

First Conference on Machine Translation (WMT 2016)

Held at ACL 2016

Berlin, Germany
7 – 12 August 2016

ISBN: 978-1-5108-2772-1

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2016) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2016)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Cross-language Projection of Dependency Trees with Constrained Partial Parsing for Tree-to-Tree Machine Translation</i>	
Yu Shen, Chenhui Chu, Fabien Cromieres and Sadao Kurohashi	1
<i>Improving Pronoun Translation by Modeling Coreference Uncertainty</i>	
Ngoc Quang Luong and Andrei Popescu-Belis	12
<i>Modeling verbal inflection for English to German SMT</i>	
Anita Ramm and Alexander Fraser	21
<i>Modeling Selectional Preferences of Verbs and Nouns in String-to-Tree Machine Translation</i>	
Maria Nadejde, Alexandra Birch and Philipp Koehn	32
<i>Modeling Complement Types in Phrase-Based SMT</i>	
Marion Weller-Di Marco, Alexander Fraser and Sabine Schulte im Walde	43
<i>Alignment-Based Neural Machine Translation</i>	
Tamer Alkhouli, Gabriel Bretschner, Jan-Thorsten Peter, Mohammed Hethnawi, Andreas Guta and Hermann Ney	54
<i>Neural Network-based Word Alignment through Score Aggregation</i>	
Joël Legrand, Michael Auli and Ronan Collobert	66
<i>Using Factored Word Representation in Neural Network Language Models</i>	
Jan Niehues, Thanh-Le Ha, Eunah Cho and Alex Waibel	74
<i>Linguistic Input Features Improve Neural Machine Translation</i>	
Rico Sennrich and Barry Haddow	83
<i>A Framework for Discriminative Rule Selection in Hierarchical Moses</i>	
Fabienne Braune, Alexander Fraser, Hal Daumé III and Aleš Tamchyna	92
<i>Fast and highly parallelizable phrase table for statistical machine translation</i>	
Nikolay Bogoychev and Hieu Hoang	102
<i>A Comparative Study on Vocabulary Reduction for Phrase Table Smoothing</i>	
Yunsu Kim, Andreas Guta, Joern Wuebker and Hermann Ney	110
<i>Examining the Relationship between Preordering and Word Order Freedom in Machine Translation</i>	
Joachim Daiber, Miloš Stanojević, Wilker Aziz and Khalil Sima'an	118
<i>Findings of the 2016 Conference on Machine Translation</i>	
Ondřej Bojar, Rajen Chatterjee, Christian Federmann, Yvette Graham, Barry Haddow, Matthias Huck, Antonio Jimeno Yepes, Philipp Koehn, Varvara Logacheva, Christof Monz, Matteo Negri, Aurelie Neveol, Mariana Neves, Martin Popel, Matt Post, Raphael Rubino, Carolina Scarton, Lucia Specia, Marco Turchi, Karin Verspoor and Marcos Zampieri	131
<i>Results of the WMT16 Metrics Shared Task</i>	
Ondřej Bojar, Yvette Graham, Amir Kamran and Miloš Stanojević	199
<i>Results of the WMT16 Tuning Shared Task</i>	
Bushra Jawaid, Amir Kamran, Miloš Stanojević and Ondřej Bojar	232

