

Prof Constantin Orăsan

Last update: 1st September 2020

Personal details

Email: C.Orasan@surrey.ac.uk
Home page: <https://dinel.org.uk/>
GitHub: <https://github.com/dinel>
Linkedin: <https://www.linkedin.com/in/constantinorasan/>
ORCID: [0000-0003-2067-8890](https://orcid.org/0000-0003-2067-8890)
Phone: +44 (0)1483 684115

Education and courses attended

Dec. 2006	awarded PhD in Computational Linguistics by University of Wolverhampton with a thesis entitled <i>Comparative evaluation of modular automatic summarisation systems using CAST</i> under the supervision of Prof. Ruslan Mitkov and Dr. Chris Paice. (PhD thesis accepted without any revision required) http://tinyurl.com/35eto42
Oct. 1998 – Oct. 2001	Postgraduate research student on automatic summarisation, School of Humanities, Languages and Social Studies, University of Wolverhampton, UK
Oct. 1994 – July 1998	Bachelour Degree in computer science from Babes-Bolyai University, Cluj-Napoca, Romania (average grade for the four years 9.60 out of 10, average grade for university degree exam 9.66 out of 10)

Employment history

Feb 2020 – present	Professor of Language and Translation Technologies, Centre for Translation Studies, University of Surrey, UK
Aug 2013 – Feb 2020	Reader in computational linguistics and deputy head of the Research Group in Computational Linguistics, Research Institute in Information and Language Processing, University of Wolverhampton, UK
Oct 2006 – July 2013	Senior lecturer in computational linguistics and deputy head of the Research Group in Computational Linguistics, Research Institute in Information and Language Processing, University of Wolverhampton, UK
Oct 2004 – Sept 2006	Lecturer in computational linguistics, University of Wolverhampton, UK
Oct 2001 – Sept 2004	Research fellow in the Research Group of Computational Linguistics, University of Wolverhampton, UK
Oct. 1998 – Sept. 2001	Postgraduate research student, University of Wolverhampton, School of Humanities, Languages and Social Studies, University of Wolverhampton, UK
Sept 1997 – Sept 1998	Software developer writing programs for PCs and handheld computers in C, C++, Visual Basic at Transart SRL, Cluj-Napoca, Romania.

Areas of expertise

Computational linguistics: Translation technology, translation memories, machine translation quality estimation, automatic summarisation, text accessibility, question answering, information extraction, text mining, opinion mining, coreference and anaphora resolution, machine learning for language processing, computer-aided language processing, natural language processing for education

Programming languages and frameworks: Python, PHP, Java, Scala, Perl, C, C++, Octave, R, NLTK, GATE, Tensorflow, sklearn, weka, keras

Corpus linguistics: corpus building, corpus annotation, tools for corpus annotation, corpora for language learning

Web development: using HTML, JavaScript, PHP, MySQL, CSS, Symfony, Smarty, Java, Glassfish, JQuery

Operating systems: Linux and Windows

Publications

Journal articles and book chapters

1. Constantin Orăsan and Ruslan Mitkov (forthcoming) Recent developments in Natural Language Processing. In Mitkov, R. (ed.) *The Oxford Handbook of Computational Linguistics*. 2nd edition.
2. Constantin Orăsan (2019) Automatic summarisation: 25 years On. *Journal of Natural Language Engineering*. 25(6), pp. 735 – 751, doi:10.1017/S1351324919000524
3. Constantin Orăsan, Carla Parra Escartín, Lianet Sepúlveda Torres, and Eduard Barbu (2019) Exploiting data-driven hybrid approaches to translation in the EXPERT project. In Meng Ji and Michael Oakes (eds) *Advances in Empirical Translation Studies: Developing Translation Resources and Technologies*. Cambridge University Press, pp. 198 – 216, doi: 10.1017/9781108525695.011
4. Richard Evans and Constantin Orăsan (2019) Identifying Signs of Syntactic Complexity for Rule-Based Sentence Simplification. *Journal of Natural Language Engineering*. Cambridge University Press. 25(1), pp. 69 - 119, doi: 10.1017/S1351324918000384
5. Constantin Orăsan, Richard Evans, and Ruslan Mitkov (2018) Intelligent text processing to help readers with autism. In Shaalan, K., Hassanien, A-E. and Tolba, M. F. (eds.) *Intelligent Natural Language Processing: Trends and Applications*, Springer, pp. 713 – 740. doi: 10.1007/978-3-319-67056-0_33
6. Carla Parra Escartín, Hanna Béchara, and Constantin Orăsan (2017) Questing for Quality Estimation A User Study. *The Prague Bulletin of Mathematical Linguistics*. No. 108, pp. 343–354. doi: 10.1515/pralin-2017-0032
7. Eduard Barbu, Carla Parra Escartín, Luisa Bentivogli, Matteo Negri, Marco Turchi, Constantin Orăsan, and Marcello Federico (2016). The First Automatic Translation Memory Cleaning Shared Task. *Machine Translation*. 30(3-4), pp. 145 – 166. doi: 10.1007/s10590-016-9183-x
8. Rohit Gupta, Constantin Orăsan, Marcos Zampieri, Mihaela Vela, Josef Van Genabith, and Ruslan Mitkov (2016) Improving translation memory matching and retrieval using paraphrases. *Machine Translation*. 30(1), pp. 19 – 40, doi: 10.1007/s10590-016-9180-0
9. Rohit Gupta, Constantin Orăsan, Qun Liu, and Ruslan Mitkov (2016). A Dynamic Programming Approach to Improving Translation Memory Matching and Retrieval using Paraphrases. In Sojka, P., Horák, A., Kopeček, I., Pala, K. (Eds.) *Text, Speech and Dialogue. 19th International Conference, (TSD)*, Brno, Czech Republic., Springer pp. 259 – 269
10. Hanna Bechara, Carla Parra Escartín, Constantin Orăsan, Lucia and Specia (2016). Semantic Textual Similarity in Quality Estimation. *Baltic Journal of Modern Computing*, 4(2), 256 – 268.

