

Jörg Tiedemann
Department of Modern Languages
Email: jorg.tiedemann@helsinki.fi
Phone: 02941



Qualifications

PhD (Computational Linguistics)
Uppsala University Uppsala Sweden
12 Dec 2003

M.Sc. (Computer Science)
Otto Von Guericke University, Magdeburg Germany
11 Sep 1997

Publications

Bootstrapping a Dependency Parser for Maltese - A Real-World Test Case

Tiedemann, J. & van der Plas, L. 12 Dec 2016 *From Semantics to Dialectometry: Festschrift in honor of John Nerbonne*. Weiling, M., Kroon, M. & Van Noord, G. (eds.). Milton Keynes: College Publications, Vol. 32, p. 355-365 11 p. (Tributes; vol. 32)

Discriminating between Similar Languages and Arabic Dialect Identification: A Report on the Third DSL Shared Task

Malmasi, S., Zampieri, M., Ljubešić, N., Nakov, P., Ali, A. & Tiedemann, J. 1 Dec 2016 p. 1-14 14 p.

Tagging Ingush - Language Technology For Low-Resource Languages Using Resources From Linguistic Field Work

Tiedemann, J., Nichols, J. & Sprouse, R. 1 Dec 2016 p. 148-155 8 p.

The Challenges of Multi-dimensional Sentiment Analysis Across Languages

Öhman, E., Honkela, T. & Tiedemann, J. 1 Dec 2016 p. 138-142 5 p.

Efficient word alignment with Markov Chain Monte Carlo

Östling, R. & Tiedemann, J. 1 Oct 2016 In : *The Prague Bulletin of Mathematical Linguistics*. 106, p. 125-146 22 p.

Large aligned treebanks for syntax-based machine translation (extended version)

Kotzé, G., Vandeghinste, V., Martens, S. & Tiedemann, J. 1 Oct 2016 In : *Language Resources and Evaluation*. p. 1-34

Detecting hospital-acquired infections: A document classification approach using support vector machines and gradient tree boosting

Ehrentraut, C., Ekholm, M., Tanushi, H., Tiedemann, J. & Dalianis, H. 4 Aug 2016 In : *Health informatics journal*.. p. 1-19 19 p.

A Linear Baseline Classifier for Cross-Lingual Pronoun Prediction

Tiedemann, J. 1 Aug 2016 p. 616-619 4 p.

Findings of the 2016 WMT Shared Task on Cross-lingual Pronoun Prediction

Guillou, L., Hardmeier, C., Nakov, P., Stymne, S., Tiedemann, J., Versley, Y., Cettolo, M., Webber, B. & Popescu-Belis, A. 1 Aug 2016 p. 525-542 18 p.

Phrase-Based SMT for Finnish with More Data, Better Models and Alternative Alignment and Translation Tools

Tiedemann, J., Cap, F., Kanerva, J., Ginter, F., Stymne, S., Östling, R. & Weller-Di Marco, M. 1 Aug 2016 p. 391-398 8 p.

Synthetic Treebanking for Cross-Lingual Dependency Parsing

Tiedemann, J. & Agi, Z. Jan 2016 In : Journal of Artificial Intelligence Research. 55, p. 209-248 40 p.

Climbing Mont BLEU: The Strange World of Reachable High-BLEU Translations

Smith, A., Hardmeier, C. & Tiedemann, J. 2016 In : Baltic Journal of Modern Computing. 4, 2, p. 269–281 13 p.

Finding Alternative Translations in a Large Corpus of Movie Subtitles

Tiedemann, J. 2016 p. 3518-3522 5 p.

OPUS – Parallel Corpora for Everyone

Tiedemann, J. 2016 In : Baltic Journal of Modern Computing. p. 384 1 p.

OpenSubtitles2015: Extracting Large Parallel Corpora from Movie and TV Subtitles

Lison, P. & Tiedemann, J. 2016

Baseline Models for Pronoun Prediction and Pronoun-Aware Translation

Tiedemann, J. 1 Sep 2015 p. 108-114 7 p.

