Jörg Tiedemann
Professor, Supervisor for doctoral programme
Department of Digital Humanities
Language Technology
Doctoral Programme in Language Studies
Type of address: Postal address.
PL 24 (Unioninkatu 40)
00014
Finland
Email: jorg.tiedemann@helsinki.fi
Mobile: +358503119440
Phone: +358294123261

Qualifications

Computational Linguistics, PhD, Recycling Translations - Extraction of Lexical Data from Parallel Corpora and their Application in Natural Language Processing, Uppsala University
2000 → 2003
Award Date: 12 Dec 2003

Computer Science, M.Sc., Automatical Lexicon Extraction from Aligned Bilingual Corpora, Otto Von Guericke University, Magdeburg
1991 → 1997
Award Date: 11 Sept 1997

Principal Investigator
Period : 27.01.2016 - 31.12.2017 in

Principal Investigator
Period : 17.02.2018 - 31.12.2025 in Department of Digital Humanities

Employment

Professor
Department of Digital Humanities
University of Helsinki
Finland
1 Jan 2018 → present

Language Technology
University of Helsinki
University of Helsinki, Finland
1 Aug 2015 → present

Supervisor for doctoral programme
Doctoral Programme in Language Studies
University of Helsinki
Finland
1 Jan 2014 → present

Projects

1. **GreenNLP: Green NLP - controlling the carbon footprint in sustainable language technology**
   Tiedemann, J. (Project manager), Attieh, J. (Participant) & Nieminen, T. J. (Participant)
   Suomen Akatemia Projektilaskutus
   01/01/2023 → 31/12/2025

2. **High Performance Language Technologies**
   Tiedemann, J. (Project manager), Aulamo, M. (Participant), De Gibert Bonet, O. (Participant), Grönroos, S.-A. (Participant), Ji, S. (Participant), Mickus, T. (Participant), Vahtola, T. (Participant), Vazquez , R. (Participant) & Virpioja, S. P. (Participant)
   Charles University in Prague Faculty of Science Department of Teaching and Didactics of Biology
   01/09/2022 → 31/08/2025

3. **Uncertainty-aware neural language models**
   Tiedemann, J. (Project manager), Celikkanat, H. (Participant), Virpioja, S. P. (Participant) & Vazquez , R. (Participant)
   Academy of Finland, Suomen Akatemia Projektilaskutus
   01/01/2022 → 01/10/2025

4. **CorCoDial: Corpus-based computational dialectology: exploiting machine translation techniques to extract, visualize and interpret dialectal patterns**
   Scherrer, Y. (Project manager), Tiedemann, J. (Project manager), Kuparinen, O. V. (Participant), Mickus, T. (Participant), Miletic Haddad, A. (Participant), Psaltaki, E. (Participant), Roemling, D. (Participant), Siewert, J. (Participant) & Siewert, J. (Participant)
   Suomen Akatemia Projektilaskutus
   01/09/2021 → 31/08/2025

5. **Behind the words: Deep neural models of language meaning for industry-grade applciations**
   Tiedemann, J. (Project manager), Creutz, M. (Principal Investigator), Creutz, M. (Participant), Itkonen, S. (Participant), Sjöblom, E. I. (Participant), Vahtola, T. (Participant), Itkonen, S. (Participant) & Vahtola, T. (Participant)
   Academy of Finland, Suomen Akatemia Projektilaskutus
   01/01/2021 → 31/12/2023

6. **OPUS-MT: Open Translation Models, Tools and Services**
   Aulamo, M. (Participant), Nieminen, T. J. (Participant), Hardwick, S. (Participant) & Tiedemann, J. (Project manager)
   01/08/2020 → 31/08/2021

7. **EOSC-nordic: The European Open Science Cloud - Nordic**
   Tiedemann, J. (Project manager), Drobac, S. (Project manager) & Raganato, A. (Project manager)
   European Commission, European Commission – Joint Research Centre
   01/09/2019 → 31/08/2022

8. **FoTran: Found in Translation - Natural Language Understanding with Cross-Lingual Grounding**
   Tiedemann, J. (Project manager), Attieh, J. (Participant), Aulamo, M. (Participant), Celikkanat, H. (Participant), De Gibert Bonet, O. (Participant), Grönroos, S.-A. (Participant), Mickus, T. (Participant), Raganato, A. (Participant), Scherrer, Y. (Participant), Silfverberg, M. (Participant), Sjöblom, E. I. (Participant), Talman, A. (Participant), Vazquez , R. (Participant), Virpioja, S. P. (Participant), Yli-Jyrä, A. (Participant), Zosa, E. (Participant), Celikkanat, H. (Participant), Raganato, A. (Participant), Silfverberg, M. (Participant), Sulubacak, U. (Participant) & Vazquez , R. (Participant)
   01/09/2018 → 31/03/2024

9. **fiskmö: Creation of a parallel corpus of translated documents and machine translation for Finnish and Swedish**
   Tiedemann, J. (Project manager), Ginter, F. (Project manager), Papula, N. (Project manager), Aulamo, M. (Participant), Nieminen, T. (Participant), Kanerva, J. (Participant) & Eskola, K. (Participant)
   Svenska kulturfonden
   01/05/2018 → 31/03/2021

10. **MeMAD: Methods for Managing Audiovisual Data: Combining Automatic Efficiency with Human Accuracy**
    Hirvonen, M. (Principal Investigator), Tiedemann, J. (Participant), Tiittula, L. (Participant), Sulubacak, U. (Participant), Vazquez , R. (Participant) & Koponen, M. (Participant)
    European Commission / Horizon 2020
    01/01/2018 → 31/03/2021

11. **NLUxLG: NLU with Cross-Lingual Grounding**
    Tiedemann, J. (Project manager), Talman, A. (Participant), Raganato, A. (Participant) & Celikkanat, H. (Participant)
    SUOMEN AKATEMIA
    01/01/2018 → 31/12/2019

12. **Development of Digital Teaching project of the BA Programme in Languages**
    Lennes, M. (Participant), Flanagan, J. (Participant), Hiltunen, T. (Participant), Creutz, M. (Participant), Aunio, L. (Participant), Buchart, M. (Participant), Gruzdeva, E. (Participant), Jakubek, I. (Participant), Keinänen, N. (Participant), Kim, J.-Y. (Participant), Kittilä, S. (Participant), Kopotev, M. (Participant), Kraenker, S. (Participant), Länsisalmi, R. (Participant), Nieuweboer, R. (Participant), Peterson, E. (Participant), Tiedemann, J. (Participant), Tuori, R. (Participant), Autere, M. J. (Participant) & Lindstedt, J. (Project manager)
    01/05/2017 → 31/12/2018

13. **NLPL: Nordic Language Processing Laboratory**
Tiedemann, J. (Participant) & Scherrer, Y. (Participant)
01/01/2017 → 31/12/2019

14. **CrossNLP: Cross-Linguistic and Multilingual Natural Language Processing with the Focus on Low-Resource Languages and Language Variants**
Tiedemann, J. (Project manager), Östling, R. (Principal Investigator) & Scherrer, Y. (Participant)
01/01/2016 → 31/12/2018

15. **Erillisrahoitus Helsingin ja Tukholman yliopistojen yhteishanke**
Lehti-Eklund, H. (Project manager), Lindström, J. (Participant), Miestamo, M. (Participant), Grünthal, R. (Participant), Palander-Collin, M. (Participant), Kalliokoski, J. (Participant), Östman, J.-O. (Participant), Ajanki, R. (Participant), Holopainen, S. (Project manager), Balode, L. (Participant), Lenk, H. E. H. (Participant), Kurhila, S. (Participant), Kotilainen, L. (Participant), Lindholm, C. (Participant), Wahlström, M. (Project manager), Solin, A. (Project manager), Fremer, M. (Project manager), Tiittula, L. (Project manager), Hamari, A. (Project manager), Kittilä, S. (Project manager), Huhtamäki, M. (Project manager), Kim, J.-Y. (Project manager), Kopotev, M. (Project manager), Peterson, E. (Project manager), Protassova, E. (Project manager), Savijärvi, M. (Project manager) & Tiedemann, J. (Project manager)
Unknown funder
01/01/2015 → 31/12/2017

16. **eSSENCE: Efficient Algorithms for Natural Language Processing Beyond Sentence Boundaries**
Nivre, J. (Project manager), Tiedemann, J. (Project manager) & Stymne, S. (Principal Investigator)
01/09/2012 → 31/08/2014

17. **DiscoMT: Discourse-Oriented Statistical Machine Translation**
Tiedemann, J. (Project manager), Hardmeier, C. (Principal Investigator), Loáiciga, S. (Participant) & Scherrer, Y. (Participant)
01/01/2012 → 31/12/2017

18. **LetsMT!: LetsMT! A Platform for Online Sharing of Training Data and Building User Tailored Machine Translation**
Tiedemann, J. (Project manager), Weinitz, P. (Principal Investigator), Saers, M. (Principal Investigator), Zumpe, M. (Principal Investigator) & Schleussner, S. (Principal Investigator)
01/03/2010 → 31/08/2012