11. Iustin Dornescu, and Constantin Orăsan (2014) Densification: Semantic Document Analysis using Wikipedia. *Journal of Natural Language Engineering*. Volume 20, Issue 4, pp. 469 – 500, doi:10.1017/S1351324913000296
12. Richard Evans, and Constantin Orăsan (2013). Annotating signs of syntactic complexity to support sentence simplification. In I. Habernal & V. Matousek (Eds.), *Text, Speech and Dialogue. Proceedings of the 16th International Conference TSD 2013*. Plzen, Czech Republic: Springer. pp. 92 – 104. doi: 10.1007/978-3-642-40585-3_13
13. Constantin Orăsan, Richard Evans, and Iustin Dornescu (2013) Text Simplification for People with Autistic Spectrum Disorders. In Dan Tufis, Vasile Rus and Corina Forascu (eds.), *Towards Multilingual Europe 2020: A Romanian Perspective*, Romanian Academy Publishing House, Bucharest, 2013, ISBN: 978-973-27-2282-4, pp. 287-312
14. Natalia Konstantinova, and Constantin Orăsan (2013) Interactive Question Answering. In S. Bandyopadhyay, S. Naskar, & A. Ekbal (Eds.), *Emerging Applications of Natural Language Processing: Concepts and New Research*, IGI Global, pp. 149-169, doi:10.4018/978-1-4666-2169-5.ch007
15. Ruslan Mitkov, Richard Evans, Constantin Orăsan, Iustin Dornescu and Miguel Rios (2012) Coreference Resolution: To What Extent Does It Help NLP Applications? In *Text, Speech and Dialogue. Lecture Notes in Computer Science Volume 7499*, 2012, pp 16-27
16. Natalia Konstantinova, Constantin Orăsan and Pedro P. Balage (2012) A Corpus-Based Method for Product Feature Ranking for Interactive Question Answering Systems. In *International Journal of Computational Linguistics and Applications*. Volume 3, Issue 1, pp. 57 – 70, ISSN 0976-0962
17. Jose G. Camargo de Souza and Constantin Orăsan (2011) Can Projected Chains in Parallel Corpora Help Coreference Resolution? In *Anaphora Processing and Applications*. Springer. pp. 59 – 69
18. Óscar Ferrández, Christian Spurk, Milen Kouylekov, Iustin Dornescu, Sergio Ferrández, Matteo Negri, Rubén Izquierdo, David Tomás, Constantin Orăsan, Günter Neumann, Bernardo Magnini, and Jose Luis Vicedo (2011) The QALL-ME Framework: A specifiable-domain multilingual Question Answering architecture, *Web Semantics: Science, Services and Agents on the World Wide Web*, Volume 9, Issue 2, Provenance in the Semantic Web, July, pp. 137 - 145, ISSN 1570-8268, doi: 10.1016/j.websem.2011.01.002.
19. Iustin Dornescu and Constantin Orăsan (2010) Interactive QA using the QALL-ME framework. In *International Journal of Computational Linguistics and Applications*, 1(1-2). pp. 233 – 247, ISSN 0976-0962
20. Constantin Orăsan (2009) The Influence of Pronominal Anaphora Resolution on Term-based Summarisation. In Nicolov, N.; Angelova, G. & Ruslan Mitkov (ed.) *Recent Advances in Natural Language Processing V, John Benjamins*, 309, 291-300
21. Constantin Orăsan (2009) Comparative evaluation of term-weighting methods for automatic summarisation. *Journal of Quantitative Linguistics*. 16(1). pp. 67 – 95, ISSN 0929-6174
22. Constantin Orăsan, Laura Hasler and Ruslan Mitkov (2009) Corpora in text summarization. In A. Ludeling and M. Kyto (eds) *Handbook of Corpus Linguistics*. Mouton de Gruyter. pp. 1271 – 1286
23. Georgiana Puscasu and Constantin Orăsan (2008) *University of Wolverhampton at CLEF 2007*. In Peters, C.; Jijkoun, V.; Mandl, Th.; Müller, H.; Oard, D.W.; Peñas, A.; Petras, V.; Santos, D. (Eds.) *Advances in Multilingual and Multimodal Information Retrieval: 8th Workshop of the Cross-Language Evaluation Forum, CLEF 2007*. Lecture Notes in Computer Science. Vol. 5152, Springer-Verlag
24. Constantin Orăsan, Le An Ha, Richard Evans, Laura Hasler, and Ruslan Mitkov (2007) Corpora for computational linguistics. *Ilha Do Desterro: A Journal of English Language, Literatures in English and Cultural Studies*. Vol. 52 (1). pp. 65 – 101, ISSN 2175-8026
25. Constantin Orăsan and Georgiana Puscasu (2007) A High Precision Information Retrieval Method for WiQA. In *Evaluation of Multilingual and Multi-modal Information Retrieval*, Lecture Notes in Artificial Intelligence (LNAI 4730), Springer-Verlag

26. Constantin Orăsan and Richard Evans (2007) NP animacy identification for anaphora resolution. *Journal of Artificial Intelligence Research*. Vol. 29 pp. 79 – 103, ISSN 1076 - 9757
27. Ruslan Mitkov, Richard Evans, Constantin Orăsan, Le An Ha and Viktor Pekar (2007) Anaphora Resolution: To What Extent Does It Help NLP Applications? In A. Branco (ed.) *Anaphora: Analysis, Algorithms and Applications*, Lecture Notes in Artificial Intelligence (LNAI 4410), Springer-Verlag, pp. 179 – 190
28. Constantin Orăsan (2005) Automatic annotation of Corpora for Text Summarisation: A Comparative Study. In *Computational Linguistics and Intelligent Text Processing*, Lecture Notes in Artificial Intelligence (LNAI 3406), Springer-Verlag, pp. 670 – 681
29. Ruslan Mitkov, Richard Evans, and Constantin Orăsan (2002): *A new, fully automatic version of Mitkov's knowledge-poor pronoun resolution method*. In *Computational Linguistics and Intelligent Text Processing*, Lecture Notes in Artificial Intelligence (LNAI 2276), Springer-Verlag
30. Constantin Orăsan and Richard Evans (2001): Experiments in Optimising the Task of Anaphora Resolution in B. Sharp, J. Filipe and J. Cordeiro (eds) *Enterprise Information Systems*, Kluwer Academic Publishers
31. Constantin Orăsan, Richard Evans, and Ruslan Mitkov (2000): Enhancing Preference-Based Anaphora Resolution with Genetic Algorithms. In Christodoulakis (Ed.) *Natural Language Processing - NLP 2000:Second International Conference*, Lecture Notes in Artificial Intelligence (LNAI 1835), Springer-Verlag, Germany, pp. 185 - 195

In conference and workshop proceedings

32. Tharindu Ranasinghe, Constantin Orăsan and Ruslan Mitkov (2020) Intelligent Translation Memory Matching and Retrieval with Sentence Encoders. In *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, Lisbon, Portugal, November, pp. 175 – 184
33. Alistair Plum, Marcos Zampieri, Constantin Orăsan, Eveline Wandler-Vogt and Ruslan Mitkov (2019) Large-scale Data Harvesting for Biographical Data. In *Proceedings of the International Conference on Biographical Data in a Digital World 2019*, Varna, Bulgaria, September
34. Alistair Plum, Tharindu Ranasinghe, Constantin Orăsan, C and Ruslan Mitkov (2019). RGCL at GermEval 2019: Offensive Language Detection with Deep Learning. In *Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)*. Erlangen, Germany. pp. 423 – 428
35. Tharindu Ranasinghe, Hadeel Saadany, Alistair Plum, Salim Mandhari, Emad Mohamed, Constantin Orăsan, Ruslan Mitkov (2019) RGCL at IDAT : Deep Learning models for Irony Detection in Arabic Language, In *Working Notes of the Forum for Information Retrieval Evaluation (FIRE 2019)*, pp. 416-425
36. Tharindu Ranasinghe, Constantin Orăsan and Ruslan Mitkov (2019) Semantic Textual Similarity with Siamese Neural Networks. In *Proceedings of Recent Advances in Natural Language Processing (RANLP2019)*, Varna, Bulgaria, September 4 – 6, pp. 1005 – 1012, doi: [10.26615/978-954-452-056-4_116](https://doi.org/10.26615/978-954-452-056-4_116)
37. Tharindu Ranasinghe, Constantin Orăsan and Ruslan Mitkov (2019) Enhancing Unsupervised Sentence Similarity Methods with Deep Contextualised Word Representations. In *Proceedings of Recent Advances in Natural Language Processing (RANLP2019)*, Varna, Bulgaria, September 4 – 6, pp. 994 – 1003, doi: [10.26615/978-954-452-056-4_115](https://doi.org/10.26615/978-954-452-056-4_115)
38. Alistair Plum, Tharindu Ranasinghe, and Constantin Orăsan (2019) Toponym Detection in the Bio-Medical Domain: A Hybrid Approach with Deep Learning. In *Proceedings of Recent Advances in Natural Language Processing (RANLP2019)*, Varna, Bulgaria, September, pp. 912-921. 10.26615/978-954-452-056-4_106. [doi:10.26615/978-954-452-056-4_106](https://doi.org/10.26615/978-954-452-056-4_106)

