

Jörg Tiedemann
professori, Ohjaaja tohtorihjelmassa
Digitaalisten ihmistieteiden osasto
Kieliteknologia
Kielentutkimuksen tohtorihjelma
Osoitetyyppi: Postiosoite.
PL 24 (Unioninkatu 40)
00014
Suomi
Sähköposti: jorg.tiedemann@helsinki.fi
Matkapuhelin: +358503119440
Puhelin: +358294123261



Pätevyudet

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

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

Vastuullinen tutkija

Ajanjakso : 27.01.2016 - 31.12.2017 organisaatiossa

Vastuullinen tutkija

Ajanjakso : 17.02.2018 - 31.12.2025 organisaatiossa Digitaalisten ihmistieteiden osasto

Työsuhteet

professori

Digitaalisten ihmistieteiden osasto
Helsingin yliopisto
Suomi
1 tammik. 2018 → present

Kieliteknologia

Helsingin yliopisto
University of Helsinki, Suomi
1 elok. 2015 → present

Ohjaaja tohtorihjelmassa

Kielentutkimuksen tohtorihjelma
Helsingin yliopisto
Suomi
1 tammik. 2014 → present

Projektit

- Open European Family of Large Language M**
Tiedemann, J. (Projektinjohtaja)
Univerzita Karlova Filozofická fakulta
01/02/2025 → 31/01/2028
- GreenNLP: GreenNLP - Att kontrollera koldioxidavtrycket inom hållbar språkteknologi**
Tiedemann, J. (Projektinjohtaja), Attieh, J. (osallistuja), Nieminen, T. J. (osallistuja) & Štefánik, M. (osallistuja)
Suomen Akatemia Projekttilaskutus
01/01/2023 → 31/12/2025

3. **High Performance Language Technologies**
Tiedemann, J. (Projektinjohtaja), Aulamo, M. (osallistuja), De Gibert Bonet, O. (osallistuja), Grönroos, S.-A. (osallistuja), Ji, S. (osallistuja), Mickus, T. (osallistuja), Vahtola, T. (osallistuja), Vazquez, R. (osallistuja) & Virpioja, S. P. (osallistuja)
Charles University in Prague Faculty of Science Department of Teaching and Didactics of Biology
01/09/2022 → 31/08/2025
4. **Uncertainty-aware neural language models**
Tiedemann, J. (Projektinjohtaja), Celikkanat, H. (Osallistuja), Virpioja, S. P. (Osallistuja) & Vazquez, R. (Osallistuja)
Academy of Finland, Suomen Akatemia Projektilaskutus
01/01/2022 → 01/10/2025
5. **CorCoDial: CorCoDial - Tekstikorpuksiin perustuva laskennallinen murre tutkimus: konekäännöstekniikoiden hyödyntäminen murrealueiden ja murrepiirteiden löytämisessä, visualisoinnissa ja tulkitsemisessä**
Scherrer, Y. (Projektinjohtaja), Tiedemann, J. (Projektinjohtaja), Kuparinen, O. V. (osallistuja), Mickus, T. (osallistuja), Miletic Haddad, A. (osallistuja), Psaltaki, E. (osallistuja), Roemling, D. (osallistuja), Siewert, J. (osallistuja) & Siewert, J. (osallistuja)
Suomen Akatemia Projektilaskutus
01/09/2021 → 31/08/2025
6. **Behind the words: Deep neural models of language meaning for industry-grade applications**
Tiedemann, J. (Projektinjohtaja), Creutz, M. (Principal Investigator), Creutz, M. (osallistuja), Itkonen, S. (osallistuja), Sjöblom, E. I. (osallistuja), Vahtola, T. (osallistuja), Itkonen, S. (Osallistuja) & Vahtola, T. (Osallistuja)
Academy of Finland, Suomen Akatemia Projektilaskutus
01/01/2021 → 31/12/2023
7. **OPUS-MT: Open Translation Models, Tools and Services**
Aulamo, M. (Osallistuja), Nieminen, T. J. (Osallistuja), Hardwick, S. (Osallistuja) & Tiedemann, J. (Projektinjohtaja)
01/08/2020 → 31/08/2021
8. **EOSC-nordic: The European Open Science Cloud - Nordic**
Tiedemann, J. (Projektinjohtaja), Drobac, S. (Projektinjohtaja) & Raganato, A. (Projektinjohtaja)
European Commission, European Commission – Joint Research Centre
01/09/2019 → 31/08/2022
9. **FoTran: Found in Translation - Natural Language Understanding with Cross-Lingual Grounding**
Tiedemann, J. (Projektinjohtaja), Attieh, J. (osallistuja), Aulamo, M. (osallistuja), Celikkanat, H. (osallistuja), De Gibert Bonet, O. (osallistuja), Grönroos, S.-A. (osallistuja), Mickus, T. (osallistuja), Raganato, A. (osallistuja), Scherrer, Y. (osallistuja), Silfverberg, M. (osallistuja), Sjöblom, E. I. (osallistuja), Talman, A. (osallistuja), Vazquez, R. (osallistuja), Virpioja, S. P. (osallistuja), Yli-Jyrä, A. (osallistuja), Zosa, E. (osallistuja), Celikkanat, H. (Osallistuja), Raganato, A. (Osallistuja), Silfverberg, M. (Osallistuja), Sulubacak, U. (Osallistuja) & Vazquez, R. (Osallistuja)
01/09/2018 → 31/03/2024
10. **fiskmö: Creation of a parallel corpus of translated documents and machine translation for Finnish and Swedish**
Tiedemann, J. (Projektinjohtaja), Ginter, F. (Projektinjohtaja), Papula, N. (Projektinjohtaja), Aulamo, M. (Osallistuja), Nieminen, T. (Osallistuja), Kanerva, J. (Osallistuja) & Eskola, K. (Osallistuja)
Svenska kulturfonden
01/05/2018 → 31/03/2021
11. **MeMAD: Methods for Managing Audiovisual Data: Combining Automatic Efficiency with Human Accuracy**
Hirvonen, M. (Principal Investigator), Tiedemann, J. (Osallistuja), Tiittula, L. (Osallistuja), Sulubacak, U. (Osallistuja), Vazquez, R. (Osallistuja) & Koponen, M. (Osallistuja)
European Commission / Horizon 2020
01/01/2018 → 31/03/2021
12. **NLUxLG: NLU with Cross-Lingual Grounding**
Tiedemann, J. (Projektinjohtaja), Talman, A. (Osallistuja), Raganato, A. (Osallistuja) & Celikkanat, H. (Osallistuja)
SUOMEN AKATEMIA
01/01/2018 → 31/12/2019
13. **Kielten kandiohjelman opettajien digiloikka**
Lennes, M. (Osallistuja), Flanagan, J. (Osallistuja), Hiltunen, T. (Osallistuja), Creutz, M. (Osallistuja), Aunio, L. (Osallistuja), Buchart, M. (Osallistuja), Gruzdeva, E. (Osallistuja), Jakubek, I. (Osallistuja), Keinänen, N. (Osallistuja), Kim, J.-Y. (Osallistuja), Kittilä, S. (Osallistuja), Kopotev, M. (Osallistuja), Kraenker, S. (Osallistuja), Länsisalmi, R. (Osallistuja), Nieuweboer, R. (Osallistuja), Peterson, E. (Osallistuja), Tiedemann, J. (Osallistuja), Tuori, R. (Osallistuja), Autere, M. J. (Osallistuja) & Lindstedt, J. (Projektinjohtaja)
01/05/2017 → 31/12/2018
14. **NLPL: Nordic Language Processing Laboratory**
Tiedemann, J. (Osallistuja) & Scherrer, Y. (Osallistuja)
01/01/2017 → 31/12/2019