Morphological Segmentation and OPUS for Finnish-English Machine Translation

Tiedemann, J., Ginter, F. & Kanerva, J. 1 Sep 2015 p. 177-183 7 p.

Part-of-Speech Driven Cross-Lingual Pronoun Prediction with Feed-Forward Neural Networks

Callin, J., Hardmeier, C. & Tiedemann, J. 1 Sep 2015 p. 59-64 6 p.

Pronoun-Focused MT and Cross-Lingual Pronoun Prediction: Findings of the 2015 DiscoMT Shared Task on Pronoun Translation

Hardmeier, C., Nakov, P., Stymne, S., Tiedemann, J., Versley, Y. & Cettolo, M. 1 Sep 2015 p. 1-16 16 p.

Cross-Lingual Dependency Parsing with Universal Dependencies and Predicted PoS Labels

Tiedemann, J. 1 Aug 2015 p. 340-349 10 p.

Boosting English-Chinese Machine Transliteration via High Quality Alignment and Multilingual Resources

Shao, Y., Tiedemann, J. & Nivre, J. 1 Jul 2015 p. 56-60 5 p.

Improving the Cross-Lingual Projection of Syntactic Dependencies

Tiedemann, J. 1 May 2015 p. 191-199 9 p.

Cross-lingual Dependency Parsing of Related Languages with Rich Morphosyntactic Tagsets

Agić, Ž., Tiedemann, J., Merkler, D., Krek, S., Dobrovoljc, K. & Može, S. 1 Oct 2014 p. 13-24 12 p.

Word's Vector Representations meet Machine Translation

Martinez Garcia, E., Tiedemann, J., España-Bonet, C. & Màrquez, L. 1 Oct 2014 p. 132-134 3 p.

A Report on the DSL Shared Task 2014

Zampieri, M., Tan, L., Ljubešić, N. & Tiedemann, J. 1 Aug 2014 p. 58-67 10 p.

Rediscovering Annotation Projection for Cross-Lingual Parser Induction

Tiedemann, J. 1 Aug 2014

Billions of Parallel Words for Free: Building and Using the EU Bookshop Corpus
Skadiņš, R., Tiedemann, J., Rozis, R. & Deksne, D. 1 May 2014 p. 1850-1855 6 p.

Merging Comparable Data Sources for the Discrimination of Similar Languages: The DSL Corpus Collection
Zampieri, M., Ljubešić, N., Tiedemann, J. & Tan, L. 1 May 2014 p. 6-10 5 p.

ParCor 1.0: A Parallel Pronoun-Coreference Corpus to Support Statistical MT
Guillou, L., Hardmeier, C., Smith, A., Tiedemann, J. & Webber, B. 1 May 2014 p. 3191-3198 8 p.

Anaphora Models and Reordering for Phrase-Based SMT
Hardmeier, C., Stymne, S., Smith, A., Tiedemann, J. & Nivre, J. 2014 p. 122-129 8 p.

Estimating Word Alignment Quality for SMT Reordering Tasks
Stymne, S., Tiedemann, J. & Nivre, J. 2014 p. 275-286 12 p.

Improved Text Extraction from PDF Documents for Large-Scale Natural Language Processing
Tiedemann, J. 2014 *Computational Linguistics and Intelligent Text Processing*. Gelbukh, A. (ed.). Berlin: Springer, Vol. 1, p. 102-112 11 p. (Lecture Notes in Computer Science LNCS 8403)

Treebank Translation for Cross-Lingual Parser Induction
Tiedemann, J., Agić, Ž. & Nivre, J. 2014 p. 130-140 11 p.

