



International Conference on
Dependency Linguistics

Proceedings of the Conference

September 18-20, 2017

Università di Pisa

Istituto di Linguistica Computazionale "A. Zampolli", CNR Pisa



UNIVERSITÀ DI PISA



Istituto di Linguistica
Computazionale "A. Zampolli"
Consiglio Nazionale delle ricerche

Edited by Simonetta Montemagni and Joakim Nivre

Cover design Chiara Mannari



Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017)

Simonetta Montemagni, Joakim Nivre (Eds.)

Linköping Electronic Conference Proceedings No. 139

ISSN: 1650-3686, eISSN: 1650-3740

ISBN: 978-91-7685-467-9

ACL Anthology W17-65

© 2017 The Authors (individual papers)

© 2017 The Editors (collection)

Inclusion of papers in this collection, electronic publication in the Linköping Electronic Conference Proceedings series, and inclusion in the ACL Anthology with permission of the copyright holders

Preface

The Depling 2017 conference in Pisa is the fourth meeting in the recently established series of international conferences on Dependency Linguistics which started in Barcelona in 2011 and continued in Prague and Uppsala in 2013 and 2015, respectively. The initiative to organize special meetings devoted to Dependency Linguistics, which is currently at the forefront of both theoretical and computational linguistics, has received great support from the community. We do hope that the present conference will manage to keep up the high standards set by the previous meetings.

This year we received 41 submissions by 93 authors from 27 countries, one of which was withdrawn before reviewing. Of the remaining 40 submissions (each reviewed by 3 members of the Program Committee), 30 were accepted, resulting in an acceptance rate of 75%. All in all, the proceedings contain a wide range of contributions to Dependency Linguistics, ranging from papers advancing new theoretical models, through empirical studies of one or more languages, as well as experimental investigations of computational systems of dependency parsing and linguistic knowledge extraction, to the design and construction of dependency-based linguistic resources (both treebanks and lexicons) for a wide range of languages.

New to Depling 2017 edition is the fact that the conference is held in conjunction with the biennial meeting of SIGPARSE, namely the International Conference on Parsing Technologies (IWPT 2017), organized by the Special Interest Group on “Natural Language Parsing” of the Association for Computational Linguistics (ACL). IWPT 2017 will take place immediately after Depling 2017, from the 20th to 22nd of September 2017. The two conferences have an overlapping event, held on September 20th and focusing on different aspects of dependency parsing, in which the results of a shared task jointly organized by Depling and IWPT are presented and discussed from different and complementary perspectives.

The shared task, named “Extrinsic Parser Evaluation” (EPE) and playing the role of “bridge event” between the two conferences, is aimed at shedding light on the downstream utility of various dependency representations (at the available levels of accuracy for different parsers), that is, to seek to contrastively isolate the relative contributions of each type of representation (and corresponding parsing systems) to a selection of state-of-the-art systems (which use different types of text and exhibit broad domain and genre variation).

In addition to the accepted papers, the core conference program also includes the contribution of two distinguished keynote speakers, Yoav Goldberg (Bar Ilan University) and Eva Hajičová (Charles University in Prague). We are honoured that they accepted to contribute to Depling 2017 and thank them for agreeing to share their knowledge and expertise on key Dependency Linguistics topics with the conference participants.

Our sincere thanks go to the members of the Program Committee who thoroughly reviewed all the submissions to the conference and provided detailed comments and suggestions, thus ensuring the quality of the published papers. Many thanks to the members of the Local Organizing Committee who took care of all matters related to the local organization of the conference. Thanks are also due to Michela Carlino, who did a great job in putting the proceedings together, and to Chiara Mannari, for designing and constructing the Depling and IWPT+Depling conference websites and continuously updating them. Last but not least, we would like to acknowledge the support from endorsing organizations and institutions and from our sponsors, who generously provided funds and

services that are crucial for the organization of this event. At the time of writing, Depling was sponsored by the newly founded “Italian Association of Computational Linguistics” (AILC) and by the University of Pisa. Special thanks are also due to the Institute for Computational Linguistics “Antonio Zampolli” of the Italian National Research Council (ILC-CNR) for the support in the organization of the event. Thanks finally to everyone who chose to submit their work to Depling 2017, without whom this volume literally would not exist.

