Proceedings of the 15th Workshop on Natural Language Processing for Computer Assisted Language Learning (NLP4CALL 2026)



University of Tartu Library
NEALT Proceedings Series Number: 61
ISBN: 978-9985-4-1522-1
ISSN: 1736-8197
eISSN 1736-6305
2026

Proceedings of the

15th Workshop on Natural Language Processing for Computer Assisted Language Learning (NLP4CALL 2026)

edited by
Arianna Masciolini, David Alfter, Herbert Lange,
Ricardo Muñoz Sánchez, Ildikó Pilán, Elena Volodina

Proceedings and all papers therein
published under a CC BY 4.0 license:
https://creativecommons.org/licenses/by/4.0

Front cover image by Pasi Mämmelä (mammela)
via Pixabay

University of Tartu Library

NEALT Proceedings Series Number: 61
ISBN: 978-9985-4-1522-1
ISSN: 1736-8197
eISSN 1736-6305

2026

Preface

The workshop series on Natural Language Processing (NLP) for Computer-Assisted Language Learning (NLP4CALL) is a meeting place for researchers working on integrating Natural Language Processing and Speech Technologies in CALL systems and exploring the theoretical and methodological issues arising in this connection. The latter includes, among others, the integration of insights from Second Language Acquisition (SLA) research and the promotion of "Computational SLA" through setting up Second Language research infrastructures.

The intersection of Natural Language Processing (or Language Technology / Computational Linguistics) and Speech Technology with Computer-Assisted Language Learning (CALL) brings "understanding" of language to CALL tools, thus making CALL intelligent. This fact has inspired the name for this area of research — Intelligent CALL, ICALL for short. As the definition suggests, apart from having excellent knowledge of Natural Language Processing and/or Speech Technology, ICALL researchers need good insights into second language acquisition theories and practices, as well as knowledge of second language pedagogy and didactics. Therefore, this workshop invites a wide range of ICALL-relevant research, including studies where NLP-enriched tools are used for testing SLA and pedagogical theories, and those where SLA theories (pedagogical practices or empirical data) are modeled using ICALL tools. The NLP4CALL workshop series is aimed at bringing together competences from these areas for sharing experiences and brainstorming around the future of the field.

Topics of Interest

We invited submissions:

- that describe research directly aimed at ICALL;
- that describe the ongoing development of resources and tools with potential use in ICALL, either directly in interactive applications, or indirectly in materials, application, or curriculum development, e.g. learning material generation, assessment of learner texts and responses, individualized learning solutions, provision of feedback;
- that discuss challenges and/or research agendas for ICALL;
- that describe empirical studies on language learner data.

This year, the workshop has a special focus on **process-oriented approaches to educational NLP**, including but not limited to work related to the collection and analysis of keystroke data from language learners.

We encouraged paper presentations and software demonstrations describing the above-mentioned themes primarily, but not exclusively, for the Nordic languages.

Invited speakers

This year, we had the pleasure to welcome two invited speakers: Andrea Horbach (Kiel University and the Leibniz Institute for Science and Mathematics Education (IPN)) and Rianne Conijn (Eindhoven University of Technology).

Andrea Horbach has been professor for Educational NLP at Kiel University and the Leibniz Institute for Science and Mathematics Education (IPN) since 2024. Her research interests include free-text scoring, argument mining, automated feedback, and the use of process data for assessment. She is particularly interested in developing explainable and fair methods that

support human agency in writing and assessment processes. The title of her talk was *A Focus on the Writing Process With and Without AI Support: Insights from the German PISA 2025 Foreign Language Assessment Data Collection.*

Rianne Conijn is an assistant professor in learning analytics in the Human-Technology Interaction group at Eindhoven University of Technology, the Netherlands. Her research focuses on developing methods to capture and model learning behavior (e.g., self-regulated learning) and writing behavior (e.g., revision processes), using fine-grained behavioral data, such as keystroke and clickstream data. In addition, she uses these methods to examine how educational technologies, such as generative AI, affect the learning process. Rianne recently received a NWO Veni grant on "Human-Centered AI in education" where she aims to improve human-AI collaboration in writing in higher education. The title of her talk was *Moving beyond the product: Keystroke Logging for CALL*

