CURRICULUM VITAE
Leo Wanner

February 2024

PERSONAL INFORMATION

Date of Birth:	May, 8th 1959
Place of Birth:	Tschernigovskij, former USSR
Citizenship:	German

DEGREES

PhD in Computational Linguistics
    (University of the Saarland, Saarbrücken, Germany, 27.04.1997)
Diploma (Ms.Sc.) in Computer Science
    (University of Karlsruhe, Germany, 31.03.1988)

PRESENT ADDRESS

Departament de Tecnologies de la Informació i les Comunicacions
Universitat Pompeu Fabra
C. Roc Boronat, 138
08018 Barcelona, Spain

tel.: ++34 93 542 2241
telefax.: ++34 93 542 2517
e-mail: leo.wanner@upf.edu

PRESENT OCCUPATION

ICREA* Research Professor and Adjunct Professor at DTIC, UPF

---

*ICREA: Institució Catalana de Recerca i Estudis Avançats 'Catalan Institution for Research and Advanced Studies'

1. Education (Transcripts, scrolls, etc. are available on request)

— University of the Saarland, Saarbrücken, Germany. SS 92–WS 96/97 Postgraduate studies at the Department of Applied Linguistics. PhD dissertation title: 'Exploring Lexical Resources for Text Generation in a Systemic Functional Language Model'.

— Summer School in Computational Linguistics, Prague, Czechoslovakia. Summer, 1991.

— Karlsruhe University, Germany. WS 1981/82–WS 1987/88. Major subject: Computer Science (Dipl. Inform., i.e., M.S. in Computer Science). Dissertation title: 'XKB: A Tool for Exploring Object-Oriented Knowledge Bases'.

— Highschool in Nagold, Germany. 1975–1981.

— Highschool in Tschernigovskij, former USSR. 1966–1974.

Languages spoken:

German, Russian:	native
English, French, Spanish:	fluent
Catalan:	basic

2. Positions held

- As staff member:

  - January 2005–*continuing*. ICREA Research Professor and Adjunct Professor at the Department of Information and Communication Technologies, Pompeu Fabra University (UPF), Barcelona, Spain.

  - September 2013–October 2014. Coordinator of the Master "Interactive Intelligent Systems" at the Department of Information and Communication Technologies, Pompeu Fabra University (UPF), Barcelona, Spain.

  - September 2006–June 2010. Coordinator of the Doctorate Programme of the Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Spain.

  - September 2008–December 2009. Associate Head of the Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Spain.

  - January 2005–December 2009. Project Manager at the Barcelona Media Research Centre, Barcelona, Spain.

  - September 2003–December 2004. Ramon y Cajal Research Fellow, Institute for Applied Linguistics, Pompeu Fabra University, Barcelona, Spain.

  - April 1999–December 2004. Assistant Professor at the Computer Science Department, University of Stuttgart, Germany.

- June 1997–March 1999 Research Fellow at the Computer Science Department, University of Stuttgart, Germany.
- March 1995–October 1996. Senior Researcher at the Computer Science Department, University of Waterloo, Canada.
- May 1993–January 1995. Research Fellow at the Romance Languages Department, University of Stuttgart.
- April 1988–April 1993. Research Fellow at the Integrated Publication and Information Systems Institute (IPSI) of the German National Research Centre for Mathematics and Computer Science (GMD, now Fraunhofer Gesellschaft), Darmstadt, Germany.

- As consultant:

  - October 1996–March 1999. LEXIQUEST (form. ERLI), Paris. Consulting in the area of text generation.

- As student:

  - April, 1987–March, 1988. Research Assistant at the Research Centre for Computer Science, University of Karlsruhe.
  - April, 1985–September, 1986. Research Assistant at the Department of Computer Science, University of Karlsruhe.

3. Research Stays (of at least four weeks)

1. 27.10.2014 – 26.11.2014. Visiting Professor at the Human-Centred Multimedia Institute, Computer Science Department, University of Augsburg, Germany.

2. 1.1.2007 – 31.3.2007. Visiting Researcher at the Center for Computational Learning Systems, Computer Science Department, Columbia University, New York.

3. March, 2002. Visiting Professor at the Linguistics Department, Université Paris 7 (Denis Diderot).

4. September, 2000. Visiting Researcher at the Linguistics Department, Université de Montréal.

5. November, 1997. Guest researcher at the Computing Research Laboratory (CRL), Computer Science Department, New Mexico State University, Las Cruces.

6. October, 1996 – December, 1996. Visiting Researcher at the Information Sciences Institute (ISI) of the University of Southern California.

7. March, 1992 – April, 1992. Guest researcher at the Linguistics Department, University of Sydney.

8. May, 1990 – June, 1990. Guest researcher at the Linguistics Department, Université de Montréal.

4. Current Research Foci

Automatic Multilingual Text Generation:

Automatic generation of multilingual spoken and written information is one of my principal fields of research since 1988. Special areas of my interest include large scale generation grammars, graph transduction-based and neural generation from fused ontological image and text data, and problems of discourse and sentence planning. In this context, I am also interested in multilevel annotation schemata of corpora for generation. A very important matter of concern is to me the development of operational generators that can be applied in practice. So far, we worked and continue to work in the scope of European R&D initiatives on applications related, among others, to the provision of air quality and meteorological information, crisis management, media production, and artist and architect support.

(Corpus-based) Computational Lexicography and Lexicology and CALL:

I am active in the field of corpus-oriented lexicography/lexicology since the early nineties. Currently, my primary area of research in this field is the automatic extraction and classification of lexical information (and, in particular, of idiosyncratic word combinations, i.e., collocations) from large multilingual text corpora exploiting machine learning techniques. Since collocations are the key for successful learning of a language, I am also interested in problems related to automatic collocation recognition and collocation error correction in Computer Assisted Language Learning (CALL) and contrastive study of the use of collocations across languages.

Concept and Conceptual Relation Extraction from Written Material:

More recently, I became interested in the extraction of concepts and concept relations from written material. The concrete goal is to develop machine learning-based techniques for the concept and relation distillation in order to obtain an accurate picture of the content of the analyzed material and thus be able to map it onto (ontological) knowledge representations for downstream applications.

Deep neural language analysis:

Another of my more recent research foci is multilingual data-driven (deep) neural natural language analysis. I am interested in developing strategies for the exploration of how and where (in which layers) neural models encode and use linguistic information. So far, we targeted syntactic information, but plan to consider in the future also lexical (and, in particular, idiosyncratic) information. The purpose of this research is to facilitate targeted improvement of the performance of the models for specific tasks and to contribute to the explainability of the neural models, which is one of the big challenges faced when using such models for practical applications.

Communicative Structure in Written and Spoken Language Production:

Information (or, in terms of the Meaning-Text Theory, *Communicative*) Structure reflects the communicative intention of the speaker. We can assume that automatic language production is in general intention-driven. In our previous work, we have shown that, e.g., in the context of domain-specific report generation, the Communicative Structure can be derived from the user profile, the discourse structure and the context using a rule-based approach. The first part of the research question that remains to be solved is how we can derive a fine-grained Communicative Structure from training samples using data-driven techniques. The second part of this research question is how the derived structure can drive written and spoken language (in particular, expressive prosody) generation.

Author profiling and abusive language detection and classification:

Inspired by the work of my former PhD Students Joan Soler and Paula Fortuna, I moved into the areas of author profiling / identification and abusive language (more specifically, hate speech) detection. In our work on hate speech detection, we explored different annotation schemata for labeling hate speech corpora and the generalizability of hate speech detection/classification models. We have shown that due to the subjective nature of hate speech, which implies that what is perceived as hate speech depends on the social, societal and educational background of the reader and the specific context in which a statement is made, a coherent annotation of hate speech data by several annotators is very hard to achieve. One of the consequences is that supervised models tend to generalize poorly beyond the data they have been trained and tested on, which, in its turn, means that their usage, e.g., for monitoring social media, can be only limited. To address this challenge, we are targeting in our current work weakly-supervised and unsupervised approaches to hate speech detection. Our work on author profiling/identification started with developing generic models, which have been tested on a broad range of different datasets. Currently, we aim to take these models further to profile authors of abusive language in social media.

5. External Funding Acquisition Record since 2005[1]

Contracts and subcontracts with companies for targeted development tasks are not listed below.

Amount (in EUR)	Title of the project	Funding Agency & Contract #	Life time
117,847.12	ALLIES: AI-based framework for supporting micro (and small) HSPs on the report and removaL of onLIne tErroriSt content	EU Commission ISF-101080090	S: 01.02.2023 E: 31.01.2025
119,062.50	RE-SILENCE: Retune the Soundscape of future cities through art and science collaboration	EU Commission HE-101070278-RIA	S: 01.09.2022 E: 31.8.2025
306,718.75	XR4DRAMA: Extended Reality for Disaster Management and Media Planning	EU Commission H2020-952133-IA	S: 01.11.2020 E: 31.10.2023
541,875.00 (3,995,709.35)	WELCOME: Multiple Intelligent Conversation Agent Services for Reception, Management and Integration of Third Country Nationals in the EU	EU Commission H2020-870930-RIA	S: 01.02.2020 E: 30.4.2023
99.468,00	Corporate Knowledge Insights (CKI): The AI based solution for Enterprise Knowledge Management	Acció INNOTECRD19-1-0013	S: 16.10.2019 E: 16.10.2021
219,562.50	INGENIOUS: The First Responder (FR) of the Future: a Next Generation Integrated Toolkit (NGIT) for Collaborative Response, increasing protection and augmenting operational capacity	EU Commission H2020–833435–RIA	S: 01.09.2019 E: 28.02.2023
282,500.00	MindSpaces: Art-driven adaptive outdoors and indoors design	EU Commission H2020-825079-STARTS	S: 01.01.2019 E: 30.06.2022
30,000.00	Pla d'impuls del creixement del grup de Tractament Automàtic del Llenguatge Natural	Pla de Recerca i Innovació de Catalunya; TECDTP18-1-0013	S: 03.05.2018 E: 02.05.2019

---

[1]In the case of projects for which the amount acquired for the whole Consortium is also indicated in parentheses, I act/acted as Project / Scientific Coordinator and main author of the proposal.

473,750.00	CONNEXIONs: InterCONnected NEXt-Generation Immersive IoT Platform of Crime and Terrorism DetectiON, PredictiON, InvestigatiON, and PreventiON Services	EU Commission; H2020-786731-RIA	S: 01.09.2018 E: 28.02.2022
86.283,02	TecSemHu. Tecnologies semàntiques per a les Humanitats Digitals	RecerCaixa: Associació Catalana Univ Publiques – GUNI 2017 ACUP 00172	S: 01.04.2018 E: 31.03.2021
476,250.00	V4Design: Visual and textual content re-purposing FOR(4) architecture, Design and video virtual reality games	EU Commission; H2020-779962-RIA	S: 01.01.2018 E: 30.09.2021
36,230.60	Pla d'impuls del creixement del grup de Tractament Automàtic del Llenguatge Natural	Pla de Recerca i Innovació de Catalunya; TECDTP16-1-0015	S: 02.05.2017 E: 02.05.2018
570,000.00	beAWARE: Enhancing decision support and management services in extreme weather climate events	EU Commission; H2020-700475-IA	S: 01.01.2017 E: 31.12.2019
566,250.00	TENSOR: Retieval and Analysis of Heterogeneous Online Content for Terrorist Activity Recognition	EU Commission; H2020–7000024-RIA	S: 01.09.2016 E: 31.08.2019
41,400.00	Pla d'actuació en transferència tecnològica del grup de recerca TALN	Pla de Recerca i Innovació de Catalunya; TECDTP15-1-0003	S: 30.10.2015 E: 01.05.2017
852,500.00 (3,633,801.00)	KRISTINA: A Knowledge-Based Information Agent with Social Competence and Human Interaction Capabilities	EU Commission; H2020–645012–RIA	S:01.03.2015 E: 28.02.2018
24,000.00	Content Oriented Patent Search and Summarization	AGAUR/FCRI EDI-PILOT-2014-14	S: 01.03.2014 E: 28.02.2015
408,640.00	MULTISENSOR: Mining and Understanding of multilinguaL contenT for Intelligent Sentiment Enriched coNtext and Social Oriented inteRpretation	EU Commission; FP7-ICT-610411	S: 01.11.2013 E: 31.10.2016
444,800.00 (1,430,000.00)	iPatDoc: Workbench for Interactive Contrastive Analysis of Patent Documentation	EU Commission; FP7-SME-606163	S: 15.11.2013 E: 14.11.2015

73,810.00	HARENES: Herramienta de ayuda a la redacción en español: Tecnologías lingüísticas para el procesamiento de colocaciones	Spanish Ministry of Economy and Competitivity: FFI2011-30219-CO2-02	S: 1.1.2012 E: 30.09.2015
431,480.00 (999,307.00)	TOPAS: Tool Platform for Intelligent Patent Analysis and Summarization	EU Commission; FP7-SME-286639	S: 1.10.2011 E: 30.9.2013
481,422.00 (2,785,300.00)	PESCaDO: Personalized Environmental Service Configuration and Delivery Orchestration	EU Commission; FP7-ICT-248594	S: 1.1.2010 E: 31.12.2012
59,000.00	COLOCATE: Colocaciones y tecnologías lingüísticas: hacia un entorno de aprendizaje basado en la web	Spanish Ministry of Science and Innovation; MCI FFI2008-06479-CO2-02/ FILO	S: 1.1.2009 E: 31.12.2011
180,000.00	*i3Media*, NLG Subproject	Spanish Ministry of Industry	S: 1.1.2007 E: 31.12.2010
625,334.00 (2,500,000.00)	PATExpert: Advanced Patent Document Processing Techniques	EU Commission; FP6-IST-028116	S: 1.1.2006 E: 15.9.2008
266,772.00 (1,000,00.00)	MARQUIS: Multimodal Air Quality Information Service for General Public	EU Commission; EDC-11258	S: 1.1.2005 E: 31.12.2006

6. Main Ongoing (externally funded) Research Projects

1. ALLIES: AI-based framework for supporting micro (and small) HSPs on the report and removaL of onLIne tErroriSt content (01.02.2023–31.01.2025), funded by the European Commission (ISF-101080090)

2. RE-SILENCE: Retune the Soundscape of future cities through art and science collaboration (01.09.2022–31.8.2025), funded by the European Commission (HE-101070278-RIA)

I am also Principal Investigator of a number of smaller projects in which my group is (sub-)contracted by the Private Sector.

