

Depling 2019

**Fifth International Conference on Dependency Linguistics  
(Depling, SyntaxFest 2019)**

**Proceedings**

27–28 August, 2019  
held within the **SyntaxFest** 2019, 26–30 August  
Paris, France

©2019 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](mailto:acl@aclweb.org)

ISBN 978-1-950737-63-5

## Preface

The Fifth edition of the International Conference on Dependency Linguistics (Depling) follows a bi-annual series that started in 2011, in Barcelona and continued in Prague (2013), Uppsala (2015), and Pisa (2017). The series responds to the growing need for linguistic meetings dedicated to approaches in syntax, semantics and the lexicon that are centered around dependency structures as a central linguistic notion.

This year's edition is special as Depling is part of the first SyntaxFest, a grouping of four events, which took place in Paris, France, during the last week of August:

- the Fifth International Conference on Dependency Linguistics (Depling 2019)
- the First Workshop on Quantitative Syntax (Quasy)
- the 18th International Workshop on Treebanks and Linguistic Theories (TLT 2019)
- the Third Workshop on Universal Dependencies (UDW 2019)

The use of corpora for NLP and linguistics has only increased in recent years. In NLP, machine learning systems are by nature data-intensive, and in linguistics there is a renewed interest in the empirical validation of linguistic theory, particularly through corpus evidence. While the first statistical parsers have long been trained on the Penn treebank phrase structures, dependency treebanks, whether natively annotated with dependencies, or converted from phrase structures, have become more and more popular, as evidenced by the success of the Universal Dependency project, currently uniting 120 treebanks in 80 languages, annotated in the same dependency-based scheme. The availability of these resources has boosted empirical quantitative studies in syntax. It has also led to a growing interest in theoretical questions around syntactic dependency, its history, its foundations, and the analyses of various constructions in dependency-based frameworks. Furthermore, the availability of large, multilingual annotated data sets, such as those provided by the Universal Dependencies project, has made cross-linguistic analysis possible to an extent that could only be dreamt of only a few years ago.

In this context it was natural to bring together TLT (Treebanks and Linguistic Theories), the historical conference on treebanks as linguistic resources, Depling (The international conference on Dependency Linguistics), the conference uniting research on models and theories around dependency representations, and UDW (Universal Dependency Workshop), the annual meeting of the UD project itself. Moreover, in order to create a point of contact with the large community working in quantitative linguistics it seemed expedient to create a workshop dedicated to quantitative syntactic measures on treebanks and raw corpora, which gave rise to Quasy, the first workshop on Quantitative Syntax. And this led us to the first SyntaxFest.

Because the potential audience and submissions to the four events were likely to have substantial overlap, we decided to have a single reviewing process for the whole SyntaxFest. Authors could choose to submit their paper to one or several of the four events, and in case of acceptance, the program co-chairs would decide which event to assign the accepted paper to.

This choice was found to be an appropriate one, as most submissions were submitted to several of the events. Indeed, there were 40 long paper submissions, with 14 papers submitted to Quasy, 31 to Depling, 13 to TLT and 16 to UDW. Among them, 28 were accepted (6 at Quasy, 10 at Depling, 6 at TLT, 6 at UDW). Note that due to multiple submissions, the acceptance rate is defined at the level of the whole SyntaxFest (around 70%). As far as short papers are concerned, 62 were submitted (24 to Quasy, 41 to

Depling, 35 to TLT and 37 to UDW), and 41 were accepted (8 were presented at Quasy, 14 at Depling, 9 at TLT and 9 at UDW), leading to an acceptance rate for short papers of around 66%.

We are happy to announce that the first SyntaxFest has been a success, with over 110 registered participants, most of whom attended for the whole week.

SyntaxFest is the result of efforts from many people. Our sincere thanks go to the reviewers who thoroughly reviewed all the submissions to the conference and provided detailed comments and suggestions, thus ensuring the quality of the published papers.

