

CURRICULUM VITAE

Véronique Hoste

Name	Véronique Hoste
Birth place and date	Tielt, October 12th 1972
Function	full professor (hoogleraar)
Research Institute	LT3 Language and Translation Technology Team Department of Translation, Interpreting and Communication Ghent University Groot-Brittanniëlaan 45 B-9000, Ghent Belgium
	Email: veronique.hoste@ugent.be Website LT3: http://www.lt3.ugent.be

(last update: Sept. 2015)

1 Academic titles and degrees

Ph.D. Computational Linguistics (Doctoraat Letteren en Wijsbegeerte), Antwerp University. “Optimization Issues in Machine Learning of Coreference Resolution”, 2005.

Postgraduate degree GGS Taalwetenschap, interuniversity degree (RUG, UIA, KUL, VUB), 1998. [cum laude]

Thesis: “Example-Based Machine Translation. The application of memory-based reasoning in translation technology”.

M.A. Higher Institute for Translators and Interpreters. German, Portugese, Danish. 1997. [magna cum laude]

Masters thesis: “Ein kontrastives niederländisch-deutsches Valenzlexikon. Zum Problem der laut- und sinnverwandten Wörter für niederländischsprachige Germanistikstudenten”.

B.A. Higher Institute for Translators and Interpreters. German, Portugese, Danish. 1995. [magna cum laude]

2 Scientific experience

October 2013 - Head of Department, full professor, Ghent University. Director of the LT3 research team. (<http://lt3.ugent.be/>).

October 2011 - October 2013 Full professor (hoogleraar), University College Ghent. Founder and Research Director of the LT3 (Language and Translation Technology Team) research group Director of the "language technology" section.

May 2006 - October 2011 Assistant Professor (docent), University College Ghent.

September 2010 - October 2013 Visiting professor, Ghent University, Faculty of Linguistics.

September 2007 - October 2013 Associate researcher, Department of Applied Mathematics and Computer Science, Ghent University. Co-founder of the association research group "Language Technology and Computational Intelligence".

May 2006 - ... Visiting professor, Antwerp University

July 2005 - April 2006 (postdoctoral research assistant, Antwerp University): 'COREA, Coreference Resolution for Extracting Answers'. (STEVIN)

April 2005 - June 2005 (postdoctoral research assistant, Antwerp University): 'BioMinT, Biological Text Mining'. (EU)

2001-2004 (research assistant, Antwerp University): "PROSIT, Prosody from Information in Text". In this project different techniques are investigated to generate a prosodic structure for a Dutch speech synthesis system. (FWO Flanders - NWO)

1998-2000 (research assistant, Antwerp University): "LINGUADUCT, An Inductive Machine Learning Environment for Corpus Annotation". This project investigates how learning algorithms can be incorporated in the annotation of existing corpora. (FWO Flanders)

3 Project management

Current project management

- Hercules
"A parsed corpus of Historical Low German"
Period: 2014-2018
Budget: 2 PhD students
- IWT SBO
"SCATE, Smart Computer-Aided Translation Environment", in cooperation with Thomas More and the Universities of Leuven and Hasselt.

Period: 01/01/2014 - 31/12/2017

Budget LT3: 519,250 euro

- IWT SBO
“AMiCA, Automatic Monitoring for Cyberspace Applications”, in cooperation with the Universities of Antwerp (CLiPS), Ghent (IBCN) and Leuven (VISICS).
Period: 01/01/2013 - 31/12/2016
Budget LT3: 793,823 euro
- IWT
“SentiFM, Sentiment Mining for the Financial Markets”
Period: 01/01/2012 - 31/12/2015
Budget LT3: PhD project of Marjan Van de Kauter, supervised by Veronique Hoste and Diane Breesch
- IWT SBO
“PARIS, Personalised AdveRtIsements built from web Sources”.
In cooperation with K.U.Leuven (LIIR (coordinator), HDMB en VISICS), Ghent University (LTCI), Antwerp University (IELP) and VLERICK.
Period: 01/05/2012 - 30/04/2016
Budget LT3: 221,205 euro

Previous projects

- HOF (Research Fund, University College Ghent)
“SUBTLE, SUBjectivity Tagging and Learning”, in cooperation with Ghent University (Johan De Caluwe) and the Suicide Prevention Centre.
Period: 01/10/2010 - 30/09/2014
Budget: 1 PhD student (Bart Desmet)
- HOF (Research Fund, University College Ghent)
“MuST: Multilingual Corpora for the automatic Structuring of Terms” ParaSense or how to use parallel corpora for Word Sense Disambiguation”, in cooperation with Ghent University (Martine De Cock).
Period: 01/01/2013-31/12/2014
Budget: postdoc project of Els Lefever
- IWT Tetra (Insititute for the Promotion of Innovation by Science and Technology in Flanders)
“TExSIS, Terminology Extraction for Semantic Interoperability and Standardization”.
Coördinator. In cooperation with Ghent University (IBCN) and with the financial support of a user group of companies (Cross Language, SD Worx, Telenet, Belga Persbureau, Selor, Yamagata Europe, Telelingua, Mentoring Systems, Xplanation Language Services, TextKernel, Docbyte, Comsof, Actonomy, ITP, Nacco Materials Handling Group, PSA Peugeot Citroën SA, Intersystems, Jabbla, MEGA-doc, Mediargus, Eurologos, Wetenschappelijk en Technisch Centrum voor het Bouwbedrijf and Jan De Nul).
Period: 01/01/2011 - 31/12/2012
Budget: 392,149 euro
- HOF (Research Fund, University College Ghent)
“HENDI, een Automatische Helderheidstool voor Engels en Nederlands Discours in Parallele

en Vergelijkbare Teksten” (“HENDI, an Automatic Clarity Tool for English and Dutch Discourse in Parallel and Comparable Texts”), in cooperation with S. Vandepitte, S. Marzo, B. Defrancq, Hildegard Vermeiren and Roland Maerivoet (University College Ghent), W. Daelemans (Antwerp University) and Gilles-Maurice de Schryver (Ghent University)
Period: 01/10/2008 - 31/10/2012
Budget: 335,785 euro.

- HOF (Research Fund, University College Ghent)
“ParaSense or how to use parallel corpora for Word Sense Disambiguation”, in cooperation with Ghent University (Martine De Cock).
Period: 01/01/2007-31/12/2012
Budget: one PhD student
- FWO (Flemish Research Fund, “Krediet voor navorsers”) “Koppeling van schrijfprocesdata aan lexica” (Merging writing process data with lexica), in cooperation with Antwerp University (Luuk Van Waes). Period: 01/01/2010 - 31/12/2012
Budget LT3: 11,700 euro
- EWI (Department of Economy, Science and Innovation, Flemish government)
“Stylene, Een web service voor stylometrie- en leesbaarheidsonderzoek voor het Nederlands” (Stylene: a web service for stylometry and readability research for Dutch), in cooperation with Antwerp University (CLiPS).
Period: 01/01/2010 - 30/09/2012
Budget: 120,000 euro.
- EWI (Department of Economy, Science and Innovation, Flemish government)
“TTNWW, TST Tools voor het Nederlands als Webservices in een Workflow” (TTNWW: web services for Dutch as web services in a work flow), in cooperation with the Universities of Antwerp, Tilburg, Leuven, Utrecht and Groningen.
Period: 01/01/2010 - 30/09/2012
Budget LT3: 50,000 euro.
- STEVIN
“SoNaR, Stevin Nederlandstalig Referentiecorpus.” (SoNaR, Stevin Dutch Reference Corpus), in cooperation with the University of Utrecht and the Catholic University of Leuven.
Period: 01/01/2009-30/11/2011
Budget LT3: 191,121 euro.
- IWT SBO pre-project
“AMiCA, Automatic Monitoring for Cyberspace Applications”, in cooperation with the Universities of Antwerp (CLiPS), Ghent (IBCN) and Leuven (VISICS).
Period: 01/10/2010 - 30/09/2011.
- HOF (Research Fund, University College Ghent)
Server for data storage and computationally expensive processing.
Budget: 21,267 euro.
- STEVIN
“DuOMAn, Dutch Language Online Media Analysis” in cooperation with the University of Amsterdam, the University of Groningen, GridLine and TrendLight.
Period: 01/05/2009-30/04/2011
Budget LT3: 44,070 euro.