6. Completed Research Projects

6.1 Projects Funded by European Agencies

1. XR4DRAMA: Extended Reality for Disaster Management and Media Planning (01.11.202–31.10.2023), funded by the European Commission (H2020-952133-IA), Principal Investigator.

2. WELCOME: Multiple Intelligent Conversation Agent Services for Reception, Management and Integration of Third Country Nationals in the EU (01.02.2020–

30.04.2023), funded by the European Commission (H2020–870930–RIA), Coordinator.

3. INGENIOUS: The First Responder (FR) of the Future: a Next Generation Integrated Toolkit (NGIT) for Collaborative Response, increasing protection and augmenting operational capacity (01.09.2019–28.02.2023), funded by the EU Commission (H2020–833435–RIA), Principal Investigator.

4. MindSpaces: Art-driven adaptive outdoors and indoors design, (01.01.2019–30.06.2022), funded by the European Commission (H2020–825079–STARTS), Principal Investigator.

5. CONNEXIONs: InterCONnected NEXt-Generation Immersive IoT Platform of Crime and Terrorism DetectiON, PredictiON, InvestigatiON, and PreventiON Services, (01.09.2018–28.02.2022), funded by the European Commission (H2020–786731–RIA), Principal Investigator.

6. V4Design: Visual and textual content re-purposing FOR(4) architecture, Design and video virtual reality games, (01.01.2018–30.09.2021), funded by the European Commission (H2020–779962–RIA), Principal Investigator.

7. beAWARE: Enhancing decision support and management services in extreme weather climate events, H2020–700475–IA, (01.01.2017–31.12.2019), Principal Investigator.

8. TENSOR: Retieval and Analysis of Heterogeneous Online Content for Terrorist Activity Recognition, H2020–7000024–RIA, (01.07.2016–30.06.2019), Principal Investigator.

9. KRISTINA: A Knowledge-Based Information Agent with Social Competence and Human Interaction Capabilities, H2020–645012–RIA (01.03.2015–28.02.2018), Coordinator.

10. Multisensor: Mining and Understanding of multilinguaL contenT for Intelligent Sentiment Enriched coNtext and Social Oriented inteRpretation, FP7-ICT-610411 (01.11.2013–31.10.2016), Principal Investigator.

11. iPatDoc: Workbench for Interactive Contrastive Analysis of Patent Documentation, FP7-SME-606163 (15.11.2013–14.11.2015), Scientific Coordinator, Principal Investigator.

12. TOPAS: Tool Platform for Intelligent Patent Analysis and Summarization, FP7-SME-286639 (01.10.2011–30.09.2013), Principal Investigator.

13. PESCaDO: Personalized Environmental Service Configuration and Delivery Orchestration, FP7-ICT-248594 (01.01.2010–30.12.2012), Coordinator.

14. PATExpert: Advanced Patent Document Processing Techniques, FP6-ICT-028116 (1.2.2006–15.9.2008), Coordinator.

15. MARQUIS: Multimodal Air Quality Information Service for General Public,

EDC-11258 (1.1.2005–31.12.2006), Coordinator.

16. Enriching EuroWordNet by Collocations, German Academic Exchange Agency (DAAD) and the Spanish Ministry of Science and Technology (1.1.2003–31.12.2004), Head of the German subproject.

17. MULTIMETEO, FP4-LE (1.10.1996–31.3.1999), Consultant

15. GREG, EC-INTAS (1.3.1999–1.5.2001), Head of the German subproject.

6.2 Projects Funded by Spanish Agencies

1. HARENES: Herramienta de ayuda a la redacción en español: Tecnologías lingüísticas para el procesamiento de colocaciones, FFI2011-30219-CO2-02 (01.01.2012–30.09.2015), Subproject Leader.

2. COLOCATE: Colocaciones y tecnologías lingüísticas: hacia un entorno de aprendizaje basado en la web, MCI FFI 2008-06479-CO2-02/FILO (1.1.2009–31.12.2011), Subproject Leader.

3. i3Media, CENIT (1.1.2007–31.12.2010), Team Leader.

4. TEXTERM 2: Fundamentos, estrategias y herramientas para el procesamiento y extracción automáticos de información especializada, Spanish Ministry of Science and Technology (1.1.2003–31.12.2005), Affiliated Researcher.

5. Optimización de la indexación semántica por el medio de información collocacional, Xunta de Galicia (1.1.2003–31.12.2005), Consultant

6.3 Projects Funded by German Agencies

1. AutoText UIS, Ministry of Environment and Traffic, Baden-Württemberg (1.1.1999–31.10.2001), Head of the Project

2. Text Corpora and Tools for their Processing, Ministry of Research and Education, Baden-Württemberg (1.5.1993–31.12.1994), Research Fellow

3. KOMET, German Federal Ministry of Research and Technology (1.4.1988–31.3.1993), Research Fellow

6.4 Projects Funded by Canadian Agencies

1. HealthDoc, Canadian Science Foundation (1.3.1995–30.9.1996), Senior Researcher

7. Publications

PhD Thesis

L. Wanner. *Exploring Lexical Resources for Text Generation in a Systemic*

Functional Language Model. PhD Dissertation. Universität des Saarlandes, 1997.

Books

Editing of volumes

1. N. Calzolari, C.-R. Huang, H. Kim, J. Pustejovsky, **L. Wanner**, . . . (eds.), *Proceedings of the 29th International Conference on Computational Linguistics*. https://aclanthology.org/2022.coling-1.0.pdf, 2022.

2. E. André, T. Bickmore, S. Vrochidis, and **L. Wanner** (eds.) *Proceedings of the Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications*. CEUR-WS Volume 2338, http://ceur-ws.org/Vol-2338/, 2018.

3. K. Gerdes, E. Hajičová, and **L. Wanner** (eds.) *Dependency Linguistics: Advances in Linguistic Theory Using Dependency Structures*, Benjamins Academic Publishers, Amsterdam, 2014.

4. K. Gerdes, E. Hajičová, and **L. Wanner** (eds.) *Computational Dependency Theory*, IOS Press, Amsterdam, 2013.

5. Y.D. Apresjan, I.M. Boguslavskij, **L. Wanner**, L.L. Iomdin, J. Milićević, M.-C. L'Homme, and A. Polguère (eds). *Meanings, Texts and Other Exciting Things. Festschrift in Honour of Igor Mel'čuk*, Škola "Jazyki Russkoj Kul'tury", Moscow, 2012.

6. K. Gerdes, T. Reuther, and **L. Wanner** (eds.) 'MTT 2007', Otto Sagner, Munich, 2007.

7. **L. Wanner** (ed.) 'Selected Lexical and Grammatical Topics in the Meaning-Text Theory. In honour of Igor Mel'čuk', Benjamins Academic Publishers, Amsterdam, 2007.

8. **L. Wanner** (ed.). 'Recent Trends in Meaning-Text Theory', Benjamins Academic Publishers, Amsterdam, 1997.

9. **L. Wanner** (ed.). 'Lexical Functions in Lexicography and Natural Language Processing'. Benjamins Academic Publishers, Amsterdam, 1996.

10. K. Haenelt and **L. Wanner** (eds.). 'Proceedings of the International Workshop on the Meaning-Text Theory'. Arbeitspapiere der GMD 671, Darmstadt, July, 1992.

Book Chapters

1. P. Fortuna, J. Soler-Companys, **L. Wanner**. 'Dataset annotation in abusive language detection'. In C. Strippel, S. Paasch-Colberg, M. Emmer & J. Trebbe (Eds.), *Challenges and perspectives of hate speech analysis*. Digital Communication Research. https://doi.org/10.48541/dcr.v12.25., 369-391, 2023.

2. **L. Wanner**. 'From lists of lexical function instances to a multilingual collocation resource'. In T. Reuther, L. Iomdin, J. Milicević, and A. Polguère. *Lifetime Linguistic Inspirations. To Igor Mel'čuk from Colleagues and Friends for his 90th Birthday*. Peter Lang. 427–442, 2022.

3. F. Huber, Y.A. Samad, T. Spathi, K.L. Hough, A. Caforio, C. Munteanu, M. Karimi, A. Eckhardt Rodriguez, E. Kantor, C. Fiadzo, **L. Wanner**, M. Moudatsou, and E. Dimopoulou. 'Shaping ICT Policies for Integrative and Inclusive Digital Services'. In Akhgar, B., Hough, K.L., Abdel Samad, Y., Saskia Bayerl, P., Karakostas, A. (eds) *Information and Communications Technology in Support of Migration. Security Informatics and Law Enforcement*, Springer, Cham, 367–380, 2022.

4. J. Soler Company and **L. Wanner**. 'Authorship Attribution using Syntactic Dependencies'. In À. Nebot, X. Binefa, R. López de Mántaras (eds.). *Artificial Intelligence Research and Development*, IOS Press, Series *Frontiers in Artificial Intelligence and Applications*, vol. 288, Amsderdam, 303–308, 2016.

5. S. Mille, M. Ballesteros, A. Burga, G. Casamayor, and **L. Wanner**. 'Towards Multilingual Natural Language Generation within Abstractive Summarization'. In À. Nebot, X. Binefa, R. López de Mántaras (eds.). *Artificial Intelligence Research and Development*, IOS Press, Series *Frontiers in Artificial Intelligence and Applications*, vol. 288, Amsderdam, 309 –314, 2016.

6. B. Bohnet, S. Mille, and **L. Wanner**. 'One Step further towards Stochastic Semantic Sentence Generation'. In K. Gerdes, E. Hajičová, and L. Wanner (eds.) *Computational Dependency Theory*, IOS Press, Amsterdam, 93–112, 2013.

7. S. Mille, A. Burga, and **L. Wanner**. 'Looking Behind the Scenes of Syntactic Dependency Corpus Annotation: Towards a Motivated Annotation Scheme of Surface Syntax in Spanish'. In K. Gerdes, E. Hajičová, and L. Wanner (eds.) *Computational Dependency Theory*, IOS Press, Amsterdam, 26–46, 2013.

8. **L. Wanner**, M. Alonso Ramos, O. Vincze, R. Nazar, G. Ferraro, E. Mosqueira, and S. Prieto. 'Annotation of collocations in a learner corpus for building a learning environment'. In S. Granger, F. Meunier, and G. Gilquin (eds.) *Twenty Years of Learner Corpus Research*, Louvain University Press, Louvain-la-Neuve, 2013.

9. **L. Wanner**. 'Electronic Dictionaries for Natural Language Generation'. In R. Gouws R., U. Heid, W. Schweickard, and E. Wiegand (eds.) *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography, Chapt. 80*. De Gruyter Mouton, Berlin, 1150–1160, 2013.

10. **L. Wanner**. 'Meaning to Text and Meaning to Meaning: Transduction and Transformation in the Meaning-Text Theory'. In Apresjan et al. (eds). *Meanings, Texts and Other Exciting Things: Festschrift in Honour of Igor Mel'čuk*, Škola "Jazyki Russkoj Kul'tury", Moscow, 2012.

11. N. Bouayad-Agha, G. Casamayor, **L. Wanner**, F. Díez, and S. López Hernández. 'FootbOWL: Using a Generic Ontology of Football Competition for Planning Match Summaries'. In G. Antoniu, M. Grobelnik, E. Simperl, B. Parsia, D. Plexousakis, P. De Leenheer and J. Pan (eds.) The Semantic Web: Research and Applications. Proceedings of the Extended Semantic Web Conference. Springer Verlag, Heidelberg, 230–244, 2011.

12. **L. Wanner**. 'Report Generation'. In N. Indurkhya and F. Damerau (eds.) 'Handbook of Natural Language Processing, second edition'. CRC Press, Taylor and Francis, London, 533–556, 2010.

13. **L. Wanner** and F. Lareau. 'Applying MTT-Models to Text Synthesis with Low- and Middle-Density Languages in Mind', In S. Nirenburg (ed.) 'Language Engineering for Lesser-Studied Languages'. IOS Press, Volume 21, NATO Science for Peace and Security Series - D: Information and Communication Security, Amsterdam, 207–242, 2009.

14. **L. Wanner** and B. Bohnet. 'Automatic Recognition of Lexical Function Instances'. In Leo Wanner (ed.). 'Selected Lexical and Grammatical Topics in the Meaning-Text Theory. In honour of Igor Mel'cuk', Benjamins Academic Publishers, Amsterdam, 231–261, 2007.

15. **L. Wanner** and M. Alonso Ramos. 'Collocation Chains: How to Deal with Them?'. In K. Gerdes, T. Reuther, and L. Wanner (eds.) 'MTT 2007'. Sagner, Munich, 5-14, 2007.

16. **L. Wanner**, B. Bohnet, M. Giereth and V. Vidal. 'The first steps towards the automatic compilation of specialized collocation dictionaries'. In F. Ibekwe-SanJuan, A. Condamines, and M.T. Cabré Castellví. Application-Driven Terminology Engineering. Benjamins Academic Publishers, Amsterdam, 143-180, 2007.

17. **L. Wanner**. 'El corpus como un diccionario de colocaciones?'. In M. Alonso Ramos (ed). *Fraseología y Colocaciones*. University of La Coruña Press. 2006.

