

WMT 2013

**8th Workshop
on
Statistical Machine Translation**

Proceedings of the Workshop

**August 8-9, 2013
Sofia, Bulgaria**

Production and Manufacturing by
Omnipress, Inc.
2600 Anderson Street
Madison, WI 53707
USA

Shared Tasks supported by the following EU Framework Programme 7 projects:

- MosesCore
- CASMACAT
- Matecat
- QTLauchPad

Additional funding provided by Microsoft Research.

©2013 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN: 978-1-937284-57-2

Preface

The ACL 2013 Workshop on Statistical Machine Translation (WMT 2013) took place on Thursday and Friday, August 8–9, 2013 in Sofia, Bulgaria, immediately following the Conference of the Association for Computational Linguistics (ACL).

This is the eighth time this workshop has been held. The first time it was held at HLT-NAACL 2006 in New York City, USA. In the following years the Workshop on Statistical Machine Translation was held at ACL 2007 in Prague, Czech Republic, ACL 2008, Columbus, Ohio, USA, EACL 2009 in Athens, Greece, ACL 2010 in Uppsala, Sweden, EMNLP 2011 in Edinburgh, Scotland, and NAACL 2012 in Montréal, Canada.

The focus of our workshop was to use parallel corpora for machine translation. Recent experimentation has shown that the performance of SMT systems varies greatly with the source language. In this workshop we encouraged researchers to investigate ways to improve the performance of SMT systems for diverse languages, including morphologically more complex languages, languages with partial free word order, and low-resource languages.

Prior to the workshop, in addition to soliciting relevant papers for review and possible presentation, we conducted three shared tasks: a translation task, a quality estimation task, and a task to test automatic evaluation metrics. The results of the shared tasks were announced at the workshop, and these proceedings also include an overview paper for the shared tasks that summarizes the results, as well as provides information about the data used and any procedures that were followed in conducting or scoring the task. In addition, there are short papers from each participating team that describe their underlying system in greater detail.

Like in previous years, we have received a far larger number of submission than we could accept for presentation. This year we have received 32 full paper submissions and 46 shared task submissions. In total WMT-2013 featured 18 full paper oral presentations and 45 shared task poster presentations.

The invited talk was given by Andreas Eisele (Directorate-General for Translation at the European Commission, Luxembourg) entitled “Machine Translation at the European Commission: Serving the multilingual needs of the European Commission”.

We would like to thank the members of the Program Committee for their timely reviews. We also would like to thank the participants of the shared task and all the other volunteers who helped with the evaluations.

Ondřej Bojar, Christian Buck, Chris Callison-Burch, Barry Haddow, Philipp Koehn, Christof Monz, Matt Post, Hervé Saint-Amand, Radu Soricut, and Lucia Specia

WMT 5-year Retrospective Best Paper Award

Each year WMT awards a 5-year Retrospective Best Paper Award. This year we selected the best paper from 2008's Workshop on Statistical Machine Translation, which was collocated with ACL in Columbus, Ohio. The goals of this retrospective award are to recognize high-quality work that has stood the test of time, and to highlight the excellent work that appears at WMT.

37 members of the WMT13 program committee voted on the best paper from a list of seven nominated papers. These were nominated by selecting the papers with the most non-self-citations in the ACL anthology network. This year the vote was very close, and was divided between several excellent papers. Ultimately, the program committee decided to award the WMT 5-year Retrospective Best Paper Award to:

Kevin Gimpel and Noah A. Smith. 2008. *Rich Source-Side Context for Statistical Machine Translation*. In Proceedings of the Workshop on Statistical Machine Translation. Pages 9-17.

In this paper, Gimpel and Smith used a variety of features, including surrounding words and part-of-speech tags, local syntactic structure, and other properties of the source language sentence to help predict each phrase's translation. They argued that source side features were easier to exploit than target side features, and that they were likely to make a bigger impact, since some target side features are already exploited via the language model. Gimpel and Smith empirically demonstrated the value of their model by augmenting the baseline Moses MT system and fielding an entry into the English-to-German shared task at WMT that year.

One of the program committee members, Preslav Nakov, commented that this work made an important contribution in the direction of context-aware SMT, which has been largely neglected in mainstream SMT research.

