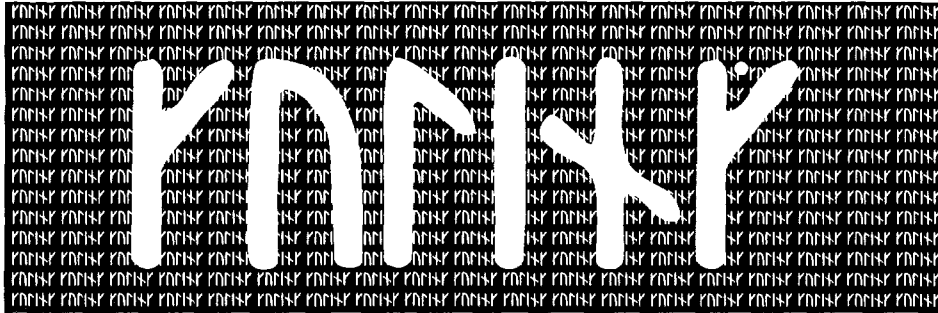
AUGUST 5 - 9, 1996

COPENHAGEN, DENMARK  
Center for Sprogteknologi

**Under the Patronage of  
Her Royal Highness Princess Alexandra of Denmark**

# COLING-96

The 16th International Conference on Computational Linguistics



## PROCEEDINGS

VOL. 1

AUGUST 5 - 9, 1996

COPENHAGEN, DENMARK  
Center for Sprogteknologi

© COLING-96 Organizing Committee

c/o Center for Sprogteknologi, [cstadm@cst.ku.dk](mailto:cstadm@cst.ku.dk)

All rights reserved. No part of this publication may be reproduced, stored in retrieval systems or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior notice.

## Preface

In my capacity as Chairman of the Program Committee of Coling 96, I wish to express my gratitude to all my colleagues who have contributed to the reading and selection of about 400 papers submitted for Coling 96. The Program Committee functioned by subcommittees, and each one of 15 subcommittees was chaired by one of the Program Committee members. The chairperson of each subcommittee chose the members of the subcommittee, and assigned two members to each paper to be reviewed. More than 150 members were finally involved in the hard and sometimes painful task of selecting (and rejecting !) papers from such high quality papers.

I also wish to express my appreciation to Dr.S.Ananiadou, Dr.H.Iida and Dr.S.Ikehara for their support in classifying the papers for subcommittees and collecting the results. They were members of the Program Committee and happened to be in Japan at the critical moment. While we classified papers and distributed them to subcommittees, quite a few papers were found relevant to more than one subject field. For such cases, we asked two subcommittees to cross-review the same paper. When two subcommittees reached different judgements, though such cases were rather rare, I took the final decisions in consultation with the chairpersons of the subcommittees. Therefore, a paper was read carefully at least by three people and in some cases by six or more reviewers.

Computational linguistics is, as its very nature, an interdisciplinary field. It is becoming more and more so with rapid expansion and development of the field. When I was asked to be the Chairman of the Program Committee, I was, honestly, scared of the vast and diverse topics to be covered and was not at all sure whether papers in such diverse fields could be judged in an appropriate way. The very nature of Coling as a single, truly international conference of the field also scared me. It implies that papers are submitted from geographically very scattered areas and that the reviewing process involves reviewers scattered around the world. In fact, we received papers from 39 countries and had reviewers from 21 different countries.

However, these fears were resolved, I believe, satisfactorily by generous devotion of the members of the Program Committee and subcommittees. Each of the 15 subcommittees played the role of Program Committee for a conference of a moderate size. The danger of independent subcommittees is that Coling may become a collection of independent smaller conferences, which in turn may not reflect the intricate mutual relationships among subfields in Computational Linguistics (think about a paper on "discourse model for machine translation" !). I believe that this danger was avoided by cross-reviewing the same papers, and by the co-operative attitude of the Program Committee members involved.

In the Call for Papers, I abolished the distinction between topical papers and project notes. This is because I observed the distinction between these two categories had become somewhat blurred, and sometimes caused unnecessary confusions. However, quite a few papers were rejected with such comments as "the paper is very informative though it does not propose a new research idea", "interesting for demonstration", etc. Since Coling, the largest conference of our field, should play the role of a platform for exchanging first-hand information, I decided to have poster/demonstration forums for these papers. I was pleased that most of the authors of these papers agreed to present their papers in these forums, even though my decision was made after the first letters of acceptance were sent out.

In the panels, we tried to cover topics which could not be covered by research papers but which are crucial for our field, such as social implications of the technology originating out from our research, the funding situation for our field, etc. For some reason, the number of papers in fields such as language processing in speech systems, new applications of NLP, etc. are much less than I expected and hoped. Some of the panels address these issues and I hope that these panels will stimulate new research in these areas.

We adopted anonymous review for the first time in Coling. It seems, though this is not certain, that this had the effect of diversifying the research groups of the accepted papers. We also allowed

e-mail submissions. Though this caused some problems and inconvenienced the authors of some papers, it worked fairly well.

