Preface

The Sixteenth International Workshop on Treebanks and Linguistic Theories (TLT16) is being held at Charles University, Czech Republic, 23–24 January 2018, for the second time in Prague, which hosted TLT already in 2006. This year, TLT16 is co-located with the Workshop on Data Provenance and Annotation in Computational Linguistics 2018 (January 22, 2018, at the same place and venue) and immediately followed by the 2nd Workshop on Corpus-based Research in the Humanities, held in Vienna, Austria.

This year, TLT16 received 32 submissions of which 15 have been selected to be presented as oral presentations, and additional seven have been asked to present as posters. We were happy that our invitation to give a plenary talk has been accepted by both Lilja Øvrelid of University of Oslo in Norway (with a talk “Downstream use of syntactic analysis: does representation matter?”) and Marie Candito, of University of Paris Diderot, France (“Annotating and parsing to semantic frames: feedback from the French FrameNet project”). We have also included a panel discussion on the very topic of Treebanks and Linguistic Theories – to discuss opinions of the future of the field and the workshop with the participants.

We are grateful to the members of the program committee, who worked hard to review the submissions and provided authors with valuable feedback. We would also like to thank various sponsors, mainly the Faculty of Mathematics and Physics of Charles University, providing local and logistical and accounting and financial support, and the Centre for Advanced Study (CAS) at the Norwegian Academy of Science and Letters for supporting one of the invited speakers.

Last but not least, we would like to thank all authors for submitting interesting and relevant papers, and we wish all participants a fruitful workshop.

December 2017

Jan Hajič (also in the name of local organizers), Charles University, Prague, Czech Republic
Sandra Kuebler, Indiana University, Bloomington, Indiana, USA
Stephan Oepen, University of Oslo, Norway, and CAS, Norwegian Academy of Science and Letters, Oslo
Markus Dickinson, Indiana University, Bloomington, Indiana, USA

Program Committee Co-chairs
Program Chairs

Jan Hajič  
Sandra Kübler  
Stephan Oepen  
Markus Dickinson

Indiana University Bloomington
Universitetet i Oslo
Indiana University Bloomington

Reviewers

Patricia Amaral  
Emily M. Bender  
Eckhard Bick  
Ann Bies  
Gosse Bouma  
Miriam Butt  
Silvie Cinková  
Marie-Catherine de Marneffe  
Koenraad De Smedt  
Tomaž Erjavec  
Filip Ginter  
Memduh Gokirmak  
Eva Hajicova  
Dag Haug  
Barbora Hladka  
Lori Levin  
Teresa Lynn  
Adam Meyers  
Jiří Mírovský  
Kaili Müürisep  
Joakim Nivre  
Petya Osenova  
Agnieszka Patejuk  
Tatjana Scheffler  
Olga Scrivner  
Djamé Seddah  
Michael White  
Nianwen Xue  
Daniel Zeman  
Heike Zinsmeister

Indiana University Bloomington
University of Washington
University of Southern Denmark
Linguistic Data Consortium, University of Pennsylvania
Rijksuniversiteit Groningen
Konstanz
Charles University, Prague
The Ohio State University
University of Bergen
Dept. of Knowledge Technologies, Jožef Stefan Institute
University of Turku
Istanbul Technical University
Charles University, Prague
University of Oslo
Charles University, Prague
Carnegie Mellon University
Dublin City University
New York University
Charles University, Prague
University of Tartu
Uppsala University
Sofia University and IICT-BAS
Institute of Computer Science, Polish Academy of Sciences
Universität Potsdam
Indiana University Bloomington
Alpage/Université Paris la Sorbonne
The Ohio State University
Brandeis University
Charles University, Prague
University of Hamburg

Local Organizing Committee

Jan Hajič (chair)  
Kateřina Bryanová  
Zdeňka Urešová  
Eduard Bejček (publication chair)

Charles University, Prague
Charles University, Prague
Charles University, Prague
Charles University, Prague
# Table of Contents

## Invited Speakers

- Annotating and parsing to semantic frames: feedback from the French FrameNet project. 
  *Marie Candito*  
  v
- Downstream use of syntactic analysis: does representation matter?  
  *Lilja Øvrelid*  
  vi

## Tuesday, 23rd January

- Distributional regularities of verbs and verbal adjectives: Treebank evidence and broader implications.  
  *Daniël de Kok, Patricia Fischer, Corina Dima and Erhard Hinrichs*  
  1
- UD Annotatrix: An annotation tool for Universal Dependencies  
  *Francis M. Tyers, Maria Sheyanova and Jonathan Washington*  
  10
- The Treebanked Conspiracy. Actors and Actions in Bellum Catilinae  
  *Marco Passarotti and Berta Gonzalez Saavedra*  
  18
- Universal Dependencies-based syntactic features in detecting human translation varieties  
  *Maria Kunilovskaya and Andrey Kutuzov*  
  27
- Graph Convolutional Networks for Named Entity Recognition  
  *Alberto Cetoli, Stefano Bragaglia, Andrew O’Harney and Marc Sloan*  
  37
- Extensions to the GreTEL Treebank Query Application  
  *Jan Odijk, Martijn van der Klis and Sheean Spoel*  
  46
- The Relation of Form and Function in Linguistic Theory and in a Multilayer Treebank  
  *Eduard Bejček, Eva Hajičová, Marie Mikulová and Jarmila Panevová*  
  56
- Literal readings of multiword expressions: as scarce as hen’s teeth  
  *Agata Savary and Silvio Ricardo Cordeiro*  
  64
- Querying Multi-word Expressions Annotation with CQL  
  *Natalia Klyueva, Anna Vernerova and Behrang Qasemizadeh*  
  73
- REALEC learner treebank: annotation principles and evaluation of automatic parsing  
  *Olga Lyashevskaya and Irina Panteleeva*  
  80
- A semiautomatic lemmatisation procedure for treebanks. Old English strong and weak verbs  
  *Marta Tío Sáenz and Dario Metola Rodriguez*  
  88
- Data point selection for genre-aware parsing  
  *Ines Rehbein and Felix Bildhauer*  
  95
- Error Analysis of Cross-lingual Tagging and Parsing  
  *Rudolf Rosa and Zdeněk Žabokrtský*  
  106

## Wednesday, 24th January

- A Telugu treebank based on a grammar book  
  *Taraka Rama and Sowmya Vajjala*  
  119
- Recent Developments within BulTreeBank  
  *Petya Osenova and Kiril Simov*  
  129
Towards a dependency-annotated treebank for Bambara
Ekaterina Aplonova and Francis M. Tyers

Merging the Trees – Building a Morphological Treebank for German from Two Resources
Petra Steiner

What I think when I think about treebanks
Anders Søgaard

Syntactic Semantic Correspondence in Dependency Grammar
Catalina Maranduc, Catalin Mittelu and Victoria Bobicev

Multi-word annotation in syntactic treebanks – Propositions for Universal Dependencies
Sylvain Kahane, Marine Courtin and Kim Gerdes

A Universal Dependencies Treebank for Marathi
Vinit Ravishankar

Dangerous Relations in Dependency Treebanks
Chiara Alzetta, Felice Dell’Orletta, Simonetta Montemagni and Giulia Venturi