Congratulations to Kevin Gimpel and Noah Smith on their excellent work!

Organizers:

Ondřej Bojar (Charles University)
Christian Buck (University of Edinburgh)
Chris Callison-Burch (Johns Hopkins University)
Barry Haddow (University of Edinburgh)
Philipp Koehn (University of Edinburgh)
Christof Monz (University of Amsterdam)
Matt Post (Johns Hopkins University)
Hervé Saint-Amand (University of Edinburgh)
Radu Soricut (Google)
Lucia Specia (University of Sheffield)

Invited Talk:

Andreas Eisele (European Commission)

Program Committee:

Lars Ahrenberg (Linköping University)
Eleftherios Avramidis (German Research Center for Artificial Intelligence (DFKI))
Daniel Beck (University of Sheffield)
Nicola Bertoldi (FBK)
Arianna Bisazza (Fondazione Bruno Kessler)
Graeme Blackwood (IBM Research)
Phil Blunsom (University of Oxford)
Chris Brockett (Microsoft Research)
Bill Byrne (University of Cambridge)
Nicola Cancedda (Xerox Research Centre Europe)
Hailong Cao (Harbin Institute of Technology)
Marine Carpuat (National Research Council)
Francisco Casacuberta (Universitat Politècnica de Valencia)
Daniel Cer (Stanford University)
Boxing Chen (NRC)
Colin Cherry (NRC)
David Chiang (USC/ISI)
Steve DeNeeffe (SDL Language Weaver)
John DeNero (Google)
Michael Denkowski (Carnegie Mellon University)
Markus Dreyer (SDL Language Weaver)
Kevin Duh (Nara Institute of Science and Technology)
Chris Dyer (Carnegie Mellon University)
Marc Dymetman (Xerox Research Centre Europe)
Stefano Faralli (Sapienza University of Rome)

Yang Feng (University of Sheffield)
Andrew Finch (NICT)
José A. R. Fonollosa (Universitat Politècnica de Catalunya)
Mikel Forcada (Universitat d'Alacant)
George Foster (NRC)
Alexander Fraser (University of Stuttgart)
Katya Garmash (University of Amsterdam)
Niyu Ge (IBM Research)
Ulrich Germann (University of Edinburgh)
Daniel Gildea (University of Rochester)
Cyril Goutte (National Research Council Canada)
Nizar Habash (Columbia University)
Jan Hajic (Charles University in Prague)
Keith Hall (Google Research)
Greg Hanneman (Carnegie Mellon University)
Christian Hardmeier (Uppsala universitet)
Xiaodong He (Microsoft Research)
Yifan He (New York University)
Kenneth Heafield (Carnegie Mellon University, University of Edinburgh)
John Henderson (MITRE)
Silja Hildebrand (CMU)
Hieu Hoang (University of Edinburgh)
Young-Sook Hwang (SKPlanet)
Gonzalo Iglesias (University of Cambridge)
Abe Ittycheriah (IBM)
Doug Jones (MIT Lincoln Laboratory)
Maxim Khalilov (TAUS Labs)
Roland Kuhn (National Research Council of Canada)
Shankar Kumar (Google)
Mathias Lambert (Amazon.com)
Qun Liu (Dublin City University)
Wolfgang Macherey (Google)
Daniel Marcu (SDL)
José B. Mariño (Polytechnic University of Catalonia)
Cettolo Mauro (FBK)
Arne Mauser (Google, Inc)
Shachar Mirkin (Xerox Research Centre Europe)
Dragos Munteanu (SDL Language Technologies)
Markos Mylonakis (Xerox Research Centre Europe)
Preslav Nakov (Qatar Computing Research Institute, Qatar Foundation)
Kemal Oflazer (Carnegie Mellon University - Qatar)
Sergio Penkale (Lingo24)
Chris Quirk (Microsoft Research)
Stefan Riezler (Heidelberg University)
Johann Roturier (Symantec)
Anoop Sarkar (Simon Fraser University)
Holger Schwenk (University of Le Mans)

