

Curriculum Vitae

Surname: Vossen
Title: Professor Dr.
Given names: Petrus Theodorus Joseph Marie
Forename: Piek
Born: 24 - 10 - 1960 in Schaesberg (The Netherlands)
Address:
Domicile:
Telephone home:
Mobile:
http: www.vossen.info
E-mail: p.vossen@let.vu.nl
Current position: Professor Computational Lexicology, Faculty of Arts, Computational Lexicology & Terminology Lab (CLTL), VU *University* Amsterdam
Civil status: married, 2 children
Languages: Dutch (native language), English (fluently), German (fluently), (French)
Programming: Pascal, Lisp, C++, Java



University Degrees

- *Kandidaats* (Bachelor) Dutch Linguistics, University of Amsterdam, 1984.
- *Doctoraal* (MA) General Linguistics, subsidiaries Cognitive Psychology and Language Philosophy, University of Amsterdam, 1986.
- *Ph.D.* (cum laude) in Linguistics on Computational Lexicology and Lexicography at the University of Amsterdam, 1995 (promoters: Simon Dik, Lachlan McKenzie).
- *Professor Computational Lexicologie* at the Vrije Universiteit Amsterdam, 2006.

Career

- *2009-2011:*
DutchSemCor
Coordinator of "DutchSemCor: a NWO-project (380-70-011)"

The goal of DutchSemCor is to deliver a one-million word Dutch corpus that is fully sense-tagged with senses and domain tags from the Cornetto database (STEVIN project STE05039). 250K words of this corpus will be manually tagged. The remainder will be automatically tagged using three different word-sense-disambiguation systems (WSD), and will be validated by human annotators. The corpus data will be based on existing corpus material collected in the projects CGN, D-Col and SoNaR. These corpora have already been automatically annotated with morpho-syntactic tags and structures. The corpora will be extended where necessary to find sufficient examples for meanings of words that are less frequent and do not appear in the above corpora. The resulting corpus, for which we aim to offer the same balance in types of text as these basic resources, will be extremely rich in terms of lexical semantic information. Its availability will enable many new lines of research and technology developments for the Dutch language. In particular, it will enable research into the relation between language form and language interpretation, and as such it will be applicable in the fields of cognitive science, (psycho-)linguistics, language learning and language teaching, semantic web applications, information retrieval, machine translation, text mining, and document interpretation (summarization, topic segmentation). We foresee that the corpus will create new directions of research and technology development on a par with current developments for English.

- *2009-2013:*
Semantics of History

The research project Semantics of History develops a historical ontology and a lexicon that are used in a new type of information system that can handle the time-based dynamics and varying perspectives in historical archives. The system will integrate new insights in the ontological and linguistic analysis of the data that will follow from empirical and fundamental research. The work is carried out by two AIOs that are situated at the Faculty of Exact Science and the Faculty of Arts and is funded by the Interfaculty research institute CAMeRA. History is typically a record of different realities in time and specifically focuses on the changes in reality. Even stronger, the perception of history can be different for different participants and for different cultural and linguistic groups. Finally, the reflection on the past can be different based on our different views: history has been and will be re-written many times. Information systems of historical archives should handle the dynamicity in time and represent all realities at an equal level while at the same time they should define the relations, the invariables and changes across the realities. The units of change are events and typically in history events can be organized at different levels of change. The most constant elements are locations, people and dates but nevertheless many different structures are still possible, which need to be related relative to these more constant elements. Such a system should also allow users to classify and structure reality from any possible perspective when accessing the archives. Vast amounts of historical data are available as free text. The text itself can be related in time just as the events. For direct reporting and communication in the same time-frame there will be little distance between the communication date and the event

date. Historical documents on the other hand have a large distance between reporting and event date. We also expect that the linguistic expression for naming these events will be different; exhibiting high abstraction and others types of perspectives in historical reports as compared to actual news reports. A historical information system requires an innovative view on the semantics of events and the ways we can conceptualize these through language in different genres of documents.

- *2009-2013:*

From sentiments and opinions in texts to positions of political parties (Text2Politics).

The research project Text2Politics combines contemporary theories and methods in linguistics and political science to develop an automated research tool for rich text-mining. The transdisciplinary relevance of the project is that a carefully constructed mining tool for language-meaning research can be applied to enhance the Kieskompas (Electoral Compass) and prove useful in the social sciences in general. The research will give new insights into the complexity of language use, the linguistic modeling of subjectivity and the representation of this knowledge in a lexicon. It will also shed new light on the complex dimensionality of competition between political parties. The work is carried out by three AIOs that are situated at the Faculty of Social Sciences Sciences and the Faculty of Arts and is funded by the Interfaculty research institute CAMeRA.

- *2009 –*

TAALBANK (GTB) of the Institute of Dutch Lexicography (INL)
Member of the Advisory Board

- *2009 –*

LOT

Member of the Netherlands Graduate School of Linguistics LOT (Landelijke Onderzoekschool Taalwetenschap).

- *2009 –*

CLARIN-NL

Member of CLARIN-NL aims to design, construct, validate, and exploit a research infrastructure that is needed to provide a sustainable and persistent eScience working environment for researchers in the Humanities, and Linguistics in particular, who want to make use of language resources and the technology to use these resources for their research. The CLARIN-NL project is a large national project in the Netherlands which aims to play a central role in the Europe-wide CLARIN infrastructure (CLARIN-EU). CLARIN-NL offers scholars the tools to allow computer-aided language processing, addressing one or more of the multiple roles language plays (i.e. carrier of cultural content and knowledge, instrument of communication, component of identity and object of study) in the Humanities and Social Sciences.

- *2008-2011:*

KYOTO: Knowledge Yielding Ontologies for Transition-based Organization ICT -211423

Coordinator of KYOTO: a 7th EU FrameWork project withing the area of Digital Libraries/Conceptual characterization of various genres in written and spoken media, across languages (IST-211423).