Tools for lexicographic use of parallel and comparable corpora
Tiedemann, J. 1 Dec 2013 *An International Encyclopedia of Lexicography -- Supplementary Volume*. Gouws, R. H., Heid, U., Schweickard, W. & Wiegand, H. E. (eds.). De Gruyter Mouton, Vol. 4, p. 1433-1444 12 p. (HSK - Dictionaries)

Bitexts as Semantic Mirrors
Tiedemann, J., van der Plas, L. & Villada Moirón, B. 1 Oct 2013

Latent Anaphora Resolution for Cross-Lingual Pronoun Prediction
Hardmeier, C., Tiedemann, J. & Nivre, J. 1 Oct 2013

Docent: A Document-Level Decoder for Phrase-Based Statistical Machine Translation
Hardmeier, C., Stymne, S., Tiedemann, J. & Nivre, J. 1 Aug 2013 p. 193-198 6 p.

Feature Weight Optimization for Discourse-Level SMT
Stymne, S., Hardmeier, C., Tiedemann, J. & Nivre, J. 1 Aug 2013 p. 60-69 10 p.

Proceedings of the Workshop on Discourse in Machine Translation
Webber, B. (ed.), Popescu-Belis, A. (ed.), Markert, K. (ed.) & Tiedemann, J. (ed.) 1 Aug 2013 Sofia, Bulgaria: Association for Computational Linguistics.

Tunable Distortion Limits and Corpus Cleaning for SMT
Stymne, S., Hardmeier, C., Tiedemann, J. & Nivre, J. 1 Aug 2013 p. 225-231 7 p.

An SMT approach to automatic annotation of historical text
Pettersson, E., Megyesi, B. B. & Tiedemann, J. 1 May 2013

Statistical Machine Translation with Readability Constraints
Stymne, S., Tiedemann, J., Hardmeier, C. & Nivre, J. 1 May 2013 p. 375-386 12 p.

Analyzing the Use of Character-Level Translation with Sparse and Noisy Datasets

Tiedemann, J. & Nakov, P. 2013 p. 676-684 9 p.

Book review of Markus Dickinson, Chris Brew and Detmar Meurers: Language and Computers

Tiedemann, J. 2013 In : Machine Translation. 27, p. 309-312 4 p.

Efficient Discrimination Between Closely Related Languages

Tiedemann, J. & Ljubesi, N. 1 Dec 2012 p. 2619-2634 16 p.

Combining Word-Level and Character-Level Models for Machine Translation Between Closely-Related Languages

Nakov, P. & Tiedemann, J. 1 Jul 2012 p. 301-305 5 p.

Document-Wide Decoding for Phrase-Based Statistical Machine Translation

Hardmeier, C., Nivre, J. & Tiedemann, J. 1 Jul 2012 p. 1179-1190 12 p.

LetsMT!: Cloud-Based Platform for Do-It-Yourself Machine Translation

Vasijevs, A., Skadis, R. & Tiedemann, J. 1 Jul 2012 p. 43-48 6 p.

Tree Kernels for Machine Translation Quality Estimation

Hardmeier, C., Nivre, J. & Tiedemann, J. 1 Jun 2012 p. 109-113 5 p.

A Distributed Resource Repository for Cloud-Based Machine Translation

Tiedemann, J., Hansen, D. H., Offersgaard, L., Olsen, S. & Zumpe, M. 1 May 2012 p. 2207-2213 7 p.

Large aligned treebanks for syntax-based machine translation

Kotzé, G., Vandeghinste, V., Martens, S. & Tiedemann, J. 1 May 2012 p. 467-473 7 p.

Parallel Data, Tools and Interfaces in OPUS

Tiedemann, J. 1 May 2012 p. 2214-2218 5 p.

Character-Based Pivot Translations for Under-Resourced Languages and Domains

Tiedemann, J. 1 Apr 2012 p. 141-151 11 p.

Detection of Hospital Acquired Infections in Sparse and Noisy Swedish Patient Records

Ehrentraut, C., Tiedemann, J., Dalianis, H. & Tanushi, H. 2012

Opening OPUS for User Contributions

Tiedemann, J., Zumpe, M. & Schleussner, S. 2012

Parse and Corpus-based Machine Translation

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

The Uppsala-FBK systems at WMT 2011

Hardmeier, C., Tiedemann, J., Saers, M., Federico, M. & Prashant, M. 1 Jul 2011 p. 372-378 7 p.