19. **OPUS: The Open Parallel Corpus**
Tiedemann, J. (Project manager)
01/06/2004 → …

Datasets

1. **Document-level Machine Translation Benchmark**
Scherrer, Y. (Creator), Tiedemann, J. (Creator) & Loáiciga, S. (Creator), University of Helsinki, 1 Nov 2019
DOI: 10.5281/zenodo.3525366
Dataset

2. **LetsMT repository**
Tiedemann, J. (Creator), Zenodo, 22 Jan 2018
DOI: 10.5281/zenodo.1157836
Dataset

3. **Uplug**
Tiedemann, J. (Creator), Zenodo, 22 Jan 2018
DOI: 10.5281/zenodo.1157180, https://github.com/Helsinki-NLP/Uplug/tree/v0.3.9
Dataset

4. **Lingua-Align**
Tiedemann, J. (Creator), Zenodo, 22 Jan 2018
DOI: 10.5281/zenodo.1157176, https://github.com/Helsinki-NLP/Lingua-Align/tree/v0.1
Dataset

5. **OPUS**
Tiedemann, J. (Creator), University of Helsinki, 2017
http://opus.nlpl.eu
Dataset

6. **WMT17 test set Finnish-English**
Tiedemann, J. (Creator), Zenodo, 21 Jan 2018
DOI: 10.5281/zenodo.1156326, https://github.com/Helsinki-NLP/WMT16-test-enfi/tree/v1.0
Dataset

7. **WMT16 test set Finnish-English**
Tiedemann, J. (Creator), Zenodo, 21 Jan 2018
DOI: 10.5281/zenodo.1156325, https://github.com/Helsinki-NLP/WMT16-test-enfi/tree/v1.0
Dataset

8. **Shared Task in Cross-Lingual Parsing**
   Tiedemann, J. (Data Manager), Atlassian, 2017
   https://bitbucket.org/hy-crossNLP/vardial2017
   Dataset

9. **DiscoMT 2015 Shared Task on Pronoun Translation**
   Tiedemann, J. (Creator), LINDAT/CLARIN, 31 Jan 2016
   DOI: http://hdl.handle.net/11372/LRT-1611, http://hdl.handle.net/11372/LRT-1611
   Dataset

Publications

1. Li, Z., Ji, S., Mickus, T., Segonne, V., & Tiedemann, J. (2024). A Comparison of Language Modeling and Translation as Multilingual Pretraining Objectives. In Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (Eds.), *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 15882-15894). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.888
2. Scherrer, Y., Jauhiainen, T., Ljubešić, N., Zampieri, M., Nakov, P., & Tiedemann, J. (Eds.) (2024). *Proceedings of the Eleventh Workshop on NLP for Similar Languages, Varieties, and Dialects (VarDial 2024)*. The Association for Computational Linguistics. https://aclanthology.org/2024.vardial-1.0
3. Mickus, T., Zosa, E., Vazquez, R., Vahtola, T., Tiedemann, J., Segonne, V., Raganato, A., & Apidianaki, M. (2024). SemEval-2024 Task 6: SHROOM, a Shared-task on Hallucinations and Related Observable Overgeneration Mistakes. In A. K. Ojha, A. S. Doğruöz, H. Tayyar Madabushi, G. Da San Martino, S. Rosenthal, & A. Rosá (Eds.), *Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)* (pp. 1979-1993). The Association for Computational Linguistics. https://aclanthology.org/2024.semeval-1.273
4. Dmitrieva, A., & Tiedemann, J. (2024). Towards Automatic Finnish Text Simplification. In G. M. Di Nunzio, F. Vezzani, L. Ermakova, H. Azarbonyad, & J. Kamps (Eds.), *Proceedings of the Workshop on DeTermIt! Evaluating Text Difficulty in a Multilingual Context @ LREC-COLING 2024* (pp. 39–50). (International conference on computational linguistics), (LREC proceedings). European Language Resources Association (ELRA). https://aclanthology.org/2024.determit-1.4/
5. Mickus, T., Grönroos, S.-A., Attieh, J., Boggia, M., De Gibert, O., Ji, S., Loppi, N. A., Raganato, A., Vázquez, R., & Tiedemann, J. (2024). MAMMOTH: Massively Multilingual Modular Open Translation @ Helsinki. In N. Aletras, & O. De Clercq (Eds.), *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations* (pp. 127-136). The Association for Computational Linguistics. https://aclanthology.org/2024.eacl-demo.14
6. Vázquez, R., Mickus, T., Tiedemann, J., Vulić, I., & Üstün, A. (Eds.) (2024). *Proceedings of the 1st Workshop on Modular and Open Multilingual NLP (MOOMIN 2024)*. The Association for Computational Linguistics. https://aclanthology.org/2024.moomin-1.0
7. de Gibert, O., Nail, G., Arefyev, N., Bañón, M., van der Linde, J., Ji, S., Zaragoza-Bernabeu, J., Aulamo, M., Ramírez-Sánchez, G., Kutuzov, A., Pyysalo, S., Oepen, S., & Tiedemann, J. (2024). A New Massive Multilingual Dataset for High-Performance Language Technologies. In N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, & N. Xue (Eds.), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 1116-1128). (International conference on computational linguistics), (LREC proceedings). European Language Resources Association (ELRA).
8. Ji, S., Mickus, T., Segonne, V., & Tiedemann, J. (2024). Can Machine Translation Bridge Multilingual Pretraining and Cross-lingual Transfer Learning? In N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, & N. Xue (Eds.), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 2809-2818). (International conference on computational linguistics), (LREC proceedings). European Language Resources Association (ELRA).
9. De Gibert Bonet, O., Aulamo, M., Scherrer, Y., & Tiedemann, J. (2024). Hybrid distillation from RBMT and NMT: Helsinki-NLP's submission to the Shared Task on Translation into Low-Resource Languages of Spain. In B. Haddow, T. Kocmi, P. Koehn, & C. Monz (Eds.), *Proceedings of the Ninth Conference on Machine Translation* (pp. 908-917). The Association for Computational Linguistics.
10. Tiedemann, J., Aulamo, M., Bakshandaeva, D., Boggia, M., Grönroos, S. A., Nieminen, T., Raganato, A., Scherrer, Y., Vázquez, R., & Virpioja, S. (2023). Democratizing neural machine translation with OPUS-MT. *Language Resources and Evaluation*. https://doi.org/10.1007/s10579-023-09704-w
11. De Gibert Bonet, O., Vázquez, R., Aulamo, M., Scherrer, Y., Virpioja, S., & Tiedemann, J. (2023). Four Approaches to Low-Resource Multilingual NMT: The Helsinki Submission to the AmericasNLP 2023 Shared Task. In M. Mager, A. Ebrahimi, & A. Oncevay, et al. (Eds.), *Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP)* (pp. 177-191). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.americasnlp-1.20
12. Vahtola, T., Creutz, M., & Tiedemann, J. (2023). Guiding Zero-Shot Paraphrase Generation with Fine-Grained Control Tokens. In A. Palmer, & J. Camacho-collados (Eds.), *Proceedings of the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023)* (pp. 323-337). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.starsem-1.29

