

CURRICULUM VITAE

ORPHÉE DE CLERCQ

Contact	Department of Translation, Interpreting and Communication Faculty of Arts and Philosophy – Ghent University Groot-Brittanniëlaan 45 B - 9000, Ghent Belgium
E-mail	orphee.declercq@ugent.be
Website	http://www.lt3.ugent.be/people/orphee-de-clercq/

Research Interests

Natural Language Processing, Readability Prediction, Automated Writing Evaluation, Intelligent Computer-Assisted Language Learning, Sentiment Analysis, Emotion Detection

1 Education

PhD Ghent University. (Computational) Linguistics (Doctoraat in de Taalkunde). 2015.
Dissertation: Tipping the scales. Exploring the added value of deep semantic processing on readability prediction and sentiment analysis.

M.A. University College Ghent. Multilingual Communication (English, French). 2008.
Master's Thesis: The use of metaphors in the language of two conservative movements: the American Christian Right and the Flemish nationalist party Vlaams Belang.

B.A. University College Ghent. Applied Linguistics (English, French). 2007.

2 Work Experience

February 2021 - Current

Assistant professor of language technology for educational applications at Ghent University. During my Tenure Track I am continuing my research on readability prediction and automated writing evaluation while focusing more explicitly on the domain of intelligent Computer-Assisted Language Learning. I am currently (co-)supervising four PhD-students on topics related to automated writing evaluation, automated writing assessment, event coreference resolution and cross-lingual transfer learning. I am also involved in various interdisciplinary research projects (SentEMO, Steunpunt Toetsen) and offer advice on incorporating NLP and

ML techniques into educational applications.

August 2015 - January 2021

Senior researcher (post-doctoral assistant) at Ghent University. Worked on readability prediction, domain portability, translatability, automated writing evaluation, fine-grained sentiment analysis and emotion detection. In 2017, I have also been able to valorize my research on aspect-based sentiment analysis during a VLAIO innovation project together with the Flemish company Hello Customer.

October 2013 - July 2015

Research Assistant, Ghent University. Continuation on the PARIS project (IWT-SBO). Responsible for text mining – deep linguistic processing – of user-generated content in order to build personalized advertisements from web sources.

February 2014 - June 2014

Visiting Researcher, University of Mannheim. Worked at the Research Group Data and Web Science in the framework of a DAAD Research Project. Supervised by prof. dr. Simone Paolo Ponzetto. Worked on content-based recommender systems and ‘semantifying’ text using Linked Open Data (LOD).

January 2012 - September 2013

Research Assistant, University College Ghent. Worked on two research projects:

1. HENDI HOF project (January – October 2012). Responsible for creating a full-fledged readability prediction system and tool for Dutch.
2. PARIS IWT-SBO project (November 2012 – September 2013). “Personalised advertisements built from web sources” (IWT-SBO), start PhD project.

January 2011 - December 2011

Research Assistant, Utrecht University (0.8 fte). Worked on the STEVIN-funded SoNaR project. Responsible for the manual annotation of 500K text with Dutch semantic roles and the creation of an automatic semantic role labeler. Recruitment and supervision of student assistants.

Research Assistant, University College Ghent (0.2 fte). Worked on the Stylene project (EWI Flanders): “A web service for research in stylometry and readability prediction”.

October 2008 - December 2010

Research Assistant, University College Ghent. Involvement in two STEVIN-funded research projects to consolidate the position of Dutch in the modern information and communication society:

1. Dutch Parallel Corpus Project DPC (October 2008 - June 2009). Responsible for data acquisition and semi-automatic control of each processing step (sentence alignment, tokenization, lemmatization, part-of-speech tagging) for the English-Dutch and Dutch-English language pairs. Supervision and recruitment of student assistants.

2. STEVIN Nederlandstalig Referentiecorpus Project SoNaR (October 2008 - December 2010). Responsible for the acquisition of Flemish text material. Responsible for the manual annotation of a one-million-word subset with coreference information and re-training and improving the automatic coreference resolution system COREA. Supervision and recruitment of student assistants.

3 Teaching Experience

- *Computer-Assisted Language Learning – Ghent University*
 - October-December 2021 - now (Master in Teaching, Master in Speech Therapy, Postgraduate Course Computer-Assisted Language Mediation, elective course, ca. 10 students): coordinator, 7 out of 9 sessions, creation of course material, individual feedback and evaluation.
 - October-December 2019, 2020 (Master in Teaching, elective course, ca. 4 students): coordinator, all sessions, creation of all course material, individual feedback and evaluation.
- *Digital Communication Management – Ghent University*
 - October-December 2016, 2017, 2019 - now (Postgraduate course Computer-Assisted Language Mediation, ca. 10 students): coordinator, 6 out of 9 sessions, creation of course material, feedback and evaluation.
- *Digital Communication – Ghent University*
 - February-May 2023 - now (Bachelor 3 Applied Language Studies, Master in Communication Science, elective course, ca. 60 students): Blended learning. Creation of 4 e-learning environments, provide individual feedback to all students, 2 theoretical sessions and 2 follow-up sessions, evaluation.
 - March-May 2017, 2018, 2019, 2020, October/December 2021 (Bachelor 3 Applied Language Studies, ca. 130 students): Blended learning. Creation of 4 e-learning environments, provide individual feedback to all students, 2 theoretical sessions and 2 follow-up sessions.
- *Introduction to Translation Technology - Ghent University*

October-December 2015 (Postgraduate course Dutch and Translation, ca. 30 students): gave all 6 hands-on sessions.
- *Introduction to Language Technology - University College Ghent*

October-December 2010 and 2013 (Postgraduate Dutch and Translation, ca. 30 students): gave 4 out of 6 hands-on sessions.
- *Introduction to Language Technology - University College Ghent*

March-May 2013 (Bachelor 3, Applied Language Studies, ca. 140 students – all 6 sessions): gave all 6 hands-on sessions.
- *Master's Thesis advisor*

- Dayana Albuin Ruiz (2024). Human-Machine Parity in Neural Machine Translation: Evaluating LLM Effectiveness in Translating Literary Colloquialisms.
- Susanna Danielyan (2024). From Complex to Clear: Assessing AI-Simplified Russian Texts for Second Language Acquisition.
- Femke De Vrieze (2023). Graffiti or gravity? Mispronunciation at its best. Improving English Pronunciation through the language learning app ELSA.
- Amber Verleyen (2022). l'Analyse des émotions présentes dans des tweets néerlandophones au cours des premiers mois de la pandémie de COVID-19.
- Nafal Ossandon Hostens (2021). Automatic Assessment of the Writing Quality and Writing Development of Dutch Learners (co-supervisor, University of Antwerp).
- Paulien Muls (2021). The influence of Dutch L1 writing classes on writing quality. Using AWE and automated text analysis to compare writing quality of language and non-language major students at Ghent University.
- Charlotte Desplenter (2021). L'analyse des sentiments et la détection de la position d'un discours haineux.
- Tessa Cornelis (2021). Stance detection and sentiment analysis in Dutch political tweets and Facebook comments.
- Jasmijn Raes (2020). Are Master students Multilingual Communication sufficiently prepared for the workplace when it comes to writing? An exploratory study into the readability of university and work placement assignments.
- Annaïs Airapetian (2020). Topical and domain-specific frameworks for emotion detection. An experimental cluster analysis of emotions in reality tv.
- Jitse Criquieliion (2020). The influence of topic interest on teens' reading comprehension.
- Liesel Maertens (2020). Human rater versus automated writing evaluation: imminent clash or harmonious fusion?
- Siel Debouver (2020). l'Analyse des sentiments appliquée sur des articles de journaux.
- Tine Dhaene (2020). Demandez-le à Siri: La variation régionale linguistique influence-t-elle sur le fonctionnement de la version Flamande de Siri?
- Senne Van Hoecke (2019). Towards automated précis grading. Linguistic features that might be good predictors: an exploratory study.
- Silke Bekaert (2019). Whether you want it or not, bias does play a role in evaluation". An assessment of the reliability and validity of the writing evaluation method used for the course Nederlands Schrijfvaardigheid at Ghent University.
- Manon Vermeersch (2018). Hotel agréable et bien situé. L'analyse des sentiments au niveau des aspects pour des avis sur des hotels en néerlandais et français.
- Marie D'Hondt (2018). Verspätet – En retard. L'analyse de sentiments appliquée aux tweets au sujet de la SNCB et de la Deutsche Bahn.
- Veera Voordeckers (2018). Un dîner interculturel. Une étude sur l'analyse de sentiments au niveau des aspects en français et russe.