Last but not least, I would like to express my gratitude to the members of the NTT machine translation research group and the members of the Computational Linguistics Laboratory of the University of Tokyo, in particular, Ms.I.Watanabe, Mr.Matsuo, Mr.K.Ogura, Mr.F.Bond, and Ms.Ootsuka, who printed out the e-mail submissions, maintained the data base of submitted papers, and helped to answer various enquiries from authors, etc. Without their efforts, the life of the Program Chairman would be intolerable.

J.TSUJII

Program Chair

CCL, UMIST and the University of Tokyo

## **Program Committee**

Program Chairman: J. Tsujii (CCL-UMIST & Univ. of Tokyo, UK/Japan: ICCL member)

A.Joshi (U.Penn, USA: ICCL member), S.G.Pulman (Univ. of Cambridge and SRI International, UK), A.Ramsay (CCL-UMIST, UK), E.Hajicova (Charles Univ., Czech.Rep.: ICCL member), E.Brill (John Hopkins Univ., USA), N.Calzolari (Istituto di Linguistica Computazionale, Italy: ICCL member), S.Ikehara (NTT, Japan), R.Grishman (New York Univ., USA), H.Iida (ATR, Japan), S.Ananiadou (MMU and CCL-UMIST, UK), G.Sabah (LIMSI-CNRS, France), R.Dale (Microsoft, Australia), W.Wahlster (DFKI, Germany), K.Koskenniemi (Univ. of Helsinki, Finland).

## **The International Committee on Computational Linguistics (ICCL):**

C. Boitet (France), N. Calzolari (Italy), E. Hajicova (Czech Rep.), B. Harris (Canada), K. Heggstad (Norway), A.Joshi (USA), H. Karlgren (Sweden), M. Kay (President, USA), O. Kulagina (Russia), W. Lenders (Germany), M. Nagao (Japan), H. Schnelle (Germany), P. Sgall (Czech Rep.), H. Tanaka (Japan), J. Tsujii (UK), H. Wada (Honorary, Japan), Y. Wilks (UK), A. Zampolli (Italy)

## **Organizing Committee**

Chair: B. Maegaard (Center for Sprogteknologi, Denmark)

Lina Henriksen (Center for Sprogteknologi, Denmark)

Lise Damsgaard Jørgensen (Center for Sprogteknologi, Denmark)

Lene Offersgaard (Center for Sprogteknologi, Denmark)

Jørgen Olsen (Institut for Humanistisk Informatik, University of Copenhagen, Denmark)

Jette Saaby (Center for Sprogteknologi, Denmark)

Nancy L. Underwood (Center for Sprogteknologi, Denmark)

B. Ørsnes (Center for Sprogteknologi, Denmark)

## **Support was received from:**

Research Council for the Humanities

European Commission, Directorate General XIII

Tele Danmark Research & Development

The Danish Ministry of Education, Tipsmidlerne

ELSNET Foundation

Scandinavian Airlines

## Subcommittee members

André, E  
Asher, N  
Bateman, J  
Bennett, P  
Bennett, W.S  
Bhatt, R  
Block, H.U  
Boisen, S  
Bourigault, D  
Bowker, L  
Brandow, R  
Briscoe, T  
Bunt, H  
Callan, J  
Campbell, N  
Carlson, L  
Carroll, J  
Cazione, R  
Chandrasekhar, R  
Choi, K-S  
Cole, R  
Collier, R  
Condamines, A  
Cowie, J  
Cunningham, H  
Daille, B  
Doran, C  
Dowding, J  
Dras, M  
Elhadad, M  
Erbach, G  
Estival, D  
Fillmore, C  
Finkler, W  
Fox, C  
Frank, F  
Friedman, C  
Gaizauskas, R  
Gale, W  
Garside, R  
Gazdar, G  
Glass, J  
Goerz, G  
Grosz, B  
Guthrie, L  
Hahn, W von  
Hajic, J  
Hayamizu, S  
Heid, U  
Hepple, M  
Hindle, D  
Hirschberg, J  
Hockey, B  
Hoppner, W  
Horacek, H  
Hovy, E  
Humphreys, K  
Hurskainen, A  
Ingram, J  
Isabelle, P  
Jacquemin, C  
Kageura, K  
Kaplan, R  
Karligen, J  
Karttunen, L  
Katz, S  
Kawahara, T  
Kayser, D  
Knight, K  
Knott, A  
Knowles, F  
Krifka, M  
Kruiffova, I  
Kulick, S  
Lauer, M  
Leacock, C  
Leech, G  
Lewin, I  
Lin, F  
Linden, K.V  
Maier, E  
Marcus, M  
Maruyama, H  
McDonald, D  
McKay, H  
McKeown, K  
Manning, C.D  
Matsumoto, Y  
Meng, H  
Mitkov, R  
Nagata, J  
Neumann, G  
Nkweni-Azeh, B  
Noord, G.V  
Nuallain, S  
Oflazer, K  
Ogura, K  
Oliva, K  
Oviatt, S  
Panevova, J  
Park, J.C  
Peregrin, J  
Petrkevici, V  
Pustejovsky, J  
Rambow, O  
Ramsay, R  
Rayner, M  
Reiter, E  
Reithinger, N  
Resnik, P  
Reyle, U  
Ristad, E  
Ritchie, G  
Roche, E  
Roeck, A. de  
Rooth, M  
Rosen, A  
Roukos, S  
Samuelsson, C  
Sanfilippo, A  
Sarkar, A  
Schabes, Y  
Schüetze, H  
Sekine, S  
Seneff, S  
Sgall, P  
Sågvall Hein, A  
Smedt, K. de  
Somers, H  
Sproat, R  
Srinivas, B  
Stede, M  
Steedman, M  
Stevenson, M  
Strzalkowski, T  
Sundheim, B  
Tapanainen, P  
Thompson, H  
Tokunaga, T  
Trost, H  
Tufis, D  
Tulloch, A  
Tzoukermann, E  
Uratani, N  
Victorri, B  
Vijayshanker, K  
Voutilainen, A  
Wang, J  
Webber, B  
Wilks, Y  
Wiren, M  
Wu, D  
Yokoyama, O  
Zaenen, A  
Zock, M  
Zue, V