- IOF StarTT (Industrial Research Fund, Ghent University)
“OptiPIL, a readability Optimizer for Patient Information Leaflets”, in cooperation with Artesis University College Antwerp
Period: 01/01/2010-30/12/2010
Budget LT3: 49,980 euro.
- Private funding: Peugeot Citroen SA (LT3 as subcontractor of Telelingua international)
“Étude doublons et synonymes thésaurus APV”.
Period: 01/02/2008- 2010
- BOF (Special Research Fund, Ghent University)
“AutoWeb: een Semantisch Web waarin Agenten zich Gedragen als Autodidact” (AutoWeb: a Semantic Web in which the agents behave as Autodidacts”), in cooperation with Ghent University (Martine De Cock).
Period: 01/11/2006-31/10/2009
Budget: one PhD student
- IWT TETRA (Institute for the Promotion of Innovation by Science and Technology in Flanders)
“ABOP, Automatische Bijsluiter Optimalisator. Een auteursomgeving voor de correctie en productie van bijsluiter teksten” (“ABOP, Automatic Leaflet Optimizer. An authoring environment for the correction and production of patient leaflets”), in cooperation with University College Antwerp, Antwerp University and the Catholic University of Leuven.
Period: 01/12/2007-30/12/2009.
Budget LT3: 157,458 euro.
- STEVIN (Stimulation programme for the development of resources for Dutch language and speech technology)
“SoNaR, Stevin Nederlandstalig Referentiecorpus. Fase 1” (SoNaR, Stevin Dutch Reference Corpus. Phase 1), in cooperation with the University of Twente, Tilburg University and the University of Nijmegen.
Period: 01/01/2008-31/12/2008.
Budget LT3: 26,400 euro.
- BOF (Special Research Fund, Antwerp University)
“APRO: Annotatie van genuscongruentie en van het pleonastisch vs. anaforisch gebruik van pronomina in het Nederlands” (“APRO: Annotation of gender agreement and of the pleonastic versus anaphoric use of pronouns in Dutch”).
Period: 2006
Budget: 7,410 euro.

4 Teaching experience

- *Language Technology / machine learning*, DARIAH Digital Humanities Summer School, August 2014.
- *Toegepaste Taalkunde* (Ba2), Faculty of Applied Language Studies, University College Ghent, 2012-. (two lectures on language technology)
- *Introduction to Language Technology* (Postgraduate Dutch and Translation), Faculty of Applied Language Studies, University College Ghent, 2007-.... In cooperation with several colleagues.

- *Taaltechnologie* (Master), Faculty of Applied Language Studies, University College Ghent, 2007-.... In cooperation with Els Lefever.
- *Inleiding tot de taaltechnologie* (Ba3), Faculty of Applied Language Studies, University College Ghent, 2007-.... In cooperation with different colleagues.
- *Advanced Methods for Corpus Linguistics*, Doctoral School of Arts, Humanities and Law, December 2009. In cooperation with Evie Coussé and Gert De Sutter.
- *Natural Text Understanding* (MA), University of Antwerp, 2009-2011. In cooperation with Guy De Pauw.
- *Computer aided translation*, March 2009. In cooperation with Telelingua International.
- *Introduction to Language Technology*, Doctoral School of Arts, Humanities and Law, Doctoral School of Natural Sciences, Ghent University, January 2009.
- *Language Technology* (MA), School of Translation Studies, University College Ghent, 2007-...
- *Introduction to Language technology* (BA3), School of Translation Studies, University College Ghent, 2006-...
- *Text Mining*, guest lecture on anaphora resolution, Tilburg University, 3 May 2005.
- *Language technology (Computerlinguïstiek)*, course on anaphora resolution, Antwerp University, 2004-2006.
- *Language technology (Computerlinguïstiek)*, course on machine translation, Antwerp University, 1999-2001.
- Flanders Language Valley, course on machine translation, Ieper, 2001.

5 Organisation of conferences and workshops

- SemEval-2013 shared task “Cross-Lingual Word Sense Disambiguation”, Atlanta, June 2013. <http://www.cs.york.ac.uk/semEval-2013/task10/>. (in cooperation with Els Lefever)
- TExSIS workshop, IWT Tetra project “Terminology Extraction for Semantic Interoperability and Standardization, Ghent, December 13, 2012.
- CLIN 2011, “Computational linguistics in the Netherlands meeting”, Ghent, February 11, 2011. (in cooperation with Peter Velaerts, Kathelijne Denturck and Els Lefever)
- SemEval-2 task “Cross-Lingual Word Sense Disambiguation”, Uppsala, July 2010. <http://semEval2.fbk.eu/semEval2.php?location=tasks>. (in cooperation with Els Lefever)
- SemEval-2 task “Coreference Resolution in Multiple languages”, Uppsala, July 2010. <http://semEval2.fbk.eu/semEval2.php?location=tasks>. (in cooperation with Lluís Marquez, M. Antonia Martí, Massimo Poesio, Marta Recasens, Emili Sapena, Mariona Taule, Yannick Versley. (Universitat de Barcelona, Universitat Politècnica de Catalunya, Università di Trento, Universität Tübingen)

- ABOP workshop (“Automatische BijsluiterOptimalisator”), Ghent, January 20, 2010.
- DIR 2008, “Dutch-Belgian Information Retrieval Workshop”, Maastricht, April 14-15, 2008. (in cooperation with Martine De Cock and Eduard Hoenkamp)
- Atila workshop, Priorij Corsendonk, November 28-29, 2005.
- “Workshop on Machine Learning of Information Extraction”, UA, Hof van Liere, Antwerp, March 10, 2005. (in cooperation with Walter Daelemans)
- “Workshop on the analysis and generation of prosody”, Tilburg, December 21, 2004. (in cooperation with Walter Daelemans, Antal van den Bosch, Erwin Marsi and Martin Reynaert)
- CLIN 2003, “Computational linguistics in the Netherlands meeting”, UA, Antwerp, December 19, 2003. (in cooperation with Bart Decadt, Guy De Pauw and Walter Daelemans)
- Atila workshop, Middelheim Castle, Antwerp, September 3, 2003.
- Benelearn 2001, “Eleventh Belgian-Dutch Conference on Machine Learning”, Hof Van Liere, Antwerp, December 21, 2001. (in cooperation with Guy De Pauw)

6 Scientific advice

Member of the Science Policy Commission (“commissie wetenschapsbeleid”) of the Flemish Council for Science and Innovation (VRWI) (2010-2014).

Member of the Ghent University Association Hercules commission (2008-2011).

Member “visitatiecommissie Taalwetenschappen” (2008), (presided by J. De Caluwe).

7 Programme committee/reviewing

- **Member of the programme committee (reviewing of full papers):**
 - 2015: WASSA Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (Lisbon, 2015), FELT-2015 Workshop on Future and Emerging Trends in Language Technology (Seville, 2015), RANLP-2015 Conference on Recent Advances in Natural Language Processing (Hissar, 2015), CICLING-2015 Conference on Intelligent Text Processing and Computational Linguistics, SemEval-2015 (Denver, 2015)
 - 2014: EACL-2014, European Chapter of the Annual Meeting of the Association of Computational Linguistics (Gothenburg, 2014), WASSA Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (Baltimore, 2014), BNAIC-2014, Benelux Conference on Artificial Intelligence (Nijmegen, 2014), COLING-2014, International Conference on Computational Linguistics (Dublin, 2014)

