

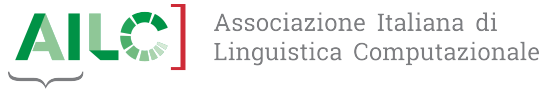
IWPT 2017

**15th International Conference on
Parsing Technologies**

Proceedings of the Conference

September 20–22, 2017
Pisa, Italy

Sponsored by:



Associazione Italiana di
Linguistica Computazionale

©2017 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 978-1-945626-73-9

Preface

Welcome to the 15th International Conference on Parsing Technologies (IWPT 2017) in Pisa, Italy. IWPT 2017 continues the tradition of biennial conferences organized by SIGPARSE, the Special Interest Group on Natural Language Parsing of the Association for Computational Linguistics (ACL), serving as the primary specialized forum for research on natural language parsing.

This year, for the first time, IWPT is co-located with the International Conference on Dependency Linguistics (DepLing), and will feature joint sessions meant to foster communication and collaboration between the two closely related communities. As part of their joint programs, DepLing and IWPT will host the First Shared Task on Extrinsic Parser Evaluation (EPE 2017), which has separate organizers, program committee and proceedings.

IWPT 2016 received 26 valid submissions (11 long and 15 short papers). Each submission was reviewed by three members of the program committee, and finally 5 long and 13 short papers were accepted for presentation. The conference will feature two invited speakers, Vera Demberg (Saarland University) and David Hall (Semantic Machines), and a conference dinner on September 21.

We are indebted to a number of people whose work made the conference possible. First and foremost, we thank the members of the local organization committee, whose outstanding contributions involved every aspect of hosting both DepLing and IWPT: Giuseppe Attardi (University of Pisa), Felice Dell'Orletta (ILC-CNR, Pisa), Alessandro Lenci (University of Pisa), Simonetta Montemagni (ILC-CNR, Pisa), and Maria Simi (University of Pisa). We thank the members of the program committee for their in-depth and constructive reviews, and the program chairs of DepLing (Simonetta Montemagni and Joakim Nivre) and EPE (Stephan Oepen) for their help and support with various aspects of the conference, including design of the joint portion of our programs. Finally, we thank the Associazione Italiana di Linguistica Computazionale (AILC) for its generous financial support.

Enjoy the conference!

Kenji Sagae (General Chair) and Yusuke Miyao (Program Chair)

Organizers:

Kenji Sagae, UC Davis, United States (General Chair)
Yusuke Miyao, National Institute of Informatics, Japan (Programme Chair)
Giuseppe Attardi, University of Pisa, Italy (Local Organizer)
Felice Dell'Orletta, ILC-CNR, Pisa, Italy (Local Organizer)
Alessandro Lenci, University of Pisa, Italy (Local Organizer)
Simonetta Montemagni, ILC-CNR, Pisa, Italy (Local Organizer)
Maria Simi, University of Pisa, Italy (Local Organizer)

Programme Committee:

Željko Agić, IT University of Copenhagen
Miguel Ballesteros, IBM Research
Anders Björkelund, IMS, Stuttgart
Philippe Blache, LPL CNRS
Harry Bunt, Tilburg University
Xavier Carreras, Naver Labs
John Carroll, University of Sussex
Wanxiang Che, Harbin Institute of Technology
David Chiang, University of Notre Dame
Stephen Clark, University of Cambridge
Berthold Crysmann, CNRS, Laboratoire de linguistique formelle
Michael Elhadad, Ben Gurion University
Federico Fancellu, University of Edinburgh
Carlos Gómez-Rodríguez, Universidade da Coruña
Sadao Kurohashi, Kyoto University
Yuji Matsumoto, Nara Institute of Science and Technology
Yusuke Miyao, National Institute of Informatics
Mark-Jan Nederhof, University of St Andrews
Giorgio Satta, University of Padova
Natalie Schluter, IT University of Copenhagen
Djamé Seddah, Université Paris Sorbonne (Paris IV)
Hiroyuki Shindo, Nara Institute of Science and Technology
Weiwei Sun, Peking University
Christoph Teichmann, Saarland University
Ivan Titov, University of Edinburgh / University of Amsterdam
Daniel Zeman, Charles University
Yi Zhang, Nuance Communications GmbH
Yue Zhang, Singapore University of Technology and Design
Hai Zhao, Shanghai Jiao Tong University

Keynote 1: Syntactic processing in humans: time course, shallow processing and processing failure

Vera Demberg
Saarland University, Germany

Abstract

Human syntactic processing is generally remarkably robust and accurate. In this talk, I will go through some recent psycholinguistic research on sentence processing which can give us a glimpse into how human parsing works. This talk will focus on some crucial properties (such as incrementality and prediction), but will also highlight some informative cases of where humans struggle to analyse a sentence correctly. I will also briefly describe current psycholinguistic frameworks for modelling human processing, which try to account for these cases.