18. **L. Wanner**. 'Introduction' (Overview of restricted lexical co-occurrence). In L. Wanner (ed.) Lexical Functions in Lexicography and Natural Language Processing. Benjamins Academic Publishers, Amsterdam, 1–36, 1996.

19. I. Mel'čuk and **L. Wanner**. 'Lexical Functions and Lexical Inheritance for Emotion Lexemes in German'. In L. Wanner (ed.) Lexical Functions in Lexicography and Natural Language Processing. Benjamins Academic Publishers, Amsterdam, 209–278, 1996.

Journal Articles

1. V.R. Xefteris, M. Dominguez, J. Grivolla, A. Tsanousa, F. Zaffanela, M. Monego, S. Symeonidis, S. Diplaris, **L. Wanner**, S. Vrochidis & I. Kompatsiaris, 'A Multimodal Late Fusion Framework for Physiological Sensor and Audio-Signal-Based Stress Detection: An Experimental Study and Public Dataset', *Electronics*, 12 - 23, 2023.

2. S- Symeonidis, ..., **L. Wanner** et al. 'V4Design: Intelligent Analysis and Integration of Multimedia Content for Creative Industries', *IEEE Systems Journal*, 17, 2, 2570-2573, 2023.

3. A. Shvets and **L. Wanner**. 'The Relation Dimension in the Identification and Classification of Lexically Restricted Word Co-Occurrences in Text Corpora'. *Mathematics*, 10(20):3831, 2022. doi: 10.3390/math10203831.

4. M. Domínguez Bajo, M. Farrús, and **L. Wanner**. 'The Information Structure Prosody Interface in Text-to-Speech Technologies. An Empirical Perspective'. *Corpus Linguistics and Linguistic Theory*, 18 (2):419-445, 2022.

5. S. Symeonidis, ...**L. Wanner**, M. Marimón, K. Valsamidou, P. Koulali, A. Tellios, J. Wuyts, and E. Lopez.'V4Design: Intelligent Analysis and Integration of Multimedia Content for Creative Industries'. *IEEE Systems Journal*, 2022. doi: 10.1109/JSYST.2022.3217655.

6. P. Fortuna, J. Soler-Company, and **L. Wanner**. 'How Well Do Hate Speech, Toxicity, Abusive and Offensive Language Classification Models Generalize Across Datasets?'. *Information Processing & Management*, 58(3), 2021.

7. S. Vrochidis, A. Moumtzidou, I. Gialampoukidis, D. Liparas, G. Casamayor, **L. Wanner**, N. Heise, T. Wagner, A. Bilous, E. Jamin, B. Simeonov, V. Alexiev, R. Busch, I. Arapakis, Y. Kompatsiaris. 'A multimodal analytics platform for journalists analysing large-scale, heterogeneous multilingual and multimedia content'. *Frontiers in Robotics and AI*, 5., 2018. https://doi.org/10.3389/frobt.2018.00123.

8. S. Dasiopoulou, G. Meditskos, **L. Wanner**, S. Vrochidis, and P. Cimiano (Eds). 'Special Issue on Expressive Natural Language Interfaces for the Semantic Web'. *Semantic Web*, 10(3), 2019.

9. J. Soler-Company and **L. Wanner**. 'On the Role of Syntactic Dependencies and Discourse Relations for Author and Gender Identification'. *Pattern Recognition Letters*, 105:87–95, 2018. https://doi.org/10.1016/j.patrec.2017.12.006

10. **L. Wanner**. 'Processament automàtic de patents: un exercici de terminologia computacional'. *Terminalia*, 16:54–56, 2017.

11. Q. Han, F. Heimerl, J. Codina-Filbà, S. Lohmann, **L. Wanner**, and T. Ertl. 'Visual Patent Trend Analysis for Informed Decision Making in Technology Management', *World Patent Information Journal*, 49:34–42, 2017.

12. **L. Wanner**, G. Ferraro, and P. Moreno. 'Towards Distributional Semantics-Based Classification of Collocations for Collocation Dictionaries'. *International Journal of Lexicography*, 30(2):167–186, 2017; published online in March 2016. doi:10.1093/ijl/ecw002.

13. J. Codina-Filbà, N. Bouayad-Agha, A, Burga, G. Casamayor, S. Mille, A. Müller, H. Saggion, and **L. Wanner**. 'Using Genre-Specific Features for Patent Summaries', *Information Processing and Management*, 53(1):151–174, 2017; published online, July 2016. doi: 10.1016/j.ipm.2016.07.002.

14. M. Ballesteros, B. Bohnet, S. Mille, and **L. Wanner**. 'Data-Driven Deep-Syntactic Dependency Parsing'. *Natural Language Engineering Journal*, 22(6):939–974, 2016; published online in August 2015.
doi: 10.1017/S1351324915000285.

15. V. Epitropou, T. Bassoukos, K. Karatzas, A. Karppinen, **L. Wanner**, S. Vrochidis, I. Kompatsiaris, J. Kukkonen. 'Environmental data extraction from heatmaps using the AirMerge system'. *Multimedia Tools and Applications*, 75(3):1589–1613, 2016; published online in May 2015. doi: 10.1007/s11042-015-2604-7.

16. **L. Wanner**, M. Rospocher, S. Vrochidis, L. Johansson, N. Bouayad-Agha, G. Casamayor, A. Karppinen, I. Kompatsiaris, S. Mille, A. Moumtzidou, L. Serafini. 'Ontology-centered environmental information delivery for personalized decision support'. *Expert Systems with Applications*, 42:5032–5046, 2015. doi: 10.1016/j.eswa.2015.02.048.

17. L. Johansson, V. Epitropou, K. Karatzas, A. Karppinen, **L. Wanner**, S. Vrochidis, A. Bassoukos, J. Kukkonen, I. Kompatsiaris. 'Fusion of meteorological and air quality data extracted from the web for personalized environmental information services'. *Environmental Modelling and Software*, 64:143–155, 2015.

18. S. Brügmann, N. Bouayad-Agha, A. Burga, S. Carrascosa, A. Ciaramella, M. Ciaramella, J. Codina-Filba, E. Escorsa, A. Judea, S. Mille, A. Müller, H. Saggion, P. Ziering, H. Schütze, and **L. Wanner**. 'Towards Content-Oriented Patent Document Processing: Intelligent Patent Analysis and Summarization'. *World Patent Information Journal*, 40:30–42, 2015. doi: 10.1016/j.wpi.2014.10.003.

19. **L. Wanner** et al. 'Getting the environmental information across: From the web to the user'. *Expert Systems*, doi: 10.1111/exsy.12100, 32(3):405–432, 2015.

20. G. Ferraro, R. Nazar, M. Alonso Ramos, and **L. Wanner**. 'Towards Advanced Collocation Error Correction in Spanish Learner Corpora'. *Language Resources and Evaluation*, 48(1):45–64, 2014.

21. N. Bouayad-Agha, G. Casamayor, and **L. Wanner**. 'Natural Language Generation in the context of the Semantic Web'. *Semantic Web Journal*, 5(6):493–513, 2014; available online since 07.08.2013 http://www.semantic-web-journal.net/system/files/swj511_0.pdf.

22. N. Bouayad-Agha, G. Casamayor, S. Mille, and **L. Wanner**. 'Perspective-Oriented Generation of Football Match Summaries: Old Tasks, New Challenges'. *ACM Transactions on Speech and Language Processing*, 9(2), 2012.

23. R. Nazar, J. Vivaldi, and **L. Wanner**. 'Automatic Taxonomy Extraction for Specialized Domains Using Distributional Semantics. *Terminology*, 18(2):188–225, 2012.

24. S. Mille, **L. Wanner**, and A. Burga. 'Treebank Annotation in the Light of the Meaning-Text Theory'. *Linguistic Issues in Language Technology*, 7(16), 2012.

25. G. Ferraro and **L. Wanner**. 'Towards the derivation of verbal content relations from patent claims using deep syntactic structures'. *Knowledge-Based Systems*, 24(8):1233–1244, 2011.

26. **L. Wanner**, B. Bohnet, N. Bouayad-Agha, F. Lareau, and D. Nicklaß. 'MARQUIS: Generation of User-Tailored Multilingual Air Quality Bulletins', *Applied Artificial Intelligence*, 24(10), 2010.

27. M. Alonso Ramos and and **L. Wanner**. 'Actants et structure communicative', *Langages*, 172(4):9–31, 2009.

28. I. Mel'čuk and **L. Wanner**. 'Morphological Mismatches in Machine Translation', *Machine Translation*, 22(3):101–152, 2008.

29. **L. Wanner** et al. 'Towards Content-Oriented Patent Document Processing'. In *World Patent Information*, 30(1), 21–33, 2008.

30. I. da Cunha, **L. Wanner**, and T. Cabré. 'Summarization of Specialized Discourse. The case of medical articles in Spanish'. *Terminology*, 13(2), 249–286, 2007.

31. A. Lohmeyer, I. Düring, M. Giereth, T. Hoffmann, D. Nicklaß, M. Klein, H. Scheu-Hachtel, C. Sörgel and **L. Wanner**. 'Kurzfrist-Feinstaub-Immissionsprognose mit den Systemen MARQUIS und ProFet'. *Gefahrstoffe. Reinhaltung der Luft*, 67(7/8) 319–326, 2007.

32. I. Mel'čuk and **L. Wanner**. 'Syntactic Mismatches in Machine Translation'. *Machine Translation*, 20(2):81-138, 2006.

33. **L. Wanner**, B. Bohnet and M. Giereth. 'Making sense of collocations'. *Computer Speech & Language*, 20(4):609–624, 2006.

34. **L. Wanner**, B. Bohnet, M. Giereth and V. Vidal. 'The first steps towards the automatic compilation of specialized collocation dictionaries'. *Terminology*, 11(1):137–174, 2005.

35. **L. Wanner**. 'Towards Automatic Fine-Grained Semantic Classification of Verb-Noun Collocations'. *Natural Language Engineering Journal*, 10(2):95–143, 2004.

36. I. Mel'čuk and **L. Wanner**. 'Towards a Lexicographic Approach to Lexical Transfer in Machine Translation (Illustrated by the German-Russian Language Pair). *Machine Translation*, 16(1):21-87, 2001.

37. **L. Wanner**. 'On the Representation of Collocations in a Multilingual Computational Lexicon'. In *t.a.l.*, 40:1, pp. 55–86, 1999.

38. **L. Wanner**. 'Lexical Choice in Text Generation and Machine Translation'. In *Machine Translation*, 19(1–3), 1996.

39. I. Mel'čuk and **L. Wanner**. 'Lexical Co-occurrence and Lexical Inheritance: A Case Study of Emotion Lexemes in German'. *Lexikos*, (4):85–161. Capetown, 1994.

40. K. Haenelt, M. Könyves-Tóth, E. Maier, D. Mörscher-Krämer, H. Müller, G. Schulz and **L. Wanner**. 'Multimediale Wissensverarbeitung in integrierten Publikations- und Informationssystemen. *KI*, 3(2):33–35, 1989.

Peer-reviewed Conferences and Workshops

2023

1. G. Cámbara, J. Grivolla, M. Farrús & **L. Wanner**. 'Automatic Speech Translation for Multinational First Responder Teams', In *Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management*, 188-196, Omaha, NE, 2023.

2. Y. Jin, **L. Wanner**, V.L. Kadam & A. Shvets. 'Towards Weakly-Supervised Hate Speech Classification Across Datasets', In *Proceedings of the 7th Workshop on Online Abuse and Harms* (WOAH), 42–59, Toronto, ON, 2023.

2022

3. P. Fortuna, M. Dimínguez, **L. Wanner**, and Z. Talat. 'Directions for NLP Practices Applied to Online Hate Speech Detection'. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Abu Dhabi, 2022.

4. L. Espinosa-Anke, A. Shvets, A. Mohammadshahi, J. Henderson, and **L. Wanner**. 'Multilingual Extraction and Categorization of Lexical Collocations with Graph-aware Transformers' In *Proceedings of the 11th Joint Conference on Lexical and Computational Semantics*, Seattle, 2022.

5. S. Mille, G. Casamayor, J. Grivolla, A. Shvets, and **L. Wanner**. 'Automatic Multilingual Incident Report Generation for Crisis Management'. In *Proceedings of the 19th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, Tarbes, France, 2022.

6. T. Pistola, N. Georgakopoulou, A. Shvets, K. Chatzistavros, V.-R. Xefteris, A. Táboas Garcia, I. Koulalis, S. Diplaris, **L. Wanner**, S. Vrochidis, I. Kompatsiaris. 'Imageability-Based Multi-modal Analysis of Urban Environments for Architects and Artists'. In *Proceedings of International Conference on Image Analysis and Processing*, pp. 198-209, Springer, Cham., 2022.

7. E.A. Stathopoulos, A. Shvets, R. Carlini, S. Diplaris, S. Vrochidis, **L. Wanner**, I. Kompatsiaris. 'Social Media and Web Sensing on Interior and Urban Design'. In *Proceedings of the 27th IEEE Symposium on Computers and Communications* (ISCC 2022), Rhodes, Greece. 2022.

2021

8. L. Pérez-Mayos, M. Ballesteros and **L. Wanner**. 'How much pretraining data do language models need to learn syntax?' In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Punta Cana, Dominican Republic, 2021.

9. **L. Wanner** et al. 'Towards a Versatile Intelligent Conversational Agent as Personal Assistant'. In *Proceedings of the 19th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS)*, Salamanca, Spain, 2021.

10. M. Domínguez Bajo, J. Soler-Company, and **L. Wanner**. 'ThemePro 2.0: Showcasing the Role of Thematic Progression in Engaging Human-Computer Interaction'. In *Proceedings of Interspeech 2021, Show and Tell Track*, Brno, Czech Republic, 2021.