Jean Senellart (SYSTRAN)
Hendra Setiawan (IBM T.J. Watson Research Center)
Kashif Shah (University of Sheffield)
Wade Shen (MIT)
Linfeng Song (ICT/CAS)
Felipe Sánchez-Martínez (Universitat d'Alacant)
Joerg Tiedemann (Uppsala University)
Christoph Tillmann (IBM Research)
Dan Tufis (Research Institute for Artificial Intelligence, Romanian Academy)
Masao Utiyama (NICT)
Josef van Genabith (Dublin City University)
David Vilar (DFKI)
Haifeng Wang (Baidu)
Taro Watanabe (NICT)
Hua Wu (Baidu)
Peng Xu (Google Inc.)
Francois Yvon (LIMSI/CNRS)
Richard Zens (Google)
Hao Zhang (Google)
Joy Ying Zhang (Carnegie Mellon University)
Liu Zhanyi (Baidu)

Table of Contents

<i>Findings of the 2013 Workshop on Statistical Machine Translation</i> Ondřej Bojar, Christian Buck, Chris Callison-Burch, Christian Federmann, Barry Haddow, Philipp Koehn, Christof Monz, Matt Post, Radu Soricut and Lucia Specia	1
<i>Results of the WMT13 Metrics Shared Task</i> Matouš Macháček and Ondřej Bojar	45
<i>The Feasibility of HMEANT as a Human MT Evaluation Metric</i> Alexandra Birch, Barry Haddow, Ulrich Germann, Maria Nadejde, Christian Buck and Philipp Koehn	52
<i>LIMSI @ WMT13</i> Alexander Allauzen, Nicolas Pécheux, Quoc Khanh Do, Marco Dinarelli, Thomas Lavergne, Aurélien Max, Hai-Son Le and François Yvon	62
<i>The CMU Machine Translation Systems at WMT 2013: Syntax, Synthetic Translation Options, and Pseudo-References</i> Waleed Ammar, Victor Chahuneau, Michael Denkowski, Greg Hanneman, Wang Ling, Austin Matthews, Kenton Murray, Nicola Segall, Alon Lavie and Chris Dyer	70
<i>Feature Decay Algorithms for Fast Deployment of Accurate Statistical Machine Translation Systems</i> Ergun Biciçi	78
<i>CUni Multilingual Matrix in the WMT 2013 Shared Task</i> Karel Bílek and Daniel Zeman	85
<i>Chimera – Three Heads for English-to-Czech Translation</i> Ondřej Bojar, Rudolf Rosa and Aleš Tamchyna	92
<i>Yandex School of Data Analysis Machine Translation Systems for WMT13</i> Alexey Borisov, Jacob Dlougach and Irina Galinskaya	99
<i>The Karlsruhe Institute of Technology Translation Systems for the WMT 2013</i> Eunah Cho, Thanh-Le Ha, Mohammed Mediani, Jan Niehues, Teresa Herrmann, Isabel Slawik and Alex Waibel	104
<i>TÜBİTAK-BİLGEM German-English Machine Translation Systems for W13</i> Ilknur Durgar El-Kahlout and Coşkun Mermer	109
<i>Edinburgh’s Machine Translation Systems for European Language Pairs</i> Nadir Durrani, Barry Haddow, Kenneth Heafield and Philipp Koehn	114
<i>Munich-Edinburgh-Stuttgart Submissions of OSM Systems at WMT13</i> Nadir Durrani, Alexander Fraser, Helmut Schmid, Hassan Sajjad and Richárd Farkas	122
<i>Towards Efficient Large-Scale Feature-Rich Statistical Machine Translation</i> Vladimir Eidelman, Ke Wu, Ferhan Ture, Philip Resnik and Jimmy Lin	128
<i>The TALP-UPC Phrase-Based Translation Systems for WMT13: System Combination with Morphology Generation, Domain Adaptation and Corpus Filtering</i> Lluís Formiga, Marta R. Costa-jussà, José B. Mariño, José A. R. Fonollosa, Alberto Barrón-Cedeño and Lluís Marquez	134