- 2013: CICLING-2013 Conference on Intelligent Text Processing and Computational Linguistics (Samos, 2013), DIR-2013 Dutch-Belgian Information Retrieval Workshop (Delft, 2013), NAACL-2013 Conference of the North American Chapter of the Association for Computational Linguistics (Atlanta, 2013), *SEM-2013 Joint Conference on Lexical and Computational Semantics (Atlanta, 2013), ACL-2013 Annual Meeting of the Association of Computational Linguistics (Sofia, 2013), WASSA-2013 4th Workshop on Computational Approaches to Subjectivity, Sentiment Social Media Analysis (Atlanta, 2013), BNAIC-2013 25th Benelux Conference on Artificial Intelligence (Delft, 2013), IJCNLP-2013 6th International Joint Conference on Natural Language Processing, RANLP-2013 Conference on Recent Advances Natural Language Processing (Hissar, 2013)
- 2012: EACL-2012, European Chapter of the Annual Meeting of the Association of Computational Linguistics (Avignon, 2012), LREC-2012, Language Resources and Evaluation Conference (Istanbul, 2012), DIR-2012, Dutch Information Retrieval Workshop (Gent, 2012), ACL-2012, Annual Meeting of the Association of Computational Linguistics (Jeju, 2012), EMNLP-2012, Conference on Empirical Methods in Natural Language Processing (Jeju, 2012), WASSA Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (Jeju, 2012), PITR-2012, Workshop on Predicting and Improving Text Readability (Montreal, 2012), ECAI-2012, 20th European Conference on Artificial Intelligence (Montpellier, 2012), ESSLLI-2012 24th European Summer School in Logic, Language and Information (Opole, 2012), EMNLP-CoNLL-2012 Shared Task (Jeju, 2012), BNAIC-2012, Benelux Conference on Artificial Intelligence (Maastricht, 2012), COLING-2012, International Conference on Computational Linguistics (Bombay, 2012), OTHLT-2012 Workshop on Optimization Techniques for Human Language Technology (Bombay, 2012)
- 2011: DIR-2011 Dutch-Belgian Information Retrieval Workshop (Amsterdam, 2011), WASSA-2011 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (Portland, US, 2011), ESSLLI -StuS 2011 (Ljubljana 2011), BNAIC-2011 23rd Benelux Conference on Artificial Intelligence (Ghent, 2011), RANLP-2011 Recent Advances in Natural Language Processing (Hissar, 2011), DAARC-2011 Discourse Anaphora and Anaphor Resolution Colloquia (Faro, 2011)
- 2010: ACL-2010 Annual Meeting of the Association of Computational Linguistics (Uppsala, 2010), LREC-2010 Language Resources and Evaluation Conference (Valletta, 2010), WASSA Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (Lisbon, 2010), NeSp-NLP 2010 Negation and Speculation in Natural Language Processing Workshop (Uppsala, 2010), EMNLP-2010 Conference on Empirical Methods in Natural Language Processing (Massachusetts, US, 2010)
- 2009: EACL-2009, European Chapter of the Annual Meeting of the Association of Computational Linguistics (Athens, 2009), co-editor of *Linguistica Antverpiensia*, New Series (8/2009), DIR-2009, Dutch-Belgian Information Retrieval Workshop (Twente, 2009), DAARC-2009, Discourse Anaphora and Anaphor Resolution Colloquia (Goa, 2009), SEW-2009, Semantic Evaluations: Recent Achievements and Future Directions (Colorado, 2009), ACL-IJCNLP 2009, Annual Meeting of the Association of Computational Linguistics (Singapore, 2009), NAACL-HLT 2009, North American Chapter of the Association for Computational Linguistics - Human Language Technologies (Colorado, 2009), ECML-PKDD 2009, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Bled, 2009), RANLP-2009, Recent Advances in Natural Language Processing (Borovets, 2009)
- 2008: LREC-2008, Language Resources and Evaluation Conference (Marrakech, 2008), ACL-2008, Annual Meeting of the Association of Computational Linguistics (Ohio, 2008),

- SEW, semantic Evaluations Workshop (Ohio, 2008), DIR-2008, Dutch-Belgian Information Retrieval Workshop (Maastricht, 2008), EMNLP-2008, Conference on Empirical Methods in Natural Language Processing (Waikiki, Honolulu, 2008)
- 2007: ICSC-2007, IEEE International Conference on Semantic Computing (Irvine, 2007), SemEval 2007, International Workshop on Semantic Evaluations (Prague, 2007), RANLP-2007, Recent Advances in Natural Language Processing (Borovets, 2007)
 - 2006: EMNLP, Empirical Methods in Natural Language Processing (Sydney, 2006), COLING-ACL, International Conference on Computational Linguistics-Annual Meeting of the Association of Computational Linguistics (Sydney, 2006), ESSLLI workshop "Ambiguity in Anaphora" (Malaga, 2006)
 - 2005 and before: CONLL-2005, Ninth Conference on Computational Natural Language Learning (Ann Arbor, 2005), ACL-2004, Annual Meeting of the Association of Computational Linguistics (Barcelona, 2004), EACL 99 workshop "Linguistically Interpreted Corpora" (Bergen, 1999)
- **Reviewing for:**
 - Journal of Artificial Intelligence Research (2003, 2006, 2007)
 - Computational Linguistics in the Netherlands (2004, 2006, 2007, 2010, 2013, 2014)
 - Journal of Machine Learning Research (2006, 2014)
 - Research on Language and Computation (2007, 2010)
 - Festschrift for David Dowty 'Theory and Evidence in Semantics' (2007)
 - Linguistica Antverpiensia New Series (2009, 2011, 2012)
 - LRE (2011, 2013, 2014)
 - Journal of the American Society for Information Science and Technology (2013, 2014)
 - Information Processing and Management (2013)
 - Expert Systems with Applications (2015)
 - **Project Proposal Refereeing:**
 - Hercules 1 and 2 (infrastructure) (2008, 2009)
 - Erasmus Hogeschool research fund (Linguistics) (2008)
 - ESRC Economic and Social Research Council (UK, 2009)
 - NWO (The Netherlands, 2009, 2012, 2014, 2015)
 - FWO (2010)
 - Swiss National Science Foundation (2011)
 - Member of the research unit of the Faculty of Translation Studies (internal proposal reviewing)
 - NTU (2013)
 - IWT (2014)

8 Ph.D. students

- Completed supervised Ph.D.'s:
 - Timur Fayruzov, *Mining and Modelling Interaction Networks for Systems Biology*, Ghent University, October 29, 2010.
 - Els Lefever, *ParaSense: Parallel Corpora for Word Sense Disambiguation*, Ghent University, September 24, 2012.

- Bart Desmet, *Finding the Online Cry for Help: Automatic Text Classification for Suicide Prevention*, Ghent University, October 17, 2014.
- Orphée De Clercq, *Tipping the scales. Exploring the added value of deep semantic processing on readability prediction and sentiment analysis*, Ghent University, July 9, 2015.
- Currently supervised Ph.D.’s:
 - Sabien Hanouille (2012-2015) prepares a Ph.D on the use of translation memories and terminology extraction in the translation of documentaries. Sabien is co-supervised by Veronique Hoste and Aline Remail (ARTESIS).
 - Cynthia Van Hee (2013-2016) prepares a Ph.D in the framework of the IWT SBO AMiCA project on the monitoring of unwanted content on social media.
 - Mariya Koleva (2014-2017) prepares a Ph. D in the framework of the Hercules project on the construction of a corpus of Historical German.
- Member of the jury:
 - Anders Noklestad: “A Machine Learning Approach to Anaphora Resolution Including Named Entity Recognition, PP Attachment Disambiguation, and Animacy Detection (University of Oslo, June 18-19, 2009).
 - Oier Lopez de Lacalle: “Domain-specific Word Sense Disambiguation” (Universidad del País Vasco, December 2009)
 - Stefan Raaijmakers: “Multinomial Language Learning. Investigations into the geometry of language” (University of Tilburg, 1st December 2009)
 - Lieve Macken: “Sub-sentential Alignment in Translational Correspondences” (Antwerp University, May 28, 2010)
 - Kim Luyckx: “Scalability Issues in Authorship Attribution” (Antwerp University, December 7, 2010)
 - Alexandra Balahur Dobrescu: “Methods and Resources for Sentiment Analysis in Multilingual Documents of Different Text Types” (Universidad de Alicante, June 13, 2011)
 - Sofie Van Landeghem: “Playing hide and seek on the genomic playground: Unveiling biological function from literature” (Ghent University, April 27, 2012)
 - Emili Sapena Masip: “A constraint-based hypergraph partitioning approach to coreference resolution” (University of Barcelona, May 15, 2012)
 - Klaar Vanopstal: “Impact of language skills and system experience on medical information retrieval” (Ghent University, August 27, 2013)
 - Ernst Roald Eiselen: “
 - Germán Hurtado Martin: “Intelligent Methods for Information Filtering of Research Resources” (Ghent University, December 6, 2013)
 - Maral Dadvar: “Experts and Machines United Against Cyberbullying” (University of Twente, September 12, 2014)
 - Frederik Vaassen: “Measuring emotion. Exploring the feasibility of automatically classifying emotional text” (Antwerp University, October 10, 2014)

