

2006



COLING • ACL

COLING • ACL 2006

INLG-06

Fourth International Natural
Language Generation Conference

Proceedings of the Conference

Programme Co-Chairs:
Nathalie Colineau, Cécile Paris,
Stephen Wan and Robert Dale

Special Session Co-Chairs:
Anja Belz and Robert Dale

15-16 July 2006
Sydney, Australia

Production and Manufacturing by
BPA Digital
11 Evans St
Burwood VIC 3125
AUSTRALIA

©2006 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 1-932432-72-8

Table of Contents

Preface	v
Organizers	vii
Conference Programme	ix
Invited Talk	
<i>Lessons Learned from Large Scale Evaluation of Systems that Produce Text: Nightmares and Pleasant Surprises</i>	
Kathleen R. McKeown	3
Main Programme	
<i>A Generation-Oriented Workbench for Performance Grammar: Capturing Linear Order Variability in German and Dutch</i>	
Karin Harbusch, Gerard Kempen, Camiel van Breugel and Ulrich Koch	9
<i>CCG Chart Realization from Disjunctive Inputs</i>	
Michael White	12
<i>Overgeneration and Ranking for Spoken Dialogue Systems</i>	
Sebastian Varges	20
<i>Individuality and Alignment in Generated Dialogues</i>	
Amy Isard, Carsten Brockmann and Jon Oberlander	25
<i>Using Distributional Similarity to Identify Individual Verb Choice</i>	
Jing Lin	33
<i>Adjective-to-Verb Paraphrasing in Japanese Based on Lexical Constraints of Verbs</i>	
Atsushi Fujita, Naruaki Masuno, Satoshi Sato and Takehito Utsuro	41
<i>Generating References to Parts of Recursively Structured Objects</i>	
Helmut Horacek	47
<i>Overspecified Reference in Hierarchical Domains: Measuring the Benefits for Readers</i>	
Ivandr� Paraboni, Judith Masthoff and Kees van Deemter	55
<i>Algorithms for Generating Referring Expressions: Do They Do What People Do?</i>	
Jette Viethen and Robert Dale	63
<i>Group-Based Generation of Referring Expressions</i>	
Kotaro Funakoshi, Satoru Watanabe and Takenobu Tokunaga	73
<i>Noun Phrase Generation for Situated Dialogs</i>	
Laura Stoia, Darla Magdalene Shockley, Donna K. Byron and Eric Fosler-Lussier	81

<i>The Clarity-Brevity Trade-off in Generating Referring Expressions</i> Imtiaz Hussain Khan, Graeme Ritchie and Kees van Deemter	89
<i>Generic Querying of Relational Databases using Natural Language Generation Techniques</i> Catalina Hallett	95
<i>Generating Intelligent Numerical Answers in a Question-Answering System</i> Véronique Moriceau	103
<i>Generating Multiple-Choice Test Items from Medical Text: A Pilot Study</i> Nikiforos Karamanis, Le An Ha and Ruslan Mitkov	111
<i>Generation of Biomedical Arguments for Lay Readers</i> Nancy Green	114
Special Session on Sharing Data and Comparative Evaluation	
<i>Introduction to the INLG'06 Special Session on Sharing Data and Comparative Evaluation</i> Anja Belz and Robert Dale	125
<i>Evaluations of NLG Systems: Common Corpus and Tasks or Common Dimensions and Metrics?</i> Cécile Paris, Nathalie Colineau and Ross Wilkinson	127
<i>Building a Semantically Transparent Corpus for the Generation of Referring Expressions.</i> Kees van Deemter, Ielka van der Sluis and Albert Gatt	130
<i>Shared-Task Evaluations in HLT: Lessons for NLG</i> Anja Belz and Adam Kilgarriff	133
<i>GENEVAL: A Proposal for Shared-task Evaluation in NLG</i> Ehud Reiter and Anja Belz	136
Author Index	139

Preface

We are pleased to introduce the technical programme of the Fourth International Natural Language Generation Conference (INLG), the Biennial Meeting of SIGGEN, the ACL Special Interest Group in Natural Language Generation. INLG is the leading international conference on research into natural language generation. It has been held at Brockenhurst (UK) in 2004, in Harriman (New York, USA) in 2002, and in Mitzpe Ramon (Israel) in 2000. Prior to 2000, the INLG meetings were International Workshops, running every other year since 1980. The INLG conference provides a forum for the discussion, dissemination and archiving of research topics and results in the field of text generation. This year, INLG is being held as a COLING/ACL 2006 workshop. It takes place on the weekend prior to the main COLING/ACL 2006 conference, on July 15-16th 2006, in Sydney, Australia.

