TLT 2023

The 21st International Workshop on Treebanks and Linguistic Theories (TLT, GURT/SyntaxFest 2023)

Proceedings of the Conference

March 9-12, 2023

The TLT organizers gratefully acknowledge the support from the following sponsors.

The Georgetown College of Arts & Sciences, the Georgetown Faculty of Languages and Linguistics, and the Georgetown Department of Linguistics

Introduction

The 21st International Workshop on Treebanks and Linguistic Theories (TLT 2023) follows an annual series that started in 2002 in Sozopol, Bulgaria. TLT addresses all aspects of treebank design, development, and use. "Treebank" is taken in a broad sense, comprising any spoken, signed, or written data augmented with computationally processable annotations of linguistic structure at various levels. For the first time, TLT is part of GURT2023, an annual linguistics conference held at Georgetown University, which this year co-locates four related but independent events:

- The Seventh International Conference on Dependency Linguistics (Depling 2023)

- The 21st International Workshop on Treebanks and Linguistic Theories (TLT 2023)

- The Sixth Workshop on Universal Dependencies (UDW 2023)

- The First International Workshop on Construction Grammars and NLP (CxGs+NLP 2023)

The Georgetown University Round Table on Linguistics (GURT) is a peer-reviewed annual linguistics conference held continuously since 1949 at Georgetown University in Washington DC, with topics and co-located events varying from year to year.

In 2023, under an overarching theme of 'Computational and Corpus Linguistics', GURT/SyntaxFest continues the tradition of SyntaxFest 2019 and SyntaxFest 2021/22 in bringing together multiple events that share a common interest in using corpora and treebanks for empirically validating syntactic theories, studying syntax from quantitative and theoretical points of view, and for training machine learning models for natural language processing. Much of this research is increasingly multilingual and cross-lingual and requires continued systematic analysis from various theoretical, applied, and practical perspectives. New this year, the CxGs+NLP workshop brings a usage-based perspective on how form and meaning interact in language.

For these reasons and encouraged by the success of the previous editions of SyntaxFest, we —the chairs of the four events— decided to facilitate another co-located event at GURT 2023 in Washington DC.

As in past co-located events involving several of the workshops, we organized a single reviewing process, with identical paper formats for all four events. Authors could indicate (multiple) venue preferences, but the ultimate assignment of papers to events for accepted papers was made by the program chairs.

33 long papers were submitted, 11 to Depling, 16 to TLT, 10 to UDW and 10 to CxGs+NLP. The program chairs accepted 27 (82%) and assigned 7 to Depling, 6 to TLT, 5 to UDW and 9 to CxGs+NLP.

16 short papers were submitted, 6 of which to Depling, 6 to TLT, 10 to UDW and 2 to CxGs+NLP. The program chairs accepted 9 (56%) and assigned 2 to Depling, 2 to TLT, 3 to UDW, and 2 to CxGs+NLP.

Our sincere thanks go to everyone who is making this event possible: everybody who submitted their papers; Georgetown University Linguistics Department students and staff—including Lauren Levine, Jessica Lin, Ke Lin, Mei-Ling Klein, and Conor Sinclair—for their organizational assistance; and of course, the reviewers for their time and their valuable comments and suggestions. Special thanks are due to Georgetown University, and specifically to the Georgetown College of Arts & Sciences and the Faculty of Languages and Linguistics for supporting the conference with generous funding. Finally, we would also like to thank ACL SIGPARSE for its endorsement and the ACL Anthology for publishing the proceedings.