9 Membership

- Member of the advisory committee of the Institute for Dutch Lexicology (Dutch Language Union) (2011-...)
- Co-promotor (in cooperation with Martine De Cock) of the Ghent University Association research group “Language Technology and Computational Intelligence”. The AOG comprises the following groups: Computational Web Intelligence (Ghent University), Language and Translation Technology Team LT3 (University College Ghent), Computational Linguistics for African Languages (Ghent University), Department of Industrial Management (University College Ghent), Department of Applied Engineering (University College Ghent)
- Member of the research unit of the faculty of translation studies (2007-2013), of the research unit of the Department of
- Member of the CLIF coordination group (2007-2012)
- Secretary of CLIF, Computational Linguistics in Flanders, FWO sponsored research community on Language and Speech Technology, 2004
- Member of ACL (Association for Computational Linguistics)

10 Publications

- Articles
 - A1
 - * Hanoulle, S., Hoste, V. & Remael, A.: 2015, ‘The translation of documentaries: can terminology-extraction reduce the translators’ workload? An experiment involving professional translators’, *New Voices in Translation Studies*.
 - * Hanoulle, S., Hoste, V. & Remael, A.: 2015, ‘The efficacy of terminology-extraction systems for the translation of documentaries’, *Perspectives: Studies in Translatology*.
 - * Van de Kauter, M., Desmet, B. & Hoste, V.: 2015, ‘The Good, the Bad and the Implicit: A Comprehensive Approach to Annotating Explicit and Implicit Sentiment’, *Language Resources and Evaluation*.
 - * Van de Kauter, M., Breesch, D. & Hoste, V.: 2015, ‘Fine-grained analysis of explicit and implicit sentiment in financial news articles’, *Expert Systems with Applications*, 42(11), 4999-5010.
 - * Lefever, E., Van de Kauter, M. & Hoste V.: 2014, ‘HypoTerm detection of hypernym relations between domain-specific terms in Dutch and English’, *Terminology*, 20(2), 250-278.
 - * De Clercq, O., Hoste, V., Desmet, B., van Oosten, P., De Cock, M., Macken, L. : 2014, Using the Crowd for Readability Prediction. *Natural Language Engineering*, 20 (3), 293-235. Cambridge Journals Online.
 - * Desmet, B., Hoste, V.: 2013, Fine-Grained Dutch Named Entity Recognition, *Language Resources and Evaluation*, 48 (2), 307-343. Springer Netherlands.
 - * Desmet, B., Hoste, V.: 2013, Emotion Detection in Suicide Notes, *Expert Systems with Applications*, 40(16), 6351-6358.

- * Macken, L., Lefever, E. & Hoste, V.: 2013, ‘TEXSIS: Bilingual Terminology Extraction from Parallel Corpora Using Chunk-based Alignment’, *Terminology*, 19(1).
 - * Lefever, E., Fayruzov, T., Hoste, V. & De Cock, M.: 2010, ‘Clustering Web People Search Results Using Fuzzy Ants’. *Information Sciences*, 180(17), 3192-3209.
 - * Hoste, V., Vanopstal, K., Lefever, E. & Delaere, I.: 2010, ‘Classification-Based Scientific Term Detection in Patient Information’. *Terminology*, 16(1).
 - * Daelemans, W. & Hoste, V.: 2009, ‘Introduction. Evaluation of Translation Technology’. *Linguistica Antverpiensia New Series 8*, University Press Antwerp, 9-13.
 - * Fayruzov, T., De Cock, M., Cornelis, C. & Hoste, V.: 2009, ‘Linguistic Feature Analysis of Protein Interaction Extraction’. *BMC Bioinformatics*, 10:374. (SCI: 3.78)
 - * Hoste, V., Daelemans, W. and Gillis, S.: 2004, ‘Using rule-induction techniques to model pronunciation variation in Dutch’, *Computer Speech and Language*, 18(1), 1-23. (SCI: 1.413)
 - * Daelemans, W., Hoste, V., De Meulder, F. and Naudts, B.: 2003, ‘Combined optimization of feature selection and algorithm parameters in machine learning of language’, *Lecture Notes in Artificial Intelligence*, 2837, 84-95. (SCI2004: 0.251)
- A2
- * Desmet, B. & Hoste, V.: 2012, ‘Combining Lexico-semantic Features for Emotion Classification in Suicide Notes’, *Biomedical Informatics Insights*, 5, 125-128, Libertas Academica.
 - * Hoste, V., Daelemans, W., Hendrickx, I. & van den Bosch, A.: 2002, ‘Parameter Optimization for Machine-Learning of Word Sense Disambiguation’, *Natural Language Engineering*, 8(4), 311-325.
- Books
- B2
- * Hoste, V.: 2012, ‘The Mention-Pair Model’, In Massimo Poesio, Roland Stuckardt & Yannick Versley (Eds.), *Anaphora Resolution: Algorithms, Resources, and Evaluation*, Springer Verlag. (to appear)
 - * Hendrickx, I., Bouma, G., Daelemans, W. & Hoste, V.: 2012, ‘COREA: Coreference Resolution for Extracting Answers for Dutch’, In Peter Spyns & Jan Odijk (Eds.), *Essential Speech and Language Technology for Dutch, Theory and Applications of Natural Language Processing*, 115-128, Springer Verlag.
 - * Oostdijk, N., Reynaert, M., Hoste, V. & Schuurman, I.: 2012, ‘The construction of a 500-million-word reference corpus of contemporary written Dutch, In Peter Spyns & Jan Odijk (Eds.), *Essential Speech and Language Technology for Dutch, Theory and Applications of Natural Language Processing*, 219-247, Springer Verlag.
 - * Daelemans, W., De Pauw, G., Durieux, G., Gillis, S., Hoste, V. & Tjong Kim Sang, E.: 2000, ‘Zelflerende Systemen als Instrument voor de Taalkunde en de Taaltechnologie, Met taal om de tuin geleid’, 81-93.
- B3
- * Daelemans, W. & Hoste, V. (eds.): 2009, *Evaluation of Translation Technology. Linguistica Antverpiensia New Series 8*. University Press Antwerp.
 - * Hoenkamp, E., De Cock, M. & Hoste, V. (eds.): 2008, *Proceedings of the 8th Belgian-Dutch Information Retrieval Workshop (DIR-2008)*, Maastricht, the Netherlands, 100 pages.