The goal of KYOTO is to develop a content enabling system that provides deep semantic search and information access to large quantities of distributed multimedia data for both experts and the general public, covering a broad range of data from wide-spread sources in a number of culturally diverse languages. In this project we will target the languages: English, Dutch, Italian, Spanish, Basque, Chinese and Japanese. This powerful system crucially rests on an ontology linked to wordnets --lexical semantic databases--in a variety of languages. Concept extraction and data mining are applied through a chain of semantic processors that re-use the knowledge for different languages and for particular domains. The shared ontology guarantees a uniform interpretation for diverse types of information from different sources and languages. The system can be maintained by field specialists using a Wiki platform. KYOTO is a generic system offering knowledge transition for any domain of knowledge and information, across different target groups in society and across linguistic, cultural and geographic borders. KYOTO will be applied to the environmental domain and span global information across European and non-European languages.

▪ *2008 -*
FLARENET

Member of an international language resources network sponsored by the European Union. It addresses the development and exploitation of Language Resources (LRs) and Language Technology (LT) that will enable easy and natural access to the wealth of information and knowledge encapsulated in (written, spoken, multimodal) digital documents. FLAReNet expectation is to achieve:

- the largest network of LR and HLT players, ready to discuss information and knowledge on LRs and LT, and with appropriate mechanisms to share such information and disseminate it widely
- an extended picture of LRs and a recasting of its definition in the light of recent scientific, methodological, technological, social developments
- a consolidation of methods and approaches, common practices, frameworks and architectures
- a “roadmap” identifying areas where consensus has been achieved or is emerging vs. areas where additional discussion and testing is required, together with an indication of priorities
- a set of recommendations in the form of a plan of coherent actions for the EU and national organizations
- an European model for the LRs of the next years.

Bringing together the diverse approaches, efforts and technologies represented by its membership, FLAReNet will enable even greater progress toward community consensus.

▪ *2008*

Pilotgrant DutchSemCor: A Dutch Corpus with Word Senses from the Cornetto Database

Startsubsidie NWO Geesteswetenschappen, jan-aug. 2008. The startsubsidie will be used to prepare an NWO middel-groot investeringsproject to be submitted in September 2008. The goal of that final project is to deliver a corpus that is fully sense-tagged with senses, ontology tags, and domain tags from the Cornetto database. This corpus will play a key role in language technology research for Dutch and also in linguistic and cognitive research that relates linguistic form to meaning. Combining the best of both worlds, the corpus will be tagged using a combination of automatic techniques and

manual editing. Automatic tagging techniques include on the one hand supervised methods, which can be trained on already tagged subcorpora as training data, enabling them to tag other subcorpora, and on the other hand unsupervised techniques that rely on other sources such as the Cornetto database itself. It is to be expected that the manual editing of the corpus will feed back in the form of adaptations to the semantic database Cornetto.

- *2006, April – ongoing*
Professor Computational Lexicology at the Vrije Universiteit Amsterdam.
- *2006 – ongoing*
Global WordNet Grid
In 2006 Vossen launched the Global Wordnet Grid: the building of a complete free worldwide wordnet grid. This grid will be build around a shared set of concepts, such as the Common Base Concepts used in many wordnet projects. These concepts will be expressed in terms of Wordnet synsets and SUMO definitions. People from all language communities are invited to upload synsets from their language to the Grid. Gradually, the Grid will then be represented by all languages. The Grid will be available to everybody and will be distributed completely free.
- *2006 – 2008*
CORNETTO
Coordinator of “Combinatorial Relational Network for Language Applications” (financially supported by the Dutch-Flemish Language Union. Cornetto build a lexical semantic database for Dutch with rich vertical and horizontal semantic relations and combinatorial lexical constraints such as multiword expressions, idioms and collocations on the one hand, and lexical functions and frames on the other. The concepts were aligned with the English Wordnet so that ontologies and domain labels can be imported. The semantic layer has been validated with a formal ontology, to make it usable in Semantic Web environments. Cornetto covers 40K entries, including the most generic and central part of the language. The database goes beyond the structure and content of Wordnet and FrameNet. Vossen was the main project coordinator.
- *2005 – 2008*
ARABIC WORDNET
Constructing an Arabic Wordnet (AWN) in Parallel with an Ontology. This project aims at the development of wordnets for the Arabic languages and is sponsored by the American government and headed by Princeton University. Vossen will be responsible for the European development and coordination of building a lexical resource in Standard Arabic. AWN will be constructed according to the methods developed for EuroWordNet (EWN;Vossen 1998) and since applied to dozens of languages around the world. The EuroWordNet approach maximizes compatibility across wordnets and focuses on manual encoding of the most complicated and important concepts. Arabic WordNet will be mappable straightforwardly onto Princeton WordNet 2.0 and EuroWordNet, enabling translation on the lexical level to English and dozens of other languages. Several tools specific to this task will be developed. AWN will be a linguistic resource with a deep formal semantic foundation. Besides the standard wordnet representation of senses, word meanings are defined with a machine understandable semantics in first order logic. The basis for this semantics is the Suggested Upper Merged Ontology (SUMO) and its associated domain ontologies.

