DepLing 2025

Eighth International Conference on Dependency Linguistics (Depling, SyntaxFest 2025)

Proceedings

August 27-28, 2025

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

VITASIS

USTANOVA PATRA STANISLAVA ŠKRABCA

UniDive

Ljubljana
Ljubljana Tourism

Mestna občina Ljubljana

Flanders
State of the Art

CLARIN.SI

cjvt
Centre for Language Resources and Technologies

AMEBIS

alpineon

AI4DH
CENTRE OF EXCELLENCE IN AI FOR DIGITAL HUMANITIES

Organized by

UNIVERSITY OF LJUBLJANA | FRI
Faculty of Computer and Information Science

UNIVERSITY OF LJUBLJANA | FF
Faculty of Arts

SDJT
Slovensko društvo za jezikovne tehnologije

As part of SyntaxFest 2025

SYNTAXFEST

Order copies of this and other ACL proceedings from:

Introduction

The Eighth edition of the International Conference on Dependency Linguistics (DepLing) follows a biannual series that started in 2011, in Barcelona and continued in Prague (2013), Uppsala (2015), Pisa (2017), Paris (2019), Sofia (2021), and Washington DC (2023). 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. DepLing (2025) took place at SyntaxFest 2025 in Ljubljana, Slovenia, which brought together five related but independent events:

- 18th International Conference on Parsing Technologies (IWPT 2025)

- 8th Universal Dependencies Workshop (UDW 2025)

- 8th International Conference on Dependency Linguistics (DepLing 2025)

- 23rd Workshop on Treebanks and Linguistic Theories (TLT 2025)

- 3rd Workshop on Quantitative Syntax (QUASY 2025)

In addition, a pre-conference workshop organized by the COST Action CA21167 – Universality, Diversity and Idiosyncrasy in Language Technology (UniDive) was held prior to the main event, with dedicated sessions on the 1st UniDive Shared Task on Morphosyntactic Parsing and the 2nd Workshop on Universal Dependencies for Turkic Languages.

SyntaxFest 2025 continues the tradition of SyntaxFest 2019 (Paris, France), SyntaxFest 2021 (Sofia, Bulgaria), and GURT/SyntaxFest 2023 (Washington DC, USA) 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 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. By co-locating these workshops under a shared umbrella, SyntaxFest fosters dialogue between overlapping research communities and supports innovation at the intersection of linguistics and language technology.

As in previous editions, all five workshops at SyntaxFest 2025 shared a common submission and reviewing process, with a unified timeline, identical submission formats, and a shared program committee. During submission, authors could indicate one or more preferred venues, but the final assignment of papers was determined by the collective program chairs, composed of the individual workshop chairs, based on thematic alignment. All accepted submissions were peer-reviewed by at least three reviewers from the shared program committee.

In total, SyntaxFest 2025 received 94 submissions, of which 73 (78%) were accepted for presentation. The final program included a total of 47 long papers, 21 short papers, and 5 non-archival contributions, distributed across the five workshops: 5 papers were presented at IWPT (2 long, 3 short); 20 at UDW (14 long, 5 short, 1 non-archival); 16 at DepLing (12 long, 2 short, 2 non-archival); 18 at TLT (10 long, 7 short, 1 non-archival); and 14 at QUASY (9 long, 4 short, 1 non-archival).

Our sincere thanks go to everyone who made this event possible. We thank all authors for their submissions and the reviewers for their time and thoughtful feedback, which contributed to a diverse and high-quality program. Special thanks go to the local organizing team at the University of Ljubljana and the Slovene Language Technologies Society for hosting the event, and to the sponsors for their generous support. Finally, we gratefully acknowledge ACL SIGPARSE for endorsing the event and the ACL Anthology for publishing the proceedings.