39. Richard Evans and Constantin Orăsan (2019) Sentence Simplification for Semantic Role Labelling and Information Extraction. In *Proceedings of Recent Advances in Natural Language Processing (RANLP2019)*, Varna, Bulgaria, September, pp. 285 – 294. doi:[10.26615/978-954-452-056-4_033](https://doi.org/10.26615/978-954-452-056-4_033)
40. Yaneva, Victoria, Constantin Orăsan, Le An Ha, and Natalia Ponomareva (2019). A Survey of the Perceived Text Adaptation Needs of Adults with Autism. In *Proceedings of Recent Advances in Natural Language Processing (RANLP 2019)*, Varna, Bulgaria, September 4 – 6, pp. 1356–1363. Varna, Bulgaria. doi: [10.26615/978-954-452-056-4_155](https://doi.org/10.26615/978-954-452-056-4_155)
41. Alistair Plum, Tharindu Ranasinghe, Pablo Calleja, Constantin Orăsan and Ruslan Mitkov (2019) RGCL-WLV at SemEval-2019 Task 12: Toponym Detection. In *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019)*, pp. 1297–1301, Minneapolis, Minnesota, USA, June 6–7, doi: [10.18653/v1/S19-2228](https://doi.org/10.18653/v1/S19-2228)
42. Reshmi Gopalakrishna Pillai, Mike Thelwall and Constantin Orăsan (2018) What Makes You Stressed? Finding Reasons From Tweets. In *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2018)*, pp. 266–272, Brussels, Belgium, October 31, doi: [10.18653/v1/W18-6239](https://doi.org/10.18653/v1/W18-6239)
43. Reshmi Gopalakrishna Pillai, Mike Thelwall and Constantin Orăsan (2018). Trouble on the Road : Finding Reasons for Commuter Stress from Tweets. In *Proceedings of the Workshop on Intelligent Interactive Systems and Language Generation (2IS&NLG)*, pp. 20–25. Tilburg, The Netherlands. doi: [10.18653/v1/W18-6705](https://doi.org/10.18653/v1/W18-6705)
44. Constantin Orăsan (2018) Aggressive language identification using word embeddings and sentiment features. In *Proceedings of the First Workshop on Trolling, Aggression and Cyberbullying*, pp. 113–119, Santa Fe, USA, August 25
45. Reshmi Gopalakrishna Pillai, Mike Thelwall, and Constantin Orăsan. (2018). Detection of Stress and Relaxation Magnitudes for Tweets. In *The 2018 Web Conference Companion (WWW'18 Companion)*, April 23–27, 2018, Lyon, France, ACM, New York, NY, 8 pages. doi: [10.1145/3184558.3191627](https://doi.org/10.1145/3184558.3191627)
46. Victoria Yaneva, Constantin Orăsan, Richard Evans, and Omid Rohanian (2017) Combining Multiple Corpora for Readability Assessment for People with Cognitive Disabilities. In *Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications*, Copenhagen, Denmark, p. 121 – 132, doi: [10.18653/v1/W17-5013](https://doi.org/10.18653/v1/W17-5013)
47. Constantin Orăsan (2016) The EXPERT Project: Training the Future Experts in Translation Technology, In *Proceedings of the 19th Annual Conference of the EAMT: Projects/Products*, p. 393
48. Hanna Bechara, Rohit Gupta, Li Ling Tan, Constantin Orăsan, Ruslan Mitkov, and Josef van Genabith (2016) WOLVESAAAR at SemEval-2016 Task 1: Replicating the Success of Monolingual Word Alignment and Neural Embeddings for Semantic Textual Similarity. *Proceedings of SemEval-2016*, San Diego, CA, USA, ACL, 2016, pp. 634–639
49. Eduard Barbu, Carla Parra Escartín, Luisa Bentivogli, Matteo Negri, Marco Turchi Marcello Federico, Luca Mastrostefano, and Constantin Orăsan (2016) 1st Shared Task on Automatic Translation Memory Cleaning Preparation and Lessons Learned. In *Proceedings of the 2nd Workshop on Natural Language Processing for Translation Memories (NLP4TM 2016)*, pages 1-6, 28 May, Portorož, Slovenia
50. Constantin Orăsan, Alessandro Cattelan, Gloria Corpas Pastor, Josef van Genabith, Manuel Herranz, Juan José Arevalillo, Qun Liu, Khalil Sima'an and Lucia Specia (2015) The EXPERT Project: Advancing the State of the Art in Hybrid Translation Technologies. In *Proceedings of Translating and the Computer 37*. London, UK
51. Rohit Gupta, Constantin Orăsan, and Josef van Genabith (2015). ReVal: A Simple and Effective Machine Translation Evaluation Metric Based on Recurrent Neural Networks. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Lisbon, Portugal, pp. 1066 – 1072

52. Rohit Gupta, Constantin Orăsan, and Josef van Genabith (2015). Machine Translation Evaluation using Recurrent Neural Networks. In *Proceedings of the Tenth Workshop on Statistical Machine Translation*, Lisbon, Portugal
53. Ismail El Maarouf, Georgiana Marsic and Constantin Orăsan (2015) Barbecued Opakapaka: Using Semantic Preferences for Ontology Population. In *Proceedings of Recent Advances in Natural Language Processing*, Hissar, Bulgaria, Sep 7–9, pp. 153–159
54. Hanna Béchara, Hernani Costa, Shiva Taslimipoor, Rohit Gupta, Constantin Orăsan, Gloria Corpas Pastor and Ruslan Mitkov (2015) MiniExperts: An SVM Approach for Measuring Semantic Textual Similarity. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*. Denver, Colorado, pp. 96 – 101
55. Rohit Gupta, Constantin Orăsan, Marcos Zampieri, Mihaela Vela, and Josef Van Genabith. (2015). Can Translation Memories afford not to use paraphrasing? In *Proceeding of European Association of Machine Translation 2015 (EAMT 2015)*, Antalya, Turkey, pp. 35 - 42
56. Hanna Béchara, Sara Moze, Ismail El-Maarouf, Constantin Orasan, Patrick Hanks and Ruslan Mitkov (2015) The Role of Corpus Pattern Analysis in Machine Translation Evaluation. In *Proceedings of the The 7th International Conference of the Iberian Association of Translation and Interpreting Studies (AIETI)*, Malaga, Spain.
57. Rohit Gupta, Hanna Bechara, and Constantin Orăsan (2014) Intelligent Translation Memory Matching and Retrieval Metric Exploiting Linguistic Technology. In *Proceedings of the thirty sixth Conference on Translating and Computer*, London, UK.
58. Justin Dornescu, Richard Evans and Constantin Orăsan (2014) Relative clause extraction for syntactic simplification. In *Proceedings of the Workshop on Automatic Text Simplification - Methods and Applications in the Multilingual Society (ATS-MA 2014)*, Dublin, Ireland, pp. 1 – 10
59. Rohit Gupta, Hanna Bechara, Ismail El Maarouf, and Constantin Orăsan (2014) UoW: NLP techniques developed at the University of Wolverhampton for Semantic Similarity and Textual Entailment, In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin, Ireland, pp. 785 – 789
60. Rohit Gupta and Constantin Orăsan (2014) Incorporating Paraphrasing in Translation Memory Matching and Retrieval. In *Proceedings of the Seventeenth Annual Conference of the European Association for Machine Translation (EAMT2014)*, Dubrovnic, Croatia, pp. 3 – 10
61. Richard Evans, Justin Dornescu, and Constantin Orăsan (2014) An evaluation of syntactic simplification rules for people with autism. In *Proceedings of of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)*. Gothenburg, Sweden, pp. 131 – 140
62. Justin Dornescu, Richard Evans, and Constantin Orăsan (2013). A Tagging Approach to Identify Complex Constituents for Text Simplification. In *Proceedings of Recent Advances in Natural Language Processing (RANLP 2013)*. Hissar, Bulgaria, pp. 221 – 229
63. Sanja Štajner, Richard Evans, Constantin Orăsan, and Ruslan Mitkov (2012) What can readability measures really tell us about text complexity?. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility (NLP4ITA)*, Istanbul, May 27, pp. 14 – 21
64. Irina Temnikova, Constantin Orăsan, and Ruslan Mitkov (2012) CLCM - A Linguistic Resource for Effective Simplification of Instructions in the Crisis Management Domain and its Evaluations. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul, Turkey, May, 23-25. pp. 3007 – 3014
65. Marta Recasens, M. Antònia Martí, and Constantin Orăsan (2012) Annotating Near-Identity from Coreference Disagreements. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul, Turkey, May, 23-25. pp. 165 – 172