13. Aulamo, M., De Gibert Bonet, O., Virpioja, S., & Tiedemann, J. (2023). Unsupervised Feature Selection for Effective Parallel Corpus Filtering. In M. Nurminen, J. Brenner, & M. Koponen, et al. (Eds.), *Proceedings of the 24th Annual Conference of the European Association for Machine Translation* (pp. 31-38). European Association for Machine Translation.
14. Scherrer, Y., Jauhiainen, T., Ljubešić, N., Nakov, P., Tiedemann, J., & Zampieri, M. (Eds.) (2023). *Tenth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2023): Proceedings of the Workshop*. The Association for Computational Linguistics. https://aclanthology.org/2023.vardial-1.0
15. Boggia, M., Grönroos, S.-A., Loppi, N., Mickus, T., Raganato, A., Tiedemann, J., & Vazquez , R. (2023). Dozens of Translation Directions or Millions of Shared Parameters? Comparing Two Types of Multilinguality in Modular Machine Translation. In T. Alumäe , & M. Fishel (Eds.), *Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)* (pp. 238–247). (NEALT Proceedings Series Publisher name; No. 52). University of Tartu Library.
16. Tuominen, T., Koponen, M., Vitikainen, K., Sulubacak, U., & Tiedemann, J. (2023). Exploring the gaps in linguistic accessibility of media: The potential of automated subtitling as a solution. *The Journal of Specialised Translation*, (39), 77-98.
17. Aulamo, M., Bogoychev, N., Ji, S., Nail, G., Ramirez-Sánchez, G., Tiedemann, J., Van Der Linde, J., & Zaragoza, J. (2023). HPLT: High Performance Language Technologies. In M. Nurminen, J. Brenner, & M. Koponen, et al. (Eds.), *Proceedings of the 24th Annual Conference of the European Association for Machine Translation* (pp. 517-518). European Association for Machine Translation.
18. Tiedemann, J., & Gibert, O. D. (2023). The OPUS-MT dashboard - A toolkit for a systematic evaluation of open machine translation models. In D. Bollegala, R. Huang, & A. Ritter (Eds.), *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics: Volume 3: System Demonstrations* (Vol. 3, pp. 315-327). Association for Computational Linguistics (ACL).
19. Talman, A., Celikkanat, H., Virpioja, S., Heinonen, M., & Tiedemann, J. (2023). Uncertainty-Aware Natural Language Inference with Stochastic Weight Averaging. In T. Alumäe , & M. Fishel (Eds.), *Proceedings of the 24th Nordic Conference on Computational Linguistics* (pp. 358-365). (NEALT Proceedings Series; No. 52). University of Tartu Library.
20. Vahtola, T., Creutz, M., & Tiedemann, J. (2022). It Is Not Easy To Detect Paraphrases: Analysing Semantic Similarity With Antonyms and Negation Using the New SemAntoNeg Benchmark. In J. Bastings, Y. Belinkov, Y. Elazar, D. Hupkes, N. Saphra, & S. Wiegreffe (Eds.), *Proceedings of the Fifth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP* (pp. 249–262). The Association for Computational Linguistics.
21. Tiedemann, J. (2022). From open parallel corpora to public translation tools: The success story of OPUS. In E. Volodina, D. Dannélls, A. Berdicevskis, M. Forsberg, & S. Virk (Eds.), *LIVE and LEARN : Festschrift in honor of Lars Borin* (pp. 133-138). (Research Reports from the Department of Swedish, Multilingualism, Language Technology; No. GU-ISS-2022-03). University of Göteborg.
22. Tiedemann, J., Aulamo, M., Hardwick, S., & Nieminen, T. (2022). Open Translation Models, Tools and Services. In G. Rehm (Ed.), *European Language Grid: A Language Technology Platform for Multilingual Europe* (pp. 325-330). (Cognitive Technologies). Springer . https://doi.org/10.1007/978-3-031-17258-8_24
23. Vazquez , R., Celikkanat, H., Ravishankar, V., Creutz, M., & Tiedemann, J. (2022). A Closer Look at Parameter Contributions When Training Neural Language and Translation Models. In N. Calzolari, C.-R. Huang, & H. Kim, et al. (Eds.), *Proceedings of the 29th International Conference on Computational Linguistics* (pp. 4788-4800). (International conference on computational linguistics; Vol. 29, No. 1). International Committee on Computational Linguistics.
24. Scherrer, Y., Jauhiainen, T., Ljubešić, N., Nakov, P., Tiedemann, J., & Zampieri, M. (Eds.) (2022). *Proceedings of the Ninth Workshop on NLP for Similar Languages, Varieties and Dialects: The 29th International Conference on Computational Linguistics*. (International conference on computational linguistics ; Vol. 29, No. 19). COLING . https://aclanthology.org/2022.vardial-1.0
25. Itkonen, S., Tiedemann, J., & Creutz, M. (2022). Helsinki-NLP at SemEval-2022 Task 2: A Feature-Based Approach to Multilingual Idiomaticity Detection. In G. Emerson, N. Schluter, G. Stanovsky, R. Kumar, A. Palmer, N. Schneider, S. Singh, & S. Ratan (Eds.), *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)* (pp. 122-134). The Association for Computational Linguistics.
26. Vahtola, T., Sjöblom, E., Tiedemann, J., & Creutz, M. (2022). Modeling Noise in Paraphrase Detection. In N. Calzolari, F. Béchet, & P. Blache, et al. (Eds.), *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)* (pp. 4324-4332). European Language Resources Association (ELRA).
27. Tiedemann, J. (2022). Språk(teknologi) är nyckeln till intelligens och rättvisa. *Språkbruk*. https://www.sprakbruk.fi/-/spr%C3%A5k-teknologi-%C3%A4r-nyckeln-till-intelligens-och-r%C3%A4ttvisa
28. Talman, A., Apidianaki, M., Chatzikyriakidis, S., & Tiedemann, J. (2022). How Does Data Corruption Affect Natural Language Understanding Models? A Study on GLUE datasets. In V. Nastase, E. Pavlick, M. Taher Pilehvar, J. Camacho-Collados, & A. Raganato (Eds.), *Proceedings of The 11th Joint Conference on Lexical and Computational Semantics* (pp. 226-233). The Association for Computational Linguistics.

29. Vazquez , R., Boggia, M., Raganato, A., Loppi, N. A., Grönroos, S.-A., & Tiedemann, J. (2022). Latest Development in the FoTran Project – Scaling Up Language Coverage in Neural Machine Translation Using Distributed Training with Language-Specific Components. In H. Moniz, L. Macken, & A. Rufener, et al. (Eds.), *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation* (pp. 311-312). European Association for Machine Translation. https://aclanthology.org/2022.eamt-1.45

30. Alnajjar, K., Hämäläinen, M., Tiedemann, J., Laaksonen, J., & Kurimo, M. (2022). When to Laugh and How Hard? A Multimodal Approach to Detecting Humor and Its Intensity. In N. Calzolari, C.-R. Huang , & H. Kim, et al. (Eds.), *Proceedings of the 29th International Conference on Computational Linguistics* (pp. 6875-6886). (International Conference on Computational Linguistics; Vol. 29, No. 1). International Committee on Computational Linguistics. https://aclanthology.org/2022.coling-1.598

31. Raganato, A., Vázquez, R., Creutz, M., & Tiedemann, J. (2021). An Empirical Investigation of Word Alignment Supervision for Zero-Shot Multilingual Neural Machine Translation. In M.-F. Moens, X. Huang, L. Specia, & S. Wen-tau Yih (Eds.), *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 8449-8456). The Association for Computational Linguistics. https://aclanthology.org/2021.emnlp-main.664.pdf

32. Nakov, P., & Tiedemann, J. (2021). Machine Translation between Similar Languages. In M. Zampieri, & P. Nakov (Eds.), *Similar Languages, Varieties, and Dialects: A Computational Perspective* (pp. 219–253). (Studies in Natural Language Processing). Cambridge University Press. https://doi.org/10.1017/9781108565080.015

33. Vazquez , R., Celikkanat, H., Creutz, M., & Tiedemann, J. (2021). On the differences between BERT and MT encoder spaces and how to address them in translation tasks. In J. Kabbara, H. Lin, A. Paullada, & J. Vamvas (Eds.), *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Student Research Workshop* (pp. 337-347). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-srw.35

34. Vázquez, R., Scherrer, Y., Virpioja, S., & Tiedemann, J. (2021). The Helsinki submission to the AmericasNLP shared task. In M. Mager [et al.] (Ed.), *Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas* (pp. 255-264). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.americasnlp-1.29

35. Aulamo, M., Virpioja, S., Scherrer, Y., & Tiedemann, J. (2021). Boosting Neural Machine Translation from Finnish to Northern Sámi with Rule-Based Backtranslation. In S. Dobnik, & L. Øvrelid (Eds.), *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)* (pp. 351-356). (Linköping Electronic Conference Proceedings ; No. 78), (NEALT Proceedings Series ; No. 45). Linköping University Electronic Press. https://www.aclweb.org/anthology/2021.nodalida-main.37.pdf

36. Dmitrieva, A., & Tiedemann, J. (2021). Creating an Aligned Russian Text Simplification Dataset from Language Learner Data. In B. Babych [et al.] (Ed.), *Proceedings of the 8th Workshop on Balto-Slavic Natural Language Processing* (pp. 73-79). ACL Anthology. https://aclanthology.org/2021.bsnlp-1.8

37. Dmitrieva, A., & Tiedemann, J. (2021). A Multi-task Learning Approach to Text Simplification. In W. M. P. van der Aalst [et al.] (Ed.), *Recent Trends in Analysis of Images, Social Networks and Texts: 9th International Conference, AIST 2020, Skolkovo, Moscow, Russia, October 15–16, 2020 Revised Supplementary Proceedings* (pp. 78-89). (Communications in Computer and Information Science; Vol. 1357). Springer. https://doi.org/10.1007/978-3-030-71214-3_7

38. Tsarfaty, R., & Tiedemann, J. (2021). Message from the program chairs. In *EACL 2021. The 16th Conference of the European Chapter of the Association for Computational Linguistics: Proceedings of the Conference* (pp. vi-viii). (Proceedings of the European Association for Computational Linguistics). The Association for Computational Linguistics. https://aclanthology.org/2021.eacl-main.0.pdf

39. Talman, A., Apidianaki, M., Chatzikyriakidis, S., & Tiedemann, J. (2021). NLI Data Sanity Check: Assessing the Effect of Data Corruption on Model Performance. In S. Dobnik , & L. Øvrelid (Eds.), *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa 2021)* (pp. 276-285). (Linköping Electronic Conference Proceedings ; No. 78), (NEALT Proceedings Series ; No. 45). Linköping University Electronic Press.

40. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Scherrer, Y., & Jauhiainen, T. (Eds.) (2021). *Proceedings of the 8th VarDial Workshop on NLP for Similar Languages, Varieties and Dialects*. The Association for Computational Linguistics. https://aclanthology.org/2021.vardial-1

41. Tiedemann, J. (2021). The Development of a Comprehensive Data Set for Systematic Studies of Machine Translation. In M. Hämäläinen, N. Partanen, & K. Alnajjar (Eds.), *Multilingual Facilitation* (pp. 248-262). University of Helsinki. https://doi.org/10.31885/9789515150257.22

42. Siewert, J., Scherrer, Y., & Tiedemann, J. (2021). Towards a balanced annotated Low Saxon dataset for diachronic investigation of dialectal variation. In K. Evang, L. Kallmeyer, R. Osswald, J. Waszczuk, & T. Zesch (Eds.), *Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021)* (pp. 242-246). KONVENS 2021 Organizers. https://aclanthology.org/2021.konvens-1.25.pdf

43. Siewert, J., Scherrer, Y., Wieling, M., & Tiedemann, J. (2020). LSDC - A comprehensive dataset for Low Saxon Dialect Classification. In M. Zampieri, P. Nakov, N. Ljubešić, J. Tiedemann, & Y. Scherrer (Eds.), *Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects* (pp. 25-35). International Committee on Computational Linguistics (ICCL).

44. Celikkanat, H., Virpioja, S., Tiedemann, J., & Apidianaki, M. (2020). Tracking the Traces of Passivization and Negation in Contextualized Representations. In A. Alishahi, Y. Belinkov, G. Chrupała, D. Hupkes, Y. Pinter, & H. Sajjad (Eds.), *Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP* (pp. 136-148). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.blackboxnlp-1.13

45. Raganato, A., Scherrer, Y., & Tiedemann, J. (2020). Fixed Encoder Self-Attention Patterns in Transformer-Based Machine Translation. In T. Cohn, Y. H. He, & Y. Liu (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 556-568). The Association for Computational Linguistics. https://www.aclweb.org/anthology/2020.findings-emnlp.49

46. Tiedemann, J., & Thottingal, S. (2020). OPUS-MT -- Building open translation services for the World. In A. Martins [et al.] (Ed.), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (pp. 479-480). European Association for Machine Translation.

47. Scherrer, Y., Raganato, A., & Tiedemann, J. (2020). The MUCOW word sense disambiguation test suite at WMT 2020. In L. Barrault [et al.] (Ed.), *Proceedings of the Fifth Conference on Machine Translation* (pp. 365-370). The Association for Computational Linguistics.

48. Tiedemann, J. (2020). The Tatoeba Translation Challenge - Realistic Data Sets for Low Resource and Multilingual MT. In L. Barrault [et al.] (Ed.), *Proceedings of the Fifth Conference on Machine Translation* (pp. 1174-1182). The Association for Computational Linguistics.

49. Koponen, M., Sulubacak, U., Vitikainen, K., & Tiedemann, J. (2020). MT for Subtitling: Investigating professional translators' user experience and feedback. In J. E. Ortega, M. Federico, C. Orasan, & M. Popovic (Eds.), *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas October 6 - 9, 2020: 1st Workshop on Post-Editing in Modern-Day Translation* (pp. 79-92). AMTA. https://www.aclweb.org/anthology/2020.amta-pemdt.6/

50. Sulubacak, U., Caglayan, O., Grönroos, S.-A., Rouhe, A., Elliott, D., Specia, L., & Tiedemann, J. (2020). Multimodal Machine Translation through Visuals and Speech. *Machine Translation*, *34*(2-3), 97-147. Article 2. https://doi.org/10.1007/s10590-020-09250-0

51. Vázquez , R., Aulamo, M., Sulubacak, U., & Tiedemann, J. (2020). The University of Helsinki Submission to the IWSLT2020 Offline Speech Translation Task. In *Proceedings of the 17th International Conference on Spoken Language Translation (IWSLT)* (pp. 95-102). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.iwslt-1.10

52. Aulamo, M., Virpioja, S., & Tiedemann, J. (2020). OpusFilter: A Configurable Parallel Corpus Filtering Toolbox. In A. Çelikyilmaz, & T.-H. Wen (Eds.), *58TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL 2020): SYSTEM DEMONSTRATIONS: System Demonstrations* (pp. 150-156). The Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-demos.20

53. Koponen, M., Sulubacak, U., Vitikainen, K., & Tiedemann, J. (2020). MT for subtitling: User evaluation of post-editing productivity. In A. Martins, H. Moniz, S. Fumega, B. Martins, F. Batista, L. Coheur, C. Parra, I. Trancoso, M. Turchi, A. Bisazza, J. Moorkens, A. Guerberof, M. Nurminen, L. Marg, & M. L. Forcada (Eds.), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation (EAMT 2020)* (pp. 115-124). European Association for Machine Translation.

54. Vazquez, R., Raganato, A., Creutz, M., & Tiedemann, J. (2020). A Systematic Study of Inner-Attention-Based Sentence Representations in Multilingual Neural Machine Translation. *Computational Linguistics*, *46*(2), 387-424. https://doi.org/10.1162/coli_a_00377

55. Aulamo, M., Sulubacak, U., Virpioja, S., & Tiedemann, J. (2020). OpusTools and Parallel Corpus Diagnostics. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)* (pp. 3782-3789). European Language Resources Association (ELRA).

56. Raganato, A., Scherrer, Y., & Tiedemann, J. (2020). An Evaluation Benchmark for Testing the Word Sense Disambiguation Capabilities of Machine Translation Systems. In N. Calzolari [et al.] (Ed.), *Proceedings of The 12th Language Resources and Evaluation Conference* (pp. 3668-3675). European Language Resources Association (ELRA). https://www.aclweb.org/anthology/2020.lrec-1.452.pdf

57. Tiedemann, J., Nieminen, T., Aulamo, M., Kanerva, J., Leino, A., Ginter, F., & Papula, N. (2020). The FISKMÖ Project: Resources and Tools for Finnish-Swedish Machine Translation and Cross-Linguistic Research. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 3808-3815). European Language Resources Association (ELRA). https://www.aclweb.org/anthology/2020.lrec-1.470

58. Mareček, D., Celikkanat, H., Silfverberg, M., Ravishankar, V., & Tiedemann, J. (2020). Are Multilingual Neural Machine Translation Models Better at Capturing Linguistic Features? *The Prague Bulletin of Mathematical Linguistics*, (115), 143-162. https://doi.org/10.14712/00326585.009

59. Kajava, K., Öhman, E., Hui, P., & Tiedemann, J. (2020). Emotion Preservation in Translation: Evaluating Datasets for Annotation Projection. In S. Reinsone, I. Skadiņa , A. Baklāne , & J. Daugavietis (Eds.), *Proceedings of Digital Humanities in Nordic Countries (DHN 2020)* (pp. 38-50). (CEUR workshop proceedings; No. 2612). CEUR.

60. Pàmies, M., Öhman, E., Kajava, K., & Tiedemann, J. (2020). LT@Helsinki at SemEval-2020 Task 12: Multilingual or language-specific BERT? In A. Herbelot, X. Zhu, A. Palmer, N. Schneider, J. May, & E. Shutova (Eds.), *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (pp. 1569-1575). International Committee for Computational Linguistics.

61. Öhman, E., Pàmies, M., Kajava, K., & Tiedemann, J. (2020). XED: A Multilingual Dataset for Sentiment Analysis and Emotion Detection. In D. Scott, N. Bel, & C. Zong (Eds.), *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 6542–6552). International Committee on Computational Linguistics. https://doi.org/10.18653/v1/2020.coling-main.575

62. Scherrer, Y., Tiedemann, J., & Loáiciga, S. (2019). Analysing concatenation approaches to document-level NMT in two different domains. In *The Fourth Workshop on Discourse in Machine Translation: Proceedings of the Workshop* (pp. 51-61). The Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-6506

63. Talman, A., Suni, A., Celikkanat, H., Kakouros, S., Tiedemann, J., & Vainio, M. (2019). Predicting Prosodic Prominence from Text with Pre-trained Contextualized Word Representations. In M. Hartmann, & B. Plank (Eds.), *22nd Nordic Conference on Computational Linguistics (NoDaLiDa): Proceedings of the Conference* (pp. 281–290). (Linköping Electronic Conference Proceedings; No. 167), (NEALT Proceedings Series; No. 42). Linköping University Electronic Press.