<i>LIMSI@WMT'16: Machine Translation of News</i>	
Alexandre Allauzen, Lauriane Aufrant, Franck Burlot, Ophélie Lacroix, Elena Knyazeva, Thomas Lavergne, Guillaume Wisniewski and François Yvon	239
<i>TÜBİTAK SMT System Submission for WMT2016</i>	
Emre Bektaş, Ertugrul Yilmaz, Coskun Mermer and İlknur Durgar El-Kahlout	246
<i>ParFDA for Instance Selection for Statistical Machine Translation</i>	
Ergun Bicici	252
<i>Sheffield Systems for the English-Romanian WMT Translation Task</i>	
Frédéric Blain, Xingyi Song and Lucia Specia	259
<i>MetaMind Neural Machine Translation System for WMT 2016</i>	
James Bradbury and Richard Socher	264
<i>NYU-MILA Neural Machine Translation Systems for WMT'16</i>	
Junyoung Chung, Kyunghyun Cho and Yoshua Bengio	268
<i>The JHU Machine Translation Systems for WMT 2016</i>	
Shuoyang Ding, Kevin Duh, Huda Khayrallah, Philipp Koehn and Matt Post	272
<i>Yandex School of Data Analysis approach to English-Turkish translation at WMT16 News Translation Task</i>	
Anton Dvorkovich, Sergey Gubanov and Irina Galinskaya	281
<i>Hybrid Morphological Segmentation for Phrase-Based Machine Translation</i>	
Stig-Arne Grönroos, Sami Virpioja and Mikko Kurimo	289
<i>The AFRL-MITLL WMT16 News-Translation Task Systems</i>	
Jeremy Gwinnup, Tim Anderson, Grant Erdmann, Katherine Young, Michael Kazi, Elizabeth Salesky and Brian Thompson	296
<i>The Karlsruhe Institute of Technology Systems for the News Translation Task in WMT 2016</i>	
Thanh-Le Ha, Eunah Cho, Jan Niehues, Mohammed Mediani, Matthias Sperber, Alexandre Allauzen and Alexander Waibel	303
<i>The Edinburgh/LMU Hierarchical Machine Translation System for WMT 2016</i>	
Matthias Huck, Alexander Fraser and Barry Haddow	311
<i>The AMU-UEDIN Submission to the WMT16 News Translation Task: Attention-based NMT Models as Feature Functions in Phrase-based SMT</i>	
Marcin Junczys-Dowmunt, Tomasz Dwojak and Rico Sennrich	319
<i>NRC Russian-English Machine Translation System for WMT 2016</i>	
Chi-kiu Lo, Colin Cherry, George Foster, Darlene Stewart, Rabib Islam, Anna Kazantseva and Roland Kuhn	326
<i>Merged bilingual trees based on Universal Dependencies in Machine Translation</i>	
David Mareček	333
<i>PROMT Translation Systems for WMT 2016 Translation Tasks</i>	
Alexander Molchanov and Fedor Bykov	339