- * Decadt, B., Hoste, V. & and Depauw, G. (eds.): 2004, *Computational Linguistics in the Netherlands 2003. Selected Papers from the Fourteenth CLIN meeting*, Antwerp, Belgium, 184 pages.
 - * Hoste, V. & De Pauw, G. (eds.): 2001, *Proceedings of the Eleventh Belgian-Dutch Conference on Machine Learning*, Antwerp, Belgium.
- Conference Proceedings
 - P1 (ISI Web of Science)
 - * De Clercq, O., Schulz, S., Desmet, B., Hoste, V. (2014). Towards Shared Datasets for Normalization Research. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk and Stelios Piperidis (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 1218-1223. European Language Resources Association (ELRA), Reykjavik, Iceland.
 - * Desmet, B., Hoste, V.: 2014. Recognising suicidal messages in Dutch social media. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk and Stelios Piperidis (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 830-835. European Language Resources Association (ELRA), Reykjavik, Iceland.
 - * Lefever, E., Van de Kauter, M., Hoste, V.: 2014. Evaluation of Automatic Hypernym Extraction from Technical Corpora in English and Dutch. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk and Stelios Piperidis (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 490-497. European Language Resources Association (ELRA), Reykjavik, Iceland.
 - * De Clercq, O., Hoste, V. & Monachesi, P.: 2012, 'Evaluating Automatic Cross-Domain Semantic Role Annotation'. *Proceedings of the 8th Language Resources and Evaluation Conference (LREC2012)*. Istanbul, Turkey.
 - * Lefever, E., Hoste, V. & De Cock, M.: 2012, 'Discovering Missing Wikipedia Inter-language Links by means of Cross-lingual Word Sense Disambiguation'. *Proceedings of the 8th Language Resources and Evaluation Conference (LREC2012)*. Istanbul, Turkey.
 - * Reynaert, M., Schuurman, I., Hoste, V., Oostdijk, N. & van Gompel, M.: 2012, 'Beyond SoNaR: towards the facilitation of large corpus building efforts', *Proceedings of the 8th Language Resources and Evaluation Conference (LREC2012)*. Istanbul, Turkey.
 - * Macken, L., Hoste, V., Leijten, M. & Van Waes, L.: 2012, 'From keystrokes to annotated process data: Enriching the output of Inputlog with linguistic information', *Proceedings of the 8th Language Resources and Evaluation Conference (LREC2012)*. Istanbul, Turkey.
 - * Hendrickx, I., De Clercq, O. & Hoste, V.: 2011, 'Analysis and Reference Resolution of Bridge Anaphora across Different Text Genres', *Lecture Notes in Artificial Intelligence*, 7099, 1-11, Springer Verlag.
 - * van Oosten, P., Hoste, V. & Tanghe, D.: 2011, 'A posteriori Agreement as a Quality Measure for Readability Prediction Systems', In Alexander Gelbukh (ed.), *Computational Linguistics and Intelligent Text Processing*, 6609: *Lecture Notes in Computer Science*, 6609, Berlin: Springer Verlag, 424-435.
 - * Hendrickx, I. & Hoste, V.: 2009, 'Coreference Resolution on Blogs and Commented News', *Lecture Notes in Artificial Intelligence*, 5847, Berlin: Springer Verlag, 43-53.

- * Hendrickx, I., Hoste, V. & Daelemans, W.: 2008, ‘Semantic and syntactic features for Dutch coreference resolution’, *Lecture Notes in Computer Science*, 4919, Berlin: Springer Verlag, 351-361.
- * Fayruzov, T., De Cock, M., Cornelis, C. & Hoste, V.: 2008, ‘DEEPER: A Full Parsing Based Approach to Protein Relation Extraction’, *Lecture Notes in Computer Science*, 4973, Berlin: Springer Verlag, 36-47.
- * Hoste, V., Lefever, E., Vanopstal, K. & Delaere, I.: 2008, ‘Learning-based Detection of Scientific Terms in Patient Information’, *Proceedings of the sixth international conference on Language Resources and Evaluation*. Marrakech, Morocco.
- * Hendrickx, I., Bouma, G., Coppens, F., Daelemans, W., Hoste, V., Kloosterman, G., Mineur, A.-M., Van Der Vloet, J. & Verschelde, J.-L.: 2008, ‘Coreference Resolution for Extracting Answers for Dutch’, *Proceedings of the sixth international conference on Language Resources and Evaluation*, Marrakech, Morocco.
- * Hoste, V., Hendrickx, I. & Daelemans, W.: 2007, ‘Disambiguation of the neuter pronoun and its effect on pronominal coreference resolution’, *Lecture Notes in Artificial Intelligence*, 4629, Berlin: Springer Verlag, 48-55.
- * Hendrickx, I., Hoste, V. & Daelemans, W.: 2007, ‘Evaluating hybrid versus data-driven coreference resolution’, *Lecture Notes in Computer Science*, 4410, Berlin: Springer Verlag, 137-150.
- * Marsi, E., Reynaert, M., van den Bosch, A., Daelemans, W. & Hoste, V.: 2003, ‘Learning to Predict Pitch Accents and Prosodic Boundaries in Dutch’, *Proceedings of the 41st annual meeting of the association for computational linguistics*, 489-496.
- * De Meulder, F., Daelemans, W. & Hoste, V.: 2002, ‘A Named Entity Recognition System for Dutch’, *Language and Computers: Studies in Practical Linguistics*, 45, 77-88.

– C1

- * Tezcan, A., Hoste, V., Desmet, B. & Macken, L.: 2015, ‘UGENT-LT3 SCATE System for Machine Translation Quality Estimation’, *WMT2015*.
- * Van Hee, C., Lefever, E. & Hoste, V.: 2015, ‘LT3: sentiment analysis of figurative tweets: piece of cake NotReally’, *SemEval : 9th International Workshop on Semantic Evaluations at Naacl 2015*, 684-688.
- * De Clercq, O., Schuhmacher, M., Ponzetto, S. & Hoste, V.: 2014, ‘Exploiting frameNet for content-based book recommendation’, *CBRecSys at ACM RecSys*, 14-21.
- * Van Hee, C., Van de Kauter, M., De Clercq, O., Lefever, E. & Hoste, V.: 2014, ‘LT3: sentiment classification in user-generated content using a rich feature set’, *Proceedings of the 8th International Workshop on Semantic Evaluation*, 406-410.
- * De Clercq, O., Hertling, S., Hoste, V., Ponzetto, S. & Paulheim, H.: 2014, ‘Identifying Disputed Topics in the News’, *Proceedings of the LD4KD Workshop at ECML/PKDD2014*, 37-48.
- * De Clercq, O., Schulz, S., Desmet, B., Lefever, E., Hoste, V.: 2013. Normalization of Dutch User-Generated Content. Proceedings of the 9th International Conference on Recent Advances in Natural Language Processing (RANLP 2013).
- * Leijten, M, Macken, L., Hoste, V., Van Horenbeeck, E. & Van Waes, L.: 2012, ‘From Character to Word Level: Enabling the Linguistic Analyses of Inputlog Process Data’, *Proceedings of the EAACL 2012 Workshop on Computational Linguistics and Writing*, Avignon, France.
- * De Clercq, O., Hoste, V., & Hendrickx, I.: 2011, ‘Cross-Domain Dutch Coreference Resolution’, *Proceedings of the 8th International Conference on Recent Advances in Natural Language Processing (RANLP 2011)*, Hissar, Bulgaria.