15. **CrossNLP: Cross-Linguistic and Multilingual Natural Language Processing with the Focus on Low-Resource Languages and Language Variants**
Tiedemann, J. (Projektinjohtaja), Östling, R. (Principal Investigator) & Scherrer, Y. (Osallistuja)
01/01/2016 → 31/12/2018
16. **Erillisrahoitus Helsingin ja Tukholman yliopistojen yhteishanke**
Lehti-Eklund, H. (Projektinjohtaja), Lindström, J. (Osallistuja), Miestamo, M. (Osallistuja), Grünthal, R. (Osallistuja), Palander-Collin, M. (Osallistuja), Kalliokoski, J. (Osallistuja), Östman, J.-O. (Osallistuja), Ajanki, R. (Osallistuja), Holopainen, S. (Projektinjohtaja), Balode, L. (Osallistuja), Lenk, H. E. H. (Osallistuja), Kurhila, S. (Osallistuja), Kotilainen, L. (Osallistuja), Lindholm, C. (Osallistuja), Wahlström, M. (Projektinjohtaja), Solin, A. (Projektinjohtaja), Fremer, M. (Projektinjohtaja), Tiittula, L. (Projektinjohtaja), Hamari, A. (Projektinjohtaja), Kittilä, S. (Projektinjohtaja), Huhtamäki, M. (Projektinjohtaja), Kim, J.-Y. (Projektinjohtaja), Kopotev, M. (Projektinjohtaja), Peterson, E. (Projektinjohtaja), Protassova, E. (Projektinjohtaja), Savijärvi, M. (Projektinjohtaja) & Tiedemann, J. (Projektinjohtaja)
Unknown funder
01/01/2015 → 31/12/2017
17. **eSSENCE: Efficient Algorithms for Natural Language Processing Beyond Sentence Boundaries**
Nivre, J. (Projektinjohtaja), Tiedemann, J. (Projektinjohtaja) & Stymne, S. (Principal Investigator)
01/09/2012 → 31/08/2014
18. **DiscoMT: Discourse-Oriented Statistical Machine Translation**
Tiedemann, J. (Projektinjohtaja), Hardmeier, C. (Principal Investigator), Loáiciga, S. (Osallistuja) & Scherrer, Y. (Osallistuja)
01/01/2012 → 31/12/2017
19. **LetsMT!: LetsMT! A Platform for Online Sharing of Training Data and Building User Tailored Machine Translation**
Tiedemann, J. (Projektinjohtaja), Weinitz, P. (Principal Investigator), Saers, M. (Principal Investigator), Zumpfe, M. (Principal Investigator) & Schleussner, S. (Principal Investigator)
01/03/2010 → 31/08/2012
20. **OPUS: The Open Parallel Corpus**
Tiedemann, J. (Projektinjohtaja)
01/06/2004 → ...

Tietoaineistot

1. **Document-level Machine Translation Benchmark**
Scherrer, Y. (Luoja), Tiedemann, J. (Luoja) & Loáiciga, S. (Luoja), University of Helsinki, 1 marrask. 2019
DOI - pysyväislinkki: 10.5281/zenodo.3525366
Tietoaineisto
2. **LetsMT repository**
Tiedemann, J. (Luoja), Zenodo, 22 tammik. 2018
DOI - pysyväislinkki: 10.5281/zenodo.1157836
Tietoaineisto
3. **Uplug**
Tiedemann, J. (Luoja), Zenodo, 22 tammik. 2018
DOI - pysyväislinkki: 10.5281/zenodo.1157180, <https://github.com/Helsinki-NLP/Uplug/tree/v0.3.9>
Tietoaineisto
4. **Lingua-Align**
Tiedemann, J. (Luoja), Zenodo, 22 tammik. 2018
DOI - pysyväislinkki: 10.5281/zenodo.1157176, <https://github.com/Helsinki-NLP/Lingua-Align/tree/v0.1>
Tietoaineisto
5. **OPUS**
Tiedemann, J. (Luoja), University of Helsinki, 2017
<http://opus.nlpl.eu>
Tietoaineisto
6. **WMT17 test set Finnish-English**
Tiedemann, J. (Luoja), Zenodo, 21 tammik. 2018
DOI - pysyväislinkki: 10.5281/zenodo.1156326, <https://github.com/Helsinki-NLP/WMT16-test-enfi/tree/v1.0>
Tietoaineisto
7. **WMT16 test set Finnish-English**
Tiedemann, J. (Luoja), Zenodo, 21 tammik. 2018
DOI - pysyväislinkki: 10.5281/zenodo.1156325, <https://github.com/Helsinki-NLP/WMT16-test-enfi/tree/v1.0>
Tietoaineisto

8. **Shared Task in Cross-Lingual Parsing**
Tiedemann, J. (Tietojen hallinnoija), Atlassian, 2017
<https://bitbucket.org/hy-crossNLP/wardial2017>
Tietoaineisto
9. **DiscoMT 2015 Shared Task on Pronoun Translation**
Tiedemann, J. (Luoja), LINDAT/CLARIN, 31 tammik. 2016
DOI - pysyväislinkki: <http://hdl.handle.net/11372/LRT-1611>, <http://hdl.handle.net/11372/LRT-1611>
Tietoaineisto

Julkaisut

1. Li, Z., Ji, S., Mickus, T., Segonne, V., & Tiedemann, J. (2024). A Comparison of Language Modeling and Translation as Multilingual Pretraining Objectives. teoksessa Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (Toimittajat), *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (Sivut 15882-15894). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.emnlp-main.888>
2. Scherrer, Y., Jauhainen, T., Ljubešić, N., Zampieri, M., Nakov, P., & Tiedemann, J. (Toimittajat) (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.wardial-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. teoksessa A. K. Ojha, A. S. Doğruöz, H. Tayyar Madabushi, G. Da San Martino, S. Rosenthal, & A. Rosá (Toimittajat), *Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)* (Sivut 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. teoksessa G. M. Di Nunzio, F. Vezzani, L. Ermakova, H. Azarbyonad, & J. Kamps (Toimittajat), *Proceedings of the Workshop on DeTermit! Evaluating Text Difficulty in a Multilingual Context @ LREC-COLING 2024* (Sivut 39–50). (International conference on computational linguistics), (LREC proceedings). European Language Resources Association (ELRA). <https://aclanthology.org/2024.determin-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. teoksessa N. Aletras, & O. De Clercq (Toimittajat), *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations* (Sivut 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. (Toimittajat) (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. teoksessa N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, & N. Xue (Toimittajat), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (Sivut 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? teoksessa N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, & N. Xue (Toimittajat), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (Sivut 2809-2818). (International conference on computational linguistics), (LREC proceedings). European Language Resources Association (ELRA).
9. Arefyev, N., Aulamo, M., Chen, P., de Gibert, O., Haddow, B., Helcl, J., Malik, B., Ramírez-Sánchez, G., Stepachev, P., Tiedemann, J., Variš, D., & Zaragoza, J. (2024). HPLT's First Release of Data and Models. teoksessa C. Scarton, C. Prescott, C. Bayliss, C. Oakley, J. Wright, S. Wrigley, X. Song, E. Gow-Smith, M. Forcada, & H. Moniz (Toimittajat), *Proceedings of the 25th Annual Conference of the European Association for Machine Translation: Volume 2: Products & Projects* (Sivut 53-54). European Association for Machine Translation. <https://eamt2024.github.io/proceedings/vol2.pdf>
10. 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. teoksessa B. Haddow, T. Kocmi, P. Koehn, & C. Monz (Toimittajat), *Proceedings of the Ninth Conference on Machine Translation* (Sivut 908-917). The Association for Computational Linguistics.
11. 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>