<i>The QT21/HimL Combined Machine Translation System</i>	
Jan-Thorsten Peter, Tamer Alkhouli, Hermann Ney, Matthias Huck, Fabienne Braune, Alexander Fraser, Aleš Tamchyna, Ondřej Bojar, Barry Haddow, Rico Sennrich, Frédéric Blain, Lucia Specia, Jan Niehues, Alex Waibel, Alexandre Allauzen, Lauriane Aufrant, Franck Burlot, elena knyazeva, Thomas Lavergne, François Yvon, Mārcis Pinnis and Stella Frank	344
<i>The RWTH Aachen University English-Romanian Machine Translation System for WMT 2016</i>	
Jan-Thorsten Peter, Tamer Alkhouli, Andreas Guta and Hermann Ney	356
<i>Abu-MaTran at WMT 2016 Translation Task: Deep Learning, Morphological Segmentation and Tuning on Character Sequences</i>	
Víctor M. Sánchez-Cartagena and Antonio Toral	362
<i>Edinburgh Neural Machine Translation Systems for WMT 16</i>	
Rico Sennrich, Barry Haddow and Alexandra Birch	371
<i>The Edit Distance Transducer in Action: The University of Cambridge English-German System at WMT16</i>	
Felix Stahlberg, Eva Hasler and Bill Byrne	377
<i>CUNI-LMU Submissions in WMT2016: Chimera Constrained and Beaten</i>	
Aleš Tamchyna, Roman Sudarikov, Ondřej Bojar and Alexander Fraser	385
<i>Phrase-Based SMT for Finnish with More Data, Better Models and Alternative Alignment and Translation Tools</i>	
Jörg Tiedemann, Fabienne Cap, Jenna Kanerva, Filip Ginter, Sara Stymne, Robert Östling and Marion Weller-Di Marco	391
<i>Edinburgh's Statistical Machine Translation Systems for WMT16</i>	
Philip Williams, Rico Sennrich, Maria Nadejde, Matthias Huck, Barry Haddow and Ondřej Bojar	399
<i>PJAiT Systems for the WMT 2016</i>	
Krzysztof Wolk and Krzysztof Marasek	411
<i>DFKI's system for WMT16 IT-domain task, including analysis of systematic errors</i>	
Eleftherios Avramidis	415
<i>ILLC-UvA Adaptation System (Scorpio) at WMT'16 IT-DOMAIN Task</i>	
Hoang Cuong, Stella Frank and Khalil Sima'an	423
<i>Data Selection for IT Texts using Paragraph Vector</i>	
Mirela-Stefania Duma and Wolfgang Menzel	428
<i>SMT and Hybrid systems of the QTLep project in the WMT16 IT-task</i>	
Rosa Gaudio, Gorka Labaka, Eneko Agirre, Petya Osenova, Kiril Simov, Martin Popel, Dieke Oele, Gertjan van Noord, Luís Gomes, João António Rodrigues, Steven Neale, João Silva, Andreia Querido, Nuno Rendeiro and António Branco	435
<i>JU-USAAR: A Domain Adaptive MT System</i>	
Koushik Pahari, Alapan Kuila, Santanu Pal, Sudip Kumar Naskar, Sivaji Bandyopadhyay and Josef van Genabith	442
<i>Dictionary-based Domain Adaptation of MT Systems without Retraining</i>	
Rudolf Rosa, Roman Sudarikov, Michal Novák, Martin Popel and Ondřej Bojar	449

<i>English-Portuguese Biomedical Translation Task Using a Genuine Phrase-Based Statistical Machine Translation Approach</i>	
José Aires, Gabriel Lopes and Luís Gomes	456
<i>The TALP–UPC Spanish–English WMT Biomedical Task: Bilingual Embeddings and Char-based Neural Language Model Rescoring in a Phrase-based System</i>	
Marta R. Costa-jussà, Cristina España-Bonet, Pranava Madhyastha, Carlos Escolano and José A. R. Fonollosa	463
<i>LIMSI’s Contribution to the WMT’16 Biomedical Translation Task</i>	
Julia Ive, Aurélien Max and François Yvon	469
<i>IXA Biomedical Translation System at WMT16 Biomedical Translation Task</i>	
Olatz Perez-de-Viñaspre and Gorka Labaka	477
<i>CobaltF: A Fluent Metric for MT Evaluation</i>	
Marina Fomicheva, Núria Bel, Lucia Specia, Iria da Cunha and Anton Malinovskiy	483
<i>DTED: Evaluation of Machine Translation Structure Using Dependency Parsing and Tree Edit Distance</i>	
Martin McCaffery and Mark-Jan Nederhof	491
<i>chrF deconstructed: beta parameters and n-gram weights</i>	
Maja Popović	499
<i>CharacTer: Translation Edit Rate on Character Level</i>	
Weiyue Wang, Jan-Thorsten Peter, Hendrik Rosendahl and Hermann Ney	505
<i>Extract Domain-specific Paraphrase from Monolingual Corpus for Automatic Evaluation of Machine Translation</i>	
Lilin Zhang, Zhen Weng, Wenyan Xiao, Jianyi Wan, Zhiming Chen, Yiming Tan, Maoxi Li and Mingwen Wang	511
<i>Particle Swarm Optimization Submission for WMT16 Tuning Task</i>	
Viktor Kocur and Ondřej Bojar	518
<i>Findings of the 2016 WMT Shared Task on Cross-lingual Pronoun Prediction</i>	
Liane Guillou, Christian Hardmeier, Preslav Nakov, Sara Stymne, Jörg Tiedemann, Yannick Versley, Mauro Cettolo, Bonnie Webber and Andrei Popescu-Belis	525
<i>A Shared Task on Multimodal Machine Translation and Crosslingual Image Description</i>	
Lucia Specia, Stella Frank, Khalil Sima’an and Desmond Elliott	543
<i>Findings of the WMT 2016 Bilingual Document Alignment Shared Task</i>	
Christian Buck and Philipp Koehn	554
<i>Cross-lingual Pronoun Prediction with Linguistically Informed Features</i>	
Rachel Bawden	564
<i>The Kyoto University Cross-Lingual Pronoun Translation System</i>	
Raj Dabre, Yevgeniy Puzikov, Fabien Cromieres and Sadao Kurohashi	571
<i>Pronoun Prediction with Latent Anaphora Resolution</i>	
Christian Hardmeier	576
<i>It-disambiguation and source-aware language models for cross-lingual pronoun prediction</i>	
Sharid Loáiciga, Liane Guillou and Christian Hardmeier	581

