

Lecture Notes in Artificial Intelligence 3123

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Anja Belz Roger Evans Paul Piwek (Eds.)

# Natural Language Generation

Third International Conference, INLG 2004  
Brockenhurst, UK, July 14-16, 2004  
Proceedings



Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA  
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Anja Belz

Roger Evans

Paul Piwek

University of Brighton, Information Technology Research Institute  
Lewes Road, Brighton BN2 4GJ, UK

E-mail: {Anja.Belz,Roger.Evans,Paul.Piwek}@itri.brighton.ac.uk

Library of Congress Control Number: 2004108214

CR Subject Classification (1998): I.2.7, I.2, F.4.3

ISSN 0302-9743

ISBN 3-540-22340-1 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable to prosecution under the German Copyright Law.

Springer-Verlag is a part of Springer Science+Business Media

[springeronline.com](http://springeronline.com)

© Springer-Verlag Berlin Heidelberg 2004

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Boller Mediendesign

Printed on acid-free paper SPIN: 11017578 06/3142 5 4 3 2 1 0

# Preface

The Third International Conference on Natural Language Generation (INLG 2004) was held from 14th to 16th July 2004 at Careys Manor, Brockenhurst, UK. Supported by the Association for Computational Linguistics Special Interest Group on Generation, the conference continued a twenty-year tradition of biennial international meetings on research into natural language generation. Recent conference venues have included Mitzpe Ramon, Israel (INLG 2000) and New York, USA (INLG 2002). It was our pleasure to invite the thriving and friendly NLG research community to the beautiful New Forest in the south of England for INLG 2004.

INLG is the leading international conference in the field of natural language generation. It provides a forum for the presentation and discussion of original research on all aspects of the generation of language, including psychological modelling of human language production as well as computational approaches to the automatic generation of language. This volume includes a paper by the keynote speaker, Ardi Roelofs of the Max Planck Institute for Psycholinguistics and the F.C. Donders Centre for Cognitive Neuroimaging, 18 regular papers reporting the latest research results and directions, and 4 student papers describing doctoral work in progress. These papers reveal a particular concentration of current research effort on statistical and machine learning methods, on referring expressions, and on variation in surface realisation. The papers were selected from 46 submissions from all over the world (27 from Europe, 13 from North America, 6 from elsewhere), which were subjected to a rigorous double-blind reviewing process undertaken by our hard-working programme committee. In addition, the conference had a poster session giving a snapshot of ongoing research projects in the field (published as ITRI technical report No. ITRI-04-01).

As always, the conference and this volume of proceedings are the result of the combined efforts of many people. First and foremost we would like to thank all the authors who submitted papers, and the Programme Committee who helped us to put together an excellent conference programme. In addition, we thank everyone who helped us put the conference together, especially the ITRI administrative team, staff at Careys Manor Hotel, and members of the SIGGEN and ACL boards, especially Owen Rambow and Kathy McCoy. Finally we thank all the staff and students at ITRI, and in the wider generation community, for their advice, support and participation in INLG 2004.

# Organisation

INLG 2004 was organised by the Information Technology Research Institute, University of Brighton, UK, on behalf of the Association for Computational Linguistics Special Interest Group on Generation (ACL SIGGEN).

## Co-chairs

Anja Belz, ITRI, Brighton  
Roger Evans, ITRI, Brighton  
Paul Piwek, ITRI, Brighton

## Administration

Domino Moore, ITRI, Brighton  
Martyn Haddock, ITRI, Brighton  
Petra Tank, ITRI, Brighton  
Amy Neale, ITRI, Brighton

## ACL SIGGEN Liaison

Owen Rambow, Columbia  
Kathy McCoy, Delaware

## Programme Committee

Anja Belz, ITRI, University of Brighton, UK (Co-chair)  
Roger Evans, ITRI, University of Brighton, UK (Co-chair)  
Paul Piwek, ITRI, University of Brighton, UK (Co-chair)  
Ion Androutsopoulos, Athens University of Economics and Business, Greece  
Srinivas Bangalore, AT&T, USA  
Regina Barzilay, CSAIL, MIT, USA  
John Bateman, Bremen University, Germany  
Tilman Becker, DFKI, Germany  
Sandra Carberry, CIS, University of Delaware, USA  
Alison Cawsey, CEE, Heriot Watt University, UK  
Robert Dale, Macquarie University, Australia  
Kees van Deemter, ITRI, University of Brighton, UK  
Michael Elhadad, Ben-Gurion University of the Negev, Israel  
Barbara di Eugenio, Computer Science, University of Illinois at Chicago, USA  
Nancy Green, University of North Carolina at Greensboro, USA