Owen Rambow, François Lareau (Depling2023 Chairs)
Daniel Dakota, Kilian Evang, Sandra Kübler, Lori Levin (TLT2023 Chairs)
Loïc Grobol, Francis Tyers (UDW2023 chairs)
Claire Bonial Harish Tayyar Madabushi (CxG+NLP2023 Chairs)
Nathan Schneider, Amir Zeldes (GURT2023 Organizers)
March 2023

Organizing Committee

Depling2023 Chairs

Owen Rambow, Stony Brook University
François Lareau, Université de Montréal

TLT2023 Chairs

Daniel Dakota, Indiana University
Kilian Evang, Heinrich Heine University Düsseldorf
Sandra Kübler, Indiana University
Lori Levin, Carnegie Mellon University

UDW2023 Chairs

Loïc Grobol, Université Paris Nanterre
Francis Tyers, Indiana University

CxGs+NLP2023 Chairs

Claire Bonial, U.S. Army Research Lab
Harish Tayyar Madabushi, The University of Bath

GURT2023 Organizers

Amir Zeldes, Georgetown University
Nathan Schneider, Georgetown University

GURT2023 Student Assistants

Lauren Levine, Georgetown University
Ke Lin, Georgetown University
Jessica Lin, Georgetown University

Program Committee

Program Committee for the Whole of GURT2023

Lasha Abzianidze, Utrecht University
Patricia Amaral, Indiana University
Valerio Basile, University of Turin
Emily Bender, University of Washington
Bernd Bohnet, Google
Claire Bonial, Army Research Lab
Gosse Bouma, University of Groningen
Miriam Butt, Universität Konstanz
Marie Candito, Université de Paris
Giuseppe G. A. Celano, Universität Leipzig
Xinying Chen, Xi'an Jiaotong University
Silvie Cinkova, Charles University Prague
Cagri Coltekin, Universität Tübingen
Stefania Degaetano-Ortlieb, Universität des Saarlandes
Éric Villemonte de la Clergerie, INRIA
Miryam de Lhoneux, KU Leuven
Valeria de Paiva, Topos Institute
Lucia Donatelli, Saarland University
Timothy Dozat, Google
Kim Gerdes, Université Paris-Saclay
Koldo Gojenola, University of the Basque Country
Loïc Grobol, Université Paris Nanterre
Bruno Guillaume, INRIA
Dag Trygve Truslew Haug, University of Oslo
Jena Hwang, Allen Institute for Artificial Intelligence
András Imrényi, Eötvös Lorand University
Alessandro Lenci, University of Pisa
Lori Levin, Carnegie Mellon University
Markéta Lopatková, Charles University Prague
Sylvain Kahane, Université Paris Nanterre
Jordan Kodner, State University of New York, Stony Brook
Sandra Kübler, Indiana University
Jan Macutek, Mathematical Institute, Slovak Academy of Sciences
Harish Tayyar Madabushi, University of Sheffield
Nicolas Mazziotta, Université de Liège
Alexander Mehler, Johann Wolfgang Goethe Universität Frankfurt am Main
Simon Mille, Dublin City University
Pierre André Ménard, Computer research institute of Montréal
Yusuke Miyao, The University of Tokyo
Simonetta Montemagni, ILC-CNR
Alexis Nasr, Aix Marseille Univ
Joakim Nivre, Uppsala University
Pierre Nugues, Lund University
Timothy John Osborne, Zhejiang University
Petya Osenova, Bulgarian Academy of Sciences
Robert Östling, Stockholm University

Simon Petitjean, Heinrich-Heine Universität Düsseldorf
Dirk Pijpops, Université de Liège
Michael Regan, University of Colorado, Boulder
Mathilde Regnault, Universität Stuttgart
Laurence Romain, University of Birmingham
Rudolf Rosa, Charles University Prague
Haruko Sanada, Rissho University
Beatrice Santorini, University of Pennsylvania
Giorgio Satta, Università degli studi di Padova
Sebastian Schuster, Universität des Saarlandes
Olga Scrivner, Rose-Hulman Institute of Technology
Ashwini Vaidya, Indian Institute of Technology, Delhi
Remi van Trijp, Sony Computer Sciences Laboratories Paris
Giulia Venturi, Institute for Computational Linguistics "A. Zampolli" (ILC-CNR)
Nianwen Xue, Brandeis University
Eva Zehentner, University of Zurich
Amir Zeldes, Georgetown University
Daniel Zeman, Charles University Prague
Heike Zinsmeister, Universität Hamburg
Hongxin Zhang, Zhejiang University

Table of Contents