Kenji Sagae, Stephan Oepen (IWPT 2025 Chairs)
Gosse Bomma, Çağrı Çöltekin (UDW 2025 Chairs)
Eva Hajičová, Sylvain Kahane (DepLing 2025 Chairs)
Heike Zinsmeister, Sarah Jablotschkin, Sandra Kübler (TLT 2025 Chairs)

Xinying Chen, Yaqin Wang (QUASY 2025 Chairs)
Kaja Dobrovoljc (SyntaxFest 2025 Organization Chair)

Ljubljana, August 2025

Organizing Committee

TLT Chairs

Heike Zinsmeister, University of Hamburg
Sarah Jablotschkin, University of Hamburg
Sandra Kübler, Indiana University

DepLing Chairs

Eva Hajičová, Charles University, Prague
Sylvain Kahane, Université Paris Nanterre

UDW Chairs

Gosse Bomma, University of Groningen
Çağrı Çöltekin, University of Tübingen

IWPT Chairs

Kenji Sagae, University of California, Davis
Stephan Oepen, University of Oslo

QUASY Chairs

Xinying Chen, University of Ostrava
Yaqin Wang, Guangdong University of Foreign Studies

Publication Chair

Sarah Jablotschkin, University of Hamburg

Local SyntaxFest 2025 Organizing Committee

Kaja Dobrovoljc, University of Ljubljana, SDJT
Špela Arhar Holdt, University of Ljubljana
Luka Terčon, University of Ljubljana
Marko Robnik-Šikonja, University of Ljubljana
Matej Klemen, University of Ljubljana
Sara Kos, University of Ljubljana
Timotej Knez, University of Ljubljana, SDJT
Tinca Lukan, University of Ljubljana

Special Thanks for designing the SyntaxFest 2025 logo to

Kim Gerdes, Université Paris-Saclay

Program Committee

Shared Program Committee

V.S.D.S.Mahesh Akavarapu, Eberhard-Karls-Universität Tübingen
Leonel Figueiredo de Alencar, Federal University of Ceará (UFC)
Patricia Amaral, Indiana University
Giuseppe Attardi, University of Pisa
John Bauer, Stanford University
David Beck, University of Alberta
Laura Becker, Albert-Ludwigs-Universität Freiburg
Aleksandrs Berdicevskis, Gothenburg University
Ann Bies, University of Pennsylvania
Igor Boguslavsky, Universidad Politécnica de Madrid
Bernd Bohnet, Google
Cristina Bosco, University of Turin
Gosse Bouma, University of Groningen
Miriam Butt, Universität Konstanz
G. A. Celano, Universität Leipzig
Heng Chen, Guangdong University of Foreign Studies
Xinying Chen, University of Ostrava
Jinho D. Choi, Emory University
Çağrı Çöltekin, University of Tuebingen
Daniel Dakota, Leidos
Stefania Degaetano-Ortlieb, Universität des Saarlandes
Kaja Dobrovoljc, University of Ljubljana
Jakub Dotlacil, Utrecht University
Gülşen Eryiğit, Istanbul Technical University
Kilian Evang, Heinrich Heine University Düsseldorf
Pegah Faghiri, CNRS
Ramon Ferrer-i-Cancho, Universidad Politécnica de Cataluna
Marcos Garcia, Universidade de Santiago de Compostela
Kim Gerdes, Université Paris-Saclay
Loïc Grobol, Université Paris Nanterre
Bruno Guillaume, INRIA
Carlos Gómez-Rodríguez, Universidade da Coruña
Eva Hajicova, Charles University
Dag Trygve Truslew Haug, University of Oslo
Santiago Herrera, University of Paris Nanterre
Richard Hudson, University College London
Maarten Janssen, Charles University Prague
Jingyang Jiang, Zhejiang University
Mayank Jobanputra, Universität des Saarlandes
Sylvain Kahane, Université Paris Nanterre
Václava Kettnerová, Charles University Prague
Sandra Kübler, Indiana University
Guy Lapalme, University of Montreal
François Lareau, Université de Montréal
Miryam de Lhoneux, KU Leuven
Zoey Liu, University of Florida