66. Jose G. Camargo de Souza and Constantin Orăsan (2011) Coreference Resolution for Portuguese using Parallel Corpora Word Alignment. In *Proceedings of the International Conference on Knowledge Engineering, Principles and Techniques (KEPT2011)*, Cluj-Napoca, Romania, July 4–6
67. Natalia Konstantinova and Constantin Orăsan (2011) Issues in topic tracking in Wikipedia articles. In *Proceedings of the International Conference on Knowledge Engineering, Principles and Techniques (KEPT2011)*, Cluj-Napoca, Romania, July 4–6
68. Iustin Dornescu, Constantin Orăsan and Tatiana Lesnikova (2010) Cross-document coreference for WePS. In *Proceedings of Web People Search (WePS3) Workshop*, Padua, Italy
69. Shiyao Ou, Dalila Mekhaldi and Constantin Orăsan. (2009). An ontology-based question answering method with the use of textual entailment. In *Proceedings of International Conference on Natural Language Processing and Knowledge Engineering 2009*, pp. 212-219. Dalian, China, 24-27 September 2009. Publisher: IEEE.
70. Constantin Orăsan, Iustin Dornescu and Natalia Ponomareva (2009) QALL-ME Needs AIR: a Portability Study. In *Proceedings of Adaptation of Language Resources and Technology to New Domains (AdaptLRTtoND) Workshop*, Borovets, Bulgaria, pp. 50 – 57
71. Constantin Orăsan and Iustin Dornescu (2009) WLV: A Confidence-based Machine Learning Method for the GREC-NEG'09 Task. In *Proceedings of the 2009 Workshop on Language Generation and Summarisation (UCNLG+Sum 2009)*, Association for Computational Linguistics, pp. 107-108
72. Andrea Varga, Georgiana Puscasu and Constantin Orăsan (2009) Identification of temporal expressions in the domain of tourism. In *Proceedings of Knowledge Engineering: Principles and Techniques Conference*, Cluj-Napoca, Romania, pp. 29 - 32
73. Irina Temnikova and Constantin Orăsan (2009). Post-editing Experiments with MT for a Controlled Language. *Proceedings of the International Symposium on Data and Sense Mining, Machine Translation and Controlled Languages (ISMTCL)*, Besancon, France, July 1-3, 2009
74. Iustin Dornescu, Georgiana Pușcașu and Constantin Orăsan (2008) University of Wolverhampton at CLEF 2008. In *Working Notes for the Cross Language Evaluation Forum (CLEF 2008)*, Aarhus, Denmark
75. Bogdan Sacaleanu, Constantin Orăsan, Christian Spurk, Shiyao Ou, Oscar Ferrandez, Milen Kouylekov and Matteo Negri (2008) Entailment-based Question Answering for Structured Data. In *Proceedings of Coling 2008: Companion volume: Posters and Demonstrations*. Manchester, UK, August, pp. 29 - 32
76. Constantin Orăsan and Oana Chiorean (2008) Evaluation of a Crosslingual Romanian-English Multi-document Summariser. *Proceedings of 6th Language Resources and Evaluation Conference (LREC2008)*, Marrakesh, Morocco, 28 – 30 May
77. Constantin Orăsan, Dan Cristea, Ruslan Mitkov, and Antonio Branco (2008) Anaphora Resolution Exercise – an overview. *Proceedings of 6th Language Resources and Evaluation Conference (LREC2008)*, Marrakesh, Morocco, 28 – 30 May
78. Elena Cabrio, Milen Kouylekov, Bernardo Magnini, Matteo Negri, Laura Hasler, Constantin Orăsan, David Tomas, Jose Luis Vicedo, Guenter Neumann and Corinna Weber (2008) The QALL-ME benchmark: a Multilingual Resource of Annotated Spoken Requests for Question Answering. *Proceedings of 6th Language Resources and Evaluation Conference (LREC2008)*, Marrakesh, Morocco, 28 – 30 May
79. Shiyao Ou, Viktor Pekar, Constantin Orăsan, Christian Spurk, and Matteo Negri (2008) Development and Alignment of a Domain-Specific Ontology for Question Answering. *Proceedings of 6th Language Resources and Evaluation Conference (LREC2008)*, Marrakesh, Morocco, 28 – 30 May
80. Shiyao Ou, Constantin Orăsan, Dalila Mekhaldi, and Laura Hasler (2008). Automatic Question Pattern Generation for Ontology-based Question Answering. In *Proceedings of the 21st International Florida Artificial Intelligence Research Society Conference (FLAIRS2008)*. Menlo Park, CA: AAAI Press., pp. 183 – 188