<i>PhraseFix: Statistical Post-Editing of TectoMT</i>	
Petra Galuščáková, Martin Popel and Ondřej Bojar	141
<i>Feature-Rich Phrase-based Translation: Stanford University's Submission to the WMT 2013 Translation Task</i>	
Spence Green, Daniel Cer, Kevin Reschke, Rob Voigt, John Bauer, Sida Wang, Natalia Silveira, Julia Neidert and Christopher D. Manning	148
<i>Factored Machine Translation Systems for Russian-English</i>	
Stéphane Huet, Elena Manishina and Fabrice Lefèvre	154
<i>Omnifluent English-to-French and Russian-to-English Systems for the 2013 Workshop on Statistical Machine Translation</i>	
Evgeny Matusov and Gregor Leusch	158
<i>Pre-Reordering for Machine Translation Using Transition-Based Walks on Dependency Parse Trees</i>	
Antonio Valerio Miceli Barone and Giuseppe Attardi	164
<i>Edinburgh's Syntax-Based Machine Translation Systems</i>	
Maria Nadejde, Philip Williams and Philipp Koehn	170
<i>Shallow Semantically-Informed PBSMT and HPBSMT</i>	
Tsuyoshi Okita, Qun Liu and Josef van Genabith	177
<i>Joint WMT 2013 Submission of the QUAERO Project</i>	
Stephan Peitz, Saab Mansour, Matthias Huck, Markus Freitag, Hermann Ney, Eunah Cho, Teresa Herrmann, Mohammed Mediani, Jan Niehues, Alex Waibel, Alexander Allauzen, Quoc Khanh Do, Bianka Buschbeck and Tonio Wandmacher	185
<i>The RWTH Aachen Machine Translation System for WMT 2013</i>	
Stephan Peitz, Saab Mansour, Jan-Thorsten Peter, Christoph Schmidt, Joern Wuebker, Matthias Huck, Markus Freitag and Hermann Ney	193
<i>The University of Cambridge Russian-English System at WMT13</i>	
Juan Pino, Aurelien Waite, Tong Xiao, Adrià de Gispert, Federico Flego and William Byrne ..	200
<i>Joshua 5.0: Sparser, Better, Faster, Server</i>	
Matt Post, Juri Ganitkevitch, Luke Orland, Jonathan Weese, Yuan Cao and Chris Callison-Burch	206
<i>The CNGL-DCU-Prompsit Translation Systems for WMT13</i>	
Raphael Rubino, Antonio Toral, Santiago Cortés Vaíllo, Jun Xie, Xiaofeng Wu, Stephen Doherty and Qun Liu	213
<i>QCRI-MES Submission at WMT13: Using Transliteration Mining to Improve Statistical Machine Translation</i>	
Hassan Sajjad, Svetlana Smekalova, Nadir Durrani, Alexander Fraser and Helmut Schmid	219
<i>Tunable Distortion Limits and Corpus Cleaning for SMT</i>	
Sara Stymne, Christian Hardmeier, Jörg Tiedemann and Joakim Nivre	225
<i>Munich-Edinburgh-Stuttgart Submissions at WMT13: Morphological and Syntactic Processing for SMT</i>	
Marion Weller, Max Kisselew, Svetlana Smekalova, Alexander Fraser, Helmut Schmid, Nadir Durrani, Hassan Sajjad and Richárd Farkas	232