- Guillaume Goemanne (2017). Précise. Research into the possibility of a consistent error typology for the précis text genre as a first step towards AWE.
 - Enzo Neve (2017). Lekker, yummy, délicieux. Towards aspect-based sentiment analysis in Dutch, English and French.
 - Lenie Van Hecke (2016). Verdict de la Cour: Illisible. Rechercher l’influence des connaissances préalables pour juger la lisibilité.
 - Sofie Vandenhoven (2016). What does the bird say? A corpus study on what Twitter can teach us about personality.
 - Margot Dedeyne (2015). LeTÕs get personal. Comparing manually created user profiles with automatic text processing techniques to derive personal interests from tweets.
 - Valerie Dumortier (2014). How can a thorough linguistic analysis contribute to automatic sentiment analysis in Twitter?
 - Rob Braet (2014). Careful what you tweet for! Can we determine personal preferences based on tweets using NLP tools and sentiment analysis.
 - Lana Eeckhout (2013). From conversation to conversion. An explorative study on the adoption of social media and the application of sentiment analysis in Flemish media and web shops.
 - Xenia Van Cotthem (2011). Est-il possible d’entraîner un étiquetteur morphosyntaxique aux textes non-normalisés? Une étude expérimentale du processus de pré-processing appliqué aux textos.”
- Guest lectures, coaching
 - Prepared sessions/workshops for prospecting high school students on “The thrills of NLP” , “Linguistics in the digital age”, “Language technology and studying languages” (2013-2019, 2022, 2024).
 - Coach Think & Talk Honours programme of Ghent University (together with Fien Desmedt) on the topic of Automated Essay Scoring and centralized testing. Podcast available at: <https://open.spotify.com/show/3wkycmfVBXvIdMqxXxSwTd> (spring semester 2024).
 - Explanatory videos which are incorporated in the Ghent University wide learning path on Generative AI for all students and employees.
 - Guest lecture on fine-grained sentiment and emotion analysis for business purposes at KU Leuven, Master Level (2024).
 - Guest lecture on sentiment analysis (together with Ellen De Geyndt) at Antwerp University, Master Level (2023, 2024).

4 Grants and Awards

- 2023: granted a Global Minds Fund – Short Research Stay 2022/2023 grant for a 1-month research stay at Ghent University and foster collaboration with a South-African PhD student.

- 2022: granted a BOF starting grant as newly appointed tenure-track professor. One full-time PhD student for 4 years.
- 2020: granted an FWO award for a senior research project fundamental research (co-promotor) - G013820N. “ENCORE: Investigating cross-document EveNt Coreference REsolution for Dutch”. One full-time PhD student for 4 years.
- 2020: granted an FWO Travel Fellowship to attend the LREC2020 conference at Marseille, France (conference cancelled because of COVID-19).
- 2016: granted an FWO Travel Fellowship to attend the NAACL2016 and SemEval2016 conference at San Diego, USA.
- 2014: granted a DAAD Contact Fellowship (Kontaktstipendien) for a 5-month research stay at the University of Mannheim to establish international collaborations between universities.

5 Organisation of Conferences, Workshops and Shared Tasks

- Main organiser of the Fourteenth WASSA workshop (2024)
- Co-organiser of the first shared task on Multilingual Grammatical Error Detection – MultiGED (2023)
- Co-organiser of the Thirteenth WASSA workshop (2023)
- Co-organiser of the workshop *The Voice of the Customer*, SentEMO project (2022).
- Co-organiser of the workshop *Feeling EmotioNL: Automatic emotion detection in Dutch texts*, on the occasion of Luna De Bruyne’s public PhD defence (2022).
- Co-organizer of the Twelfth WASSA workshop (2022)
- Co-organizer of the Computational Linguistics in the Netherlands Conference (2021).
- Main organizer of the Eleventh WASSA workshop (2021)
- Co-organizer of the WASSA 2021 Shared Task: Predicting Empathy and Emotion in Reaction to News Stories.
- Co-organizer of the Tenth WASSA workshop (2019)
- Co-organizer of the WASSA 2018 Implicit Emotions Shared Task. WASSA stands for Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis.
- Co-organizer of SemEval 2016 - Task 5 Aspect-Based Sentiment Analysis
- Co-organiser of the PARIS project workshop (2016)
- Co-organiser of the Atila 2014 conference (2014)
- Co-organiser of the Atila 2010 conference (2014)
- Co-organiser of the DPC (“Dutch Parallel Corpus”) project workshop (2009)

6 Reviewing and Program Committees

- Full paper reviewing
 - 2024: senior area chair LREC-COLING 2024, demonstration chair EMNLP 2024, reviewer (ACL 2024 main and demo, CMC corpora 2024), PC member of the BEA workshop.
 - 2023: area chair for 2 tracks (Multidisciplinary and other NLP Applications & Sentiment Analysis and Argument Mining) at EACL 2023. Reviewer for EMNLP 2023 Demo track, EACL 2023 Student Research Workshop, ACL 2023, WASSA 2023, CMC corpora 2023, Computational Linguistics Journal.
 - 2022: Reviewer for COLING 2022, Journal of Natural Language Engineering, Journal of IEEE Transactions on Affective Computing. Academic Editor PLOS ONE journal. PC member of the BEA workshop.
 - 2021: area chair Sentiment, Stylistic Analysis, Argumentation Mining for EMNLP2021. Reviewer for EACL 2021, IJCAI 2021, ACL 2021, ECML/PKDD 2021, the Northern European Journal of Language Technology, a Routledge book proposal. PC member of the BEA workshop.
 - 2020: Reviewer for LREC2020, IJCAI 2020, EMNLP 2020, Language@Internet journal, Future Generation Computer Systems journal, PC member of the BEA 2020 workshop.
 - 2019: Reviewer for Applied Computing and Informatics journal. PC member of IJCAI 2019, ACL 2019, WASSA 2019, ESSLLI 2019 Student Session, RANLP 2019, EMNLP-IJCNLP 2019.
 - 2018: Reviewer for LREC2018, the ISI-NLP2 workshop, PLOS ONE journal, SemEval 2018. PC member of IJCAI-ECAI 2018, ESSLLI 2018 Student Session, COLING 2018, CMC-corpora 2018, WASSA 2018. Guest Editor for PLOS ONE journal.
 - 2017: PC member of The 3rd International Workshop on Natural Language Processing for Informal Text, PC member of the EACL 2017 Student Research Workshop, Reviewer for HUSO2017, RANLP2017, Journal of Information Processing and Management.
 - 2016: Reviewer for EMOSEDE Session at HUSO2016 conference, PC member of the ISWC Poster & Demos Session, Reviewer for LREC2016, SemEval2016, PC member of The 2nd International Workshop on Natural Language Processing for Informal Text, Dutch data provider for SemEval 2016 Task 5 on Aspect-based Sentiment Analysis.
 - 2015: Reviewer for SemEval2015, RANLP2015, Journal of Natural Language Engineering.
 - 2014: Reviewer for LREC2014, Journal of Natural Language Engineering.
 - 2013: Reviewer for RANLP2013, secondary reviewer for CICLING2013, ACL2013, SocInfo2013.
 - 2012: PC member of Lingvisticae Investigationes (Special Issue on SMS Communication: A linguistic approach), secondary reviewer for PITR2012.

- Project reviewing
 - 2021: Review of a thesis proposal of the North-West University (Research unit: Language and Literature in the South African Context).
 - 2016: Scientific reviewer for the Flanders Innovation & Entrepreneurship government agency (VLAIO).

7 PhD Guidance and Commissions

- Guidance
 - Current
 - * Promotor of Joni Kruijsbergen who prepares a PhD on Dutch Automated Writing Evaluation with a focus on L2 Dutch Learners.
 - * Promotor of Serafina Van Geertruyen who prepares a PhD on Text-level automated writing evaluation for Dutch (TUTORED).
 - * Co-promotor of Pranaydeep Singh who prepares a PhD on Low-Shot Cross-Lingual Knowledge Transfer (LoCKe).
 - * Co-promotor of Loïc De Langhe who prepares a PhD in the framework of the ENCORE project (Investigating cross-document Event Coreference Resolution for Dutch).
 - Defended
 - * Co-promotor of Luna De Bruyne who defended her PhD “Feeling EmotioNL: Automatically detecting emotions in Dutch texts” on 2 December 2022. (Currently tenure track professor at Antwerp University)
 - * Co-promotor of Bram Vanroy who defended his PhD “Syntactic difficulties in translation” on 27 May 2021. (Currently researcher at KU Leuven and ILT)
 - Member of the Doctoral Advisory Committee (DBC) of Tess Dejaeghere who prepares a PhD on multilingual NLP for historical literature.
- Examination Commission
 - Jury member, examination commission of Semere Kiros Bitew (UGent - Computer Science Engineering): Language Model Adaptation with Applications in AI for Education. (2024)
 - Jury member, examination commission of Amir Hadifar (UGent - Computer Science Engineering): Towards Efficient NLP Models: Strategies for Improving Data Selection, Training, and Inference. (2023)
 - Jury member, examination commission of Ekaterina Loginova (UGent - Business Economics): Studies on Advanced Natural Language Processing Representations and Deep Learning in Decision Making Systems. (2022).
 - Jury member, examination commission of Gilles Jacobs (UGent - Linguistics): Extracting Fine-Grained Events and Sentiment from Economic News. (2021).
 - Jury member, examination commission of Nils Smeuninckx (UGent - Linguistics): Dear Stakeholder. Exploring the language of sustainability reporting, a closer look at readability, sentiment and perception. (2018)

