

INLG 2016

**The 9th International
Natural Language Generation conference**

Proceedings of the Conference

September 5-8, 2016
Edinburgh, UK

©2016 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

Introduction

Welcome to the 9th International Natural Language Generation Conference (INLG 2016)! INLG is the biennial meeting of the ACL Special Interest Group on Natural Language Generation (SIGGEN). The INLG conference provides the premier forum for the discussion, dissemination, and archiving of research and results in the field of Natural Language Generation (NLG). Previous INLG conferences have been held in Ireland, the USA, Australia, the UK, and Israel. Prior to 2000, INLG meetings were held as international workshops with a history stretching back to 1983. In 2016, INLG is organized in Edinburgh, UK.

The INLG 2016 program includes presentations of substantial, original, and previously unpublished results on all topics related to natural language generation. This year, INLG received 65 submissions (29 full papers, 30 short papers and 6 demo proposals) from 12 different countries. 16 submissions were accepted as full papers (10 presented orally, 6 as posters), 20 as short papers (4 presented orally, and 16 as posters), and 3 demos. In addition, INLG 2016 includes an invited talk by Yejin Choi (University of Washington) and Vera Demberg (Saarland University).

The organizing committee would like to offer their thanks to our invited speakers for agreeing to join us and to the authors of all submitted papers. We have also received sponsorship from Arria and WebNLG, for which we are extremely grateful. Finally, we would like to welcome you to Edinburgh and hope that you have an enjoyable and inspiring visit!

Amy Isard, Verena Rieser and Dimitra Gkatzia
INLG 2016 Co-Chairs

Organizers:

Amy Isard
Verena Rieser
Dimitra Gkatzia
Jekaterina Novikova
Simon Keizer

Program Committee:

Anja Belz
Aoife Cahill
Nina Dethlefs
Katja Filippova
Claire Gardent
Albert Gatt
Pablo Gervás
Alistair Knott
Ioannis Konstas
Emiel Kraemer
Guy Lapalme
David McDonald
Cecile Paris
Paul Piwek
Ehud Reiter
Kristina Striegnitz
Mariët Theune
Ielka van der Sluis
Keith Vander Linden
Leo Wanner
Michael White
Sandra Williams
Helen Hastie
Barbara Di Eugenio
Vera Demberg
Laura Perez-Beltrachini
Saad Mahamood
Andreas Vlachos
Kees van Deemter
David Howcroft
Yejin Choi
Kathy McKeown
Diana Bental

François Portet
Simon Mille
Cyril Labbe
John Kelleher
Sander Wubben
Somayajulu Sripada
Nikolaos Engonopoulos
Amanda Stent
David Schlangen
Sina Zarriß

Invited Speakers:

Yejin Choi (University of Washington)
Vera Demberg (Saarland University)

Table of Contents

<i>Summarising News Stories for Children</i> Iain Macdonald and Advait Siddharthan	1
<i>Discourse-Driven Narrative Generation With Bipartite Planning</i> David Winer and R. Michael Young	11
<i>Generating English from Abstract Meaning Representations</i> Nima Pourdamghani, Kevin Knight and Ulf Hermjakob	21
<i>Generating summaries of hospitalizations: A new metric to assess the complexity of medical terms and their definitions</i> Sabita Acharya, Barbara Di Eugenio, Andrew D Boyd, Karen Dunn Lopez, Richard Cameron and Gail M Keenan	26
<i>Designing Algorithms for Referring with Proper Names</i> Kees van Deemter	31
<i>When to Plummet and When to Soar: Corpus Based Verb Selection for Natural Language Generation</i> Charese Smiley, Vassilis Plachouras, Frank Schilder, Hiroko Bretz, Jochen Leidner and Dezhao Song	36
<i>Sketch-to-Text Generation: Toward Contextual, Creative, and Coherent Composition</i> Yejin Choi	40
<i>Abstractive Compression of Captions with Attentive Recurrent Neural Networks</i> Sander Wubben, Emiel Kraemer, Antal van den Bosch and Suzan Verberne	41
<i>Infusing NLU into Automatic Question Generation</i> Karen Mazidi and Paul Tarau	51
<i>Automatic label generation for news comment clusters</i> Ahmet Aker, Monica Paramita, Emina Kurtic, Adam Funk, Emma Barker, Mark Hepple and Rob Gaizauskas	61
<i>Improving Fluency in Narrative Text Generation With Grammatical Transformations and Probabilistic Parsing</i> Emily Ahn, Fabrizio Morbini and Andrew Gordon	70
<i>The Multilingual Affective Soccer Corpus (MASC): Compiling a biased parallel corpus on soccer reportage in English, German and Dutch</i> Nadine Braun, Martijn Goudbeek and Emiel Kraemer	74
<i>Challenges of Argument Mining: Generating an Argument Synthesis based on the Qualia Structure</i> Patrick Saint-Dizier	79

