

4th Workshop on Statistical Machine Translation 2009

(WMT 09)

Held in Conjunction with EACL 2009

**Athens, Greece
30-31 March 2009**

ISBN: 978-1-61738-558-2

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© (2009) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2010)

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: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Findings of the 2009 Workshop on Statistical Machine Translation</i> Chris Callison-Burch, Philipp Koehn, Christof Monz and Josh Schroeder	1
<i>Syntax-Oriented Evaluation Measures for Machine Translation Output</i> Maja Popović and Hermann Ney	29
<i>A Simple Automatic MT Evaluation Metric</i> Petr Homola, Vladislav Kuboň and Pavel Pecina	33
<i>Machine Translation Evaluation with Textual Entailment Features</i> Sebastian Padó, Michel Galley, Daniel Jurafsky and Christopher D. Manning	37
<i>Combining Multi-Engine Translations with Moses</i> Yu Chen, Michael Jellinghaus, Andreas Eisele, Yi Zhang, Sabine Hunsicker, Silke Theison, Christian Federmann and Hans Uszkoreit	42
<i>CMU System Combination for WMT'09</i> Almut Silja Hildebrand and Stephan Vogel	47
<i>The RWTH System Combination System for WMT 2009</i> Gregor Leusch, Evgeny Matusov and Hermann Ney	51
<i>Machine Translation System Combination with Flexible Word Ordering</i> Kenneth Heafield, Greg Hanneman and Alon Lavie	56
<i>Incremental Hypothesis Alignment with Flexible Matching for Building Confusion Networks: BBN System Description for WMT09 System Combination Task</i> Antti-Veikko Rosti, Bing Zhang, Spyros Matsoukas and Richard Schwartz	61
<i>The RWTH Machine Translation System for WMT 2009</i> Maja Popović, David Vilar, Daniel Stein, Evgeny Matusov and Hermann Ney	66
<i>Translation Combination using Factored Word Substitution</i> Christian Federmann, Silke Theison, Andreas Eisele, Hans Uszkoreit, Yu Chen, Michael Jellinghaus and Sabine Hunsicker	70
<i>NUS at WMT09: Domain Adaptation Experiments for English-Spanish Machine Translation of News Commentary Text</i> Preslav Nakov and Hwee Tou Ng	75
<i>The Universität Karlsruhe Translation System for the EACL-WMT 2009</i> Jan Niehues, Teresa Herrmann, Muntsin Kolss and Alex Waibel	80
<i>The TALP-UPC Phrase-Based Translation System for EACL-WMT 2009</i> José A. R. Fonollosa, Maxim Khalilov, Marta R. Costa-jussà, José B. Mariño, Carlos A. Henríguez Q., Adolfo Hernández H. and Rafael E. Banchs	85
<i>Deep Linguistic Multilingual Translation and Bilingual Dictionaries</i> Eric Wehrli, Luka Nerima and Yves Scherrer	90
<i>MATREX: The DCU MT System for WMT 2009</i> Jinhua Du, Yifan He, Sergio Penkale and Andy Way	95

<i>LIMSI's Statistical Translation Systems for WMT'09</i> Alexandre Allauzen, Josep Crego, Aurélien Max and François Yvon	100
<i>NICT@WMT09: Model Adaptation and Transliteration for Spanish-English SMT</i> Michael Paul, Andrew Finch and Eiichiro Sumita	105
<i>Statistical Post Editing and Dictionary Extraction: Systran/Edinburgh Submissions for ACL-WMT2009</i> Loïc Dugast, Jean Senellart and Philipp Koehn	110
<i>Experiments in Morphosyntactic Processing for Translating to and from German</i> Alexander Fraser	115
<i>Improving Alignment for SMT by Reordering and Augmenting the Training Corpus</i> Maria Holmqvist, Sara Stymne, Jody Foo and Lars Ahrenberg	120
<i>English-Czech MT in 2008</i> Ondřej Bojar, David Mareček, Václav Novák, Martin Popel, Jan Ptáček, Jan Rouš and Zdeněk Žabokrtský	125
<i>SMT and SPE Machine Translation Systems for WMT'09</i> Holger Schwenk, Sadaf Abdul Rauf, Loïc Barrault and Jean Senellart	130
<i>Joshua: An Open Source Toolkit for Parsing-Based Machine Translation</i> Zhifei Li, Chris Callison-Burch, Chris Dyer, Sanjeev Khudanpur, Lane Schwartz, Wren Thornton, Jonathan Weese and Omar Zaidan	135
<i>An Improved Statistical Transfer System for French-English Machine Translation</i> Greg Hanneman, Vamshi Ambati, Jonathan H. Clark, Alok Parlikar and Alon Lavie	140
<i>The University of Maryland Statistical Machine Translation System for the Fourth Workshop on Machine Translation</i> Chris Dyer, Hendra Setiawan, Yuval Marton and Philip Resnik	145
<i>Toward Using Morphology in French-English Phrase-Based SMT</i> Marine Carpuat	150
<i>MorphoLogic's Submission for the WMT 2009 Shared Task</i> Attila Novák	155
<i>Edinburgh's Submission to all Tracks of the WMT 2009 Shared Task with Reordering and Speed Improvements to Moses</i> Philipp Koehn and Barry Haddow	160
<i>Mining a Comparable Text Corpus for a Vietnamese-French Statistical Machine Translation System</i> Thi Ngoc Diep Do, Viet Bac Le, Brigitte Bigi, Laurent Besacier and Eric Castelli	165
<i>Improving Arabic-Chinese Statistical Machine Translation using English as Pivot Language</i> Nizar Habash and Jun Hu	173
<i>Domain Adaptation for Statistical Machine Translation with Monolingual Resources</i> Nicola Bertoldi and Marcello Federico	182
<i>Chinese Syntactic Reordering for Adequate Generation of Korean Verbal Phrases in Chinese-to-Korean SMT</i> Jin-Ji Li, Jungi Kim, Dong-II Kim and Jong-Hyeok Lee	190

<i>A Quantitative Analysis of Reordering Phenomena</i> Alexandra Birch, Phil Blunsom and Miles Osborne	197
<i>A POS-Based Model for Long-Range Reorderings in SMT</i> Jan Niehues and Muntsin Kolss	206
<i>Disambiguating "DE" for Chinese-English Machine Translation</i> Pi-Chuan Chang, Daniel Jurafsky and Christopher D. Manning	215
<i>A Systematic Analysis of Translation Model Search Spaces</i> Michael Auli, Adam Lopez, Hieu Hoang and Philipp Koehn	224
<i>A Deep Learning Approach to Machine Transliteration</i> Thomas Deselaers, Saša Hasan, Oliver Bender and Hermann Ney	233
<i>Stabilizing Minimum Error Rate Training</i> George Foster and Roland Kuhn	242
<i>On the Robustness of Syntactic and Semantic Features for Automatic MT Evaluation</i> Jesús Giménez and Lluís Màrquez	250
<i>Fluency, Adequacy, or HTER? Exploring Different Human Judgments with a Tunable MT Metric</i> Matthew Snover, Nitin Madnani, Bonnie Dorr and Richard Schwartz	259