We would also like to warmly extend our thanks to the five invited speakers,

- Ramon Ferrer i Cancho - Universitat Politècnica de Catalunya (UPC)
- Emmanuel Dupoux - ENS/CNRS/EHESS/INRIA/PSL Research University, Paris
- Barbara Plank - IT University of Copenhagen
- Paola Merlo - University of Geneva
- Adam Przepiórkowski - University of Warsaw / Polish Academy of Sciences / University of Oxford

We are grateful to the Université Sorbonne Nouvelle for generously making available the Amphithéâtre du Monde Anglophone, a very pleasant venue in the heart of Paris. We would like to thank the ACL SIGPARSE group for its endorsement and all the institutions who gave financial support for SyntaxFest:

- the "Laboratoire de Linguistique formelle" (Université Paris Diderot & CNRS)
- the "Laboratoire de Phonétique et Phonologie" (Université Sorbonne Nouvelle & CNRS)
- the Modyco laboratory (Université Paris Nanterre)
- the "École Doctorale Connaissance, Langage, Modélisation" (CLM) - ED 139
- the "Université Sorbonne Nouvelle"
- the "Université Paris Nanterre"
- the Empirical Foundations of Linguistics Labex (EFL)
- the ATALA association
- Google
- Inria and its Almanach team project.

Finally, we would like to express special thanks to the students who have been part of the local organizing committee. We warmly acknowledge the enthusiasm and community spirit of:

Danrun Cao, Université Paris Nanterre

Marine Courtin, Sorbonne Nouvelle

Chuanming Dong, Université Paris Nanterre

Yoann Dupont, Inria

Mohammed Galal, Sohag University

Gaël Guibon, Inria  
Yixuan Li, Sorbonne Nouvelle  
Lara Perinetti, Inria et Fortia Financial Solutions  
Mathilde Regnault, Lattice and Inria  
Pierre Rochet, Université Paris Nanterre  
Chunxiao Yan, Université Paris Nanterre

Marie Candito, Kim Gerdes, Sylvain Kahane, Djamé Seddah (local organizers and co-chairs),  
and Xinying Chen, Ramon Ferrer-i-Cancho, Alexandre Rademaker, Francis Tyers (co-chairs)

September 2019

## Program co-chairs

The chairs for each event (and co-chairs for the single SyntaxFest reviewing process) are:

- Quasy:
  - Xinying Chen (Xi’an Jiaotong University / University of Ostrava)
  - Ramon Ferrer i Cancho (Universitat Politècnica de Catalunya)
- Depling:
  - Kim Gerdes (LPP, Sorbonne Nouvelle & CNRS / Almanach, INRIA)
  - Sylvain Kahane (Modyco, Paris Nanterre & CNRS)
- TLT:
  - Marie Candito (LLF, Paris Diderot & CNRS)
  - Djamé Seddah (Paris Sorbonne / Almanach, INRIA)
  - with the help of Stephan Oepen (University of Oslo, previous co-chair of TLT) and Kilian Evang (University of Düsseldorf, next co-chair of TLT)
- UDW:
  - Alexandre Rademaker (IBM Research, Brazil)
  - Francis Tyers (Indiana University and Higher School of Economics)
  - with the help of Teresa Lynn (ADAPT Centre, Dublin City University) and Arne Köhn (Saarland University)

## Local organizing committee of the SyntaxFest

Marie Candito, Université Paris-Diderot (co-chair)  
Kim Gerdes, Sorbonne Nouvelle (co-chair)  
Sylvain Kahane, Université Paris Nanterre (co-chair)  
Djamé Seddah, University Paris-Sorbonne (co-chair)  
Danrun Cao, Université Paris Nanterre  
Marine Courtin, Sorbonne Nouvelle  
Chuanming Dong, Université Paris Nanterre  
Yoann Dupont, Inria  
Mohammed Galal, Sohag University  
Gaël Guibon, Inria  
Yixuan Li, Sorbonne Nouvelle  
Lara Perinetti, Inria et Fortia Financial Solutions  
Mathilde Regnault, Lattice and Inria  
Pierre Rochet, Université Paris Nanterre  
Chunxiao Yan, Université Paris Nanterre