81. Constantin Orăsan (2007) Pronominal anaphora resolution for text summarisation. In *Proceedings of RANLP2007*, 27 – 29 September, Borovets, Bulgaria
82. Constantin Orăsan and Laura Hasler (2007) Computer-aided summarisation: How much does it really help? In *Proceedings of RANLP2007*, 27 – 29 September, Borovets, Bulgaria
83. Georgiana Puscasu and Constantin Orăsan (2007) University of Wolverhampton at CLEF 2007. In *Working Notes for the CLEF 2007 Workshop*, 19 – 21 September, Budapest, Hungary
84. Constantin Orăsan (2007) The role of linguistic information for shallow language processing. In *Proceedings of Knowledge engineering principles and technologies conference (KEPT2007)*, 6 – 8 June, Cluj-Napoca, Romania (invited speech)
85. Georgiana Puscasu, Adrian Iftene, Ionel Pistol, Diana Trandabat, Dan Tufiş, Alex Ceausu, Dan Stefanescu, Radu Ion, Constantin Orăsan, Iustin Dornescu, Alex Moruz, and Dan Cristea (2006) Developing a Question Answering System for the Romanian-English Track at CLEF 2006. In *Working Notes for the CLEF 2006 Workshop*, 20 – 22 September, Alicante, Spain
86. Constantin Orăsan and Georgiana Puscasu (2006) A high precision information retrieval method for WiQA. In *Working Notes for the CLEF 2006 Workshop*, 20 – 22 September, Alicante, Spain
87. Constantin Orăsan and Laura Hasler (2006) Computer-aided summarisation – what the user really wants. In *Proceedings of the 5th Language Resources and Evaluation Conference (LREC2006)*, Genoa, Italy, May 24 – 26, pp. 1548 – 1551
88. Laura Hasler, Constantin Orăsan and Karin Naumann (2006) NPs for Events: Experiments in Coreference Annotation. In *Proceedings of the 5th Language Resources and Evaluation Conference (LREC2006)*, Genoa, Italy, May 24 – 26, pp.1167 – 1172
89. Oana Postolache, Dan Cristea and Constantin Orăsan (2006). Transferring coreference chains through word alignment. In *Proceedings of the 5th Language Resources and Evaluation Conference (LREC2006)*, Genoa, Italy, May 24 – 26, pp. 889 – 892
90. Constantin Orăsan, Ted Marshall, Rob Clark, Le An Ha, and Ruslan Mitkov (2005) Building a WSD Module within an MT system to enable interactive resolution in the user's source language, In *Proceedings of EAMT2005*, Budapest, Hungary
91. Le An Ha and Constantin Orăsan (2005) Concept-centred summarisation: producing glossary entries for terms using summarisation methods. In *Proceedings of RANLP2005*, Borovetz, Bulgaria
92. Constantin Orăsan (2004) The influence of personal pronouns for automatic summarization of scientific articles. In *Proceedings of DAARC2004*, Furnas, S. Migue, Azores, Portugal, September, 23 – 24, pp. 127 – 132
93. Ruslan Mitkov and Constantin Orăsan (2004) Discourse and coherence: revising the claims and conventions in centering theory. In *Proceedings of DAARC2004*, Furnas, S. Migue, Azores, Portugal, September, 23 – 24
94. Constantin Orăsan, Viktor Pekar and Laura Hasler (2004) A comparison of summarization methods based on term specificity estimation. In *Proceedings of the 4th Language Resources and Evaluation Conference (LREC2004)*, Lisbon, Portugal
95. Agnes Tutin, Meriam Haddara, Ruslan Mitkov, and Constantin Orăsan (2004) Annotation of anaphoric expressions in an aligned bilingual corpus. In *Proceedings of the 4th Language Resources and Evaluation Conference (LREC2004)*, Lisbon, Portugal
96. Constantin Orăsan (2003) Human centered evaluation of coherence and cohesion in automatic extracts. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing*. Borovets, Bulgaria. 10 – 12 September, pp. 352 – 356
97. Constantin Orăsan (2003) An evolutionary approach for Improving the Quality of Automatic Summaries. In *Proceedings of the Multilingual Summarisation and Question Answering – Machine Learning and Beyond Workshop*. Sapporo, Japan, 11th July, pp. 37 – 45

98. Constantin Orăsan (2003) PALinkA: a highly customizable tool for discourse annotation. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialog*, Sapporo, Japan, 5 – 6 July, pp. 39 – 43
99. Constantin Orăsan, Doina Tatar, Gabi Serban, Dana Lupsa and Adrian Onet (2003): How to build a QA system in your back-garden: application for Romanian. In *Proceedings of Research Notes of the 10th Conference of the European Chapter of the Association for Computational Linguistics*, Budapest, Hungary, pp. 139 – 142
100. Constantin Orăsan, Ruslan Mitkov, and Laura Hasler (2003): CAST: a Computer-Aided Summarisation Tool. In *Proceedings of Research Notes Sessions of the 10th Conference of The European Chapter of the Association for Computational Linguistics (EACL2003)*, Budapest, Hungary, pp. 135 – 138
101. Laura Hasler, Constantin Orăsan and Ruslan Mitkov (2003): Building better corpora for summarisation. In *Proceedings of Corpus Linguistics 2003*, Lancaster, UK, March, pp. 309 – 319
102. Constantin Orăsan (2002): A study of anaphoric expressions in human produced abstracts, in *Proceedings of the 4th Discourse Anaphora and Anaphor Resolution Colloquium*, Lisbon, Portugal, 18 – 20 September, pp. 147 – 152
103. Constantin Orăsan and Richard Evans (2002): Assessing the difficulty of finding people in texts, in *Proceedings of 3rd Conference on Language Resources and Evaluation (LREC2002)*, Las Palmas de Grand Canaria, May 29 – 31
104. Constantin Orăsan and Ramesh Krishnamurthy (2002): A corpus-based investigation of junk emails, in *Proceedings of 3rd Conference on Language Resources and Evaluation (LREC2002)*, Las Palmas de Grand Canaria, May 29 – 31
105. Constantin Orăsan (2002): Building annotated resources for automatic text summarisation, in *Proceedings of 3rd Conference on Language Resources and Evaluation (LREC2002)*, Las Palmas de Grand Canaria, May 29 – 31
106. Constantin Orăsan and Richard Evans (2001): Learning to identify animate references, in *Proceedings of the Workshop Computational Natural Language Learning (CoNLL-2001)*, Toulouse, France, 6th – 7th July 2001, pp. 129 – 137
107. Constantin Orăsan (2001): Patterns in scientific abstracts, in *Proceedings of Corpus Linguistics 2001 Conference*, Lancaster University, Lancaster, UK, pp. 433 – 443
108. Richard Evans and Constantin Orăsan (2000): Improving anaphora resolution by identifying animate entities in texts, in P. Baker, A. Hardie, T. McEnery and A. Siewierska (eds) *Proceedings of the Discourse Anaphora and Reference Resolution Conference (DAARC2000)*, Lancaster University, Lancaster, UK, 16 – 18 November 2000, pp. 154 – 163
109. Ruslan Mitkov, Richard Evans, Constantin Orăsan, Catalina Barbu, Lisa Jones, and Violeta Sotirova (2000): Coreference and anaphora: developing annotating tools, annotated resources and annotation strategies, in P. Baker, A. Hardie, T. McEnery and A. Siewierska (eds) *Proceedings of the Discourse Anaphora and Reference Resolution Conference (DAARC2000)*, Lancaster University, Lancaster, UK, 16 – 18 November 2000, pp. 49 – 58
110. Constantin Orăsan (2000): CLinkA a Coreferential Links Annotator in *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC'2000)*, Athens, Greece, pp. 491 – 496
111. Constantin Orăsan, and Ramesh Krishnamurthy (2000): An Open Architecture for the Construction and Administration of Corpora in *Proceedings of Proceedings of the Second International Conference on Language Resources and Evaluation (LREC'2000)*, Athens, Greece, pp. 793 – 800
112. Constantin Orăsan (2000): A hybrid method for clause splitting in unrestricted English texts, in *Proceedings of ACIDCA '2000, Corpora and Natural Language Processing*, March 22-24, Monastir, Tunisia, pp. 129 – 134.

