# Standardizing Annotations in Turkic Languages

**Objective:** Standardize annotations in Universal Dependencies treebanks of Turkic languages.

**Introduction:**

- Growing number of treebanks in UD.
- Importance of annotation consistency.

**UD Turkic Workshop 2023 (UDTW23):**

- September 2023, Istanbul.
- Discussions on issues and examples.

**Issues Discussed:**

- Tokenization, tense/mood tags, oblique/object distinction.
- Question particle, code switching, transcription, pronominalized nouns.

**Conclusions:**

- Workshop as a catalyst for unified annotation approach.
- Future steps: comprehensive paper on issues and decisions.

**References and Acknowledgements:** See poster for details.

Languages in UD: Kazakh, Kyrgyz, Tatar, Turkish (9), Uyghur, Yakut, Old Turkish, and Turkish-German

Coming to UD: Uzbek, Ottoman Turkish, Turkish (1)

# Annotation of MWEs and NEs in the Serbian extension of ELEXIS-WSD: comparisons, solutions and open questions

WG1&2

cvetana@jerteh.rs, ranka@rgf.rs, aleksandra.markovic@isj.sanu.ac.rs

## 1. The extension of ELEXIS-WSD

- automatically translated SS, checked, proofread; automatically tokenized, lemmatized, POS-tagged, manually corrected.

**To do:**

- annotation (MWEs, NEs & syntactic)
- linking with the sense repository.

- **2. All MWEs & NEs from the WSD**

automatically translated into SR.

- *lingua franca* (6 lang. sets)
- 'Greece' the most freequent NE

## 3. The comparison of MWEs & NEs accross languages

- automatic translation of MWEs sometimes imprecise: *издигам се* (BG) *nastati* 'become' (**ustati* (SR) 'get up')
- sometimes the translation of MWEs was good, but not annotated in SSS (društvena mreža 'social network')
- the lack of annotation is not unusual in other sets: *heavy water* (EN) & *межка вода* (BG) weren't annotated neither.

# Towards a Dutch Parseme Corpus (WG1)

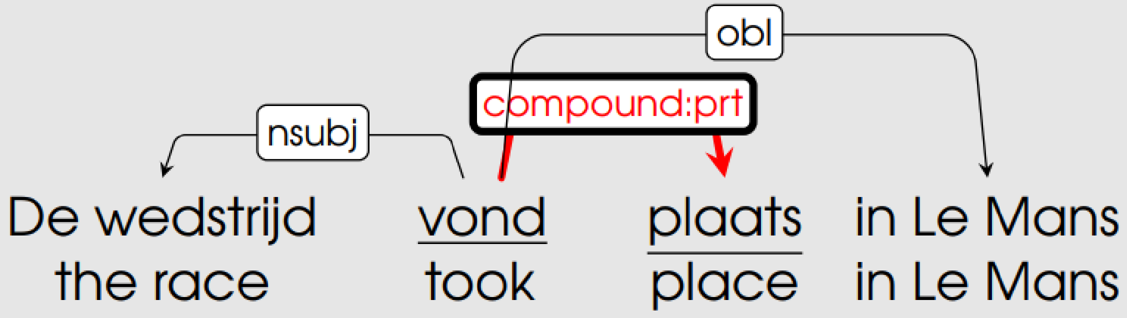
Gosse Bouma
University of Groningen

Jan Odijk
Utrecht University

Carole Tiberius
Dutch Language Institute

## Automatic Conversion

### Light Verb Constructions



De wedstrijd vond plaats in Le Mans
the race took place in Le Mans

## Results

| Class | Alpino | LassySmall |
|---|---|---|
| VPC | 2937 | 1025 |
| IAV | 1372 | 570 |
| VID | 1347 | 419 |
| LVC | 354 | 98 |
| IRV | 188 | 90 |
| MVC | 41 | 4 |
| Total | 6239 | 2206 |

# An English-Bulgarian Comparable Corpus Annotated with FrameNet Valence Patterns

UniDive | COST European Cooperation in Science & Technology | Funded by the European Union

WG2 **Ivelina Stoyanova** | iva@dcl.bas.bg | https://dcl.bas.bg/
Department of Computational Linguistics, Institute for Bulgarian Language, Bulgarian Academy of Sciences

2$^{nd}$ General Meeting
University of Naples "L'Orientale", Naples, Italy
8-9 February 2024

## Objectives

❖ **Practical:**

- ➢ To build a bilingual corpus that demonstrates the syntactic realisation of the conceptual description of verbs in English and Bulgarian.
- ➢ To combine information from various resources for the extensive semantic and syntactic description of verbs.