11. A. Shvets, P. Fortuna, J. Soler-Company, and **L. Wanner**. 'Targets and Aspects in Social Media Hate Speech'. In *Proceedings of the 5th WS on Online Abuse and Harms*, collocated with ACL-IJCNLP, online, 2021.

12. P. Fortuna, L. Pérez-Mayos, A. AbuRa'ed, J. Soler-Company, and **L. Wanner**. 'Cartography of Natural Language Processing for Social Good: Definitions, Statistics and White Spots'. In *Proceedings of the First WS on NLP for Positive Impact*, collocated with ACL-IJCNLP, online, 2021.

13. L. Pérez-Mayos, A. Táboas García, S. Mille, and **L. Wanner**. 'Assessing the Syntactic Capabilities of Transformer-based Multilingual Language Models'. In *Findings of ACL*, 2021.

14. L. Espinosa Anke, J. Codina-Filbà and **L. Wanner**. 'Evaluating language models for the retrieval and categorization of lexical collocations'. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Kyiv, Ukraine, 2021.

15. L. Pérez-Mayos, R. Carlini, M. Ballesteros and **L. Wanner**. 'On the evolution of syntactic information encoded by BERT's contextualized representations'. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Kyiv, Ukraine, 2021.

2020

16. A. Shvets and **L. Wanner**. 'Concept Extraction Using Pointer-Generator Networks'. In *Proceedings of the 22nd International Conference on Knowledge Engineering and Knowledge Management*, Bolzano, Italy, 2020.

17. S. Mille, S. Symeonidis, M. Rousi, M. Marimon, K. Stavrothanasopoulos, P. Alvanitopoulos, R. Carlini, J. Grivolla, G. Meditskos, S. Vrochidis and **L. Wanner**. 'A Case Study of NLG from Multimedia Data Sources: Generating Architectural Landmark Descriptions'. In *Proceedings of the WebNLG Workshop*, collocated with the International Conference on Natural Language Generation (INLG), Dublin, Ireland, 2020.

18. S. Mille, A. Belz, B. Bohnet, T. Castro Ferreira, Y. Graham, **L. Wanner**. 'The Third Multilingual Surface Realisation Shared Task: Overview and Evaluation Results'. In *Proceedings of the Third Workshop on Multilingual Surface Realisation*, collocated with COLING 2020, Barcelona, Spain, 2020.

19. B. Fisas, L. Espinosa-Anke, J. Codina-Filbà, **L. Wanner**. 'CollFrEn: Rich Bilingual English–French Collocation Resource'. In *Proceedings of the Workshop on Multiword Expressions and Electronic Lexicons*, collocated with COLING 2020, Barcelona, Spain, 2020.

20. D. Chatzakou, J. Soler-Company, T. Tsikrika, **L. Wanner**, S. Vrochidis, and I. Kompatsiaris. 'User Identity Linkage in Social Media Using Linguistic and Social Interaction Features'. In *Proceedings of the 12th ACM Web Science Conference*, Southampton, UK, 2020.

21. M. Domínguez, J. Soler, and **L. Wanner**. 'ThemePro: A Toolkit for the Analysis of Thematic Progression'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation* (LREC), Marseille, France, 2020.

22. P. Fortuna, J. Soler, and **L. Wanner**. Toxic, Hateful, Offensive or Abusive? What Are We Really Classifying? An Empirical Analysis of Hate Speech Datasets. In *Proceedings of the International Conference on Linguistic Resources and Evaluation* (LREC), Marseille, France, 2020.

2019

23. S. Mille, A. Belz, B. Bohnet, Y. Graham, and **L. Wanner**. 'The Second Multilingual Surface Realisation Shared Task (SR'19): Overview and Evaluation Results'. In *Proceedings of the Second Workshop on Multilingual Surface Realisation at EMNLP*, pp. 1–17, Hong Kong, China, 2019.

24. S. Mille, B. Fisas, S. Dasiopoulou, and **L. Wanner**. Teaching FORGe to Verbalize DBpedia Properties in Spanish. In *Proceedings of the 12th International Conference on Natural Language Generation*. Tokyo, Japan, 2019.

25. L. Espinosa Anke, S. Schockaert and **L. Wanner**. Collocation Classification with Unsupervised Relation Vectors. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*, Short papers. Florence, Italy, 2019.

26. P. Fortuna, J. Rocha da Silva, J. Soler-Company, **L. Wanner** and S. Nunes. A Hierarchically-Labeled Portuguese Hate Speech Dataset. In *Proceedings of the 3rd WS on Abusive Language Online*, collocated with ACL. Florence, Italy, 2019.

27. J. Soler-Company and **Leo Wanner**. Automatic Classification and Linguistic Analysis of Extremist Online Material. In *Proceedings of the 25th International Conference on Multimedia Modeling (MMM)*, pp. 577 – 582. Thessaloniki, Greece, 2019.

28. S. Mille, S. Dasiopoulou, and **Leo Wanner**. A portable grammar-based NLG system for verbalization of structured data. In *Proceedings of ACM SAC 2019, Special Track on Knowledge and Language Processing*, Limassol, Cyprus, 2019.

29. P. Alvanitopoulos, . . . , **L. Wanner**, T. Klein, S. Vrochidis, I. Kompatsiaris. 'MindSpaces: Art-driven Adaptive Outdoors and Indoors Design'. In *Proceedings of the International Conference on Digital Presentation and Preservation of Cultural and Scientific Heritage*, 9, pp. 391–402, 2019.

2018

30. G. Eschweiler and **L. Wanner**. How Can Intelligent Conversational Agents Help? The Needs of Geriatric Patients and Their Caregivers. In *Proceedings of the Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications*. CEUR-WS Volume 2338, pp. 1–9. http://ceur-ws.org/Vol-2338/.

31. M. Domínguez Bajo, A. Burga, M. Farrús, and **L. Wanner**. On the Role of Communicative Structure in Read Aloud. Applications for the Elderly. In *Proceedings of the Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications*. CEUR-WS Volume 2338, pp. 40–44. http://ceur-ws.org/Vol-2338/.

32. M. Domínguez Bajo, A. Burga, M. Farrús, and **L. Wanner**. Towards expressive prosody generation in TTS for reading aloud applications. In *Proceedings of the IberSpeech Conference*. Barcelona, Spain, 2018.

33. S. Mille, A. Belz, B. Bohnet and **L. Wanner**. 'Underspecified Universal Dependency Structures as Inputs for Multilingual Surface Realisation'. In *Proceedings of the 11th International Conference on Natural Language Generation*. Tilburg, The Netherlands, 2018.

34. A. Shvets, S. Mille and **L. Wanner**. 'Sentence Packaging in Text Generation from Semantic Graphs as a Community Detection Problem'. In *Proceedings of the 11th International Conference on Natural Language Generation*. Tilburg, The Netherlands, 2018.

35. S. Mille, A. Belz, B. Bohnet, Y. Graham, E. Pitler and **L. Wanner**. 'The First Multilingual Surface Realisation Shared Task (SR'18): Overview and Evaluation Results'. In *Proceedings of the First Workshop on Multilingual Surface Realisation at ACL*, pp. 1–12, Melbourne, Australia, 2018.

36. J. Soler-Company and **L. Wanner**. 'Automatic Identification of Texts Written by Authors with Alzheimer's Disease'. In *Proceedings of the 40th Annual Cognitive Science Society Meeting*, Madison, WI, 2018.

37. M. Domínguez Bajo, M. Farrús, and **L. Wanner**. 'Thematicity-based Prosody Enrichment for Text-to-Speech Applications'. In *Proceedings of the 9th Speech Prosody Conference*, Poznań, Poland, 2018.

38. A. Burga, M. Domínguez, M. Farrús, and **L. Wanner**. 'Compilation of Corpora for the Study of the Information Structure–Prosody Interface'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation* (LREC), Miyazaki, Japan, 2018.

39. S. Rodríguez Fernández, R. Carlini, and **L. Wanner**. 'Generation of a Spanish Artificial Collocation Error Corpus'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation* (LREC), Miyazaki, Japan, 2018.

40. L. Pérez Mayos, F. Sukno, and **L. Wanner**. 'Improving the Quality of Video-to-Language Models by Optimizing Annotation of the Training'. In *Proceedings of the 24th International Conference on Multimedia Modeling (MMM)*, Bangkok, Thailand, 2018.

2017

41. A. Öktem, M. Farrús and **L. Wanner**. 'Attentional Parallel RNNs for Generating Punctuation in Transcribed Speech'. In *Proceedings of the 5th International Conference on Statistical Language and Speech Processing (SLSP 2017)*, Le Mans, France, 2017.

42. **L. Wanner** et al. 'Design of a Knowledge-Based Agent as a Social Companion'. In *Proceedings of the International Conference on Health and Social Care Information Systems and Technologies (HCIct)*, Barcelona, Spain, 2017.

43. S. Mille, B. Bohnet, **L. Wanner**, and A. Belz. 'Shared Task: Multilingual Surface Realization Using Universal Dependency Trees'. In Proceedings of the *International Conference on Natural Language Generation (INLG)*, Santiago de Compostela, Spain, 2017.

44. S. Mille and **L. Wanner**. 'A demo of FORGe: the Pompeu Fabra Open Rule-based Generator'. In Proceedings of the *International Conference on Natural Language Generation (INLG)*, Santiago de Compostela, Spain, 2017.

45. S. Mille, R. Carlini, I. Latorre, and **L. Wanner**. 'UPF at EPE 2017: Transduction-based Deep Analysis'. In *Proceedings of the 2017 Shared Task on Extrinsic Parser Evaluation at the Fourth International Conference on Dependency Linguistics and the 15th International Conference on Parsing Technologies*. Pisa, Italy, 2017.

46. C. Ten Ventura, R. Carlini, S. Dasioupoulou, G.Llorach Tó, and **L. Wanner**. 'Towards Reasoned Modality Selection in an Embodied Conversation Agent'. In *Proceedings of the Intelligent Virtual Agents (IVA) Conference*, Stockholm, Sweden, 2017.

47. J. Soler Company and **L. Wanner**. 'On the Relevance of Syntactic and Discourse Features for Author Profiling and Identification'. In *Proceedings of the Meeting of the European Chapter for Computational Linguistics (EACL), Short paper track*, Valencia, Spain, 2017.

48. **L. Wanner** et al. 'KRISTINA: A Knowledge-Based Virtual Conversation Agent'. In *Proceedings of the 15th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS)*, Porto, Portugal, 2017.

49. J. Kleinhans, M. Farrús, A. Gravano, J.M. Pérez, C. Lai, and **L. Wanner**. 'Using Prosody to Classify Discourse Relations'. In *Proceedings of Interspeech 2017*, Stockholm, Sweden, 2017.

50. M. Domínguez Bajo, M. Farrús and **L. Wanner**. 'A Thematicity-based Prosody Enrichment Tool for CTS'. In *Proceedings of Interspeech 2017, Show and Tell Track*, Stockholm, Sweden, 2017.

51. A. Öktem, M. Farrús, and **L. Wanner**. 'ProsodyScore: A Tool for Prosody Visualisation of Large Speech Corpora'. In *Proceedings of Interspeech 2017, Show and Tell Track*, Stockholm, Sweden, 2017.

52. S. Mille, A. Burga, R. Carlini, and **L. Wanner**. 'FORGe at SemEval-2017 Task 9: Deep sentence generation based on a sequence of graph transducers'. In *Proceedings of SemEval*, Vancouver, Canada 2017.

53. A. Burga, A. Öktem and **L. Wanner**. 'Revising the METU-Sabancı Turkish treebank: An Exercise in Surface-Syntactic Annotation of Agglutinative Languages'. In *Proceedings of the Fourth International Conference on Dependency Linguistics (DepLing)*, Pisa, Italy, 2017.

54. A. Öktem, M. Farrús, and **L. Wanner**. 'Automatic Extraction of Parallel Speech Corpora from Dubbed Movies'. In *Proceedings of 10th BUCC, Workshop on Building and Using Comparable Corpora*, collocated with ACL, Vancouver, Canada, 2017.

2016

55. L. Espinosa Anke, J. Camacho-Collados, S. Rodríguez-Fernández, H. Saggion and **L. Wanner**. 'Extending WordNet with Fine-Grained Collocational Information'. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, Osaka, Japan, 2016.

56. M. Domínguez, M. Farrús, and **L. Wanner**. 'An Automatic Prosody Tagger for Spontaneous Speech'. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, Osaka, Japan, 2016.

57. M. Domínguez, I. Latorre, M. Farrús, J. Codina-Filba, and **L. Wanner**. 'Praat on the Web: An Upgrade of Praat for Semi-Automatic Speech Annotation'. In *Proceedings of the International Conference on Computational Linguistics (COLING), Demo track*, Osaka, Japan, 2016.

58. M. Ballesteros and **L. Wanner**. 'A Neural Network Architecture for Multilingual Punctuation Generation'. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Short paper track*, Austin, TX, 2016.

59. G. Meditskos, S. Dasiopoulou, S. Vrochidis, **L. Wanner**, and Y. Kompatsiaris. 'Question Answering over Pattern-Based User Models'. In *Proceedings of the Semantics Conference*, Leipzig, Germany, 2016.

60. S. Rodríguez Fernández, L. Espinosa Anke, R. Carlini, and **L. Wanner**. 'Semantics-Driven Recognition of Collocations Using Word Embeddings'. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), Short paper track*, Berlin, Germany, 2016.

61. J. Soler Company and **L. Wanner**. 'A Semi-Supervised Approach to Gender Identification'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation* (LREC), Portorož, Slovenia, 2016.

