

Quasy 2025

**Third Workshop on Quantitative Syntax (QUASY,
SyntaxFest 2025)**

Proceedings

August 29, 2025

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

VITASIS



Ljubljana Tourism



Mestna občina
Ljubljana



Flanders
State of the Art



CJVT Centre for
Language Resources
and Technologies



AI4DH CENTRE OF EXCELLENCE IN AI
FOR DIGITAL HUMANITIES

Organized by



As part of SyntaxFest 2025



©2025 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-293-0

Introduction

Quantitative Syntax (QUASY) is a workshop series dedicated to advancing quantitative, statistical, and computational methods in syntactic research. The workshops bring together researchers to discuss and explore quantitative, statistical, and computational methods in syntax research, responding to the growing need for linguistic meetings that focus on empirical data-driven approaches to syntactic theory. Since the first QUASY workshop in 2019 in Paris, the series has brought together researchers working at the intersection of syntax, corpus linguistics, computational methods, and related empirical approaches. This year’s workshop, the third QUASY 2025, is held as part of SyntaxFest 2025 in Ljubljana, Slovenia, which brings 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 Bouma, 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 Bouma, 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 Catalunya
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
Ruo Chen 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 Samardžić, 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
Zdeněk Zabokrtský, 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

Keynote

The rhetorical and pragmatic functions of syntactically complex structures in academic and second language writing

Xiaofei Lu

The Pennsylvania State University



Abstract: Previous studies of linguistic complexity in academic and second language (L2) writing has often focused on quantitative differences across different writer groups and/or longitudinal changes over time, without systematic attention to the rhetorical or pragmatic functions that complex forms are used to convey. This talk argues for the importance of and delineates the scope of the function dimension of linguistic complexity analysis in L2 writing research, reviews the methods and findings of emerging efforts on this dimension, and discusses how future L2 writing research could attend to this dimension.

Bio: Xiaofei Lu is the George C. and Jane G. Greer Professor of Applied Linguistics and Asian Studies at The Pennsylvania State University. His research has long centered on computational and quantitative analyses of linguistic complexity in reading materials, second language production, and academic writing. His current work explores mappings between linguistic forms and rhetorical/pragmatic functions in language production and sense-aware measurements of linguistic complexity that account for the specific meanings of polysemous linguistic forms in context. He has published over 90 peer-reviewed articles in leading journals, including *Applied Linguistics*, *Behavior Research Methods*, *Computer Assisted Language Learning*, *Language Learning*, *Studies in Second Language Acquisition*, *TESOL Quarterly*, and *The Modern Language Journal*. He received the 2023 Ken Hyland Best Paper Award from the *Journal of English for Academic Purposes*. His latest book, *Corpus Linguistics and Second Language Acquisition: Perspective, Issues, and Findings*, was published by Routledge in 2023.

Non-Archival Abstract

Syntactic Complexity and News Credibility in Czech Media

Miroslav Kubát, Xinying Chen, Michaela Nogolová and Michal Místecký
University of Ostrava

This study examines how syntactic complexity varies across news articles differing in credibility, using a Czech-language corpus annotated with five credibility levels: credible, partially credible, misleading, manipulative, and unclassifiable. We apply a dependency parsing pipeline and compute five syntactic metrics measuring features such as sentence length, clause density, and hierarchical depth. Results show that manipulative texts are structurally the most complex, while misleading and unclassifiable texts are simpler and more fragmented. Credible texts display balanced complexity consistent with journalistic norms. These findings highlight the role of syntax in shaping rhetorical strategies and contribute to the linguistic understanding of news credibility.

Table of Contents

<i>Subject-Verb Agreement Alternations in Spanish Pseudopartitive Constructions: A Corpus Study</i> Marina Cerebrinsky	1
<i>Degree centrality as a measure of robustness of dependency structures of the sentences in a large-scale learner corpus of English</i> Masanori Oya	9
<i>Application of Existing Readability Methods to the Ukrainian Language: A Comprehensive Study</i> Serhii D. Prykhodchenko and Oksana Yu. Prykhodchenko	17
<i>Extraction of Contrastive Rules from Syntactic Treebanks: A Case Study in Romance Languages</i> Santiago Herrera, Ioana-Madalina Silai, Bruno Guillaume and Sylvain Kahane	26
<i>A Quantitative Study of Syntactic Complexity across Genres: Dependency Distance in English and Chinese</i> Yaqin Wang	39
<i>Syntactic Complexity in L2 Reading: A Comparison of Adapted and Original Czech Texts</i> Žaneta Stiborská, Michaela Nogolová, Xinying Chen and Miroslav Kubát	47
<i>Modeling the Law of Abbreviation in Classical, Modern, and ChatGPT-Generated Chinese: A Power-Law Analysis of Structural Economy</i> Jianwei Yan and Heng Chen	56
<i>A Computational Method for Analyzing Syntactic Profiles: The Case of the ELEXIS-WSD Parallel Sense-Annotated Corpus</i> Jaka Čibej	63
<i>The Interplay of Noun Phrase Complexity and Modification Type in Scientific Writing</i> Isabell Landwehr	72
<i>Predictability Effects of Spanish-English Code-Switching: A Directionality and Part of Speech Analysis</i> Josh Higdon, Valeria Pagliai and Zoey Liu	83
<i>On the Flatness, Non-linearity, and Branching Direction of Natural Language and Random Constituency Trees: Analyzing Structural Variation within and across Languages</i> Taiga Ishii and Yusuke Miyao	90
<i>First Insights into the Syntax of Slovene Student Writing: A Statistical Analysis of Šolar 3.0 vs. Učbeniki 1.0</i> Tina Munda and Špela Arhar Holdt	105
<i>Syntactic units and their length distributions: A case study in Czech</i> Michaela Nogolová, Michaela Koščová, Jan Macutek and Radek Cech	115
<i>Do Multilingual Transformers Encode Paninian Grammatical Relations? A Layer-wise Probing Study</i> Akshit Kumar, Dipti Sharma and Parameswari Krishnamurthy	124