- * Lefever, E., Hoste, V. & De Cock, M.: 2011, ‘ParaSense or how to use Parallel Corpora for Word Sense Disambiguation’, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Portland, US.
- * Recasens, M., Màrquez, L., Sapena, E., Martí, A., Taulé, M., Hoste, V., Poesio, M. & Versley, Y.: 2010, ‘SemEval-2010 Task 1: Coreference Resolution in Multiple Languages’, *Proceedings of the ACL Workshop on Semantic Evaluations (SemEval-2010)*, Uppsala, Sweden.
- * Lefever, E. & Hoste, V.: 2010, ‘SemEval-2010 Task 3: Cross-Lingual Word Sense Disambiguation’, *Proceedings of the ACL Workshop on Semantic Evaluations (SemEval-2010)*, Uppsala, Sweden.
- * Desmet, B. Hoste, V.: 2010, ‘Dutch Named Entity Recognition using Classifier Ensembles’, *Computational Linguistics in the Netherlands 2010: selected papers from the twentieth CLIN meeting*, Utrecht, Netherlands.
- * Vanopstal, K. & Hoste, V.: 2010, ‘Towards a learning approach for abbreviation detection and resolution’, *Proceedings of the International Conference on Language Resources and Evaluation*, Valletta, Malta.
- * Van Oosten, P., Tanghe, D. & Hoste, V.: 2010, ‘Towards an Improved Methodology for Automated Readability Prediction’, *Proceedings of the International Conference on Language Resources and Evaluation*, Valletta, Malta.
- * Desmet, B. & Hoste, V.: 2010, ‘Towards a Balanced Named Entity Corpus for Dutch’, *Proceedings of the International Conference on Language Resources and Evaluation*, Valletta, Malta.
- * Lefever, E. & Hoste, V.: 2010, ‘Construction of a Benchmark Data Set for Cross-lingual Word Sense Disambiguation’, *Proceedings of the International Conference on Language Resources and Evaluation*, Valletta, Malta.
- * Ineke Schuurman, I., Hoste, V. & Monachesi, P.: 2010, ‘Interacting semantic layers of annotation in SoNaR, a reference corpus of contemporary written Dutch’, *Proceedings of the International Conference on Language Resources and Evaluation*, Valletta, Malta.
- * Lefever, E., Macken, L. & Hoste, V.: 2009, ‘Language-independent bilingual terminology extraction from a multilingual parallel corpus’, *Proceedings of the European Chapter of the Annual Meeting of the Association of Computational Linguistics*. Athens, Greece, 496-504.
- * Delaere, I., Hoste, V., Peersman, C., Van Vaerenbergh, L. & Velaerts, P.: 2009, ‘ABOP, Automatic Optimization of Patient Information Leaflets’, *Proceedings of the International Symposium on Data and Sense Mining, Machine Translation and Controlled Languages*, Besancon, France.
- * Schuurman, I., Hoste, V. & Monachesi, P.: 2009, ‘Cultivating Trees: Adding Several Semantic Layers to the Lassy Treebank in SoNaR’, *Proceedings of the 7th International Workshop on Treebanks and Linguistic Theories*, Groningen, The Netherlands, 135-146.
- * Lefever, E., Fayruzov, T., Hoste, V. & De Cock, M.: 2009, ‘Fuzzy Ants Clustering for Web People Search’, *Proceedings of the 2nd Web People Search Evaluation Workshop (WePS 2009)*, 18th WWW Conference, Madrid, Spain.
- * Lefever, E. & Hoste, V.: 2009, ‘SemEval-2010 Task 3: Cross-lingual Word Sense Disambiguation’, *Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions (SEW-2009)*, Boulder, Colorado, 82-87.
- * Macken, L., Lefever, E. & Hoste, V.: 2008, ‘Linguistically-based sub-sentential alignment for terminology extraction from a bilingual automotive corpus’, *Proceedings*

- of the 22nd International Conference on Computational Linguistics (Coling 2008), Manchester, United Kingdom, 529-536.
- * Fayruzov, T., De Cock, M., Cornelis, C. & Hoste, V.: 2008, 'Exploiting Grammatical Relations for Protein Relation Extraction and Role Labeling', *Proceedings of the Dutch-Belgian Information Retrieval workshop (DIR-2008)*, Maastricht, the Netherlands, 37-44.
 - * Lefever, E., Fayruzov, T. & Hoste, V.: 2007, 'A combined classification and clustering approach for web people disambiguation', *Proceedings of the 4th International Workshop on Semantic Evaluations*, Prague, Czech Republic, 105-108.
 - * Hoste, V., Vanopstal, K. & Lefever, E.: 2007, 'The Automatic Detection of Scientific Terms in Patient Information', *Proceedings of the Workshop on Acquisition and Management of Multilingual Lexicons (in conjunction with RANLP-2007)*, Borovetz, Bulgaria, 41-48.
 - * Hoste, V. & van den Bosch, A.: 2007, 'A Modular Approach to Learning Dutch Co-reference Resolution', *Proceedings from the first Bergen Workshop on Anaphora Resolution (WAR I)*, Mjolfell, Norway, 51- 75.
 - * Hoste, V. & Daelemans, W.: 2006, 'Comparing Learning Approaches to Language Learning. There is More to it Than Bias', *Proceedings of Benelearn 2006*, Ghent, Belgium, 131-138.
 - * Hoste, V. & De Pauw, G.: 2006, 'KNACK-2002: a Richly Annotated Corpus of Dutch Written Text', *Proceedings of the fifth international conference on Language Resources and Evaluation*, Genova, Italy.
 - * Hoste, V. & Daelemans, W.: 2005, 'Learning Dutch Coreference Resolution', *Proceedings of the Fifteenth Computational Linguistics in the Netherlands Meeting (CLIN 2004)*.
 - * Hoste, V. & Daelemans, W.: 2005, 'Comparing Learning Approaches to Coreference Resolution. There is More to it Than 'Bias'', *Proceedings of the Workshop on Meta-Learning (held in conjunction with ICML-2005)*, Bonn, Germany, 20-27.
 - * Decadt, B., Hoste, V., Daelemans, W. & van den Bosch, A.: 2004, 'GAMBL, Genetic Algorithm Optimization of Memory-Based WSD', *Proceedings of SENSEVAL-3*, Barcelona, Spain, 108-112.
 - * Hoste, V., Daelemans, W., Hendrickx, I. & van den Bosch, A.: 2002, 'Evaluating the Results of a Memory-Based Word-Expert Approach to Unrestricted Word Sense Disambiguation', *Proceedings of the Workshop on Word Sense Disambiguation: Recent Successes and Future Directions*, Philadelphia, PA, USA, 95-101.
 - * Daelemans, W. & Hoste, V.: 2002, 'Evaluation of Machine Learning Methods for Natural Language Processing Tasks', *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002)*, Las Palmas, Gran Canaria, 755-760.
 - * Hendrickx, I., van den Bosch, A., Hoste, V. & Daelemans, W.: 2002, 'Dutch Word Sense Disambiguation: Optimizing the Localness of Context', *Proceedings of the Workshop on Word Sense Disambiguation: Recent Successes and Future Directions*, Philadelphia, PA, USA, 61-66.
 - * Marsi, E., Busser, B., Daelemans, W., Hoste, V., Reynaert, M., van den Bosch, A.: 2002, 'Combining Information Sources for Memory-Based Pitch Accent Placement', *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP 2002)*.
 - * Hoste, V., Kool, A. & Daelemans, W.: 2001, 'Classifier Optimization and Combination in the English All Words Task', *Proceedings of SENSEVAL-2. Second*

International Workshop on Evaluating Word Sense Disambiguation Systems, New Brunswick: ACL, 83-86.

- * Hoste, V., Daelemans, W., Tjong Kim Sang, E. & Gillis, S.: 2000, 'Meta-Learning for Phonemic Annotation of Corpora', *Proceedings of the 17th International Conference on Machine Learning*, San Francisco: Morgan Kaufmann, 375-382.
 - * Hoste, V., Gillis, S. & Daelemans, W.: 2000, 'A Rule Induction Approach to Modeling Regional Pronunciation Variation', *Proceedings of the 18th International Conference on Computational Linguistics*, Saarbrücken, Germany, 327-333.
 - * Hoste, V., Gillis, S. & Daelemans, W.: 2000, 'Machine learning for modeling Dutch pronunciation variation', *Computational Linguistics in the Netherlands 1999. Selected papers from the Tenth CLIN Meeting*, Utrecht, the Netherlands, 73-84.
 - * Hoste, V., Daelemans, W., Tjong Kim Sang, E. & Gillis, S.: 2000, 'Meta-Learning for Phonemic Annotation of Corpora', *Proceedings of the Twelfth Belgium-Netherlands Artificial Intelligence Conference*, Tilburg, the Netherlands, 331-332.
 - * Hoste, V. & Daelemans, W.: 2000, 'Comparing Bagging and Boosting for Natural Language Processing Tasks, a Typicality Approach', *Proceedings of the Tenth Belgian-Dutch Conference on Machine Learning*, Tilburg, the Netherlands, 101-109.
- Dissertation D1
Véronique Hoste, *Optimization Issues in Machine Learning of Coreference Resolution*, Ph.D. thesis, 246 pages, Antwerp University, 2005.

