

1st UniDive Training School - accepted projects

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

Projects

A

- Larger Treebank for Amharic Annotated with Semi-automatic Method - [Abnet Shimeles Ibssa](#) (abstract)
-) - **Not eligible for reimbursement**
- Syntactic analysis and syntactic patterns of MWEs - [Agute Klints](#) (abstract)
-)
- Developing accurate dependency parsing tool for Galician as low-resource language - [Albina Sarymsakova](#) , Xulia Sánchez-Rodríguez, Laura Castro (abstract)
-) - canceled
- Towards universal annotation guidelines for coreference resolution and information status - [Andrew Thomas Dyer](#) (abstract
Poster)
-)
- Development of an Enhanced Universal Dependencies Treebank and Multiword Expression Annotation for Standard Albanian Language - [Anila Çepani, Adelina Çerpja, Nelda Kote, Alba Haveriku](#) (abstract
Poster)
-)
- SwELL-UD: A Treebank of L2 Swedish Essays - [Arianna Masciolini](#) (abstract
Poster)
-)
- Universal Dependencies for Uzbek - [Arofat Akhundjanova](#) (abstract)
-)

B

- Kyrgyz Universal Dependency Treebank - [Bermet Chontaeva](#) (abstract)
-)

C

- An Annotated Corpus for Buranello - [Chiara Di Maio](#) (abstract)
) - **Not eligible for reimbursement**
- Annotating the Divine Comedy in Universal Dependency: A Journey through Inferno, Purgatorio and Paradiso. How to (...) ellipses, and other annotation issues - [Claudia Corbetta](#) (abstract)
)
- Beyond Standardization: Crafting a Comprehensive Annotated Corpus for Corsican and Poitevin-Saintongeais - [Cristina García Holgado](#) (abstract)
)
- Experimental UD Treebank for Mansi - [Csilla Horváth](#) (abstract)
)

D

- Development of an Amharic Dependency Treebank: Annotated Corpus Construction and Linguistic Annotation - [Dawit TJ](#) (abstract)
)

E

- Creating an annotated corpus for Puno Quechua language - [Elwin Huaman](#) (abstract)
)
- Expanding the Macedonian Language Tree Bank - [Emilija Mitovska](#) (abstract)
) - **Not eligible for reimbursement**
- Exploring the possibility for a multiword expression corpus for Southern Nambikwara - [Emilia Roosvall](#) (abstract)
)
- Advancing Universal Dependencies: Towards Enhanced Syntactic Annotation in the Albanian Treebank- [Enriketa Sogutlu](#) (abstract)
)

F

- UD and POS tagging in Uyghur - [Faruk Mardan](#) (abstract)
)
- A Collaborative Grammatical Annotation Tool, BoAT - [Furkan Akkurt](#) (abstract
Poster)
)

G

- Tetun Universal Dependencies Treebank - [Gabriel de Jesus](#) (

abstract

)

- Measuring and Mitigating Bias for Maltese as a Low-Resource Language - Galea Melanie (abstract

) - canceled

- Annotation of Multiword Expressions in Lithuanian Corpus of User-Generated Comments - Giedrė Valūnaitė-Oleškevičienė (

abstract

)

H

- Construction of a UD Treebank for Mukri Kurdish and Modification of Existing UDs for Sorani and Kurmanji Kurdish and Persian with Emphasis on Multi-Word Expressions - Hiwa Asadpour (abstract

)

I

- Developing a Lithuanian Textual Corpus for Identifying Propaganda and Detecting Propaganda Techniques - Ieva Rizgelienė (

abstract
Poster

)

- Towards Building Recursive and Complete Discourse Structures of UD/PARSEME corpus - Iskandar keskes, Ines boujelben (

abstract
Poster

)

J

- Integrating Geographically Diverse Low-Resourced Spanish Varieties into Universal Dependencies - Johnatan E. Bonilla, Laura M. Merino Hernández, Lucía Amorós-Poveda (

abstract
Poster

)

K

- Tautological Predicative Constructions with Lexical Repetitions (TPCLRs) in the Ukrainian Language - Kristina Husenko (

abstract

) - **Not eligible for reimbursement**

- Enhancing and improving accuracy of the Ukrainian UD treebank using resources of the Ukrainian Brown Corpus - Kyrylo Popik (

abstract

)

L

- Creation of a Multi-Word Expression Dataset in Galician for Language Model Evaluation - [Laura Castro](#), Xulia Sánchez-Rodríguez, Albina Sarymsakova (abstract)
-)
- Expansion of UCxN: Typologically Informed Annotation of Constructions Atop Universal~Dependencies - [Leonie Weissweiler](#) (abstract)
-)
- Annotation of Multiword Expressions in the Lithuanian Corpus of the Cyber-Security Domain - [Liudmila Mockienė](#) (abstract)
-)
- Towards the first UD Treebank of Spoken Italian: the KIParla forest - [Ludovica Pannitto](#) (abstract
Poster)
-)

M

- Expanding the Ukrainian UD treebank and exploring the possibilities of using it for annotation of large corpora and in teaching- [Maria Shvedova](#) (abstract)
-)
- Italian L1 learner treebank - [Marta Sartor](#) (abstract
Poster)
-)
- UD_Gwichin-TueCL: Annotating an endangered polysynthetic Athabaskan language in UD - [Matthew Kirk Andrews](#) (abstract
Poster)
-)
- Bantu PARSEME Corpora: Annotation for Swahili and Bapuku MWEs - [Matthew Malone](#) (abstract)
-) - **Not eligible for reimbursement**
- Developing a Potentially Idiomatic Expressions (PIE) Corpus for Turkish - [Murathan Kurfali](#) (abstract)
-) - canceled

N

- Enhancing Croatian idiom recognition through contextual analysis - [Nika Marinović](#) (abstract)
-) - **Not eligible for reimbursement**
- Starting a UD Treebank of Tundra Nenets - [Nikolett Mus](#) (abstract
Poster)
-)
- Improving HDT-UD - [Nina Böbel](#) (abstract)
-)
- Automatic identification of Potential Multi-word Expressions using Naturalness of scenes -

Nishan Chatterjee (

abstract

)

O

- UD annotation tools and best practices for the creation of morpho-syntactic treebanks for low-resource languages - Omer Goldman (

abstract

) canceled

P

- Annotation of Multiword Expressions for Hate Speech Identification in the Domain of Sports over Social Media - Panagiotis Grigoriadis (

abstract

Poster

)

- Advancing Language Technology for Jola - Papa Ousmane Diedhiou (

abstract

)

R

- The annotation of the Pyramid Texts in the Egyptian-Ujaen treebank - Roberto A. Diaz Hernandez (

abstract

Poster

)

S

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

abstract

)

- A New UD Treebank for Thai - Siriluck Rattananiyomkul (

abstract

Poster

)

- Azerbaijani-TueCL: Universal Dependency Treebank for Azerbaijani Language - Soudabeh Eslami (

abstract

)

- Creating a UD treebank for a low-resourced Modern Greek dialect: the case of Lesbian - Stavros Bompolas (

abstract

)

T

- Integrating PARSEME-MWEs into the Spoken Slovene SST UD Corpus - Tina Munda (

abstract

) - **Not eligible for reimbursement**

V

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

X

- A proposal for the accurate syntactic annotation of Galician - [Xulia Sánchez-Rodríguez](#), Albina Sarymsakova, Laura Castro (
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

Last update: **2024/07/07 20:50**