We welcome you all to Depling 2017 in Pisa and wish you an enjoyable conference!

Simonetta Montemagni and Joakim Nivre

Program Co-Chairs, Depling 2017

Organizers

Program Co-Chairs

- Simonetta Montemagni, Istituto di Linguistica Computazionale “A. Zampolli” - CNR
- Joakim Nivre, Uppsala University

Local Organizing Committee

- Giuseppe Attardi, Università di Pisa
- Felice Dell’Orletta, Istituto di Linguistica Computazionale “A. Zampolli” - CNR
- Alessandro Lenci, Università di Pisa
- Simonetta Montemagni, Istituto di Linguistica Computazionale “A. Zampolli” - CNR
- Maria Simi, Università di Pisa

Program Committee

- Giuseppe Attardi, Università di Pisa
- Miguel Ballesteros, IBM Research Watson
- Xavier Blanco, Universitat Autònoma de Barcelona
- Igor Boguslavsky, Universidad Politecnica de Madrid and Russian Academy of Sciences
- Bernd Bohnet, Google
- Cristina Bosco, Università di Torino
- Marie Candito, Université Paris Diderot
- Jinho Choi, University of Colorado at Boulder
- Benoit Crabbé, Université Paris Diderot
- Eric De La Clergerie, INRIA
- Felice Dell’Orletta, Istituto di Linguistica Computazionale “A. Zampolli” - CNR
- Marie-Catherine de Marneffe, The Ohio State University
- Kim Gerdes, Sorbonne Nouvelle
- Filip Ginter, University of Turku
- Koldo Gojenola, University of the Basque Country UPV/EHU
- Carlos Gómez-Rodríguez, Universidade da Coruña
- Eva Hajičová, Charles University in Prague
- Richard Hudson, University College London
- Leonid Iomdin, Russian Academy of Sciences
- Sylvain Kahane, Université Paris Ouest Nanterre
- Marco Kuhlmann, Linköping University
- François Lareau, Université de Montréal
- Alessandro Lenci, Università di Pisa
- Beth Levin, Stanford University
- Haitao Liu, Zhejiang University
- Marketa Lopatkova, Charles University in Prague
- Ryan McDonald, Google

- Igor Mel'čuk, University of Montreal
- Wolfgang Menzel, Hamburg University
- Paola Merlo, Université de Genève
- Jasmina Milicevic, Dalhousie University
- Henrik Høeg Müller, Copenhagen Business School
- Alexis Nasr, Université de la Méditerranée
- Pierre Nugues, Lund University
- Kemal Oflazer, Carnegie Mellon University Qatar
- Timothy Osborne, Zhejiang University
- Jarmila Panevova, Charles University in Prague
- Alain Polguère, Université de Lorraine ATILF CNRS
- Prokopis Prokopidis, Institute for Language and Speech Processing/Athena RC, Greece
- Owen Rambow, Columbia University
- Ines Rehbein, Potsdam University
- Dipti Sharma, IIT, Hyderabad
- Maria Simi, Università di Pisa
- Reut Tsarfaty, Open University of Israel
- Giulia Venturi, Istituto di Linguistica Computazionale “A. Zampolli” - CNR
- Leo Wanner, Pompeu Fabra University
- Daniel Zeman, Charles University in Prague
- Yue Zhang, Singapore University of Technology and Design

Supporting Institutions

- Università degli Studi di Pisa
 - Dipartimento di Filologia, Letteratura e Linguistica
 - Dipartimento di Informatica
- Istituto di Linguistica Computazionale “A. Zampolli”, Consiglio Nazionale delle Ricerche

Sponsor

- Associazione Italiana di Linguistica Computazionale (AILC)



