

1st UniDive Training School - accepted projects

* Back to the UniDive 1st UniDive training school [training page](#)

Projects

1. Construction of a UD Treebank for Mukri Kurdish and Modification of Existing UD for Sorani and Kurmanji Kurdish and Persian with Emphasis on Multi-Word Expressions - [Hiwa Asadpour](#) (abstract)
-)
2. The annotation of the Pyramid Texts in the Egyptian-Ujaen treebank - [Roberto A. Diaz Hernandez](#) (abstract)
-)
3. Towards the first UD Treebank of Spoken Italian: the KIParla forest - [Ludovica Pannitto](#) (abstract)
-)
4. Integrating Geographically Diverse Low-Resourced Spanish Varieties into Universal Dependencies - **[Johnatan E. Bonilla](#), [Laura M. Merino Hernández](#), [Lucía Amorós-Poveda](#)** (abstract)
-)
5. Integrating PARSEME-MWEs into the Spoken Slovene SST UD Corpus - [Tina Munda](#) (abstract)
-) - **Not eligible for reimbursement**
6. Annotation of Multiword Expressions in Lithuanian Corpus of User-Generated Comments - [Giedrė Valūnaitė-Oleškevičienė](#) (abstract)
-)
7. Enhancing and improving accuracy of the Ukrainian UD treebank using resources of the Ukrainian Brown Corpus - [Kyrylo Popik](#) (abstract)
-)
8. UD annotation tools and best practices for the creation of morpho-syntactic treebanks for low-resource languages - [Omer Goldman](#) (abstract)
-)
9. Developing accurate dependency parsing tool for Galician as low-resource language - [Albina Sarymsakova](#) , [Xulia Sánchez-Rodríguez](#), [Laura Castro](#) (abstract)
-)
10. Measuring and Mitigating Bias for Maltese as a Low-Resource Language - [Melanie Galea](#) (abstract)
-)

11. Annotating the Divine Comedy in Universal Dependency: A Journey through Inferno, Purgatorio and Paradiso. How to (...) ellipses, and other annotation issues - [Claudia Corbetta](#) (
abstract
)
12. Universal Dependencies for Uzbek - [Arofat Akhundjanova](#) (
abstract
)
13. UD and POS tagging in Uyghur - [Faruk Mardan](#) (
abstract
)
14. Beyond Standardization: Crafting a Comprehensive Annotated Corpus for Corsican and Poitevin-Saintongeais - [Cristina García Holgado](#) (
abstract
)
15. Creating a UD treebank for a low-resourced Modern Greek dialect: the case of Lesbian - [Stavros Bompolas](#) (
abstract
)
16. Expanding the Ukrainian UD treebank and exploring the possibilities of using it for annotation of large corpora and in teaching- [Maria Shvedova](#) (
abstract
)
17. Development of an Enhanced Universal Dependencies Treebank and Multiword Expression Annotation for Standard Albanian Language - [Anila Çepani, Adelina Çerpja, Nelda Kote, Alba Haveriku](#) (
abstract
)
18. SweLL-UD: A Treebank of L2 Swedish Essays - [Arianna Masciolini](#) (
abstract
)
19. An Annotated Corpus for Buranello - [Chiara Di Maio](#) (
abstract
) - **Not eligible for reimbursement**
20. Expansion of UCxN: Typologically Informed Annotation of Constructions Atop Universal~Dependencies - [Leonie Weissweiler](#) (
abstract
)
21. A Collaborative Grammatical Annotation Tool, BoAT - [Furkan Akkurt](#) (
abstract
)
22. Syntactic analysis and syntactic patterns of MWEs - [Agute Klints](#) (
abstract
)
23. Expanding the Macedonian Language Tree Bank - [Emilija Mitovska](#) (
abstract
) - **Not eligible for reimbursement**
24. Kyrgyz Universal Dependency Treebank - [Bermet Chontaeva](#) (
abstract
)
25. Azerbaijani-TueCL: Universal Dependency Treebank for Azerbaijani Language - [Soudabeh Eslami](#) (
)

abstract

)