12. 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. teoksessa M. Mager, A. Ebrahimi, & A. Oncevay, et al. (Toimittajat), *Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP)* (Sivut 177-191). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.americasnlp-1.20>
13. Vahtola, T., Creutz, M., & Tiedemann, J. (2023). Guiding Zero-Shot Paraphrase Generation with Fine-Grained Control Tokens. teoksessa A. Palmer, & J. Camacho-collados (Toimittajat), *Proceedings of the 12th Joint Conference on Lexical and Computational Semantics (*SEM 2023)* (Sivut 323-337). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.starsem-1.29>
14. Aulamo, M., De Gibert Bonet, O., Virpioja, S., & Tiedemann, J. (2023). Unsupervised Feature Selection for Effective Parallel Corpus Filtering. teoksessa M. Nurminen, J. Brenner, & M. Koponen, et al. (Toimittajat), *Proceedings of the 24th Annual Conference of the European Association for Machine Translation* (Sivut 31-38). European Association for Machine Translation.
15. Scherrer, Y., Jauhainen, T., Ljubešić, N., Nakov, P., Tiedemann, J., & Zampieri, M. (Toimittajat) (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>
16. 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. teoksessa T. Alumäe, & M. Fishel (Toimittajat), *Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)* (Sivut 238–247). (NEALT Proceedings Series Publisher name; Nro 52). University of Tartu Library.
17. 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.
18. 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. teoksessa M. Nurminen, J. Brenner, & M. Koponen, et al. (Toimittajat), *Proceedings of the 24th Annual Conference of the European Association for Machine Translation* (Sivut 517-518). European Association for Machine Translation.
19. Tiedemann, J., & Gibert, O. D. (2023). The OPUS-MT dashboard - A toolkit for a systematic evaluation of open machine translation models. teoksessa D. Bollegala, R. Huang, & A. Ritter (Toimittajat), *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics: Volume 3: System Demonstrations* (Vuosikerta 3, Sivut 315-327). Association for Computational Linguistics (ACL). <https://doi.org/10.18653/v1/2023.acl-demo.30>
20. Talman, A., Celikkanat, H., Virpioja, S., Heinonen, M., & Tiedemann, J. (2023). Uncertainty-Aware Natural Language Inference with Stochastic Weight Averaging. teoksessa T. Alumäe, & M. Fishel (Toimittajat), *Proceedings of the 24th Nordic Conference on Computational Linguistics* (Sivut 358-365). (NEALT Proceedings Series; Nro 52). University of Tartu Library.
21. 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. teoksessa J. Bastings, Y. Belinkov, Y. Elazar, D. Hupkes, N. Saphra, & S. Wiegrefe (Toimittajat), *Proceedings of the Fifth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP* (Sivut 249–262). The Association for Computational Linguistics.
22. Tiedemann, J. (2022). From open parallel corpora to public translation tools: The success story of OPUS. teoksessa E. Volodina, D. Dannélls, A. Berdicevskis, M. Forsberg, & S. Virk (Toimittajat), *LIVE and LEARN : Festschrift in honor of Lars Borin* (Sivut 133-138). (Research Reports from the Department of Swedish, Multilingualism, Language Technology; Nro GU-ISS-2022-03). University of Göteborg.
23. Tiedemann, J., Aulamo, M., Hardwick, S., & Nieminen, T. (2022). Open Translation Models, Tools and Services. teoksessa G. Rehm (Toimittaja), *European Language Grid: A Language Technology Platform for Multilingual Europe* (Sivut 325-330). (Cognitive Technologies). Springer . https://doi.org/10.1007/978-3-031-17258-8_24
24. Vazquez, R., Celikkanat, H., Ravishankar, V., Creutz, M., & Tiedemann, J. (2022). A Closer Look at Parameter Contributions When Training Neural Language and Translation Models. teoksessa N. Calzolari, C.-R. Huang, & H. Kim, et al. (Toimittajat), *Proceedings of the 29th International Conference on Computational Linguistics* (Sivut 4788-4800). (International conference on computational linguistics; Vuosikerta 29, Nro 1). International Committee on Computational Linguistics.
25. Scherrer, Y., Jauhainen, T., Ljubešić, N., Nakov, P., Tiedemann, J., & Zampieri, M. (Toimittajat) (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 ; Vuosikerta 29, Nro 19). COLING . <https://aclanthology.org/2022.vardial-1.0>
26. Itkonen, S., Tiedemann, J., & Creutz, M. (2022). Helsinki-NLP at SemEval-2022 Task 2: A Feature-Based Approach to Multilingual Idiomaticity Detection. teoksessa G. Emerson, N. Schluter, G. Stanovsky, R. Kumar, A. Palmer, N. Schneider, S. Singh, & S. Ratan (Toimittajat), *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)* (Sivut 122-134). The Association for Computational Linguistics.

27. Vahtola, T., Sjöblom, E., Tiedemann, J., & Creutz, M. (2022). Modeling Noise in Paraphrase Detection. teoksessa N. Calzolari, F. Béchet, & P. Blache, et al. (Toimittajat), *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)* (Sivut 4324-4332). European Language Resources Association (ELRA).
28. Tiedemann, J. (2022). Språk(tekologi) är nyckeln till intelligens och rättvisa. *Språkbruk*. <https://www.sprakbruk.fi/-/spr%C3%A5k-tekologi-%C3%A4r-nyckeln-till-intelligens-och-r%C3%A4ttvisa>
29. Talman, A., Apidianaki, M., Chatzikiyakidis, S., & Tiedemann, J. (2022). How Does Data Corruption Affect Natural Language Understanding Models? A Study on GLUE datasets. teoksessa V. Nastase, E. Pavlick, M. Taher Pilehvar, J. Camacho-Collados, & A. Raganato (Toimittajat), *Proceedings of The 11th Joint Conference on Lexical and Computational Semantics* (Sivut 226-233). The Association for Computational Linguistics.
30. 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. teoksessa H. Moniz, L. Macken, & A. Rufener, et al. (Toimittajat), *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation* (Sivut 311-312). European Association for Machine Translation. <https://aclanthology.org/2022.eamt-1.45>
31. Alnajjar, K., Hämmäläinen, M., Tiedemann, J., Laaksonen, J., & Kurimo, M. (2022). When to Laugh and How Hard? A Multimodal Approach to Detecting Humor and Its Intensity. teoksessa N. Calzolari, C.-R. Huang, & H. Kim, et al. (Toimittajat), *Proceedings of the 29th International Conference on Computational Linguistics* (Sivut 6875-6886). (International Conference on Computational Linguistics; Vuosikerta 29, Nro 1). International Committee on Computational Linguistics. <https://aclanthology.org/2022.coling-1.598>
32. Raganato, A., Vázquez, R., Creutz, M., & Tiedemann, J. (2021). An Empirical Investigation of Word Alignment Supervision for Zero-Shot Multilingual Neural Machine Translation. teoksessa M.-F. Moens, X. Huang, L. Specia, & S. Wen-tau Yih (Toimittajat), *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (Sivut 8449-8456). The Association for Computational Linguistics. <https://aclanthology.org/2021.emnlp-main.664.pdf>
33. Nakov, P., & Tiedemann, J. (2021). Machine Translation between Similar Languages. teoksessa M. Zampieri, & P. Nakov (Toimittajat), *Similar Languages, Varieties, and Dialects: A Computational Perspective* (Sivut 219–253). (Studies in Natural Language Processing). Cambridge University Press. <https://doi.org/10.1017/9781108565080.015>
34. 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. teoksessa J. Kabbara, H. Lin, A. Paullada, & J. Vamvas (Toimittajat), *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* (Sivut 337-347). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.acl-srw.35>
35. Vázquez, R., Scherrer, Y., Virpioja, S., & Tiedemann, J. (2021). The Helsinki submission to the AmericasNLP shared task. teoksessa M. Mager [et al.] (Toimittaja), *Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas* (Sivut 255-264). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.americasnlp-1.29>
36. Aulamo, M., Virpioja, S., Scherrer, Y., & Tiedemann, J. (2021). Boosting Neural Machine Translation from Finnish to Northern Sámi with Rule-Based Backtranslation. teoksessa S. Dobnik, & L. Øvrelid (Toimittajat), *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)* (Sivut 351-356). (Linköping Electronic Conference Proceedings ; Nro 78), (NEALT Proceedings Series ; Nro 45). Linköping University Electronic Press. <https://www.aclweb.org/anthology/2021.nodalida-main.37.pdf>
37. Dmitrieva, A., & Tiedemann, J. (2021). Creating an Aligned Russian Text Simplification Dataset from Language Learner Data. teoksessa B. Babych [et al.] (Toimittaja), *Proceedings of the 8th Workshop on Balto-Slavic Natural Language Processing* (Sivut 73-79). ACL Anthology. <https://aclanthology.org/2021.bsnlp-1.8>
38. Dmitrieva, A., & Tiedemann, J. (2021). A Multi-task Learning Approach to Text Simplification. teoksessa W. M. P. van der Aalst [et al.] (Toimittaja), *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* (Sivut 78-89). (Communications in Computer and Information Science; Vuosikerta 1357). Springer. https://doi.org/10.1007/978-3-030-71214-3_7
39. Tsarfaty, R., & Tiedemann, J. (2021). Message from the program chairs. teoksessa *EACL 2021. The 16th Conference of the European Chapter of the Association for Computational Linguistics: Proceedings of the Conference* (Sivut vi-viii). (Proceedings of the European Association for Computational Linguistics). The Association for Computational Linguistics. <https://aclanthology.org/2021.eacl-main.0.pdf>
40. Talman, A., Apidianaki, M., Chatzikiyakidis, S., & Tiedemann, J. (2021). NLI Data Sanity Check: Assessing the Effect of Data Corruption on Model Performance. teoksessa S. Dobnik, & L. Øvrelid (Toimittajat), *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa 2021)* (Sivut 276-285). (Linköping Electronic Conference Proceedings ; Nro 78), (NEALT Proceedings Series ; Nro 45). Linköping University Electronic Press.
41. Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Scherrer, Y., & Jauhainen, T. (Toimittajat) (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>