Previous workshops

This workshop follows a series of workshops on NLP4CALL organized by the NEALT Special Interest Group on Intelligent Computer-Assisted Language Learning (SIG-ICALL)[1]. The workshop series has previously been financed by:

- the Center for Language Technology at the University of Gothenburg;
- the SweLL project;[2]
- the Swedish Research Council's conference grant, Språkbanken Text;[3]
- the L2 profiling project;[4]
- itec;[5]
- the CENTAL;[6]
- the Analytics for Language Learning (A4LL) project[7] at LIDILE - Univ Rennes;
- the Mormor Karl project;[8] and
- the project "Expanding the scope of a multi-purpose lexicographic resource to grammar and L2 competence".[9]

Submissions to the fifteen workshop editions have targeted a wide range of languages, ranging from well-resourced languages (Chinese, German, English, French, Portuguese, Russian, Spanish) to lesser-resourced languages (Erzya, Arabic, Estonian, Irish, Komi-Zyrian, Meadow Mari, Saami, Udmurt, Võro). Among these, several Nordic languages have been targeted, namely Danish, Estonian, Finnish, Icelandic, Norwegian, Saami, Swedish and Võro. The wide scope of the workshop is also evident in the affiliations of the participating authors as illustrated in Table 1.

[1] https://spraakbanken.gu.se/en/research/themes/icall/sig-icall
[2] https://spraakbanken.gu.se/en/projects/swell
[3] https://spraakbanken.gu.se
[4] https://spraakbanken.gu.se/en/projects/l2profiles
[5] https://itec.kuleuven-kulak.be
[6] https://cental.uclouvain.be
[7] https://sites-recherche.univ-rennes2.fr/lidile/articles/a4all/
[8] https://mormor-karl.github.io/
[9] https://eki.ee/prg-1978/

The acceptance rate has varied between 44% and 83%, the average being 62% (see Table 2). Although the acceptance rate is rather high, the reviewing process has always been very rigorous with two to three double-blind reviews per submission. This indicates that submissions to the workshop have usually been of high quality.

Country	**Count**	**Country**	**Count**
Algeria	1	Italy	16
Australia	2	Japan	7
Belgium	21	Lithuania	1
Canada	4	Netherlands	4
China	5	Norway	16
Cyprus	3	Poland	1
Czech Republic	2	Portugal	8
Denmark	5	Romania	1
Egypt	1	Russia	10
Estonia	6	Slovakia	1
Finland	15	Spain	5
France	51	Sweden	101
Germany	145	Switzerland	16
Greece	6	Ukraine	2
Iceland	6	UK	35
Iran	2	Uruguay	5
Ireland	7	US	15
Israel	1	Vietnam	3

Table 1: NLP4CALL speakers' and co-authors' affiliations, 2012–2026

Workshop year	Submitted	Accepted	Acceptance rate
2012	12	8	67%
2013	8	4	50%
2014	13	10	77%
2015	9	6	67%
2016	14	10	71%
2017	13	7	54%
2018	16	11	69%
2019	16	10	63%
2020	7	4	57%
2021	11	6	55%
2022	23	13	56%
2023	18	12	67%
2024	23	19	83%
2025	16	7	44%
2026	51	25	49%

Table 2: Submissions and acceptance rates, 2012–2026

Program committee

We would like to thank our Program Committee for providing detailed feedback for the reviewed papers:

- Soroosh Akef, Leibniz-Institut für Wissensmedien, Germany
- David Alfter, University of Gothenburg, Sweden
- Stephen Bodnar, University of Tübingen, Germany
- Christopher Bryant, University of Cambridge, UK
- Andrew Caines, University of Cambridge, UK
- Xiaobin Chen, University of Tübingen, Germany
- Orphée De Clercq, Ghent University, Belgium
- Kordula De Kuthy, Leibniz-Institut für Wissensmedien, Germany
- Piet Desmet, KU Leuven, Belgium
- Irene Elmerot, University of Gothenburg, Sweden
- Thomas Francois, Université catholique de Louvain, Belgium
- Thomas Gaillat, Université Rennes 2, France
- Johannes Graën, University of Zurich, Switzerland
- Andrea Horbach, FernUniversität in Hagen, Germany
- Nikolai Ilinykh, University of Gothenburg, Sweden
- Jelena Kallas, Institute of the Estonian Language, Estonia
- Joni Kruijsbergen, Ghent University, Belgium
- Murathan Kurfalı, RISE Research Institutes of Sweden, Sweden
- Ronja Laarmann-Quante, Ruhr-Universität Bochum, Germany
- Herb Lange, University of Gothenburg, Sweden
- Peter Ljunglöf, Chalmers University of Technology, Sweden
- Arianna Masciolini, University of Gothenburg, Sweden
- Felix Morger, University of Gothenburg, Sweden
- Ricardo Muñoz Sánchez, University of Gothenburg, Sweden
- Lionel Nicolas, Eurac Research, Italy
- Robert Östling, Stockholm University, Sweden
- Ulrike Pado, HFT Stuttgart, Germany

- Magali Paquot, Université catholique de Louvain, Belgium
- Ildikó Pilán, University of Gothenburg, Sweden
- Egon W. Stemle, Eurac Research, Italy
- Maria Irena Szawerna, University of Gothenburg, Sweden
- Anaïs Tack, KU Leuven, Belgium
- Irina Temnikova, Sofia University "St. Kliment Ohridski", Bulgaria
- Ahmet Yavuz Uluslu, University of Zurich, Switzerland
- Sowmya Vajjala, National Research Council, Canada
- Elena Volodina, University of Gothenburg, Sweden
- Ekaterina Voloshina, University of Gothenburg and Chalmers University of Technology, Sweden
- Niklas Zechner, University of Gothenburg, Sweden
- Torsten Zesch, FernUniversität in Hagen, Germany

We intend to continue this workshop series, which so far has been the only ICALL-related recurring event based in the Nordic countries. Our intention is to co-locate the workshop series with the two major LT events in Scandinavia, the Swedish Language Technology Conference (SLTC) and the Nordic Conference on Computational Linguistics (NoDaLiDa), thus making this workshop an annual event. Through this workshop, we intend to profile ICALL research in Nordic countries as well as beyond, and we aim at providing a dissemination venue for researchers active in this area.

Workshop website

`https://nlp4call.github.io/`

Workshop organizers

- Arianna Masciolini, Språkbanken, University of Gothenburg, Sweden
- David Alfter, Gothenburg Research Infrastructure in Digital Humanities (GRIDH), University of Gothenburg, Sweden
- Herbert Lange, Språkbanken, University of Gothenburg, Sweden
- Ricardo Muñoz Sánchez, Språkbanken, University of Gothenburg, Sweden
- Ildikó Pilán, Norwegian Computing Center, Norway
- Elena Volodina, Språkbanken, University of Gothenburg, Sweden

Acknowledgments

The 2026 edition of this workshop was supported jointly by:

- Språkbanken – jointly funded by its 10 partner institutions and the Swedish Research Council (2025–2028; project id 2023-00161);[10]
- Huminfra;[11]
- the Department of Swedish, Multilingualism, Language Technology;[12]
- the IncluEdS project[13]

[10] https://spraakbanken.gu.se/
[11] https://www.huminfra.se/
[12] https://www.gu.se/en/swedish
[13] https://inclueds.org/

Content

Preface i
Arianna Masciolini, David Alfter, Herbert Lange, Ricardo Muñoz Sánchez, Ildikó Pilán and Elena Volodina

How Much of Native Language Identification Is Corpus Artifact? A Cross-Corpus Audit of SVMs and LLMs on TOEFL11 and ICLEv3 1
Shima Rahimi