## CONTENTS

### INVITED SPEAKERS

Barbara J. <u>Grosz</u> Discovering the Sounds of Discourse Structure . . . . .	1
Willem J.M. <u>Levelt</u> A Theory of Lexical Access in Speech Production . . . . .	3

### PAPERS

Hang Li, Naoki <u>Abe</u> Clustering Words with the MDL Principle . . . . .	4
Hang Li, Naoki <u>Abe</u> Learning Dependencies between Case Frame Slots . . . . .	10
Eneko <u>Agirre</u> , German Rigau Word Sense Disambiguation Using Conceptual Density . . . . .	16
Hiroyuki Kaji and Toshiko <u>Aizono</u> Extracting Word Correspondences from Bilingual Corpora Based on Word Co-occurrence Information . . . . .	23
Pascal <u>Amsili</u> , Nabil Hathout Computational Semantics of Time/Negation Interaction . . . . .	29
Jan W. <u>Amtrup</u> , Jörg Benra Communication in Large distributed AI Systems for Natural Language Processing . . . . .	35
Katerina T. Frantzi, Sophia <u>Ananiadou</u> Extracting Nested Collocations . . . . .	41
Jean-Yves <u>Antoine</u> Parsing Spoken Language Without Syntax . . . . .	47
Chinatsu <u>Aone</u> , Kevin Hausman Unsupervised Learning of a Rule-based Spanish Part of Speech Tagger . . . . .	53
Fujii <u>Atsushi</u> , Inui Kentaro, Tokunaga Takenobu, Tanaka Hozumi To what extent does Case contribute to Verb Sense Disambiguation? . . . . .	59
Caroline <u>Barrière</u> , Fred Popowich Concept clustering and knowledge integration from a children's dictionary . . . . .	65
Anton <u>Batliner</u> , Anke Feldhaus, Stefan Geissler, Andreas Kiessling, Tibor Kiss, Ralf Kompe, Elmar Nöth Integrating Syntactic and Prosodic Information for the Efficient Detection of Empty Categories . . . . .	71
Marco <u>Battista</u> , Vito Pirrelli Monotonic Paradigmatic Schemata in Italian Verb Inflection . . . . .	77
Sergei Nirenburg, Kavi Mahesh, Stephen <u>Beale</u>	

Measuring Semantic Coverage . . . . .	83
Kenneth R. <u>Beesley</u> Arabic Finite-State Morphological Analysis and Generation . . . . .	89
Walter Daelemans, Peter <u>Berck</u> , Steven Gillis Unsupervised Discovery of Phonological Categories through Supervised Learning of Morphological Rules . . . . .	95
Arturo Trujillo, Simon <u>Berry</u> Connectivity in Bag Generation . . . . .	101
Ezra <u>Black</u> , Hideki Kashioka, Stephen Eubank, Roger Garside, Geoffrey Leech, David Magerman Beyond Skeleton Parsing : Producing a Comprehensive Large-Scale General-English Treebank With Full Grammatical Analysis . . . . .	107
Christopher Kennedy, Branimir <u>Boguraev</u> Anaphora for Everyone: Pronominal Anaphora Resolution without a Parser . . . . .	113
Christian <u>Boitet</u> , Mutsuko Tomokiyo Theory and Practice of Ambiguity labelling with a View to interactive Disambiguation in Text and Speech MT . . . . .	119
Francis <u>Bond</u> , Kentaro Ogura, Satoru Ikehara Classifiers in Japanese-to-English Machine Translation . . . . .	125
Johan <u>Bos</u> , Björn Gambäck, Christian Lieske, Yoshiki Mori, Manfred Pinkal, Karsten Worm Compositional Semantics in Verbmobil . . . . .	131
Jacques <u>Bouaud</u> , Bruno Bachimont, and Pierre Zweigenbaum Processing Metonymy: a Domain-Model Heuristic Graph Traversal Approach . . . . .	137
Pierrette <u>Bouillon</u> Mental State Adjectives: the Perspective of Generative Lexicon . . . . .	143
António H. <u>Branco</u> Branching Split Obliqueness at the Syntax-Semantics Interface . . . . .	149
Carl Vogel, Ulrike Hahn, Holly <u>Branigan</u> Cross-Serial Dependencies Are Not Hard to Process . . . . .	157
Andrew <u>Bredenkamp</u> , Stella Markantonatou, Louisa Sadler Lexical Rules: What are they? . . . . .	163
Ralf D. <u>Brown</u> Example-Based Machine Translation in the Pangloss System . . . . .	169
Flora Ramírez <u>Bustamante</u> , Fernando Sánchez León GramCheck: A Grammar and Style Checker . . . . .	175
Miriam <u>Butt</u> , Christian Fortmann, Christian Rohrer Syntactic Analyses for Parallel Grammars: Auxiliaries and Genitive NPs . . . . .	182
Finn Dag <u>Buø</u> , Alex Waibel FeasPar - A Feature Structure Parser Learning to Parse Spoken Language . . . . .	188
Marie-Hélène <u>Candito</u> A principle-based hierarchical representation of LTAGs . . . . .	194