42. Tiedemann, J. (2021). The Development of a Comprehensive Data Set for Systematic Studies of Machine Translation. teoksessa M. Härmäläinen, N. Partanen, & K. Alnajjar (Toimittajat), *Multilingual Facilitation* (Sivut 248-262). University of Helsinki. <https://doi.org/10.31885/9789515150257.22>
43. Siewert, J., Scherrer, Y., & Tiedemann, J. (2021). Towards a balanced annotated Low Saxon dataset for diachronic investigation of dialectal variation. teoksessa K. Evang, L. Kallmeyer, R. Osswald, J. Waszczuk, & T. Zesch (Toimittajat), *Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021)* (Sivut 242-246). KONVENS 2021 Organizers. <https://aclanthology.org/2021.konvens-1.25.pdf>
44. Siewert, J., Scherrer, Y., Wieling, M., & Tiedemann, J. (2020). LSDC - A comprehensive dataset for Low Saxon Dialect Classification. teoksessa M. Zampieri, P. Nakov, N. Ljubešić, J. Tiedemann, & Y. Scherrer (Toimittajat), *Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects* (Sivut 25-35). International Committee on Computational Linguistics (ICCL).
45. Celikkanat, H., Virpioja, S., Tiedemann, J., & Apidianaki, M. (2020). Tracking the Traces of Passivization and Negation in Contextualized Representations. teoksessa A. Alishahi, Y. Belinkov, G. Chrupala, D. Hupkes, Y. Pinter, & H. Sajjad (Toimittajat), *Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP* (Sivut 136-148). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.blackboxnlp-1.13>
46. Raganato, A., Scherrer, Y., & Tiedemann, J. (2020). Fixed Encoder Self-Attention Patterns in Transformer-Based Machine Translation. teoksessa T. Cohn, Y. H. He, & Y. Liu (Toimittajat), *Findings of the Association for Computational Linguistics: EMNLP 2020* (Sivut 556-568). The Association for Computational Linguistics. <https://www.aclweb.org/anthology/2020.findings-emnlp.49>
47. Tiedemann, J., & Thottingal, S. (2020). OPUS-MT -- Building open translation services for the World. teoksessa A. Martins [et al.] (Toimittaja), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (Sivut 479-480). European Association for Machine Translation.
48. Scherrer, Y., Raganato, A., & Tiedemann, J. (2020). The MUCOW word sense disambiguation test suite at WMT 2020. teoksessa L. Barrault [et al.] (Toimittaja), *Proceedings of the Fifth Conference on Machine Translation* (Sivut 365-370). The Association for Computational Linguistics.
49. Tiedemann, J. (2020). The Tatoeba Translation Challenge - Realistic Data Sets for Low Resource and Multilingual MT. teoksessa L. Barrault [et al.] (Toimittaja), *Proceedings of the Fifth Conference on Machine Translation* (Sivut 1174-1182). The Association for Computational Linguistics.
50. Koponen, M., Sulubacak, U., Vitikainen, K., & Tiedemann, J. (2020). MT for Subtitling: Investigating professional translators' user experience and feedback. teoksessa J. E. Ortega, M. Federico, C. Orasan, & M. Popovic (Toimittajat), *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* (Sivut 79-92). AMTA. <https://www.aclweb.org/anthology/2020.amta-pemdt.6/>
51. 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. Artikkel 2. <https://doi.org/10.1007/s10590-020-09250-0>
52. Vázquez, R., Aulamo, M., Sulubacak, U., & Tiedemann, J. (2020). The University of Helsinki Submission to the IWSLT2020 Offline Speech Translation Task. teoksessa *Proceedings of the 17th International Conference on Spoken Language Translation (IWSLT)* (Sivut 95-102). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.iwslt-1.10>
53. Aulamo, M., Virpioja, S., & Tiedemann, J. (2020). OpusFilter: A Configurable Parallel Corpus Filtering Toolbox. teoksessa A. Çelikyilmaz, & T.-H. Wen (Toimittajat), *58TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL 2020): SYSTEM DEMONSTRATIONS: System Demonstrations* (Sivut 150-156). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-demos.20>
54. Koponen, M., Sulubacak, U., Vitikainen, K., & Tiedemann, J. (2020). MT for subtitling: User evaluation of post-editing productivity. teoksessa 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 (Toimittajat), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation (EAMT 2020)* (Sivut 115-124). European Association for Machine Translation.
55. 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
56. Aulamo, M., Sulubacak, U., Virpioja, S., & Tiedemann, J. (2020). OpusTools and Parallel Corpus Diagnostics. teoksessa 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 (Toimittajat), *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)* (Sivut 3782-3789). European Language Resources Association (ELRA).
57. Raganato, A., Scherrer, Y., & Tiedemann, J. (2020). An Evaluation Benchmark for Testing the Word Sense Disambiguation Capabilities of Machine Translation Systems. teoksessa N. Calzolari [et al.] (Toimittaja), *Proceedings of The 12th Language Resources and Evaluation Conference* (Sivut 3668-3675). European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/2020.lrec-1.452.pdf>

