

18th Annual Conference of the European Association for Machine Translation (EAMT 2015)

Antalya, Turkey
11 – 13 May 2015

Editors:

**İlknur Durgar El-Kahlout
Mehmed Özkan
Felipe Sánchez-Martínez**

**Gema Ramírez-Sánchez
Fred Hollowood
Andy Way**

ISBN: 978-1-5108-1187-4

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

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

List of papers

Invited Talk

What we want, what we need, what we absolutely can't do without – an enterprise user's perspective on machine translation technology and stuff around it

Olga Beregovaya

1

Research papers

2

Exploiting portability to build an RBMT prototype for a new source language

Nora Aranberri, Gorka Labaka, Arantza Díaz de Ilarraza and Kepa Sarasola

3

Building hybrid machine translation systems by using an EBMt preprocessor to create partial translations

Mikel Artetxe, Gorka Labaka, Kepa Sarasola

11

Using on-line available sources of bilingual information for word-level machine translation quality estimation

Miquel Esplá-Gomis, Felipe Sánchez-Martínez, Mikel L. Forcada

19

A general framework for minimizing translation effort: towards a principled combination of translation technologies in computer-aided translation

Mikel L. Forcada and Felipe Sánchez-Martínez

27

Can Translation Memories afford not to use paraphrasing?

Rohit Gupta, Constantin Orasan, Marcos Zampieri, Mihaela Vela, Josef van Genabith

35

Dependency-based Reordering Model for Constituent Pairs in Hierarchical SMT

Arefeh Kazemiy, Antonio Toral, Andy Way, Amirhassan Monadjemiy, Mohammadali Nematbakhshy

43

The role of artificially generated negative data for quality estimation of machine translation

Varvara Logacheva and Lucia Specia

51

Document-Level Machine Translation with Word Vector Models

Eva Martinez Garcia, Cristina Espana-Bonet, Lluís Marquez

59

The potential and limits of lay post-editing in an online community

Linda Mitchell

67

Post-Editing Evaluations: Trade-offs between Novice and Professional Participants

Joss Moorkens and Sharon O'Brien

75

Benchmarking SMT Performance for Farsi Using the TEP++ Corpus

Peyman Passban, Andy Way, Qun Liu

82

THE 18TH ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR MACHINE
TRANSLATION (EAMT 2015)

<i>Dynamic Terminology Integration Methods in Statistical Machine Translation</i> Marcis Pinnis	89
<i>Identifying main obstacles for statistical machine translation of morphologically rich South Slavic languages</i> Maja Popovic and Mihael Arcan	97
<i>Poor man's lemmatisation for automatic error classification</i> Maja Popovic, Mihael Arcan, Eleftherios Avramidis, Aljoscha Burchardt, Arle Lommel	105
<i>Truly Exploring Multiple References for Machine Translation Evaluation</i> Ying Qin and Lucia Specia	113
<i>Searching for Context: a Study on Document-Level Labels for Translation Quality Estimation</i> Carolina Scarton, Marcos Zampieri, Mihaela Vela, Josef van Genabith, Lucia Specia	121
<i>Stripping Adjectives: Integration Techniques for Selective Stemming in SMT Systems</i> Isabel Slawik, Jan Niehues, Alex Waibel	129
<i>Evaluating machine translation for assimilation via a gap-filling task</i> Ekaterina Ageeva, Francis M. Tyers, Mikel L. Forcada, Juan Antonio Pérez-Ortiz	137
<i>Unsupervised training of maximum-entropy models for lexical selection in rule-based machine translation</i> Francis M. Tyers, Felipe Sánchez-Martinez, Mikel L. Forcada	145
<i>Assessing linguistically aware fuzzy matching in translation memories</i> Tom Vanallemeersch and Vincent Vandeghinste	153
<i>Re-assessing the WMT2013 Human Evaluation with Professional Translators Trainees</i> Mihaela Vela and Josef van Genabith	161
<i>Integrating a Large, Monolingual Corpus as Translation Memory into Statistical Machine translation</i> Katharina Wäschle and Stefan Riezler	169
<i>TargetSide Generation of Prepositions for SMT</i> Marion Weller, Alexander Fraser, Sabine Schulte im Walde	177
<i>Word Alignment Based Parallel Corpora Evaluation and Cleaning Using Machine Learning Techniques</i> Ieva Zariņa, Pēteris Nikiforovs, Raivis Skadiņš	185