11 Lectures/conference contributions

(to be updated)

- session chair at CLIN 2012, Tilburg, The Netherlands, January 2012.
- session chair at ABLA 2012, Antwerp, Belgium, February 2012.
- presentation: 'ABOP. A (semi) automatic patient information leaflet optimiser' (with Leona Van Vaernebergh), ABLA-2012, Antwerp, Belgium, February 2012.
- paper presentation: 'Cross-Domain Dutch Coreference Resolution' (with Orphee De Clercq and Iris Hendrickx), RANLP-2011, Hissar, Bulgaria, September 2011.
- paper presentation: 'Analysis and Reference Resolution of Bridge Anaphora across Different Text Genres' (with Iris Hendrickx and Orphee De Clercq), DAARC-2011, October 2011, Faro, Portugal, 2011.
- paper presentation: 'ParaSense or how to use Parallel Corpora for Word Sense Disambiguation' (with Els lefever and Martine De Cock), ACL-2011, Portland, US, 2011.
- paper presentation: 'Examining the Validity of Cross-Lingual Word Sense Disambiguation' (with Els Lefever), CICLing-2011, Tokio, Japan, March 2011.
- paper presentation: 'A posteriori Agreement as a Quality Measure for Readability Prediction Systems' (with Philip van Oosten), CICLing-2011, Tokio, Japan, March 2011.

- poster presentation: ‘ParaSense or how to use parallel corpora for Cross-Lingual Word Sense Disambiguation’ (with Els Lefever), CLIN-2011, Ghent, Belgium, February 2011.
- poster presentation: ‘Cross-Domain Dutch Coreference Resolution’ (with Orphee De Clercq), CLIN-2011, Ghent, Belgium, February 2011.
- Session chair at CLIN2011, Ghent, Belgium, February 2011.
- presentation: ‘Sentiment mining for suicide prevention’, The Political Economy of Surveillance: A Research Workshop, Milton Keynes, UK, September 2010.
- paper presentation: ‘Construction of a Benchmark Data Set for Cross-lingual Word Sense Disambiguation (with Els Lefever), LREC-2010, Valletta, Malta, May 2010.
- paper presentation: ‘Interacting semantic layers of annotation in SoNaR, a reference corpus of contemporary written Dutch, LREC-2010, Valletta, Malta, May 2010.
- presentation: ‘Sentiment mining for suicide prevention’, ESSSB-13, Rome, Italy, September 2010.
- presentation: ‘Fusion of Readability assessments for ranking in a possibilistic logic framework’ (with Dries Tanghe, Philip van Oosten), ATILA meeting, Oostende, September 2010.
- paper presentation: ‘Towards a learning approach for abbreviation detection and resolution (with Klaar Vanopstal and Bart Desmet), LREC-2010, Valletta, Malta, May 2010.
- paper presentation: ‘Towards an Improved Methodology for Automated Readability Prediction (with Philip Van Oosten en Dries Tanghe), LREC-2010, Valletta, Malta, May 2010.
- paper presentation: ‘Towards a Balanced Named Entity Corpus for Dutch (with Bart Desmet), LREC-2010, Valletta, Malta, May 2010.
- presentation ‘Hybrid Shallow Parsing Based Information Extraction for Glycosylation Curation (with Violaine Pillet, Anne-Lise Veuthey and Walter Daelemans), BioTM workshop, Ghent, Belgium, May 2010.
- presentation: ‘Towards a Gold Standard for Readability Prediction’, ATILA-2009, Corsendonk, Belgium, November 2009.
- presentation: ‘ABOP, an Automatic optimizer for Patient Information Leaflets’ (with Isabelle Delaere and Peter Velaerts), ATILA-2009, Corsendonk, Belgium, November 2009.
- presentation: ‘SemEval-2: Cross-lingual Word Sense Disambiguation’ (with Els Lefever), ATILA-2009, Corsendonk, Belgium, 10 november 2009.
- paper presentation: ‘Coreference Resolution on Blogs and Commented News’ (with Iris Hendrickx), 7th Discourse Anaphora and Anaphor Resolution Colloquium, Goa, India, November 2009.
- Session chair, RANLP-2009, Borovetz, Bulgaria, September 2009.
- presentation: ‘PRISE, a PRotein Interaction Search Engine’ (with Timur Fayruzov, Martine De Cock, Chris Cornelis et al.) , LTCI-2009, Ghent, Belgium, September 2009.
- presentation: ‘Coreference Resolution for Dutch Online Media Analysis’ (with Iris Hendrickx), LTCI-2009, Ghent, Belgium, September 2009.

- presentation: ‘Automatic Scoring of Text Readability’ (with Philp Van Oosten en Dries Tanghe), LTCI-2009, Ghent, Belgium, September 2009.
- paper presentation: ‘ABOP, Automatic Optimization of Patient Information Leaflets’ (with Isabelle Delaere, Claudia Peersman, Leona Van Vaerenbergh and Peter Velaerts), International Symposium on Data and Sense Mining, Machine Translation and Controlled Languages, Besançon, France, July 2009.
- paper presentation: ‘Task 3: Cross-lingual Word Sense Disambiguation’ (with Els Lefever), Workshop on Semantic Evaluations: Recent Achievements and Future Directions (SEW-2009), Boulder, VS, June 2009.
- paper presentation: ‘Fuzzy Ants Clustering for Web People Search’ (with Els Lefever, Timur Fayruzov and Martine De Cock), Second Web People Search Evaluation Workshop, Madrid, Spain, April 2009.
- paper presentation: ‘Language-independent bilingual terminology extraction from a multilingual parallel corpus’ (with Els Lefever and Lieve Macken), Annual Meeting of the Association of Computational Linguistics, Athens, Greece, April 2009.
- paper presentation: ‘Cultivating Trees: Adding Several Semantic Layers to the Lassy Treebank in SoNaR’ (with Ineke Schuurman and Paola Monachesi), International Workshop on Treebanks and Linguistic Theories, Groningen, The Netherlands, January 2009.
- presentation: ‘Language-independent bilingual terminology extraction from a multilingual parallel corpus’ (with Els Lefever en Lieve Macken), CLIN 2009, Groningen, Groningen, The Netherlands, January 2009.
- paper presentation: ‘Linguistically-based sub-sentential alignment for terminology extraction from a bilingual automotive corpus’ (with Lieve Macken and Els Lefever), Coling 2008, Manchester, UK, August 2008.
- presentation: ‘ABOP: Automatic PIL Optimizer’ (with Isabelle Delaere), LTCI-2008 (Second AUGent Workshop on Language Technology and Computational Intelligence), Ghent, August 2008.
- paper presentation: ‘Learning-based Detection of Scientific Terms in Patient Information’ (with Els Lefever, Klaar Vanopstal and Isabelle Delaere), LREC-2008, Marrakech, Morocco, May 2008.
- paper presentation: ‘Coreference Resolution for Extracting Answers for Dutch’ (with Iris Hendrickx, Gosse Bouma, Frederik Coppens, Walter Daelemans, Geert Kloosterman, Anne-Marie Mineur, Joeri Van Der Vloet and Jean-Luc Vershelde), LREC-2008, Marrakech, Morocco, May 2008.
- paper presentation: ‘Exploiting Grammatical Relations for Protein Relation Extraction and Role Labeling’ (with Timur Fayruzov, Martine De Cock and Chris Cornelis), DIR-2008, Maas-tricht, The Netherlands, April 2008.
- presentation: ‘Learning-based Detection of Scientific Terms in Patient Information’ (with Els Lefever and Klaar Vanopstal), CLIN 2007, Nijmegen, The Netherlands, December 2007.
- presentation: ‘COREA Wrap up: the corpus and evaluation’ (with Iris Hendrickx, Gosse Bouma, Frederik Coppens, Walter Daelemans, Anne-Marie Mineur, Joeri Van Der Vloet, Jean-Luc Vershelde), CLIN 2007, Nijmegen, The Netherlands, December 2007.