- Jury member, examination commission of Suzanne Kleijn (Utrecht University - Linguistics): Clozing in on readability. How linguistic features affect and predict text comprehension and on-line processing. (2018)
- Rapporteur, examination commission of Arda Tezcan (UGent - Linguistics): Informative Quality Estimation of Machine Translation Output. (2018)
- Jury member, examination commission of Cynthia Van Hee (UGent - Linguistics): Can machines sense irony? Exploring automatic irony detection on social media. (2017)
- Rapporteur, examination commission of Sabien Hanouille (UGent - Translation Studies): Translating documentaries: Does the integration of a bilingual glossary of domain-specific terminology into the translation process reduce the translator’s workload? (2017)
- Rapporteur, examination commission of Joke Daems (UGent - Translation Studies): A translation robot for each translator? A comparative study of manual translation and post-editing of machine translations: process, quality and translator attitude. (2016)

8 Research interns

- PhD

- Suna Şeyma Uçar, HiTZ Research Center, University of the Basque Country: 3 month research stay to collaborate on exploring LLMs’ potential for education and, more specifically, question generation (March-June 2024).
- Johannes Sibeko, Doctor of Philosophy, Language and Literature with a special focus on Digital Humanities: short research stay to collaborate on ”Measuring Text Readability in Sesotho” (February 2023).

- Master

- Joni Kruijsbergen, Advanced Master in AI - Speech and Language Technology, KU Leuven: research internship on automated writing evaluation of Dutch L2 learners (Spring 2023).
- Nafal Ossandon Hostens, Master of Linguistics, Antwerp University: research internship on automatic text analysis of writing products of Dutch L2 learners (Spring 2021).

9 Additional Training

- 2023: *Agile leadership for PhD supervisors* (UGent)
- 2022: *Research Track for newly appointed professors* (UGent)
- 2018: *Research School ‘Facets of Writing and Writing Research: Substantive and Methodological Issues’* (Sig Writing 2018, EARLI special interest group on writing)

- 2017: *Online Marketing Fundamentals Certificate*, *Google Analytics Certificate* (Google DMA), *Basisdocententraining* (UGent), *Academic Leadership Programme for Postdocs* (True Colours cvba, UGent postdoc training)
- 2016: *Creativity, Pitching & Resilience* (Streetwize, UGent postdoc training)
- 2015: *Technical Writing* (Teka Info Pilots)
- 2013: *Project Management*, *Effective Slide Design* (UGent Doctoral Schools)
- 2012: *Reference and Coherence in Discourse* and *Using Bayesian methods to model cognition* (LOT Summer School); *Introduction to R* and *Introductory Statistics* (IPVW – UGent Doctoral Schools)
- 2009: *Introduction to Language Technology*, *Advanced Methods for Corpus Linguistics* (UGent Doctoral Schools).

10 Membership and Service

- 2022-: Member of the Faculty board, Faculty of Arts and Philosophy, Ghent University
- 2018-: Member of the Departmental Board, Department of Translation, Interpreting and Communication, Ghent University
- 2018-2021: Member of the Study Programme Committee, Department of Translation, Interpreting and Communication, Ghent University
- 2018-2020: Member of DH Research Lab, Ghent Centre for Digital Humanities, Ghent University

11 Publications

- Journal articles
 - A1 (ISI Web of Science)
 - * Van Ostaeyen, Sofie, Loic De Langhe, **Orphée De Clercq**, Mieke Embo, Tammy Schellens, and Martin Valcke (2023). Automating the Identification of Feedback Quality Criteria and the CanMEDS Roles in Written Feedback Comments Using Natural Language Processing. *PERSPECTIVES ON MEDICAL EDUCATION*, 12(1). <https://doi.org/10.5334/pme.1056>
 - * Van Ostaeyen, Sofie, Mieke Embo, Tijs Rotsaert, **Orphée De Clercq**, Tammy Schellens, and Martin Valcke (2023). A Qualitative Textual Analysis of Feedback Comments in Eportfolios: Quality and Alignment with the CanMEDS Roles. *PERSPECTIVES ON MEDICAL EDUCATION*, 12(1). <https://doi.org/10.5334/pme.1050>
 - * De Langhe, Loic, Thierry Desot, **Orphée De Clercq** and Veronique Hoste (2023). A benchmark for Dutch end-to-end cross-document event coreference resolution. *ELECTRONICS*, 12(4). <https://doi.org/10.3390/electronics12040850>

- * Singh, Pranaydeep **Orphée De Clercq** and Els Lefever (2023). Distilling monolingual models from large multilingual transformers. *ELECTRONICS*, 12(4). <https://doi.org/10.3390/electronics12041022>
- * Colruyt Camiel, **Orphée De Clercq**, Thierry Desot and Veronique Hoste (2023) EventDNA: a dataset for Dutch news event extraction as a basis for news diversification. *LANGUAGE RESOURCES AND EVALUATION*. <https://doi.org/10.1007/s10579-022-09623-2>
- * De Langhe, Loic, **Orphée De Clercq**, and Veronique Hoste (2023) Constructing a Cross-Document Event Coreference Corpus for Dutch. *LANGUAGE RESOURCES AND EVALUATION*. <https://doi.org/10.1007/s10579-022-09597-1>.
- * Joris, Glen, Stefaan Vercoetere, **Orphée De Clercq**, Kristin Van Damme, Peter Mechant, and Lieven De Marez (2023) Nudging towards Exposure Diversity: Examining the Effects of News Recommender Design on Audiences' News Exposure Behaviours and Perceptions. *DIGITAL JOURNALISM*, 1–22. <https://doi.org/10.1080/21670811.2022.2106445>.
- * De Bruyne, Luna, **Orphée De Clercq**, and Veronique Hoste (2021) Mixing and Matching Emotion Frameworks: Investigating Cross-Framework Transfer Learning for Dutch Emotion Detection. *ELECTRONICS* 10 (21). <https://doi.org/10.3390/electronics10212643>.
- * Matos Veliz, Claudia, **Orphée De Clercq**, and Veronique Hoste (2021) Is Neural Always Better? SMT versus NMT for Dutch Text Normalization. *EXPERT SYSTEMS WITH APPLICATIONS* 170. <https://doi.org/10.1016/j.eswa.2020.114500>.
- * De Bruyne, Luna, **Orphée De Clercq**, and Veronique Hoste (2021) Annotating Affective Dimensions in User-Generated Content: Comparing the Reliability of Best-Worst Scaling, Pairwise Comparison and Rating Scales for Annotating Valence, Arousal and Dominance. *LANGUAGE RESOURCES AND EVALUATION* 55 (4): 1017–45. <https://doi.org/10.1007/s10579-020-09524-2>.
- * Bouko, Catherine, July De Wilde, Sofie Decock, **Orphée De Clercq**, Valentina Manchia, and David Garcia (2021) Reactions to Brexit in Images: A Multimodal Content Analysis of Shared Visual Content on Flickr. *VISUAL COMMUNICATION* 20 (1): 4–33. <https://doi.org/10.1177/1470357218780530>.
- * Vanroy, Bram, **Orphée De Clercq**, and Lieve Macken (2019) Correlating Process and Product Data to Get an Insight into Translation Difficulty.” *PERSPECTIVES-STUDIES IN TRANSLATION THEORY AND PRACTICE* 27 (6): 924–41. <https://doi.org/10.1080/0907676X.2019.1594319>.
- * De Sutter, Gert, Bert Cappelle, **Orphée De Clercq**, Rudy Loock, and Koen Plevoets (2017) Towards a Corpus-Based, Statistical Approach of Translation Quality: Measuring and Visualizing Linguistic Deviance in Student Translations. Edited by Geoffrey S. Koby and Isabel Lacruz. *LINGUISTICA ANTVERPIENSIA NEW SERIES-THEMES IN TRANSLATION STUDIES* 16: 25–39.

