DeriMo 2025

Fifth International Workshop on Resources and Tools for Derivational Morphology

University of Fribourg, Switzerland

September 4-5, 2025

**Swiss National
Science Foundation**

UNI
FR

UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

DeriMo workshops

- The DeriMo workshops aim to bring together researchers in derivational morphology to discuss language resources and tools that support their research

  - Methodological focus: DeriMo provides a platform for presenting advancements in existing resources and tools, and for introducing new ones

  - Theoretical impact: DeriMo explores how both established and emerging topics in derivational morphology can benefit from diverse methodological approaches

- Previous editions were held in Milan (DeriMo 2017), Prague (DeriMo 2019), Nancy (DeriMo 2021), and Dubrovnik (DeriMo 2023)

- The 2025 workshop will build on the discussions initiated in the previous editions

DeriMo 2025 organization

- Program committee chairs

  Richard Huyghe
  Megan Prudent
  Justine Salvadori
  Matea Filko
  Krešimir Šojat

- Local organization committee

  Raphaël Cornaz
  Richard Huyghe
  Megan Prudent
  Justine Salvadori

- Program committee members

Alexandra Bagasheva	Lívia Körtvélyessy	Ingo Plag
Olivier Bonami	Eleonora Litta	Jan Radimský
Maria Copot	Claudia Marzi	Andrea Sims
Cristina Fernández-Alcaina	Fabio Montermini	Pavol Štekauer
Jesús Fernández-Dominguez	Akiko Nagano	Pavel Štichauer
Nabil Hathout	Fiammetta Namer	Marko Tadić
Martin Hilpert	Renáta Panocová	Salvador Valera
Gianina Iordăchioaia	Marco Passarotti	Zdeněk Žabokrtský

Financial support

- DeriMo 2025 is supported by:

  - The University of Fribourg

  - The Swiss National Science Foundation

Program

- 2 Invited speakers
  - Sabine Arndt-Lappe (Universität Trier)
  - Magda Ševčíková (ÚFAL, Charles University, Prague)
- 12 papers selected for presentation
- Workshop proceedings are available online

  `https://events.unifr.ch/derimo2025/en/`

Research focus

- DeriMo 2025 is intended to cover a wide range of topics, including:
    - Resources for word-formation research in single or multiple languages
    - Representation of word-formation processes in lexical resources
    - Statistical and computational modelling of derivational data
    - Diversity of information (morphological, phonological, semantic, syntactic) encoded in lexical databases, and enrichment of morphological resources with additional features
    - Compilation and standardization of existing resources, and integration of morphological resources with other types of linguistic data
    - Data-intensive research on all aspects of word formation (for individual languages, with a comparative focus, from a diachronic perspective, etc.)
    - Theoretical accounts of word formation based on quantitative and computational methods

Overview of the presentations: themes

- Presentations at the workshop will address the following themes:

    - Lexical and morphological borrowing (Ševčíková; Stephen & John)
    - Productivity (Lacić; Pakerys, Navickaitė-Klišauskienė & Dadurkevičius; Zampetta)
    - Morphological competition (Schäfer)
    - Overabundance (Pellegrini, Litta & Iurescia)
    - Stress assignment (Arndt-Lappe)
    - Morphophonological alternation (Namer, Lignon & Hathout)
    - Compound structuration (Šojat)
    - Semantics of word-formation processes (Barque)
    - Valency in derivation (Hledíková)
    - Morphological segmentation (Olbrich & Žabokrtský)
    - Base word recognition (John & Žabokrtský)

Overview of the presentations: data collection

- The linguistic data examined in the presentations are sourced from:

    - Corpus data (e.g., Pakerys et al.)

    - Existing lexical resources (e.g., Olbrich & Žabokrtský)

    - A combination of corpus and lexical resources (e.g., Pellegrini et al.)

    - Experimentally elicited data (e.g., Arndt-Lappe)

- Data collection and analysis may result in the development of new resources and databases (e.g., Barque; Hledíková; Lacić; Namer et al.; Ševčíková; Šojat; Zampetta)

Overview of the presentations: data analysis

- A large diversity of methods is employed to analyze the collected data, including (but not limited to):

    - Manual annotation (Lacić; Ševčíková)

    - Semi-automatic annotation (Hledíková; Pakerys et al.; Namer et al.)

    - (Semi-)supervised classification (Barque; Stephen & John; Olbrich & Žabokrtský)

    - Distributional semantics (Schäfer)

    - Analogical modelling (Arndt-Lappe)

    - Frequency distribution modelling (Zampetta)

    - Network analysis (Arndt-Lappe)

Overview of the presentations: languages

- The languages under study include:
  - Ancient Greek (Zampetta)
  - Croatian (Sojat)
  - Czech (Hledíková; Stephen & John; Ševčíková)
  - English (Arndt-Lappe; Schäfer)
  - French (Barque; Namer et al.)
  - Italian (Lacić)
  - Latin (Pellegrini et al.)
  - Lithuanian (Pakerys et al.)
  - Etc. (Olbrich & Žabokrtský; John & Žabokrtský)

Have a productive and rewarding workshop!