Pak-kwong Wong, Chorkin <u>Chan</u> Chinese Word Segmentation based on Maximum Matching and Word Binding Force	200
Hubert Hin-Cheung Law, Chorkin <u>Chan</u> N-th Order Ergodic Multigram HMM for Modeling of Languages without Marked Word Boundaries	204
Jason J. S. <u>Chang</u> , Sur-Jin Ker Aligning More Words with High Precision for Small Bilingual Corpora	210
Kuang-hua Chen, Hsin-Hsi <u>Chen</u> A Rule-Based and MT-Oriented Approach to Prepositional Phrase Attachment	216
Hsin-Hsi <u>Chen</u> , Jen-Chang Lee Identification and Classification of Proper Nouns in Chinese Texts	222
Jung H. Shin, Key-Sun <u>Choi</u> , Young S. Han Bilingual Knowledge Acquisition from Korean-English Parallel Corpus Using Alignment Method ( Korean-English Alignment at Word and Phrase Level )	230
Sung-Young Jung, Young C. Park, Key-Sun <u>Choi</u> , Youngwhan Kim Markov Random Field based English Part-of-Speech Tagging System	236
Salvador <u>Climent</u> Semantics of Portions and Partitive Nouns for NLP	243
José <u>Coch</u> Evaluating and comparing three Text-production Techniques	249
Marc Dymetman, Max <u>Copperman</u> Extended Dependency Structures and their Formal Interpretation	255
Josef van Genabith, Dick <u>Crouch</u> Direct and Underspecified Interpretations of LFG f-structures	262
R.I. Damper, J.F.G. <u>Eastmond</u> Pronouncing Text by Analogy	268
Marc B. Vilain and David <u>Day</u> Finite-state Phrase Parsing by Rule Sequences	274
Thierry <u>Declerck</u> Dealing with Cross-Sentential Anaphora Resolution in ALEP	280
Paul Schmidt, Sibylle Rieder, Axel Theofilidis, Thierry <u>Declerck</u> Lean Formalisms, Linguistic Theory, and Applications. Grammar Development in ALEP	286
Judy <u>Delin</u> , Donia Scott, Anthony Hartley Language-Specific Mappings from Semantics to Syntax	292
Jerneja Gros, Ivo Ipšič, Simon <u>Dobrišek</u> , France Mihelič, Nikola Pavešić Segmentation and Labelling of Slovenian Diphone Inventories	298
Kohji <u>Dohsaka</u> , Akira Shimazu A Computational Model of Incremental Utterance Production in Task-Oriented Dialogues	304

Shin-ichiro Kamei, Kazunori Muraki, Shin'ichi <u>Doi</u> Lexical Information for Determining Japanese Unbounded Dependency . . . . .	310
Michael <u>Dorna</u> , Martin C. Emele Semantic-based Transfer . . . . .	316
Bonnie <u>Dorr</u> and Doug Jones Role of Word Sense Disambiguation in Lexical Acquisition: Predicting Seman- tics from Syntactic Cues . . . . .	322
Laila <u>Dybkjær</u> , Niels Ole Bernsen, Hans Dybkjær Grice Incorporated. Cooperativity in Spoken Dialogue . . . . .	328
Kurt <u>Eberle</u> Disambiguation by Information Structure in DRT . . . . .	334
Jason <u>Eisner</u> Three New Probabilistic Models for Dependency Parsing: An Exploration . . . . .	340
Keith Vander Linden, Barbara Di <u>Eugenio</u> A Corpus Study of Negative Imperatives in Natural Language Instructions . . . . .	346
Barbara Di <u>Eugenio</u> The Discourse Functions of Italian Subjects: a centering Approach . . . . .	352
Yan Qu, Carolyn P. Rosé, Barbara Di <u>Eugenio</u> Using Discourse Predictions for Ambiguity Resolution . . . . .	358
Cécile <u>Fabre</u> Interpretation of Nominal Compounds: Combining Domain-Independent and Domain-Specific Information . . . . .	364
Laurel <u>Fais</u> Lexical Accommodation in Machine-Mediated Interactions . . . . .	370
Simonetta Montemagni, Stefano <u>Federici</u> , Vito Pirrelli Resolving syntactic Ambiguities with Lexico-semantic Patterns: an Analogy-- based Approach . . . . .	376
Beate <u>Firzlaff</u> , Daniela S. Kunz Discourse Semantics Meets Lexical Field Semantics . . . . .	382
Ingrid <u>Fischer</u> , Martina Keil Parsing decomposable Idioms . . . . .	388
George <u>Foster</u> , Pierre Isabelle, Pierre Plamondon Word Completion: A First Step Toward Target-Text Mediated IMT . . . . .	394
Keiichi Sakai, Tsuyoshi Yagisawa, and Minoru <u>Fujita</u> A CD-ROM Retrieval System With Multiple Dialogue Agents . . . . .	400
Fumiyo <u>Fukumoto</u> , Yoshimi Suzuki An Automatic Clustering of Articles Using Dictionary Definitions . . . . .	406
Osamu <u>Furuse</u> , Hitoshi Iida Incremental Translation Utilizing Constituent Boundary Patterns . . . . .	412
Takahiro Wakao, Robert J. <u>Gaizauskas</u> , Yorick Wilks Evaluation of An Algorithm for the Recognition and Classification of Proper Names . . . . .	418