<i>Coping with the Subjectivity of Human Judgements in MT Quality Estimation</i> Marco Turchi, Matteo Negri and Marcello Federico	240
<i>Online Polylingual Topic Models for Fast Document Translation Detection</i> Kriste Krstovski and David A. Smith	252
<i>Combining Bilingual and Comparable Corpora for Low Resource Machine Translation</i> Ann Irvine and Chris Callison-Burch	262
<i>Generating English Determiners in Phrase-Based Translation with Synthetic Translation Options</i> Yulia Tsvetkov, Chris Dyer, Lori Levin and Archna Bhatia	271
<i>Dramatically Reducing Training Data Size Through Vocabulary Saturation</i> William Lewis and Sauleh Eetemadi	281
<i>Multi-Task Learning for Improved Discriminative Training in SMT</i> Patrick Simianer and Stefan Riezler	292
<i>Online Learning Approaches in Computer Assisted Translation</i> Prashant Mathur, Cettolo Mauro and Marcello Federico	301
<i>Length-Incremental Phrase Training for SMT</i> Joern Wuebker and Hermann Ney	309
<i>Positive Diversity Tuning for Machine Translation System Combination</i> Daniel Cer, Christopher D. Manning and Dan Jurafsky	320
<i>Selecting Feature Sets for Comparative and Time-Oriented Quality Estimation of Machine Translation Output</i> Eleftherios Avramidis and Maja Popovic	329
<i>SHEF-Lite: When Less is More for Translation Quality Estimation</i> Daniel Beck, Kashif Shah, Trevor Cohn and Lucia Specia	337
<i>Referential Translation Machines for Quality Estimation</i> Ergun Bicici	343
<i>FBK-UEdin Participation to the WMT13 Quality Estimation Shared Task</i> José Guilherme Camargo de Souza, Christian Buck, Marco Turchi and Matteo Negri	352
<i>The TALP-UPC Approach to System Selection: Asiya Features and Pairwise Classification Using Random Forests</i> Lluís Formiga, Meritxell González, Alberto Barrón-Cedeño, José A. R. Fonollosa and Lluís Marquez	359
<i>Quality Estimation for Machine Translation Using the Joint Method of Evaluation Criteria and Statistical Modeling</i> Aaron Li-Feng Han, Yi Lu, Derek F. Wong, Lidia S. Chao, Liangye He and Junwen Xing	365
<i>MT Quality Estimation: The CMU System for WMT'13</i> Silja Hildebrand and Stephan Vogel	373
<i>LORIA System for the WMT13 Quality Estimation Shared Task</i> David Langlois and Kamel Smaili	380

<i>LIG System for WMT13 QE Task: Investigating the Usefulness of Features in Word Confidence Estimation for MT</i>	
Ngoc Quang Luong, Benjamin Lecouteux and Laurent Besacier	386
<i>DCU-Symantec at the WMT 2013 Quality Estimation Shared Task</i>	
Raphael Rubino, Joachim Wagner, Jennifer Foster, Johann Roturier, Rasoul Samad Zadeh Kaljahi and Fred Hollowood	392
<i>LIMSI Submission for the WMT'13 Quality Estimation Task: an Experiment with N-Gram Posteriors</i>	
Anil Kumar Singh, Guillaume Wisniewski and François Yvon	398
<i>Ranking Translations using Error Analysis and Quality Estimation</i>	
Mark Fishel	405
<i>Are ACT's Scores Increasing with Better Translation Quality?</i>	
Najeh Hajlaoui	408
<i>A Description of Tunable Machine Translation Evaluation Systems in WMT13 Metrics Task</i>	
Aaron Li-Feng Han, Derek F. Wong, Lidia S. Chao, Yi Lu, Liangye He, Yiming Wang and Jiaji Zhou	414
<i>MEANT at WMT 2013: A Tunable, Accurate yet Inexpensive Semantic Frame Based MT Evaluation Metric</i>	
Chi-kiu Lo and Dekai Wu	422
<i>An Approach Using Style Classification Features for Quality Estimation</i>	
Erwan Moreau and Raphael Rubino	429
<i>DCU Participation in WMT2013 Metrics Task</i>	
Xiaofeng Wu, Hui Yu and Qun Liu	435
<i>Efficient Solutions for Word Reordering in German-English Phrase-Based Statistical Machine Translation</i>	
Arianna Bisazza and Marcello Federico	440
<i>A Phrase Orientation Model for Hierarchical Machine Translation</i>	
Matthias Huck, Joern Wuebker, Felix Rietig and Hermann Ney	452
<i>A Dependency-Constrained Hierarchical Model with Moses</i>	
Yvette Graham	464
<i>Investigations in Exact Inference for Hierarchical Translation</i>	
Wilker Aziz, Marc Dymetman and Sriram Venkatapathy	472
<i>Evaluating (and Improving) Sentence Alignment under Noisy Conditions</i>	
Omar Zaidan and Vishal Chowdhary	484
<i>Multi-Rate HMMs for Word Alignment</i>	
Elif Eyigöz, Daniel Gildea and Kemal Oflazer	494
<i>Hidden Markov Tree Model for Word Alignment</i>	
Shuhei Kondo, Kevin Duh and Yuji Matsumoto	503
<i>An MT Error-Driven Discriminative Word Lexicon using Sentence Structure Features</i>	
Jan Niehues and Alex Waibel	512