62. S. Rodríguez Fernández, R. Carlini, L. Espinosa Anke, and **L. Wanner**. 'Example-based Acquisition of Fine-grained Collocation Resources'. *Proceedings of the International Conference on Linguistic Resources and Evaluation* (LREC), Portorož, Slovenia, 2016

63. A. Burga, S. Cajal, J. Codina-Filbà, and **L. Wanner**. 'Towards Multiple Antecedent Coreference Resolution in Specialized Discourse'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation* (LREC), Portorož, Slovenia, 2016

64. M. Domínguez, A. Burga, M. Farrús, and **L. Wanner**. 'Using hierarchical information structure for prosody prediction in content-to-speech applications'. In *Proceedings of the 8th Speech Prosody Conference*, Boston, MA, 2016.

65. M. Domínguez, M. Farrús, and **L. Wanner**. 'Combining acoustic and linguistic features in phrase-oriented prosody prediction'. In *Proceedings of the 8th Speech Prosody Conference*, Boston, MA, 2016.

66. S. Rodríguez-Fernández, L. Espinosa-Anke, R. Carlini, and **L. Wanner**. 'Semantics-Driven Collocation Discovery'. In *Proceedings of the Spanish Conference on Computational Linguistics*, Salamanca, Spain, 2016.

67. **L. Wanner** et al. 'Towards a Multimedia Knowledge-Based Agent with Social Competence and Human Interaction Capabilities'. In *Proceedings of the MARMI 2016 Workshop*, New York City, NY, 2016.

68. G. Meditskos, S. Dasiopoulou, L. Pragst, S. Ultes, S. Vrochidis, I. Kompatsiaris, and **L. Wanner**. 'Towards an Ontology-Driven Adaptive Dialogue Framework'. In *Proceedings of the MARMI 2016 Workshop*, New York City, NY, 2016.

69. J. Codina-Filbà and **L. Wanner**. 'Combining Dictionary- and Corpus-Based Concept Extraction'. In *Proceedings of the First International Workshop on Multimodal Multimedia Data Analytics (MMDA), held in conjunction with the European Conference on Artificial Intelligence (ECAI)*, CEUR Workshop Proceedings 1801, The Hague, The Netherlands, 2016.

70. S. Mille, M. Ballesteros, A. Burga, G. Casamayor, and **L. Wanner**. 'Multilingual Natural Language Generation within Abstractive Summarization'. In *Proceedings of the First International Workshop on Multimodal Multimedia Data Analytics (MMDA), held in conjunction with the European Conference on Artificial Intelligence (ECAI)*, CEUR Workshop Proceedings 1801, The Hague, The Netherlands, 2016.

2015

71. S. Dasiopoulou, S. Lohmann, J. Codina, and **L. Wanner**. 'Representing and Visualizing Text as Ontologies: A Case from the Patent Domain'. In *Proceedings of the Visualizations and User Interfaces for Ontologies and Linked Data (VOILA!2015) Workshop*, co-located with ISWC 2015, Bethlehem, PA., 2015.

72. M. Ballesteros, B. Bohnet, S. Mille, and **L. Wanner**. 'Data-driven sentence generation with non-isomorphic trees'. In *Proceedings of the Annual Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL)*, Denver, CO, 2015.

73. J. Soler Company, M. Ballesteros, B. Bohnet, S. Mille, and **L. Wanner**. 'Visualizing Deep Syntactic Parser Output'. In *Proceedings of the Annual Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL), Demo Track*, Denver, CO, 2015.

74. J. Soler Company and **L. Wanner**. 'Multiple Language Gender Identification for Blog Posts'. In *Proceedings of the 37th Annual Cognitive Science Society Meeting*, Pasadena, CA, 2015.

75. A. Burga, S. Mille, A. Granvik and **L. Wanner**. 'Towards a multi-layered dependency annotation of Finnish'. In *Proceedings of the Third International Dependency Linguistics Conference (DepLing)*, Uppsala, 2015.

76. S. Rodríguez Fernández, R. Carlini, and **L. Wanner**. 'Classification of Grammatical Collocation Errors in the Writings of Learners of Spanish'. In *Proceedings of the Annual Spanish Computational Linguistics Conference (SEPLN)*, Alicante, 2015.

77. S. Rodríguez Fernández, R. Carlini, and **L. Wanner**. 'Classification of Lexical Collocation Errors in the Writings of Learners of Spanish' (poster). In *Proceedings of the Conference on Recent Advances in Natural Language Processing (RANLP)*, Hissar, Bulgaria, 2015.

78. R. Carlini, J. Codina Filbà, and **L. Wanner**. 'Improving the Use of Electronic Collocation Resources by Visual Analytics Techniques'. In *Proceedings of the eLex 2015 Conference: Electronic lexicography in the 21st century: Linking lexical data in the digital age*, Herstmonceux Castle, UK, 2015.

79. M. Alonso Ramos, R. Carlini, J. Codina Filbà, A. Orol, O. Vincze, and **L. Wanner**. 'Towards a Learner Need-Oriented Second Language Writing Assistant'. In *Proceedings of the European Conference on Computer Assisted Language Learning (CALL)*, Padova, Italy, 2015.

80. S. Vrochidis, I. Kompatsiaris, G. Casamayor, I. Arapakis, R. Busch, V. Alexiev, E. Jean Jacques Jamin, M. Jugov, N. Heise, T. Forrellat, D. Liparas, **L. Wanner**, I. Miliaraki, V. Aleksic, K. Simov, A. Mas Soro, M. Eckhoff, T. Wagner, M. Puigbó. 'MULTISENSOR: Development of Multimedia Content Integration Technologies for Journalism, Media Monitoring and International Exporting Decision Support'. In *Proceedings of the IEEE International Conference on Multimedia and Expo, European Project Track*, Torino, Italy, 2015.

81. S. Mille and **L. Wanner**. 'Towards large-coverage detailed lexical resources for data-to-text generation'. In *Proceedings of the 1st International Workshop on Data-to-Text Generation*, Edinburgh, Scotland, 2015.

2014

82. R. Carlini, J. Codina Filbà, and **L. Wanner**. 'Improving Collocation Correction by Ranking Suggestions Using Linguistic Knowledge'. In Proceedings of the 3rd Workshop on NLP for computer-assisted language learning. Uppsala, Sweden, 2014.

83. M. Ballesteros, B. Bohnet, S. Mille, and **L. Wanner**. 'Deep-Syntactic Parsing'. In Proceedings of the International Conference on Computational Linguistics (COLING), Dublin, Ireland, 2014.

84. M. Ballesteros, S. Mille, and **L. Wanner**. 'Classifiers for Deep Data-Driven Sentence Generation'. In Proceedings of the International Conference on Natural Language Generation, Philadelphia, PA, 2014.

85. S. Vrochidis, I. Kompatsiaris, V.Alexiev, G. Casamayor, E. Staromiejski, I. Arapakis, R. Busch, M. Jugov, N. Heise, T. Forrellat, A. Moumtzidou, M. Grinberg, **L. Wanner**, A. Mas Soro, I. Miliaraki, V. Alexic, M. Eckhoff, T. Wagner, S. Pedraz. 'MULTISENSOR: Mining and Understanding of multilinguaL contenT for Intelligent Sentiment Enriched coNtext and Social Oriented inteRpretation'. In EU Project Networking Session at the ESWC, Crete, Greece, 2014.

86. L. Johansson, V. Epitropou, K. Karatzas, **L. Wanner**, A. Karppinen, J. Kukkonen, S. Vrochidis and I. Kompatsiaris, 'The extraction and fusion of meteorological and air quality information for orchestrated services'. In Proceedings of the International Workshop on Environmental Multimedia Retrieval held in Conjunction with ACM International Conference on Multimedia Retrieval (ICMR), Glasgow, UK, 2014.

87. M. Domínguez, M. Farrús, A. Burga and **L. Wanner**. 'Towards Automatic Extraction of Prosodic Patterns for Speech Synthesis'. In *Proceedings of the 7th Speech Prosody Conference*, Dublin, Ireland, 2014.

88. M. Domínguez, M. Farrús, A. Burga and **L. Wanner**. 'The Information Structure–Prosody Language Interface Revisited'. In *Proceedings of the 7th Speech Prosody Conference*, Dublin, Ireland, 2014.

89. J. Soler and **L. Wanner**. 'How to Use Less Features and Reach Better Performance in Author Gender Identification'. In Proceedings of the International Conference on Linguistic Resources and Evaluation (LREC), Reykjavik, Iceland, 2014.

90. N. Bouayad-Agha, A. Burga, G. Casamayor, J. Codina, R. Nazar, and **L. Wanner**. 'An Exercise in Reuse of Resources: Adapting General Discourse Coreference Resolution for Detecting Lexical Chains in Patent Documentation'. In Proceedings of the International Conference on Linguistic Resources and Evaluation (LREC), Reykjavik, Iceland, 2014.

2013

91. B. Bohnet, A. Burga, and **L. Wanner**. 'Towards the Annotation of Penn TreeBank with Information Structure'. In Proceedings of *International Joint Conference on Natural Language Processing*, Nagoya, Japan, 2013.

92. S. Mille, A. Burga, and **L. Wanner**. 'AnCora-UPF: A Multi-Level Annotation of Spanish'. In *Proceedings of the 2nd International Conference on Dependency Linguistics*, Prague, Czech Republic, 2013.

93. L. Johansson, A. Karppinen, and **L. Wanner**. 'The Fusion of Meteorological- and Air Quality Information for Orchestrated Services Using Environmental Profiling'. In *Proceedings of the 16th International Conference on Information Fusion*, Istanbul, Turkey, 2013.

94. N. Bouayad-Agha, G. Casamayor, **L. Wanner**, and C. Mellish. 'Overview of the First Content Selection Challenge from Open Semantic Web Data'. In *Proceedings of the European Workshop on Natural Language Generation*, Sofia, Bulgaria, 2013.

95. A. Burga, J. Codina, G. Ferraro, H. Saggion, and **L. Wanner**. 'The Challenge of Syntactic Dependency Parsing Adaptation for the Patent Domain'. In *Proceedings of the ESSLI Workshop on Extrinsic Parse Improvement (EPI)*, Düsseldorf, Germany, 2013.

96. **L. Wanner**, H. Bosch, S. Vrochidis, N. Bouayad-Agha, G. Casamayor, L. Johansson, A. Karppinen, A. Moumtzidou, I. Kompatsiaris, T. Ertl. 'Involving the Expert in the Delivery of Environmental Information from the Web'. In *Proceedings of EnviroInfo 2013*, Hamburg, Germany, September, 2013.

97. S. Verlinde, M. Alonso Ramos, and **L. Wanner**. 'Writing Assistants and Automated Lexical Error Correction: Word Combinatorics'. In *Proceedings of eLex 2013: Electronic Lexicography in the 21st Century*, Tallinn, Estonia, October 2013.

98. P. Moreno, G. Ferraro, and **L. Wanner**. 'Can we determine the semantics of collocations without using semantics?'. In *Proceedings of eLex 2013: Electronic Lexicography in the 21st Century*, Tallinn, Estonia, October 2013.

2012

99. S. Mille, A. Burga, G. Ferraro, and **L. Wanner**. 'How Does the Granularity of an Annotation Scheme Influence Dependency Parsing Performance?'. In *Proceedings of International Conference on Computational Linguistics (COLING)*, Mumbay, India, 2012.

100. G. Ferraro and **L. Wanner**. 'Labeling Semantically Motivated Clusters of Verbal Relations'. In *Proceedings of SEPLN*, 2012.

101. R. Nazar, J. Vivaldi, and **L. Wanner**. 'Co-occurrence Graphs Applied to Taxonomy Extraction in Scientific and Technical Corpora'. In *Proceedings of SEPLN*, 2012.

102. **L. Wanner**. 'Generation of Patent Abstracts: A Challenge for Automatic Text Summarization' (keynote talk). In *Proceedings of the 2nd International Workshop on Exploiting Large Knowledge Repositories at SEPLN*, 2012.

103. **L. Wanner** et al., 'Personalized Environmental Service Orchestration for Quality of Life Improvement. In *Proceedings of AIAI* (2):351-360, 2012.

104. N. Bouayad-Agha, G. Casamayor, S. Mille, M. Rospocher, L. Serafini, J. Moßgraber, and **L. Wanner**. 'Generation of Multilingual Personalized Environmental Bulletins from an OWL-based Ontology'. In *Proceedings of the 26th EnviroInfo Conference*, Dessau, Germany, 2012.

105. **L. Wanner**, S. Mille, and B. Bohnet. 'Towards a Surface Realization-Oriented Corpus Annotation'. In *Proceedings of the 7th International Conference on Natural Language Generation*. Starved Rock, IL., 2012.

106. A. Belz, B. Bohnet, S. Mille, **L. Wanner**, M. White. 'The Surface Realization Task: Recent Developments and Future Plans'. In *Proceedings of the 7th International Conference on Natural Language Generation*. Starved Rock, IL., 2012.

107. N. Bouayad-Agha, G. Casamayor, **L. Wanner**, and C. Mellish. 'Content Selection from Semantic Web Data'. In *Proceedings of the 7th International Conference on Natural Language Generation*. Starved Rock, IL., 2012.

108. **L. Wanner** et al., 'Personalized Environmental Service Configuration and Delivery Orchestration: The PESCaDO Demonstrator'. In *Demo Proceedings of the Extended Semantic Web Conference (ESWC)*, Heraklion, Greece, 2012.

109. N. Bouayad-Agha, G. Casamayor, S. Mille, M. Rospocher, L. Serafini, and **L. Wanner**. 'From Ontology to NL Generation of Multilingual User-Oriented Environmental Reports'. In *Proceedings of NLDB 2012: 17th International Conference on Applications of Natural Language Processing to Information Systems*, Groningen, 2012.