## **Program committee for the whole SyntaxFest**

Patricia Amaral (Indiana University Bloomington)  
Miguel Ballesteros (IBM)  
David Beck (University of Alberta)  
Emily M. Bender (University of Washington)  
Ann Bies (Linguistic Data Consortium, University of Pennsylvania)  
Igor Boguslavsky (Universidad Politécnica de Madrid)  
Bernd Bohnet (Google)  
Cristina Bosco (University of Turin)  
Gosse Bouma (Rijksuniversiteit Groningen)  
Miriam Butt (University of Konstanz)  
Radek Āech (University of Ostrava)  
Giuseppe Giovanni Antonio Celano (University of Pavia)  
Āagrı Āöltekin (University of Tuebingen)  
Benoit Crabbé (Paris Diderot University)  
Éric De La Clergerie (INRIA)  
Miryam de Lhoneux (Uppsala University)  
Marie-Catherine de Marneffe (The Ohio State University)  
Valeria de Paiva (Samsung Research America and University of Birmingham)  
Felice Dell'Orletta (Istituto di Linguistica Computazionale "Antonio Zampolli" - ILC CNR)  
Kaja Dobrovoljc (Jožef Stefan Institute)  
Leonel Figueiredo de Alencar (Universidade federal do Ceará)  
Jennifer Foster (Dublin City University, Dublin 9, Ireland)  
Richard Futrell (University of California, Irvine)  
Filip Ginter (University of Turku)  
Koldo Gojenola (University of the Basque Country UPV/EHU)  
Kristina Gulordava (Universitat Pompeu Fabra)  
Carlos Gómez-Rodríguez (Universidade da Coruña)  
Memduh Gökirmak (Charles University, Prague)  
Jan HajiĀ (Charles University, Prague)  
Eva HajiĀov (Charles University, Prague)  
Barbora Hladk (Charles University, Prague)  
Richard Hudson (University College London)  
Leonid Iomdin (Institute for Information Transmission Problems, Russian Academy of Sciences)  
Jingyang Jiang (Zhejiang University)  
Sandra Kbler (Indiana University Bloomington)  
Franois Lareau (OLST, Universit de Montral)  
John Lee (City University of Hong Kong)  
Nicholas Lester (University of Zurich)  
Lori Levin (Carnegie Mellon University)  
Haitao Liu (Zhejiang University)  
Jn MaĀutek (Comenius University, Bratislava, Slovakia)  
Nicolas Mazziotta (Universit)  
Ryan Mcdonald (Google)  
Alexander Mehler (Goethe-University Frankfurt am Main, Text Technology Group)

Wolfgang Menzel (Department of Informatik, Hamburg University)  
Paola Merlo (University of Geneva)  
Jasmina Milićević (Dalhousie University)  
Simon Mille (Universitat Pompeu Fabra)  
Simonetta Montemagni (ILC-CNR)  
Jiří Mírovský (Charles University, Prague)  
Alexis Nasr (Aix-Marseille Université)  
Anat Ninio (The Hebrew University of Jerusalem)  
Joakim Nivre (Uppsala University)  
Pierre Nugues (Lund University, Department of Computer Science Lund, Sweden)  
Kemal Oflazer (Carnegie Mellon University-Qatar)  
Timothy Osborne (independent)  
Petya Osenova (Sofia University and IICT-BAS)  
Jarmila Panevová (Charles University, Prague)  
Agnieszka Patejuk (Polish Academy of Sciences / University of Oxford)  
Alain Polguère (Université de Lorraine)  
Prokopis Prokopidis (Institute for Language and Speech Processing/Athena RC)  
Ines Rehbein (Leibniz Science Campus)  
Rudolf Rosa (Charles University, Prague)  
Haruko Sanada (Rissho University)  
Sebastian Schuster (Stanford University)  
Maria Simi (Università di Pisa)  
Reut Tsarfaty (Open University of Israel)  
Zdenka Uresova (Charles University, Prague)  
Giulia Venturi (ILC-CNR)  
Veronika Vincze (Hungarian Academy of Sciences, Research Group on Artificial Intelligence)  
Relja Vulcanovic (Kent State University at Stark)  
Leo Wanner (ICREA and University Pompeu Fabra)  
Michael White (The Ohio State University)  
Chunshan Xu (Anhui Jianzhu University)  
Zhao Yiyi (Communication University of China)  
Amir Zeldes (Georgetown University)  
Daniel Zeman (Univerzita Karlova)  
Hongxin Zhang (Zhejiang University)  
Heike Zinsmeister (University of Hamburg)  
Robert Östling (Department of Linguistics, Stockholm University)  
Lilja Øvrelid (University of Oslo)

