

# **10th Workshop on Statistical Machine Translation 2015**

Held at EMNLP 2015

Lisbon, Portugal  
17 – 18 September 2015

ISBN: 978-1-5108-1188-1

**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© (2015) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2015)

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](mailto: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: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Findings of the 2015 Workshop on Statistical Machine Translation</i> Ondřej Bojar, Rajen Chatterjee, Christian Federmann, Barry Haddow, Matthias Huck, Chris Hokamp, Philipp Koehn, Varvara Logacheva, Christof Monz, Matteo Negri, Matt Post, Carolina Scarton, Lucia Specia and Marco Turchi . . . . .	1
<i>Statistical Machine Translation with Automatic Identification of Translationese</i> Naama Twitto, Noam Ordan and Shuly Wintner . . . . .	47
<i>Data Selection With Fewer Words</i> Amittai Axelrod, Philip Resnik, Xiaodong He and Mari Ostendorf . . . . .	58
<i>DFKI's experimental hybrid MT system for WMT 2015</i> Eleftherios Avramidis, Maja Popović and Aljoscha Burchardt . . . . .	66
<i>ParFDA for Fast Deployment of Accurate Statistical Machine Translation Systems, Benchmarks, and Statistics</i> Ergun Bicici, Qun Liu and Andy Way . . . . .	74
<i>CUNI in WMT15: Chimera Strikes Again</i> Ondřej Bojar and Aleš Tamchyna . . . . .	79
<i>CimS - The CIS and IMS Joint Submission to WMT 2015 addressing morphological and syntactic differences in English to German SMT</i> Fabienne Cap, Marion Weller, Anita Ramm and Alexander Fraser . . . . .	84
<i>The Karlsruhe Institute of Technology Translation Systems for the WMT 2015</i> Eunah Cho, Thanh-Le Ha, Jan Niehues, Teresa Herrmann, Mohammed Mediani, Yuqi Zhang and Alex Waibel . . . . .	92
<i>New Language Pairs in TectoMT</i> Ondřej Dušek, Luís Gomes, Michal Novák, Martin Popel and Rudolf Rosa . . . . .	98
<i>Tuning Phrase-Based Segmented Translation for a Morphologically Complex Target Language</i> Stig-Arne Grönroos, Sami Virpioja and Mikko Kurimo . . . . .	105
<i>The AFRL-MITLL WMT15 System: There's More than One Way to Decode It!</i> Jeremy Gwinnup, Tim Anderson, Grant Erdmann, Katherine Young, Christina May, Michael Kazi, Elizabeth Salesky and Brian Thompson . . . . .	112
<i>The KIT-LIMSI Translation System for WMT 2015</i> Thanh-Le Ha, Quoc-Khanh DO, Eunah Cho, Jan Niehues, Alexandre Allauzen, François Yvon and Alex Waibel . . . . .	120
<i>The Edinburgh/JHU Phrase-based Machine Translation Systems for WMT 2015</i> Barry Haddow, Matthias Huck, Alexandra Birch, Nikolay Bogoychev and Philipp Koehn . . . . .	126
<i>Montreal Neural Machine Translation Systems for WMT'15</i> Sébastien Jean, Orhan Firat, Kyunghyun Cho, Roland Memisevic and Yoshua Bengio . . . . .	134
<i>GF Wide-coverage English-Finnish MT system for WMT 2015</i> Prasanth Kolachina and Aarne Ranta . . . . .	141

<i>LIMSI@WMT'15 : Translation Task</i>	
Benjamin Marie, Alexandre Allauzen, Franck Burlot, Quoc-Khanh Do, Julia Ive, elena knyazeva, Matthieu Labeau, Thomas Lavergne, Kevin Löser, Nicolas Pécheux and François Yvon . . . . .	145
<i>UdS-Sant: English–German Hybrid Machine Translation System</i>	
Santanu Pal, Sudip Naskar and Josef van Genabith . . . . .	152
<i>The RWTH Aachen German-English Machine Translation System for WMT 2015</i>	
Jan-Thorsten Peter, Farzad Toutouchi, Joern Wuebker and Hermann Ney . . . . .	158
<i>Exact Decoding with Multi Bottom-Up Tree Transducers</i>	
Daniel Quernheim . . . . .	164
<i>Sheffield Systems for the Finnish-English WMT Translation Task</i>	
David Steele, Karin Sim Smith and Lucia Specia . . . . .	172
<i>Morphological Segmentation and OPUS for Finnish-English Machine Translation</i>	
Jörg Tiedemann, Filip Ginter and Jenna Kanerva . . . . .	177
<i>Abu-MaTran at WMT 2015 Translation Task: Morphological Segmentation and Web Crawling</i>	
Raphael Rubino, Tommi Pirinen, Miquel Esplà-Gomis, Nikola Ljubešić, Sergio Ortiz Rojas, Vasilis Papavassiliou, Prokopis Prokopidis and Antonio Toral . . . . .	184
<i>The University of Illinois submission to the WMT 2015 Shared Translation Task</i>	
Lane Schwartz, Bill Bryce, Chase Geigle, Sean Massung, Yisi Liu, Haoruo Peng, Vignesh Raja, Subhro Roy and Shyam Upadhyay . . . . .	192
<i>Edinburgh's Syntax-Based Systems at WMT 2015</i>	
Philip Williams, Rico Sennrich, Maria Nadejde, Matthias Huck and Philipp Koehn . . . . .	199
<i>The FBK Participation in the WMT15 Automatic Post-editing Shared Task</i>	
Rajen Chatterjee, Marco Turchi and Matteo Negri . . . . .	210
<i>USAAR-SAPE: An English–Spanish Statistical Automatic Post-Editing System</i>	
Santanu Pal, Mihaela Vela, Sudip Kumar Naskar and Josef van Genabith . . . . .	216
<i>Why Predicting Post-Edition is so Hard? Failure Analysis of LIMSI Submission to the APE Shared Task</i>	
Guillaume Wisniewski, Nicolas Pécheux and François Yvon . . . . .	222
<i>Hierarchical Machine Translation With Discontinuous Phrases</i>	
Miriam Kaeshammer . . . . .	228
<i>Discontinuous Statistical Machine Translation with Target-Side Dependency Syntax</i>	
Nina Seemann and Andreas Maletti . . . . .	239
<i>ListNet-based MT Rescoring</i>	
Jan Niehues, Quoc-Khanh DO, Alexandre Allauzen and Alex Waibel . . . . .	248
<i>Results of the WMT15 Metrics Shared Task</i>	
Miloš Stanojević, Amir Kamran, Philipp Koehn and Ondřej Bojar . . . . .	256
<i>Results of the WMT15 Tuning Shared Task</i>	
Miloš Stanojević, Amir Kamran and Ondřej Bojar . . . . .	274
<i>Extended Translation Models in Phrase-based Decoding</i>	
Andreas Guta, Joern Wuebker, Miguel Graca, Yunsu Kim and Hermann Ney . . . . .	282