58. 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. teoksessa 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 (Toimittajat), *Proceedings of the 12th Language Resources and Evaluation Conference* (Sivut 3808-3815). European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/2020.lrec-1.470>
59. 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>
60. Kajava, K., Öhman, E., Hui, P., & Tiedemann, J. (2020). Emotion Preservation in Translation: Evaluating Datasets for Annotation Projection. teoksessa S. Reinsone, I. Skadiņa, A. Baklāne, & J. Daugavietis (Toimittajat), *Proceedings of Digital Humanities in Nordic Countries (DHN 2020)* (Sivut 38-50). (CEUR workshop proceedings; Nro 2612). CEUR.
61. Pàmies, M., Öhman, E., Kajava, K., & Tiedemann, J. (2020). LT@Helsinki at SemEval-2020 Task 12: Multilingual or language-specific BERT? teoksessa A. Herbelot, X. Zhu, A. Palmer, N. Schneider, J. May, & E. Shutova (Toimittajat), *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (Sivut 1569-1575). International Committee for Computational Linguistics.
62. Öhman, E., Pàmies, M., Kajava, K., & Tiedemann, J. (2020). XED: A Multilingual Dataset for Sentiment Analysis and Emotion Detection. teoksessa D. Scott, N. Bel, & C. Zong (Toimittajat), *Proceedings of the 28th International Conference on Computational Linguistics* (Sivut 6542-6552). International Committee on Computational Linguistics. <https://doi.org/10.18653/v1/2020.coling-main.575>
63. Scherrer, Y., Tiedemann, J., & Loáiciga, S. (2019). Analysing concatenation approaches to document-level NMT in two different domains. teoksessa *The Fourth Workshop on Discourse in Machine Translation: Proceedings of the Workshop* (Sivut 51-61). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-6506>
64. Talman, A., Suni, A., Celikkanat, H., Kakouros, S., Tiedemann, J., & Vainio, M. (2019). Predicting Prosodic Prominence from Text with Pre-trained Contextualized Word Representations. teoksessa M. Hartmann, & B. Plank (Toimittajat), *22nd Nordic Conference on Computational Linguistics (NoDaLiDa): Proceedings of the Conference* (Sivut 281-290). (Linköping Electronic Conference Proceedings; Nro 167), (NEALT Proceedings Series; Nro 42). Linköping University Electronic Press.
65. Raganato, A., Vázquez, R., Creutz, M., & Tiedemann, J. (2019). An Evaluation of Language-Agnostic Inner-Attention-Based Representations in Machine Translation. teoksessa I. Augenstein, S. Gella, S. Ruder, K. Kann, B. Can, J. Welbl, A. Conneau, X. Ren, & M. Rei (Toimittajat), *The 4th Workshop on Representation Learning for NLP (Repl4NLP-2019): Proceedings of the Workshop* (Sivut 27-32). The Association for Computational Linguistics. <https://www.aclweb.org/anthology/W19-4304>
66. 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. teoksessa O. Bojar, R. Chatterjee, C. Federmann, & E. A. (Toimittajat), *Fourth Conference on Machine Translation: Proceedings of the Conference (Volume 2: Shared Task Papers, Day 1)* (Sivut 470-480). The Association for Computational Linguistics. <https://www.aclweb.org/anthology/W19-5354>
67. 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. teoksessa *Fourth Conference on Machine Translation (WMT 2019): Proceedings of the Conference: Volume 2* (Sivut 412-423). The Association for Computational Linguistics. <http://www.statmt.org/wmt19/papers.html>
68. 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>
69. Vazquez, R., Sulubacak, U., & Tiedemann, J. (2019). The University of Helsinki submission to the WMT19 Parallel Corpus Filtering Task. teoksessa O. Bojar, R. Chatterjee, C. Federmann, & E. A. (Toimittajat), *Fourth Conference on Machine Translation: Proceedings of the Conference: Volume 3: Shared Task Papers, Day 2* (Sivut 294-300). The Association for Computational Linguistics.
70. Tiedemann, J., & Scherrer, Y. (2019). Measuring Semantic Abstraction of Multilingual NMT with Paraphrase Recognition and Generation Tasks. teoksessa A. Rogers, A. Drozd, A. Rumshisky, & Y. Goldberg (Toimittajat), *Proceedings of the 3rd Workshop on Evaluating Vector Space Representations for NLP* (Sivut 35-42). The Association for Computational Linguistics. <https://www.aclweb.org/anthology/W19-2005>
71. Vazquez Carrillo, J. R., Raganato, A., Tiedemann, J., & Creutz, M. (2019). Multilingual NMT with a language-independent attention bridge. teoksessa I. Augenstein, S. Gella, S. Ruder, K. Kann, B. Can, J. Welbl, A. Conneau, X. Ren, & M. Rei (Toimittajat), *The 4th Workshop on Representation Learning for NLP (Repl4NLP-2019): Proceedings of the Workshop* (Sivut 33-39). The Association for Computational Linguistics.
72. Lison, P., Tiedemann, J., & Kouylekov, M. (2019). Open subtitles 2018: Statistical rescoring of sentence alignments in large, noisy parallel corpora. teoksessa N. Calzolari ... [et al.] (Toimittaja), *LREC 2018, Eleventh International Conference on Language Resources and Evaluation* (Sivut 1742-1748). European Language Resources Association (ELRA).

73. Härmäläinen, M., Säily, T., Rueter, J., Tiedemann, J., & Mäkelä, E. (2019). Revisiting NMT for normalization of early English letters. teoksessa B. Alex, S. Degaetano-Ortlieb, A. Kazantseva, N. Reiter, & S. Szpakowicz (Toimittajat), *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (Sivut 71–75). (ACL Anthology; Nro W19-25). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/W19-2509>
74. Aulamo, M., & Tiedemann, J. (2019). The OPUS Resource Repository: An Open Package for Creating Parallel Corpora and Machine Translation Services. teoksessa M. Hartmann, & B. Plank (Toimittajat), *22nd Nordic Conference on Computational Linguistics (NoDaLiDa): Proceedings of the Conference* (Sivut 389–394). (Linköping Electronic Conference Proceedings; Nro 167), (NEALT Proceedings Series; Nro 42). Linköping University Electronic Press.
75. 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
76. 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. teoksessa O. Bojar, R. Chatterjee, C. Federmann, M. Fishel, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, C. Monz, M. Negri, A. Nèveol, M. Neves, M. Post, L. Specia, M. Turchi, & K. Verspoor (Toimittajat), *Proceedings of the Third Conference on Machine Translation (WMT): Shared Task Papers* (Sivut 603-611). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/W18-6439>
77. 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. teoksessa A. Balahur, S. M. Mohammad, V. Hoste, & R. Klinger (Toimittajat), *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (Sivut 24-30). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/W18-6205>
78. Sulubacak, U., Tiedemann, J., Rouhe, A., Stig-Arne, G., & Kurimo, M. (2018). The MeMAD Submission to the IWSLT 2018 Speech Translation Task. teoksessa M. Turchi, J. Niehues, & M. Federico (Toimittajat), *Proceedings of the 15th International Workshop on Spoken Language Translation (IWSLT 2018)* (Sivut 89-94).
79. 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>
80. Raganato, A., & Tiedemann, J. (2018). An Analysis of Encoder Representations in Transformer-Based Machine Translation. teoksessa L. Tal, G. Chrupala, & A. Alishahi (Toimittajat), *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP* (Sivut 287-297). The Association for Computational Linguistics. <http://aclweb.org/anthology/W18-5431>
81. Tiedemann, J. (2018). Emerging Language Spaces Learned From Massively Multilingual Corpora. teoksessa E. Mäkelä, M. Tolonen, & J. Tuominen (Toimittajat), *Proceedings of the Digital Humanities in the Nordic Countries 3rd Conference (DHN 2018)* (Vuosikerta 2084, Sivut 188-197). (CEUR Workshop Proceedings). CEUR Workshop Proceedings.
82. 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. teoksessa M. Zampieri, P. Nakov, N. Ljubešić, J. Tiedemann, S. Malmasi, & A. Ali (Toimittajat), *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects* (Sivut 1-17). The Association for Computational Linguistics. <http://web.science.mq.edu.au/~smalmasi/varDial5/pdf/W18-391.pdf>
83. Härmäläinen, M., Säily, T., Rueter, J., Tiedemann, J., & Mäkelä, E. (2018). Normalizing early English letters to Present-day English spelling. teoksessa B. Alex, S. Degaetano-Ortlieb, A. Feldman, A. Kazantseva, N. Reiter, & S. Szpakowicz (Toimittajat), *Proceedings of the 2nd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (Sivut 87-96). (ACL Anthology; Nro W18-45). The Association for Computational Linguistics. <http://aclweb.org/anthology/W18-4510>
84. Lison, P., Tiedemann, J., & Kouylekov, M. (2018). OpenSubtitles2018: Statistical Rescoring of Sentence Alignments in Large, Noisy Parallel Corpora. teoksessa 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 (Toimittajat), *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)* (Sivut 1742-1748). European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2018/summaries/294.html>
85. Bozovic, P., Erjavec, T., Tiedemann, J., Ljubesic, N., & Gorjanc, V. (2018). Opus-MontenegrinSubs 1.0: First electronic corpus of the Montenegrin language. teoksessa D. Fišer, & A. Pančur (Toimittajat), *Proceedings of the conference on Language Technologies & Digital Humanities 2018* (Sivut 24-28). Ljubljana University Press. <http://nl.ijs.si/jtdh18/proceedings-en.html>
86. Zampieri, M., Nakov, P., Ljubesic, N., Tiedemann, J., Malmasi, S., & Ali, A. (Toimittajat) (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/W17/W17-12.pdf>