- *2002 – 2005*
MEANING project
Developing Multilingual Web-scale Language Technologies, Proposal/Contract no.: IST-2001-34460. The project tries to solve two complementary and interrelated tasks: Word Sense Disambiguation (WSD) and large-scale enrichment of Lexical Knowledge Bases. On the long term, the results of MEANING will be also useful for the purposes of the semantic web.
- *2002 – 2004*
PIDGIN project (financially supported by the Dutch government within the framework of the CIC-programme (Compete with ICT-competences) of the Ministries of Economic Affairs and of Educational and Cultural Affairs) Five organisations had started a project with the aim to improve the cross-lingual communication on the Internet. This name has been derived from the pidgin languages that were developed between people with different language backgrounds during colonial times. The Pidgin-project is an unique project which combines the sciences of IT and linguistics. By using special techniques of both sciences, Pidgin enables internet users all over the world: 1. to retrieve multi-lingual information from the internet (human-machine); 2. to communicate with their fellow users in their own native languages (human-human). Pidgin combines advanced search techniques with translation strategies. To improve the searching, the computer has to be able to "understand" language by way of syntactic and semantic analysis.
- *March 2001 - October 2009*
CTO at Irion Technologies, Delft. Irion Technologies develops multilingual language technology for many different languages. The technology combines state-of-the-art statistical engines developed at TNO with intelligent language modules. It uses a rich semantic network for 7 languages that is unique in the world, which makes it possible to develop concept-based information and knowledge management systems. Current main products are: cross-lingual concept-based search engine, document classification system, and a question-answer system. All products are available for English, Dutch, German, French, Italian and Spanish.
- *2001 – ongoing:*
Global Wordnet Association: Vossen is Founder and President of the Global WordNet Association. He founded GWA (with Christiane Fellbaum of Princeton University) as a public and non-commercial organization that provides a platform for discussing, sharing and connecting wordnets for all languages in the world. In this respect he organized 3 conferences (2002 India, 2004 Czech Republic, 2006 Korea, 2008 Hungary, 2010, Mumbai, India)
- *September 1999 – March 2001*
Senior researcher and manager at Sail Labs, Antwerp: Sail Labs is a long-term research laboratory developing language-technology of the future. I was responsible for developing novel technologies and strategies, and as such leading a team of 4 developers and 8 linguists. Core technologies were: cross-lingual information retrieval, classification, semantic ontologies and wordnet, disambiguation, corpus-based tuning and tailoring of resources and applications.

- *November 1997 – July 1999*
Member of the Ansi-committee on Ontology Standards. The aim of this group was to develop a world-wide standard for Ontologies and to provide the upper-level as a public Reference Ontology (c.a. 3,000 concepts).
- *May 1997 – September 1999*
Member of the Eagles Lexicon Interest Group. In this group we were developing Standards for encoding Lexical Semantic Information for Natural Language Processing products. Co-editor of the Preliminary Recommendations on Semantic Encoding.
- *March 1996 – September 1999*
EuroWordNet I and II
Project coordinator of the EuroWordNet project (LE2 4003 and LE 8328) and site-manager for the Dutch wordnet. The aim of the EuroWordNet project was to develop a multilingual database with wordnets for Dutch, Italian, Spanish, French, German, Czech, Estonian and English. Each wordnet was also linked to a so-called Inter-Lingual-Index, based on WordNet1.5. The database has been tested in an Information Retrieval application by Novell Linguistic Development.
- *May 1996 – February 1998*
Coordinator of the Dutch-Functional Grammar group in which a group of linguists from the research school IFOTT were developing a discourse grammar for Dutch based on Functional Grammar.
- *January 1994 - July 1996*
Project-manager of the Sift project (LRE 62030) for Amsterdam, in charge of a team of 4 researchers and programmers. The aim of the Sift project was to develop a text retrieval system that makes use of distributive semantic representations of words. Distributive representations consist of simple semantic features with weights indicating the relevance of these features for the concept associated with a word. Using these representations equivalence can be measured in a flexible and computationally tractable way. The main result for Amsterdam has been the development of the first version of the Amsterdam Lexicon System (ALS). This is a very efficient and fast object-oriented lexical database system, developed in C and running on Unix, Windows and Macs.
- *March 1995*
PhD (cum laude): "Grammatical and Conceptual Individuation in the Lexicon" (IFOTT, Amsterdam) describes a new model, the so-called Anchored Relational Model, for defining the syntactic and semantic properties of English and Dutch nouns based on differences in their pragmatic use. Using formal computer representations, the semantic and syntactic properties can be correlated in complex ways, making a further distinction between a cognitive level of meaning and a lexical semantic level.
- *August 1992 - October 1995*
Senior researcher on the Acquilex-2 project (BRA 7315) which was a sequel to Acquilex-1 in which also other resources such as corpora were considered to build up lexicons, and in which we cooperated with the publishers Van Dale in The Netherlands, Cambridge University Press in the UK, Bibliograf in Spain.

- *January 1990 – April 1994*
Researcher for LIKE (Linguistic Instruments in Knowledge Engineering), which was a research group studying information systems based on Linguistic modules and programs. LIKE had regular meetings, workshops and presentations.
- *Augustus 1989 - April 1992*
Senior researcher on the Aquilex-1 project (Esprit BRA-3030): Aquilex-1 was a joint enterprise of the Universities of Cambridge, Dublin, Pisa, Barcelona and Amsterdam in which we examined the feasibility of building a multilingual knowledge base usable in Natural Language Processing on the basis of information extracted from several Machine Readable Dictionaries. My work focused on LDOCE and a Van Dale monolingual Dutch dictionary, which I linked using Van Dale bilingual Dutch-English and English-Dutch dictionaries. I developed definition parsers for Dutch definitions and made programs for automatically converting the analyzed information into formal typed feature structure representations, which can be loaded into a lexical knowledge base. Finally, I developed some techniques and tools to (cross-linguistically) compare the semantic organization of different dictionaries.
- *October 1986 - July 1989*
Researcher in the Links-project in which I developed a parser for the definitions of nouns, verbs and adjectives of the Longman Dictionary of Contemporary English and stored the output (65,000 parse-trees) in the Nijmegen Linguistic Database. The Links-project was funded by the Dutch Council of Research (NWO) as project no. 300 - 169 - 007.
- **Post University Training:**
 - July 1988 - Augustus 1988: Attended the ESF Summer school on Computational Lexicology and Lexicography in Pisa, 25 July - 12 August 1988.
 - March 1990: Attended a LOT AIO-course on Psycholinguistics at the Max-Planck Institute in Nijmegen.
- *September 1983 - September 1986*
Studied **General Linguistics** at the University of Amsterdam with subsidiaries Cognitive Psychology and Language Philosophy. Specializations: Lexical Semantics, Computational Linguistics, Computational Lexicology.
- *September 1979 - January 1984*
Studied **Dutch Linguistics** at the University of Amsterdam.