2011

110. A. Burga, S. Mille, and **L. Wanner**. 'Looking Behind the Scenes of Syntactic Dependency Corpus Annotation: Towards a Motivated Annotation Schema of Surface-Syntax in Spanish'. In *Proceedings of the International Dependency Linguistics Conference*, Barcelona, Spain, 104-114, 2011.

111. B. Bohnet, **L. Wanner**, and S. Mille. 'Statistical Language Generation from Semantic Structures'. In *Proceedings of the International Dependency Linguistics Conference*, Barcelona, Spain, 251-261, 2011.

112. A. Burga, S. Mille, and **L. Wanner**. 'Towards the Annotation of Communicative Structure in Corpora'. In *Proceedings of the 5th International Conference on Meaning-Text Theory*, Barcelona, Spain, 37-47, 2011.

113. G. Ferraro, R. Nazar, and **L. Wanner**. 'Collocations: A Challenge in Computer-Assisted Language Learning'. In *Proceedings of the 5th International Conference on Meaning-Text Theory*, Barcelona, Spain, 69-79, 2011.

114. N. Bouayad-Agha, G. Casamayor, and **L. Wanner**. 'Content selection from an ontology-based knowledge base for the generation of football summaries'. In *Proceedings of the 13th European Workshop on Natural Language Generation*. Nancy, France, 72-81, 2011.

115. **L. Wanner** et al. 'Building an Environmental Information System for Personalized Content Delivery'. In *Proceedings of the International Symposium on Environmental Software Systems*. Brno, Czech Republic, 2011.

116. B. Bohnet, S. Mille, B. Favre, and **L. Wanner**. 'StuMaBa: From Deep Representation to Surface'. In *Proceedings of the 13th European Workshop on Natural Language Generation*. Nancy, France, 2011.

117. M. Alonso Ramos, **L. Wanner**, O. Vincze, R. Nazar, G. Ferraro, E. Mosqueira and S. Prieto. 'Annotation of collocations in a learner corpus for building a learning environment'. In *Proceedings of the Learner Corpus Research 2011 Conference*. Louvain-la-Neuve, Belgium, 2011.

118. N. Bouayad-Agha, G. Casamayor, **L. Wanner**, F. Díez and S. López Hernández. 'FootbOWL: Using a generic ontology of football competition for planning match summaries'. In *Proceedings of the Extended Semantic Web Conference (ESWC)*, Heraklion, Greece, 2011.

2010

119. **L. Wanner** et al. 'Service-Based Infrastructure for User-Oriented Environmental Information Delivery'. In *Proceedings of the ENVIP2010 Workshop at the EnviroInfo 2010*, Bonn, Germany, 2010.

120. B. Bohnet, **L. Wanner**, S. Mille, and A. Burga. 'Broad Coverage Multilingual Deep Sentence Generation with a Stochastic Multi-Level Realizer'. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING)*, Beijing, China. 2010.

121. S. Mille and **L. Wanner**. 'Syntactic Dependencies for Multilingual and Multilevel Corpus Annotation'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation (LREC)*, Malta. 2010.

122. M. Alonso Ramos, **L. Wanner**, N. Vázquez, O. Vincze, E. Mosqueira, S. Prieto, and G. Casamayor. 'Towards a Motivated Annotation Schema of Collocation Errors in Learner Corpora'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation (LREC)*, Malta, 2010.

123. B. Bohnet and **L. Wanner**. 'Open Source Graph Transducer Interpreter and Grammar Development Environment'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation (LREC)*, Malta, 2010.

2009

124. S. Mille, A. Burga, V. Vidal, and **L. Wanner**. 'Towards a Rich Dependency Annotation of Spanish Corpora'. In *Proceedings of the Spanish Conference on Computational Linguistics (SEPLN*, San Sebastian, 2009.

125. **L. Wanner**. 'Explanatory Combinatorial Lexicology and NLP Applications in Specialized Discourse' (invited paper). In *Proceedings of the Workshop Terminology and Lexical Semantics*, Montreal, 2009.

126. S. Mille, A. Burga, V. Vidal, and **L. Wanner**. 'Creating an MTT Treebank for Spanish'. In Proceedings of the Fourth International Conference on the Meaning-Text Theory, Montreal, 2009.

127. M. Alonso Ramos, **L. Wanner**, N. Vázquez Veiga, E. Mosqueira Suárez, and S. Prieto González. 'Tagging Collocations for Learners'. In *Proceedings of the International Conference on eLexicography in the 21st Century: New Applications, New Challenges*. Louvain-la-Neuve, Belgium, 2009.

128. N. Bouayad-Agha, G. Casamayor, G. Ferraro, S. Mille, V. Vidal, and **L. Wanner**. 'Improving the Comprehension of Legal Documentation: The Case of Patent Claims'. In *Proceedings of the International Conference on Artificial Intelligence in Law*. Barcelona, 2009.

129. N. Bouayad-Agha, G. Casamayor, G. Ferraro, and **L. Wanner**. 'Simplification of Patent Claim Sentences for their Paraphrasing and Summarization'. In *Proceedings of the 22nd International FLAIRS Conference*, Sanibel Island, FL, 2009.

2008

130. S. Mille and **L. Wanner**. 'Multilingual Summarization in Practice: The Case of Patent Claims'. In *Proceedings of the European Machine Translation Summit (EAMT 2008)*, Hamburg, 2008.

131. R. Nazar, **L. Wanner**, and J. Vivaldi. 'Two-Step Flow in Bilingual Lexicon Extraction from Unrelated Corpora'. In *Proceedings of the European Machine Translation Summit (EAMT 2008)*, Hamburg, 2008.

31

132. M. Alonso Ramos, O. Rambow, and **L. Wanner**. 'Using Semantically Annotated Corpora to Build Collocation Resources'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation (LREC)*, Marrakesh, 2008.

133. S. Mille and **L. Wanner**. 'Making Text Resources Accessible to the Reader: The Case of Patent Claims'. In *Proceedings of the International Conference on Linguistic Resources and Evaluation (LREC)*, Marrakesh, 2008.

2007

134. M. Giereth, S. Koch, I. Kompatsiaris, S. Papadopoulos, E. Pianta, L. Serafini, and **L. Wanner**. 'A Modular Framework for Ontology-based Representation of Patent Information'. In *Proceedings of JURIX 2007: The 20th Anniversary International Conference on Legal Knowledge and Information Systems*, Leiden, 2007.

135. F. Lareau and **L. Wanner**. 'Towards a generic multilingual dependency grammar for text generation'. In Tracy Holloway King and Emily Bender (eds.) In *Proceedings of Grammatical Engineering Across Frameworks (GEAF), CSLI Publications in Computational Linguistics ONLINE*, Stanford, 2007.

136. **L. Wanner**, Bohnet B., N. Bouayad-Agha, F. Lareau, A. Lohmeyer, and D. Nicklaß. 'On the Challenge of Creating and Communicating Air Quality Information A Case for Environmental Engineers'. In *Proceedings of Environmental Software Systems: Dimensions of Environmental Informatics*, Vol. 7, Prague, 2007.

137. **L. Wanner** et al. 'From Measurement Data to Environmental Information - MARQUIS - A Multimodal AiR QUality Information Service for the General Public'. In *Proceedings of Environmental Software Systems: Dimensions of Environmental Informatics*, Vol. 7, Prague, 2007.

138. **L. Wanner**. 'Towards New Generation Environmental Information Services'. In *Proceedings of EnviroInfo 2007*, Warsaw, 2007.

139. N. Bouayad and **L. Wanner**. 'Text Planning of Air Quality Information'. In *Proceedings of EnviroInfo 2007*, Warsaw, 2007.

140. D. Nicklaß, N. Bouayad, and **L. Wanner**. 'Addressee-Tailored Interpretation of Air Quality Data'. In *Proceedings of EnviroInfo 2007*, Warsaw, 2007.

141. B. Bohnet, F. Lareau, and **L. Wanner**. 'Automatic Production of Multilingual Environmental Information'. In *Proceedings of EnviroInfo 2007*, Warsaw, 2007.

142. R. Nazar, J. Vivaldi, and **L. Wanner**. 'Towards Quantitative Concept Analysis'. In *Proceedings of the Spanish Conference on Computational Linguistics (SEPLN)*, Sevilla, 39:139-146, 2007.

2006

143. V. Vidal Méndez, M. T. Cabré, and **L. Wanner**. 'Hacia una tipología de los fenómenos de combinatoria restringida del discurso especializado: implicaciones en su tratamiento terminográfico'. In *Proceedings of the II Congreso Internacional de Lexicografía Hispánica*. Alicante, 2006.

144. **L. Wanner**, B. Bohnet and M. Giereth. 'What Is Beyond Collocations? Insights from Machine Learning Experiments'. In *Proceedings of the EURALEX Conference*, Turin, 2006.

145. N. Bouayad-Agha, **L. Wanner** and D. Nicklass. 'Discourse Structuring of Dynamic Content' In *Proceedings of the Spanish Conference on Computational Linguistics (SEPLN)*. Zaragoza, 2006.

146. D. Burgos and **L. Wanner**. 'Using CBIR for Multilingual Terminological Glossary Compilation and Cross-Language Image Indexing. In *Proceedings of the Workshop OntoImage at the International Conference on Linguistic Resources and Evaluation (LREC)*, Genoa, 2006.

147. **L. Wanner** and M. Alonso Ramos. 'Local Document Relevance Clustering in IR Using Collocation Information'. In *Proceedings of the International Conference on Linguistic Resoures and Evaluation (LREC)*, Genoa, 2006.

148. I. da Cunha and **L. Wanner**. 'Resumen automático de articulos médicos en español: integración de técnicas de análisis textual, léxico, discursivo y sintáctico-communicativo. In *Proceedings of VII Congreso de Lingüística General*, Barcelona, 2006.

2005

149. I. da Cunha and **L. Wanner**. 'Towards the Automatic Summarization of Medical Articles in Spanish: Integration of textual, lexical, discursive and syntactic criteria. In *Proceedings of the Workshop "Crossing Barriers in Text Summarization Research" at the Conference "Recent Advances in Natural Language Processing"*, Borovets, Bulgaria, September, 2005.

150. **L. Wanner** *et al.* 'Das multimodale Luftqualitäts-Informations-System MARQUIS'. In *Proceedings of the International Symposium "Münsteraner GI-Tage"*, Münster, Germany. July, 2005.

151. **L. Wanner** and M. Alonso Ramos. 'What Kind of Entity is a Lexical Function?' In *Proceedings of the Second International Conference on Meaning-Text Linguistics*. Moscow, June 2005.

152. **L. Wanner**, B. Bohnet, M. Alonso Ramos and N. Vázquez Veiga. 'The True Deep Happiness: Towards the Automatic Classification of Adjective-Noun Collocations'. In *Proceedings of the COMPLEX-2005 Conference*, Budapest. June 2005.

2004

153. **L. Wanner**. 'El corpus como un diccionario de colocaciones?' In *Proceedings of the First International Conference on Hispanic Lexicography*, A Coruña. 2004.

154. **L. Wanner**, M. Alonso Ramos and A. Martí. Enriching the EuroWordNet by Collocations. In *Proceedings of the LREC '04*, Lisbon, 2004.

2003

155. **L. Wanner**. 'Definitions of Lexical Meanings: Some Reflections on Purpose and Structure. In *Proceedings of the First International Conference on Meaning-Text Theory*, Paris, 2003.

156. **L. Wanner**, B. Bohnet and M. Giereth. 'Deriving the Communicative Structure in Applied NLG'. In *Proceedings of the 9th European Workshop on Natural Language Generation at the Biannual Meeting of the European Chapter of the Association for Computational Linguistics*. Budapest. 2003.

157. M. Alonso Ramos and **L. Wanner**. 'Información colocacional y Recuperación de la Información. In *Proceedings of the Spanish Conference on Computational Linguistics (SEPLN)*. Alcala. 2003.

2002

158. **L. Wanner**, S. Klatt, O. Kapanadze and N. Kapanadze. 'Towards a Semantically Motivated Organization of a Valency Lexicon for NLP: The GREG Proposal'. In *Proceedings of the EURALEX Conference*, Copenhagen, 2002.

159. B. Bohnet, S. Klatt, and **L. Wanner**. 'A Bootstrapping Approach to Automatic Annotation of Functional Information to Adjectives with an Application to German'. In *Proceedings of the LREC '02 Workshop* Linguistic Knowledge Acquisition and Representation: Bootstrapping Annotated Language Data. Las Palmas, Canary Islands, 2002.

2001

160. B. Bohnet and **L. Wanner**. 'On Using a Parallel Graph Rewriting Grammar Formalism in Generation'. In *Proceedings of the 8th European Natural Language Generation Workshop at the Annual Meeting of the Association for Computational Linguistics*, Toulouse, France, 2001.

161. B. Bohnet, **L. Wanner** *et al.* 'Autotext-UIS: Automatische Produktion von Ozonkurzberichten im Umweltinformationssystem Baden-Württemberg'. In *Proceedings of the Workshop Hypermedia und Umweltschutz*. Ulm, May 2001.

162. **L. Wanner** and M. Alonso Ramos. 'Vers une approche sémantique pour l'identification des collocations en corpus. In *Proceedings of the Journée d'études de l'ATALA: La Collocation*. Paris, January 2001.

2000

163. B. Bohnet, A. Langjahr, and **L. Wanner**. 'A Development Environment for an MTT-Based Sentence Generator'. In *Proceedings of the First International Conference on Natural Language Generation*. Mitzpe Ramon, Israel, June 2000.

164. **L. Wanner**. 'On the Relevance of Semantics for the Representation of Collocations in a Multilingual Computational Lexicon'. In *Proceedings of the Annual Meeting of the German Linguistics Society*. Marburg, Germany, March, 2000.

165. B. Bohnet, A. Langjahr, and **L. Wanner**. 'A Development Environment for MTT-Based Sentence Generators'. In *Proceedings of Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*, September 2000.