87. Raganato, A., Scherrer, Y., Nieminen, T., Hurskainen, A., & Tiedemann, J. (2018). The University of Helsinki submissions to the WMT18 news task. teoksessa O. Bojar, R. Chatterjee, C. Federmann, M. Fishel, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, C. Monz, M. Negri, A. N ev ol, M. Neves, M. Post, L. Specia, M. Turchi, & K. Verspoor (Toimittajat), *Proceedings of the Third Conference on Machine Translation: Shared Task Papers* (Sivut 488-495). (WMT 2018 - 3rd Conference on Machine Translation, Proceedings of the Conference; Vuosikerta 2). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/W18-64052>
88. Shao, Y., Hardmeier, C., Tiedemann, J., & Nivre, J. (2017). Character-based Joint Segmentation and POS Tagging for Chinese using Bidirectional RNN-CRF. teoksessa G. Kondrak, & T. Watanabe (Toimittajat), *The Eighth International Joint Conference on Natural Language Processing: Proceedings of the Conference, Vol. 1 (Long Papers)* (Sivut 173-183). Asian Federation of Natural Language Processing. <http://aclweb.org/anthology/I17/I17-1018.pdf>
89. Tiedemann, J. (Toimittaja) (2017). *Proceedings of the 21st Nordic Conference on Computational Linguistics (NoDaLiDa)*. (Link oping Electronic Conference Proceedings; Nro 131). Link oping University Electronic Press. <http://www.ep.liu.se/ecp/131/ecp17131.pdf>
90.  stling, R., & Tiedemann, J. (2017). Continuous multilinguality with language vectors. teoksessa *15th Conference of the European Chapter of the Association for Computational Linguistics: Proceedings of Conference, volume 2: Short Papers* (Sivut 644-649). The Association for Computational Linguistics. <https://www.aclweb.org/anthology/E17-2102/>
91. Tiedemann, J. (2017). Cross-Lingual Dependency Parsing for Closely Related Languages: Helsinki's Submission to VarDial 2017. teoksessa *Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial'2017): Proceedings of the Workshop* (Sivut 131-136). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/w17-1216>
92. Zampieri, M., Malmasi, S., Ljube i , N., Nakov, P., Ali, A., Tiedemann, J., Scherrer, Y., & Aepli, N. (2017). Findings of the VarDial Evaluation Campaign 2017. teoksessa *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects* (Sivut 1-15). The Association for Computational Linguistics. <http://aclweb.org/anthology/W17/W17-1201.pdf>
93. Tiedemann, J., & van der Plas, L. (2017). Bootstrapping a Dependency Parser for Maltese - A Real-World Test Case. teoksessa M. Wieling, M. Kroon, G. Van Noord, & G. Bouma (Toimittajat), *From Semantics to Dialectometry: Festschrift in honor of John Nerbonne* (Sivut 355-365). (Tributes; Nro 32). College publications. <http://www.let.rug.nl/festschriftnerbonne/36.%20Tiedemann%20&%20van%20der%20Plas.pdf>
94. Lo iciga, 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. teoksessa *Discourse in Machine Translation (DiscoMT 2017): Proceedings of the Workshop* (Sivut 1-16). The Association for Computational Linguistics. <http://aclweb.org/anthology/W17/W17-4801.pdf>
95. Kotz , 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>
96. Tiedemann, J., & Scherrer, Y. (2017). Neural Machine Translation with Extended Context. teoksessa *Proceedings of the Third Workshop on Discourse in Machine Translation* (Sivut 82-92). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/w17-4811>
97. Hurskainen, A., & Tiedemann, J. (2017). Rule-based Machine Translation from English to Finnish. teoksessa *Proceedings of the Second Conference on Machine Translation (WMT2017)* (Sivut 323-329). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/w17-4731>
98. Tjong Kim Sang, E., Bollmann, M., Boschker, R., Casacuberta, F., Dietz, F., Dipper, S., Domingo, M., van der Goot, R., van Koppen, M., Ljube i , N.,  stling, 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>
99.  stling, R., Scherrer, Y., Tiedemann, J., Tang, G., & Nieminen, T. (2017). The Helsinki Neural Machine Translation System. teoksessa O. Bojar, C. Buck, R. Chatterjee, C. Federmann, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, P. Koehn, & J. Kreutzer (Toimittajat), *Proceedings of the Second Conference on Machine Translation (WMT2017)* (Sivut 338-347). (WMT 2017 - 2nd Conference on Machine Translation, Proceedings). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/w17-4733>
100. Malmasi, S., Zampieri, M., Ljube i , N., Nakov, P., Ali, A., & Tiedemann, J. (2016). Discriminating between Similar Languages and Arabic Dialect Identification: A Report on the Third DSL Shared Task. teoksessa *Third Workshop on NLP for Similar Languages, Varieties and Dialects: Proceedings of the Workshop* (Sivut 1-14). The COLING 2016 Organizing Committee . <http://web.science.mq.edu.au/~smalmasi/vardial3/pdf/VarDial301.pdf>
101. Tiedemann, J., Nichols, J., & Sprouse, R. (2016). Tagging Ingush - Language Technology For Low-Resource Languages Using Resources From Linguistic Field Work. teoksessa *Language Technology Resources and Tools for Digital Humanities (LT4DH): Proceedings of the Workshop* (Sivut 148-155). <http://aclweb.org/anthology/W16-4020>

102. Öhman, E., Honkela, T., & Tiedemann, J. (2016). The Challenges of Multi-dimensional Sentiment Analysis Across Languages. teoksessa *Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (PEOPLES): Proceedings of the Workshop* (Sivut 138-142). The COLING 2016 Organizing Committee . <https://peoples2016.github.io/peoples2016.pdf>
103. 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>
104. Tiedemann, J. (2016). A Linear Baseline Classifier for Cross-Lingual Pronoun Prediction. teoksessa *The 54th Annual Meeting of the Association for Computational Linguistics: Proceedings of the First Conference on Machine Translation (WMT)* (Sivut 616-619). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/w16-2356>
105. 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. teoksessa *The 54th Annual Meeting of the Association for Computational Linguistics: Proceedings of the First Conference on Machine Translation (WMT)* (Sivut 525-542). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W16-2345>
106. 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. teoksessa *The 54th Annual Meeting of the Association for Computational Linguistics: Proceedings of the First Conference on Machine Translation (WMT)* (Sivut 391-398). The Association for Computational Linguistics. <https://doi.org/10.18653/v1/w16-2326>
107. 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>
108. 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.
109. Tiedemann, J. (2016). Finding Alternative Translations in a Large Corpus of Movie Subtitles. teoksessa N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Toimittajat), *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC-2016)* (Sivut 3518-3522) http://www.lrec-conf.org/proceedings/lrec2016/pdf/62_Paper.pdf
110. Lison, P., & Tiedemann, J. (2016). OpenSubtitles2015: Extracting Large Parallel Corpora from Movie and TV Subtitles. teoksessa N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Toimittajat), *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC-2016)* http://www.lrec-conf.org/proceedings/lrec2016/pdf/947_Paper.pdf
111. 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
112. Tiedemann, J. (2015). Baseline Models for Pronoun Prediction and Pronoun-Aware Translation. teoksessa *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT)* (Sivut 108-114). The Association for Computational Linguistics. <http://aclweb.org/anthology/W15-2515>
113. Tiedemann, J., Ginter, F., & Kanerva, J. (2015). Morphological Segmentation and OPUS for Finnish-English Machine Translation. teoksessa *Proceedings of the Tenth Workshop on Statistical Machine Translation* (Sivut 177-183). The Association for Computational Linguistics. <http://aclweb.org/anthology/W15-3021>
114. Callin, J., Hardmeier, C., & Tiedemann, J. (2015). Part-of-Speech Driven Cross-Lingual Pronoun Prediction with Feed-Forward Neural Networks. teoksessa *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT)* (Sivut 59-64). The Association for Computational Linguistics. <http://aclweb.org/anthology/W15-2508>
115. 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. teoksessa *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT)*, (Sivut 1-16). The Association for Computational Linguistics. <http://aclweb.org/anthology/W15-2501>
116. Tiedemann, J. (2015). Cross-Lingual Dependency Parsing with Universal Dependencies and Predicted PoS Labels. teoksessa *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)* (Sivut 340-349). Uppsala University. <http://www.aclweb.org/anthology/W15-2137>
117. Shao, Y., Tiedemann, J., & Nivre, J. (2015). Boosting English-Chinese Machine Transliteration via High Quality Alignment and Multilingual Resources. teoksessa X. Duan, R. E. Banchs, M. Zhangs, H. Li, & A. Kumara (Toimittajat), *Proceedings of the Fifth Named Entity Workshop, joint with 53rd ACL and the 7th IJCNLP* (Sivut 56-60). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W15-3908>
118. Tiedemann, J. (2015). Improving the Cross-Lingual Projection of Syntactic Dependencies. teoksessa B. Megyesi (Toimittaja), *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)* (Sivut 191-199). (Linköping electronic conference proceedings; Nro 109), (NEALT Proceedings Series). Linköping University Electronic Press. <http://www.aclweb.org/anthology/W15-1824>