❖ **Theoretical:**

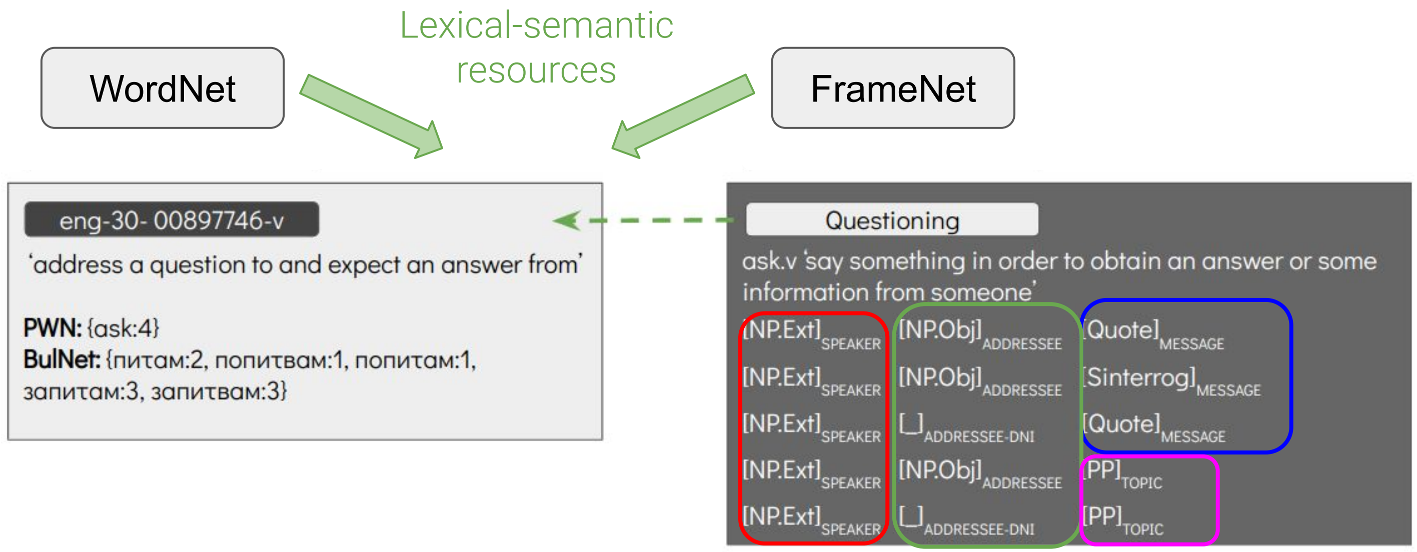
- ➢ To study universality and the possible cross-language linking and transfer of information (English- Bulgarian).

For English:

- ❖ 13,295 annotated examples
- ❖ 3,577 different valence patterns

For Bulgarian (work in progress):

- ❖ 2,050 annotated examples
- ❖ 272 different valence patterns



eng-30- 00897746-v

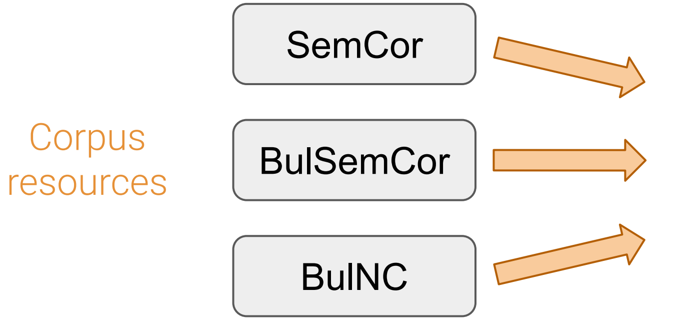
'address a question to and expect an answer from'

**PWN:** {ask:4}
**BulNet:** {питам:2, попитвам:1, попитам:1, запитам:3, запитвам:3}

Questioning

ask.v 'say something in order to obtain an answer or some information from someone'

| | | |
|---|---|---|
| [NP.Ext]$_{\text{SPEAKER}}$ | [NP.Obj]$_{\text{ADDRESSEE}}$ | [Quote]$_{\text{MESSAGE}}$ |
| [NP.Ext]$_{\text{SPEAKER}}$ | [NP.Obj]$_{\text{ADDRESSEE}}$ | [Sinterrog]$_{\text{MESSAGE}}$ |
| [NP.Ext]$_{\text{SPEAKER}}$ | [_]$_{\text{ADDRESSEE-DNI}}$ | [Quote]$_{\text{MESSAGE}}$ |
| [NP.Ext]$_{\text{SPEAKER}}$ | [NP.Obj]$_{\text{ADDRESSEE}}$ | [PP]$_{\text{TOPIC}}$ |
| [NP.Ext]$_{\text{SPEAKER}}$ | [_]$_{\text{ADDRESSEE-DNI}}$ | [PP]$_{\text{TOPIC}}$ |



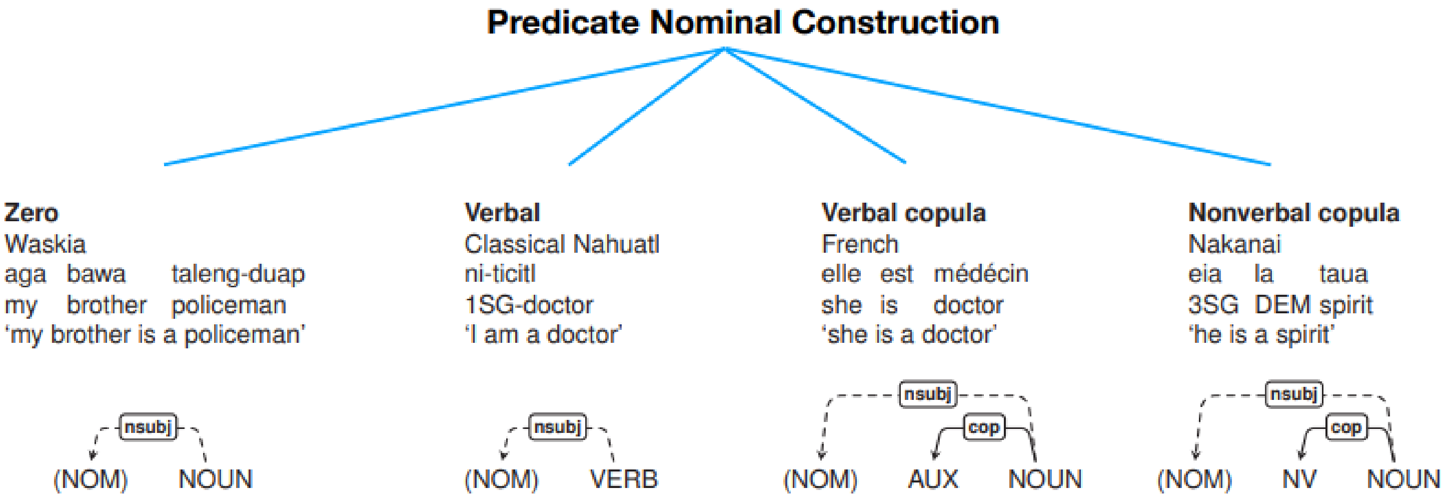
| | | | |
|---|---|---|---|
| [ Престъпникът ]$_{\text{SPEAKER:NP.Ext}}$ | **попита** | [ полицая ]$_{\text{ADDRESSEE:NP.Obj}}$ | [ дали може да си купи цигари ]$_{\text{MESSAGE.Clause}}$ |
| [ Criminal-DEF ]$_{\text{SPEAKER:NP.Ext}}$ | **asked** | [ policeman-DEF ]$_{\text{ADDRESSEE:NP.Obj}}$ | [ whether can to himself buy cigarettes ]$_{\text{MESSAGE.Clause}}$ |
| [ The criminal ]$_{\text{SPEAKER:NP.Ext}}$ | **asked** | [ the policeman ]$_{\text{ADDRESSEE:NP.Obj}}$ | [ whether he could buy some cigarettes. ]$_{\text{MESSAGE.Clause}}$ |