<i>Tense and Aspect in Runyankore Using a Context-Free Grammar</i> Joan Byamugisha, C. Maria Keet and Brian DeRenzi	84
<i>Task demands and individual variation in referring expressions</i> Adriana Baltaretu and Thiago Castro Ferreira	89
<i>Category-Driven Content Selection</i> Rania Mohammed, Laura Perez-Beltrachini and Claire Gardent	94
<i>Evaluative Pattern Extraction for Automated Text Generation</i> Chia-Chen Lee and Shu-Kai HSIEH	99
<i>Statistics-Based Lexical Choice for NLG from Quantitative Information</i> Xiao Li, Kees van Deemter and Chenghua Lin	104
<i>Incremental Generation of Visually Grounded Language in Situated Dialogue (demonstration system)</i> Yanchao Yu, Arash Eshghi and Oliver Lemon	109
<i>Unsupervised Sentence Simplification Using Deep Semantics</i> Shashi Narayan and Claire Gardent	111
<i>Enabling text readability awareness during the micro planning phase of NLG applications</i> Priscilla Moraes, Kathleen Mccoy and Sandra Carberry	121
<i>How can we adapt generation to the user’s cognitive load?</i> Vera Demberg	132
<i>Selecting Domain-Specific Concepts for Question Generation With Lightly-Supervised Methods</i> Yiping Jin and Phu Le	133
<i>Statistical Natural Language Generation from Tabular Non-textual Data</i> Joy Mahapatra, Sudip Kumar Naskar and Sivaji Bandyopadhyay	143
<i>Paraphrase Generation from Latent-Variable PCFGs for Semantic Parsing</i> Shashi Narayan, Siva Reddy and Shay B. Cohen	153
<i>The WebNLG Challenge: Generating Text from DBPedia Data</i> Emilie Colin, Claire Gardent, Yassine Mrabet, Shashi Narayan and Laura Perez-Beltrachini ..	163
<i>The aNALoGuE Challenge: Non Aligned Language GEneration</i> Jekaterina Novikova and Verena Rieser	168
<i>A Challenge Proposal for Narrative Generation Using CNLs</i> Eugenio Concepción, Gonzalo Méndez, Pablo Gervás and Carlos León	171
<i>On the verbalization patterns of part-whole relations in isiZulu</i> C. Maria Keet and Langa Khumalo	174
<i>SimpleNLG-IT: adapting SimpleNLG to Italian</i> Alessandro Mazzei, Cristina Battaglini and Cristina Bosco	184

<i>Don't Mention the Shoe! A Learning to Rank Approach to Content Selection for Image Description Generation</i>	
Josiah Wang and Robert Gaizauskas	193
<i>Good Automatic Authentication Question Generation</i>	
Simon Woo, Zuyao Li and Jelena Mirkovic	203
<i>Automatic Generation of Student Report Cards</i>	
Amy Isard and Jeremy Knox	207
<i>Collecting Reliable Human Judgements on Machine-Generated Language: The Case of the QG-STEC Data</i>	
Keith Godwin and Paul Piwek	212
<i>Ranking Automatically Generated Questions Using Common Human Queries</i>	
Yllias Chali and Sina Golestanirad	217
<i>Towards proper name generation: a corpus analysis</i>	
Thiago Castro Ferreira, Sander Wubben and Emiel Krahmer	222
<i>An Analysis of the Ability of Statistical Language Models to Capture the Structural Properties of Language</i>	
Aneiss Ghodsi and John DeNero	227
<i>Enhancing PTB Universal Dependencies for Grammar-Based Surface Realization</i>	
David L. King and Michael White	232
<i>Effect of Data Annotation, Feature Selection and Model Choice on Spatial Description Generation in French</i>	
Anja Belz, Adrian Muscat, Brandon Birmingham, Jessie Levacher, Julie Pain and Adam Quinquenel	237
<i>QGASP: a Framework for Question Generation Based on Different Levels of Linguistic Information</i>	
Hugo Patinho Rodrigues, Luisa Coheur and Eric Nyberg	242
<i>Automatic Reports from Spreadsheets: Data Analysis for the Rest of Us</i>	
Pablo Duboue	244
<i>Towards Generating Colour Terms for Referents in Photographs: Prefer the Expected or the Unexpected?</i>	
Sina Zarriß and David Schlangen	246
<i>Absolute and Relative Properties in Geographic Referring Expressions</i>	
Rodrigo de Oliveira, Somayajulu Sripada and Ehud Reiter	256
<i>Crowd-sourcing NLG Data: Pictures Elicit Better Data.</i>	
Jekaterina Novikova, Oliver Lemon and Verena Rieser	265