- * Daems, Joke, **Orphée De Clercq**, and Lieve Macken (2017) Translationese and Post-Editese: How Comparable Is Comparable Quality? Edited by Geoffrey S. Koby and Isabel Lacruz. *LINGUISTICA ANTVERPIENSIA NEW SERIES-THEMES IN TRANSLATION STUDIES* 16: 89–103.
- * Bouko, Catherine, Laura Calabrese, and **Orphée De Clercq** (2017) Cartoons as Interdiscourse: A Quali-Quantitative Analysis of Social Representations Based on Collective Imagination in Cartoons Produced after the Charlie Hebdo Attack. *DISCOURSE CONTEXT & MEDIA* 15: 24–33.
<https://doi.org/10.1016/j.dcm.2016.12.001>.
- * **De Clercq, Orphée**, and Veronique Hoste (2016) All Mixed up? Finding the Optimal Feature Set for General Readability Prediction and Its Application to English and Dutch. *COMPUTATIONAL LINGUISTICS* 42 (3): 457–90.
<https://doi.org/10.1162/COLL.a.00255>.
- * Schulz, Sarah, Guy De Pauw, **Orphée De Clercq**, Bart Desmet, Veronique Hoste, Walter Daelemans, and Lieve Macken (2016) Multi-Modular Text Normalization of Dutch User-Generated Content. *ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY* 7 (4).
<https://doi.org/10.1145/2850422>.
- * **De Clercq, Orphée**, Veronique Hoste, Bart Desmet, Philip van Oosten, Martine De Cock, and Lieve Macken (2014) Using the Crowd for Readability Prediction. *NATURAL LANGUAGE ENGINEERING* 20 (3): 293–325.
<https://doi.org/10.1017/S1351324912000344>.
- * Macken, Lieve, **Orphée De Clercq**, and Hans Paulussen (2011) Dutch Parallel Corpus: A Balanced Copyright-Cleared Parallel Corpus. *META* 56 (2): 374–90.

– VABB/A2

- * De Langhe, Loic, **Orphée De Clercq**, and Veronique Hoste (2022). Towards Fine(r)-grained Identification of Event Coreference Resolution Types. *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 12: 183–205.
- * **De Clercq, Orphée**, Gert De Sutter, Rudy Loock, Bert Cappelle, and Koen Plevoets (2021) Uncovering Machine Translationese Using Corpus Analysis Techniques to Distinguish between Original and Machine-Translated French. *TRANSLATION QUARTERLY* (101): 21–45.
- * De Bruyne, Luna, **Orphée De Clercq**, and Veronique Hoste (2021) Prospects for Dutch Emotion Detection: Insights from the New EmotioNL Dataset. *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 11: 231–255.
- * **De Clercq, Orphée**, Luna De Bruyne, and Veronique Hoste (2020) News Topic Classification as a First Step towards Diverse News Recommendation. *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 10: 37–55.
- * Van Hee, Cynthia, Marjan Van de Kauter, **Orphée De Clercq**, Els Lefever, Bart Desmet, Veronique Hoste (2017). Noise or Music? Investigating the Usefulness of Normalisation for Robust Sentiment Analysis on Social Media Data. *TRAITEMENT AUTOMATIQUE DES LANGUES* 58 (1): 63–87.

– A2

- * De Langhe, Loic, Aaron Maladry, Bram Vanroy, Luna De Bruyne, Pranaydeep Singh, Els Lefever, and **Orphée De Clercq** (2024). Benchmarking Zero-Shot Text Classification for Dutch. *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 13: 63–90.
- * Kruijsbergen, Joni, Serafina Van Geertruyen, Veronique Hoste and **Orphée De Clercq** (2024). Exploring LLMs’ Capabilities for Error Detection in Dutch L1 and L2 Writing Products. *COMPUTATIONAL LINGUISTICS IN THE NETHERLANDS JOURNAL* 13: 173–191.
- * De Bruyne, Luna, Toni G.L.A. van der Meer, **Orphée De Clercq**, and Veronique Hoste (2024). Using State-of-the-Art Emotion Detection Models in a Crisis Communication Context.” *COMPUTATIONAL COMMUNICATION RESEARCH* 6 (1): 1–45. <https://doi.org/10.5117/ccr2024.1.1.debr>.

– A3

- * Montero Perez, Maribel, **Orphée De Clercq**, Piet Desmet, Geert Peeters and Serge Verlinde (2009) Dutch parallel corpus: un nouveau corpus parallèle multilingue disponible en ligne. *Romanesque*, 34 (4), 2-8.

– A4

- * **Orphée De Clercq** (2016). Van tekstherkenning naar tekstbegrip. *Computers en plagiaatdetectie*. Karakter, 53.

• Books

– VABB/B2

- * Vanroy, Bram **Orphée De Clercq**, Arda Tezcan, Joke Daems and Lieve Macken (2021). Metrics of syntactic equivalence to assess translation difficulty. In M. Carl (ed.), *Explorations in empirical translation process research*. Springer.
- * Vandeghinste, Vincent Adeline Müller, Thomas François, and **Orphée De Clercq** (2021). Easy Language in Belgium. In C. Lindholm and U. Vanhatalo (eds.), *Handbook of Easy Languages in Europe*. Frank & Timme.
- * Joris, Glen, Camiel Colruyt, Judith Vermeulen, Stefaan Vercoutere, Frederik De Grove, Kristin Van Damme, **Orphée De Clercq**, Cynthia Van Hee, Lieven De Marez, Veronique Hoste, Eva Lievens, Toon De Pessemier, and Luc Martens (2020) News Diversity and Recommendation Systems: Setting the Interdisciplinary Scene. In *Privacy and Identity Management: Data for Better Living: AI and Privacy*, ed by. Michael Friedewald, Melek Önen, Eva Lievens, Stephan Krenn, and Samuel Fricker, 576:90–105. Springer. https://doi.org/10.1007/978-3-030-42504-3_7.
- * Daelemans, Walter, **Orphée De Clercq** and Véronique Hoste (2018). *STYLENE: an Environment for Stylometry and Readability Research for Dutch*. In *CLARIN in the Low Countries*, edited by Jan Odijk and Arjan Van Hessen, 195–210. London: Ubiquity Press.

– B2

- * Van Hee, Cynthia, Veronique Hoste, Lieve Macken, **Orphée De Clercq**, and Els Lefever (2024). “Hoe Leert Een Computer Taal?” In *Taaltechnologie Ontrafeld*, ed by. Cynthia Van Hee and Veronique Hoste, 30–55. Kalmthout: Pelckmans.
- * Van Hee, Cynthia, Veronique Hoste, **Orphée De Clercq**, Els Lefever, and Lieve Macken (2024). “Wat Is Taaltechnologie?” In *Taaltechnologie Ontrafeld*, ed by. Cynthia Van Hee and Veronique Hoste, 12–16. Kalmthout: Pelckmans.
- * Van Hee, Cynthia, Veronique Hoste, **Orphée De Clercq**, Els Lefever, and Lieve Macken (2024). “Waarom Is Taal Zo Moeilijk?” In *Taaltechnologie Ontrafeld*, ed by. Cynthia Van Hee and Veronique Hoste, 17–29. Kalmthout: Pelckmans.
- * Van Hee, Cynthia, Veronique Hoste, and **Orphée De Clercq** (2024). “Sentiment-En Emotieanalyse.” In *Taaltechnologie Ontrafeld*, ed by. Cynthia Van Hee and Veronique Hoste, 102–112. Kalmthout: Pelckmans.
- * Daems, Joke, Lieve Macken, **Orphée De Clercq**, and Cynthia Van Hee (2024). “Zelf Aan de Slag.” In *Taaltechnologie Ontrafeld*, ed by. Cynthia Van Hee and Veronique Hoste, 123–136. Kalmthout: Pelckmans.
- * **De Clercq, Orphée** and Véronique Hoste (2014). Hoe meetbaar is leesbaarheid? In S. Evenepoel, P. Goethals and L. Jooen (eds.), *Beschouwingen uit een talenhuis*, 147-155. Academia Press, Ghent, Belgium.
- * Desmet, Bart, **Orphée De Clercq**, Marjan Van de Kauter, Sarah Schulz, Cynthia Van Hee and Véronique Hoste (2014). *Taaltechnologie 2.0: sentimentanalyse en normalisatie*. In S. Evenepoel, P. Goethals and L. Jooen (eds.), *Beschouwingen uit een talenhuis*, 157-161. Academia Press, Ghent, Belgium.
- * Macken, Lieve, **Orphée De Clercq**, Bart Desmet and Véronique Hoste (2014). Dutch Parallel Corpus en SoNaR. In S. Evenepoel, P. Goethals and L. Jooen (eds.), *Beschouwingen uit een talenhuis*, 163-170. Academia Press, Ghent, Belgium.

– VABB/B3

- * Balahur, Alexandra, Roman Klinger, Veronique Hoste, Carlo Strapparava, and **Orphée De Clercq**, eds. (2019) *Proceedings of the 10th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Minneapolis, USA: Association for Computational Linguistics.

– B3

- * **De Clercq, Orphée**, Valentin Barriere, Jeremy Barnes, Roman Klinger, Joao Sedoc and Shabnam Tafreshi eds. (2023). *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Toronto, Canada: Association for Computational Linguistics.
- * Aletras, Nikos and **Orphée De Clercq** (2024). *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. Valletta, Malta: Association for Computational Linguistics
- * Barnes, Jeremy, **Orphée De Clercq** and Roman Klinger, eds. (2023). *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity,*

Sentiment and Social Media Analysis. Toronto, Canada: Association for Computational Linguistics.