<i>Investigations on Phrase-based Decoding with Recurrent Neural Network Language and Translation Models</i>	
Tamer Alkhouli, Felix Rietig and Hermann Ney .....	294
<i>Referential Translation Machines for Predicting Translation Quality and Related Statistics</i>	
Ergun Bicici, Qun Liu and Andy Way .....	304
<i>UAlacant word-level machine translation quality estimation system at WMT 2015</i>	
Miquel Esplà-Gomis, Felipe Sánchez-Martínez and Mikel Forcada .....	309
<i>QUality Estimation from ScraTCH (QUETCH): Deep Learning for Word-level Translation Quality Estimation</i>	
Julia Kreutzer, Shigehiko Schamoni and Stefan Riezler .....	316
<i>LORIA System for the WMT15 Quality Estimation Shared Task</i>	
David Langlois .....	323
<i>Data enhancement and selection strategies for the word-level Quality Estimation</i>	
Varvara Logacheva, Chris Hokamp and Lucia Specia .....	330
<i>USHEF and USAAR-USHEF participation in the WMT15 QE shared task</i>	
Carolina Scarton, Liling Tan and Lucia Specia .....	336
<i>SHEF-NN: Translation Quality Estimation with Neural Networks</i>	
Kashif Shah, Varvara Logacheva, Gustavo Paetzold, Frédéric Blain, Daniel Beck, Fethi Bougares and Lucia Specia .....	342
<i>Strategy-Based Technology for Estimating MT Quality</i>	
Liugang Shang, Dongfeng Cai and Duo Ji .....	348
<i>UGENT-LT3 SCATE System for Machine Translation Quality Estimation</i>	
Arda Tezcan, Veronique Hoste, Bart Desmet and Lieve Macken .....	353
<i>Multi-level Evaluation for Machine Translation</i>	
Boxing Chen, Hongyu Guo and Roland Kuhn .....	361
<i>VERTa: a Linguistically-motivated Metric at the WMT15 Metrics Task</i>	
Elisabet Comelles and Jordi Atserias .....	366
<i>UPF-Cobalt Submission to WMT15 Metrics Task</i>	
Marina Fomicheva, Núria Bel, Iria da Cunha and Anton Malinovskiy .....	373
<i>Machine Translation Evaluation using Recurrent Neural Networks</i>	
Rohit Gupta, Constantin Orasan and Josef van Genabith .....	380
<i>Alignment-based sense selection in METEOR and the RATATOUILLE recipe</i>	
Benjamin Marie and Marianna Apidianaki .....	385
<i>chrF: character n-gram F-score for automatic MT evaluation</i>	
Maja Popović .....	392
<i>BEER 1.1: ILLC UvA submission to metrics and tuning task</i>	
Miloš Stanojević and Khalil Sima'an .....	396
<i>Predicting Machine Translation Adequacy with Document Embeddings</i>	
Mihaela Vela and Liling Tan .....	402

<i>LeBLEU: N-gram-based Translation Evaluation Score for Morphologically Complex Languages</i> Sami Virpioja and Stig-Arne Grönroos .....	411
<i>CASICT-DCU Participation in WMT2015 Metrics Task</i> Hui Yu, Qingsong Ma, Xiaofeng Wu and Qun Liu .....	417
<i>Drem: The AFRL Submission to the WMT15 Tuning Task</i> Grant Erdmann and Jeremy Gwinnup .....	422
<i>MT Tuning on RED: A Dependency-Based Evaluation Metric</i> Liangyou Li, Hui Yu and Qun Liu .....	428
<i>Improving evaluation and optimization of MT systems against MEANT</i> Chi-kiu Lo, Philipp Dowling and Dekai Wu .....	434
<i>An Investigation of Machine Translation Evaluation Metrics in Cross-lingual Question Answering</i> Kyoshiro Sugiyama, Masahiro Mizukami, Graham Neubig, Koichiro Yoshino, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura .....	442
<i>Dependency Analysis of Scrambled References for Better Evaluation of Japanese Translation</i> Hideki Isozaki and Natsume Kouchi .....	450
<i>How do Humans Evaluate Machine Translation</i> Francisco Guzmán, Ahmed Abdelali, Irina Temnikova, Hassan Sajjad and Stephan Vogel .....	457
<i>Local System Voting Feature for Machine Translation System Combination</i> Markus Freitag, Jan-Thorsten Peter, Stephan Peitz, Minwei Feng and Hermann Ney .....	467