Catalina Hallett, ITRI, University of Brighton, UK  
Helmut Horacek, Saarland University, Germany  
Eduard Hovy, ISI, University of Southern California, USA  
Aravind Joshi, CIS, University of Pennsylvania, USA  
Min-Yen Kan, National University of Singapore, Singapore  
Emiel Krahmer, Computational Linguistics, Tilburg University, The Netherlands  
Rodger Kibble, Goldsmiths College, University of London, UK  
Inderjeet Mani, Georgetown University, USA  
Daniel Marcu, ISI, University of Southern California, USA  
Colin Matheson, Informatics, University of Edinburgh, UK  
Kathleen McCoy, CIS, University of Delaware, USA  
Kathleen McKeown, CS, Columbia University, USA  
Chris Mellish, Computer Science, University of Aberdeen, UK  
Detmar Meurers, Linguistics, Ohio State University, USA  
Johanna Moore, Informatics, University of Edinburgh, UK  
Mick O'Donnell, Wagsoft Systems, UK  
Jon Oberlander, Informatics, University of Edinburgh, UK  
Shimei Pan, IBM T.J. Watson Research Center, USA  
Richard Power, ITRI, University of Brighton, UK  
Ehud Reiter, Aberdeen University, UK  
Matthew Stone, CS, Rutgers University, USA  
Sebastian Varges, ITRI, University of Brighton, UK  
Nigel Ward, CS, University of Texas at El Paso, USA  
Ingrid Zukerman, CSSE, Monash University, Australia

# Table of Contents

## Keynote Paper

The Seduced Speaker: Modeling of Cognitive Control . . . . .	1
<i>Ardi Roelofs</i>	

## Regular Papers

Generating Intensional Answers in Intelligent Question Answering Systems . . . . .	11
<i>Farah Benamara</i>	
Saliency-Driven Text Planning . . . . .	21
<i>Christian Chiarcos, Manfred Stede</i>	
Finetuning NLG Through Experiments with Human Subjects: The Case of Vague Descriptions . . . . .	31
<i>Kees van Deemter</i>	
Indirect Supervised Learning of Content Selection Logic . . . . .	41
<i>Pablo A. Duboue</i>	
Generating Referring Expressions Using Perceptual Groups . . . . .	51
<i>Kotaro Funakoshi, Satoru Watanabe, Naoko Kuriyama, Takenobu Tokunaga</i>	
The Use of a Structural N-gram Language Model in Generation-Heavy Hybrid Machine Translation . . . . .	61
<i>Nizar Habash</i>	
On Referring to Sets of Objects Naturally . . . . .	70
<i>Helmut Horacek</i>	
An ATMS Approach to Systemic Sentence Generation . . . . .	80
<i>Hasan Kamal, Chris Mellish</i>	
A Corpus-Based Methodology for Evaluating Metrics of Coherence for Text Structuring . . . . .	90
<i>Nikiforos Karamanis, Chris Mellish, Jon Oberlander, Massimo Poesio</i>	
Classification-Based Generation Using TAG . . . . .	100
<i>Tomasz Marciniak, Michael Strube</i>	
Resolving Structural Ambiguity in Generated Speech . . . . .	110
<i>Chris Mellish</i>	

A Framework for Stylistically Controlled Generation . . . . .	120
<i>Daniel S. Paiva, Roger Evans</i>	
SEGUE: A Hybrid Case-Based Surface Natural Language Generator . . . .	130
<i>Shimei Pan, James Shaw</i>	
Modelling Politeness in Natural Language Generation . . . . .	141
<i>Kaška Porayska-Pomsta, Chris Mellish</i>	
Context-Based Incremental Generation for Dialogue . . . . .	151
<i>Matthew Purver, Ruth Kempson</i>	
Contextual Influences on Near-Synonym Choice . . . . .	161
<i>Ehud Reiter, Somayajulu Sripada</i>	
Overgenerating Referring Expressions Involving Relations and Booleans . .	171
<i>Sebastian Varges</i>	
Reining in CCG Chart Realization . . . . .	182
<i>Michael White</i>	
<b>Student Papers</b>	
Categorization of Narrative Semantics for Use in Generative Multidocument Summarization . . . . .	192
<i>David K. Elson</i>	
Corpus-Based Planning of Deictic Gestures in COMIC . . . . .	198
<i>Mary Ellen Foster</i>	
Hybrid NLG in a Generic Dialog System . . . . .	205
<i>Martin Klärner</i>	
Two Kinds of Alternative Sets and a Marking Principle – When to Say <i>Also</i> . . . . .	212
<i>Kristina Striegnitz</i>	
<b>Author Index</b> . . . . .	219