Lexical Profiling for Dutch: From Resource to Real-Life Implementation 16
Jasper Degraeuwe and Natalie Finlayson

Making Feedback Visible: Character-Level Error Localization in Handwritten CALL Input 34
Said Yasin and Torsten Zesch

Towards Semi-Automatic Benchmarking of Large Language Models for Self-Regulated Language Learning 40
Émile Alexandre, Yannick Parmentier and Guillaume Nassau

Adversarial Disentanglement of Speaker Identity and L2 Phonological Errors via Raw-Waveform Architectures 58
Amirhossein Ghanipour Amirhandeh

Investigating the role of revisions for learner level predictions 65
Ahood Al Sawar, Léo Nebel and Nicolas Ballier

Universal, but Grounded? Anchoring Cross-Lingual CEFR Assessment in the Framework's Own Descriptors 78
David Alfter

Assessing the Reliability of Automatic Speech-Processing Methods for L2 Speech: Evidence from an Automatic Pipeline, Posteriorgrams, and Whisper Probability Scores 99
Mehak Sohail and Nicolas Ballier

Multilingual language learning with Grammatical Framework 111
Adrian De Lon and Inari Listenmaa

Using LLMs for feature extraction: the case of language functions with deixis 119
Thomas Gaillat, Xiaobin Chen and Nelly Sagirov

Exploring the syntactic parsing of learner texts in Modern Greek: Variety trumps size, genre trumps grammaticality 130
Christina Klironomou, Arianna Masciolini and Aleksandrs Berdicevskis

Predicting Personalized Vocabulary Learning Burden from Intralexical and Interlexical Factors 140
Mykhailo Diachenko and John P. McCrae

A Multidimensional Evaluation of LLM-Generated Learning Materials for Greek Primary School Learners 160
Eleftherios Kalogeros, Sofia Stamou, Markella Kolocheri, Thaleia-Dionysia Kolovou, Manolis Gergatsoulis and Matthew Damigos

Etymology Can Inform Lexical Complexity Prediction 173
Noah-Manuel Michael, Anna Hülsing and Andrea Horbach

Data Augmentation for L2 English Speaking Assessment using TTS 197
Stefano Bannò, Penny Karanasou, Mengjie Qian, Kate M. Knill and Mark J. F. Gales

Automatic Analysis of Noun Phrase Complexity in Learner spoken Productions 214
Saint Joannes Monengue Essengo, Bi Yoman Ghislain Aymard Toa, Thomas Gaillat, Christine Evain and Fabienne Moreau

A corpus of written Swedish texts and keystroke data from deaf and hard-of-hearing children and adolescents 229
Moa Gärdenfors, Robert Östling, Lisa Rudebeck and Mats Wirén

CineLexi: A corpus-based bilingual resource of evaluative vocabulary in film reviews (German–Russian) 244
Mariia Kopylova

Translation Surprisal Predicts Learner Error in L2 French Production: A Language-Pair-Specific Difficulty Feature for ICALL 256
Matteo Lanza

Behind the Smokescreen: Uncovering Cross-LLM Bias and Reasoning Contradictions of LLM-as-a-judge for Generated Feedback 269
Céline Leuzinger

Reading Ahead: Pace-Based Prediction for Real-Time Tracking of Children's Oral Reading in the Browser 288
Bohdan Podziubanchuk, Einar Meister and Tanel Alumäe

ABC Diagnostics: A Web-Based Tool for Explainable CEFR Classification 295
Ricardo Muñoz Sánchez, Kostiantyn Kucher, Filip Kayar and Elena Volodina

Acoustic and Lexical Signals for Native Language Identification of Language Learners 306
Ahmet Yavuz Uluslu, Mengjie Qian, Kate Knill and Mark Gales

Poor Generalisation of Proficiency Classifiers Trained on Linguistic Complexity Features to New Writing Tasks 314
Emre Güçer, Soroosh Akef and Detmar Meurers

Automated L2 Speech Assessment with Audio-Based Self-Supervised Neural Networks 334
Eli Stafford and Sylvain Coulange