- * Barnes, Jeremy, **Orphée De Clercq**, Valentin Barriere, Shabnam Tafreshi, Sawsan Alqahtani, Joao Sedoc, Roman Klinger, Alexandra Balahur, eds. (2022) Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Dublin, Ireland: Association for Computational Linguistics.
- * **De Clercq, Orphée**, Alexandra Balahur, Joao Sedoc, Valentin Barriere, Shabnam Tafreshi, Sven Buechel and Véronique Hoste, eds. (2021) Proceedings of the 11th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Online: Association for Computational Linguistics.

- Conference Proceedings

- P1 (ISI Web of Science)

- * De Geyndt, Ellen, **Orphée De Clercq**, Cynthia Van Hee, Els Lefever, Pranaydeep Singh, Olivier Parent, and Veronique Hoste (2022) SentEMO: A Multilingual Adaptive Platform for Aspect-Based Sentiment and Emotion Analysis. In *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, 51–61. Association for Computational Linguistics (ACL).
 - * De Bruyne, Luna, **Orphée De Clercq**, and Veronique Hoste (2020) An Emotional Mess! Deciding on a Framework for Building a Dutch Emotion-Annotated Corpus. In *Proceedings of the 12th Language Resources and Evaluation Conference (LREC 2020)*, 1636–44. European Language Resources Association (ELRA).
 - * **De Clercq, Orphée**, and Senne Van Hoecke (2020) An Exploratory Study into Automated Précis Grading. In *Proceedings of the 12th Language Resources and Evaluation Conference (LREC 2020)*, 397–404. European Language Resources Association (ELRA).
 - * **De Clercq, Orphée**, and Véronique Hoste (2016), Rude waiter but mouth-watering pastries! An exploratory study into Dutch Aspect-Based Sentiment Analysis. In *Proceedings of the 10th Language Resources and Evaluation Conference (LREC2016)*, 2910–2917. European Language Resources Association (ELRA).
 - * **De Clercq, Orphée**, Sarah Schulz, Bart Desmet and Véronique Hoste (2014), Towards Shared Datasets for Normalization Research. In *Proceedings of the 9th Language Resources and Evaluation Conference (LREC2014)*, 1218–1223. European Language Resources Association (ELRA).
 - * **De Clercq, Orphée**, Véronique Hoste and Paola Monachesi (2012) Evaluating Automatic Cross-Domain Semantic Role Annotation. In *Proceedings of the 8th Language Resources and Evaluation Conference (LREC2012)*, 88–93. European Language Resources Association (ELRA).
 - * Maaske Treurniet, **Orphée De Clercq**, Henk van den Heuvel and Nelleke Oostdijk (2012), Collecting a Corpus of Dutch SMS. *Proceedings of the 8th*

Language Resources and Evaluation Conference (LREC2012), 2268–73. European Language Resources Association (ELRA).

- * **De Clercq, Orphée** and Maribel Montero Perez (2010), Data Collection and IPR in Multilingual Parallel Corpora: Dutch Parallel Corpus. In *Proceedings of the 7th Language Resources and Evaluation Conference (LREC2010)*, 3383–3388. European Language Resources Association (ELRA).
- * Martin Reynaert, Nelleke Oostdijk, **Orphée De Clercq**, Henk van den Heuvel and Franciska de Jong (2010), Balancing SoNaR: IPR versus Processing Issues in a 500-Million-Word Written Dutch Reference Corpus. In *Proceedings of the 7th Language Resources and Evaluation Conference (LREC2010)*, 2693–2698. European Language Resources Association (ELRA).

– VABB/C1

- * Van Hee, Cynthia, **Orphée De Clercq**, and Veronique Hoste. (2021) Exploring Implicit Sentiment Evoked by Fine-Grained News Events. In *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (EACL 2021)*, 138–148. Association for Computational Linguistics.
- * Desot, Thierry, **Orphée De Clercq**, and Veronique Hoste (2021) Event Prominence Extraction Combining a Knowledge-Based Syntactic Parser and a BERT Classifier for Dutch. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, 346–357. INCOMA. doi:10.26615/978-954-452-072-4.040.
- * De Bruyne, Luna, **Orphée De Clercq** and Véronique Hoste (2021). Emotional RobBERT and Insensitive BERTje: Combining Transformers and Affect Lexica for Dutch Emotion Detection. In *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2021)*.
- * Tafreshi, Shabnam, **Orphée De Clercq**, Valentin Barriere, Sven Buechel, Joao Sedoc and Alexandra Balahur (2021). WASSA 2021 Shared Task: Predicting Empathy and Emotion in Reaction to News Stories. In *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2021)*.
- * **De Clercq, Orphée**, and Veronique Hoste. 2020. It’s Absolutely Divine! Can Fine-Grained Sentiment Analysis Benefit from Coreference Resolution? In *Proceedings of the Third Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC 2020)*, 11–21. Association for Computational Linguistics.
- * Klinger, Roman **Orphée De Clercq**, Saif M. Mohammad and Alexandra Balahur (2018). IEST: WASSA-2018 Implicit Emotions Shared Task. In *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*.
- * De Langhe, Loic, **Orphée De Clercq**, and Veronique Hoste (2022) Investigating Cross-Document Event Coreference for Dutch. In *Proceedings of the Fifth Workshop on Computational Models of Reference, Anaphora and Coreference*, 88–98. Association for Computational Linguistics.

- * Matoz Veliz, Claudia, **Orphée De Clercq** and Véronique Hoste (2019). Benefits of Data Augmentation for NMT-Based Text Normalization of User-Generated Content. Proceedings of the 2019 EMNLP-IJCNLP Workshop W-NUT: The 5th Workshop on Noisy User-Generated Text.
- * Colruyt, Camiel, **Orphée De Clercq** and Véronique Hoste (2019). Leveraging syntactic parsing to improve event annotation matching. Proceedings of the 2019 EMNLP-IJCNLP Workshop AnnoNLP: The First Workshop on Aggregating and Analysing Crowdsourced Annotations for NLP.
- * Matos Veliz, Claudia, **Orphée De Clercq** and Véronique Hoste (2019). Comparing MT Approaches for Text Normalization. Proceedings of the 12th International Conference on Recent Advances in Natural Language Processing (RANLP2019).
- * De Bruyne, Luna, **Orphée De Clercq** and Véronique Hoste (2019). Towards an Empirically-Grounded Framework for Emotion Analysis. Proceedings of the Second International Conference on Human and Social Analytics (HUSO).
- * Vanroy, Bram, **Orphée De Clercq** and Lieve Macken (2018). Predicting difficulty in translation?: a pilot study. In T. Read, S. Montaner, & B. Sedano (Eds.), Technological innovation for specialized linguistic domains: languages for digital lives and cultures, proceedings of TISLID'18 (pp. 163-174). Presented at the Third conference on Technological Innovation for Specialized Linguistic Domains (TISLID'18), Beau Bassin, Mauritius: Éditions Universitaires Européennes.
- * **De Clercq, Orphée** , Els Lefever, Gilles Jacobs, Tjil Carpels and Véronique Hoste (2017). Towards an integrated pipeline for aspect-based sentiment analysis in various domains. Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis at EMNLP 2017.
- * **De Clercq, Orphée** (2016). The many aspects of fine-grained sentiment analysis. An overview of the task and its main challenges. Proceedings of the Second International Conference on Human and Social Analytics (HUSO). ISSN 2519-8351.
- * Vandenhoven, Sofie and **Orphée De Clercq** (2016). What does the bird say? Exploring the Link Between Personality and Language Use in Dutch Tweets. Proceedings of HUSO 2016, The Second International Conference on Human and Social Analytics. ISSN 2519-8351.
- * Pontiki, Maria, Dimitrios Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, **Orphée De Clercq**, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukakevitch, Evgeny Kotelnikov, Nuria Bel, Salud María Jiménez-Zafra and Gülşen Eryiğit (2016). SemEval-2016 Task 5: Aspect Based Sentiment Analysis. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016) at NAACL-HTL.
- * **De Clercq, Orphée** , Marjan Van de Kauter, Els Lefever and Véronique Hoste (2015), LT3: Applying Hybrid Terminology Extraction to Aspect-Based Sentiment Analysis. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval-2015) at NAACL.

- * **De Clercq, Orphée**, Sven Hertling, Véronique Hoste, Simone Paolo Ponzetto and Heiko Paulheim (2014), Identifying Disputed Topics in the News. Proceedings of the LD4KD workshop at ECML/PKDD2014. CEUR-1232.
- * Van Hee, Cynthia, Marjan Van de Kauter, **Orphée De Clercq**, Els Lefever and Véronique Hoste (2014), LT3: Sentiment Classification in user-generated content using a rich feature set. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval-2014), Dublin, Ireland.
- * **De Clercq, Orphée** , Sarah Schulz, Bart Desmet, Els Lefever and Véronique Hoste (2013), Normalisation of Dutch User-Generated Content. Proceedings of the 9th International Conference on Recent Advances in Natural Language Processing (RANLP2013).
- * **De Clercq, Orphée**, Véronique Hoste and Iris Hendrickx (2011), Cross-Domain Dutch Coreference Resolution. Proceedings of the 8th International Conference on Recent Advances in Natural Language Processing (RANLP 2011).
- * Hendrickx, Iris, **Orphée De Clercq** and Véronique Hoste (2011), Analysis and Reference Resolution of Bridge Anaphora across Different Genres. Lecture Notes in Artificial Intelligence, 7099. Springer – Verlag.