166. **L. Wanner**, B. Bohnet and C. Gerstenberger. 'Generation of Air Pollutant Information'. In *Proceedings of Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*, September 2000.

1998–1990

167. S. Beale, S. Nirenburg, E. Viegas, and **L. Wanner**. 'De-Constraining Text Generation'. In *Proceedings of the 9th International Natural Language Generation Workshop*. Niagara-on-the-Lake, Ontario, Canada, June 1998.

168. **L. Wanner**. 'Nutzung eines konventionellen Valenzwörterbuches in einem MÜ-System: Insegnare l'italiano a CAT2'. IAI Technical Report. IAI, Saarbrücken, Germany, 1997.

169. E. Hovy and **L. Wanner**. 'Managing Sentence Planning Requirements'. In *Proceedings of the Workshop 'Gaps and Bridges: New Directions in Planning and Natural Language Generation' at ECAI '96*. Budapest, Hungary, August 1996.

170. **L. Wanner** and E. Hovy. 'The HealthDoc Sentence Planner'. In *Proceedings of the 8th International Natural Language Generation Workshop*. Herstmonceux, England, June 1996.

171. C. DiMarco, G. Hirst, **L. Wanner**, and J. Wilkinson. 'HealthDoc: Customizing Patient Information and Health Education by Medical Condition and Personal Characteristics'. In *Proceedings of the International Workshop on Patient Education*. Glasgow, July, 1995.

172. I. Mel'čuk and **L. Wanner**. 'Towards an Efficient Representation of Lexical Co-occurrence'. In *Proceedings of the European Conference on Lexicography (*EURALEX*)*, Amsterdam, September, 1994.

173. **L. Wanner**. 'On Lexically Biased Discourse Organization for Text Generation'. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, Kyoto, Japan, August, 1994.

174. **L. Wanner**. 'Building Another Bridge over the Generation Gap'. In *Proceedings of the 7th International Natural Language Generation Workshop*. Kennebunkport, MA, June 1994.

175. **L. Wanner**. 'Lexical Choice and the Organization of Lexical Resources in Text Generation'. In *Proceedings of the European Conference on Artificial Intelligence (ECAI)*, Vienna, Austria, August, 1992.

176. J.A. Bateman, E.A. Maier, E. Teich, and **L. Wanner**. 'Towards an Architecture for Situated Text Generation'. In *Proceedings of the International Conference on Current Issues in Computational Linguistics*, Penang, Malaysia, June, 1991.

177. **L. Wanner** and E. Maier. 'Lexical Choice as an Integrated Component of Situated Text Planning'. In *Proceedings of the 3rd European Workshop on Natural Language Generation*, Judenstein, Austria, April, 1991.

178. J.A. Bateman and **L. Wanner**. 'Towards a lexicon for German organized by communicative function: An application of 'lexical functions'. In *Proceedings of the 14th. German Workshop on Artificial Intelligence (GWAI)*, September, 1990.

179. **L. Wanner** and J.A. Bateman. 'Lexical Co-occurrence Relations in Text Generation'. In *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, Boston, MA., July, 1990.

180. **L. Wanner** and J.A. Bateman. 'A collocational based approach to salience-sensitive lexical selection'. In *Proceedings of the 5th International Workshop on Natural Language Generation*, Dawson, PA, June, 1990.

8. Invited Lectures

I gave invited lectures at the following universities / research institutions (at some of them, on several occasions):

Department of Computer Science, University of Augsburg • Catalan Association of Artificial Intelligence • Institute of Formal and Applied Linguistics, Charles University of Prague • Patent Documentation Group • Institut d'Estudis Catalans, Barcelona • ETI, Université de Genève • Institut für maschinelle Sprachverarbeitung, Universität Stuttgart • Fakultät für Linguistik und Literaturwissenschaft, Universität Bielefeld • Departamento de Informática, Universidad Politécnica de Madrid • Computer Science Department, Columbia University • Université Bordeaux 3 • Facultat de Filologia, Universitat Autonoma de Barcelona • LIDILEM, Université Grenoble 3 • TALANA, Université Paris 7 • Département de Linguistique, Université de Montréal • Institut für angewandte Linguistik, Universität des Saarlandes • Institut für Informatik, Technische Universität Berlin • CIS, Universität München.

9. Co-organization of Conferences/Workshops/Seminars

I am Member of the International Committee on Computational Linguistics (a permanent body for organization of the International Conferences on Computational Linguistics (COLING)).

Furthermore, I was involved in the co-organization of the following conferences/ workshops/seminars:

1. International Conference on Computational Linguistics (COLING), Gyeongju, Korea, Program Committee Co-Chair, 2022.

2. Special Session on Multimedia and Multimodal Analytics in the Medical Domain and Pervasive Environments at the 28th International Conference on Multimedia Modeling, Qui Nhon, Vietnam, 2022.

3. Special Session on Multimedia and Multimodal Analytics in the Medical Domain and Pervasive Environments at the 27th International Conference on Multimedia Modeling, Prague, Czech Republic, 2021.

4. International Conference on Computational Linguistics (COLING), Barcelona, Spain, Local Organization Co-Chair, 2020.

5. Third Multilingual Surface Realization Workshop: Shared Task and Beyond, co-located with the International Conference on Computational Linguistics, Barcelona, Spain, 2020.

6. Special Session on Multimedia and Multimodal Analytics in the Medical Domain and Pervasive Environments at the 26th International Conference on Multimedia Modeling, Daejeon, Korea, 2020.

7. Second Multilingual Surface Realization Workshop: Shared Task and Beyond, co-located with the International Conference on Empirical Methods in NLP, Hong Kong, 2019.

8. WS on Multimedia Analysis for Architecture, Design and Virtual Reality Games, co-located with the 17th International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, 2018.

9. First Multilingual Surface Realization Workshop: Shared Task and Beyond, co-located with the Annual Meeting of the Association for Computational Linguistics (ACL), Melbourne, Australia, 2018.

10. Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications, co-located with the Federated AI Meeting, Stockholm, Sweden, 2018.

11. International Conference on Computational Linguistics (COLING), Santa Fe, USA, Sponsoring Co-Chair, 2018.

12. Dagstuhl Seminar on Socially Competent Conversation Agents, Dagstuhl, Germany, 2017.

13. Special Session on Multimedia and Multimodal Interaction for Health and Basic Care Applications at the 23rd International Conference on Multimedia Modeling, Reykjavik, Iceland, 2017.

14. 1st International Workshop on Multimodal Data Analytics, in conjunction with the European Conference on Artificial Intelligence, The Hague, Netherlands, 2016.

15. 1st International Workshop on Multimedia Analysis and Retrieval for Multimodal Interaction, in conjunction with the ACM Conference on Multimedia Retrieval, NYC, NY, USA, 2016.

16. Third International Conference on Dependency Linguistics, Uppsala, Sweden, 2015.

17. Annual Conference of the Association for Computational Linguistics (ACL), Mentoring Chair, 2013.

18. Second International Conference on Dependency Linguistics, Prague, Czech Republic, 2013.

19. International Conference on Dependency Linguistics, Barcelona, Spain, 2011.

20. Fifth International Conference on Meaning-Text Theory, Barcelona, Spain, 2011.

21. NATO AIS Seminar on Language Engineering for Low- and Middle Density Languages, Batumi, Georgia, 2007.

22. Third International Conference on Meaning-Text Theory, Klagenfurt, Austria, 2007.

23. European Workshop on Natural Language Generation, Toulouse, France, July, 2001.

24. European Conference on Lexicography (EURALEX), Stuttgart, Germany, August, 2000.

25. Colloquium on the Meaning-Text Theory, St. Just, France, August, 1999.

26. Workshop "Discourse Relations and Discourse Markers" at the Annual Conference of the Association for Computational Linguistics (ACL) and the International Conference on Computational Linguistics (COLING), Montreal, Canada, 1998.

27. International Workshop on the Meaning-Text Theory, Darmstadt, Germany, July, 1992.

10. Membership in Editorial Boards, Programme or Selection Committees

Editorial Board Member of the *Computational Intelligence* and the *Frontiers in AI, Language and Computation* Journals.

I acted as research proposal reviewer of the European Commission's FP7, the Heidelberg Academy of Sciences, the Swiss Research Foundation, the Leibniz Foundation and the Czech National Research Council.

Starting from 2023, I attempt to reduce my reviewing workload to the minimum in order to have more time for research and project management activities.

Over the years, I have been member of the Programme Committees/Reviewer Boards of the following scientific events:

2023: ACL Rolling Review Action Editor and Reviewer (for ACL, EMNLP, NAACL and EACL)

2022: ACL Rolling Review Action Editor and Reviewer (for ACL, EMNLP, NAACL and EACL) • International Joint Conference on Artificial Intelligence (IJCAI) • IVA 2022

2021: Annual Conference of the Association for Computational Linguistics (ACL) (Area Chair for the Social Media track) • International Joint Conference on Artificial Intelligence (IJCAI) • Knowledge and Language Processing track at ACM SAC2021 • Annual Conference of the North-American Chapter of the ACL: Human Language Technologue (NAACL-HLT) • Biannual Conference of the European Chapter of the ACL (EACL)

2020: 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural

Language Processing • International Conference on Computational Linguistics (COLING) • Empirical Methods in Natural Language Processing Conference (EMNLP), both Main and Demo Tracks (outstanding reviewer of the Main Conf.) • Intelligent Virtual Agents (IVA) • Annual Conference of the Association for Computational Linguistics (ACL) • International Joint Conference on Artificial Intelligence (IJCAI) • International Conference on Linguistic Resources and Evaluation (LREC) • Knowledge and Language Processing track at ACM SAC2020 • International Conference on Natural Language Generation (INLG)

2019: SIGNLL Conference on Computational Natural Language Learning (CoNLL) • ACM CHI Conference on Human Factors in Computing Systems (CHI) • Annual Conference of the North-American Chapter of the ACL: Human Language Technologue (NAACL-HLT) • ACM Symposium on Applied Computing (SAC) • International Joint Conference on Artificial Intelligence (IJCAI) • Interlligent Virtual Agents (IVA) • Annual Conference of the Association for Computational Linguistics (ACL) (Area Chair for Text Generation) • Annual Conference of the Association for Computational Linguistics, SRW • Empirical Methods in Natural Language Processing Conference (EMNLP) (Area Chair for Text Generation and Summarization) and Demo Track of EMNLP-IJCNLP 2019 • Catalan Conference on Artificial Intelligence • International Conference on Dependency Linguistics (DepLing)

2018: Empirical Methods in Natural Language Processing Conference (EMNLP), both Main and Demo Tracks • Intelligent Virtual Agents (IVA) • International Conference on Natural Language Generation (INLG) • International Joint Conference on Artificial Intelligence (IJCAI) • Annual Conference of the Association for Computational Linguistics (ACL) • International Conference on Linguistic Resources and Evaluation (LREC) • Annual Conference of the North-American Chapter of the ACL: Human Language Technologue (NAACL-HLT) • International Conference on Autonomous Agents and Multiagent Systems (AAMAS) • Intelligent Environments • Catalan Conference on Artificial Intelligence

2017: International Joint Conference on Natural Language Processing (IJCNLP) • Empirical Methods in Natural Language Processing Conference (EMNLP), both Main and Demo Tracks • Annual Conference of the Association for Computational Linguistics (ACL) • International Conference on Dependency Linguistics (DepLing) • SemEval '17 • International Conference on Natural Language Generation (INLG) • Intelligent Virtual Agents (IVA) • Canadian Conference on Artificial Intelligence • Catalan Conference on Artificial Intelligence

2016: International Conference on Computational Linguistics (COLING) • Annual Conference of the North-American Chapter of the ACL: Human Language

Technologue (NAACL-HLT) • International Conference on Linguistic Resources and Evaluation (LREC) • Canadian Conference on Artificial Intelligence • International Conference on Natural Language Generation (INLG) • Intelligent Virtual Agents (IVA) • Catalan Conference on Artificial Intelligence • Gramm&Lex Workshop

2015: Annual Conference of the Association for Computational Linguistics (ACL) • Third International Conference on Dependency Linguistics (DepLing) • European Workshop on Natural Language Generation (ENLGW) • Environmental Multimedia Retrieval 2015 (EMR 2015) Workshop • 6th Symposium on Future Directions in Information Access (FDIA) • Canadian Conference on Artificial Intelligence

2014: Empirical Methods in Natural Language Processing Conference (EMNLP) • Annual Conference of the Association for Computational Linguistics (ACL), main conference and demo sessions • International Conference on Computational Linguistics (COLING) • International Conference on Natural Language Generation (INLG) • International Conference on Linguistic Resources and Evaluation (LREC) • Environmental Multimedia Retrieval 2014 (EMR 2014) Workshop • Spanish Annual Computational Linguistics Conference (SEPLN)

2013: International Joint Conference on Natural Language Processing (IJCNLP) • Annual Conference of the Association for Computational Linguistics (ACL) • Second International Conference on Dependency Linguistics (DepLing) • European Workshop on Natural Language Text Generation • NLP4ITA • 4th Intelligent Systems for Quality of Life Information Services Workshop

2012: Annual Conference of the Association for Computational Linguistics (ACL) • International Conference on Computational Linguistics (COLING) • International Conference on Linguistic Resources and Evaluation (LREC) • International Conference on Natural Language Generation (INLG) • LREC Workshop on Natural Language Processing for Improved Textual Accessibility • 3rd Intelligent Systems for Quality of Life Information Services Workshop

2011: CICLing • International Conference on Dependency Linguistics (DepLing) • International Conference on Meaning-Text Theory

2010: CICLing • International Conference on Computational Linguistics (COLING) • International Conference on Linguistic Resources and Evaluation (LREC)