Anthony F. <u>Gallippi</u> Learning to Recognize Names Across Languages . . . . .	424
Claire <u>Gardent</u> , Michael Kohlhase Focus and Higher-Order Unification . . . . .	430
Short, S., Shiu S. <u>Garigliano</u> , R. Distributedness and Non-Linearity of LOLITA's Semantic Network . . . . .	436
Alon Lavi, Donna <u>Gates</u> , Marsal Gavaldà, Laura Mayfield, Alex Waibel, Lori Levin Multi-lingual Translation of Spontaneously Spoken Language in a Limited Domain . . . . .	442
John <u>Griffith</u> Modularizing Contexted Constraints . . . . .	448
Edmund <u>Grimley-Evans</u> , George Anton Kiraz, Stephen G. Pulman Compiling a Partition-Based Two-Level Formalism . . . . .	454
Adam Meyers, Roman Yangarber and Ralph <u>Grishman</u> Alignment of Shared Forests for Bilingual Corpora . . . . .	460
Ralph <u>Grishman</u> , Beth Sundheim Message Understanding Conference - 6: A Brief History . . . . .	466
Catherine Macleod, Adam Meyers, Ralph <u>Grishman</u> The Influence of Tagging on the Classification of Lexical Complements . . . . .	472
Jon Atle <u>Gulla</u> , Sjur Nørstebø Moshagen A Sign Expansion Approach to Dynamic, Multi-Purpose Lexicons . . . . .	478
Günther <u>Görz</u> , Marcus Kessler, Hans Weber, Jörg Spilker Research on Architectures for Integrated Speech/Language Systems in Verb- mobil . . . . .	484
Benoit <u>Habert</u> , Elie Naulleau, Adeline Nazarenko Symbolic Word Clustering for Medium-size Corpora . . . . .	490
Udo <u>Hahn</u> , Michael Strube, Katja Markert Bridging Textual Ellipsis . . . . .	496
Peter Neuhaus, Udo <u>Hahn</u> Restricted Parallelism in Object-Oriented Lexical Parsing . . . . .	502
Hideki Hirakawa, Zhonghui Xu, Kenneth <u>Haase</u> Inherited Feature-based Similarity Measure Based on Large Semantic Hierar- chy and Large Text Corpus . . . . .	508
Young S. <u>Han</u> , Hyouk R. Park, Key-Sun Choi, Kang H. Lee A Probabilistic Approach to Compound Noun Indexing in Korean Texts . . . . .	514
Daniel <u>Hardt</u> Centering in Dynamic Semantics . . . . .	519
Masahiko <u>Haruno</u> , Satoru Ikehara, Takefumi Yamazaki Learning Bilingual Collocations by Word-level Sorting . . . . .	525
Hasida <u>Kôiti</u> Issues in Communication Game . . . . .	531

Mark <u>Hepple</u>	
A Compilation-Chart Method for Linear Categorical Deduction . . . . .	537
Erhard W. <u>Hinrichs</u> , Tsuneko Nakazawa	
Applying Lexical Rules under Subsumption . . . . .	543
Toru <u>Hisamitsu</u> , Yoshihiko Nitta	
Analysis of Japanese Compound Nouns By Direct Text Scanning . . . . .	550
Beryl <u>Hoffman</u>	
Translating into Free Word Order Languages . . . . .	556
Wide R. <u>Hogenhout</u> , Yuji Matsumoto	
Towards a More Careful Evaluation of Broad Coverage Parsing Systems . . . . .	562
Arvi <u>Hurskainen</u>	
Disambiguation of morphological Analysis in Bantu Languages . . . . .	568
Satoru <u>Ikehara</u> , Satoshi Shirai, Hajime Uchino	
A Statistical Method for Extracting Uninterrupted and Interrupted Colloca- tions from Very Large Corpora . . . . .	574
Kumiko Tanaka, Hideya <u>Iwasaki</u>	
Extraction of Lexical Translations from Non-Aligned Corpora . . . . .	580
Shiho Nobesawa, Junya Tsutsumi, Da Jiang <u>Sun</u> , Tomohisa Sano, Kengo Sato, Masa- kazu Nakanishi	
Segmenting Sentences into Linky Strings Using D-bigram Statistics . . . . .	586
Christer <u>Johansson</u>	
Good Bigrams . . . . .	592
Kristiina <u>Jokinen</u>	
Goal Formulation based on Communicative Principles . . . . .	598
Bernard <u>Jones</u>	
Towards a Syntactic Account of Punctuation . . . . .	604
Anoop Sarkar, Aravind <u>Joshi</u>	
Coordination in Tree Adjoining Grammars: Formalization and Implementation . . . . .	610
Masayuki <u>Kameda</u>	
A Portable & Quick Japanese Parser : QJP . . . . .	616
André Kempe, Lauri <u>Karttunen</u>	
Parallel Replacement in Finite State Calculus . . . . .	622
Walter <u>Kasper</u> , Hans-Ulrich Krieger	
Modularizing Codescriptive Grammars for Efficient Parsing . . . . .	628
Naoto <u>Katoh</u> , Tsuyoshi Morimoto	
Statistical Method of Recognizing Local Cohesion in Spoken Dialogues . . . . .	634
Naoyuki Nomura and Kazunori <u>Muraki</u>	
An Empirical Architecture for Verb Subcategorization Frame -- a Lexicon for a Real-world Scale Japanese-English Interlingual MT . . . . .	640
Bill <u>Keller</u>	
An Evaluation Semantics for DATR Theories . . . . .	646