113. Ruslan Mitkov, Constantin Orăsan, and Richard Evans (1999): The importance of annotated corpora for NLP: the cases of anaphora resolution and clause splitting in *Proceeding of "Corpora and NLP: Reflecting on Methodology Workshop"*, TALN'99

PhD Thesis

C. Orasan (2006) *Comparative Evaluation of Modular Automatic Summarisation Systems Using CAST*. PhD thesis. University of Wolverhampton

Edited proceedings

Irina Temnikova, Constantin Orăsan, Gloria Corpas and Ruslan Mitkov (eds.) (2019) *The Second Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT 2019)* at RANLP 2019, Sept 5-6, Varna, Bulgaria

<https://www.aclweb.org/anthology/volumes/W19-87/>

Guest editor of the double special issue on *NLP for Translation memories* in the Machine translation journal <http://rdcu.be/x71E>

Irina Temnikova, Constantin Orăsan, Gloria Corpas and Stephan Vogel (eds.) (2017) *The First Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT)* at RANLP 2017, Sept 7, Varna, Bulgaria

<https://aclanthology.coli.uni-saarland.de/volumes/proceedings-of-the-workshop-human-informed-translation-and-interpreting-technology>

Constantin Orăsan, Carla Parra Escartín, Eduard Barbu, and Marcello Federico (2016) *Proceedings of the 2nd Workshop on Natural Language Processing for Translation Memories (NLP4TM 2016)* at LREC 2016, 28 May, Portorož, Slovenia

http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-NLP4TM_Proceedings.pdf

Wolfgang Maier, Sandra Kübler, and Constantin Orasan (eds.) (2016) *Proceedings of Workshop on Discontinuous Structures in Natural Language Processing (DiscoNLP)*, San Diego, California, USA, June 17

<https://aclanthology.coli.uni-saarland.de/volumes/proceedings-of-the-workshop-on-discontinuous-structures-in-natural-language-processing>

Constantin Orasan and Rohit Gupta (eds.) (2015) *Proceedings of the First Workshop on Natural Language Processing for Translation Memories (NLP4TM 2015)*. Hissar, Bulgaria, 11 September 2015.

<http://lml.bas.bg/ranlp2015/docs/nlp4tm2015.pdf>

Projects

Sept 2018 - Feb 2020	Local coordinator and course leader of <i>the Erasmus Mundus Masters in Technology for Translation and Interpreting</i> (https://em-tti.eu)
Dec 2018 - Dec 2019	Co-investigator in the APIS project which develops automatic methods for mining biographical information about people who were important in the Austro-Hungarian Empire.
Dec 2016 - present	Co-investigator in the project on <i>Identification and analysis of metadiscourse strategies in Spanish and in English</i> , Reference: FFI2016-77941-P, Ministerio de Economía y Competitividad, Spain. The annotation tool which I am developing as a contribution to this project can be found at https://github.com/dinel/metadiscourse-annotator
Oct 2012 - Sept 2016	Coordinator of the EXPERT project, an FP7 EU-funded ITN network which trained early stage researchers in the field of hybrid language translation technologies. In addition to coordinating the project, I also trained

- researchers in the project and carried out research on translation memories and quality estimation for professional translators. (<http://expert-itn.eu/>)
- Oct 2011 – Sept 2014 Deputy coordinator of the FP7 EU-funded FIRST project which developed technologies for simplifying texts for people with Autism Spectrum Disorders (ASD) with the help of language technology (LT). The main output of the project is the OpenBook tool which enables carers of people with ASD to convert documents to a more accessible form (<http://first-asd.eu>)
- Jan 2011 – Jan 2012 Principal investigator of the Wolverhampton team in *Coreference as a continuum: Cross-linguistic validation and computational framework*, a project funded by the Catalan government, on the notion of identity. The project was coordinated by University of Barcelona, Spain. The main output of this project is the <http://clic.ub.edu/corpus/en/nident>
- June 2010 – Sept 2014 Local coordinator of the Erasmus Mundus Masters in Natural Language Processing & Human Language Technology (<http://mastermundusnlp-hlt.univ-fcomte.fr/>)
- July 2010 – Jan 2011 Consultant on knowledge transfer engagement within the West Midlands region by managing and delivering a specific small business programme using funding sourced through the regional development agency and Europe known as the Innovation Voucher scheme. This builds links between universities and SME's by creating a demand led approach to support innovation. The voucher enabled 30 plus hours of academic support to investigate integration of dictionaries with online translation engines.
- Oct 2006 – Sept 2009 Principal investigator of the Wolverhampton team in FP6 EU-funded project on *Question Answering Learning technologies in a multiLingual and Multimodal Environment*, (QALL-ME). In addition to the managing the project, worked on the Data access work package (development of the QALL-ME ontology, the IE modules and encoding of the data), implementation of the QALL-ME framework and evaluation of the project. (<http://clg.wlv.ac.uk/projects/QALL-ME/>)
- Jan 2008 – Aug 2009 Principal investigator of the Wolverhampton team in an EU-funded project on *Controlled languages standards for messages, alerts and protocols arising from terrorism and other security related risks* (<http://message-project.univ-fcomte.fr/>)
- Jan 2006 – July 2006 Principal investigator of a project on enhancing language teaching using corpora funded by University of Wolverhampton (<http://clg.wlv.ac.uk/projects/C4T/>)
- July 2005 – June 2007 Project manager of the Wolverhampton team in a project for development of syntax based anaphora resolution methods for German and English with University of Tuebingen, Germany
- Mar 2005 – June 2007 Programming consultant for Aston Corpus Network (ACORN), Aston University, UK. Provided consultancy on development of online tools for corpus building and access implemented in Java and PHP. (<http://corpus.aston.ac.uk/>)
- June 2005 – March 2006 Project manager of a British Academy funded project on *Annotation of Cross-document coreference* (<http://clg.wlv.ac.uk/projects/NP4E/>)
- June 2004 – Sept 2005 Project manager of projects on translation aids funded by Translution LTD. Supervised the development of various dictionary based tools that can help translation engines. Supervised a team of linguists to develop a WordNet-based multilingual dictionary that can be used by translation tools. Worked with C#, PHP and MySQL.

Jan 2001 – Dec 2003	Participant in an British Council funded project on multilingual anaphora resolution with Université de Grenoble 3, Grenoble, France
Oct 2001 – Sept 2004	Research fellow in the Research Group of Computational Linguistics, University of Wolverhampton, UK working on the AHRC funded project on Computer-Aided Summarisation Tool (CAST). The main role in this project was to carry out day to day research, implement all the necessary software in Java and Perl, maintain the project's website, prepare reports for the PI and supervise the research student.
Nov 2000 – Oct 2001	Participant in “Language, Computers and Style: Investigating Text Genres” project, University of Wolverhampton, UK
Oct 1999 – Oct 2002	Participant in a project for development of an integrated corpus-based anaphora resolution system for Spanish and English with University of Alicante, Spain
Oct 1999 – Apr. 2000	Computer officer for ELDA business corpus: writing software for text conversion from different formats; encoding with SGML, POS; and validating to CES Standards. The corpus is available for purchase from ELDA.
Oct. 1998 – Apr. 1999	Participant in British Council/Research Council of Hong Kong funded project on practical automatic abstracting with City University of Hong-Kong

