

## Preface

**Jens Lemmens\***

JENS.LEMMENS@UANTWERPEN.BE

**Lisa Hilte\*\***

LISA.HILTE@VAIA.BE

**Jens Van Nooten\***

JENS.VANNOOTEN@UANTWERPEN.BE

**Maxime De Bruyn\***

MAXIME.DEBRUYN@UANTWERPEN.BE

**Pieter Fivez\***

PIETER.FIVEZ@UANTWERPEN.BE

**Ine Gevers\***

INE.GEVERS@UANTWERPEN.BE

**Jeska Buhmann\***

JESKA.BUHMANN@UANTWERPEN.BE

**Ehsan Lotfi\***

EHSAN.LOTFI@UANTWERPEN.BE

**Nicolae Banari\***

NICOLAE.BANARI@UANTWERPEN.BE

**Nerves Yuzbashyan\***

NERSES.YUZBASHYAN@UANTWERPEN.BE

**Walter Daelemans\***

WALTER.DAELEMANS@UANTWERPEN.BE

\* University of Antwerp (CLiPS)  
Prinsstraat 13, 2000 Antwerp (Belgium)

\*\* Vlaamse AI Academie (VAIA)

The 13<sup>th</sup> volume of Computational Linguistics in the Netherlands (CLIN) contains finished and reviewed work that was presented at the 33rd CLIN conference<sup>1</sup>. CLIN is organized annually, and took place at the city campus of the University of Antwerp in 2023. In total, we received 122 abstracts, of which 120 were accepted and 118 were presented. Thirty-nine abstracts were presented as talks distributed over various parallel sessions, whereas the remaining 79 abstracts were presented as poster presentations throughout two afternoon sessions. In addition, it was our pleasure to welcome Prof. Dr. Irena Gurevych of the UKP lab (Darmstadt, Germany) and current president of the Association for Computational Linguistics (ACL) as our keynote speaker. In her talk, she discussed the importance of automated fact-checking (AFC), exemplified by the COVID-19 infodemic and the increasing usage of hallucinating large language models (LLMs). The gap between academic research in AFC and applied AFC was highlighted, and a selection of interesting future research directions were proposed, including the development of high-quality datasets and models that can reconstruct fallacious arguments in misinformation.



1. Credits of pictures go to Jeroen Van Sweeveldt.

After the conference, we received 20 submissions to the current volume of CLIN journal, of which 10 were accepted for publication after reviewing. The most dominant trend in these papers was a focus on large language models. First, Fivez et al. describe the results of the CLIN33 shared task. This task consisted of a scientifically and societally relevant problem: the detection of texts generated by large language models. Concretely, models for binary classification across six domains and two languages had to be designed. Four teams participated, and a prize consisting of 500 euro and a certificate was awarded to the teams that achieved highest performance per language (averaged across the domains). For the Dutch and English data, respectively, this was the Elsevier team (Yury Kashnitsky and Savvas Chamezopoulos) and the Radboud University delegation (Hans Van Halteren). Congratulations!



Additionally, Kosar et al. conducted a study in which they compare the ability of humans to assign topics to news texts with LLMs. The main takeaways of their work are that LLMs can generate topic labels that are on par or even of higher quality than human-generated labels, and that human annotators are biased by individual cognitive processes and linguistics preferences when performing topic classification. They also emphasize that future work should focus on developing metrics that evaluate attributes such as completeness, concreteness, objectivity, and length when assigning names to topics.

In other work, Wolters and Van Craenenburgh use T5 models to transform historical Dutch spelling to contemporary spelling. They show that their approach does not only outperform an existing method based on linguistic rules, but also that the character-based model variants generalize to words that were not seen in the training data, indicating their usefulness for the task of historical spelling normalization.