Conference Program

Wednesday, September 7, 2016

8:00–9:15 **Registration**

9:15–9:30 **Welcome / Opening remarks**

9:30–10:30 **Oral Session: Summarisation / Narrative**

09:30–10:00 *Summarising News Stories for Children*
Iain Macdonald and Advaith Siddharthan

10:00–10:30 *Discourse-Driven Narrative Generation With Bipartite Planning*
David Winer and R. Michael Young

10:30–11:00 **Coffee break**

11:00–12:20 **Oral Session: short papers**

11:00–11:20 *Generating English from Abstract Meaning Representations*
Nima Pourdamghani, Kevin Knight and Ulf Hermjakob

11:20–11:40 *Generating summaries of hospitalizations: A new metric to assess the complexity of medical terms and their definitions*
Sabita Acharya, Barbara Di Eugenio, Andrew D Boyd, Karen Dunn Lopez, Richard Cameron and Gail M Keenan

11:40–12:00 *Designing Algorithms for Referring with Proper Names*
Kees van Deemter

12:00–12:20 *When to Plummet and When to Soar: Corpus Based Verb Selection for Natural Language Generation*
Charese Smiley, Vassilis Plachouras, Frank Schilder, Hiroko Bretz, Jochen Leidner and Dezhao Song

Wednesday, September 7, 2016 (continued)

12:20–13:30 Lunch

13:30–14:30 *Invited Talk*
Yejin Choi, session chair: Amy Isard

–13:30 *Sketch-to-Text Generation: Toward Contextual, Creative, and Coherent Composition*
Yejin Choi

14:30–16:00 Posters and Demos with coffee

15:00–15:30 *Abstractive Compression of Captions with Attentive Recurrent Neural Networks*
Sander Wubben, Emiel Krahmer, Antal van den Bosch and Suzan Verberne

15:00–15:30 *Infusing NLU into Automatic Question Generation*
Karen Mazidi and Paul Tarau

15:00–15:30 *Automatic label generation for news comment clusters*
Ahmet Aker, Monica Paramita, Emina Kurtic, Adam Funk, Emma Barker, Mark Hepple and Rob Gaizauskas

15:00–15:30 *Improving Fluency in Narrative Text Generation With Grammatical Transformations and Probabilistic Parsing*
Emily Ahn, Fabrizio Morbini and Andrew Gordon

15:00–15:30 *The Multilingual Affective Soccer Corpus (MASC): Compiling a biased parallel corpus on soccer reportage in English, German and Dutch*
Nadine Braun, Martijn Goudbeek and Emiel Krahmer

15:00–15:30 *Challenges of Argument Mining: Generating an Argument Synthesis based on the Qualia Structure*
Patrick Saint-Dizier

15:00–15:30 *Tense and Aspect in Runyankore Using a Context-Free Grammar*
Joan Byamugisha, C. Maria Keet and Brian DeRenzi

15:00–15:30 *Task demands and individual variation in referring expressions*
Adriana Baltaretu and Thiago Castro Ferreira

15:00–15:30 *Category-Driven Content Selection*
Rania Mohammed, Laura Perez-Beltrachini and Claire Gardent

Wednesday, September 7, 2016 (continued)

- 15:00–15:30 *Evaluative Pattern Extraction for Automated Text Generation*
Chia-Chen Lee and Shu-Kai HSIEH
- 15:00–15:30 *Statistics-Based Lexical Choice for NLG from Quantitative Information*
Xiao Li, Kees van Deemter and Chenghua Lin
- 15:00–15:30 *Incremental Generation of Visually Grounded Language in Situated Dialogue (demonstration system)*
Yanchao Yu, Arash Eshghi and Oliver Lemon

16:00–17:00 Oral Session: Simplification and Readability

- 16:00–16:30 *Unsupervised Sentence Simplification Using Deep Semantics*
Shashi Narayan and Claire Gardent
- 16:30–17:00 *Enabling text readability awareness during the micro planning phase of NLG applications*
Priscilla Moraes, Kathleen Mccoy and Sandra Carberry