<i>Pronoun Language Model and Grammatical Heuristics for Aiding Pronoun Prediction</i> Ngoc Quang Luong and Andrei Popescu-Belis	589
<i>Cross-Lingual Pronoun Prediction with Deep Recurrent Neural Networks</i> Juhani Luotolahti, Jenna Kanerva and Filip Ginter	596
<i>Pronoun Prediction with Linguistic Features and Example Weighing</i> Michal Novák	602
<i>Feature Exploration for Cross-Lingual Pronoun Prediction</i> Sara Stymne	609
<i>A Linear Baseline Classifier for Cross-Lingual Pronoun Prediction</i> Jörg Tiedemann	616
<i>Cross-lingual Pronoun Prediction for English, French and German with Maximum Entropy Classification</i> Dominikus Wetzel	620
<i>Does Multimodality Help Human and Machine for Translation and Image Captioning?</i> Ozan Caglayan, Walid Aransa, Yaxing Wang, Marc Masana, Mercedes García-Martínez, Fethi Bougares, Loïc Barrault and Joost van de Weijer	627
<i>DCU-UvA Multimodal MT System Report</i> Iacer Calixto, Desmond Elliott and Stella Frank	634
<i>Attention-based Multimodal Neural Machine Translation</i> Po-Yao Huang, Frederick Liu, Sz-Rung Shiang, Jean Oh and Chris Dyer	639
<i>CUNI System for WMT16 Automatic Post-Editing and Multimodal Translation Tasks</i> Jindřich Libovický, Jindřich Helcl, Marek Tlustý, Ondřej Bojar and Pavel Pecina	646
<i>WMT 2016 Multimodal Translation System Description based on Bidirectional Recurrent Neural Networks with Double-Embeddings</i> Sergio Rodríguez Guasch and Marta R. Costa-jussà	655
<i>SHEF-Multimodal: Grounding Machine Translation on Images</i> Kashif Shah, Josiah Wang and Lucia Specia	660
<i>DOCAL - Vicomtech's Participation in the WMT16 Shared Task on Bilingual Document Alignment</i> Andoni Azpeitia and Thierry Etchegoyhen	666
<i>Quick and Reliable Document Alignment via TF/IDF-weighted Cosine Distance</i> Christian Buck and Philipp Koehn	672
<i>YODA System for WMT16 Shared Task: Bilingual Document Alignment</i> Aswarth Abhilash Dara and Yiu-Chang Lin	679
<i>Bitextor's participation in WMT'16: shared task on document alignment</i> Miquel Esplà-Gomis, Mikel Forcada, Sergio Ortiz Rojas and Jorge Ferrández-Tordera	685
<i>Bilingual Document Alignment with Latent Semantic Indexing</i> Ulrich Germann	692
<i>First Steps Towards Coverage-Based Document Alignment</i> Luís Gomes and Gabriel Pereira Lopes	697