Table of Contents

Invited Talk: <i>Capturing Dependency Syntax with “Deep” Sequential Models</i> Yoav Goldberg	1
Invited Talk: <i>Syntax-Semantics Interface: A Plea for a Deep Dependency Sentence Structure</i> Eva Hajičová	2
<i>The Benefit of Syntactic vs. Linear N-Grams for Linguistic Description</i> Melanie Andresen and Heike Zinsmeister	4
<i>On the Predicate-Argument Structure: Internal and Absorbing Scope</i> Igor Boguslavsky	15
<i>On the Order of Words in Italian: A Study on Genre vs Complexity</i> Dominique Brunato and Felice Dell’Orletta	25
<i>Revising the METU-Sabancı Turkish Treebank: An Exercise in Surface-Syntactic Annotation of Agglutinative Languages</i> Alicia Burga, Alp Öktem and Leo Wanner	32
<i>Enhanced UD Dependencies with Neutralized Diathesis Alternation</i> Marie Candito, Bruno Guillaume, Guy Perrier and Djamé Seddah	42
<i>Classifying Languages by Dependency Structure. Typologies of Delexicalized Universal Dependency Treebanks</i> Xinying Chen and Kim Gerdes	54
<i>A Dependency Treebank for Kurmanji Kurdish</i> Memduh Gökırmak and Francis M. Tyers	64

<i>What are the Limitations on the Flux of Syntactic Dependencies? Evidence from UD Treebanks</i> Sylvain Kahane, Chunxiao Yan and Marie-Amélie Botalla	73
<i>Fully Delexicalized Contexts for Syntax-Based Word Embeddings</i> Jenna Kanerva, Sampo Pyysalo and Filip Ginter	83
<i>Universal Dependencies for Dargwa Mehweb</i> Alexandra Kozhukhar	92
<i>Menzerath-Altmann Law in Syntactic Dependency Structure</i> Ján Mačutek, Radek Čech and Jiří Milička	100
<i>Assessing the Annotation Consistency of the Universal Dependencies Corpora</i> Marie-Catherine de Marneffe, Matias Grioni, Jenna Kanerva and Filip Ginter	108
<i>To What Extent is Immediate Constituency Analysis Dependency-Based? A Survey of Foundational Texts</i> Nicolas Mazziotta and Sylvain Kahane	116
<i>Dependency Structure of Binary Conjunctions (of the IF..., THEN... Type)</i> Igor Mel'čuk	127
<i>Non-Projectivity in Serbian: Analysis of Formal and Linguistic Properties</i> Aleksandra Miletic and Assaf Urieli	135
<i>Prices Go Up, Surge, Jump, Spike, Skyrocket, Go through the Roof... Intensifier Collocations with Parametric Nouns of Type PRICE</i> Jasmina Milićević	145
<i>Chinese Descriptive and Resultative V-de Constructions. A Dependency-based Analysis</i> Ruochen Niu	154

<i>The Component Unit. Introducing a Novel Unit of Syntactic Analysis</i>	
Timothy Osborne and Ruochen Niu	165
<i>Control vs. Raising in English. A Dependency Grammar Account</i>	
Timothy Osborne and Matthew Reeve	176
<i>Segmentation Granularity in Dependency Representations for Korean</i>	
Jungyeul Park	187
<i>Universal Dependencies for Portuguese</i>	
Alexandre Rademaker, Fabricio Chalub, Livy Real, Cláudia Freitas, Eckhard Bick and Valeria de Paiva	197
<i>UDLex: Towards Cross-language Subcategorization Lexicons</i>	
Giulia Rambelli, Alessandro Lenci and Thierry Poibeau	207
<i>Universal Dependencies are Hard to Parse – or are They?</i>	
Ines Rehbein, Julius Steen, Bich-Ngoc Do and Anette Frank	218
<i>Annotating Italian Social Media Texts in Universal Dependencies</i>	
Manuela Sanguinetti, Cristina Bosco, Alessandro Mazzei, Alberto Lavelli and Fabio Tamburini	229
<i>Hungarian Copula Constructions in Dependency Syntax and Parsing</i>	
Katalin Ilona Simkó and Veronika Vincze	240
<i>Semgrex-Plus: a Tool for Automatic Dependency-Graph Rewriting</i>	
Fabio Tamburini	248
<i>Unity in Diversity: a Unified Parsing Strategy for Major Indian Languages</i>	
Juhi Tandon and Dipti Misra Sharma	255

<i>Quantitative Comparative Syntax on the Cantonese-Mandarin Parallel Dependency Treebank</i>	
Tak-sum Wong, Kim Gerdes, Herman Leung and John Lee	266
 <i>Understanding Constraints on Non-Projectivity Using Novel Measures</i>	
Himanshu Yadav, Ashwini Vaidya and Samar Husain	276
 <i>Core Arguments in Universal Dependencies</i>	
Daniel Zeman	287