26. Bantu PARSEME Corpora: Annotation for Swahili and Bapuku MWEs - [Matt Malone](#) (

abstract

) - **Not eligible for reimbursement**

27. Creation of a Multi-Word Expression Dataset in Galician for Language Model Evaluation - [Laura Castro](#), Xulia Sánchez-Rodríguez, Albina Sarymsakova (

abstract

)

28. Misinformation detection for South African Low Resource Language - **Seani Rananga** (

abstract

)

29. Starting a UD Treebank of Tundra Nenets - [Nikolett Mus](#) (

abstract

)

30. Annotation of Multiword Expressions in the Lithuanian Corpus of the Cyber-Security Domain - [Liudmila Mockienė](#) (

abstract

)

31. Correcting and Enlarging the Slovak UD Treebank - [Vladimír Benko](#) (

abstract

)

32. Experimental UD Treebank for Mansi - [Csilla Horváth](#) (

abstract

)

33. Developing a Potentially Idiomatic Expressions (PIE) Corpus for Turkish - [Murathan Kurfali](#) (

abstract

)

34. Advancing Universal Dependencies: Towards Enhanced Syntactic Annotation in the Albanian Treebank- [Enriketa Sogutlu](#) (

abstract

)

35. Developing a Lithuanian Textual Corpus for Identifying Propaganda and Detecting Propaganda Techniques - [Ieva Rizgelienė](#) (

abstract

)

36. Improving HDT-UD - [Nina Böbel](#) (

abstract

)

37. Annotation of Multiword Expressions for Hate Speech Identification in the Domain of Sports over Social Media - [Panagiotis Grigoriadis](#) (

abstract

)

38. Tautological Predicative Constructions with Lexical Repetitions (TPCLRs) in the Ukrainian Language - [Kristina Husenko](#) (

abstract

) - **Not eligible for reimbursement**

39. A New UD Treebank for Thai - [Siriluck Rattananiyomkul](#) (

abstract

)

40. Tetun Universal Dependencies Treebank - [Gabriel de Jesus](#) (
abstract
)
41. Towards Building Recursive and Complete Discourse Structures of UD/PARSEME corpus - [Iskandar keskes, Ines boujelben](#) (
abstract
)
42. Enhancing Croatian idiom recognition through contextual analysis - [Nika Marinović](#) (
abstract
) - **Not eligible for reimbursement**
43. Automatic identification of Potential Multi-word Expressions using Naturalness of scenes - [Nishan Chatterjee](#) (
abstract
)
44. Creating an annotated corpus for Puno Quechua language - [Elwin Huaman](#) (
abstract
)
45. A proposal for the accurate syntactic annotation of Galician - [Xulia Sánchez-Rodríguez](#), Albina Sarymsakova, Laura Castro (
abstract
)
46. Larger Treebank for Amharic Annotated with Semi-automatic Method - [Abnet Shimeles Ibssa](#) (
abstract
) - **Not eligible for reimbursement**
47. UD_Gwichin-TueCL: Annotating an endangered polysynthetic Athabascan language in UD - [Matthew Kirk Andrews](#) (
abstract
)
48. Towards universal annotation guidelines for coreference resolution and information status - [Andrew Thomas Dyer](#) (
abstract
)
49. Advancing Language Technology for Jola - [Papa Ousmane Diedhiou](#) (
abstract
)
50. Exploring the possibility for a multiword expression corpus for Southern Nambikwara - [Emilia Roosvall](#) (
abstract
)
51. Development of an Amharic Dependency Treebank: Annotated Corpus Construction and Linguistic Annotation - [Dawit TJ](#) (
abstract
)
52. Italian L1 learner treebank - [Marta Sartor](#) (
abstract
)

From:

<https://unidive.lisn.upsaclay.fr/> - **Universality, diversity and idiosyncrasy
in language technology
CA21167 COST Action**

Permanent link:

https://unidive.lisn.upsaclay.fr/doku.php?id=accepted_projects&rev=1718984107

Last update: **2024/06/21 17:35**