64. Raganato, A., Vázquez, R., Creutz, M., & Tiedemann, J. (2019). An Evaluation of Language-Agnostic Inner-Attention-Based Representations in Machine Translation. In I. Augenstein, S. Gella, S. Ruder, K. Kann, B. Can, J. Welbl, A. Conneau, X. Ren, & M. Rei (Eds.), *The 4th Workshop on Representation Learning for NLP (RepL4NLP-2019): Proceedings of the Workshop* (pp. 27-32). The Association for Computational Linguistics. https://www.aclweb.org/anthology/W19-4304

65. Raganato, A., Scherrer, Y., & Tiedemann, J. (2019). The MuCoW Test Suite at WMT 2019: Automatically Harvested Multilingual Contrastive Word Sense Disambiguation Test Sets for Machine Translation. In O. Bojar, R. Chatterjee, C. Federmann, & E. A. (Eds.), *Fourth Conference on Machine Translation : Proceedings of the Conference (Volume 2: Shared Task Papers, Day 1)* (pp. 470-480). The Association for Computational Linguistics. https://www.aclweb.org/anthology/W19-5354

66. Talman, A., Sulubacak, U., Vazquez , R., Scherrer, Y., Virpioja, S., Raganato, A., Hurskainen, A., & Tiedemann, J. (2019). The University of Helsinki submissions to the WMT19 news translation task. In *Fourth Conference of Conference on Machine Translation (WMT 2019): Proceedings of the Conference: Volume 2* (pp. 412-423). The Association for Computational Linguistics. http://www.statmt.org/wmt19/papers.html

67. Talman, A. J., Yli-Jyrä, A., & Tiedemann, J. (2019). Sentence Embeddings in NLI with Iterative Refinement Encoders. *Natural Language Engineering*, *25*(4), 467-482. https://doi.org/10.1017/s1351324919000202

68. Vazquez , R., Sulubacak, U., & Tiedemann, J. (2019). The University of Helsinki submission to the WMT19 Parallel Corpus Filtering Task. In O. Bojar, R. Chatterjee, C. Federmann, & E. A. (Eds.), *Fourth Conference on Machine Translation: Proceedings of the Conference: Volume 3: Shared Task Papers, Day 2* (pp. 294-300). The Association for Computational Linguistics.

69. Tiedemann, J., & Scherrer, Y. (2019). Measuring Semantic Abstraction of Multilingual NMT with Paraphrase Recognition and Generation Tasks. In A. Rogers, A. Drozd, A. Rumshisky, & Y. Goldberg (Eds.), *Proceedings of the 3rd Workshop on Evaluating Vector Space Representations for NLP* (pp. 35-42). The Association for Computational Linguistics. https://www.aclweb.org/anthology/W19-2005

70. Vazquez Carrillo, J. R., Raganato, A., Tiedemann, J., & Creutz, M. (2019). Multilingual NMT with a language-independent attention bridge. In I. Augenstein, S. Gella, S. Ruder, K. Kann, B. Can, J. Welbl, A. Conneau, X. Ren, & M. Rei (Eds.), *The 4th Workshop on Representation Learning for NLP (RepL4NLP-2019): Proceedings of the Workshop* (pp. 33-39). The Association for Computational Linguistics.

71. Lison, P., Tiedemann, J., & Kouylekov, M. (2019). Open subtitles 2018: Statistical rescoring of sentence alignments in large, noisy parallel corpora. In N. Calzolari ... [et al.] (Ed.), *LREC 2018, Eleventh International Conference on Language Resources and Evaluation* (pp. 1742-1748). European Language Resources Association (ELRA).

72. Hämäläinen, M., Säily, T., Rueter, J., Tiedemann, J., & Mäkelä, E. (2019). Revisiting NMT for normalization of early English letters. In B. Alex, S. Degaetano-Ortlieb, A. Kazantseva, N. Reiter, & S. Szpakowicz (Eds.), *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 71–75). (ACL Anthology; No. W19-25). The Association for Computational Linguistics. https://doi.org/10.18653/v1/W19-2509

73. Aulamo, M., & Tiedemann, J. (2019). The OPUS Resource Repository: An Open Package for Creating Parallel Corpora and Machine Translation Services. In M. Hartmann, & B. Plank (Eds.), *22nd Nordic Conference on Computational Linguistics (NoDaLiDa): Proceedings of the Conference* (pp. 389–394). (Linköping Electronic Conference Proceedings; No. 167), (NEALT Proceedings Series; No. 42). Linköping University Electronic Press.

74. Bjerva, J., Östling, R., Han Veiga, M., Tiedemann, J., & Augenstein, I. (2019). What do Language Representations Really Represent? *Computational Linguistics*, *45*(2), 381-389. https://doi.org/10.1162/coli_a_00351

75. Grönroos, S.-A., Huet, B., Kurimo, M., Laaksonen, J., Merialdo, B., Pham, P., Sjöberg, M., Sulubacak, U., Tiedemann, J., Troncy, R., & Vázquez Carrillo, J. R. (2018). The MeMAD Submission to the WMT18 Multimodal Translation Task. In O. Bojar, R. Chatterjee, C. Federmann, M. Fishel, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, C. Monz, M. Negri, A. Névéol, M. Neves, M. Post, L. Specia, M. Turchi, & K. Verspoor (Eds.), *Proceedings of the Third Conference on Machine Translation (WMT): Shared Task Papers* (pp. 603-611). The Association for Computational Linguistics. https://doi.org/10.18653/v1/W18-6439

76. Öhman, E. S., Tiedemann, J., Honkela, T. U., & Kajava, K. S. A. (2018). Creating a Dataset for Multilingual Fine-grained Emotion-detection Using Gamification-based Annotation. In A. Balahur, S. M. Mohammad, V. Hoste, & R. Klinger (Eds.), *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 24-30). The Association for Computational Linguistics. https://doi.org/10.18653/v1/W18-6205

77. Sulubacak, U., Tiedemann, J., Rouhe, A., Stig-Arne, G., & Kurimo, M. (2018). The MeMAD Submission to the IWSLT 2018 Speech Translation Task. In M. Turchi, J. Niehues, & M. Frederico (Eds.), *Proceedings of the 15th International Workshop on Spoken Language Translation (IWSLT 2018)* (pp. 89-94).

78. Ehrentraut, C., Ekholm, M., Tanushi, H., Tiedemann, J., & Dalianis, H. (2018). Detecting hospital-acquired infections: A document classification approach using support vector machines and gradient tree boosting. *Health informatics journal*., *24*(1), 24-42. https://doi.org/10.1177/1460458216656471

79. Raganato, A., & Tiedemann, J. (2018). An Analysis of Encoder Representations in Transformer-Based Machine Translation. In L. Tal, G. Chrupała , & A. Alishahi (Eds.), *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP* (pp. 287-297). The Association for Computational Linguistics. http://aclweb.org/anthology/W18-5431

80. Tiedemann, J. (2018). Emerging Language Spaces Learned From Massively Multilingual Corpora. In E. Mäkelä, M. Tolonen, & J. Tuominen (Eds.), *Proceedings of the Digital Humanities in the Nordic Countries 3rd Conference (DHN 2018)* (Vol. 2084, pp. 188-197). (CEUR Workshop Proceedings). CEUR Workshop Proceedings.

81. Zampieri, M., Malmasi, S., Nakov, P., Ali, A., Shon, S., Glass, J., Scherrer, Y., Samardžić, T., Ljubešić, N., Tiedemann, J., van der Lee, C., Grondelaers, S., Oostdijk, N., Speelman, D., van den Bosch, A., Kumar, R., Lahiri, B., & Jain, M. (2018). Language Identification and Morphosyntactic Tagging: The Second VarDial Evaluation Campaign. In M. Zampieri, P. Nakov, N. Ljubešić, J. Tiedemann, S. Malmasi , & A. Ali (Eds.), *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects* (pp. 1-17). The Association for Computational Linguistics. http://web.science.mq.edu.au/~smalmasi/vardial5/pdf/W18-391.pdf

82. Hämäläinen, M., Säily, T., Rueter, J., Tiedemann, J., & Mäkelä, E. (2018). Normalizing early English letters to Present-day English spelling. In B. Alex, S. Degaetano-Ortlieb, A. Feldman, A. Kazantseva, N. Reiter, & S. Szpakowicz (Eds.), *Proceedings of the 2nd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 87-96). (ACL Anthology; No. W18-45). The Association for Computational Linguistics. http://aclweb.org/anthology/W18-4510

83. Lison, P., Tiedemann, J., & Kouylekov, M. (2018). OpenSubtitles2018: Statistical Rescoring of Sentence Alignments in Large, Noisy Parallel Corpora. In N. Calzolari, C. Khalid, C. Christopher, D. Thierry, G. Sara, H. Koiti, I. Hitoshi, M. Bente, M. Joseph, M. Hélène, M. Asuncion, O. Jan, P. Stelios, & T. Takenobu (Eds.), *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)* (pp. 1742-1748). European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2018/summaries/294.html

84. Bozovic, P., Erjavec, T., Tiedemann, J., Ljubesic, N., & Gorjanc, V. (2018). Opus-MontenegrinSubs 1.0: First electronic corpus of the Montenegrin language. In D. Fišer, & A. Pančur (Eds.), *Proceedings of the conference on Language Technologies & Digital Humanities 2018* (pp. 24-28). Ljubljana University Press. http://nl.ijs.si/jtdh18/proceedings-en.html

85. Zampieri, M., Nakov, P., Ljubesic, N., Tiedemann, J., Malmasi, S., & Ali, A. (Eds.) (2018). *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*. The Association for Computational Linguistics. http://aclweb.org/anthology/W/W17/W17-12.pdf