Invited talks

June 2018	Invited talk on <i>Are computers useful for professional translators?</i> at the Annual Research Conference, University of Wolverhampton, UK
Nov 2017	Invited talk on Quality Estimation for Professional Translators at “New Translation and Interpreting Technologies” workshop, Málaga, Spain
June 2017	Invited talk on <i>Adding semantic information to translation memories</i> at the 1st Conference on Recent Advances in Artificial Intelligence (RAAI 2017), Bucharest, Romania
March 2016	Invited talk on <i>Rewriting syntactically complex English sentences in the FIRST project</i> , Birmingham University, UK
Feb 2016	Invited talk on <i>Corpus annotation using experts and non-experts: Examples from coreference annotation projects</i> , Polytechnic University of Valencia, Spain
Feb 2016	Invited talk on <i>A hybrid method for rewriting syntactically complex English sentences</i> , “Heinrich Heine” University Düsseldorf, Germany
Sept 2015	Tutorial on <i>New trends in NLP applications</i> at RANLP 2015, Hissar, Bulgaria
July 2014	Invited talk on <i>The FIRST project and lessons learnt from writing it</i> at Horizon 2020 Future and Emerging Technologies Funding Programme, University of Wolverhampton, UK
June 2013	Invited talk on <i>Research in computational linguistics at University of Wolverhampton</i> , Toyohashi University, Japan
Sept 2012	Invited talk on <i>Text simplification for people with Autistic Spectrum Disorders</i> at ' Diaspora in Scientific Research and Higher Education in Romania ' Conference, Bucharest, Romania,
July 2011	Invited talk on <i>From TREC to Watson: is open domain question answering a solved problem?</i> at KEPT 2011 , Cluj-Napoca, Romania http://tinyurl.com/4x8tgma
August 2010	Invited talk on <i>Annotation of anaphora and coreference for automatic processing</i> at Copenhagen Dependency Treebank Workshop , Copenhagen, Denmark http://tinyurl.com/24kaxb7

March 2010	Invited talk on <i>Porting the QALL-ME framework to Romanian</i> at Processing Romanian in Multilingual, Interoperational and Scalable Environments (PROMISE) , Iasi, Romania http://tinyurl.com/2bjhfc2
Sept 2009	Invited talk on <i>QALL-ME: Ontology and Semantic Web</i> at Driving Future Question Answering: Research Trends And Market Perspectives Workshop . Trento, Italy. http://tinyurl.com/3yr6ef3
Sept 2009	Tutorial on <i>Automatic text summarisation</i> at RANLP 2009 , Borovets, Bulgaria http://tinyurl.com/29nbt8
Sept 2009	Invited talk on <i>Multilingual information access using the QALL-ME framework: An example for Romanian</i> at Multilingual resources, technologies and evaluation for central and Eastern European languages Workshop , Borovets, Bulgaria
June 2007	Invited talk on <i>The role of linguistic information for shallow language processing</i> at Knowledge Engineering: Principles and Techniques (KEPT2007) , Cluj-Napoca, Romania http://tinyurl.com/36kknz
May 2007	Invited talk on <i>What is Computer-Aided Summarisation and does it really work?</i> , Department of linguistics, Tuebingen University, Germany http://tinyurl.com/22w37xx
Nov 2006	Seminar presentation on <i>Comparative evaluation of modular automatic summarisation systems using CAST</i> , Research Group in Computational Linguistics, University of Wolverhampton, UK
Jan 2006	Invited talk on <i>Automatic text summarisation and corpora</i> , School of Languages and Social Sciences, Aston University, UK
April 2004	Seminar presentation on <i>An integrated architecture for automatic summarisation</i> , Research Institute in Advanced Technologies, University of Wolverhampton, UK
Oct 1999	Presentation on <i>Tools for coreferential annotation</i> , at The Tuscan Word Centre, Italy
March 1999	Presentation on <i>A memory-based learning method for clause splitting</i> , City University, Hong Kong

Research supervision

PhD level

Completed

- Director of studies of [Richard Evans](#)' PhD thesis on *Sentence simplification for text processing* (completed June 2020)
- Director of studies of Hanna Bechara' PhD thesis on *A Semantic Textual Similarity Enhanced Quality Estimation Method and its Applications to Natural Language Processing Task* (completed Jan 2020)
- Director of studies of Rohit Gupta's PhD thesis on *Use of language technology to improve matching and retrieval in translation memories* (completed Sept 2016)
- Supervisor of Sanja Stajner's PhD thesis on *New Data-Driven Approaches to Text Simplification* (completed March 2015)
- Supervisor of Natalia Konstantinova's PhD thesis on *Knowledge acquisition from user reviews for Interactive Question Answering* (completed June 2013)
- Director of studies of Iustin Dornescu's PhD thesis on *Encyclopedic question answering* (completed Oct 2012)

- Supervisor of Georgiana Marsic's PhD thesis on [Temporal Processing of News: Annotation of Temporal Expressions, Verbal Events and Temporal Relations](#) (completed in Dec 2011)
- Supervisor of Laura Hasler's PhD thesis entitled [From extracts to abstracts: Human summary production operations for computer-aided summarisation](#) (completed July 2007)

Currently supervising

- Supervisor of Reshmi Gopalakrishna Pillai's PhD thesis on *An enhanced method to detect strength of stress expressed in social media content* (expected submission date autumn 2019)
- Director of studies of Alistair Plum's PhD thesis on *Wikipedia-based Approaches to Multi-Lingual Information Extraction* (expected submission date autumn 2021)
- Supervisor of Tharindu Ranasinghe's PhD thesis on *A semantic textual similarity metric based on deep learning for effective translation memories* (expected submission date autumn 2021)

Other related professional activities

Events organised

- | | |
|-------------|--|
| Sept 2019 | The Second Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT 2019) at RANLP 2019 (with Irina Temnikova, Gloria Corpas and Ruslan Mitkov) |
| Sept 2017 | The First Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT) at RANLP 2017 (with Irina Temnikova, Gloria Corpas and Stephan Vogel) |
| June 2016 | Workshop on Discontinuous Structures in Natural Language Processing at NAACL 2016, San Diego, California, USA (with Wolfgang Maier, Sandra Kübler) |
| May 2016 | Second Workshop on Natural Language Processing for Translation Memories (NLP4TM 2016) at LREC 2016, Portorož, Slovenia (with Carla Parra Escartín, Eduard Barbu and Marcello Federico) |
| May 2016 | First Shared Task on Cleaning of Translation Memories at NLP4TM2016 Portorož, Slovenia (with Eduard Barbu, Carla Parra Escartín, Luca Mastrostefano, Matteo Negri, Marco Turchi and Luisa Bentivogli) |
| Sept 2015 | First Workshop on Natural Language Processing for Translation Memories (NLP4TM 2015) at RANLP 2015, Hissar, Bulgaria (with Rohit Gupta) |
| Aug 2014 | Workshop on Automatic Text Simplification: Methods and Applications in the Multilingual Society at Coling 2014, Dublin, Ireland (with Cristina Vertan and Petya Osenova) |
| Sept 2009 | Workshop Events in Emerging Text Types (eETTs) at RANLP 2009, Borovets, Bulgaria (with Laura Hasler and Corina Forascu) |
| Sept 2007 | Workshop on Computer-Aided Language Processing (CALP07) in conjunction with Recent Advances in Natural Language Processing (RANLP07), Borovets, Bulgaria (with Sandra Kuebler) |
| March 2007 | First Anaphora Resolution Exercise (ARE2007) in conjunction with the 6 th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC2007), Lagos, Portugal (with Antonio Branco, Dan Cristea, Ruslan Mitkov) |
| August 2001 | <i>Workshop on Modular Programming applied to Natural Language Processing</i> in conjunction with EUROLAN01 Summer School, Iasi, Romania (with Catalina Barbu and Richard Evans) |