UPPSALA UNIVERSITET

# A Constructicon for Universal Dependencies

Joakim Nivre

RI. SE



**Predicate Nominal Construction**

**Zero**
Waskia
aga bawa taleng-duap
my brother policeman
‘my brother is a policeman’

**Verbal**
Classical Nahuatl
ni-ticitl
1SG-doctor
‘I am a doctor’

**Verbal copula**
French
elle est médécin
she is doctor
‘she is a doctor’

**Nonverbal copula**
Nakanai
eia la taua
3SG DEM spirit
‘he is a spirit’

UniDive

UniDive

2nd General Meeting

University of Naples L'Orientale, Naples, Italy, 8-9 February 2024

https://unidive.lisn.upsaclay.fr/

WG1, WG3

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Funded by
the European Union

# Universal Dependencies Treebank for Standard Albanian

Nelda Kote[1], Anila Çepani Sema[2], Alba Haveriku[1]

[1]Polytechnic University of Tirana, Tirana, Albania

[2]University of Tirana, Tirana, Albania

## CONTRIBUTION

- A UD treebank for the Standard Albanian language, created in collaboration between linguistics and information technology experts.
- 25,000 tokens, 1,300 sentences.

## ANNOTATED CORPORA

- Sentence segmentation;
- Words segmentation within a sentence;
- Lemmatization;
- Part-of-speech tags;
- Morphological features;
- Syntactic annotation.

nkote@fti.edu.al; anila.cepani@unitir.edu.al; alba.haveriku@fti.edu.al

*This project was funded by the National Agency for Scientific Research and Innovation as part of the National Research and Development Programs.*

# Creativity, productivity and diversity:

## The case of Hebrew possessive constructions

Ittamar Erb & Nurit Melnik
The Open University of Israel

### Goals:

Predict construction *extensibility*:

i. model construction *diversity* as attested in linguistic corpora.

ii. predict speakers' evaluation of *coinages*: unattested combinations of constructions with lexical items.

### Two competing constructions:

(1) ha-ʃem **ʃel-i**
the-name **of-POSS.1S**
'my name'

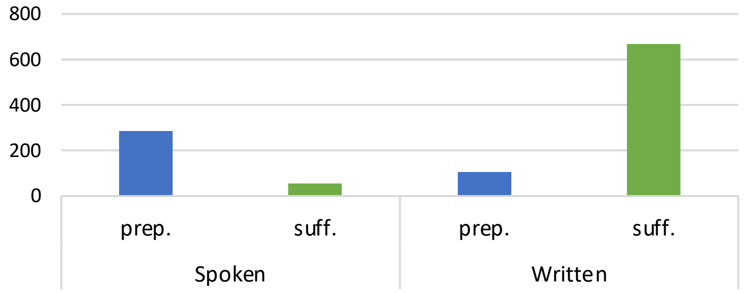
Prepositional construction

(2) **ʃm-i**
name-**POSS.1S**
'my name'

Suffixed construction

### Diversity measures: Variety | Balance | Disparity

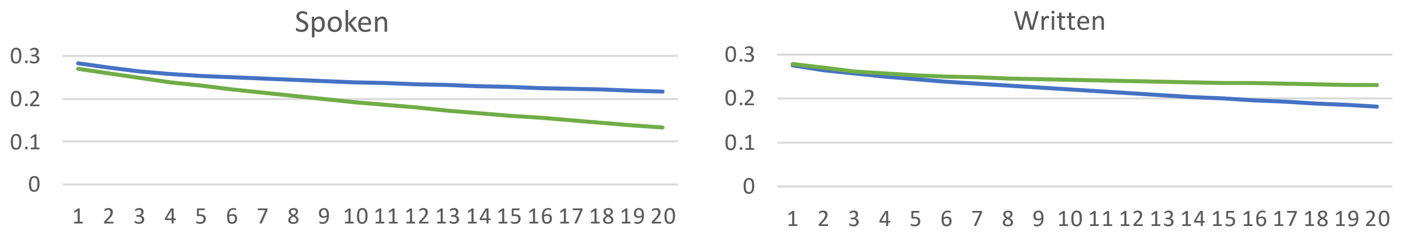
**Variety**: Noun Types Counts



**Balance**:
Type-token ratio

| Spoken | | Written | |
|---|---|---|---|
| prep. | suff. | prep. | suff. |
| 0.438 | 0.320 | 0.890 | 0.319 |

**Disparity:** mean distance of lemma vector from centroids in a Semantic Vector clusters



Work in progress

UniDive
2nd General Meeting
University of Naples “L’Orientale
Naples, Italy, 8-9 February 2024
https://unidive.lisn.upsaclay.fr/

cost
EUROPEAN COOPERATION

Funded by
the European Union

# REVITALIZING THE HISTORICAL ROMANIAN TEXTS WITH CYRILLIC SCRIPTS

CAFTANATOV OLESEA, MALAHOV LUDMILA AND BUMBU TUDOR
Moldova State University, Vladimir Andrunachievici Institute of Mathematics and Computer Science