Teresa Lynn, Dublin City University
Jan Macutek, Slovak Academy of Sciences
Robert Malouf, San Diego State University
Marie-Catherine de Marneffe, UCLouvain
Nicolas Mazziotta, Université de Liège
Alexander Mehler, Johann Wolfgang Goethe Universität Frankfurt am Main
Maitrey Mehta, University of Utah
Wolfgang Menzel, Universität Hamburg
Marie Mikulová, Charles University
Aleksandra Miletić, University of Helsinki
Jasmina Milićević, Dalhousie University
Simon Mille, Dublin City University
Yusuke Miyao, The University of Tokyo
Noor Abo Mokh, Indiana University
Simonetta Montemagni, Institute for Computational Linguistics "A. Zampolli" (ILC-CNR)
Jiří Mírovský, Charles University Prague
Kaili Müürisep, Institute of computer science, University of Tartu
Anna Nedoluzhko, Charles University Prague
Ruochen Niu, Beijing Language and Culture University
Joakim Nivre, Uppsala University
Stephan Oepen, University of Oslo
Timothy John Osborne, Zhejiang University
Petya Osenova, Sofia University "St. Kliment Ohridski"
Agnieszka Patejuk, Polish Academy of Sciences
Lucie Poláková, Charles University Prague
Prokopis Prokopidis, Athena Research Center
Mathilde Regnault, Universität Stuttgart
Kateřina Rysová, University of South Bohemia
Magdaléna Rysová, Charles University Prague
Tanja Samardzic, University of Zurich
Giuseppe Samo, Beijing Language and Culture University
Haruko Sanada, Rissho University
Nathan Schneider, Georgetown University
Djamé Seddah, Sorbonne University
Anastasia Shimorina, Orange
Maria Simi, University of Pisa
Achim Stein, University of Stuttgart
Daniel G. Swanson, Indiana University
Luka Terčon, Faculty of Arts, University of Ljubljana
Giulia Venturi, Institute for Computational Linguistics "A. Zampolli" (ILC-CNR)
Veronika Vincze, University of Szeged
Yaqin Wang, Guangdong University of Foreign Studies
Pan Xiaxing, Huaqiao University
Chunshan Xu, Anhui Jianzhu University
Nianwen Xue, Brandeis University
Jianwei Yan, Zhejiang University
Zdenek Zabokrtsky, Faculty of Mathematics and Physics, Charles University Prague
Eva Zehentner, University of Zurich
Amir Zeldes, Georgetown University
Daniel Zeman, Charles University Prague
Šárka Zikánová, Charles University Prague

Heike Zinsmeister, Universität Hamburg

DepLing Keynote
Auxiliaries across Languages and Frameworks

Daniel Zeman

Charles University, Prague

Abstract: In my talk, I will discuss the status of auxiliaries (i.e., auxiliary verbs as well as uninflected non-verbal particles with auxiliary function) in dependency treebanks. The focus will be on two frameworks, Universal Dependencies (UD) and the Prague family of treebanks, rooted in the Functional Generative Description. However, I will occasionally show examples from other treebanks and frameworks, encountered during the HamleDT harmonization effort.

Besides looking at various treatments of auxiliaries in different annotation schemes, I will also discuss the question of delimiting the set of auxiliaries in individual languages (or, more exactly, the set of words that receive the special treatment in the respective annotation schemes). Various grammatical tests may be available, but sometimes the auxiliaries are simply enumerated by traditional school grammar. Moreover, there is a scale of categories ranging from pure grammatical auxiliaries through modals and phase verbs to various semantically bleached verbs that take other verbs as complements, yet their contribution is lexical rather than grammatical and their syntactic behavior shows no anomalies. All these aspects complicate finding a unified definition that would be applicable in a multi-lingual dataset, such as HamleDT or UD.

In the last part of the talk, I will show some examples of contrastive cross-linguistic studies that would benefit from comparably defined auxiliaries. I will show how we encourage comparability in UD using a common database of auxiliaries, and I will argue that it has the potential to become a useful resource of its own.