19:00 Drinks, dinner, ceilidh at Ghillie Dhu - Announcement of best paper

Thursday, September 8, 2016

- 09:00–10:00 *Invited Talk*
Vera Demberg, session chair: Verena Rieser
- 9:00–10:00 *How can we adapt generation to the user's cognitive load?*
Vera Demberg

Thursday, September 8, 2016 (continued)

10:00–10:30 Oral Session, Statistical NLG 1

10:00–10:30 *Selecting Domain-Specific Concepts for Question Generation With Lightly-Supervised Methods*
Yiping Jin and Phu Le

10:30–11:00 Coffee break

11:00–12:00 Oral Session, Statistical NLG 2

11:00–11:30 *Statistical Natural Language Generation from Tabular Non-textual Data*
Joy Mahapatra, Sudip Kumar Naskar and Sivaji Bandyopadhyay

11:30–12:00 *Paraphrase Generation from Latent-Variable PCFGs for Semantic Parsing*
Shashi Narayan, Siva Reddy and Shay B. Cohen

12:00–12:45 GenChal

12:00–12:15 *The WebNLG Challenge: Generating Text from DBPedia Data*
Emilie Colin, Claire Gardent, Yassine Mrabet, Shashi Narayan and Laura Perez-Beltrachini

12:15–12:30 *The aNALoGuE Challenge: Non Aligned Language Generation*
Jekaterina Novikova and Verena Rieser

12:30–12:45 *A Challenge Proposal for Narrative Generation Using CNLs*
Eugenio Concepción, Gonzalo Méndez, Pablo Gervás and Carlos León

Thursday, September 8, 2016 (continued)

12:45–13:45 Lunch

13:45–15:15 Posters and demos with coffee

14:45–15:15 *On the verbalization patterns of part-whole relations in isiZulu*
C. Maria Keet and Langa Khumalo

14:45–15:15 *SimpleNLG-IT: adapting SimpleNLG to Italian*
Alessandro Mazzei, Cristina Battaglini and Cristina Bosco

14:45–15:15 *Don't Mention the Shoe! A Learning to Rank Approach to Content Selection for Image Description Generation*
Josiah Wang and Robert Gaizauskas

14:45–15:15 *Good Automatic Authentication Question Generation*
Simon Woo, Zuyao Li and Jelena Mirkovic

14:45–15:15 *Automatic Generation of Student Report Cards*
Amy Isard and Jeremy Knox

14:45–15:15 *Collecting Reliable Human Judgements on Machine-Generated Language: The Case of the QG-STEAC Data*
Keith Godwin and Paul Piwek

14:45–15:15 *Ranking Automatically Generated Questions Using Common Human Queries*
Yllias Chali and Sina Golestanirad

14:45–15:15 *Towards proper name generation: a corpus analysis*
Thiago Castro Ferreira, Sander Wubben and Emiel Krahmer

14:45–15:15 *An Analysis of the Ability of Statistical Language Models to Capture the Structural Properties of Language*
Aneiss Ghodsi and John DeNero

14:45–15:15 *Enhancing PTB Universal Dependencies for Grammar-Based Surface Realization*
David L. King and Michael White

14:45–15:15 *Effect of Data Annotation, Feature Selection and Model Choice on Spatial Description Generation in French*
Anja Belz, Adrian Muscat, Brandon Birmingham, Jessie Levacher, Julie Pain and Adam Quinquenel

Thursday, September 8, 2016 (continued)

14:45–15:15 *QGASP: a Framework for Question Generation Based on Different Levels of Linguistic Information*

Hugo Patinho Rodrigues, Luisa Coheur and Eric Nyberg

14:45–15:15 *Automatic Reports from Spreadsheets: Data Analysis for the Rest of Us*

Pablo Duboue

15:15–16:45 Oral Session, REG and Vision

16:15–16:45 *Towards Generating Colour Terms for Referents in Photographs: Prefer the Expected or the Unexpected?*

Sina Zarriß and David Schlangen

15:15–15:45 *Absolute and Relative Properties in Geographic Referring Expressions*

Rodrigo de Oliveira, Somayajulu Sripada and Ehud Reiter

15:45–16:15 *Crowd-sourcing NLG Data: Pictures Elicit Better Data.*

Jekaterina Novikova, Oliver Lemon and Verena Rieser

Closing Remarks