## The Aim:

The aim of our work is revitalizing the historical Romanian texts with Cyrillic Scripts from the XVII – XX century.

to create linguistic resources of the Historical Romanian

to reissue a folkloric books in latin script that will be used for educational purposes

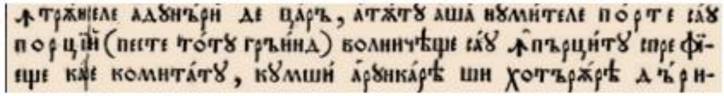
for philological research, for instance at lexicographical diachronic analysis and others

## The Digitization Platform Architecture:

## The Challenges:

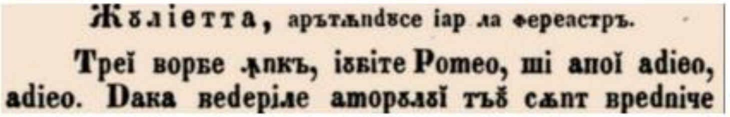
We researched various types of historical documents such as: manuscripts, religions books, dialectal text and others. The endeavor to address the linguistic heritage of Romanian history involves tackling several specific challenges, including:

1) dealing with a multitude of language evolution periods;
2) coping with the scarcity of widely available resources;
3) managing the diverse array of alphabets used in historical printings, including mixed Cyrillic-Latin “transition alphabets”;
4) overcoming the absence of reliable tools for accurately recognizing Cyrillic letters from various historical eras;
5) addressing the shortage of lexicons suitable for the time periods of these resources.



ѫтръжирѣле адꙋнъри де царъ, атътꙋ аша нꙋмителе порте саꙋ
порції (песте тотꙋ грънинд) болничѣще саꙋ ѫпърцитꙋ спре фї-
еще каре комитатꙋ, кꙋмши ѧрꙋнкарѣ ши хотърърѣ дъри-

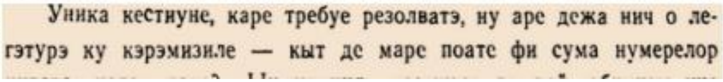
XVIII Century



Жȣліетта, арътѫndȣсе iар ла фереастръ.

Треї ворбе ѫнкъ, iȣвiте Ромео, шi апоï adieo, adieo. Dака веderiле аморȣлȣї тъȣ сѫпт вреdniче

XIX century, mixed alphabet

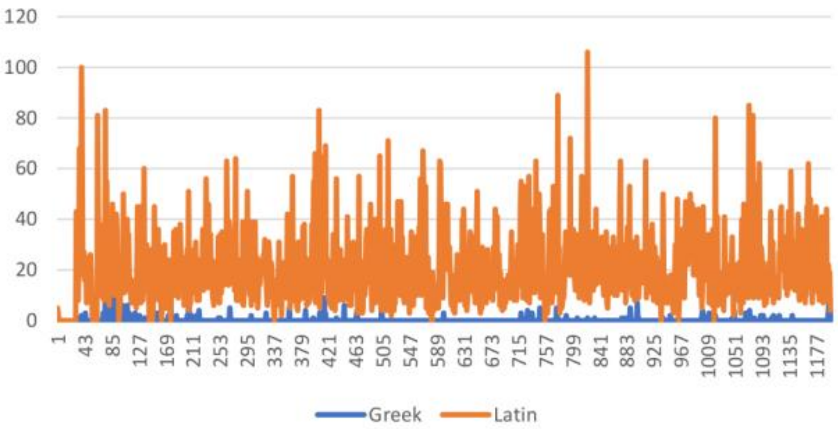


Уника кестиуне, каре требуе резолватэ, ну аре дежа нич о легэтурэ ку кэрэмизиле — кыт де маре поате фи сума нумерелор

XX Century

# The “tongueprint” as language identification tool: Elaborating the proof-of-concept

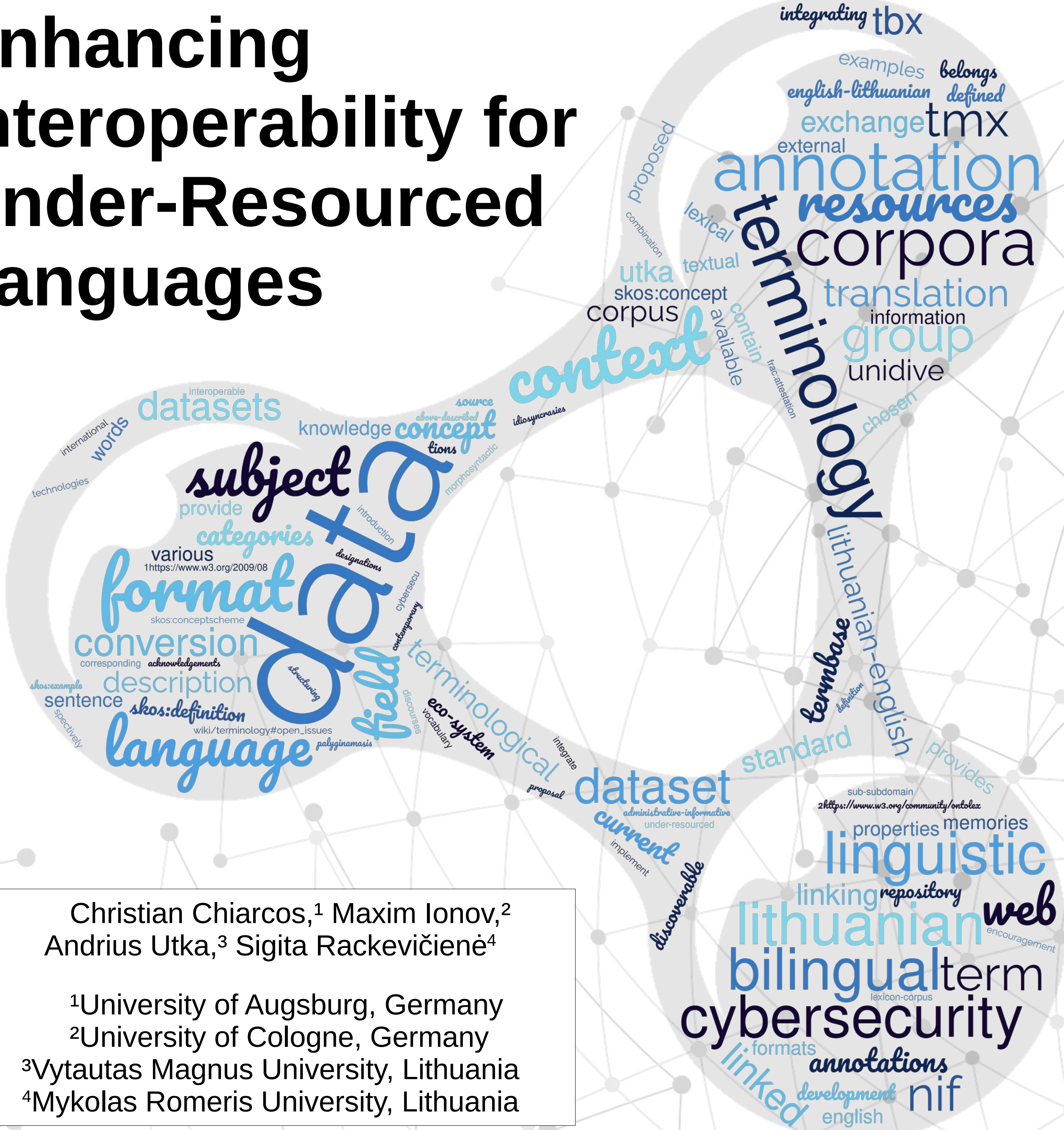
## When was the great age of Latin-Greek bilingualism?