Biography

Vera Demberg is professor for Computer Science and Computational Linguistics at Saarland University, Saarbrücken, Germany. From 2010 till 2016, she held a position as an independent research group leader at the Cluster of Excellence "Multimodal Computing and Interaction", Saarland University. She received her PhD in 2010 from the University of Edinburgh, for which she was awarded the Glushko Dissertation Prize by the Cognitive Science Society and the runner-up prize of the CPHC / BCS Distinguished Dissertations Competition. Her research interests include psycholinguistic experimental research and computational modelling on human sentence processing at the levels of syntax, thematic role assignment, event cognition and coherence relations in discourse.

Keynote 2: What good is a grammar anyway?

David Hall

Semantic Machines, United States

Abstract

Until recently, a large fraction of constituency parsing research consisted of finding clever ways of “augmenting” a base treebank grammar with extra information to work around the limitations of dynamic programming-based parsing algorithms. Nowadays, the art of grammar engineering for statistical parsing is slipping away, as neural network models are now able to easily obtain state of the art performance with basically no grammar engineering. What’s going on? In this talk, I’ll explore this trend and reflect on the importance of grammar in the modern era. Along the way, I’ll also touch on some related issues affecting parsing (both syntactic and semantic) that I’ve encountered in my time in industry and discuss a few lessons learned.

Biography

David Hall is a senior research scientist working on conversational computing at Semantic Machines. He received his PhD in Computer Science from UC Berkeley advised by Dan Klein. He is the recipient of the 2012 Google PhD Fellowship in Natural Language Processing, an NSF graduate research fellowship, the 2011 EECS Outstanding Graduate Student Instructor award, the journal *Language*’s 2016 best paper award, and a distinguished paper at EMNLP 2012. He has authored fifteen publications at top conferences and has built and released numerous software systems, including a fast GPU-based constituency parser, state-of-the-art parsers for ten languages, the Breeze scientific computing library, and the award-winning Overmind StarCraft agent. He has a B.S. and M.S. from Stanford University, both in Symbolic Systems.

Table of Contents

<i>Automatically Acquired Lexical Knowledge Improves Japanese Joint Morphological and Dependency Analysis</i>	
Daisuke Kawahara, Yuta Hayashibe, Hajime Morita and Sadao Kurohashi	1
<i>Dependency Language Models for Transition-based Dependency Parsing</i>	
Juntao Yu and Bernd Bohnet	11
<i>Lexicalized vs. Delexicalized Parsing in Low-Resource Scenarios</i>	
Agnieszka Falenska and Özlem Çetinoğlu	18
<i>Improving neural tagging with lexical information</i>	
Benoît Sagot and Héctor Martínez Alonso	25
<i>Positional Phrase Attachment over Word Embedding Products</i>	
Pranava Swaroop Madhyastha, Xavier Carreras and Ariadna Quattoni	32
<i>L1-L2 Parallel Dependency Treebank as Learner Corpus</i>	
John Lee, Keying Li and Herman Leung	44
<i>Splitting Complex English Sentences</i>	
John Lee and J. Buddhika K. Pathirage Don	50
<i>Hierarchical Word Structure-based Parsing: A Feasibility Study on UD-style Dependency Parsing in Japanese</i>	
Takaaki Tanaka, Katsuhiko Hayashi and Masaaki Nagata	56
<i>Leveraging Newswire Treebanks for Parsing Conversational Data with Argument Scrambling</i>	
Riyaz A. Bhat, Irshad Bhat and Dipti Sharma	61
<i>Using hyperlinks to improve multilingual partial parsers</i>	
Anders Søgaard	67
<i>Correcting prepositional phrase attachments using multimodal corpora</i>	
Sebastien Delecraz, Alexis Nasr, Frederic Bechet and Benoit Favre	72
<i>Exploiting Structure in Parsing to 1-Endpoint-Crossing Graphs</i>	
Robin Kurtz and Marco Kuhlmann	78
<i>Effective Online Reordering with Arc-Eager Transitions</i>	
Ryosuke Kohita, Hiroshi Noji and Yuji Matsumoto	88
<i>Arc-Hybrid Non-Projective Dependency Parsing with a Static-Dynamic Oracle</i>	
Miryam de Lhoneux, Sara Stymne and Joakim Nivre	99
<i>Encoder-Decoder Shift-Reduce Syntactic Parsing</i>	
Jiangming Liu and Yue Zhang	105
<i>Arc-Standard Spinal Parsing with Stack-LSTMs</i>	
Miguel Ballesteros and Xavier Carreras	115
<i>Coarse-To-Fine Parsing for Expressive Grammar Formalisms</i>	
Christoph Teichmann, Alexander Koller and Jonas Groschwitz	122