The INLG programme consists of substantial, original, and previously unpublished results on all topics related to natural language generation. This year, as in previous years, each submission was reviewed as a full paper by at least three members of an international programme committee of leading researchers in the field, listed on the next page. We received 38 submissions (both long and short papers) from all over the world, from which we accepted 11 long papers (including two student papers) and five short papers. We would like to thank all who submitted papers and our programme committee for their hard work.

This year, the programme centers around a variety of research issues around the realisation component of a natural language system, including the use of statistical techniques. In particular, there are a substantial number of contributions on the generation of referring expressions. The programme also includes an invited talk by Professor Kathleen McKeown (Columbia University, New York, USA) entitled "Lessons Learned from Large Scale Evaluation from Systems that Produce Text: Nightmares and Pleasant Surprises" and a special session on "Sharing Data and Comparative Evaluation", organised by Dr Anja Belz (University of Brighton, UK) and Professor Robert Dale (Macquarie University, Australia).

We would like to thank the whole COLING/ACL 2006 committee, in particular: Suzanne Stevenson, the COLING/ACL 2006 Workshop Chair; Olivia Kwong, the Publications Chair; Priscilla Rasmussen, the ACL Business manager; and the local organising committee, including Judy Potter of Well Done Events, for their help and advice in organising INLG 2006. We are also grateful to the ACL Executive Committee and the SIGGEN board for their support and advice.

We welcome you to the conference and hope you will find it to be a productive and enjoyable experience.

Nathalie Colineau
Cécile Paris
Stephen Wan
Robert Dale
INLG 2006 Programme Co-chairs

Organizers

Programme Co-Chairs:

Nathalie Colineau, CSIRO - ICT Centre, Australia
Cécile Paris, CSIRO - ICT Centre, Australia
Stephen Wan, CSIRO - ICT Centre and Macquarie University, Australia
Robert Dale, Macquarie University, Australia

Special Session Co-Chairs:

Anja Belz, University of Brighton, UK
Robert Dale, Macquarie University, Australia

Programme Committee:

Regina Barzilay, Columbia University, USA
Kalina Bontcheva, University of Sheffield, UK
Joyce Y. Chai, Michigan State University, USA
Nathalie Colineau, CSIRO, Australia
Robert Dale, Macquarie University, Australia
Laurence Danlos, University of Paris 7, France
Noemie Elhadad, City College of New York, USA
Sabine Geldof, Namahn, Belgium
Graeme Hirst, University of Toronto, Canada
Kentarō Inui, Nara Institute of Science and Technology, Japan
Elena Not, IRST, Italy
Cecile Paris, CSIRO, Australia
Ehud Reiter, University of Aberdeen, UK
Norbert Reithinger, DFKI, Germany
Rolf Schwitter, Macquarie University, Australia
Donia Scott, Open University, UK
Mariet Theune, University of Twente, Netherlands
Keith Vander Linden, Calvin College, USA
Ingrid Zukerman, Monash University, Australia

Student Session Program Committee:

Bernd Bohnet, University of Stuttgart, Germany
Matt Huenerfauth, University of Pennsylvania, Philadelphia, USA
Eric Kow, Loria, France.
Tomasz Marciniak, Language Computer Corporation, Richardson, Texas.
Ani Nenkova, Columbia University, USA
David Reitter, University of Edinburgh, Scotland, UK
Stephen Wan, University of Macquarie, Australia

Invited Speaker:

Kathleen R. McKeown, Columbia University, USA

Conference Programme

Saturday, 15 July 2006

9:15–9:30 Opening Remarks

Surface Realisation Session 1:

9:30–9:45 *A Generation-Oriented Workbench for Performance Grammar: Capturing Linear Order Variability in German and Dutch*

Karin Harbusch, Gerard Kempen, Camiel van Breugel and Ulrich Koch

9:45–10:15 *CCG Chart Realization from Disjunctive Inputs*

Michael White

10:15–10:30 *Overgeneration and Ranking for Spoken Dialogue Systems*

Sebastian Varges

10:30–11:00 Morning Tea

Surface Realisation Session 2:

11:00–11:30 *Individuality and Alignment in Generated Dialogues*

Amy Isard, Carsten Brockmann and Jon Oberlander

11:30–12:00 *Using Distributional Similarity to Identify Individual Verb Choice*

Jing Lin

12:00–12:15 *Adjective-to-Verb Paraphrasing in Japanese Based on Lexical Constraints of Verbs*

Atsushi Fujita, Naruaki Masuno, Satoshi Sato and Takehito Utsuro

12:15–2:00 Lunch

Referring Expressions Session 1:

2:00–2:30 *Generating References to Parts of Recursively Structured Objects*

Helmut Horacek

2:30–3:00 *Overspecified Reference in Hierarchical Domains: Measuring the Benefits for Readers*

Ivandr  Paraboni, Judith Masthoff and Kees van Deemter

3:00–3:30 *Algorithms for Generating Referring Expressions: Do They Do What People Do?*

Jette Viethen and Robert Dale

3:30–4:00 Afternoon Tea

Saturday, 15 July 2006 (continued)

Special Session on Sharing Data and Comparative Evaluation:

- 4:00–4:15 Special Session Opening Remarks
- 4:15–4:30 *Evaluations of NLG Systems: Common Corpus and Tasks or Common Dimensions and Metrics?*
Cécile Paris, Nathalie Colineau and Ross Wilkinson
- 4:30–4:45 *Building a semantically transparent corpus for the generation of referring expressions.*
Kees van Deemter, Ielka van der Sluis and Albert Gatt
- 4:45–5:00 *Shared-task Evaluations in HLT: Lessons for NLG*
Anja Belz and Adam Kilgarriff
- 5:00–5:15 *GENEVAL: A Proposal for Shared-task Evaluation in NLG*
Ehud Reiter and Anja Belz
- 5:15–6:15 Open Mic Session

Participants:

Kees van Deemter, Albert Gatt, Ielka van der Sluis
Michael White
David Reitter and Charles Callaway
Helmut Horacek
Sebastian Varges
Donia Scott and Johanna Moore
.. and other contributions from the floor.

Sunday, 16 July 2006

9:00–10:30 Invited Talk

*Lessons Learned from Large Scale Evaluation of Systems that Produce Text:
Nightmares and Pleasant Surprises*

Kathleen R. McKeown

10:30–11:00 Morning Tea

Referring Expressions Session 2:

11:00–11:30 *Group-Based Generation of Referring Expressions*

Kotaro Funakoshi, Satoru Watanabe and Takenobu Tokunaga

11:30–12:00 *Noun Phrase Generation for Situated Dialogs*

Laura Stoia, Darla Magdalene Shockley, Donna K. Byron and Eric Fosler-Lussier

12:00–12:15 *The Clarity-Brevity Trade-off in Generating Referring Expressions*

Imtiaz Hussain Khan, Graeme Ritchie and Kees van Deemter

12:15–1:45 Lunch

Session on Querying, Answering and Arguing:

1:45–2:15 *Generic Querying of Relational Databases using Natural Language Generation Techniques*

Catalina Hallett

2:15–2:45 *Generating Intelligent Numerical Answers in a Question-Answering System*

Véronique Moriceau

2:45–3:00 *Generating Multiple-Choice Test Items from Medical Text: A Pilot Study*

Nikiforos Karamanis, Le An Ha and Ruslan Mitkov

3:00–3:30 *Generation of Biomedical Arguments for Lay Readers*

Nancy Green

3:30–4:00 Afternoon Tea