86. Raganato, A., Scherrer, Y., Nieminen, T., Hurskainen, A., & Tiedemann, J. (2018). The University of Helsinki submissions to the WMT18 news task. In O. Bojar, R. Chatterjee, C. Federmann, M. Fishel, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, C. Monz, M. Negri, A. Névéol, M. Neves, M. Post, L. Specia, M. Turchi, & K. Verspoor (Eds.), *Proceedings of the Third Conference on Machine Translation: Shared Task Papers* (pp. 488-495). (WMT 2018 - 3rd Conference on Machine Translation, Proceedings of the Conference; Vol. 2). The Association for Computational Linguistics. https://doi.org/10.18653/v1/W18-64052

87. Shao, Y., Hardmeier, C., Tiedemann, J., & Nivre, J. (2017). Character-based Joint Segmentation and POS Tagging for Chinese using Bidirectional RNN-CRF. In G. Kondrak, & T. Watanabe (Eds.), *The Eighth International Joint Conference on Natural Language Processing: Proceedings of the Conference, Vol. 1 (Long Papers)* (pp. 173-183). Asian Federation of Natural Language Processing. http://aclweb.org/anthology/I/I17/I17-1018.pdf

88. Tiedemann, J. (Ed.) (2017). *Proceedings of the 21st Nordic Conference on Computational Linguistics (NoDaLiDa)*. (Linköping Electronic Conference Proceedings; No. 131). Linköping University Electronic Press. http://www.ep.liu.se/ecp/131/ecp17131.pdf

89. Östling, R., & Tiedemann, J. (2017). Continuous multilinguality with language vectors. In *15th Conference of the European Chapter of the Association for Computational Linguistics: Proceedings of Conference, volume 2: Short Papers* (pp. 644-649). The Association for Computational Linguistics. https://www.aclweb.org/anthology/E17-2102/

90. Tiedemann, J. (2017). Cross-Lingual Dependency Parsing for Closely Related Languages: Helsinki's Submission to VarDial 2017. In *Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial'2017): Proceedings of the Workshop* (pp. 131-136). The Association for Computational Linguistics. https://doi.org/10.18653/v1/w17-1216

91. Zampieri, M., Malmasi, S., Ljubešić, N., Nakov, P., Ali, A., Tiedemann, J., Scherrer, Y., & Aepli, N. (2017). Findings of the VarDial Evaluation Campaign 2017. In *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects* (pp. 1-15). The Association for Computational Linguistics. http://aclweb.org/anthology/W/W17/W17-1201.pdf

92. Tiedemann, J., & van der Plas, L. (2017). Bootstrapping a Dependency Parser for Maltese - A Real-World Test Case. In M. Wieling, M. Kroon, G. Van Noord, & G. Bouma (Eds.), *From Semantics to Dialectometry: Festschrift in honor of John Nerbonne* (pp. 355-365). (Tributes; No. 32). College publications. http://www.let.rug.nl/festschriftnerbonne/36.%20Tiedemann%20&%20van%20der%20Plas.pdf

93. Loáiciga, S., Stymne, S., Nakov, P., Hardmeier, C., Tiedemann, J., Cettolo, M., & Versley, Y. (2017). Findings of the 2017 DiscoMT Shared Task on Cross-lingual Pronoun Prediction. In *Discourse in Machine Translation (DiscoMT 2017): Proceedings of the Workshop* (pp. 1-16). The Association for Computational Linguistics. http://aclweb.org/anthology/W/W17/W17-4801.pdf

94. Kotzé, G., Vandeghinste, V., Martens, S., & Tiedemann, J. (2017). Large aligned treebanks for syntax-based machine translation. *Language Resources and Evaluation*, *51*(2), 249-282. https://doi.org/10.1007/s10579-016-9369-0

95. Tiedemann, J., & Scherrer, Y. (2017). Neural Machine Translation with Extended Context. In *Proceedings of the Third Workshop on Discourse in Machine Translation* (pp. 82-92). The Association for Computational Linguistics. https://doi.org/10.18653/v1/w17-4811

96. Hurskainen, A., & Tiedemann, J. (2017). Rule-based Machine Translation from English to Finnish. In *Proceedings of the Second Conference on Machine Translation (WMT2017)* (pp. 323-329). The Association for Computational Linguistics. https://doi.org/10.18653/v1/w17-4731

97. Tjong Kim Sang, E., Bollmann, M., Boschker, R., Casacuberta, F., Dietz, F., Dipper, S., Domingo, M., van der Goot, R., van Koppen, M., Ljubešić, N., Östling, R., Petran, F., Pettersson, E., Scherrer, Y., Schraagen, M., Sevens, L., Tiedemann, J., Vanallemeersch, T., & Zervanou, K. (2017). The CLIN27 Shared Task: Translating Historical Text to Contemporary Language for Improving Automatic Linguistic Annotation. *Computational Linguistics in the Netherlands*, *7*, 53-64. http://www.clinjournal.org/sites/clinjournal.org/files/04.clin27-shared-task.pdf

98. Östling, R., Scherrer, Y., Tiedemann, J., Tang, G., & Nieminen, T. (2017). The Helsinki Neural Machine Translation System. In O. Bojar, C. Buck, R. Chatterjee, C. Federmann, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, P. Koehn, & J. Kreutzer (Eds.), *Proceedings of the Second Conference on Machine Translation (WMT2017)* (pp. 338-347). (WMT 2017 - 2nd Conference on Machine Translation, Proceedings). The Association for Computational Linguistics. https://doi.org/10.18653/v1/w17-4733

99. Malmasi, S., Zampieri, M., Ljubešić, N., Nakov, P., Ali, A., & Tiedemann, J. (2016). Discriminating between Similar Languages and Arabic Dialect Identification: A Report on the Third DSL Shared Task. In *Third Workshop on NLP for Similar Languages, Varieties and Dialects: Proceedings of the Workshop* (pp. 1-14). The COLING 2016 Organizing Committee . http://web.science.mq.edu.au/~smalmasi/vardial3/pdf/VarDial301.pdf

100. Tiedemann, J., Nichols, J., & Sprouse, R. (2016). Tagging Ingush - Language Technology For Low-Resource Languages Using Resources From Linguistic Field Work. In *Language Technology Resources and Tools for Digital Humanities (LT4DH): Proceedings of the Workshop* (pp. 148-155). http://aclweb.org/anthology/W16-4020

101. Öhman, E., Honkela, T., & Tiedemann, J. (2016). The Challenges of Multi-dimensional Sentiment Analysis Across Languages. In *Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (PEOPLES): Proceedings of the Workshop* (pp. 138-142). The COLING 2016 Organizing Committee . https://peoples2016.github.io/peoples2016.pdf

102. Östling, R., & Tiedemann, J. (2016). Efficient word alignment with Markov Chain Monte Carlo. *The Prague Bulletin of Mathematical Linguistics*, *106*, 125-146. https://doi.org/10.1515/pralin-2016-0013

103. Tiedemann, J. (2016). A Linear Baseline Classifier for Cross-Lingual Pronoun Prediction. In *The 54th Annual Meeting of the Association for Computational Linguistics: Proceedings of the First Conference on Machine Translation (WMT)* (pp. 616-619). The Association for Computational Linguistics. https://doi.org/10.18653/v1/w16-2356

104. Guillou, L., Hardmeier, C., Nakov, P., Stymne, S., Tiedemann, J., Versley, Y., Cettolo, M., Webber, B., & Popescu-Belis, A. (2016). Findings of the 2016 WMT Shared Task on Cross-lingual Pronoun Prediction. In *The 54th Annual Meeting of the Association for Computational Linguistics: Proceedings of the First Conference on Machine Translation (WMT)* (pp. 525-542). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W16-2345

105. Tiedemann, J., Cap, F., Kanerva, J., Ginter, F., Stymne, S., Östling, R., & Weller-Di Marco, M. (2016). Phrase-Based SMT for Finnish with More Data, Better Models and Alternative Alignment and Translation Tools. In *The 54th Annual Meeting of the Association for Computational Linguistics: Proceedings of the First Conference on Machine Translation (WMT)* (pp. 391-398). The Association for Computational Linguistics. https://doi.org/10.18653/v1/w16-2326

106. Tiedemann, J., & Agi, Z. (2016). Synthetic Treebanking for Cross-Lingual Dependency Parsing. *Journal of Artificial Intelligence Research*, *55*, 209-248. https://doi.org/10.1613/jair.4785

107. Smith, A., Hardmeier, C., & Tiedemann, J. (2016). Climbing Mont BLEU: The Strange World of Reachable High-BLEU Translations. *Baltic Journal of Modern Computing*, *4*(2), 269–281.