Publications

- see attached document Publications or www.vossen.info/publications
- [Papers indexed by Google Scholar](#) (**h-index: 19**: Citations in [Bibtex](#), measured using [Harzing's Publish or Perish](#)), October 1, 2009.
- [Top 5 of my favourite publications](#)

Referents

- Prof. Dr. W. Martin, Department of Lexicology, Free University Amsterdam. Telephone: +31 20 444 6451
- Prof. Dr. J.L. Mackenzie (<http://www.lachlanmackenzie.com>).
- Prof. Dr. R. Scha, Computational Linguistics Department, University of Amsterdam. Telephone: +31 20 525 2077
- Prof. Dr. Y. Wilks, Computer Science Department, University of Sheffield. Telephone: +44 1142 221 805

Other Activities

- *Advisory & Research Committees:*
 - 2000 - ongoing: Co-Founder and Co-President of the Global WordNet Association.
 - 2010: Conference Chair and member of the Scientific Committee of the 5th Global Wordnet Conference, Mumbai, India, January 21-February 4, 2010.
 - 2010: Member of the Committee of the SemEval-2010: 5th International Workshop on Semantic Evaluations on Kyoto's subtask WSD17: All-words Word Sense Disambiguation on a Specific Domain.
 - 2009: Member of FLaReNet (Fostering Language Resources Network), 7th EU Framework Project.
 - 2009 - ongoing: Member of the Scientific Advisory Board (SAB) of CAMeRA: the Interfaculty research institute. The Scientific Advisory Board consists of prominent members from the four (i.e. Sociology, Computer Science, Clinical Psychology and Language & Communication) participating faculties and research groups.
 - 2009 - ongoing: Member of the Advisory Board of the Taalbank (GTB) of the Institute of Dutch Lexicography (INL).
 - 2009: Member of the committee for the Nederlandse TaalUnie to revise the Referentie Bestand Nederlands.
 - 2009: Member of the Scientific Committee of the Third International Workshop on Distributed Agent-based Retrieval Tools (DART'09), that will be co-located with WI'09 and IAT'09, Milan, Italy, September 15, 2009.
 - 2009: Member of the National Advisory Panel of CLARIN-NL.
 - 2009: Member of the Netherlands Graduate School of Linguistics LOT (Landelijke Onderzoekschool Taalwetenschap).
 - 2009: Member of the committee for the Nederlandse TaalUnie to revise the Referentie Bestand Nederlands.
 - 2008: Member of the Research Center for Advanced Media Research Amsterdam CAMeRA, VU University Amsterdam.
 - 2004 - 2009: Member of the Research Committee STEVIN (Sprak-en Taaltechnologische Essentiële Voorzieningen In het Nederlands); een

- Vlaams/Nederlands meerjarenprogramma voor Nederlandstalige taal- en spraaktechnologie door SenterNovem in samenwerking met de Nederlandse Taalunie, NWO, AWI, FWO en IWT.
- 2005 - 2007: Expert consultant for Constructing Arabic Wordnet in Parallel with an Ontology: This project aims at the development of wordnets for the Arabic languages and is sponsored by the American government and headed by Princeton University. Vossen will be responsible for the European development and coordination.
 - 2002 - 2008: Member of the Program Committee of IMIX: Interactieve Multimodale Informatie Extractie, NWO.
 - 2006: Member of the Scientific Committee for the 5th Conference on Language Resources and Evaluation LREC 2006.
 - 2006: Expert reviewer for the European Union.
 - 2004: Member of the Scientific Committee for the 4th Conference on Language Resources and Evaluation LREC 2004
 - 2002 - 2004: Expert consultant for BalkaNet: the EC-project BalkaNet aims at the development of wordnets for the Balkan languages.
 - 2001 - 2003: Member of the EACL-board
 - 2001 - 2002: Expert consultant for EuroTerm: the EC-project EuroTerm aims at the development of domain extension module for wordnets.
 - January 2000 - 2003: Expert consultant for Corpus Gesproken Nederlands.
 - February 2000 - 2002: Expert consultant in the Eagles/Isle project on standardization of multilingual resources for NLP.
 - June 2000: Expert consultant for Van Dale, Groot Woordenboek der Nederlandse Taal, CR-ROM edition.
 - April 1998: Referent for the IFOTT research school for appointing new PhD students.
 - November 1997 - July 1999: Member of the Ansi-committee on Ontology Standards. The aim of this group was to develop a world-wide standard for Ontologies and to provide the upper-level as a public Reference Ontology (c.a. 3,000 concepts).
 - May 1997 - September 1999: Member of the Eagles Lexicon Interest Group. In this group we were developing Standards for encoding Lexical Semantic Information for Natural Language Processing products. Co-editor of the Preliminary Recommendations on Semantic Encoding (report).
 - May 1996 -February 1998: Coordinator of the Dutch-Functional Grammar group in which a group of linguists from the research school IFOTT (now called ACLIC), were developing a discourse grammar for Dutch based on Functional Grammar.
 - April 1996: Referent for the IFOTT research school for appointing new PhD students.
 - 1993 August - 1995: Expert-consultant in the area of language and information-engineering for the EU (PB/EG nr. C24/10, 28193).
 - June 1992 - September 1994: Freelance consultant for Cambridge University Press and Van Dale Lexicography to set up lexical semantic databases.
 - January 1990 - April 1994: Member of LIKE (Linguistic Instruments in Knowledge Engineering), which was a research group studying information systems based on Linguistic modules and programs. LIKE had regular meetings, workshops and presentations.