Automatic extraction of medical term variants from multilingual parallel translations

Plas, L. V. D., Tiedemann, J. & Fahmi, I. 2011 *Theory and Applications of Natural Language Processing*. 21 p. (Theory and Applications of Natural Language Processing)

Bitext Alignment

Tiedemann, J. & Hirst, G. (ed.) 2011 San Rafael, California (USA): Morgan & Claypool. (Synthesis Lectures on Human Language Technologies)

Synonym acquisition across domains and languages

Plas, L. V. D., Tiedemann, J. & Manguin, J-L. 2011 *Advances in Distributed Agent-based Retrieval Tools*. Berlin: Springer , (Studies in Computational Intelligence)

Finding Medical Term Variations using Parallel Corpora and Distributional Similarity

van der Plas, L. & Tiedemann, J. 1 Aug 2010 p. 28-37 10 p.

Context Adaptation in Statistical Machine Translation Using Models with Exponentially Decaying Cache

Tiedemann, J. 1 Jul 2010 p. 8-15 8 p.

Proceedings of the 2010 Workshop on Domain Adaptation for Natural Language Processing

III, H. D. (ed.), Deoskar, T. (ed.), McClosky, D. (ed.), Plank, B. (ed.) & Tiedemann, J. (ed.) 1 Jul 2010 Uppsala, Sweden: Association for Computational Linguistics.

To Cache or Not To Cache? Experiments with Adaptive Models in Statistical Machine Translation

Tiedemann, J. 1 Jul 2010 p. 189-194 6 p.

English to Bangla Phrase-Based Machine Translation

Islam, M. Z., Tiedemann, J. & Eisele, A. 1 May 2010

Lingua-Align: An Experimental Toolbox for Automatic Tree-to-Tree Alignment

Tiedemann, J. 1 May 2010

Automatic acquisition of synonyms for French using parallel corpora

Plas, L. V. D., Tiedemann, J. & Manguin, J-L. 2010

Let's MT! - A Platform for Sharing SMT Training Data

Tiedemann, J. & Weijnitz, P. 2010

Proceedings of the Workshop on Annotation and Exploitation of Parallel Corpora

Ahrenberg, L. (ed.), Tiedemann, J. (ed.) & Volk, M. (ed.) 2010 Northern European Association for Language Technology . (NEALT Proceedings Series)

Translating Transliterations

Tiedemann, J. & Nabende, P. 1 Oct 2009 In : International journal of computing and ICT research . 3, 1, p. 33-41 9 p.

A Discriminative Approach to Tree Alignment

Tiedemann, J. & Kotzé, G. 1 Sep 2009 p. 33-39 7 p.

Evidence-Based Word Alignment

Tiedemann, J. 1 Sep 2009 p. 28-32 5 p.

Character-based PSMT for Closely Related Languages

Tiedemann, J. 1 May 2009 p. 12 - 19

Translating Questions for Cross-Lingual QA

Tiedemann, J. 1 May 2009 p. 112 - 119

Building a Large Machine-Aligned Parallel Treebank

Tiedemann, J. & Kotzé, G. 2009 12 p.

Corpus parallèles et corpus ciblés au secours du dictionnaire de traduction

Bertels, A., Fairon, C., Tiedemann, J. & Verlinde, S. 2009 *Cahiers de lexicologie*. Classiques Garnier, p. 199-219 21 p. (Revue)

News from OPUS - A Collection of Multilingual Parallel Corpora with Tools and Interfaces

Tiedemann, J. 2009 *Recent Advances in Natural Language Processing*. Nicolov, N., Bontcheva, K., Angelova, G. & Mitkov, R. (eds.). Vol. V, p. 237-248 12 p.