119. Agić, Ž., Tiedemann, J., Merkler, D., Krek, S., Dobrovoljc, K., & Može, S. (2014). Cross-lingual Dependency Parsing of Related Languages with Rich Morphosyntactic Tagsets. teoksessa *Language Technology for Closely Related Languages and Language Variants (LT4CloseLang)* (Sivut 13-24). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W14-4203>
120. Martinez Garcia, E., Tiedemann, J., España-Bonet, C., & Màrquez, L. (2014). Word's Vector Representations meet Machine Translation. teoksessa D. Wu, M. Carpuat, X. Carreras, & E. M. Vecchi (Toimittajat), *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation* (Sivut 132-134). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W14-4015>
121. Zampieri, M., Tan, L., Ljubešić, N., & Tiedemann, J. (2014). A Report on the DSL Shared Task 2014. teoksessa *Proceedings of the First Workshop on Applying NLP Tools to Similar Languages, Varieties and Dialects* (Sivut 58-67). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W14-5307>
122. Tiedemann, J. (2014). Rediscovering Annotation Projection for Cross-Lingual Parser Induction. teoksessa *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*
123. Skadiňš, R., Tiedemann, J., Rozis, R., & Deksnė, D. (2014). Billions of Parallel Words for Free: Building and Using the EU Bookshop Corpus. teoksessa *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC-2014)* (Sivut 1850-1855) <http://www.lrec-conf.org/proceedings/lrec2014/summaries/846.html>
124. Zampieri, M., Ljubešić, N., Tiedemann, J., & Tan, L. (2014). Merging Comparable Data Sources for the Discrimination of Similar Languages: The DSL Corpus Collection. teoksessa *Proceedings of the 7th Workshop on Building and Using Comparable Corpora (BUCC)* (Sivut 6-10)
125. Guillou, L., Hardmeier, C., Smith, A., Tiedemann, J., & Webber, B. (2014). ParCor 1.0: A Parallel Pronoun-Coreference Corpus to Support Statistical MT. teoksessa *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)* (Sivut 3191-3198) <http://www.lrec-conf.org/proceedings/lrec2014/summaries/298.html>
126. Hardmeier, C., Stymne, S., Smith, A., Tiedemann, J., & Nivre, J. (2014). Anaphora Models and Reordering for Phrase-Based SMT. teoksessa *Proceedings of the Ninth Workshop on Statistical Machine Translation* (Sivut 122-129). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W14-3312>
127. Stymne, S., Tiedemann, J., & Nivre, J. (2014). Estimating Word Alignment Quality for SMT Reordering Tasks. teoksessa *Proceedings of the Ninth Workshop on Statistical Machine Translation* (Sivut 275-286). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W14-3334.pdf>
128. Tiedemann, J. (2014). Improved Text Extraction from PDF Documents for Large-Scale Natural Language Processing. teoksessa A. Gelbukh (Toimittaja), *Computational Linguistics and Intelligent Text Processing* (Vuosikerta 1, Sivut 102-112). (Lecture Notes in Computer Science LNCS 8403). Springer.
129. Tiedemann, J., Agić, Ž., & Nivre, J. (2014). Treebank Translation for Cross-Lingual Parser Induction. teoksessa *Proceedings of the Eighteenth Conference on Computational Language Learning* (Sivut 130-140). The Association for Computational Linguistics.
130. Tiedemann, J. (2013). Tools for lexicographic use of parallel and comparable corpora. teoksessa R. H. Gouws, U. Heid, W. Schweickard, & H. E. Wiegand (Toimittajat), *An International Encyclopedia of Lexicography -- Supplementary Volume* (Vuosikerta 4, Sivut 1433-1444). (Handbücher zur Sprach- und Kommunikationswissenschaft). De Gruyter Mouton. <http://www.degruyter.com/view/product/175228>
131. Tiedemann, J., van der Plas, L., & Villada Moirón, B. (2013). Bitexts as Semantic Mirrors. teoksessa *Unknown host publication* <http://stp.lingfil.uu.se/~joerg/paper/bitext13.pdf>
132. Hardmeier, C., Tiedemann, J., & Nivre, J. (2013). Latent Anaphora Resolution for Cross-Lingual Pronoun Prediction. teoksessa *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* (Sivut 380-391). The Association for Computational Linguistics.
133. Hardmeier, C., Stymne, S., Tiedemann, J., & Nivre, J. (2013). Docent: A Document-Level Decoder for Phrase-Based Statistical Machine Translation. teoksessa *Unknown host publication* (Sivut 193-198) <http://www.aclweb.org/anthology/P13-4033>
134. Stymne, S., Hardmeier, C., Tiedemann, J., & Nivre, J. (2013). Feature Weight Optimization for Discourse-Level SMT. teoksessa *Unknown host publication* (Sivut 60-69). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W13-3308>
135. Webber, B., Popescu-Belis, A., Markert, K., & Tiedemann, J. (Toimittajat) (2013). *Proceedings of the Workshop on Discourse in Machine Translation*. The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W13-33>
136. Stymne, S., Hardmeier, C., Tiedemann, J., & Nivre, J. (2013). Tunable Distortion Limits and Corpus Cleaning for SMT. teoksessa *Proceedings of the Eighth Workshop on Statistical Machine Translation* (Sivut 225-231). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W13-2229>
137. Pettersson, E., Megyesi, B. B., & Tiedemann, J. (2013). An SMT approach to automatic annotation of historical text. teoksessa *Unknown host publication*