- *PhD Committees:*

- fc. Promotor of Agata Cybulska, Amsterdam, the Netherlands.
- fc. co-Promotor of Roxane Segers, Amsterdam, the Netherlands
- fc. Promotor of **Isa Maks**, Amsterdam, the Netherlands.
- fc.: Promotor of **Bertie Kaal**, Amsterdam, the Netherlands.
- fc: Promotor of Annemieke van Elfrinkhof, Amsterdam, the Netherlands.
- fc: Co-promotor of **Francesco Ronzano**, Pisa, Italy / Amsterdam, the Netherlands.
- fc: Member of the PhD Committee of Marcel Thelen, Amsterdam, the Netherlands.
- 2009: Member of the PhD Committee of of **Nicole Grégoire**, Utrecht, the Netherlands, November 10, 2009.
- 2009: Member of the PhD Committee of **Ismail Fahmi**, Groningen, mei 2009.
- 2009: Member of the PhD Committee of **Willem Robert van Hage**, Amsterdam, January, 2009
- 2008: Member of the PhD Committee of **Lonneke van der Plas**, Groningen, October, 2008
- 2008: Member of the PhD Committee of **Ales Horac**, Masaryk, Czech Republic, 2008
- 2003: Member of the PhD Committee of **Frank Joosten**, KU Leuven, October, 2003.
- 2002: Member of the PhD Committee of **Anselmo Peñas**, Dpto. Lenguajes y Sistemas Informáticos, Universidad Nacional de Educación a Distancia (UNED)
- 1999: Member of the PhD Committee of **Eneko Agirre**, University of Bilbao.
- 1998: Member of the PhD Committee **German Rigau**, University Polytechnica de Catalunya, Barcelona.
- 1998: Member of the PhD Committee **Henny van der Vliet**, Faculty of Arts, Free University, Amsterdam.

- *Invited Speaker:*

2009

- Keynote speaker on "Kyoto" at eLexicography in the 21st century: new challenges, new applications, Louvain-la-Neuve, Belgium, October 22-24, 2009
- Invited Speaker on "Division of semantic labor over vocabulary and ontology layers", FLaReNet/SILT Workshop on Lexicon-Ontology Relationship in connection with GL 2009 - 5th International Conference on Generative Approaches to the Lexicon, Pisa, Italy, September 17-19, 2009.
- Presentation on "KAF: a generic semantic annotation format", 5th International Conference on Generative Approaches to the Lexicon (GL 2009). Pisa, Italy, September 17-19, 2009.
- Keynote speaker at the 25th Annual Conference of the Spanish Society for Natural Language Processing 2009 (SEPLN'09), San Sebastian, Spain, September 8-10, 2009

- Invited Speaker on "Kyoto" at EU-Japan ICT Cooperation Forum on ICT, Brussels, Belgium, July 2, 2009.
- Invited speaker on "Kyoto" at the EU Information and Networking Event for the 5th Call (ICT FP7 SO 4.3) on Intelligent Information Management, Luxemburg, May 12, 2009
- Keynote speaker on "Kyoto" at the Asian Language Resource Summit (ALRS), Phuket, Thailand, March 20-21, 2009
- Guest Lecture "From WordNet to Global WordNet for language technology", Course Language Engineering Applications, Center for Computational Linguistics, University of Leuven, Belgium, February 26, 2009
- Lecture Course Masters (College Wordnets, College Ontologies, College Kyoto): Theme Language Acquisition & Documentation: Multilingual Lexical Framing, MA, Faculty of Arts, VU University Amsterdam, February 2 - June 10, 2009
- Presentation on "The Semantics of History", Seminar over Semantisch Netwerk Technologie, VU University Amsterdam, January 22, 2009

2008

- Guest Lecture for the ICT Onderwijscentrum of the Vrije University Amsterdam, Amsterdam, the Netherlands, December 10, 2007
- Keynote speaker at Kyoto at the 10th Lustrum of NL Term, Amsterdam, The Netherlands, October 25, 2008
- Invited speaker at workshop HLT Collaboration between South Africa and the Low Lands, MonkeyValley, South Africa, November 24-25, 2008
- Invited speaker at VKS / KNAW-workshop, Amsterdam, The Netherlands, October 14, 2008
- Invited speaker at LCC, Amsterdam, the Netherlands, October 9, 2008
- Presentation on Cornetto at the Stevin Programmadag, Nederlandse TaalUnie, Hoeven, the Netherlands, September 21, 2008.
- Keynote speaker on Kyoto at DART 2008: 2nd Workshop on Distributed Agent-based Retrieval Tools, Cagliari, Italy, 10 September 2008
- Panel on Lexical Units Compounds as Words at the Fourth International GlobalWordNet Conference - GWC 2008, Szeged, Hungary, January 22-25, 2008

2007

- Guest Lecture on Corpusgebaseerd tekstonderzoek for the ICT Onderwijscentrum of the Vrije University Amsterdam, December 13, 2007.
- Guest Lecture on Corpus-based Methods: "Automatic term extraction from domain corpora", University Nijmegen, November 26, 2007
- Invited speaker on Taal, Intelligentie en Betekenis" on the Lustrum symposium Kunstmatige Intelligentie, Ontwikkelingen en Toepassingen in Muziek, Taal en Geneeskunde" van het Genootschap Physica, Alkmaar, October 27, 2007
- Invited Speaker on Global Wordnet Grid on the 6th International PLAIN Language Conference, Amsterdam, Netherlands, October 11-14, 2007.
- Invited speaker on Global Wordnet Grid at the Politechnika Koszalin, Koszalin, Poland, October 8, 2007.
- Keynote Speaker on Global WordNet at the 3rd Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, Poznan, Poland, October 5-7, 2007
- Invited Speaker on the Stevin Programmadag, Nederlandse TaalUnie, Hoeven, September 21, 2007

- Invited speaker on "Cornetto: Een Combinatorische Semantische Database voor het Nederlands" for the Capita Selecta of the Instituut voor Nederlandse Lexicologie (INL), Leiden, May 22, 2007.
- Invited speaker on Global WordNet and Global WordNet Grid at the University of Potchefstroom, South Africa, March 19, 2007.
- Invited speaker on Cornetto and Wordnets at the University of Potchefstroom, South Africa, March 20, 2007.
- Invited speaker at the "African Languages WordNet Workshop/Training" of the Department African Languages of the University of South Africa, Unisa), Pretoria, South Africa, March 21-23, 2007.
- Invited speaker on "Connecting the Universal to the Specific: Towards the Global Grid", at the The First International Workshop on Intercultural Collaboration (IWIC 2007), Kyoto, Japan, January 25-26, 2007.
- Presentation on "The design of the Cornetto Database", In: Proceedings of CLIN 17, Centre for Computational Linguistics, Faculty of Arts of the University of Leuven, January 12, 2007