Raf van Rooy, Flavio Massimiliano Cecchini, Isabelle Maes (KU Leuven)

Relevant UniDive working groups: WG1, WG3, WG4

# Enhancing Interoperability for Under-Resourced Languages



Christian Chiarcos,[1] Maxim Ionov,[2]
Andrius Utka,[3] Sigita Rackevičienė[4]

[1]University of Augsburg, Germany
[2]University of Cologne, Germany
[3]Vytautas Magnus University, Lithuania
[4]Mykolas Romeris University, Lithuania

- **Languages**: Lithuanian - English
- **Data**: cybersecurity domain
  1) lexicon (terminology: TBX)
  2) corpora (parallel: TMX / annotated: CoNLL)
- **Challenge**: Publish that such that
  1) All data can be easily re-used
  2) We integrate lexical data, linguistic annotations and parallel corpus
  3) We access / query / interlink / process all data with off-the-shelf technology
- **UniDive**: WG2 (mostly)

Nexus Linguarum

- **Status:** On-going

**We have a solution that works nicely :)**

for us ..., but

1) Can we do better?
   Can we improve data modelling?
2) Can we do more?
   What needs to be done to apply this to other use cases? Where would it be beneficial?

# Old Egyptian Multiword Expressions consisting of a head word + 𓄣 *ib* "heart"

Roberto A. Díaz Hernández

University of Jaén

Work in progress

WG1, WG2, WG3, WG4

## Phases of Egyptian

Egyptian is one of the longest lived languages in history. This Afroasiatic language knew the following phases:

1) **Old Egyptian** (ca. 2700–2000 BC)

2) Middle Egyptian (ca. 2000–1400 BC)

3) Late Egyptian (ca. 1300–700 BC)

4) Demotic (7th century BC to 5th century BC)

5) Coptic (4th to 14th century CE)

## Aim of this research work

It is a semantic and syntactic analysis of Old Egyptian Multiword Expressions (MWEs) consisting of a head word + *ib* "heart".

Old Egyptian uses the noun "heart" with a metonymic meaning to form Multiword Expressions as do some modern languages, for example:

"Listen to your heart" / (G.) "Höre auf dein Herz" / (Sp.) "Escucha a tu corazón"

It is an opportunity to check the validity of the universal categorization of MWEs based mostly on modern Indo-European languages.

## Applying the definition of a MWE to Old Egyptian

(see Savary et al. 2018: 92–93 and Baldwin/Kim 2010: 269)

A MWE is a sequence of words with the following properties:

a) It shows some degree of orthographic, morphological, syntactic and semantic idiosyncrasy.

b) It has at least two lexicalized components including a head word and another syntactically related word.

## Typology

Old Egyptian MWEs consisting of a head word + *ib* can be classified into:

1) Nominal Multiword Expressions (NMWEs) if the head word is a noun. There are two types of NMWEs:

a) Noun/infinitive + *ib*.

b) Adjective/participle + *ib*. This type corresponds to the Sanskrit construction known as *bahuvrīhi*.

2) Prepositional Multiword Expressions (PMWEs) if the head word is a preposition.

3) Verbal Multiword Expressions (VMWEs) if the head word is a verb. There are two types of VMWEs:

a) Light-verb constructions (LVCs).

b) Verbal idioms (IDs).

The poster shows one of the **earliest occurrences of MWEs** in a cross-linguistic perspective. It also contains **a list of 63 ib-MWEs** in Old Egyptian.

## Idiosyncrasy of Old Egyptian MWEs

A word stem can be used in different types of MWEs

1) A verb stem in a VMWE can be transformed into an infinitive in a NMWE.

2) Most of NMWEs derive from a verb stem.

3) A preposition in a PMWE can be used as a nisba adjective in a NMWE.

4) The meaning of a MWE can change due to syntactic reasons.