Bio: Daniel Zeman is an associate professor of computational linguistics at the Charles University in Prague. He obtained his PhD (also from Charles University) in 2005 with a dissertation on statistical methods for syntactic parsing of Czech. He then worked on cross-lingual transfer techniques for low-resource languages, and led several projects focused on multilingual NLP and harmonization of linguistic resources, including Interset (for morphological tagsets) and HamleDT (for dependency treebanks). He is one of the founders and leading personalities of the Universal Dependencies initiative, and vice-chair of the COST Action "Universality, Diversity and Idiosyncrasy in Language Technology" (UniDive). His current work extends to harmonized datasets for coreference resolution (CorefUD) and Uniform Meaning Representation (UMR).

Local SyntaxFest Keynote
What we learn about syntax when dependencies fail: Experimental insights into syntactic locality constraints

Artur Stepanov
University of Nova Gorica

Abstract: This talk examines a class of syntactic dependencies that cannot be formed: classic island violations (extraction from adjuncts, complex NPs, wh-islands etc.). I survey psycho- and neurolinguistic evidence quantifying the cognitive cost of breaching locality constraints, showing how these findings expose limits on dependency formation that remain invisible in standard treebanks yet are central to real-time sentence processing. I consider implications for parsing, dependency representations, and cross-linguistic variation, with suggestions for incorporating experimental diagnostics into syntactic annotation and parser-evaluation frameworks.

Bio: Artur Stepanov is a professor of psycholinguistics at the University of Nova Gorica. His work focuses on the cognitive representation and real-time processing of syntactic dependencies in monolingual and multilingual speakers, exploring how internal grammatical competence maps onto observable linguistic behavior. He combines psycholinguistic experimentation with insights from generative syntax, with particular emphasis on lesser-studied Slavic languages. He is involved in multiple international collaborations on projects related to sentence comprehension and production, the linguistic and cognitive dimensions of multilingualism, and, more recently, the compositional aspects of animal (marine mammal) vocalization sequences.

Non-Archival Abstract

Dependency Analysis of Chinese Comparative Sentences
Zexin Liu
Zhejiang University

This paper examines the dependency structures of comparative sentences across various Chinese dialects. In equality comparatives, the comparative result is post-posed (R-back) in all Chinese dialects, which contrasts with English. Although Mandarin also follows the R-back pattern for superiority comparatives, dialects such as Hong Kong Cantonese and Southern Min adopt an R-front type, similar to English. However, Southern Min lacks a comparative marker, while English's comparative marker than dominates the standard of comparison. In contrast, the comparative marker in Cantonese does not dominate the standard. Through the calculation of dependency distances and syn-tactic tests, we argue that when the comparative result is preposed, it dominates the standard of comparison. Conversely, when the comparative construction fol-lows an R-back type, the comparative marker dominates the comparative result. This analysis aligns closely with the annotation choices of the Surface-Syntactic Universal Dependencies (SUD), differing significantly from those of the Universal Dependencies (UD) model.

Non-Archival Abstract

A Quantitative Study of Subject-Predicate-Object Word Class Composition in vernacular Chinese Based on Dependency Grammar

Bingli Liu[1] and Yiyi Zhao[2]

[1]Huaqiao University Quanzhuo

[2]Xiamen University

The paper aims at studying the evolution of lexical composition of subject-verb-object sentences in vernacular Chinese. Five corpora are constructed for the Tang and Five Dynasties, Song Dynasty, Yuan and Ming Dynasties, Qing Dynasty, and the present contemporary era witch lasts for more than 1,000 years. The syntactic structures of these sentences are labeled, counted, and analyzed based on the theoretical foundation of dependency grammar, with the aim of investigating the evolution of the lexical category composition of the subject-predicate-object in vernacular Chinese over time. The results show that the ratio of nouns and pronouns in each period occupies the majority of the total number of subject lexemes, and the lexical composition of predicates has been very stable since ancient times, with verbal predicates accounting for the vast majority of predicates. Compared with the subject lexical composition, objects are richer and the lexical composition changes more slowly.

Table of Contents