108. Tiedemann, J. (2016). Finding Alternative Translations in a Large Corpus of Movie Subtitles. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC-2016)* (pp. 3518-3522) http://www.lrec-conf.org/proceedings/lrec2016/pdf/62_Paper.pdf

109. Lison, P., & Tiedemann, J. (2016). OpenSubtitles2015: Extracting Large Parallel Corpora from Movie and TV Subtitles. In N. Calzolari , K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC-2016) http://www.lrec-conf.org/proceedings/lrec2016/pdf/947_Paper.pdf*

110. Tiedemann, J. (2016). OPUS -- Parallel Corpora for Everyone. *Baltic Journal of Modern Computing*, 384. http://www.bjmc.lu.lv/fileadmin/user_upload/lu_portal/projekti/bjmc/Contents/4_2_28_Products.pdf

111. Tiedemann, J. (2015). Baseline Models for Pronoun Prediction and Pronoun-Aware Translation. In *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT)* (pp. 108-114). The Association for Computational Linguistics. http://aclweb.org/anthology/W15-2515

112. Tiedemann, J., Ginter, F., & Kanerva, J. (2015). Morphological Segmentation and OPUS for Finnish-English Machine Translation. In *Proceedings of the Tenth Workshop on Statistical Machine Translation* (pp. 177-183). The Association for Computational Linguistics. http://aclweb.org/anthology/W15-3021

113. Callin, J., Hardmeier, C., & Tiedemann, J. (2015). Part-of-Speech Driven Cross-Lingual Pronoun Prediction with Feed-Forward Neural Networks. In *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT)* (pp. 59-64). The Association for Computational Linguistics. http://aclweb.org/anthology/W15-2508

114. Hardmeier, C., Institute, Q. C., Stymne, S., Tiedemann, J., Versley, Y., & Cettolo, M. (2015). Pronoun-Focused MT and Cross-Lingual Pronoun Prediction: Findings of the 2015 DiscoMT Shared Task on Pronoun Translation. In *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT),* (pp. 1-16). The Association for Computational Linguistics. http://aclweb.org/anthology/W15-2501

115. Tiedemann, J. (2015). Cross-Lingual Dependency Parsing with Universal Dependencies and Predicted PoS Labels. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)* (pp. 340-349). Uppsala University. http://www.aclweb.org/anthology/W15-2137

116. Shao, Y., Tiedemann, J., & Nivre, J. (2015). Boosting English-Chinese Machine Transliteration via High Quality Alignment and Multilingual Resources. In X. Duan, R. E. Banchs, M. Zhangs, H. Li, & A. Kumara (Eds.), *Proceedings of the Fifth Named Entity Workshop, joint with 53rd ACL and the 7th IJCNLP* (pp. 56-60). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W15-3908

117. Tiedemann, J. (2015). Improving the Cross-Lingual Projection of Syntactic Dependencies. In B. Megyesi (Ed.), *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)* (pp. 191-199). (Linköping electronic conference proceedings; No. 109), (NEALT Proceedings Series). Linköping University Electronic Press. http://www.aclweb.org/anthology/W15-1824

118. Agić, Ž., Tiedemann, J., Merkler, D., Krek, S., Dobrovoljc, K., & Može, S. (2014). Cross-lingual Dependency Parsing of Related Languages with Rich Morphosyntactic Tagsets. In *Language Technology for Closely Related Languages and Language Variants (LT4CloseLang)* (pp. 13-24). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W14-4203

119. Martinez Garcia, E., Tiedemann, J., España-Bonet, C., & Màrquez, L. (2014). Word's Vector Representations meet Machine Translation. In D. Wu, M. Carpuat, X. Carreras, & E. M. Vecchi (Eds.), *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation* (pp. 132-134). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W14-4015

120. Zampieri, M., Tan, L., Ljubešić, N., & Tiedemann, J. (2014). A Report on the DSL Shared Task 2014. In *Proceedings of the First Workshop on Applying NLP Tools to Similar Languages, Varieties and Dialects* (pp. 58-67). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W14-5307

121. Tiedemann, J. (2014). Rediscovering Annotation Projection for Cross-Lingual Parser Induction. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*

122. Skadiņš, R., Tiedemann, J., Rozis, R., & Deksne, D. (2014). Billions of Parallel Words for Free: Building and Using the EU Bookshop Corpus. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC-2014)* (pp. 1850-1855) http://www.lrec-conf.org/proceedings/lrec2014/summaries/846.html

123. Zampieri, M., Ljubešić, N., Tiedemann, J., & Tan, L. (2014). Merging Comparable Data Sources for the Discrimination of Similar Languages: The DSL Corpus Collection. In *Proceedings of the 7th Workshop on Building and Using Comparable Corpora (BUCC)* (pp. 6-10)

124. Guillou, L., Hardmeier, C., Smith, A., Tiedemann, J., & Webber, B. (2014). ParCor 1.0: A Parallel Pronoun-Coreference Corpus to Support Statistical MT. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)* (pp. 3191-3198) http://www.lrec-conf.org/proceedings/lrec2014/summaries/298.html

125. Hardmeier, C., Stymne, S., Smith, A., Tiedemann, J., & Nivre, J. (2014). Anaphora Models and Reordering for Phrase-Based SMT. In *Proceedings of the Ninth Workshop on Statistical Machine Translation* (pp. 122-129). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W/W14/W14-3312

126. Stymne, S., Tiedemann, J., & Nivre, J. (2014). Estimating Word Alignment Quality for SMT Reordering Tasks. In *Proceedings of the Ninth Workshop on Statistical Machine Translation* (pp. 275-286). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W14-3334.pdf

127. Tiedemann, J. (2014). Improved Text Extraction from PDF Documents for Large-Scale Natural Language Processing. In A. Gelbukh (Ed.), *Computational Linguistics and Intelligent Text Processing* (Vol. 1, pp. 102-112). (Lecture Notes in Computer Science LNCS 8403). Springer.

128. Tiedemann, J., Agić, Ž., & Nivre, J. (2014). Treebank Translation for Cross-Lingual Parser Induction. In *Proceedings of the Eighteenth Conference on Computational Language Learning* (pp. 130-140). The Association for Computational Linguistics.

129. Tiedemann, J. (2013). Tools for lexicographic use of parallel and comparable corpora. In R. H. Gouws, U. Heid, W. Schweickard, & H. E. Wiegand (Eds.), *An International Encyclopedia of Lexicgraphy -- Supplementary Volume* (Vol. 4, pp. 1433-1444). (Handbücher zur Sprach- und Kommunikationswissenschaft). De Gruyter Mouton. http://www.degruyter.com/view/product/175228

130. Tiedemann, J., van der Plas, L., & Villada Moirón, B. (2013). Bitexts as Semantic Mirrors. In *Unknown host publication http://stp.lingfil.uu.se/~joerg/paper/bitext13.pdf*

131. Hardmeier, C., Tiedemann, J., & Nivre, J. (2013). Latent Anaphora Resolution for Cross-Lingual Pronoun Prediction. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, EMNLP 2013, 18-21 October 2013, Grand Hyatt Seattle, Seattle, Washington, USA, A meeting of SIGDAT, a Special Interest Group of the ACL* (pp. 380-391). The Association for Computational Linguistics.

132. Hardmeier, C., Stymne, S., Tiedemann, J., & Nivre, J. (2013). Docent: A Document-Level Decoder for Phrase-Based Statistical Machine Translation. In *Unknown host publication* (pp. 193-198) http://www.aclweb.org/anthology/P13-4033

133. Stymne, S., Hardmeier, C., Tiedemann, J., & Nivre, J. (2013). Feature Weight Optimization for Discourse-Level SMT. In *Unknown host publication* (pp. 60-69). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W13-3308

134. Webber, B., Popescu-Belis, A., Markert, K., & Tiedemann, J. (Eds.) (2013). *Proceedings of the Workshop on Discourse in Machine Translation*. The Association for Computational Linguistics. http://www.aclweb.org/anthology/W13-33

135. Stymne, S., Hardmeier, C., Tiedemann, J., & Nivre, J. (2013). Tunable Distortion Limits and Corpus Cleaning for SMT. In *Proceedings of the Eighth Workshop on Statistical Machine Translation* (pp. 225-231). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W13-2229

136. Pettersson, E., Megyesi, B. B., & Tiedemann, J. (2013). An SMT approach to automatic annotation of historical text. In *Unknown host publication*

137. Stymne, S., Tiedemann, J., Hardmeier, C., & Nivre, J. (2013). Statistical Machine Translation with Readability Constraints. In *Unknown host publication* (pp. 375-386)

138. Tiedemann, J., & Nakov, P. (2013). Analyzing the Use of Character-Level Translation with Sparse and Noisy Datasets. In *Unknown host publication* (pp. 676-684)

139. Tiedemann, J. (2013). Book review of Markus Dickinson, Chris Brew and Detmar Meurers: Language and Computers. *Machine Translation*, *27*, 309-312. https://doi.org/10.1007/s10590-013-9143-7

140. Tiedemann, J., & Ljubesi, N. (2012). Efficient Discrimination Between Closely Related Languages. In *Proceedings of COLING 2012: technical papers* (pp. 2619-2634). Indian Institute of Technology. http://www.aclweb.org/anthology/C12-1160

141. Nakov, P., & Tiedemann, J. (2012). Combining Word-Level and Character-Level Models for Machine Translation Between Closely-Related Languages. In *50th Annual Meeting of the Association for Computational Linguistics: Volume 2, Short papers : proceedings of the Conference : July 8 - 14, 2012, Jeju Island, Korea.* (pp. 301-305). The Association for Computational Linguistics. http://www.aclweb.org/anthology/P12-2059