Gen-itiro <u>Kikui</u> Identifying the Coding System and Language of On-line Documents on the Internet . . . . .	652
James <u>Kilbury</u> Top-Down Predictive Linking and Complex-Feature-Based Formalisms . . . . .	658
George A. <u>Kiraz</u> Computing Prosodic Morphology . . . . .	664
Jonas <u>Kuhn</u> An Underspecified HPSG Representation for Information Structure . . . . .	670
Gunnel <u>Källgren</u> Linguistic Indeterminacy as a Source of Errors in Tagging . . . . .	676
Christopher <u>Laenzlinger</u> , Martin S. Ulmann, Eric Wehrli Arguments desperately seeking Interpretation: Parsing German Infinitives . . . . .	681
Shalom <u>Lappin</u> , Hsue-Hueh Shih A Generalized Reconstruction Algorithm for Ellipsis Resolution . . . . .	687
Dong-Young <u>Lee</u> Computation of Relative Social Status on the Basis of Honorification in Korean . . . . .	693
Mark <u>Lee</u> , Yorick Wilks An ascription-based approach to Speech Acts . . . . .	699
Clifford Weinstein, Dinesh Tummala, Young-Suk <u>Lee</u> , Stephanie Seneff Automatic English-to-Korean Text Translation of Telegraphic Messages in a Limited Domain . . . . .	705
Sabine <u>Lehmann</u> , Stephan Oepen, Sylvie Regnier-Prost, Klaus Netter, Veronika Lux, Judith Klein, Kirsten Falkedal, Frederik Fouvry, Dominique Estival, Eva Dauphin, Hervé Compagnion, Judith Baur, Lorna Balkan, Doug Arnold TSNLP --- Test Suites for Natural Language Processing . . . . .	711
Yves <u>Lepage</u> , Ando Shin-ichi Saussurian Analogy: A theoretical Account and its Application . . . . .	717
Vincenzo Lombardo, Leonardo <u>Lesmo</u> An Earley-type Recognizer for Dependency Grammar . . . . .	723
Dekang <u>Lin</u> On the Structural Complexity of Natural Language Sentences . . . . .	729
Cécile Paris, Keith Vander <u>Linden</u> Building Knowledge Bases for the Generation of Software Documentation . . . . .	734
Stefan Wermter, Matthias <u>Löchel</u> Learning Dialog Act Processing . . . . .	740
Isa <u>Maks</u> and Willy Martin MULTITALE: Linking medical Concepts by Means of Frames . . . . .	746
Johannes <u>Matiassek</u> , Harald Trost An HPSG-Based Generator for German An Experiment in the Reusability of Linguistic Resources . . . . .	752

Graham Wilcock, Yuji <u>Matsumoto</u> Reversible delayed lexical Choice in a bidirectional Framework . . . . .	758
I. Dan <u>Melamed</u> Automatic Detection of Omissions in Translations . . . . .	764
Andrei <u>Mikheev</u> Learning Part-Of-Speech Guessing Rules from Lexicon: Extension to Non- Concatenative Operations . . . . .	770
Leonid <u>Mitjushin</u> An Agreement Corrector for Russian . . . . .	776
Tatsunori <u>Mori</u> , Hiroshi Nakagawa Zero Pronouns and Conditionals in Japanese Instruction Manuals . . . . .	782
Yoshiki <u>Mori</u> Multiple Discourse Relations on the Sentential Level in Japanese . . . . .	788
Yasuhiko Watanabe, Masaki <u>Murata</u> , Masahito Takeuchi, Makoto Nagao Document Classification Using Domain Specific Kanji Characters Extracted by X2 Method . . . . .	794
Stefan <u>Müller</u> Yet Another Paper About Partial Verb Phrase Fronting in German . . . . .	800
Masaaki <u>Nagata</u> Context-Based Spelling Correction for Japanese OCR . . . . .	806
Hiromi <u>Nakaiwa</u> , Satoshi Shirai Anaphora Resolution of Japanese Zero Pronouns with Deictic Reference . . . . .	812
Takehiro <u>Nakayama</u> Content-Oriented Categorization of Document Images . . . . .	818
Tetsuya <u>Nasukawa</u> Full-text Processing: Improving a practical NLP System based on Surface Information within the Context . . . . .	824
John <u>Nerbonne</u> , Petra Smit GLOSSER-RuG: in Support of Reading . . . . .	830
Stephan Vogel, Hermann <u>Ney</u> , Christoph Tillmann HMM-Based Word Alignment in Statistical Translation . . . . .	836
Victor Raskin, Sergei <u>Nirenburg</u> Adjectival Modification in Text Meaning Representation . . . . .	842
Elena <u>Not</u> A Computational Model for Generating Referring Expressions in a Multi- lingual Application Domain . . . . .	848
Kemal <u>Oflazer</u> , Okan Yilmaz A Constraint-based Case Frame Lexicon . . . . .	854
Kemal <u>Oflazer</u> Error-tolerant Tree Matching . . . . .	860