## **Additional reviewers**

James Barry  
Ivan Vladimir Meza Ruiz  
Rebecca Morris  
Olga Sozinova  
He Zhou

## Table of Contents

SyntaxFest 2019 Invited talk - Inductive biases and language emergence in communicative agents <i>Emmanuel Dupoux</i>	1
SyntaxFest 2019 Invited talk - Transferring NLP models across languages and domains . . . . . <i>Barbara Plank</i>	2
Syntactic dependencies correspond to word pairs with high mutual information . . . . . <i>Richard Futrell, Peng Qian, Edward Gibson, Evelina Fedorenko and Idan Blank</i>	3
Reflexives in Czech from a Dependency Perspective . . . . . <i>Vaclava Kettnerova and Marketa Lopatkova</i>	14
Coordination of Unlike Grammatical Functions . . . . . <i>Agnieszka Patejuk and Adam Przepiórkowski</i>	26
Predicate Catenae: A Dependency Grammar Analysis of It-Clefts . . . . . <i>Timothy Osborne</i>	38
Noun Phrases Rooted by Adjectives: A Dependency Grammar Analysis of the Big Mess Construction <i>Timothy Osborne</i>	49
Cliticization of Serbian Personal Pronouns and Auxiliary Verbs. A Dependency-Based Account . . . . . <i>Jasmina Milicevic</i>	60
The evolution of spatial rationales in Tesnière’s stemmas . . . . . <i>Nicolas Mazziotta</i>	69
Toward a cognitive dependency grammar of Hungarian . . . . . <i>András Imrényi</i>	81
Interpreting and defining connections in a dependency structure . . . . . <i>Sylvain Kahane</i>	89
Identifying Grammar Rules for Language Education with Dependency Parsing in German . . . . . <i>Eleni Metheniti, Pomi Park, Kristina Kolesova and Günter Neumann</i>	100
How to Parse Low-Resource Languages: Cross-Lingual Parsing, Target Language Annotation, or Both? . . . . . <i>Ailsa Meechan-Maddon and Joakim Nivre</i>	112
Word order variation in Mbyá Guaraní . . . . . <i>Angelika Kiss and Guillaume Thomas</i>	121
Examining MDD and MHD as Syntactic Complexity Measures with Intermediate Japanese Learner Corpus Data . . . . . <i>Saeko Komori, Masatoshi Sugiura and Wenping Li</i>	130
Dependency Parsing as Sequence Labeling with Head-Based Encoding and Multi-Task Learning . . . . . <i>Ophélie Lacroix</i>	136
Towards Deep Universal Dependencies . . . . . <i>Kira Droganova and Daniel Zeman</i>	144
Delimiting Adverbial Meanings. A corpus-based comparative study on Czech spatial prepositions and their English equivalents . . . . .	153

<i>Marie Mikulová, Veronika Kolářová, Jarmila Panevová and Eva Hajičová</i>	
<b>A Spanish E-dictionary of Collocations</b> . . . . .	160
<i>Maria Auxiliadora Barrios Rodriguez and Igor Boguslavsky</i>	
<b>Exceptive constructions. A Dependency-based Analysis</b> . . . . .	168
<i>Mohamed Galal, Sylvain Kahane and Yomna Safwat</i>	
<b>Quantitative Analysis on verb valence evolution of Chinese</b> . . . . .	175
<i>Bingli Liu and Chunshan Xu</i>	
<b>Association Metrics in Neural Transition-Based Dependency Parsing</b> . . . . .	181
<i>Patricia Fischer, Sebastian Pütz and Daniël de Kok</i>	
<b>Presenting TWITTIRÒ-UD: An Italian Twitter Treebank in Universal Dependencies</b> . . . . .	190
<i>Alessandra Teresa Cignarella, Cristina Bosco and Paolo Rosso</i>	
<b>Pāninian Syntactico-Semantic Relation Labels</b> . . . . .	198
<i>Amba Kulkarni and Dipti Sharma</i>	
<b>Experiments on human incremental parsing</b> . . . . .	209
<i>Leonid Mityushin and Leonid Iomdin</i>	
<b>Character-level Annotation for Chinese Surface-Syntactic Universal Dependencies</b> . . . . .	216
<i>Yixuan Li, Gerdes Kim and Dong Chuanming</i>	