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
## 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

### Research papers

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploiting portability to build an RBMT prototype for a new source language</td>
<td>Nora Aranberri, Gorka Labaka, Arantza Díaz de Ilarraza and Kepa Sarasola</td>
<td>3</td>
</tr>
<tr>
<td>Building hybrid machine translation systems by using an EBMT preprocessor to create partial translations</td>
<td>Mikel Artetxe, Gorka Labaka, Kepa Sarasola</td>
<td>11</td>
</tr>
<tr>
<td>Using on-line available sources of bilingual information for word-level machine translation quality estimation</td>
<td>Miquel Esplá-Gomis, Felipe Sánchez-Martínez, Mikel L. Forcada</td>
<td>19</td>
</tr>
<tr>
<td>A general framework for minimizing translation effort: towards a principled combination of translation technologies in computer-aided translation</td>
<td>Mikel L. Forcada and Felipe Sánchez-Martínez</td>
<td>27</td>
</tr>
<tr>
<td>Can Translation Memories afford not to use paraphrasing?</td>
<td>Rohit Gupta, Constantin Orasan, Marcos Zampieri, Mihaela Vela, Josef van Genabith</td>
<td>35</td>
</tr>
<tr>
<td>Dependency-based Reordering Model for Constituent Pairs in Hierarchical SMT</td>
<td>Arefeh Kazemiy, Antonio Toral, Andy Way, Amirhassan Monadjemiy, Mohammadali Nematbakhshy</td>
<td>43</td>
</tr>
<tr>
<td>The role of artificially generated negative data for quality estimation of machine translation</td>
<td>Varvara Logacheva and Lucia Specia</td>
<td>51</td>
</tr>
<tr>
<td>The potential and limits of lay post-editing in an online community</td>
<td>Linda Mitchell</td>
<td>67</td>
</tr>
<tr>
<td>Post-Editing Evaluations: Trade-offs between Novice and Professional Participants</td>
<td>Joss Moorkens and Sharon O’Brien</td>
<td>75</td>
</tr>
<tr>
<td>Benchmarking SMT Performance for Farsi Using the TEP++ Corpus</td>
<td>Peyman Passban, Andy Way, Qun Liu</td>
<td>82</td>
</tr>
</tbody>
</table>
**Dynamic Terminology Integration Methods in Statistical Machine Translation**
Marcis Pinnis 89

**Identifying main obstacles for statistical machine translation of morphologically rich South Slavic languages**
Maja Popovic and Mihael Arcan 97

**Poor man’s lemmatisation for automatic error classification**
Maja Popovic, Mihael Arcan, Eleftherios Avramidis, Aljoscha Burchardt, Arle Lommel 105

**Truly Exploring Multiple References for Machine Translation Evaluation**
Ying Qin and Lucia Specia 113

**Searching for Context: a Study on Document-Level Labels for Translation Quality Estimation**
Carolina Scarton, Marcos Zampieri, Mihaela Vela, Josef van Genabith, Lucia Specia 121

**Stripping Adjectives: Integration Techniques for Selective Stemming in SMT Systems**
Isabel Slawik, Jan Niehues, Alex Waibel 129

**Evaluating machine translation for assimilation via a gap-filling task**
Ekaterina Ageeva, Francis M. Tyers, Mikel L. Forcada, Juan Antonio Pérez-Ortiz 137

**Unsupervised training of maximum-entropy models for lexical selection in rule-based machine translation**
Francis M. Tyers, Felipe Sánchez-Martinez, Mikel L. Forcada 145

**Assessing linguistically aware fuzzy matching in translation memories**
Tom Vanallemeersch and Vincent Vandehгинste 153

**Re-assessing the WMT2013 Human Evaluation with Professional Translators Trainees**
Mihaela Vela and Josef van Genabith 161

**Integrating a Large, Monolingual Corpus as Translation Memory into Statistical Machine translation**
Katharina Wäschle and Stefan Riezler 169

**TargetSide Generation of Prepositions for SMT**
Marion Weller, Alexander Fraser, Sabine Schulte im Walde 177

**Word Alignment Based Parallel Corpora Evaluation and Cleaning Using Machine Learning Techniques**
Ieva Zariņa, Pēteris Nikiforovs, Raivis Skadiņš 185
**User papers**

*Content Translation: Computer-assisted translation tool for Wikipedia articles*
Niklas Laxström, Pau Giner, Santhosh Thottingal

*Pre-reordering for Statistical Machine Translation of Non-fictional Subtitles.*
Magdalena Plamada, Gion Linder, Phillip Ströbel, Martin Volk

*SMT at the International Maritime Organization: experiences with combining in-house corpus with more general corpus*
Bruno Pouliquen, Marcin Junczys-Dowmunt, Blanca Pinero, Michał Ziemski

*Evaluation of the domain adaptation of MT systems in ACCURAT*
Gregor Thurmair

**Project/product descriptions**

*MixedEmotions: Social Semantic Emotion Analysis for Innovative Multilingual Big Data Analytics Markets*
Mihael Arcan and Paul Buitelaar

*The ACCEPT Academic Portal: Bringing Together Pre-editing, MT and Post-editing into a Learning Environment*
Pierrette Bouillon, Johanna Gerlach, Asheesh Gulati, Victoria Porro, Violeta Seretan

*Russian-Chinese Sentence-level Aligned News Corpus*
Wenjun Du, Wuying Liu, Junting Yu, Mianzhu Yi

*HimL (Health in my Language)*
Barry Haddow

*MT-enhanced fuzzy matching with Transit NXT and STAR Moses*
Nadira Hofmann

*HandyCAT - An Open-Source Platform for CAT Tool Research*
Chris Hokamp and Qun Liu

*TraMOOC: Translation for Massive Open Online Courses*
Valia Kordoni, Kostadin Cholakov, Markus Egg, Andy Way, Lexi Birch, Katia Kermanidis, Vilelmini Sosoni, Dimitrios Tsoumakos, Antal van den Bosch, Iris Hendrickx, Michael Papadopoulos, Panayota Georgakopoulou, Maria Gialama, Menno van Zaanen, Ioana Buliga, Mitja Jermol and Davor Orlic

*Streamlining Translation Workflows with Style Scorer*
David Landan and Olga Beregovaya
Estonian-English Reversible Smart Phone Dictionary of Military Terms and Relevant Vocabulary
Epp Leete 219

FALCON: Federated Active Linguistic data CuratiON
David Lewis 220

Tapadóir
Eimear Maguire, John Judge and Teresa Lynn 221

Okapi+QuEst: Translation Quality Estimation within Okapi
Gustavo Henrique Paetzold, Lucia Specia, Yves Savourel 222

CRACKER: Cracking the Language Barrier
Georg Rehm 223

LTC KnowHow: Empowering the Social Enterprise in the Language Industry
Adriane Rinsche and Sabine Rinsches 224

Multi-Dialect Machine Translation (MuDMat)
Fatiha Sadat 226

Abu-MaTran: Automatic building of Machine Translation
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

MNH-TT: A Platform to Support Collaborative Translator Training
Masao Utiyama, Kyo Kageura, Martin Thomas, Anthony Hartley 228

Smart Computer Aided Translation Environment
Vincent Vandeghinste, Tom Vanallemeursch, 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