138. Stymne, S., Tiedemann, J., Hardmeier, C., & Nivre, J. (2013). Statistical Machine Translation with Readability Constraints. teoksessa *Unknown host publication* (Sivut 375-386)
139. Tiedemann, J., & Nakov, P. (2013). Analyzing the Use of Character-Level Translation with Sparse and Noisy Datasets. teoksessa *Unknown host publication* (Sivut 676-684)
140. 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>
141. Tiedemann, J., & Ljubesic, N. (2012). Efficient Discrimination Between Closely Related Languages. teoksessa *Proceedings of COLING 2012: technical papers* (Sivut 2619-2634). Indian Institute of Technology. <http://www.aclweb.org/anthology/C12-1160>
142. Nakov, P., & Tiedemann, J. (2012). Combining Word-Level and Character-Level Models for Machine Translation Between Closely-Related Languages. teoksessa *50th Annual Meeting of the Association for Computational Linguistics: Volume 2, Short papers : proceedings of the Conference : July 8 - 14, 2012, Jeju Island, Korea.* (Sivut 301-305). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/P12-2059>
143. Hardmeier, C., Nivre, J., & Tiedemann, J. (2012). Document-Wide Decoding for Phrase-Based Statistical Machine Translation. teoksessa *Unknown host publication* (Sivut 1179-1190) <http://www.aclweb.org/anthology/D12-1108>
144. Vasijevs, A., Skadis, R., & Tiedemann, J. (2012). LetsMT! Cloud-Based Platform for Do-It-Yourself Machine Translation. teoksessa M. Zhang (Toimittaja), *50th Annual Meeting of the Association for Computational Linguistics: ACL 2012 : proceedings of the System Demonstrations : 10 July 2012, Jeju Island, Korea* (Sivut 43-48). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/P12-3008>
145. Hardmeier, C., Nivre, J., & Tiedemann, J. (2012). Tree Kernels for Machine Translation Quality Estimation. teoksessa *Unknown host publication* (Sivut 109-113) <http://www.aclweb.org/anthology/W12-3112>
146. Tiedemann, J., Hansen, D. H., Offersgaard, L., Olsen, S., & Zumpe, M. (2012). A Distributed Resource Repository for Cloud-Based Machine Translation. teoksessa N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (Toimittajat), *Unknown host publication* (Sivut 2207-2213) <http://www.lrec-conf.org/proceedings/lrec2012/pdf/457Paper.pdf>
147. Kotzé, G., Vandeghinste, V., Martens, S., & Tiedemann, J. (2012). Large aligned treebanks for syntax-based machine translation. teoksessa N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (Toimittajat), *Unknown host publication* (Sivut 467-473) <http://www.lrec-conf.org/proceedings/lrec2012/pdf/924Paper.pdf>
148. Tiedemann, J. (2012). Parallel Data, Tools and Interfaces in OPUS. teoksessa N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (Toimittajat), *Unknown host publication* (Sivut 2214-2218) <http://www.lrec-conf.org/proceedings/lrec2012/pdf/463Paper.pdf>
149. Tiedemann, J. (2012). Character-Based Pivot Translations for Under-Resourced Languages and Domains. teoksessa *Unknown host publication* (Sivut 141-151) <http://www.aclweb.org/anthology/E12-1015>
150. Ehrentraut, C., Tiedemann, J., Dalanis, 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 . teoksessa *Proceedings of the Sixth Workshop on Analytics for Noisy Unstructured Text Data, 2012 Association for Computing Machinery.*
151. Tiedemann, J., Zumpe, M., & Schleussner, S. (2012). Opening OPUS for User Contributions. teoksessa *Unknown host publication*
152. 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. teoksessa P. Spyns, & J. Odijk (Toimittajat), *Essential Speech and Language Technology for Dutch: resources, tools and applications* (Theory and Applications of Natural Language Processing). Springer.
153. Hardmeier, C., Tiedemann, J., Saers, M., Kessler, F., & Kessler, F. (2011). The Uppsala-FBK systems at WMT 2011. teoksessa *Sixth Workshop on Statistical Machine Translation: WMT 2011 : proceedings of the Workshop : July 30–31, 2011* (Sivut 372-378). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W11-2144>
154. Plas, L. V. D., Tiedemann, J., & Fahmi, I. (2011). Automatic extraction of medical term variants from multilingual parallel translations. teoksessa A. Bosch, & G. Bouma (Toimittajat), *Interactive Multi-modal Question-Answering* (Theory and Applications of Natural Language Processing). Springer.
155. Tiedemann, J. (2011). *Bitext Alignment*. (Synthesis Lectures on Human Language Technologies; Vuosikerta 4, Nro 2). Morgan & Claypool publishers. <http://dx.doi.org/10.2200/S00367ED1V01Y201106HLT014>
156. Plas, L. V. D., Tiedemann, J., & Manguin, J.-L. (2011). Synonym acquisition across domains and languages. teoksessa V. Pallotta, A. Soro, & E. Vargiu (Toimittajat), *Advances in Distributed Agent-based Retrieval Tools* (Sivut 41-57). (Studies in Computational Intelligence; Vuosikerta 361). Springer.
157. van der Plas, L., & Tiedemann, J. (2010). Finding Medical Term Variations using Parallel Corpora and Distributional Similarity. teoksessa A. Oltramari, P. Vossen, & Q. Lu (Toimittajat), *Proceedings of OntoLex 2010: 6th Workshop on Ontologies and Lexical Resources* (Sivut 28-37). Coling 2010 Organizing Committee. <http://www.aclweb.org/anthology/W10-3304>

158. Tiedemann, J. (2010). Context Adaptation in Statistical Machine Translation Using Models with Exponentially Decaying Cache. teoksessa H. Daumé, T. Deoskar, D. McClosky, B. Plank, & J. Tiedemann (Toimittajat), *DANLP 2010: 2010 Workshop on Domain Adaptation for Natural Language Processing : proceedings of the Workshop : 15 July 2010, Uppsala University, Uppsala, Sweden.* (Sivut 8-15). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W10-2602>
159. Ill, H. D., Deoskar, T., McClosky, D., Plank, B., & Tiedemann, J. (Toimittajat) (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>
160. Tiedemann, J. (2010). To Cache or Not To Cache? Experiments with Adaptive Models in Statistical Machine Translation. teoksessa *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR* (Sivut 189-194). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W10-1728>
161. Islam, M. Z., Tiedemann, J., & Eisele, A. (2010). English to Bangla Phrase-Based Machine Translation. teoksessa F. Yvon, & V. Hansen (Toimittajat), *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>
162. Tiedemann, J. (2010). Lingua-Align: An Experimental Toolbox for Automatic Tree-to-Tree Alignment. teoksessa N. C. Chair, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, & D. Tapias (Toimittajat), *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, 17-23 May 2010, Valletta, Malta* (Sivut 736-743). European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2010/pdf/144Paper.pdf>
163. Plas, L. V. D., Tiedemann, J., & Manguin, J.-L. (2010). Automatic acquisition of synonyms for French using parallel corpora. teoksessa *Unknown host publication* <http://www.dart-project.org/dart2010/doku.php?id=program>
164. Tiedemann, J., & Weijnitz, P. (2010). Let's MT! A Platform for Sharing SMT Training Data. teoksessa *Unknown host publication*
165. Ahrenberg, L., Tiedemann, J., & Volk, M. (Toimittajat) (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>
166. Tiedemann, J., & Nabende, P. (2009). Translating Transliterations. *International journal of computing and ICT research* , 3(1), 33-41. <http://www.ijcir.org/specialissue2009/>
167. Tiedemann, J., & Kotzé, G. (2009). A Discriminative Approach to Tree Alignment. teoksessa *Unknown host publication* (Sivut 33-39) <http://www.aclweb.org/anthology/W09-4206>
168. Tiedemann, J. (2009). Evidence-Based Word Alignment. teoksessa I. Ilisei, V. Pekar, & S. Bernardini (Toimittajat), *Unknown host publication* (Sivut 28-32) <http://www.aclweb.org/anthology/W09-4205>
169. Tiedemann, J. (2009). Character-based PSMT for Closely Related Languages. teoksessa L. Marqués, & H. Somers (Toimittajat), *Unknown host publication* (Sivut 12 - 19) <http://www.mt-archive.info/EAMT-2009-Tiedemann-1.pdf>
170. Tiedemann, J. (2009). Translating Questions for Cross-Lingual QA. teoksessa L. Marqués, & H. Somers (Toimittajat), *Unknown host publication* (Sivut 112 - 119) <http://www.mt-archive.info/EAMT-2009-Tiedemann-2.pdf>
171. Tiedemann, J., & Kotzé, G. (2009). Building a Large Machine-Aligned Parallel Treebank. teoksessa *Unknown host publication*
172. Bertels, A., Fairon, C., Tiedemann, J., & Verlinde, S. (2009). Corpus parallèles et corpus ciblés au secours du dictionnaire de traduction. teoksessa *Cahiers de lexicologie* (Sivut 199-219). (Revue). Classiques Garnier.
173. Tiedemann, J. (2009). News from OPUS - A Collection of Multilingual Parallel Corpora with Tools and Interfaces. teoksessa N. Nicolov, K. Bontcheva, G. Angelova, & R. Mitkov (Toimittajat), *Recent Advances in Natural Language Processing* (Vuosisikerta V, Sivut 237-248)