2006

- Presentation on "Er kan meer dan men doet met de Cornetto Database" at the TST-dag, Rotterdam, November 30, 2006.
- Presentation on Cornetto at the Stevin Programmadag, Antwerpen, September 11, 2006.
- Invited speaker: Lexical ambiguity within an application perspective at the Language Technology and Resources Colloquium of the The Department of Language and Speech, Radboud University Nijmegen, June 6, 2006.
- Invited speaker: Communicative evolution: from strings to words to expressions to concepts to intentions at the ICT-Kenniscongres 2006, Amsterdam, RAI, April 10 and 11, 2006.

2005

- Invited speaker: Demo on the search engine of Irion at the CLIN dag : CLIN 2005: The 16th Meeting of Computational Linguistics in the Netherlands, Amsterdam, December 16, 2005.
- Invited speaker: "Building Wordnets", International Seminar on Wordnets in general and DanNet: a Danish WordNet, based on the Princeton WordNet and EuroWordNet, University of Copenhagen, Denmark, 13 October 2005.
- Invited speaker: Kickoff-meeting of a new project "Constructing Arabic Wordnet in Parallel with an Ontology", University of Princeton, USA, August 30-September 5, 2005.
- Videopresentation on "Het RBN als lexicon in een classificatiesysteem" for the TST-dag, Antwerpen, July 4, 2005.
- Invited speaker: "A centralized approach to language resources", Multilingualism Meeting EU, Luxembourg, June 6, 2005.
- Invited speaker: "Crossing the chasm between statistic processing and language understanding", University of Amsterdam, Amsterdam, May 19, 2005.
- Invited speaker: Wat wil de industrie van de kennisinfrastructuur?, Tweede brokerage event - Domein Taaltechnologie: "Taal- en spraaktechnologie en de ICT-markt", Nederlandse Taalunie, Antwerpen, 2 maart 2005.
- Invited speaker: Meaning-2 for the 5th EU Call IST 2005-2006 Work Programme 2.5.7 Multimodal Interfaces, Luxembourg, February 28, 2005.

2004

- Invited speaker: Cornetto: a Combinatoric and relational network for language technology, International Colloquium on "Word structure and lexical systems: models and applications", University of Pavia, Italy, 16-17 December 2004.
- Invited speaker: Wordnets for Information Retrieval: a hole in one, organizer: Omnipaper, Leuven. Belgium, 2 december 2004.
- Invited speaker: Interactieve Workshop Taal & Spraak, Senter, The Hague, The Netherlands, 15 september 2004.
- Invited speaker: Search Engine Meeting, The Hague, The Netherlands, 19-20 April 2004.

2003

- Invited speaker Wordnet, EuroWordNet and Global Wordnet at Dublin University College Ireland, March 22, 2003.
- Invited speaker: Senter, Zwolle
- Tutorial Wordnet, EuroWordNet and Global Wordnet, International Conference RANLP-2003 (Recent Advances in Natural Language Processing), Borovets, Bulgaria, 10-12 September 2003.

2002

- Invited Speaker Content for content. XVIII Conference of SEPLN in Valladolid, Spain. 11-13 September 2002

2001

- Invited Speaker "Tuning Document-Based Hierarchies with Generative Principles at the ISSCO workshop on the generative lexicon, Geneva, April 2001 (ISSCO)
- Invited Speaker WordNet, EuroWordNet and Global Wordnet at the 'Computing with Logic, Language, and Information' Seminar. Institute for Logic, Language and Computation University of Amsterdam (ILLC)
- Invited Speaker Multilingual Language Technology as a Commercial Business , SPINN seminar, Copenhagen, Denmark, 26 -27 October 2001.

2000

- Invited Speaker for the SNLP 2000 Conference in Chiang Mai, Thailand, May 2000.
- Invited Speaker for the International Symposium Computational Linguistics for the New Millennium: Divergence or Synergy?" Heidelberg, 21-22 July 2000.

< 2000

- Course on the use of Wordnets in NLP at the Faculty of Humanities, University of Amsterdam, 1999.
- Lecture on Functional Lexicology at the Amsterdam Summer School, University of Amsterdam, 1992.

- *Organized Conferences and Workshops*

2010

- Organizer of the 5th Global WordNet Conference, Mumbai, India, January 31 - February 4, 2010.

- Co-Organizer of the SemEval-2010: 5th International Workshop on Semantic Evaluations on Kyoto's subtask WSD17: All-words Word Sense Disambiguation on a Specific Domain.

2009

- Organizer of the 1st Workshop of Kyoto, Amsterdam, the Netherlands, February 2-3, 2009.

2008

- Organizer of the Workshop on "Linguistic Studies of Ontology: From Lexical Semantics to Formal Ontologies and Back" at the 18th International Congress of Linguists (CIL18), Seoul, Republic of Korea, July 21-26, 2008.
- Organizer of the Ontolex workshop at LREC 2008, Marrakech, Turkey, May 31, 2008.
- Organizer of the DutchSemCor Workshop, Vrije University Amsterdam, May 16, 2008.
- Organizer of the 4th Global Wordnet Conference, Szeged, Hungary, January 22-25, 2008.

2007

- Organizer of the workshop on L&TC WordNet Special Track :the development of new WordNets, related tools and WordNet applications in AI, during the 3rd Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, Poznan, Poland, October 5-7, 2007.
- Organizer of the workshop on Evaluating WSD on Cross-Language Information Retrieval on Sem-Eval 2007, the 4th International Workshop on Semantic Evaluations, during the ACL, Prague, Czech Republic, June 23-30, 2007.
- Organizer of the Cornetto workshop, Vrije Universiteit Amsterdam, June 7, 2007.
- Organizer of the the "African Languages WordNet Workshop/Training" for the Department African Languages of the University of South Africa, Unisa, Pretoria, South Africa, March 19-23, 2007.
- Organizer of the The First International Workshop on Intercultural Collaboration (IWIC 2007), Kyoto, Japan, January 25-26, 2007.