Masahiro <u>Oku</u>	Analyzing Japanese Double-Subject Construction having an Adjective Predicate . . . . .	865
<u>Okumura</u> Manabu, Tamura Kouji	Zero Pronoun Resolution in Japanese Discourse. Based on Centering Theory . . . . .	871
Lluís <u>Padró</u>	POS Tagging Using Relaxation Labelling . . . . .	877
Vladimir <u>Pericliev</u>	Learning Linear Precedence Rules . . . . .	883
Allan <u>Ramsay</u>	Aspect and Aktionsart: Fighting or Cooperating? . . . . .	889
Christer <u>Samuelsson</u>	Handling Sparse Data by Successive Abstraction . . . . .	895
Ken <u>Satoh</u>	Disambiguation by Prioritized Circumscription . . . . .	901
Michael <u>Schiehlen</u>	Semantic Construction from Parse Forests . . . . .	907
Satoshi <u>Sekine</u>	Modeling Topic Coherence for Speech Recognition . . . . .	913
Hadar <u>Shemtov</u>	Generation of Paraphrases from Ambiguous Logical Forms . . . . .	919
Hideo <u>Shimazu</u> and Yosuke Takashima	Multi-Modal-Method: A Design Method for Building Multi-Modal Systems . . . . .	925
Tomek <u>Strzalkowski</u> , Jin Wang	A Self-Learning Universal Concept Spotter . . . . .	931
Roland <u>Stuckardt</u>	Anaphor Resolution and the Scope of Syntactic Constraints . . . . .	937
Hideki <u>Tanaka</u>	Decision Tree Learning Algorithm with Structured Attributes: Application to Verbal Case Frame Acquisition . . . . .	943
Kentaro <u>Torisawa</u> , Jun'ichi Tsujii	Computing Phrasal-Signs in HPSG Prior to Parsing . . . . .	949
Naohiko <u>Uramoto</u>	Positioning Unknown Words in a Thesaurus by Using Information Extracted from a Corpus . . . . .	956
Akira <u>Utsumi</u>	A Unified Theory of Irony and Its Computational Formalization . . . . .	962
Takehito <u>Utsuro</u>	Sense Classification of Verbal Polysemy based-on Bilingual Class/Class Association . . . . .	968

<u>Hideo Watanabe</u>	
A Method for Abstracting Newspaper Articles by Using Surface Clues . . . . .	974
<u>Jürgen Wedekind</u>	
On Inference-Based Procedures for Lexical Disambiguation . . . . .	980
<u>Klaus Zechner</u>	
Fast Generation of Abstracts from General Domain Text Corpora by Extrac- ting Relevant Sentences . . . . .	986
<u>Michael Zock</u>	
The Power of Words in Message Planning . . . . .	990
 <u>PANELS</u>	
 <u>PANEL 1</u>	
<u>Steven Krauwer, Joseph Mariani</u>	
"Is Speech Language?" . . . . .	996
 <u>PANEL 2</u>	
<u>Hans Karlgren</u>	
Distortion or Improvement - Effects of Information Technology on the Deve- lopment of Natural Languages . . . . .	997
 <u>PANEL 3</u>	
<u>Bente Maegaard</u>	
Evaluation of NLP Systems . . . . .	1000
 <u>PANEL 4</u>	
<u>Michael Zock</u>	
Computational Linguistics and its Use in real World: the Case of Computer Assisted Language Learning . . . . .	1002
<u>Gerard Kempen</u>	
Human Language Technology can modernize Writing and Grammar Instruc- tion . . . . .	1005
<u>Henry Hamburger</u>	
" Yes! NLP-based FL-ITS will be Important" . . . . .	1007
<u>Dan Tufis</u>	
Call: The Potential of Lingware and the Use of Empirical Linguistic Data . . . . .	1010
<u>Inui Kentaro</u>	
The Internet a "natural" Channel for Language Learning . . . . .	1012
<u>Robert Debski</u>	
Instrumental Technologies vs Teaching Instruments: A challenge for Compu- tional Linguists . . . . .	1014