# An Empirical Study of Multilingual Representations from Language Modeling and Translation

- a principled standpoint and train comparable MT and LM systems to contrast their cross-lingual and monolingual downstream performances;
- an empirical study on publicly available pretrained LM and MT systems and study whether continued training on MT helps or hinders the emergence of cross-lingual capabilities.
- Data: UNPC (Ziemski et al., 2016) and OpenSubtitles (Tiedemann, 2012)
- Languages: Arabic, Chinese, English, French, Russian, and Spanish
- Models
  1. Masked Language Modeling (MLM) with the BERT architecture (Devlin et al., 2019);
  2. Causal Language Modeling (CLM) with the GPT-2 architecture (Radford et al., 2019);
  3. Translation Language Modeling (TLM) with the GPT-2 architecture, where the input is the concatenation of a language pair following a setup similar to Conneau and Lample (2019);
  4. Denoising Sequenece-to-Sequence Langauge Modeling with BART architecture (Lewis et al., 2020);
  5. Machine Translation (MT) with the classic encoder-decoder transformer architecture (Vaswani et al., 2017) and the BART architecture (Lewis et al., 2020).

Babelscape

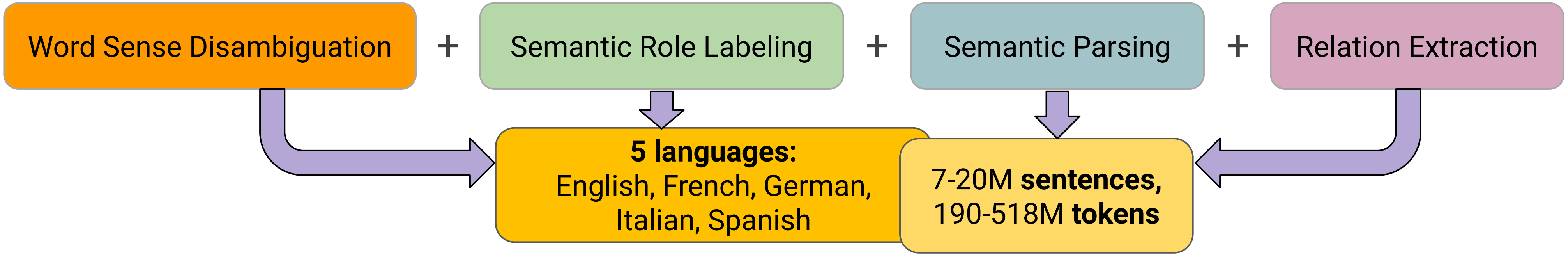
SAPIENZA
UNIVERSITÀ DI ROMA

FUTURE AI RESEARCH

PNRR MUR project
PE0000013-FAIR

# Creating a Multilingual Wide-Coverage Multi-Layered Semantically Annotated Corpus

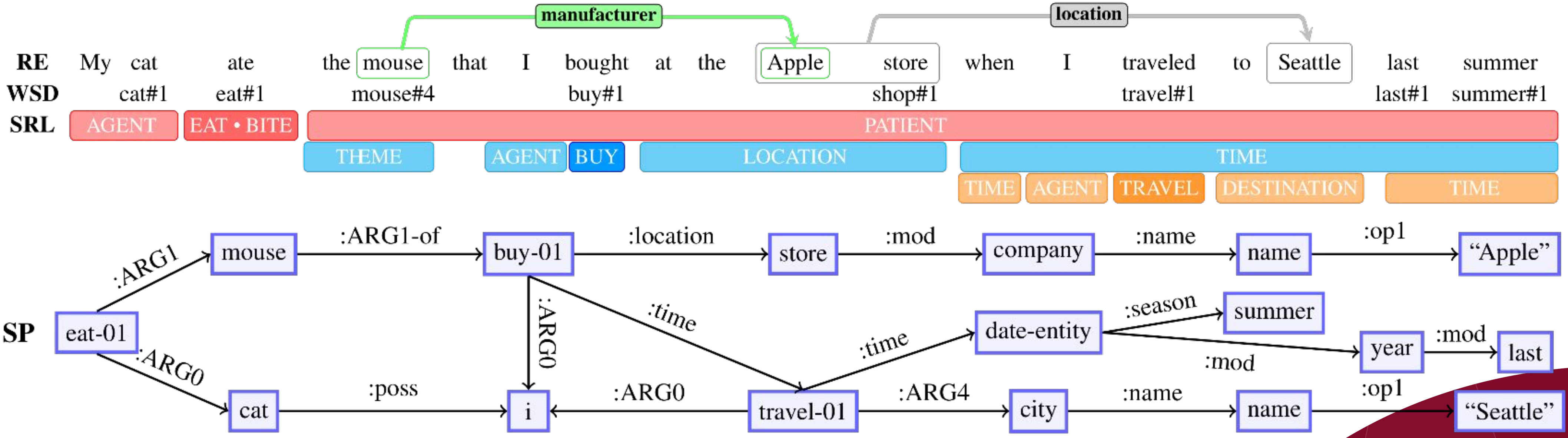
S. Conia, E. Barba, A. Carlos Martinez Lorenzo, P. Huguet Cabot, R. Orlando, L. Procopio, R. Navigli



Word Sense Disambiguation + Semantic Role Labeling + Semantic Parsing + Relation Extraction

**5 languages:**
English, French, German, Italian, Spanish

7-20M **sentences,**
190-518M **tokens**



SAPIENZA
NLP

# Treating Multiword Expressions with a view to Morphologically Rich Languages

UniDive
cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY
Funded by the European Union

WG1
WG2

**Svetlozara Leseva |** zarka@dcl.bas.bg | https://dcl.bas.bg/
Department of Computational Linguistics, Institute for Bulgarian Language, Bulgarian Academy of Sciences

2nd General Meeting
University of Naples "L'Orientale", Naples, Italy
8-9 February 2024

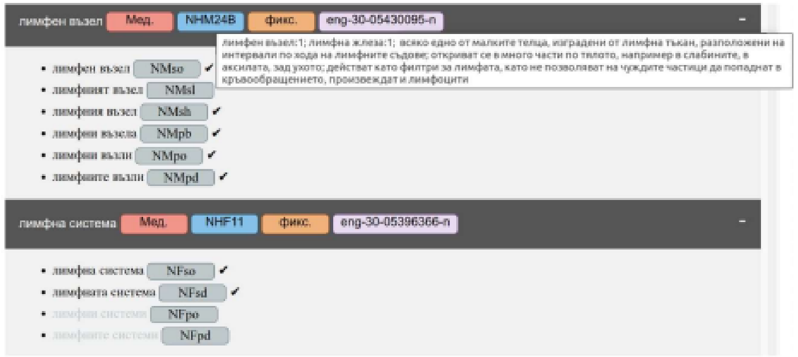
## Objective

- ❖ A uniform linguistic description focusing on:
  - ➢ representation of the structural, morphological, morphosyntactic, word-order, etc. features of Bulgarian MWEs;
  - ➢ extension to MWEs for other languages;
  - ➢ with a view to the automatic recognition and annotation of MWEs in running text.