2006

- Organizer of the 3rd International Global Wordnet Conference, Jeju Island, South Jeju (Seogwipo), Korea, January 22-26, 2006.

2005

- Organizer of the 2nd Workshop on the MEANING Project, Trento, Italy, February 3-4, 2005.

2004

- Organizer of the 2nd International Global Wordnet Conference, Brno, Czech, 20-23 January 2004.

2002

- Organizer of the 1st International Global Wordnet Conference, Mysore, India, 19-23 January 2002.

2001

- Organizer and panelist of the ACL Panel Global Wordnet, Standardizing and Linking wordnets with particular attention to Word Sense Inventory, Representation, and Distinction, EACL/ACL Panel, 6 July 2001.
- Panelist on ACL Panel Senseval domain specific sense disambiguation, ACL 2001, Toulouse, EACL/ACL Panel, 6 July 2001.

1998

- Member of the organizing committee and attendant of the Ontology Standards workshop at Heidelberg, on behalf of the Ansi-committee on Ontology Standards. Heidelberg, Germany.

1997

- Coordinator and chair of the ACL/EACL-97 workshop: Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications, Madrid, 12 July 1997.

1992

- Member of the organizing committee of the 6th Euralex International Conference on Lexicography in Amsterdam, August 1994.

- *Program Committees for conferences and workshops:*

2010

- Member of the organizing and program committee of the 5th Global Wordnet Conference, Mumbai, India, January 31-February 4, 2010.
- Member of the program committee of the SemEval-2010: 5th International Workshop on Semantic Evaluations on Kyoto's subtask WSD17: All-words Word Sense Disambiguation on a Specific Domain.

2009

- Member of the program committee of the 4th Language and Technology Conference, Poznan, Poland, November 6-8, 2009.
- Member of the program committee of the Third International Workshop on Distributed Agent-based Retrieval Tools (DART'09), co-located with WI'09 and IAT'09, Milan, Italy, September 15, 2009.
- Member of the program committee of the joint ACL-IJCNLP 2009 conference in the area "Semantics", sub-area "Word Sense Disambiguation", Singapore, August 2-7, 2009.
- Member of the program committee of the Conference on Empirical Methods in NLP (EMNLP 2009) in the area: "Multilinguality", Singapore, August 6-7, 2009.
- Member of the program committee of the Conference on Empirical Methods in NLP (EMNLP 2009) in the area: "NLP Applications", Singapore, August 6-7, 2009.
- Member of the program committee of EACL 2009 in the area: "Lexicon, lexical databases, language resources, ontologies", Athens, Greece, March 30 - April 3, 2009.
- Member of the program committee of EACL 2009 in the area: "Compositional semantics and textual entailment", Athens, Greece, March 30 - April 3, 2009.

- Member of the program and organizing committee of the 1st Workshop of Kyoto, Amsterdam, the Netherlands, February 2-3, 2009.

2008

- Member of the program committee of the Fifth International Conference on Formal Ontology in Information Systems FOIS 2008, Saarbrücken, Germany, October 31 - November 3, 2008.
- Member of the program committee of the Conference on Empirical Methods in Natural Language Processing EMNLP08, Waikiki, Honolulu, Hawaii, from October 25-27, 2008.
- Member of the program committee of COLING 2008, Manchester, August 18-22, 2008.
- Member of the organizing and program committee of the Workshop on "Linguistic Studies of Ontology: From Lexical Semantics to Formal Ontologies and Back" at the 18th International Congress of Linguists (CIL18), Seoul, Republic of Korea, July 21-26, 2008.
- Member of the program Committee of the 46th Annual Meeting of the Association for Computational Linguistics, (ACL-08), Columbus, Ohio, USA, June 15-20, 2008.
- Member of the scientific committee of LREC 2008, Marrakech, Turkey, May 26-June 1, 2008.
- Member of the organizing and program committee of the Ontolex workshop at LREC 2008, Marrakech, Turkey, May 31, 2008.
- Member of the organizing and program committee of the DutchSemCor Workshop, Vrije University Amsterdam, May 16, 2008.
- Member of the organizing and program committee of the 4th Global Wordnet Conference, Szeged, Hungary, January 22-25, 2008
- Member of the program committee of ICGL 2008, the First International Conference on Global Interoperability for Language Resources, Hong Kong, January 9-12, 2008.

2007

- Member of the organizing and program committee of the L&TC WordNet Special Track at the 3rd Language & Technology Conference LTC'07: Human Language Technologies as a Challenge for Computer Science and Linguistics, Poznan, Poland, October 5-7, 2007.
- Member of the program committee of the RANLP 2007, Borovets, Bulgaria, 27-29 September 2007.
- Member of the program committee of the Workshop Acquisition and Management of Multilingual Lexicons (AMML'2007) held at RANLP 2007, Borovets, Bulgaria, 26 September 2007.
- Member of the program committee of the ACL 2007, Prague, Czech Republic, June 23-30, 2007.
- Member of the organizing and program committee of the workshop on Evaluating WSD on Cross-Language Information Retrieval on Sem-Eval 2007, the 4th International Workshop on Semantic Evaluations, during the ACL 2007, Prague, Czech Republic, June 23-30, 2007.
- Member of the program committee of the EMNLP-CoNLL 2007, Conference on Empirical Methods in Natural Language Processing and Conference on Computational Natural Language Learning (joint meeting following ACL 2007), Prague, Czech Republic, June 28-30, 2007.