142. Hardmeier, C., Nivre, J., & Tiedemann, J. (2012). Document-Wide Decoding for Phrase-Based Statistical Machine Translation. In *Unknown host publication* (pp. 1179-1190) http://www.aclweb.org/anthology/D12-1108

143. Vasijevs, A., Skadis, R., & Tiedemann, J. (2012). LetsMT! Cloud-Based Platform for Do-It-Yourself Machine Translation. In M. Zhang (Ed.), *50th Annual Meeting of the Association for Computational Linguistics: ACL 2012 : proceedings of the System Demonstrations :10 July 2012, Jeju Island, Korea* (pp. 43-48). The Association for Computational Linguistics. http://www.aclweb.org/anthology/P12-3008

144. Hardmeier, C., Nivre, J., & Tiedemann, J. (2012). Tree Kernels for Machine Translation Quality Estimation. In *Unknown host publication* (pp. 109-113) http://www.aclweb.org/anthology/W12-3112

145. Tiedemann, J., Hansen, D. H., Offersgaard, L., Olsen, S., & Zumpe, M. (2012). A Distributed Resource Repository for Cloud-Based Machine Translation. In N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (Eds.), *Unknown host publication* (pp. 2207-2213) http://www.lrec-conf.org/proceedings/lrec2012/pdf/457Paper.pdf

146. Kotzé, G., Vandeghinste, V., Martens, S., & Tiedemann, J. (2012). Large aligned treebanks for syntax-based machine translation. In N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (Eds.), *Unknown host publication* (pp. 467-473) http://www.lrec-conf.org/proceedings/lrec2012/pdf/924Paper.pdf

147. Tiedemann, J. (2012). Parallel Data, Tools and Interfaces in OPUS. In N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (Eds.), *Unknown host publication* (pp. 2214-2218) http://www.lrec-conf.org/proceedings/lrec2012/pdf/463Paper.pdf

148. Tiedemann, J. (2012). Character-Based Pivot Translations for Under-Resourced Languages and Domains. In *Unknown host publication* (pp. 141-151) http://www.aclweb.org/anthology/E12-1015

149. Ehrentraut, C., Tiedemann, J., Dalianis, H., & Tanushi, H. (2012). Detection of Hospital Acquired Infections in Sparse and Noisy Swedish Patient Records: A machine learning approach using Naïve Bayes, Support Vector Machines and C4.5 . In *Proceedings of the Sixth Workshop on Analytics for Noisy Unstructured Text Data, 2012* Association for Computing Machinery.

150. Tiedemann, J., Zumpe, M., & Schleussner, S. (2012). Opening OPUS for User Contributions. In *Unknown host publication*

151. Vandeghinste, V., Martens, S., Kotzé, G., Tiedemann, J., Bogaert, J. V. D., Smet, K. D., Eynde, F. V., & Noord, G. V. (2012). Parse and Corpus-based Machine Translation. In P. Spyns, & J. Odijk (Eds.), *Essential Speech and Language Technology for Dutch: resources, tools and applications* (Theory and Applications of Natural Language Processing). Springer.

152. Hardmeier, C., Tiedemann, J., Saers, M., Kessler, F., & Kessler, F. (2011). The Uppsala-FBK systems at WMT 2011. In *Sixth Workshop on Statistical Machine Translation: WMT 2011 : proceedings of the Workshop : July 30–31, 2011* (pp. 372-378). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W11-2144

153. Plas, L. V. D., Tiedemann, J., & Fahmi, I. (2011). Automatic extraction of medical term variants from multilingual parallel translations. In A. Bosch, & G. Bouma (Eds.), *Interactive Multi-modal Question-Answering* (Theory and Applications of Natural Language Processing). Springer.

154. Tiedemann, J. (2011). *Bitext Alignment*. (Synthesis Lectures on Human Language Technologies; Vol. 4, No. 2). Morgan & Claypool publishers. http://dx.doi.org/10.2200/S00367ED1V01Y201106HLT014

155. Plas, L. V. D., Tiedemann, J., & Manguin, J.-L. (2011). Synonym acquisition across domains and languages. In V. Pallotta, A. Soro, & E. Vargiu (Eds.), *Advances in Distributed Agent-based Retrieval Tools* (pp. 41-57). (Studies in Computational Intelligence; Vol. 361). Springer.

156. van der Plas, L., & Tiedemann, J. (2010). Finding Medical Term Variations using Parallel Corpora and Distributional Similarity. In A. Oltramari, P. Vossen, & Q. Lu (Eds.), *Proceedings of OntoLex 2010: 6th Workshop on Ontologies and Lexical Resources* (pp. 28-37). Coling 2010 Organizing Committee. http://www.aclweb.org/anthology/W10-3304

157. Tiedemann, J. (2010). Context Adaptation in Statistical Machine Translation Using Models with Exponentially Decaying Cache. In H. Daumé, T. Deoskar, D. McClosky, B. Plank, & J. Tiedemann (Eds.), *DANLP 2010: 2010 Workshop on Domain Adaptation for Natural Language Processing : proceedings of the Workshop : 15 July 2010, Uppsala University, Uppsala, Sweden.* (pp. 8-15). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W10-2602

158. III, H. D., Deoskar, T., McClosky, D., Plank, B., & Tiedemann, J. (Eds.) (2010). *Proceedings of the 2010 Workshop on Domain Adaptation for Natural Language Processing*. The Association for Computational Linguistics. http://www.aclweb.org/anthology/W10-26

159. Tiedemann, J. (2010). To Cache or Not To Cache? Experiments with Adaptive Models in Statistical Machine Translation. In *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR* (pp. 189-194). The Association for Computational Linguistics. http://www.aclweb.org/anthology/W10-1728

160. Islam, M. Z., Tiedemann, J., & Eisele, A. (2010). English to Bangla Phrase-Based Machine Translation. In F. Yvon, & V. Hansen (Eds.), *Proceedings of the 14th annual conference of the European Association for machine translation: EAMT 2010 : 27-28 May 2010 Saint-Raphaël Congrès, Saint-Raphaël, France http://www.mt-archive.info/EAMT-2010-Islam.pdf*

161. Tiedemann, J. (2010). Lingua-Align: An Experimental Toolbox for Automatic Tree-to-Tree Alignment. In N. C. Chair), K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, & D. Tapias (Eds.), *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, 17-23 May 2010, Valletta, Malta* (pp. 736-743). European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2010/pdf/144Paper.pdf

162. Plas, L. V. D., Tiedemann, J., & Manguin, J.-L. (2010). Automatic acquisition of synonyms for French using parallel corpora. In *Unknown host publication http://www.dart-project.org/dart2010/doku.php?id=program*

163. Tiedemann, J., & Weijnitz, P. (2010). Let's MT! A Platform for Sharing SMT Training Data. In *Unknown host publication*

164. Ahrenberg, L., Tiedemann, J., & Volk, M. (Eds.) (2010). *Proceedings of the Workshop on Annotation and Exploitation of Parallel Corpora*. (NEALT Proceedings Series). Northern European Association for Language Technology . http://hdl.handle.net/10062/15893

165. Tiedemann, J., & Nabende, P. (2009). Translating Transliterations. *International journal of computing and ICT research* , *3*(1), 33-41. http://www.ijcir.org/specialissue2009/

166. Tiedemann, J., & Kotzé, G. (2009). A Discriminative Approach to Tree Alignment. In *Unknown host publication* (pp. 33-39) http://www.aclweb.org/anthology/W09-4206

167. Tiedemann, J. (2009). Evidence-Based Word Alignment. In I. Ilisei, V. Pekar, & S. Bernardini (Eds.), *Unknown host publication* (pp. 28-32) http://www.aclweb.org/anthology/W09-4205

168. Tiedemann, J. (2009). Character-based PSMT for Closely Related Languages. In L. Marqués, & H. Somers (Eds.), *Unknown host publication* (pp. 12 - 19) http://www.mt-archive.info/EAMT-2009-Tiedemann-1.pdf

169. Tiedemann, J. (2009). Translating Questions for Cross-Lingual QA. In L. Marqués, & H. Somers (Eds.), *Unknown host publication* (pp. 112 - 119) http://www.mt-archive.info/EAMT-2009-Tiedemann-2.pdf

170. Tiedemann, J., & Kotzé, G. (2009). Building a Large Machine-Aligned Parallel Treebank. In *Unknown host publication*

171. Bertels, A., Fairon, C., Tiedemann, J., & Verlinde, S. (2009). Corpus parallèles et corpus ciblés au secours du dictionnaire de traduction. In *Cahiers de lexicologie* (pp. 199-219). (Revues). Classiques Garnier.

172. Tiedemann, J. (2009). News from OPUS - A Collection of Multilingual Parallel Corpora with Tools and Interfaces. In N. Nicolov, K. Bontcheva, G. Angelova, & R. Mitkov (Eds.), *Recent Advances in Natural Language Processing* (Vol. V, pp. 237-248)