Kruisbergen et. al investigate how accurate zero-shot language error detection with GPT3.5 is compared to fine-tuned (Ro)BERT(a) models. They focus on Dutch writing products of L1 and L2 speakers and report results on different types of errors, such as spelling errors, grammatical errors, and capitalization errors. The findings indicate that fine-tuning smaller models leads to higher accuracy than prompting GPT, which also has a high post-processing load. Additionally, the authors found that the error annotations in the utilized data were inconsistent, and emphasize that creating high-quality annotations is essential in future work.

Seidl and Vandeghinste introduce a new sentence simplification method for Dutch using T5 and control tokens to manage complexity (e.g. sentence length and syntactic structures). Synthetic datasets were created to evaluate the approach, and the results show that using control tokens substantially increase sentence simplicity (using the SARI metric). This work provides stepping stones towards future research in Dutch sentence simplification.

Luden et al. present work on a well-known and relevant issue in LLMs, namely performance deterioration due to lacking temporal generalization capabilities. Rather than using perplexity as a metric for this phenomenon, they propose a contextual word definition generation task for evaluation. Their results show that the proposed approach can be used to evaluate temporal generalization, not only for new words, but also for words that are subject to semantic change over time.

Two submissions that were accepted present resources for Dutch. First, Delobelle & Remy introduce RobBERT-2023 (base and large), the latest iterations of their Dutch RoBERTa-based models. The authors report results on various tasks, such as document classification and sequence labeling, and compare them with previous versions of RobBERT and other (Ro)BERT(a) architectures. Secondly, De Langhe et al. present a benchmark dataset for zero-shot Dutch-language text classification. The benchmark includes (aspect-based) sentiment analysis, emotion detection, irony detection, news topic classification, and an event co-reference resolution task. The authors provide results for various generative LLMs, but also for NLI- and MLM-based zero-shot approaches.

Finally, the two remaining papers focused more on linguistics. Sung and Prokić propose a method for the automatic extraction of linguistic features through normalized pointwise mutual information (nPMI). They compare the method to Fischer's linear discriminant and factor analysis, and evaluate the methods based on exclusivity and representativeness. The results suggest that nPMI outperforms the other approaches with respect to exclusivity, of which they argue that it is the most relevant metric. Bitew et al., on the other hand, hypothesize that personality, which is essential information in psychopathology, can be inferred more effectively from patients' language than from questionnaires. Various methodologies are investigated, and it is shown that features based on LIWC significantly outperform classifiers based on questionnaire answers. The best results were achieved by combining both methods.

To conclude, we would like to thank everyone who has played a role in turning CLIN33 into a grand success, which first and foremost are all presenters, authors and 231(!) participants. We also thank our sponsors for their generous support: Instituut voor de Nederlandse Taal (INT), Taalunie, Textgain, Textkernel, Stichting Toepassing Inductieve Leerteknieken (STIL), Faktion, Vlaamse AI Academie (VAIA), Zeta Alpha, CLARIAH, CrossLang, De Taalsector, Nederlandse Organisatie voor Taal- en Spraaktechnologie (NOTaS), and Erhnmyr. Additionally, we want to convey our appreciation towards the students of the Digital Text Analysis master's program who kindly volunteered to assist during the conference: Siddharth Singh, Saida Karmass Farcou, Franci Haest, Oliver Bogaerts, Ellen Jansen, and Jeroen Van Sweeveldt. Further, we would like to express our gratitude towards the editorial board of CLIN journal and the external reviewers for their insightful comments and thorough reviewing. Finally, CLIN33 would not have been possible without the organizational committee, consisting of Walter Daelemans, Lisa Hilte, Jens Lemmens, Jens Van Nooten, Maxime De Bruyn, Pieter Fivez, Ine Gevers, Jeska Buhmann, Ehsan Lotfi, Nicolae Banari, and Nerves Yuzbashyan. We wish you an interesting reading experience, and we are looking forward to seeing you again at the 34<sup>th</sup> edition of CLIN, organized in Leiden on the 30<sup>th</sup> of August 2024.

