

EACL 2009

**Fourth Workshop
on
Statistical
Machine Translation**

Proceedings of the Workshop

30 March – 31 March 2009
Megaron Athens International Conference Centre
Athens, Greece

Production and Manufacturing by
TEHNOGRAFIA DIGITAL PRESS
7 Ektoros Street
152 35 Vrilissia
Athens, Greece



©2009 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

Introduction

The EACL 2009 Workshop on Statistical Machine Translation (WMT09) took place on March 30 and 31 in Athens, Greece, immediately preceding the 12th conference of the European Chapter of the Association for Computational Linguistics (EACL), which was organized by the Greek National Centre for Scientific Research with support from Athens University of Economics and Business—Department of Informatics and the Institute for Language and Speech Processing.

This is the fifth time this workshop has been held. The first time was in 2005 as part of the ACL 2005 Workshop on Building and Using Parallel Texts. In the following years the Workshop on Statistical Machine Translation was held at HLT-NAACL 2006 in New York City, US, at ACL 2007 in Prague, Czech Republic, and at ACL 2008 in Columbus, Ohio, US.

The focus of our workshop was to evaluate the state of the art in machine translation (MT) for a variety of languages. Recent experimentation has shown that the performance of machine translation systems varies greatly with the source language. In this workshop we encouraged researchers to investigate ways to improve the performance of MT systems for diverse languages.

Prior to the workshop, in addition to soliciting relevant papers for review and possible presentation we conducted a shared task that brought together machine translation systems for an evaluation on previously unseen data. The shared task also included a track for evaluation metrics and system combination methods.

The results of the shared task were announced at the workshop, and these proceedings also include an overview paper for the shared task 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 some 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 21 full paper submissions. 12 full papers were selected for oral presentation.

We received 3 short paper submissions for the evaluation task, 5 short paper submissions for the system combination task, and 20 short paper submissions for the translation task. Due to the large number of high quality submission for the full paper track, shared task submissions were presented as posters. The poster session gave participants of the shared task the opportunity to present their approaches.

The invited talk was given by Martin Kay (Stanford University and University of the Saarland).

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 manual evaluations.

Chris Callison-Burch, Philipp Koehn, Christof Monz, and Josh Schroeder
Co-Organizers

EACL 2009 Organizers

Organizers:

Chris Callison-Burch, Johns Hopkins University (United States of America)
Philipp Koehn, University of Edinburgh (United Kingdom)
Christof Monz, University of Amsterdam (The Netherlands)
Josh Schroeder, University of Edinburgh (United Kingdom)

Invited Talk:

Martin Kay (Stanford University and University Saarbrücken)

Program Committee:

Lars Ahrenberg, Linköping University (Sweden)
Yaser Al-Onaizan, IBM Research (United States of America)
Necip Fazil Ayan, SRI (United States of America)
Thorsten Brants, Google (United States of America)
Chris Brockett, Microsoft Research (United States of America)
Francisco Casacuberta, University of Valencia (Spain)
David Chiang, ISI/University of Southern California (United States of America)
Colin Cherry, Microsoft Research (United States of America)
Stephen Clark, Cambridge University (United Kingdom)
Trevor Cohn, Edinburgh University (United Kingdom)
Brooke Cowan, MIT (United States of America)
Mona Diab, Columbia University (United States of America)
Andreas Eisele, University Saarbrücken (Germany)
Marcello Federico, FBK-irst (Italy)
George Foster, Canada National Research Council (Canada)
Alex Fraser, University of Stuttgart (Germany)
Michel Galley, Columbia University (United States of America)
Jesus Gimenez, Technical University of Catalonia (Spain)
Keith Hall, Google (United States of America)
John Henderson, MITRE (United States of America)
Rebecca Hwa, University of Pittsburgh (United States of America)
Doug Jones, Lincoln Labs MIT (United States of America)
Damianos Karakos, Johns Hopkins University (United States of America)
Katrín Kirchoff, University of Washington (United States of America)
Kevin Knight, ISI/University of Southern California (United States of America)
Shankar Kumar, Google (United States of America)
Philippe Langlais, University of Montreal (Canada)
Alon Lavie, Carnegie Mellon University (United States of America)
Adam Lopez, Edinburgh University (United Kingdom)
Daniel Marcu, ISI/University of Southern California (United States of America)
Lambert Mathias, Johns Hopkins University (United States of America)
Bob Moore, Microsoft Research (United States of America)
Smaranda Muresan, Rutgers University (United States of America)
Franz Josef Och, Google (United States of America)