2009: International Conference on Meaning-Text Theory • International Workshop on Terminology and Lexical Semantics • Workshop on Content Analysis in Web 2.0 at the 18th International World Wide Web Conference • 7th Brazilian Symposium in Information and Human Language Technology

2008: International Conference on Linguistic Resources and Evaluation (LREC) • European Conference on Lexicography (EURALEX) • LREC Workshop on Multilingual and Comparative Perspectives in Specialized Language Resources

2007: Colloque International "Le lexique des émotions et sa combinatoire" • International Conference on Meaning-Text Theory

2006: European Conference on Lexicography (EURALEX) • International Conference on Linguistic Resources and Evaluation (LREC)

2005: International Conference on Meaning-Text Theory • Conférence sur le Traitement Automatique des Langues Naturelles (TALN)

2004: International Conference on Computational Linguistics (COLING) • European Conference on Lexicography (EURALEX) • Workshop "Enhancing and Using Electronic Dictionaries" at COLING

2003: Conférence sur le Traitement Automatique des Langues Naturelles (TALN) • International Conference on Meaning-Text Theory • European Workshop on Natural Language Generation

2002: European Conference on Lexicography (EURALEX) • International Conference on Natural Language Generation

2001: Annual Conference of the Association for Computational Linguistics (ACL) and Biannual Conference of the European Chapter of the Association for Computational Linguistics (EACL) • Annual Conference of the Northamerican Chapter of the Association for Computational Linguistics (NAACL) • European Workshop on Natural Language Generation at ACL/EACL '01 • Workshop "Collocations" at ACL/EACL '01

2000: European Conference on Lexicography (EURALEX)

1998: Annual Conference of the Association for Computational Linguistics (ACL) and Biannual International Computational Linguistics Conference (COLING) • Workshop "Discourse Relations and Discourse Markers" at COLING/(ACL '98 • LREC Workshop "Minimizing the Effort for the Language Resource Acquisition"

1997: Annual Conference of the Association for Computational Linguistics (ACL) and Biannual Conference of the European Chapter of the Association for Computational Linguistics (EACL)

1996: Annual Conference of the Association for Computational Linguistics (ACL) • International Workshop on Natural Language Generation

11. Journal Reviewer Activities

I have been acting as reviewer of the following journals:[2]

Affective Computing (1) • *Computer Speech and Language* (4) • Corpus Linguistics and Linguistic Theory (1) • *Dialogue & Discourse* (1) • *Education Sciences* (1) • *Environmental Modeling and Software* (1) • *Expert Systems with Applications* (5) • *Functions of Language* (2) • *Information Processing & Management* (14) • *Information Sciences* (7) • *Information Systems* (1) • *International Journal of Human–Computer Studies* (3) • *Journal of Web Semantics* (2) • *Knowledge Based Systems* (10) • *Lingua* (1) • *Machine Translation* (6) • *Natural Language Engineering* (5) • *Revista Española de Lingística Aplicada* (1) • *The Modern Language Journal* (1) • *Semantic Web Journal* (2) • *Traitement automatique des langues* (5) • *Transactions of the Association for Computational Linguistics* (1) • *World Patent Information* (1)

12. Directed PhD Theses

1. Paula Fortuna, Topic: *Re-thinking Large Scale Hate Speech Identification: Beyond Common NLP Conventions and Supervised Machine Learning*, defended in March 2023. Pompeu Fabra University.

2. Laura Pérez Mayos, Topic: *In-depth exploration of the syntactic capabilities of autoencoding language models for downstream applications*, defended in June 2022. Pompeu Fabra University.

3. Gerard Casamayor del Bosque, Topic: *Semantically-oriented Text Planning for Automatic Summarization*, defended in April 2021. Pompeu Fabra University.

4. Sara Rodríguez Fernández, Topic: *Collocation and Collocation Error Processing in the Context of Second Language Learning*, defended in March 2018. Pompeu Fabra University

5. Mónica Domínguez Bajo, Topic: *The Information Structure – Prosody Interface: On the role of hierarchical thematicity in an empirically-grounded model*, defended in November 2017. Pompeu Fabra University.

6. Juan Soler Company, Topic: *Feature Engineering in Author Profiling and Identification: The Role of Syntax and Discourse*, defendeded in July 2017. Pompeu Fabra University.

7. Simon Mille, Topic: *Deep Stochastic Sentence Generation: Resources and Strategies*, defended in July 2014. Pompeu Fabra University.

---

[2]The number in parentheses behind a journal indicates the number of reviewed submissions for the journal in question.

8. Diego Burgos Herrera, Topic : *Towards an Image-Term Co-occurrence Model for Multilingual Terminology Alignment and Cross-Language Image Indexing*, defended in May 2014. Pompeu Fabra University.

9. Mark Giereth, Topic: *Visual Patent Analysis*, defended in October 2012. University of Stuttgart.

10. Gabriela Ferraro, Topic: *Towards deep content extraction from specialized discourse: The case of verbal relations in patent claims*, defended in July 2012, Pompeu Fabra University.

11. Rogelio Nazar, Topic: *An Approach to Quantitative Conceptual Analysis*, defended in October 2010, Pompeu Fabra University.

12. Iria da Cunha, Topic: *Hacia un modelo lingüístico de resúmen automático de artículos en español*, defended in April 2008, Pompeu Fabra University.

13. Bernd Bohnet, Topic: *Graph Transformation and Graph Grammars in Natural Language Processing*, defended in December 2005, University of Stuttgart.

Currently, I am directing several ongoing PhD Theses on the following topics:

1. *Personalized Hate Speech Analysis* (Department of Information and Communication Technologies, UPF).

2. *Evaluation of Neural Language Models for Automatic Text Generation* (Department of Information and Communication Technologies, UPF).

3. *Transformer-based Document Classification* (Department of Information and Communication Technologies, UPF).

4. *Lexical Processing for a Spanish Computer-Assisted Vocabulary Learning Application* (Department of Linguistics and Translation Studies, UPF).

13. Directed Master Theses

1. Dongjun Tang. "Experiments on Prompt Engineering for Combined Knowledge-to-Text and Text-to-Text Generation". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2023.

2. Igor Kuzmin. "Identification of communicative intents of speakers in spontaneous dialogues". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2023.

3. Kosuke J. Nishio. "Extraction and Categorization of Japanese Lexical Collocations with Graph-Aware Transformers" (co-supervision with Alexander Shvets). Master in Intelligent Interaction Systems, Pompeu Fabra University, 2022.

4. Victor Šimić de Torres: "Reinforcement Learning-Driven Dialogue Repair" (co-supervision with Alexander Shvets), Master in Intelligent Interaction Systems, Pompeu Fabra University, 2021.

5. Xiao Li: "Natural Language Paragraph Generation" (co-supervision with Simon Mille), Master in Intelligent Interaction Systems, Pompeu Fabra University, 2021.

6. Shixiao Du: "Exploring Neural Paraphrasing to Improve Fluency of Rule-Based Generation" (co-supervision with Simon Mille), Master in Intelligent Interaction Systems, Pompeu Fabra University, 2021.

7. José Pablo Umaña: "The Language of Art and Architecture" (co-supervision with Alexander Shvets), Master in Cognitive Systems and Interactive Media, Pompeu Fabra University, 2021.

8. Ana Gabriela Mestre Borges: "Towards Automatic Blurb Extraction from Novels" (co-supervision with Juan Soler-Company), Master in Intelligent Interaction Systems, Pompeu Fabra University, 2021.

9. Cristian Morales: "Neural Graph-to-Graph Transduction" (co-supervision with Bernd Bohnet, Google AI), Master in Intelligent Interaction Systems, Pompeu Fabra University, 2020.

10. Hamit Kavas (co-supervision with Juan Soler Company): "Political Bias in Media Coverage", Master in Intelligent Interaction Systems, Pompeu Fabra University, 2020.

11. Roberto Carlini: "Transformer Linearizer" (co-supervision with Vicenç Gómez). Master in Intelligent Interaction Systems, Pompeu Fabra University, 2019.

12. Peimin Wang: "Using Encoder-Decoder Architecture For Automatic Lyrics Generation". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2019.

13. Jessica Pérez Guijarro (co-supervision with Luis Espinosa-Anke): "Aion and Classification of Collocations with Word Embeddings". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2019.

14. Guo Cheng: "Resources and algorithms for supporting Chinese learners of English". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2019.

15. Albert Marin (co-supervision with Jérôme Noailly): "Automatic concept extraction from biomedical material". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2018.

16. Laura Pérez Mayos (co-supervision with Federico Sucno): "Improving the quality of video-to-language models by an optimized annotation of the training material". Master in Computer Vision, Pompeu Fabra University, 2017.

17. Fenglu Zhang: "Identifying the native tongue of authors of scientific publications written in English". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2017.

18. Sònia Gudayol Marquès: "Deep data-driven machine translation". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2017.

19. Carla Ten Ventura: "Communication Mode Selection in Humanoid Conversation Agents". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2017.

20. Joan Pere Sánchez Pellicer (co-supervision with Mireia Farrús): "Improvement of Speech Synthesis through Emotion Pitch Markers". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2014.

21. Joan Soler: "Author Profiling: Gender Identification". Master in Intelligent Interaction Systems, Pompeu Fabura University, 2013.

22. Raimon Bosch: "Sentiment Analysis: Incremental Learning to Build Domain Models". Master in Intelligent Interaction Systems, Pompeu Fabra University, 2013.

23. Johannes Graën: "Incremental generation of interaction moves for actors in a dynamic game environment". Master in Computational Linguistics, University of Stuttgart, 2012.

24. Gerard Casamayor: "Ontology-Based Content Selection as a Task of Automatic Text Generation: The Case of Football Commentaries". TIC Master, Pompeu Fabra University. 2010.

25. Gabriela Ferraro: "Approximación a la estructura del contenido de patentes" 'Approximation to the content structure of patents'. Master of the Institute for Applied Linguistics, Pompeu Fabra University. 2008.

26. Ivan Herreros: "Morphology Processing with Self Organising Networks". TIC Master, Pompeu Fabra University. 2007.

27. Diego Burgos: "Recuperación de terminología multilingüe a través de la imagen como mediación interlingüística" 'Retrieval of multilingual terminology via images as interlingual mediation'. Master of the Institute for Applied Linguistics, Pompeu Fabra University. 2005.

28. Iria da Cunha: "Hacia un modelo lingüístico de resumen automático de artículos médicos en español" 'Towards a linguistic model of automatic summarization of medical articles in Spanish'. Master of the Institute for Applied Linguistics, Pompeu Fabra University. 2005.

29. Vanesa Vidal: "Combinaciones verbo-nominales especializadas: análisis y aplicaciones" 'Specialized Verb-Noun Combinations: Analysis and Applications'. Master of the Institute for Applied Linguistics, Pompeu Fabra University. 2004.

30. Jan Finzen: "Lexikalisierung im Paraphasierungsparadigma" 'Lexicalization in the Paraphrasing Paradigm'. Diploma Degree, Computer Science Department, University of Stuttgart. 2004.

31. Vitali Specht: "Entwicklung eines dependenz-orientierten Parsers für die syntaktische Analyse deutscher Sätze" 'Development of a Dependency-Oriented Parser for the Syntactic Analysis of German Sentences', Diploma Degree, Computer Science Department, University of Stuttgart. 2003.

32. Konstantia Nalmpanti: "Entwicklung eines linguistisch motivierten Textzusammenfassungssystems" 'Development of a Linguistically Motivated Text summarization System'. Diploma Degree, Computer Science Department, University of Stuttgart. 2002.

33. Bernd Bohnet: "Entwurf und Implementierung eines intelligenten Unterstützungswerkzeugs für die Entwicklung von Generierungsgrammatiken" 'Design and Implementation of an intelligent Tool for the Development of Generation Grammars'. Diploma Degree, Computer Science Department, University of Stuttgart. 1999.

34. Mark Giereth: "Entwicklung eines Compilers für eine in der automatischen Generierung natürlichsprachlicher Texte eingesetzten Dependenzgrammatik" 'Development of a Compiler for a Dependency Grammar in Automatic Text Generation'. Diploma Degree, Computer Science Department, University of Stuttgart. 1999.

14. Teaching

At the Department of Information and Communication Technologies, UPF, I am currently teaching the following courses:

1. *Natural Language Interaction* (Master in Intelligent Interactive Systems).

2. *Natural Language Processing Techniques for Everyday Applications* (undergraduate level).

Between 1997 and 2005, I was teaching at the Computer Science and Linguistics Departments of the University of Stuttgart and at the Institute for Applied Linguistics, Pompeu Fabra University. Below, courses regularly taught at the University of Stuttgart and Pompeu Fabra and courses taught once one semester at the University of the Saarland and the University of Frankfurt/Main, respectively, are listed.

- **Courses in Computational Linguistics**:

  1. *Introduction to Computational Linguistics* (UPF), **2.** *Applied Natural Language Processing* (U of Stuttgart), **3.** *Current Syntactic Theories in Computational Linguistics* (U of Frankfurt), **4.** *Text Generation* (U of Stuttgart, U of Augsburg), **5.** *Machine Translation* (U of Stuttgart), **6.** *Theoretical Aspects of Natural Language Processing* (UPF), **7.** *Practical Aspects of Natural Language Processing* (UPF).

- **Courses in Linguistics**

  1. *Modern Syntactic Theories* (U of Stuttgart), **2.** *Lexicology* (U of Stuttgart), **3.** *Lexical Combinatorics in Theory and Practice* (U of Stuttgart), **4.** *Multilingual Specialized Dictionaries* (U of the Saarland)

- **Courses in Artificial Intelligence** (all at U of Stuttgart):

  1. *Introduction to AI*, **2.** *Selected Aspects of AI*, **3.** *Intelligent Information Retrieval*, **4.** *Intelligent Software Agents*, **5.** *Knowledge Representation*.