<i>BAD LUC@WMT 2016: a Bilingual Document Alignment Platform Based on Lucene</i> Laurent Jakubina and Phillippe Langlais	703
<i>Using Term Position Similarity and Language Modeling for Bilingual Document Alignment</i> Thanh Le, Hoa Trong Vu, Jonathan Oberländer and Ondřej Bojar	710
<i>The ADAPT Bilingual Document Alignment system at WMT16</i> Pintu Lohar, Haithem Affi, Chao-Hong Liu and Andy Way	717
<i>WMT2016: A Hybrid Approach to Bilingual Document Alignment</i> Sainik Mahata, Dipankar Das and Santanu Pal	724
<i>English-French Document Alignment Based on Keywords and Statistical Translation</i> Marek Medved', Miloš Jakubíček and Vojtech Kovár	728
<i>The ILSP/ARC submission to the WMT 2016 Bilingual Document Alignment Shared Task</i> Vassilis Papavassiliou, Prokopis Prokopidis and Stelios Piperidis	733
<i>Word Clustering Approach to Bilingual Document Alignment (WMT 2016 Shared Task)</i> Vadim Shchukin, Dmitry Khristich and Irina Galinskaya	740
<i>The FBK Participation in the WMT 2016 Automatic Post-editing Shared Task</i> Rajen Chatterjee, José G. C. de Souza, Matteo Negri and Marco Turchi	745
<i>Log-linear Combinations of Monolingual and Bilingual Neural Machine Translation Models for Automatic Post-Editing</i> Marcin Junczys-Dowmunt and Roman Grundkiewicz	751
<i>USAAR: An Operation Sequential Model for Automatic Statistical Post-Editing</i> Santanu Pal, Marcos Zampieri and Josef van Genabith	759
<i>Bilingual Embeddings and Word Alignments for Translation Quality Estimation</i> Amal Abdelsalam, Ondřej Bojar and Samhaa El-Beltagy	764
<i>SHEF-MIME: Word-level Quality Estimation Using Imitation Learning</i> Daniel Beck, Andreas Vlachos, Gustavo Paetzold and Lucia Specia	772
<i>Referential Translation Machines for Predicting Translation Performance</i> Ergun Bicici	777
<i>UAlacant word-level and phrase-level machine translation quality estimation systems at WMT 2016</i> Miquel Esplà-Gomis, Felipe Sánchez-Martínez and Mikel Forcada	782
<i>Recurrent Neural Network based Translation Quality Estimation</i> Hyun Kim and Jong-Hyeok Lee	787
<i>YSDA Participation in the WMT'16 Quality Estimation Shared Task</i> Anna Kozlova, Mariya Shmatova and Anton Frolov	793
<i>USFD's Phrase-level Quality Estimation Systems</i> Varvara Logacheva, Frédéric Blain and Lucia Specia	800
<i>Unbabel's Participation in the WMT16 Word-Level Translation Quality Estimation Shared Task</i> André F. T. Martins, Ramón Astudillo, Chris Hokamp and Fabio Kepler	806

<i>SimpleNets: Quality Estimation with Resource-Light Neural Networks</i> Gustavo Paetzold and Lucia Specia	812
<i>Translation Quality Estimation using Recurrent Neural Network</i> Raj Nath Patel and Sasikumar M	819
<i>The UU Submission to the Machine Translation Quality Estimation Task</i> Oscar Sagemo and Sara Stymne	825
<i>Word embeddings and discourse information for Quality Estimation</i> Carolina Scarton, Daniel Beck, Kashif Shah, Karin Sim Smith and Lucia Specia	831
<i>SHEF-LIUM-NN: Sentence level Quality Estimation with Neural Network Features</i> Kashif Shah, Fethi Bougares, Loïc Barrault and Lucia Specia	838
<i>UGENT-LT3 SCATE Submission for WMT16 Shared Task on Quality Estimation</i> Arda Tezcan, Véronique Hoste and Lieve Macken	843