Miles Osborne, Edinburgh University (United Kingdom)
Kay Peterson, NIST (United States of America)
Mark Przybocki, NIST (United States of America)
Chris Quirk, Microsoft Research (United States of America)
Antti-Veikko Rosti, BBN Technologies (United States of America)
Holger Schwenk, LIUM (France)
Jean Senellart, Systran (France)
Libin Shen, BBN Technologies (United States of America)
Wade Shen, Lincoln Labs MIT (United States of America)
Michel Simard, National Research Council Canada (Canada)
David Talbot, Google (United States of America)
Jörg Tiedemann, University of Groningen (The Netherlands)
Christoph Tillmann, IBM Research (United States of America)
Dan Tufis, Romanian Academy (Romania)
Clare Voss, Army Research Labs (United States of America)
Taro Watanabe, NTT (Japan)
Andy Way, Dublin City University (Ireland)
Jinxi Xu, BBN Technologies (United States of America)
Richard Zens, Google (United States of America)

Additional Reviewers

Nicola Bertoldi, FBK-irst (Italy)
Mauro Cettolo, FBK-irst (Italy)
Jeffrey Micher, Carnegie Mellon University (United States of America)

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 Popovic and Hermann Ney	29
<i>A Simple Automatic MT Evaluation Metric</i> Petr Homola, Vladislav Kubon and Pavel Pecina	33
<i>Machine Translation Evaluation with Textual Entailment Features</i> Sebastian Pado, 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 Popovic, 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-jussá, 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, Loic 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 WMT2009 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-Il 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

Conference Program

Monday, March 30, 2009

9:00–9:15 Opening Remarks

Overview of the Shared Tasks

9:15–9:45 *Findings of the 2009 Workshop on Statistical Machine Translation*
Chris Callison-Burch, Philipp Koehn, Christof Monz and Josh Schroeder

Shared Task: Evaluation and System Combination

9:45–10:15 Boaster Session

10:15–11:30 Poster Session: Shared Evaluation Task

Syntax-Oriented Evaluation Measures for Machine Translation Output
Maja Popovic and Hermann Ney

A Simple Automatic MT Evaluation Metric
Petr Homola, Vladislav Kubon and Pavel Pecina

Machine Translation Evaluation with Textual Entailment Features
Sebastian Pado, Michel Galley, Daniel Jurafsky and Christopher D. Manning

10:15–11:30 Poster Session: Shared System Combinations

Combining Multi-Engine Translations with Moses
Yu Chen, Michael Jellinghaus, Andreas Eisele, Yi Zhang, Sabine Hunsicker, Silke Theison, Christian Federmann and Hans Uszkoreit

CMU System Combination for WMT'09
Almut Silja Hildebrand and Stephan Vogel

The RWTH System Combination System for WMT 2009
Gregor Leusch, Evgeny Matusov and Hermann Ney

Machine Translation System Combination with Flexible Word Ordering
Kenneth Heafield, Greg Hanneman and Alon Lavie

Monday, March 30, 2009 (continued)

Incremental Hypothesis Alignment with Flexible Matching for Building Confusion Networks: BBN System Description for WMT09 System Combination Task

Antti-Veikko Rosti, Bing Zhang, Spyros Matsoukas and Richard Schwartz

11:30-12:30 Invited Talk by Martin Kay

12:30-14:00 Lunch break

14:00-14:30 Panel Discussion

Shared Task: Translation

14:30-15:30 Boaster Session

15:30-17:30 Poster Session

The RWTH Machine Translation System for WMT 2009

Maja Popovic, David Vilar, Daniel Stein, Evgeny Matusov and Hermann Ney

Translation Combination using Factored Word Substitution

Christian Federmann, Silke Theison, Andreas Eisele, Hans Uszkoreit, Yu Chen, Michael Jellinghaus and Sabine Hunsicker

NUS at WMT09: Domain Adaptation Experiments for English-Spanish Machine Translation of News Commentary Text

Preslav Nakov and Hwee Tou Ng

The Universität Karlsruhe Translation System for the EACL-WMT 2009

Jan Niehues, Teresa Herrmann, Muntsin Kolss and Alex Waibel

The TALP-UPC Phrase-Based Translation System for EACL-WMT 2009

José A. R. Fonollosa, Maxim Khalilov, Marta R. Costa-jussá, José B. Mariño, Carlos A. Henrández Q., Adolfo Hernández H. and Rafael E. Banchs

Deep Linguistic Multilingual Translation and Bilingual Dictionaries

Eric Wehrli, Luka Nerima and Yves Scherrer

MATREX: The DCU MT System for WMT 2009

Jinhua Du, Yifan He, Sergio Penkale and Andy Way

Monday, March 30, 2009 (continued)