– C1

- * De Langhe, Loic, **Orphée De Clercq**, and Veronique Hoste (2023). Leveraging Structural Discourse Information for Event Coreference Resolution in Dutch. In Proceedings of the 4th Workshop on Computational Approaches to Discourse (CODI 2023), 48–53. Association for Computational Linguistics.
- * Volodina, Elena, Chris Bryant, Andrew Caines, **Orphée De Clercq**, Jennifer-Carmen Frey, Elizaveta Ershova, Alexander Rosen and Olga Vinogradova (2023). MultiGED-2023 shared task at NLP4CALL: Multilingual Grammatical Error Detection. In *emphProceedings of the 12th Workshop on NLP for Computer Assisted Language Learning*, 1–16. Torshavn, Faroe Islands: LiU Electronic Press.
- * De Langhe, Loic, **Orphée De Clercq** and Veronique Hoste (2023). What Does BERT actually Learn about Event Coreference? Probing Structural Information in a Fine-Tuned Dutch Language Model. In *Proceedings of the Fourth Workshop on Insights from Negative Results in NLP*, 53–58. Dubrovnik, Croatia: Association of Computational Linguistics.
- * Sibeko, Johannes and **Orphée De Clercq** (2023). A Corpus-Based List of Frequently Used Words in Sesotho. In *Proceedings of the Fourth Workshop on Resources for African Indigenous Languages (RAIL 2023)*, 32–41. Dubrovnik, Croatia: Association for Computational Linguistics.
- * De Bruyne, Luna, Pranaydeep Singh, **Orphée De Clercq**, Els Lefever, and Veronique Hoste (2022) How Language-Dependent Is Emotion Detection? Evidence from Multilingual BERT. In *Proceedings of The 2nd Workshop on Multilingual Representation Learning (MRL)*, 76–85. Association for Computational Linguistics.
- * Desot, Thierry, **Orphée De Clercq**, and Veronique Hoste (2022) A Hybrid Knowledge and Transformer-Based Model for Event Detection with Automatic

Self-Attention Threshold, Layer and Head Selection. In *Proceedings of the 5th Workshop on Challenges and Applications of Automated Extraction of Sociopolitical Events from Text*. Association for Computational Linguistics.

- * De Bruyne, Luna, Akbar Karimi, **Orphée De Clercq**, Andrea Prati, and Veronique Hoste (2022) Aspect-Based Emotion Analysis and Multimodal Coreference: A Case Study of Customer Comments on Adidas Instagram Posts. In *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, 574–580. European Language Resources Association (ELRA).
- * Singh, Pranaydeep, **Orphée De Clercq**, and Els Lefever (2022) Investigating the Quality of Static Anchor Embeddings from Transformers for Under-Resourced Languages. In *Proceedings of the 1st Annual Meeting of the ELRA/ISCA Special Interest Group on Under-Resourced Languages*, edited by Maite Melero, Sakriani Sakti, and Claudia Soria, 176–84. Marseille, France: European Language Resources Association (ELRA).
- * De Bruyne, Luna **Orphée De Clercq** and Véronique Hoste (2018). LT3 at SemEval-2018 Task 1: A classifier chain to detect emotions in tweets. In *Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018)*. at NAACL 2018.
- * **De Clercq, Orphée** , Michael Schuhmacher, Simone Paolo Ponzetto and Véronique Hoste (2014), Exploiting FrameNet for Content-Based Book Recommendation. Proceedings of the CBRecsys workshop at ACM RecSys2014. CEUR-1245.

– Dissertation

- * **Orphée De Clercq** (2015), Tipping the scales. Exploring the added value of deep semantic processing on readability prediction and sentiment analysis. Ghent University, Faculty of Arts and Philosophy.

– Other

- * **De Clercq, Orphée**, Bart Desmet and Véronique Hoste (2014). Guidelines for Normalizing Dutch and English User Generated Content. LT3 Technical Report LT3 14.01. Version 1.0
- * Oostdijk, Nelleke, Véronique Hoste, Franciska de Jong, Martin Reynaert, **Orphée De Clercq**, Bart Desmet, Henk van den Heuvel (2014). SoNaR-500. Centrale voor Taal- en Spraaktechnologie, ISBN 9789491325014.
- * Oostdijk, Nelleke, Martin Reynaert, Véronique Hoste, Henk van den Heuvel, Paola Monachesi, Bart Desmet, **Orphée De Clercq**, Ineke Schuurman (2014). SoNaR-1. Centrale voor Taal- en Spraaktechnologie, ISBN 9789491325021.
- * Oostdijk, Nelleke, Martin Reynaert, Véronique Hoste, Henk van den Heuvel, **Orphée De Clercq**, Eric Sanders (2014). SoNaR Nieuw Media Corpus. Centrale voor Taal- en Spraaktechnologie, ISBN 9789491325038.
- * Desmet, Piet, Willy Vandeweghe, Hans Paulussen, Lieve Macken, Maribel Montero Perez, **Orphée De Clercq**, Lidia Rura and Julia Trushkina (2011). Dutch Parallel Corpus. Instituut voor Nederlandse Lexicologie, ISBN 9789491325007.
- * **De Clercq, Orphée** and Martin Reynaert (2010), SoNaR Acquisition Manual. LT3 Technical Report LT3 10-02. Version 1.0

12 Presentations, Conference Contributions

– Invited:

- * presentation on “Generatieve AI voor tekstgeneratie: een nieuwe realiteit” (with Joni Kruijsbergen) at Taalunie Zomercursus Nederlands, Ghent, Belgium, August 14, 2024.
- * presentation on “AES for Flemish Centralized Testing” at T2K research brainstorm day, Ghent University - imec, Ghent, Belgium, January 26, 2024.
- * presentation on “The Quest for Rich Models of Text Meaning” (with Els Lefever) at HiTz Basque Center for Language Technology - Ixa NLP Group, University of the Basque Country, Donostia, December 12, 2023.
- * presentation on “AI voor tekstgeneratie” at a Ghent-university wide info session on generative AI, online via Teams, December 4, 2023.
- * presentation on “Zei iemand ChatGPT? Generatieve AI: een nieuwe realiteit” (with Arda Tezcan) at the language technology demystified study evening for language teachers, Ghent, Belgium, May 4, 2023.
- * poster presentation on “Intricate Sentiment Analysis of Customer Feedback” at AI4Growth Session 4 - Retail, B2C & Customer exp., Ghent, Belgium, January 28, 2019.
- * presentation on “Aspect-based sentiment analysis for customer reviews” at Hands-on Digital Day 2019 (Department of Communication Sciences, Ghent University), Ghent, Belgium, May 6, 2019.
- * presentation on “Generic Readability Prediction” at the workshop on Predicting and Improving Readability, Utrecht, The Netherlands, April 5, 2018.
- * presentation on “Staan linguïsten en juristen lijnrecht tegenover elkaar wanneer het aankomt op leesbaarheid?” at the symposium Begrijpelijke taal in het privaatrecht, Leiden, December 1, 2017.
- * tutorial on “From Easy-to-calculate Formulas to Holistic Readability Prediction” (together with Véronique Hoste) at the RANLP conference, Varna, Bulgaria, September 2, 2017.
- * presentation on “Lekker, yummy, délicieux. Fine-grained sentiment analysis of customer reviews” at the 3rd Belgium NLP Meetup, Ghent, Belgium, May 18, 2017.
- * keynote on “The many aspects of fine-grained sentiment analysis. An overview of the task and its main challenges” at HUSO conference, Barcelona, Spain, November 15, 2016.
- * presentation on “*UCL: great beer at the ‘cercles’ but very dirrrty!* Aspect-based sentiment analysis of customer reviews.” at Séminaires du CENTAL, UCL, Louvain-la-Neuve, Belgium, October 28, 2016.
- * presentation on “Overview of my readability research”, DWS Colloquium, Mannheim, Germany, February 14, 2014.

– Main presenter:

- * accepted abstract, poster presentation on “Writing Support for Adult Learners of Dutch in the Era of Generative AI” (co-presentation with Joni Kruijsbergen) at SigWriting, Paris, France, June 27, 2024.