The lexicon includes:

- ❖ over 10,000 nominal MWEs
  - ➢ including 5,000 NEs;
- ❖ 6,500 verbal MWEs
  - ➢ 1,200 light verb constructions;
  - ➢ 1,800 verbal idioms;
  - ➢ 3,400 reflexives and others.

### Nominal MWEs with validation of forms from corpora



лимфен възел Мед. NHM24B фикс. eng-30-05430095-n −

лимфен възел:1; лимфна жлеза:1; всяко едно от малките телца, изградени от лимфна тъкан, разположени на интервали по хода на лимфните съдове; откриват се в много части по тялото, например в слабините, в аксилата, зад ухото; действат като филтри за лимфата, като не позволяват на чуждите частици да попаднат в кръвообращението, произвеждат и лимфоцити

- лимфен възел NMso ✓
- лимфният възел NMsl
- лимфния възел NMsh ✓
- лимфни възела NMpb ✓
- лимфни възли NMpo ✓
- лимфните възли NMpd ✓

лимфна система Мед. NHF11 фикс. eng-30-05396366-n −

- лимфна система NFso ✓
- лимфната система NFsd ✓
- лимфни системи NFpo
- лимфните системи NFpd

### Description of verbal MWEs

| BG: *удрям джакпота* – EN: *hit the jackpot 'succeed by luck'* | |
|---|---|
| **Synset ID / MWE ID** | eng-30-02524739-v / bg_2291 |
| **MWE lemma / Abstract lemma** | *удрям джакпота / удрям джакпот* |
| **Morphosyntactic features** | *удрям.V_IMPERF_r1s джакпота.Nsh* |
| **Head and head inflection type** | *удрям*.V_IM_TT_S3_01 |
| **Head restrictions** | none |
| **Dependent and dependent restrictions** | *джакпота* / fixed; **N**(umber) = s; **D**(efiniteness) = h |
| **Syntactic structure** | Constituent: V N(P) \| UD: V + obj |
| **Possible modifiers of the head** | regular |
| **Possible modifiers of dependent** | regular; A(P); Ex.: *удрям* ***голямия/Ash*** *джакпот/Ns0* |
| **External elements** | regular (question particle \| subj \| AdvP...) |
| **PARSEME type** | VID |
| **Register and connotation** | Colloquial; −0.125 +0.25 |
| **Derivational relations** | *удряне на джакпота* |

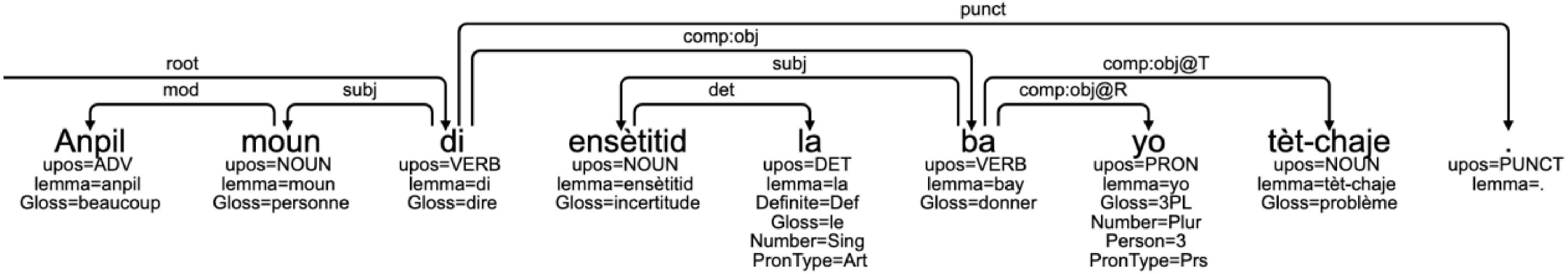
developped in SUD
(Surface-syntactic UD)
converted in UD
(Universal Dependencies)

144 sentences
3279 tokens
glossed and translated
in French

TAM particles
Double object construction
Serial Verb Construction
Relative clause
Determiner

# The first Haitian Creole treebank

Sylvain Kahane, Claudel Pierre-Louis, Sandra Jagodzińska, Agata Savary



‘Many people say that uncertainty is a problem for them.’, lit. gives them head-change

# MULTINCI [WIP] - A MULTILINGUAL NOUN COMPOUND IDIOMATICITY DATASET

THOMAS PICKARD
UNIVERSITY OF SHEFFIELD
tmrpickard1@sheffield.ac.uk

B. MĂDĂLINA ZGREABĂN
UNIVERSITEIT UTRECHT
b.zgreaban@uu.nl

ALINE VILLAVICENCIO
UNIVERSITY OF SHEFFIELD
a.villavicencio@sheffield.ac.uk

## NCTTI

- The Noun Compound Type and Token Idiomaticity dataset [7]: 280 English (en) and 180 Portuguese (pt) nominal compounds (NCs).
- Human annotated in three context sentences (type and token-level annotation).
- PROs: effects of context on annotation judgements, comparison for language models.