- paper presentation: ‘The Automatic Detection of Scientific Terms in Patient Information’, Workshop on Acquisition and Management of Multilingual Lexicons (collocated with RANLP-2007), Borovetz, Bulgaria, October 2007.
- paper presentation: ‘Disambiguation of the neuter pronoun and its effect on pronominal coreference resolution’ (with Walter Daelemans en Iris Hendrickx), 10th International Conference TSD 2007, Plzen, Czech Republic, September 2007.
- paper presentation: ‘Evaluating hybrid versus data-driven coreference resolution.’ (with Iris Hendrickx, Walter Daelemans), 6th Discourse Anaphora and Anaphor Resolution Colloquium, Lagos, Portugal, March 2007.
- poster presentation: ‘A combined classification and clustering approach for web people disambiguation.’ (with Els Lefever, Timur Fayruzov), Workshop Machine Learning for Natural Language Processing, Amsterdam, the Netherlands, May 2007.
- presentation: ‘The automatic resolution of “het” in a machine learning approach to Dutch coreference resolution’ (with Walter Daelemans and Iris Hendrickx), CLIN17, Leuven, January 2007.
- paper presentation: ‘Comparing Learning Approaches to Learning Language. There is More to it Than Bias’ (with Walter Daelemans), Benelearn conference, Ghent, May 2006.
- paper presentation: ‘Comparing Learning Approaches to Coreference Resolution. There is More to it Than Bias’ (with Walter Daelemans), Workshop on Meta-Learning, Bonn, Germany, August 2005.
- presentation: ‘Learning Dutch Coreference Resolution’ (with Walter Daelemans), Fifteenth Meeting of Computational Linguistics in the Netherlands, Leiden, The Netherlands, December 2004.
- paper presentation: ‘GAMBL Genetic Algorithm Optimization of Memory-Based WSD’ (with Bart Decadt, Walter Daelemans, Antal van den Bosch), Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (Senseval- 3), Barcelona, Spain, July 2004.
- paper presentation: ‘Combined Optimization of Feature Selection and Algorithm Parameters in Machine Learning of Language’, (with Walter Daelemans, Fien De Meulder and Bart Naudts), 14th European Conference On Machine Learning, Cavtat, Croatia, September 2003.
- presentation: ‘Optimization in Machine Learning of Natural Language’ (with Walter Daelemans), CLIF symposium, Brussels, April 2003.
- presentation: ‘Predicting pitch accents and phrase breaks from text. (with Erwin Marsi, Martin Reynaert, Walter Daelemans, Antal van den Bosch), 13th Meeting of Computational Linguistics in the Netherlands, Groningen, The Netherlands, November 2002.
- presentation: ‘Limitations of comparative machine learning methodology for language technology (with Walter Daelemans), 13th Meeting of Computational Linguistics in the Netherlands, Groningen, The Netherlands, November 2002.
- paper presentation: ‘Evaluating the results of a memory-based word-expert approach to unrestricted WSD. (with iris Hendrickx, Walter Daelemans, Antal van den Bosch), ACL 2002 Workshop on Word Sense Disambiguation: Recent Successes and Future Directions, Philadelphia, US, July 2002.

- paper presentation: ‘Evaluation of Machine Learning Methods for Natural Language Processing Tasks. (with Walter Daelemans), LREC 2002, Las Palmas de Gran Canaria, May 2002.
- paper presentation: ‘Dutch Word Sense Disambiguation: Optimizing the Localness of Context. (with Iris Hendrickx, Walter Daelemans, Antal van den Bosch), ACL 2002 Workshop on Word Sense Disambiguation: Recent Successes and Future Directions, Philadelphia, US, July 2002.
- ‘Limitations of comparative machine learning methodology. (with Walter Daelemans), ATILA meeting, Tilburg, The Netherlands, June 2002.
- presentation: ‘Interaction between named-entity recognition and base NP chunking for Dutch (with Walter Daelemans and Fien De Meulder), Twelfth Meeting of Computational Linguistics in the Netherlands, Enschede, The Netherlands, November 2001.
- presentation: ‘Memory-based all-words word-sense disambiguation for Dutch and English’ (with Walter Daelemans, Antal van den Bosch en Iris Hendrickx), Twelfth Meeting of Computational Linguistics in the Netherlands, Enschede, The Netherlands, November 2001.
- presentation: ‘Comparing Bagging and Boosting for Natural Language Processing Tasks: A Typicality Approach’ (with Walter Daelemans), Eleventh Meeting of Computational Linguistics in the Netherlands, Tilburg, The Netherlands, December 2000.
- paper presentation: ‘Comparing Bagging and Boosting for Natural Language Processing Tasks: A Typicality Approach’ (with Walter Daelemans, Steven Gillis), BNAIC, Tilburg, The Netherlands, November 2000.
- paper presentation: ‘A Rule Induction Approach to Modeling Regional Pronunciation Variation’ (with Walter Daelemans, Steven Gillis), 18th International Conference on Computational Linguistics, Saarbrücken, Germany, August 2000.
- paper presentation: ‘Meta-Learning for Phonemic Annotation of Corpora’ (with Walter Daelemans, Steven Gillis), 17th International Conference on Machine Learning, Stanford, US, July 2000.
- presentation: ‘A Machine Learning Approach to Phonemic Corpus Annotation’ (with Walter Daelemans, Steven Gillis), Tenth Meeting of Computational Linguistics in the Netherlands, Utrecht, The Netherlands, December 1999.

Invited

- ‘Automatic monitoring for cyberspace applications’, Citizen Science Conference, November 13, 2014. (with Karolien Poels)
- ‘Natural language processing and its role in digital humanities’, BAAHE Conference, November 28, 2014.
- ‘Sociale media en zelfmoordpreventie’, Mechelen, December 5, 2013.
- ‘Big Brother en de social media’, several lectures in fall 2013 and spring 2014.
- ‘Introducing language technology’, Madrid, May 28, 2013. (invitation by Julia Lavid)
- ‘Automatische termextractie’, Translation Symposium Belgacom-De Post-NMBS, June 4, 2013.

- ‘Big Brother en de social media’, Innerwheel Deinze, December 11, 2012.
- Solutions-2, June 27, 2012. (invitation by Anita Prinzie)
- ‘Cross-Lingual Word Sense Disambiguation. How to use parallel corpora for WSD?’, NLP seminar, CENTAL, University of Louvain La Neuve, December 15, 2011. (invitation by Richard Beaufort)
- ‘Slim zoeken met taaltechnologie’, Innovatielunch, Innovatiecentrum Oost-Vlaanderen, November 4, 2010. (invitation by Peter Rutten)
- ‘ABOP, Automatische BijsluiterOptimalisator’, workshop rond ‘Valorisatie van onderzoek in de HSW’. Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, February 12, 2010. (invitation by Godelieve Laureys)
- ‘Taaltechnologie’, Rotary Deinze, February 1, 2010.
- ‘ABOP, Automatische Bijsluiter-Optimalisator’, Taal en Bedrijf, November 19, 2008. (invitation by Peter Spyns)
- ‘Dutch coreference resolution. Issues and applications’, Corpus-Based Approaches to Coreference Resolution in Romance Languages (CBA-2008), November 14, 2008. (invitation by Toni Marti)
- ‘Language independent bilingual terminology extraction from a parallel corpus. The PSA test case’, University of Wolverhampton, November 5, 2008. (invitation by Ruslan Mitkov)
- ‘Introduction into to the Foundations and Challenges of Text Data Mining’, LTCI-2008, August 25, 2008. (invitation by Martine de Cock and Patricia Victor)
- ‘NLP challenges in biomedical text mining’, VIB, Ghent, November 21, 2007. (invitation by Yvan Saeys)
- ‘Framing Discourse as a classification approach’, Machine Learning for NLP 2007, Amsterdam, the Netherlands, May 16, 2007. (invitation by Maarten van Someren and Sophia Katrenko)
- ‘State-of-the-art translation technology: introduction and perspectives’, Telelingua Seminar, keynote speaker, Brussels, June 15, 2006. (invitation by Luc Van Haute)
- ‘Optimization in machine learning of language. A focus on the task of coreference resolution’, ILPS seminar, University of Amsterdam, Amsterdam, the Netherlands, June 9, 2006. (invitation by David Ahn)
- ‘Optimization issues in Machine Learning of Coreference Resolution’, Workshop on Anaphora Resolution (WAR), Mjolfjell, Norway, September 28, 2005. (invitation by Christer Johansson)
- ‘Learning Noun Phrase Coreference Resolution’, TKI Colloquium, Tilburg, the Netherlands, May 19, 2004. (invitation by Erwin Marsi)