Conference Program

Thursday, August 8, 2013

9:00–9:10 Opening Remarks

Session 1: Shared Tasks and their Evaluation

9:10–10:10 *Findings of the 2013 Workshop on Statistical Machine Translation*
Ondřej Bojar, Christian Buck, Chris Callison-Burch, Christian Federmann, Barry Haddow, Philipp Koehn, Christof Monz, Matt Post, Radu Soricut and Lucia Specia

Results of the WMT13 Metrics Shared Task
Matouš Macháček and Ondřej Bojar

10:10–10:30 *The Feasibility of HMEANT as a Human MT Evaluation Metric*
Alexandra Birch, Barry Haddow, Ulrich Germann, Maria Nadejde, Christian Buck and Philipp Koehn

10:30–11:00 Coffee

Session 2: Poster Session

11:00–12:30 Shared Task: Translation

LIMSI @ WMT13
Alexander Allauzen, Nicolas Pécheux, Quoc Khanh Do, Marco Dinarelli, Thomas Lavergne, Aurélien Max, Hai-Son Le and François Yvon

The CMU Machine Translation Systems at WMT 2013: Syntax, Synthetic Translation Options, and Pseudo-References
Waleed Ammar, Victor Chahuneau, Michael Denkowski, Greg Hanneman, Wang Ling, Austin Matthews, Kenton Murray, Nicola Segall, Alon Lavie and Chris Dyer

Feature Decay Algorithms for Fast Deployment of Accurate Statistical Machine Translation Systems
Ergun Biciçi

CUni Multilingual Matrix in the WMT 2013 Shared Task
Karel Bílek and Daniel Zeman

Chimera – Three Heads for English-to-Czech Translation
Ondřej Bojar, Rudolf Rosa and Aleš Tamchyna

Thursday, August 8, 2013 (continued)

Yandex School of Data Analysis Machine Translation Systems for WMT13

Alexey Borisov, Jacob Dlougach and Irina Galinskaya

The Karlsruhe Institute of Technology Translation Systems for the WMT 2013

Eunah Cho, Thanh-Le Ha, Mohammed Mediani, Jan Niehues, Teresa Herrmann, Isabel Slawik and Alex Waibel

TÜBİTAK-BİLGEM German-English Machine Translation Systems for W13

Ilknur Durgar El-Kahlout and Coşkun Mermer

Edinburgh's Machine Translation Systems for European Language Pairs

Nadir Durrani, Barry Haddow, Kenneth Heafield and Philipp Koehn

Munich-Edinburgh-Stuttgart Submissions of OSM Systems at WMT13

Nadir Durrani, Alexander Fraser, Helmut Schmid, Hassan Sajjad and Richárd Farkas

Towards Efficient Large-Scale Feature-Rich Statistical Machine Translation

Vladimir Eidelman, Ke Wu, Ferhan Ture, Philip Resnik and Jimmy Lin

The TALP-UPC Phrase-Based Translation Systems for WMT13: System Combination with Morphology Generation, Domain Adaptation and Corpus Filtering

Lluís Formiga, Marta R. Costa-jussà, José B. Mariño, José A. R. Fonollosa, Alberto Barrón-Cedeño and Lluís Marquez

PhraseFix: Statistical Post-Editing of TectoMT

Petra Galuščáková, Martin Popel and Ondřej Bojar

Feature-Rich Phrase-based Translation: Stanford University's Submission to the WMT 2013 Translation Task

Spence Green, Daniel Cer, Kevin Reschke, Rob Voigt, John Bauer, Sida Wang, Natalia Silveira, Julia Neidert and Christopher D. Manning

Factored Machine Translation Systems for Russian-English

Stéphane Huet, Elena Manishina and Fabrice Lefèvre

Omnifluent English-to-French and Russian-to-English Systems for the 2013 Workshop on Statistical Machine Translation

Evgeny Matusov and Gregor Leusch

Pre-Reordering for Machine Translation Using Transition-Based Walks on Dependency Parse Trees