- * accepted abstract, presentation on “Exploring challenges and opportunities of automated essay scoring in beginner L2 English writers” (co-presentation with Vanessa De Wilde) at Anéla 2024, Ghent, Belgium, June 6, 2024.
- * accepted paper, poster presentation on “Enhancing Unrestricted Cross-Document Event Coreference with Graph Reconstruction Networks” (co-presentation with Loïc De Langhe at LREC-COLING 2025, Torino, Italy, May 22, 2024.
- * accepted abstracts, poster presentation on “Towards a Valid and Transparent ML Methodology for Dutch AES” (co-presentation with Geert De Meyer) and poster presentation on “Benchmarking Zero-Shot Text Classification for Dutch” (co-presentation with Loïc De Langhe and Pranaydeep Singh) at CLIN 2023, Antwerp, Belgium, September 22, 2023.
- * accepted paper, oral presentation on “It’s absolutely divine, Can fine-grained sentiment analysis benefit from coreference resolution?” at COLING 2020 – CRAC workshop, online, December 12, 2020
- * presentation on “Uncovering Machine Translationese: an experiment on 4 MT systems for English-French translations” (co-presentation with Rudy Look) at TQ2020, Roubaix, France, January 31, 2020.
- * accepted abstract, presentation on “Towards Dutch Automated Writing Evaluation” at CLIN30, Utrecht, the Netherlands, January 30, 2020.
- * accepted abstract, presentation on “EventDNA: Identifying event mention spans in Dutch-language news text” at CLIN30, Utrecht, the Netherlands, January 30, 2020.
- * accepted paper, poster presentation on “Benefits of Data Augmentation for NMT-Based Text Normalization of User-Generated Content” at EMNLP-IJCNLP 2019 - W-NUT workshop, Hong Kong, November 4, 2019.
- * accepted abstract (on invitation), presentation on “Machine learning for readability prediction” at DICOEN 2019 conference, Panel 5: Readability: methodological advances and practical applications, Leuven, Belgium, June 4, 2019.
- * accepted abstract, presentation on “OMG het was zooo lekker! Can fine-grained sentiment analysis benefit from coreference resolution?” at CLIN 28, Nijmegen, The Netherlands, January 26, 2018.
- * poster presentation on “Towards an integrated pipeline for aspect-based sentiment analysis” at the LW Research Day, Ghent, Belgium, November 29, 2017.
- * accepted abstract, presentation on “Hello ABSA! Collaboration between LT3 and Hello Customer to develop an aspect-based sentiment analysis pipeline” at the Atila Meeting, Tilburg, The Netherlands, November 17, 2017.
- * presentation of demo “Readability prediction” at the Atila Meeting, Tilburg, The Netherlands, November 16, 2017.
- * accepted paper, poster presentation on “Towards an integrated pipeline for aspect-based sentiment analysis in various domains” at the 8th WASSA workshop at EMNLP2017, Kopenhagen, Denmark, September 8, 2017.
- * accepted abstract, poster presentation on “Court verdict: unreadable” and presentation on “Translationese and Posteditese: Can users and computers

- detect the difference?” (co-presentation with Joke Daems) at CLIN 27, Leuven, Belgium, February 9, 2017.
- * two accepted papers, presentation on “What does the bird say? Exploring the Link Between Personality and Language Use in Dutch Tweets” at HUSO conference, EMOSEDE session, Barcelona, Spain, November 15, 2016.
 - * acceptor paper, poster presentation, shared task co-organizer of SemEval 2016 ABSA: aspect-based sentiment analysis at NAACL2016, San Diego, USA, June 16, 2016.
 - * accepted paper, poster presentation on “Rude waiter but mouthwatering pastries! An exploratory study into Dutch Aspect-Based Sentiment Analysis” at LREC2016, Portoroz, Slovenia, May 26, 2016.
 - * two presentations on “Wot r U talkin bout?! Transform noisy user-generated text to something readable” and “Lovely dress but such a mess: Discover how customers talk about your company/store/product at PARIS workshop, Heverlee, Belgium, May 10, 2016.
 - * accepted abstract, poster presentation on “Sustainability report readability prediction: an NLP expansion on the tried-and-true formulae” (co-presentation with Nils Smeuninx) at CLIN 26, Amsterdam, The Netherlands, December 18, 2015.
 - * lightning talk on “Participate in this year’s SemEval 2016 Task on multilingual aspect-based sentiment analysis” at Atila meeting, Antwerp, Belgium, October 15, 2015.
 - * presentation on “Terminology Extraction to Aspect-Based Sentiment Analysis” at CLIN 2015, Antwerp, Belgium, February 6, 2015.
 - * presentation on, “Content-Based Book Recommendation”, Atila meeting, Ghent, Belgium, November 20, 2014.
 - * accepted paper, presentation on, “Exploiting FrameNet for Content-Based Book Recommendation” at the CBRecsys workshop, Recsys 2014, Foster City, USA, October 5, 2014.
 - * accepted paper, presentation on “Identifying Disputed Topics in the News at the LD4KD workshop, ECML 2014, Nancy, France, September 19, 2014.
 - * accepted paper, presentation on “Towards Shared Datasets for Normalization Research” at (co-presentation with Sarah Schulz) at LREC 2014, Reykjavik, Iceland, May 28, 2014.
 - * presentation on “Crash-testing text processing on user-generated content” at CLIN 2014, Leiden, The Netherlands, January 17, 2014.
 - * poster presentation on “Normalization of Dutch User-Generated Content” (co-presentation with Sarah Schulz and Bart Desmet) at CLIN 2014, Leiden, The Netherlands, January 17, 2014.
 - * presentation on “Automatic Normalization of Dutch User-Generated Content” at Onderzoeksforum Faculty of Applied Language Studies, Ghent, Belgium, December 10, 2013.
 - * presentation on “Mining through the noise” at LTCI seminar, Ghent, Belgium, December 6, 2013.

- * presentation on “Plain Language: Investigating the effect of different text characteristics on text readability” at 3rd Translation Studies Day@European Commission, Brussels, Belgium, October 25, 2013.
- * presentation on “First steps in normalizing Dutch UGC” at Atila Meeting, Corsendonk, Belgium, September 27, 2013.
- * accepted paper, presentation on “Normalisation of Dutch User-Generated Content” at RANLP2013, Hissar, Bulgaria, September 10, 2013.
- * presentation on “HENDI: holistic readability prediction” at Onderzoekforum Faculty of Applied Language Studies, Ghent, Belgium, April 22, 2013.
- * accepted abstract, presentation on “Features for automatic readability prediction” at CLIN 2013, Enschede, The Netherlands, January 18, 2013.
- * presentation on “Dutch readability prediction” at LTCI Meeting, Ghent, Belgium, December 14, 2012.
- * presentation on “Comparing features for automatic readability prediction” at Atila Meeting, Groesbeek, The Netherlands, November 23, 2012.
- * poster presentation on “SoNaR1: a one-million-word Dutch semantic resource”, LOT Summer School, Utrecht, The Netherlands, July 2, 2012.
- * accepted paper, poster presentation on “Collecting a corpus of Dutch SMS” at LREC 2012, Istanbul, Turkey, May 24, 2012.
- * accepted paper, presentation on “Evaluating automatic cross-domain Dutch semantic role annotation” at LREC 2012, Istanbul, Turkey, May 23, 2012.
- * accepted abstract, presentation on “Evaluating automatic cross-domain Dutch semantic role annotation” at CLIN 2012, Tilburg, The Netherlands, January 20, 2012.
- * presentation on “Coreference resolution across different genres and the special case of bridge relations” at Atila Meeting, Antwerp, Belgium, December 1, 2011.
- * accepted paper, presentation on “Cross-Domain Dutch Coreference Resolution” at RANLP 2011, Hissar, Bulgaria, September 13, 2011.
- * poster presentation on “Semantic annotation in SoNaR” at Doctorandidag Hogeschool Gent, Ghent, Belgium, March 10, 2011.
- * accepted abstract, presentation on “Cross-Domain Dutch Coreference Resolution” at CLIN 2011, Belgium, Ghent, February 11, 2011.
- * presentation on “SoNaR Stevin Nederlandstalig Referentiecorpus” at STEVIN-programmadag, Antwerp, Belgium, October 14, 2010.
- * accepted paper, poster presentation on “Data Collection and IPR in Multilingual Parallel Corpora” at LREC 2010, Valletta, Malta, May 21, 2010.
- * accepted paper, poster presentation on “Balancing SoNaR: IPR versus Processing Issues in a 500-Million-Word Written Dutch Reference Corpus” at LREC 2010, Valletta, Malta, May 20, 2010.
- * accepted abstract, presentation on “Acquiring a Reference Corpus of Written Dutch” at CLIN 2010, Utrecht, The Netherlands, February 5, 2010.
- * accepted abstract, presentation on “Dutch Parallel Corpus: a new parallel corpus with Dutch as central language” at CLIN 2010, Utrecht, The Netherlands, February 5, 2010.

- * presentation on “Acquisition and IPR” (co-presentation with Maribel Montero Perez), DPC workshop, Antwerp, Belgium, September 18, 2009.