## POSTER SESSIONS

Galja <u>Angelova</u> , Kalina Bontcheva NL Domain Explanations in Knowledge Based MAT	1016
Hiroshi Echizen-ya, Kenji <u>Araki</u> , Toshio Momouchi, Koji Tochinai Machine Translation Method Using Inductive Learning with Genetic Algorithms	1020
Louissette Emirkanian, Lyne Da Sylva, Lorne H. <u>Bouchard</u> The Implementation of a Computational Grammar of French Using the Grammar Development Environment	1024
Andrew <u>Bredenkamp</u> , Frederik Fouvry, Thierry Declerck, Bradley Music Efficient Integrated Tagging of Word Constructs	1028
Gian Piero <u>Zarri</u> NKRL, a Knowledge Representation Language for "Narrative" Natural Language Processing	1032
Frédérique Segond, Giuseppe Valetto, Elisabeth <u>Breidt</u> Formal Description of Multi-Word Lexemes with the Finite-State Formalism IDAREX	1036
R. <u>Chandrasekar</u> , Christine Doran, B. Srinivas Motivations and Methods for Text Simplification	1041
Chu-Ren Huang, Keh-jiann Chen, Li-Li <u>Chang</u> Segmentation Standard for Chinese Natural Language Processing	1045
Kim Seongyong, Key-Sun <u>Choi</u> Korean Language Engineering: Current Status of the Information Platform	1049
Damien Genthial, Jacques <u>Courtin</u> , Jacques Menézo Distributing and Porting General Linguistic Tools	1053
Hamish <u>Cunningham</u> , Yorick Wilks, Robert J. Gaizauskas GATE - a General Architecture for Text Engineering	1057
Eva <u>Dauphin</u> , Veronika Lux Corpus-based annotated Test Set for Machine Translation Evaluation by an Industrial User	1061
Eric M. Visser, Masaru <u>Fuji</u> Using Sentence Connectors for evaluating MT Output	1066
Haodong Wu, Teiji <u>Furugori</u> Prepositional Phrase Attachment Through a Hybrid Disambiguation Model	1070
Hitoshi Iida, Eiichiro Sumita, Osamu <u>Furuse</u> Spoken-Language Translation Method Using Examples	1074
Joakim Nivre, Leif <u>Grönqvist</u> , Malin Gustafsson, Torbjörn Lager, Sylvana Sofkova Tagging Spoken Language Using Written Language Statistics	1078
Wakako Kuwahata, Minako <u>Hashimoto</u> Senses of Polysemous Nouns: Building a Computational Lexicon of Basic Japanese Nouns	1082

Kawtrakul Asanee, Thumkanon Chalutip, Jamjanya Thitima, Muangyunnan Parinee, <u>Poolwan Kritsada</u> , Inagaki Yasuyoshi A Gradual Refinement Model for A Robust Thai Morphological Analyzer . . . . .	1086
Hideko Miyoshi, Kenji Sugiyama, Masahiro <u>Kobayashi</u> , Takano Oginō An Overview of the EDR Electronic Dictionary and the Current Status of its Utilization . . . . .	1090
Kiyoshi <u>Kogure</u> , Akira Shimazu, Mikio Nakano Parsing Plans Situation-Dependently in Dialogues . . . . .	1094
Karsten <u>Konrad</u> , Holger Maier, David Milward, Manfred Pinkal An Education and Research Tool for Computational Semantics . . . . .	1098
Evelin <u>Kuusik</u> Learning Morphology: Algorithms for the Identification of Stem Changes . . . . .	1102
Mathieu <u>Lafourcade</u> Structured lexical Data: how to make them widely available, useful and reasonably protected? - A practical Example with a trilingual Dictionary . . . . .	1106
Robert W.P. <u>Luk</u> Chinese String Searching using the KMP Algorithm . . . . .	1111
Bjarne <u>Ørsnes</u> , Bradley Music, Bente Maegaard PaTrans - A Patent Translation System . . . . .	1115
Shinsuke <u>Mori</u> , Makoto Nagao Word Extraction from Corpora and Its Part-of-Speech Estimation Using Distributional Analysis . . . . .	1119
Gábor <u>Prószéky</u> Morphological Analyzer as Syntactic Parser . . . . .	1123
Patrick <u>Saint-Dizier</u> Constructing Verb Semantic Classes for French: Methods and Evaluation . . . . .	1127
Hiroyuki <u>Shinnou</u> Redefining similarity in a Thesaurus by using Corpora . . . . .	1131
Masahito Takahashi, Tsuyoshi Shinchu, Kenji Yoshimura, Kosho <u>Shudo</u> Processing Homonyms in the Kana-to-Kanji Conversion . . . . .	1135
Lahcene <u>Si Ameur</u> and Jacques Rouault How the Linguistic Negation can have an Effect in Object-Based Knowledge Representation Model . . . . .	1139
Virach <u>Sornlertlamvanich</u> , Hozumi Tanaka The Automatic Extraction of Open Compounds from Text Corpora . . . . .	1143
Peter <u>Spyns</u> A Tagger/Lemmatiser for Dutch medical Language . . . . .	1147
Marie-Hélène <u>Stefanini</u> and Karine Warren A Distributed Architecture for Text Analysis in French: An Application to Complex Linguistic Phenomena Processing . . . . .	1151
Koichi <u>Takeda</u> Pattern-Based Machine Translation . . . . .	1155

Akira <u>Ushioda</u>	
Hierarchical Clustering of Words . . . . .	1159
G. Jan <u>Wilms</u>	
Using a Hybrid System of Corpus- and Knowledge-Based Techniques to Automate the Induction of a Lexical Sublanguage Grammar . . . . .	1163
Author's Index . . . . .	1167