## MULTINCI - OBJECTIVES

- Extended NCTTI dataset having core NCs common cross-linguistically, as well as language-specific compounds.
- Include languages with limited MWE resources.
- Increase cross-lingual applications by having correspondences across languages.

## LANGUAGES

- **English** (en):
  - Cleaned & updated context sentences
  - Extended compound list to increase potentially idiomatic expressions
- **Romanian** (ro):
  - Test case for protocol
  - 260 NCs
  - 36 directly equivalent to en; 39 exclusive to Romanian, and 185 that have en translations (not part of the original NCTTI)
- **Georgian** (ka), Irish (ga):
  - Initial work underway (funded by  UniDvie STSMs)
- **Modern Greek** (el), **Ukrainian** (uk), **Brazilian Portuguese** (pt-br):
  - Potential collaborations identified

## FUTURE WORK

- **Protocols** to be refined and completed.
- **Data collection**, translation and its annotation for in-progress and planned languages.
- **Annotations** from human volunteers.
- Extend MultiNCI to **more languages** and their varieties
  - Collaborations welcome
  - See UniDive STSM call

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

## REFERENCES

Funded by the European Union

UniDive

# Morpheme-level Coreference Annotations for Pro-dropped Languages

## Motivation and Goal:

- Coreferential relations of dropped pronouns is necessary for Pro-Dropped languages in Coreference Resolution
    - Null-subjects and omitted possessive pronouns
- Information about dropped pronouns are easily deducible from morphology, *morphemes*.
- Representation and Evaluation Scheme

Sen [benim] anne**[m]**in geldiğ**[i]**ni gördü**[n]** mü?

|  |  |  |  |  |
|---|---|---|---|---|
| ~~Sen benim~~ | **annemin** | geldiğini | gördü**n** mü? |
| You my | mother | came | see_did |

*Did you see the coming of my mother ?*

## Approach:

- Each pronominal marker ~ coreferential mention
    - `Possessive marker' for nouns, and `Personal marker' for verbs.
- No added any artifically inserted token (e.g. empty node)
    - Eliminates difficulty in determining the most accurate and natural position of the empty node in the sentence
- Validated on Turkish Coreference Resolution

## Key Observations:

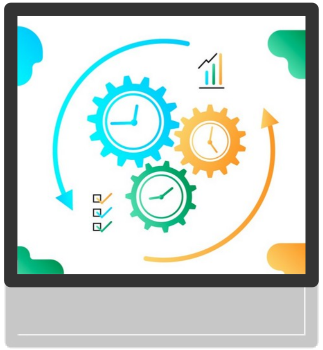
- Representation Scheme: Multiple annotation over a single token is allowed, no need to add artificially inserted token.
- Evaluation Scheme: Pre and post-processors to enhance available CR evaluator to cover dropped pronouns (i.e., multiple annotations over a single word) are developed.

Tuğba Pamay Arslan, Gülşen Eryiğit
{pamay, gulsen.cebiroglu}@itu.edu.tr

#17, Session B

# Adding Semantics to UD *markers*

Tudor Voicu, Verginica Barbu Mititelu



- focus on functional MWEs: conjunctions
- consistent annotation of multiword conjunctions in a UD treebank
- adding a semantic layer to the UD treebank using PDTB 3.0 inventory of discourse relations

# The Philotis Platform: Empowering Low-Resource Language Processing

Vivian Stamou Vasileios Arampatzakis Dimitrios Karamatskos Vasileios Sevetlidis Nicolaos Valeontis Stella Markantonatou George Pavlidis

Philotis web-based platform:

- Full pipeline for speech-, image- and text-to-text development of raw and annotated corpora and models (UD-framed)
- Key-board development facility
- Addressed to language specialists with varying technical expertise
- Openly available technology

2nd General Meeting
University of Naples "L'Orientale Naples, Italy,
8-9 February 2024
https://unidive.lisn.upsaclay.fr/

# Après Toi: Scoring Systems based on Dataset Votes

Yuval Pinter
Ben-Gurion University
uvp@cs.bgu.ac.il

WG4

## The Problem

- When running a multi-dataset competition between systems, we use the **averaging** aggregate metric for deciding the winner
- On one hand, this ensures that small datasets, typically representing low-resource languages, are viewed as equivalent to large datasets
- On the other hand, language differences may lead to lower points of saturation for some, leading to focus work on "easy" languages (or on one's "comfort zone")

System 1 wins on the average metric, but is it really the best?

| | Dataset A | Dataset B | Dataset C |
|---|---|---|---|
| System 1 | 90 | 53 | 79 |
| System 2 | 80 | 56 | 82 |
| System 3 | 70 | 57 | 83 |
| System 4 | 60 | 58 | 84 |

## The Alternative(s)

**Voting-based scoring**

- Each dataset has a budget of "votes" which it distributes among the systems
- In Eurovision-style voting (ESC):
  - Top system gets 12 points
  - #2 gets 10
  - #3 gets 8, and #4-#10 get one point less each
- Final scores are the accumulations from the datasets

| | | |
|---|---|---|
| SWEDEN | 4 | 281 |
| ITALY | 8 | 150 |
| ISRAEL | 7 | 149 |
| FINLAND | 1 | 142 |
| ESTONIA | 5 | 115 |
| AUSTRALIA | 2 | 97 |
| BELGIUM | 12 | 91 |
| CZECHIA | | 88 |
| AUSTRIA | | 86 |
| SPAIN | | 84 |
| LITHUANIA | 6 | 65 |
| CYPRUS | 3 | 59 |
| NORWAY | | 52 |
| SWITZERLAND | | 51 |
| ARMENIA | 10 | 48 |
| FRANCE | | 48 |
| UKRAINE | | 40 |
| PORTUGAL | | 33 |
| SLOVENIA | | 26 |
| MOLDOVA | | 20 |
| UNITED KINGDOM | | 15 |
| ALBANIA | | 14 |
| SERBIA | | 13 |
| POLAND | | 12 |
| CROATIA | | 11 |
| GERMANY | | 2 |