User papers	193
<i>Content Translation: Computer-assisted translation tool for Wikipedia articles</i> Niklas Laxström, Pau Giner, Santhosh Thottingal	194
<i>Pre-reordering for Statistical Machine Translation of Non-fictional Subtitles.</i> Magdalena Plamada, Gion Linder, Phillip Ströbel, Martin Volk	198
<i>SMT at the International Maritime Organization: experiences with combining in-house corpus with more general corpus</i> Bruno Pouliquen, Marcin Junczys-Dowmunt, Blanca Pinero, Michał Ziemski	202
<i>Evaluation of the domain adaptation of MT systems in ACCURAT</i> Gregor Thurmair	206
Project/product descriptions	210
<i>MixedEmotions: Social Semantic Emotion Analysis for Innovative Multilingual Big Data Analytics Markets</i> Mihael Arcan and Paul Buitelaar	211
<i>The ACCEPT Academic Portal: Bringing Together Pre-editing, MT and Post-editing into a Learning Environment</i> Pierrette Bouillon, Johanna Gerlach, Asheesh Gulati, Victoria Porro, Violeta Seretan	212
<i>Russian-Chinese Sentence-level Aligned News Corpus</i> Wenjun Du, Wuying Liu, Junting Yu, Mianzhu Yi	213
<i>HimL (Health in my Language)</i> Barry Haddow	214
<i>MT-enhanced fuzzy matching with Transit NXT and STAR Moses</i> Nadira Hofmann	215
<i>HandyCAT - An Open-Source Platform for CAT Tool Research</i> Chris Hokamp and Qun Liu	216
<i>TraMOOC: Translation for Massive Open Online Courses</i> Valia Kordoni, Kostadin Cholakov, Markus Egg, Andy Way, Lexi Birch, Katia Keramanidis, Vilemini Sisoni, Dimitrios Tsoumakos, Antal van den Bosch, Iris Hendrickx, Michael Papadopoulos, Panayota Georgakopoulou, Maria Gialama, Menno van Zaanen, Ioana Buliga, Mitja Jermol and Davor Orlic	217
<i>Streamlining Translation Workflows with Style Scorer</i> David Landan and Olga Beregovaya	218

THE 18TH ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR MACHINE
TRANSLATION (EAMT 2015)

<i>Estonian-English Reversible Smart Phone Dictionary of Military Terms and Relevant Vocabulary</i> Epp Leete	219
<i>FALCON: Federated Active Linguistic data CuratiON</i> David Lewis	220
<i>Tapadóir</i> Eimear Maguire, John Judge and Teresa Lynn	221
<i>Okapi+QuEst: Translation Quality Estimation within Okapi</i> Gustavo Henrique Paetzold, Lucia Specia, Yves Savourel	222
<i>CRACKER: Cracking the Language Barrier</i> Georg Rehm	223
<i>LTC KnowHow: Empowering the Social Enterprise in the Language Industry</i> Adriane Rinsche and Sabine Rinsches	224
<i>Multi-Dialect Machine Translation (MuDMat)</i> Fatiha Sadat	226
<i>Abu-MaTran: Automatic building of Machine Translation</i> Antonio Toral, Tommi A Pirinen, Andy Way, Gema Ramírez-Sánchez, Sergio Ortiz Rojas, Raphael Rubino, Miquel Esplà, Mikel Forcada, Vassilis Papavassiliou, Prokopis Prokopidis and Nikola Ljubešić	227
<i>MNH-TT: A Platform to Support Collaborative Translator Training</i> Masao Utiyama, Kyo Kageura, Martin Thomas, Anthony Hartley	228
<i>Smart Computer Aided Translation Environment</i> Vincent Vandeghinste, Tom Vanallemeersch, Frank Van Eynde, Geert Heyman, Sien Moens, Joris Pelemans, Patrick Wambacq, Iulianna Van der Lek - Ciudin, Arda Tezcan, Lieve Macken, Véronique Hoste, Eva Geurts and Mieke Haesen	229