Area chair at international conferences

- Chair of the *Recent NLP applications* area at IJCNLP 2013
- Chair of the *Question Answering* area at COLING 2012

PhD theses examinations

- Alina Maria Ciobanu's PhD thesis on *Problems of Similarity in Computational Linguistics*, University of Bucharest, Romania
- Patrick Kenekayoro's PhD thesis on *Collaboration between UK universities: A machine-learning based webometric analysis*, University of Wolverhampton, UK
- Ehsan Mohammadi's PhD thesis on *Identifying the invisible impact of scholarly publications: A multi-disciplinary analysis using altmetrics*, University of Wolverhampton, UK
- Naveed Afzal's PhD thesis on *Unsupervised Relation Extraction for E-Learning Applications*. University of Wolverhampton, UK

Programme committees

in 2020: Language Resources and Evaluation Conference (LREC 2020), CoNLL 2020 (Multilingual work and translation track), EMNLP 2020, COLING 2020

in 2019: ACL 2019 (Resources and Evaluation Track and Applications Track), CONLL 2019, The Web Conference 2019 (Web Mining and Content Analysis and Semantics and Knowledge areas), EMNLP 2019, Cicling 2019, MT Summit, EAMT 2019, NAACL 2019 (Information Extraction Track), the Second Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC 2019)

in 2018: ACL 2018 (Resources and Evaluation, and Demos areas), COLING 2018, EAMT 2018, EMNLP 2018 (Social Media, Computational Social Science, Sentiment, Opinion Area and Demo Track), LREC 2018, WASSA2018, WWW2018 (Web Content Analysis, Semantics and Knowledge Track),

in 2017: ACL 2017 (IE, QA and Text Mining Applications Area), CICLing 2017, EACL 2017 (Generation and summarisation track), EMNLP 2017 (Sentiment Analysis and Opinion Mining track), FLAIRS 2017, RANLP 2017 (Main conference), WASSA 2017, [Workshop on Coreference Resolution Beyond OntoNotes](#) (CARBON 2017)

in 2016: ACL 2016 (IE and QA area), COLING 2016 (Paraphrasing and Textual Entailment Area), FLAIRS 2016, LREC 2016, MeTa4eS 2016, WASSA 2016, WWW2016 (Content Analysis track), [Workshop on Coreference Resolution Beyond OntoNotes](#) (CARBON 2016)

in 2015: ACL-IJCNLP 2015 (Information Extraction and Discourse, Coreference and Pragmatics Track), [CICLing 2015](#), EMNLP 2015 (Summarization and Generation Area), [NoDaLiDa 2015](#), RANLP 2015 (Main conference and student workshop), WASSA 2015

in 2014: [CICLing 2014](#), [COLING 2014](#) (Information retrieval and question answering area), [EMNLP 2014](#) (Summarisation and Generation Track), [LREC 2014](#), [SEPLN 2014](#), [WASSA 2014](#),

in 2013: [ACL 2013](#) (Question answering track), [CICLing 2013](#), FLAIRS 26 (Applied Natural Language Processing (ANLP) Special Track), [NLP4ITA 2013](#), [RANLP 2013](#) (main conference and student research workshop), [WASSA 2013](#)

in 2012: [ACL 2012](#) (Summarisation and Generation area), [CICLing 2012](#), [FLAIRS 2012](#) (Applied Natural Language Processing (ANLP) Special Track), [EACL 2012](#) (Discourse, Dialogue, and Pragmatics Area), [EACL 2012 Student Research Workshop](#), [LREC 2012](#), [NAACL 2012](#) (Social Media area), Natural Language Processing for Improving Textual Accessibility ([NLP4ITA](#)) Workshop, COLING 2012

in 2011: [CICLing 2011](#), [FLAIRS 2011](#) (Applied Natural Language Processing track), [ACL 2011](#) (Text Mining and Information Extraction, Resources, Question Answering tracks), [ESSLLI 2011 Student Session](#), [Workshop on Computational Approaches to Subjectivity and Sentiment Analysis \(WASSA 2.011\)](#), [DAARC 2011](#), [IJCNLP 2011](#) (Language resources track), [RANLP 2011](#) (main conference and student research workshop), [MICAI 2011](#), [STIL 2011](#)

in 2010: [Processing ROmanian in Multilingual, Interoperational and Scalable Environments \(PROMISE 2010\)](#), [WASSA 2010](#), [NAACL 2010](#) (discourse area), [ACL 2010](#) (language generation and summarisation area), [LREC 2010](#), [CICLing 2010](#), [FLAIRS 2010](#) (Applied Natural Language Processing track), [MICAI 2010](#), [ConsILR 2010](#), [IBERAMIA 2010](#)

in 2009: [STIL 2009](#), [RANLP 2009](#) (main conference and student research workshop), [ACL 2009](#) ("Discourse, Dialogue and Pragmatics" area), [DAARC2009](#), [EACL 2009](#) (question answering, information extraction and information retrieval area), [FLAIRS 2009](#) (Applied Natural Language Processing track), [CLEF 2009](#) (GikiCLEF track), [MICAI 2009](#)

in 2008: [The 2nd MMIES Workshop: Multi-source, Multilingual Information Extraction and Summarization](#) at COLING2008, [LREC2008](#), [MICAI2008](#), [FLAIRS2008](#) (Applied Natural Language Processing track), [ACL 2008](#) (student research workshop), [CICLing2008](#), [ConsILR 2008](#)

in 2007: [FLAIRS2007](#) (Applied Natural Language Processing track), [CICLing2007](#), [ACL 2007](#) (student research workshop), [RANLP 2007](#), [Knowledge Engineering: Principles and Techniques \(KEPT2007\)](#), [EUROLAN2007](#) (Doctoral Consortium), [ConsILR 2007](#), [DAARC 2007](#)

in 2006: [FLAIRS 2006](#) (Trends in Natural Language Processing track), [CICLing 2006](#), [ConsILR 2006](#), [EACL 2006](#) (student research workshop), [ESSLLI 2006](#) (student research workshop)

before 2006 (selected): [SIGIR 2005](#) (additional reviewer), [RANLP 2005](#), [DAARC 2004](#), [COLING 2004](#) (additional reviewer), [RANLP 2003](#), [ACL 2003](#) (student research workshop), [EACL 2003](#) (Research Notes and Demos), PorTAL 2002, RANLP 2001 (additional reviewer)

Peer reviewing activities

Member of the editorial board of the [Book series on Natural Language Processing](#) published by [John Benjamins](#).

Member of the editorial board of the [Journal of Natural Language Engineering](#)

Reviewer for [Corpora](#), [Computer Speech and Language Journal](#), [Journal of Interesting Negative Results in Natural Language Processing and Machine Learning](#), [Journal of Natural Language Engineering](#), [International Journal of Human-Computer Studies](#), [Knowledge and Information Systems \(KAIS\)](#), [Language Resources and Evaluation \(LREV\)](#), [ACM Computing Surveys](#), [BMC Bioinformatics](#), [Information Processing & Management](#) and [IEEE Transactions on Affective Computing](#)