

# **2nd Workshop on Statistical Machine Translation 2007**

**(WMT 07)**

**Held in conjunction with ACL 2007**

**Prague, Czech Republic  
23 June 2007**

**ISBN: 978-1-61738-518-6**

**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© (2007) 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](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>Using Dependency Order Templates to Improve Generality in Translation</i> Arul Menezes and Chris Quirk .....	1
<i>CCG Supertags in Factored Statistical Machine Translation</i> Alexandra Birch, Miles Osborne and Philipp Koehn .....	9
<i>Integration of an Arabic Transliteration Module into a Statistical Machine Translation System</i> Mehdi M. Kashani, Eric Joanis, Roland Kuhn, George Foster and Fred Popowich .....	17
<i>Exploring Different Representational Units in English-to-Turkish Statistical Machine Translation</i> Kemal Oflazer and Ilknur Durgar El-Kahlout .....	25
<i>Can We Translate Letters?</i> David Vilar, Jan-Thorsten Peter and Hermann Ney .....	33
<i>A Dependency Treelet String Correspondence Model for Statistical Machine Translation</i> Deyi Xiong, Qun Liu and Shouxun Lin .....	40
<i>Word Error Rates: Decomposition over POS classes and Applications for Error Analysis</i> Maja Popovic and Hermann Ney .....	48
<i>Speech-Input Multi-Target Machine Translation</i> Alicia Pérez, M. Teresa González, M. Inés Torres and Francisco Casacuberta .....	56
<i>Meta-Structure Transformation Model for Statistical Machine Translation</i> Jiadong Sun, Tiejun Zhao and Huashen Liang .....	64
<i>Training Non-Parametric Features for Statistical Machine Translation</i> Patrick Nguyen, Milind Mahajan and Xiaodong He .....	72
<i>Using Word-Dependent Transition Models in HMM-Based Word Alignment for Statistical Machine Translation</i> Xiaodong He .....	80
<i>Efficient Handling of N-gram Language Models for Statistical Machine Translation</i> Marcello Federico and Mauro Cettolo .....	88
<i>Human Evaluation of Machine Translation Through Binary System Comparisons</i> David Vilar, Gregor Leusch, Hermann Ney and Rafael E. Banchs .....	96
<i>Labelled Dependencies in Machine Translation Evaluation</i> Karolina Owczarzak, Josef van Genabith and Andy Way .....	104
<i>An Iteratively-Trained Segmentation-Free Phrase Translation Model for Statistical Machine Translation</i> Robert Moore and Chris Quirk .....	112

<i>Using Paraphrases for Parameter Tuning in Statistical Machine Translation</i> Nitin Madnani, Necip Fazil Ayan, Philip Resnik and Bonnie Dorr .....	120
<i>Mixture-Model Adaptation for SMT</i> George Foster and Roland Kuhn .....	128
<i>(Meta-) Evaluation of Machine Translation</i> Chris Callison-Burch, Cameron Fordyce, Philipp Koehn, Christof Monz and Josh Schroeder ...	136
<i>Context-aware Discriminative Phrase Selection for Statistical Machine Translation</i> Jesús Giménez and Lluís Màrquez .....	159
<i>Ngram-Based Statistical Machine Translation Enhanced with Multiple Weighted Reordering Hypotheses</i> Marta R. Costa-jussà, Josep M. Crego, Patrik Lambert, Maxim Khalilov, José A. R. Fonollosa, José B. Mario and Rafael E. Banchs .....	167
<i>Analysis of Statistical and Morphological Classes to Generate Weighted Reordering Hypotheses on a Statistical Machine Translation System</i> Marta R. Costa-jussà and José A. R. Fonollosa .....	171
<i>Domain Adaptation in Statistical Machine Translation with Mixture Modelling</i> Jorge Civera and Alfons Juan .....	177
<i>Getting to Know Moses: Initial Experiments on German-English Factored Translation</i> Maria Holmqvist, Sara Stymne and Lars Ahrenberg .....	181
<i>NRC's PORTAGE System for WMT 2007</i> Nicola Ueffing, Michel Simard, Samuel Larkin and Howard Johnson .....	185
<i>Building a Statistical Machine Translation System for French Using the Europarl Corpus</i> Holger Schwenk .....	189
<i>Multi-Engine Machine Translation with an Open-Source SMT Decoder</i> Yu Chen, Andreas Eisele, Christian Federmann, Eva Hasler, Michael Jellinghaus and Silke Theison .....	193
<i>The ISL Phrase-Based MT System for the 2007 ACL Workshop on Statistical Machine Translation</i> Matthias Paulik, Kay Rottmann, Jan Niehues, Silja Hildebrand and Stephan Vogel .....	197
<i>Rule-Based Translation with Statistical Phrase-Based Post-Editing</i> Michel Simard, Nicola Ueffing, Pierre Isabelle and Roland Kuhn .....	203
<i>The "Noisier Channel": Translation from Morphologically Complex Languages</i> Christopher J. Dyer .....	207
<i>UCB System Description for the WMT 2007 Shared Task</i> Preslav Nakov and Marti Hearst .....	212

<i>The Syntax Augmented MT (SAMT) System at the Shared Task for the 2007 ACL Workshop on Statistical Machine Translation</i>	
Andreas Zollmann, Ashish Venugopal, Matthias Paulik and Stephan Vogel . . . . .	216
<i>Statistical Post-Editing on SYSTRAN's Rule-Based Translation System</i>	
Loïc Dugast, Jean Senellart and Philipp Koehn . . . . .	220
<i>Experiments in Domain Adaptation for Statistical Machine Translation</i>	
Philipp Koehn and Josh Schroeder . . . . .	224
<i>METEOR: An Automatic Metric for MT Evaluation with High Levels of Correlation with Human Judgments</i>	
Alon Lavie and Abhaya Agarwal . . . . .	228
<i>English-to-Czech Factored Machine Translation</i>	
Ondřej Bojar . . . . .	232
<i>Sentence Level Machine Translation Evaluation as a Ranking</i>	
Yang Ye, Ming Zhou and Chin-Yew Lin . . . . .	240
<i>Localization of Difficult-to-Translate Phrases</i>	
Behrang Mohit and Rebecca Hwa . . . . .	248
<i>Linguistic Features for Automatic Evaluation of Heterogenous MT Systems</i>	
Jesús Giménez and Lluís Màrquez . . . . .	256