Antonio Valerio Miceli Barone and Giuseppe Attardi

Thursday, August 8, 2013 (continued)

Edinburgh's Syntax-Based Machine Translation Systems

Maria Nadejde, Philip Williams and Philipp Koehn

Shallow Semantically-Informed PBSMT and HPBSMT

Tsuyoshi Okita, Qun Liu and Josef van Genabith

Joint WMT 2013 Submission of the QUAERO Project

Stephan Peitz, Saab Mansour, Matthias Huck, Markus Freitag, Hermann Ney, Eunah Cho, Teresa Herrmann, Mohammed Mediani, Jan Niehues, Alex Waibel, Alexander Allauzen, Quoc Khanh Do, Bianka Buschbeck and Tonio Wandmacher

The RWTH Aachen Machine Translation System for WMT 2013

Stephan Peitz, Saab Mansour, Jan-Thorsten Peter, Christoph Schmidt, Joern Wuebker, Matthias Huck, Markus Freitag and Hermann Ney

The University of Cambridge Russian-English System at WMT13

Juan Pino, Aurelien Waite, Tong Xiao, Adrià de Gispert, Federico Flego and William Byrne

Joshua 5.0: Sparser, Better, Faster, Server

Matt Post, Juri Ganitkevitch, Luke Orland, Jonathan Weese, Yuan Cao and Chris Callison-Burch

The CNGL-DCU-Prompsit Translation Systems for WMT13

Raphael Rubino, Antonio Toral, Santiago Cortés Vaíllo, Jun Xie, Xiaofeng Wu, Stephen Doherty and Qun Liu

QCRI-MES Submission at WMT13: Using Transliteration Mining to Improve Statistical Machine Translation

Hassan Sajjad, Svetlana Smekalova, Nadir Durrani, Alexander Fraser and Helmut Schmid

Tunable Distortion Limits and Corpus Cleaning for SMT

Sara Stymne, Christian Hardmeier, Jörg Tiedemann and Joakim Nivre

Munich-Edinburgh-Stuttgart Submissions at WMT13: Morphological and Syntactic Processing for SMT

Marion Weller, Max Kisselew, Svetlana Smekalova, Alexander Fraser, Helmut Schmid, Nadir Durrani, Hassan Sajjad and Richárd Farkas

12:30–14:00 Lunch Break

Thursday, August 8, 2013 (continued)

Session 3: Invited Talk

14:00–15:10 Andreas Eisele: MT@EC: Serving the multilingual needs of the European Commission

Session 4: Quality Estimation

15:10–15:30 *Coping with the Subjectivity of Human Judgements in MT Quality Estimation*
Marco Turchi, Matteo Negri and Marcello Federico

15:30–16:00 Coffee Break

Session 5: Translation Models

16:00–16:20 *Online Polylingual Topic Models for Fast Document Translation Detection*
Kriste Krstovski and David A. Smith

16:20–16:40 *Combining Bilingual and Comparable Corpora for Low Resource Machine Translation*
Ann Irvine and Chris Callison-Burch

16:40–17:00 *Generating English Determiners in Phrase-Based Translation with Synthetic Translation Options*
Yulia Tsvetkov, Chris Dyer, Lori Levin and Archana Bhatia

17:00–17:20 *Dramatically Reducing Training Data Size Through Vocabulary Saturation*
William Lewis and Sauleh Eetemadi

Friday, August 9, 2013

Friday, August 9, 2013 (continued)

Session 6: Learning

- 9:00–9:20 *Multi-Task Learning for Improved Discriminative Training in SMT*
Patrick Simianer and Stefan Riezler
- 9:20–9:40 *Online Learning Approaches in Computer Assisted Translation*
Prashant Mathur, Cettolo Mauro and Marcello Federico
- 9:40–10:00 *Length-Incremental Phrase Training for SMT*
Joern Wuebker and Hermann Ney
- 10:00–10:20 *Positive Diversity Tuning for Machine Translation System Combination*
Daniel Cer, Christopher D. Manning and Dan Jurafsky
- 10:20–11:00 Coffee Break