- Member of the organizing and program committee of the Cornetto workshop, Vrije Universiteit Amsterdam, June 7, 2007.
- Member of the organizing and program committee of the "African Languages WordNet Workshop/Training" of the Department of African Languages of the University of South Africa, Unisa), Pretoria, South Africa, March 19-23, 2007.
- Member of the organizing and program Committee of the The First International Workshop on Intercultural Collaboration (IWIC 2007), Kyoto, Japan, January 25-26, 2007.

2006

- Member of the program committee for the COLING/ACL 2006 Conference, Sydney, Australia, July 17-21, 2006.
- Member of the scientific committee for the 5th Conference on Language Resources and Evaluation LREC2006, Genoa, Italy, May 22-28, 2006.
- Member of the program committee of the Workshop "Lexicon, lexical databases, ontologies, language resources" at the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL), April 3-7, 2006, Trento, Italy.
- Member of the organizing and program committee of the 3rd Global Wordnet Conference, Jeju Island, Korea, January, 22-26, 2006

2005

- Member of the program committee of RANLP 2005, September 21-23, 2005, Borovets, Bulgaria.
- Member of the program committee of the 2nd Workshop organized by the MEANING Project, February 3-4, 2005, Trento, Italy

2004

- Member of the program committee of Coling 2004, August 2004, Geneva.
- Member of the organizing and program committee of the 2nd Global Wordnet Conference, January 2004, Brno, Czech.
- Member of the scientific committee of LREC 2004, May 2004.
- Member of the program committee of First Workshop on International Proofing Tools and Language Technologies in Patras, Greece, June 2004.

2003

- Member of the programm committee of the 1st MEANING Workshop on Word Sense Disambiguation and Lexical Acquisition, Donostia, Basque Country, April 11-12, 2003
- Member of the program committee of GL2003: 2nd International Workshop on Generative Approaches to the Lexicon, University of Geneva, Geneva, Switzerland, 15-17 May 2003.
- Member of the program committee of International Conference RANLP03 (Recent Advances in Natural Language Processing), Borovets, Bulgaria, 10-12 September 2003.

2002

- Member of the program committee of Coling 2002, Area Computational Lexicons.
- Member of the program and organizing committee of the 1st Global Wordnet Conference, Bangalore, India (report).
- Member of the program committee of Coling 2002 workshop Building and Using Semantic Networks, COLING 2002, Taipei, Taiwan, 1 September 2002.

- Member of the program committee of Workshop LREC 2002: Ontologies and Lexical Knowledge Bases, Las Palmas, Canary Islands - Spain, 27-28 May 2002.
- Member of the program committee of LREC 2002 Workshop on WordNet Structures and Applications for the Lesser-Studied Languages (WSA).
- Member of the program committee of LREC 2002 Workshop on Wordnet Structures and Standardization and how these affect Wordnet Applications and Evaluation (report)
- Member of the program committee of LREC 2002 Workshop Customizing Knowledge in NLP Applications: Strategies, Issues and Evaluation.
- Member of the program committee of Workshop on Using semantics for information retrieval and filtering, state of the art and future research" at LREC 2002, Las Palmas, Canary Islands, 29 May - 2 June 2002.
- Member of the program committee of OntoLex02
- Member of the program committee of NLP symposium called SNLP-2002, as a member of the International Advisory Committee, Thailand.

2001

- Member of the program committee of the ACL Panel Global Wordnet, Standardizing and Linking wordnets with particular attention to Word Sense Inventory, Representation, and Distinction, EACL/ACL Panel, July 2001.
- Member of the Senseval-2 program committee.
- Member of the program committee for the NAACL workshop WordNet and Other Lexical Resources: Applications, Extensions and Customizations, Pittsburg, June 2001.
- Member of the program committee ESSLLI 2001 workshop on Semantic Knowledge Acquisition and Categorization.
- Member of the program committee FOIS-2001, International Conference on Formal Ontology in Information Systems, (<http://www.fois.org>), Ogunquit, Maine, 17-19 October 2001.
- Member of the program committee of NLPRS 2001, 6th Natural Language Processing, Pacific Rim Symposium, National Center of Science, Tokyo, Japan, 27 - 29 November, 2001.
- Member of the program committee of RANLP 2001.
- Member of the program committee of ACL Panel: Senseval domain specific sense disambiguation, ACL 2001, Toulouse, EACL/ACL Panel, Friday, 6 July 2001.

2000

- Referent for NAANLP 2000.
- Referent for ACL 2000.

1999

- Member of the program committee of ELDA project proposals.
- Member of the program committee of ACL Siglex workshop, University of Pittsburgh, May 1999.

1998

- Referent for the ACL/Coling-98 workshop The Computational Treatment of Nominals. Organised by Federica Busa (Brandeis University), Inderjeet Mani (The MITRE Corporation), Patrick Saint Dizier (IRIT, Université Paul Sabatier)

- *Journal and Book Reviews:*

- 2007/2008: Referent for *Literator*, Tydskrif vir besondere en vergelykende taal en literatuurstudie, ed. H. Viljoen. Potchefstroom, Zuid-Afrika.
- 2007/2008: Referent for Semantic representation in Functional Discourse Grammar, a special issue of *Lingua* by K. Hengeveld and G. Wanders (eds).
- 2005: Referent for the Internation Journal of Lexicography, "Semantic Frames".
- 2004: Referent for the Journal of Logic, Language and Information: "Handbook for Language Engineers", edited by Ali Farghaly, Center for the Study of Language and Information, Stanford, California, December 2003.
- 2004: Referent for the Journal Traitement Automatique des Langues, Special issue on "Electronic Dictionaries".
- 2003: Referent for the Journal Computer Speech and Language, "Word Sense Disambiguation Using Untagged Corpus at Sense Level", Authors: I Nica, M A Marti, A Montoyo, S Vasquez.
- 2003: Referent for the Journal Computer & the Humanities, "Semantic Roles as Slots in OIL Ontologies", Authors: Cybulka/Martinek
- 2001: Referent for the Journal Traitement Automatique des Langues (TAL).
- November 1998: Referent for the Journal Information Processing & Management, School of Communication, Information and Library Studies and Infoproman