– **Not main presenter**

- * accepted abstracts, poster presentation on “EmotionNL Demo: A Framework for Dutch Emotion Detection in Crisis Communication” (with **Luna De Bruyne** and Veronique Hoste) and poster presentation on “Exploring LLMs’ capabilities for Dutch Grammatical Error Detection in L1 and L2 texts” (with **Serafina Van Geertruyen, Joni Kruijsbergen** and Veronique Hoste) at CLIN 2023, Antwerp, Belgium, September 22, 2023.
- * accepted abstract, presentation on “Man and machine: match made in heaven? (status: ‘It’s complicated...’): exploring Dutch Automated Essay Scoring for young learners’ written output” (with **Geert De Meyer**, Fauve De Backer and Maribel Montero Perez) at ALTE 2023, Madrid, Spain, April 28, 2023. *For this abstract we received a Registration Scholarship.*
- * accepted abstract, poster presentation on “Dutch Automated Writing Evaluation” (with **Serafina Van Geertruyen** and Veronique Hoste) at Linghientian Doctorials, Ghent, December 21, 2022.
- * accepted paper, presentation on (with Johannes Sibeko) at EACL 2023 - workshop, 2023.
- * accepted paper, presentation on (with Loic De Langhe and Veronique Hoste) at EACL 2023 - CRAC workshop, 2023.
- * accepted paper, poster presentation on “How Language-Dependent is Emotion Detection? Evidence from Multilingual BERT” (with **Luna De Bruyne**, Pranaydeep Singh, Els Lefever and Veronique Hoste) at EMNLP 2022 - MLR workshop, online, December 8, 2022.
- * acceptor paper, presentation on “A Hybrid Knowledge and Transformer-Based Model for Event Detection with Automatic Self-Attention Threshold, Layer and Head Selection” (with Thierry Desot and Veronique Hosete) at EMNLP 2022 - CASE workshop, online, December 7, 2022.
- * accepted paper, presentation on “Investigating Cross-Document Event Coreference for Dutch” (with **Loic De Langhe** and Veronique Hoste) at COLING 2022 - CRAC workshop, Gyeongju, October 16, 2022.
- * accepted abstract, presentation on “Unravelling language use that determines the quality of narrative comments in ePortfolios: a text analysis approach” (with **Sofie Van Ostaeyen**, Mieke Embo, Tammy Schellens and Martin Valcke) at AMEE Conference, Lyon, August 30, 2022).
- * accepted abstract, poster presentation on “Unravelling language use of narrative comments in ePortfolios: a text analysis approach” (with **Sofie Van Ostaeyen**, Mieke Embo, Tammy Schellens and Martin Valcke) at EARLI SIG 1 & 4 Joint Conference, Madrid, June 27, 2022.
- * accepted paper, poster presentation on “Investigating the Quality of Static Anchor Embeddings from Transformers for Under-Resourced Languages” (with **Pranaydeep Singh** and Els Lefever) at the 1st Annual Meeting of the Special Interest Group on Under-resourced Languages, Marseille, June 25, 2022.

- * acceptor paper, presentation on “Aspect-Based Emotion Analysis and Multimodal Coreference: A Case Study of Customer Comments on Adidas Instagram Posts” (with **Luna De Bruyne**, Akbar Karimi, Andrea Prati and Veronique Hoste) at LREC 2022, Marseille, June 21, 2022.
- * acceptor paper, poster presentation on “SentEMO: A Multilingual Adaptive Platform for Aspect-based Sentiment and Emotion Analysis” (with **Ellen De Geyndt**, Cynthia Van Hee, Els Lefever, Pranaydeep Singh, Olivier Parent and Veronique Hoste) at ACL 2022 - WASSA workshop, Dublin, Ireland, May 26, 2022.
- * accepted abstract, presentation on “Measuring cross-lingual syntactic similarity” (with **Bram Vanroy**, Arda Tezcan and Lieve Macken) at NITS 2022, Groningen, The Netherlands, May 13, 2022.
- * accepted abstract, presentation on “An in-depth analysis of written feedback found in ePortfolios of undergraduate midwifery students” (with **Sofie Van Ostaeyen**, Mieke Embo, Tammy Schellens and Martin Valcke) at CARE4 Conference, online, February 9, 2022.
- * accepted abstract, presentation on “Investigating Cross-Document Event Coreference Resolution for Dutch” (with **Loic De Langhe** and Veronique Hoste) at CLIN31, online, July 9, 2021.
- * accepted abstract, poster presentation on “Assessing Dutch as a Second Language Learners’ Writing Through Syntactic Complexity” (with **Nafal Ossandon Hostens** and Sarah Bernolet) at CLIN31, online, July 9, 2021.
- * accepted abstract, poster presentation on “Combining Transformers and Affect Lexica for Dutch Emotion Detection: Insights from the New EmotionNL Dataset” (with **Luna De Bruyne** and Veronique Hoste) at CLIN31, online, July 9, 2021.
- * acceptor paper, poster presentation on “Emotional RobBERT and Insensitive BERTje: Combining Transformers and Affect Lexica for Dutch Emotion Detection” (with **Luna De Bruyne** and Veronique Hoste) at EACL 2021 - WASSA 2021 workshop, online, April 19, 2021.
- * acceptor paper, poster presentation on “Exploring Implicit Sentiment Evoked by Fine-grained News Events” (with **Cynthia Van Hee** and Veronique Hoste) at EACL 2021 - WASSA 2021 workshop, online, April 19, 2021.
- * acceptor paper, presentation on “Leveraging syntactic parsing to improve event annotation matching” (with **Camiel Colruyt** and Veronique Hoste) at EMNLP-IJCNLP 2019 - AnnoNLP workshop, Hong Kong, November 3, 2019.
- * accepted paper, poster presentation on “Comparing MT Approaches for Text Normalization” (with **Claudia Matoz Veliz** and Veronique Hoste) at RANLP 2019, Varna, Bulgaria, September 2, 2019.
- * accepted abstract, presentation on “Modelling Translation Entropy and Syntactic Equivalence” (with **Bram Vanroy** and Lieve Macken) at the MEMENTO workshop - Machine Translation Summit, Dublin, Ireland, August 20, 2019.
- * accepted paper, presentation on “Towards an Empirically Grounded Framework for Emotion Analysis” (with **Luna De Bruyne** and Veronique Hoste)

at EMOSEDE - HUSO, Rome, Italy, July 1, 2019.

- * accepted abstract, poster presentation on “Transfer Learning for Spanish Emotion Detection” (with **Luna De Bruyne** and Veronique Hoste) at CLIN 29, Groningen, The Netherlands, January 31, 2019.
- * accepted abstract, presentation on “Varying Background Corpora for SMT-based Text Normalization” (with **Claudia Matos Veliz** and Veronique Hoste) at CLIN 29, Groningen, The Netherlands, January 31, 2019.
- * accepted abstract, poster presentation on “Varying Background Corpora for SMT-based Text Normalization” (with **Claudia Matos Veliz** and Veronique Hoste) at CMC and Social Media Corpora 2018, Antwerp, Belgium, September 18, 2018.
- * accepted paper, poster presentation on “LT3 at SemEval-2018 Task 1: A classifier chain to detect emotions in tweets.” (with **Luna De Bruyne**, Véronique Hoste) at SemEval 2018, New Orleans, USA, June 5, 2018.
- * accepted abstract, presentation on “Predicting Difficulty in Translation: A Pilot Study” (with **Bram Vanroy** and Lieve Macken) at the TISLID’18 conference, Ghent, Belgium, May 24, 2018.
- * poster presentation on “Predicting difficulty in translation” (with **Bram Vanroy** and Lieve Macken) at the LW Research Day, Ghent, Belgium, November 29, 2017.
- * accepted abstract, presentation on “A corpus-based, statistical approach to translation quality” (with Bert Cappelle, Gert De Sutter, **Rudy Loock** and Koen Plevoets) at Traduction Qualité 2016: corpus qualité, Lille, France, February 5, 2016.
- * accepted paper, poster presentation on “LT3: Applying Hybrid Terminology Extraction to Aspect-Based Sentiment Analysis” (with **Els Lefever** and Veronique Hoste) at SemEval 2015, Denver, USA, June 4, 2015.
- * accepted paper, poster presentation on “LT3: Sentiment Classification in user-generated content using a rich feature set” (with **Cynthia Van Hee**, Marjan Van de Kauter, Els Lefever and Véronique Hoste) at SemEval 2014, Dublin, Ireland, August 24, 2014.
- * presentation on “Crowdsourcing for Dutch Readability Prediction” (with **Veronique Hoste**), Workshop on Writing and Readability Tools, Utrecht, The Netherlands, January 31, 2014.
- * accepted paper, presentation on “Analysis and Reference Resolution of Bridge Anaphora across Different Genres” (with **Iris Hendrickx** and Véronique Hoste), DAARC 2011, Faro, Portugal, October 7, 2011.

Last update: August 2024