Session 7: Poster Session

- 11:00–12:30 Shared Task: Quality Estimation
- Selecting Feature Sets for Comparative and Time-Oriented Quality Estimation of Machine Translation Output*
Eleftherios Avramidis and Maja Popovic
- SHEF-Lite: When Less is More for Translation Quality Estimation*
Daniel Beck, Kashif Shah, Trevor Cohn and Lucia Specia
- Referential Translation Machines for Quality Estimation*
Ergun Bici
- FBK-UEdin Participation to the WMT13 Quality Estimation Shared Task*
José Guilherme Camargo de Souza, Christian Buck, Marco Turchi and Matteo Negri
- The TALP-UPC Approach to System Selection: Asiya Features and Pairwise Classification Using Random Forests*
Lluís Formiga, Meritxell González, Alberto Barrón-Cedeño, José A. R. Fonollosa and Lluís Marquez
- Quality Estimation for Machine Translation Using the Joint Method of Evaluation Criteria and Statistical Modeling*
Aaron Li-Feng Han, Yi Lu, Derek F. Wong, Lidia S. Chao, Liangye He and Junwen Xing

Friday, August 9, 2013 (continued)

MT Quality Estimation: The CMU System for WMT'13

Silja Hildebrand and Stephan Vogel

LORIA System for the WMT13 Quality Estimation Shared Task

David Langlois and Kamel Smaili

LIG System for WMT13 QE Task: Investigating the Usefulness of Features in Word Confidence Estimation for MT

Ngoc Quang Luong, Benjamin Lecouteux and Laurent Besacier

DCU-Symantec at the WMT 2013 Quality Estimation Shared Task

Raphael Rubino, Joachim Wagner, Jennifer Foster, Johann Roturier, Rasoul Samad Zadeh Kaljahi and Fred Hollowood

LIMSI Submission for the WMT'13 Quality Estimation Task: an Experiment with N-Gram Posteriors

Anil Kumar Singh, Guillaume Wisniewski and François Yvon

11:00–12:30 Shared Task: Evaluation

Ranking Translations using Error Analysis and Quality Estimation

Mark Fishel

Are ACT's Scores Increasing with Better Translation Quality?

Najeh Hajlaoui

A Description of Tunable Machine Translation Evaluation Systems in WMT13 Metrics Task

Aaron Li-Feng Han, Derek F. Wong, Lidia S. Chao, Yi Lu, Liangye He, Yiming Wang and Jiaji Zhou

MEANT at WMT 2013: A Tunable, Accurate yet Inexpensive Semantic Frame Based MT Evaluation Metric

Chi-kiu Lo and Dekai Wu

An Approach Using Style Classification Features for Quality Estimation

Erwan Moreau and Raphael Rubino

DCU Participation in WMT2013 Metrics Task

Xiaofeng Wu, Hui Yu and Qun Liu

12:30–14:00 Lunch Break

Friday, August 9, 2013 (continued)

Session 8: Reordering and Hierarchical Models

- 14:00–14:20 *Efficient Solutions for Word Reordering in German-English Phrase-Based Statistical Machine Translation*
Arianna Bisazza and Marcello Federico
- 14:20–14:40 *A Phrase Orientation Model for Hierarchical Machine Translation*
Matthias Huck, Joern Wuebker, Felix Rietig and Hermann Ney
- 14:40–15:00 *A Dependency-Constrained Hierarchical Model with Moses*
Yvette Graham
- 15:00–15:20 *Investigations in Exact Inference for Hierarchical Translation*
Wilker Aziz, Marc Dymetman and Sriram Venkatapathy
- 15:20–16:00 Coffee Break

Session 9: Alignment and Word Translation Models

- 16:00–16:20 *Evaluating (and Improving) Sentence Alignment under Noisy Conditions*
Omar Zaidan and Vishal Chowdhary
- 16:20–16:40 *Multi-Rate HMMs for Word Alignment*
Elif Eyigöz, Daniel Gildea and Kemal Oflazer
- 16:40–17:00 *Hidden Markov Tree Model for Word Alignment*
Shuhei Kondo, Kevin Duh and Yuji Matsumoto
- 17:00–17:20 *An MT Error-Driven Discriminative Word Lexicon using Sentence Structure Features*
Jan Niehues and Alex Waibel

