

1st UniDive Training School - accepted projects

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

Projects

A

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

B

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

C - [An Annotated Corpus for Buranello](#) - [Chiara Di Maio](#) (

abstract

) - **Not eligible for reimbursement**

1. Annotating the Divine Comedy in Universal Dependency: A Journey through Inferno, Purgatorio and Paradiso. How to (...) ellipses, and other annotation issues - [Claudia Corbetta](#) (abstract)
)
2. Beyond Standardization: Crafting a Comprehensive Annotated Corpus for Corsican and Poitevin-Saintongeais - [Cristina García Holgado](#) (abstract)
)
3. Experimental UD Treebank for Mansi - [Csilla Horváth](#) (abstract)
)

D

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

E

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

F

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

G

1. Tetun Universal Dependencies Treebank - [Gabriel de Jesus](#) (abstract)
)
2. Measuring and Mitigating Bias for Maltese as a Low-Resource Language - [Galea Melanie](#) (abstract)

) - canceled

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

abstract

)

H

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

)

I

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

abstract

)

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

abstract

)

J

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

abstract

)

K

1. 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

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

abstract

)

2. Expansion of UCxn: Typologically Informed Annotation of Constructions Atop Universal~Dependencies - Leonie Weissweiler (

abstract

)

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

abstract

)

4. Towards the first UD Treebank of Spoken Italian: the KIParla forest - [Ludovica Pannitto](#) (

abstract

)

M

1. Expanding the Ukrainian UD treebank and exploring the possibilities of using it for annotation of large corpora and in teaching- [Maria Shvedova](#) (

abstract

)

2. Italian L1 learner treebank - [Marta Sartor](#) (

abstract

)

3. UD_Gwichin-TueCL: Annotating an endangered polysynthetic Athabascan language in UD - [Matthew Kirk Andrews](#) (

abstract

)

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

abstract

) - **Not eligible for reimbursement**

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

abstract

) - canceled

N

1. Enhancing Croatian idiom recognition through contextual analysis - [Nika Marinović](#) (

abstract

) - **Not eligible for reimbursement**

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

abstract

)

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

abstract

)

4. 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

) P

1. Annotation of Multiword Expressions for Hate Speech Identification in the Domain of Sports over

Social Media - [Panagiotis Grigoriadis](#) (

abstract

)

2. Advancing Language Technology for Jola - [Papa Ousmane Diedhiou](#) (

abstract

)

[Q R - The annotation of the Pyramid Texts in the Egyptian-Ujaen treebank](#) - [Roberto A. Diaz Hernandez](#) (

abstract

) S

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

abstract

)

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

abstract

)

3. Azerbaijani-TueCL: Universal Dependency Treebank for Azerbaijani Language - [Soudabeh Eslami](#)

(

abstract

)

4. Creating a UD treebank for a low-resourced Modern Greek dialect: the case of Lesbian - [Stavros Bompolas](#) (

abstract

)

T

1. Integrating PARSEME-MWEs into the Spoken Slovene SST UD Corpus - [Tina Munda](#) (

abstract

) - **Not eligible for reimbursement**

U V

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

abstract

)

W X

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

abstract

)

Y Z

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=1720253228

Last update: **2024/07/06 10:07**