Evaluating LSTM models for grammatical function labelling

Bich-Ngoc Do and Ines Rehbein 128

Conference Programme

Wednesday, September 20, 2017

08:30–09:30 *Registration*

09:30–11:00 **Depling Long Talk Session**

11:00–11:30 *Break*

11:30–12:00 **Report on the CoNLL 2017 Shared Task on "Multilingual Parsing from Raw Text to Universal Dependencies"**

12:00–13:00 **First Shared Task on Extrinsic Parser Evaluation (EPE 2017)**

13:00–14:30 *Lunch*

14:30–16:00 **First Shared Task on Extrinsic Parser Evaluation (EPE 2017)**

16:00–16:30 *Break*

16:30–18:00 **Joint Depling & IWPT Panel Discussion**

Thursday, September 21, 2017

08:30–09:30 *Registration*

09:30–10:00 *Opening*

10:00–11:00 **Keynote 1**

10:00–11:00 *Syntactic processing in humans: time course, shallow processing and processing failure*
Vera Demberg

11:00–11:30 *Break*

11:30–13:00 **Long and short talks**

11:30–12:00 *Automatically Acquired Lexical Knowledge Improves Japanese Joint Morphological and Dependency Analysis*
Daisuke Kawahara, Yuta Hayashibe, Hajime Morita and Sadao Kurohashi

12:00–12:20 *Dependency Language Models for Transition-based Dependency Parsing*
Juntao Yu and Bernd Bohnet

12:20–12:40 *Lexicalized vs. Delexicalized Parsing in Low-Resource Scenarios*
Agnieszka Falenska and Özlem Çetinoğlu

12:40–13:00 *Improving neural tagging with lexical information*
Benoît Sagot and Héctor Martínez Alonso

13:00–14:30 *Lunch*

Thursday, September 21, 2017 (continued)

14:30–16:00 Long and short talks

14:30–15:00 *Prepositional Phrase Attachment over Word Embedding Products*
Pranava Swaroop Madhyastha, Xavier Carreras and Ariadna Quattoni

15:00–15:20 *L1-L2 Parallel Dependency Treebank as Learner Corpus*
John Lee, Keying Li and Herman Leung

15:20–15:40 *Splitting Complex English Sentences*
John Lee and J. Buddhika K. Pathirage Don

15:40–16:00 *Hierarchical Word Structure-based Parsing: A Feasibility Study on UD-style Dependency Parsing in Japanese*
Takaaki Tanaka, Katsuhiko Hayashi and Masaaki Nagata

16:00–16:30 Break

16:30–17:30 Short talks

16:30–16:50 *Leveraging Newswire Treebanks for Parsing Conversational Data with Argument Scrambling*
Riyaz A. Bhat, Irshad Bhat and Dipti Sharma

16:50–17:10 *Using hyperlinks to improve multilingual partial parsers*
Anders Søgaard

17:10–17:30 *Correcting prepositional phrase attachments using multimodal corpora*
Sebastien Delecraz, Alexis Nasr, Frederic Bechet and Benoit Favre

Friday, September 22, 2017

08:30–09:30 *Registration*

09:30–10:00 *IWPT Business Meeting*

10:00–11:00 **Keynote 2**

10:00–11:00 *What good is a grammar anyway?*
David Hall

11:00–11:30 *Break*

11:30–13:00 **Long and short talks**

11:30–12:00 *Exploiting Structure in Parsing to 1-Endpoint-Crossing Graphs*
Robin Kurtz and Marco Kuhlmann

12:00–12:30 *Effective Online Reordering with Arc-Eager Transitions*
Ryosuke Kohita, Hiroshi Noji and Yuji Matsumoto

12:30–12:50 *Arc-Hybrid Non-Projective Dependency Parsing with a Static-Dynamic Oracle*
Miryam de Lhoneux, Sara Stymne and Joakim Nivre

13:00–14:30 *Lunch*

Friday, September 22, 2017 (continued)

14:30–16:00 Long and short talks

14:30–15:00 *Encoder-Decoder Shift-Reduce Syntactic Parsing*

Jiangming Liu and Yue Zhang

15:00–15:20 *Arc-Standard Spinal Parsing with Stack-LSTMs*

Miguel Ballesteros and Xavier Carreras

15:20–15:40 *Coarse-To-Fine Parsing for Expressive Grammar Formalisms*

Christoph Teichmann, Alexander Koller and Jonas Groschwitz

15:40–16:00 *Evaluating LSTM models for grammatical function labelling*

Bich-Ngoc Do and Ines Rehbein

16:00–16:30 Closing