LIMSI's Statistical Translation Systems for WMT'09

Alexandre Allauzen, Josep Crego, Aurélien Max and François Yvon

NICT@WMT09: Model Adaptation and Transliteration for Spanish-English SMT

Michael Paul, Andrew Finch and Eiichiro Sumita

Statistical Post Editing and Dictionary Extraction: Systran/Edinburgh submissions for ACL-WMT2009

Loïc Dugast, Jean Senellart and Philipp Koehn

Experiments in Morphosyntactic Processing for Translating to and from German

Alexander Fraser

Improving Alignment for SMT by Reordering and Augmenting the Training Corpus

Maria Holmqvist, Sara Stymne, Jody Foo and Lars Ahrenberg

English-Czech MT in 2008

Ondřej Bojar, David Mareček, Václav Novák, Martin Popel, Jan Ptáček, Jan Rouš and Zdeněk Žabokrtský

SMT and SPE Machine Translation Systems for WMT'09

Holger Schwenk, Sadaf Abdul Rauf, Loic Barrault and Jean Senellart

Joshua: An Open Source Toolkit for Parsing-Based Machine Translation

Zhifei Li, Chris Callison-Burch, Chris Dyer, Sanjeev Khudanpur, Lane Schwartz, Wren Thornton, Jonathan Weese and Omar Zaidan

An Improved Statistical Transfer System for French-English Machine Translation

Greg Hanneman, Vamshi Ambati, Jonathan H. Clark, Alok Parlikar and Alon Lavie

The University of Maryland Statistical Machine Translation System for the Fourth Workshop on Machine Translation

Chris Dyer, Hendra Setiawan, Yuval Marton and Philip Resnik

Toward Using Morphology in French-English Phrase-Based SMT

Marine Carpuat

MorphoLogic's Submission for the WMT 2009 Shared Task

Attila Novák

Monday, March 30, 2009 (continued)

Edinburgh's Submission to all Tracks of the WMT2009 Shared Task with Reordering and Speed Improvements to Moses

Philipp Koehn and Barry Haddow

Tuesday, March 31, 2009

Full Papers Session 1: Use of training data

9:00–9:30 *Mining a Comparable Text Corpus for a Vietnamese-French Statistical Machine Translation System*

Thi Ngoc Diep Do, Viet Bac Le, Brigitte Bigi, Laurent Besacier and Eric Castelli

9:30–10:00 *Improving Arabic-Chinese Statistical Machine Translation using English as Pivot Language*

Nizar Habash and Jun Hu

10:00–10:30 *Domain Adaptation for Statistical Machine Translation with Monolingual Resources*

Nicola Bertoldi and Marcello Federico

10:30-11:00 Coffee break

Full Papers Session 2: Reordering

11:00–11:30 *Chinese Syntactic Reordering for Adequate Generation of Korean Verbal Phrases in Chinese-to-Korean SMT*

Jin-Ji Li, Jungi Kim, Dong-Il Kim and Jong-Hyeok Lee

11:30–12:00 *A Quantitative Analysis of Reordering Phenomena*

Alexandra Birch, Phil Blunsom and Miles Osborne

12:00–12:30 *A POS-Based Model for Long-Range Reorderings in SMT*

Jan Niehues and Muntsin Kolss

12:30-14:30 Lunch break

Tuesday, March 31, 2009 (continued)

Full Papers Session 3: Linguistic modeling

14:30–15:00 *Disambiguating "DE" for Chinese-English Machine Translation*
Pi-Chuan Chang, Daniel Jurafsky and Christopher D. Manning

15:00–15:30 *A Systematic Analysis of Translation Model Search Spaces*
Michael Auli, Adam Lopez, Hieu Hoang and Philipp Koehn

15:30–16:00 *A Deep Learning Approach to Machine Transliteration*
Thomas Deselaers, Saša Hasan, Oliver Bender and Hermann Ney

16:00–16:30 Coffee break

Full Papers Session 4: Error metrics and tuning

16:30–17:00 *Stabilizing Minimum Error Rate Training*
George Foster and Roland Kuhn

17:00–17:30 *On the Robustness of Syntactic and Semantic Features for Automatic MT Evaluation*
Jesús Giménez and Lluís Màrquez

17:30–18:00 *Fluency, Adequacy, or HTER? Exploring Different Human Judgments with a Tunable MT Metric*
Matthew Snover, Nitin Madnani, Bonnie Dorr and Richard Schwartz

