Machine Translation Summit XIV

2–6 September 2013, Nice, France

Workshop Proceedings:

Workshop on Post-editing Technology and Practice

(WPTP-2)

Editors: Sharon O'Brien, Michel Simard and Lucia Specia

EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION

IAMT

Workshop Proceedings for:
The Workshop on Post-editing Technology and Practice
(Organised at the 14th Machine Translation Summit)

Editors:	Sharon O'Brien	CNGL/DCU
	Michel Simard	National Research Council Canada
	Lucia Specia	University of Sheffield

Published by
The European Association for Machine Translation
Schützenweg 57
CH-4123 Allschwil / Switzerland

ISBN: 978-3-9524207-2-0

©2013 The authors.
These proceedings are licensed under a Creative Commons 3.0 licence, no derivative works, attribution, CC-BY-ND
(For certain papers of these proceedings there may be stated other copyrights)

MT Summit 2013 Workshop Chair

Svetlana SHEREMETYEVA

Workshop Organisers

Sharon O'BRIEN – CNGL / Dublin City University
Michel SIMARD – National Research Council Canada
Lucia SPECIA – University of Sheffield

Program Committee

Nora ARANBERRI – TAUS
Diego BARTOLOME – tauyou language technology
Michael CARL – Copenhagen Business School
Francisco CASACUBERTA – Universitat Politècnica de València
Mike DILLINGER – eBay
Stephen DOHERTY – Dublin City University
Andreas EISELE – European Commission
Jakob ELMING – Copenhagen Business School
Marcello FEDERICO – FBK Trento, Italy
Mikel L. FORCADA – Universitat d'Alacant
Ana GUERBEROF – Pactera Technology International Ltd
Nizar HABASH – Columbia University
Kristian Tangsgaard HVELPLUND – Copenhagen Business School
Maxim KHALILOV – TAUS
Philipp KOEHN – University of Edinburgh
Roland KUHN – National Research Council Canada
Philippe LANGLAIS – RALI / Université de Montréal
Alon LAVIE – Carnegie Mellon University
Elliott MACKLOVITCH – MT Consultant
Daniel MARCU – SDL and USC/ISI
John MORAN – Transpiral Translation Services
Kristen PARTON – Facebook
Maja POPOVIĆ – DFKI
Johann ROTURIER – Symantec
Midori TATSUMI – Dublin City University
Jörg TIEDEMANN – Department of Linguistics and Philology, Uppsala University
Andy WAY – Lingo24

Workshop Programme

9:00 – 10:30: Session 1

9:00 Word of welcome

9:15 *This translation is not too bad: An analysis of post-editor choices in a machine translation post-editing task*
Maarit Koponen

9:40 *What can we learn about selection mechanism for post-editing?*
Maja Popović, Aljoscha Burchardt, Eleftherios Avramidis and David Vilar

10:05 *User Attitudes to the Post-Editing Interface*
Joss Moorkens and Sharon O'Brien

Coffee Break: 10:30 – 10:45

10:45 – 12:00: Session 2

10:45 *Integrated Post-Editing and Translation Management for Lay User Communities*
Adrian Laurenzi, Megumu Brownstein, Anne M. Turner and Katrin Kirchhoff

11:10 *Community-based post-editing of machine-translated content: monolingual vs. bilingual*
Linda Mitchell, Johann Roturier and Sharon O'Brien

11:35 *Combining pre-editing and post-editing to improve SMT of user-generated content*
Johanna Gerlach, Victoria Porro, Pierrette Bouillon and Sabine Lehmann

12:00 – 13:30: Lunch Break

13:30 – 14:30: Invited Talk

Organizational Ergonomics of Post-Editing
Mirko Plitt

14:30 – 16:25: Posters and Demos

Advanced Computer Aided Translation with a Web-Based Workbench
(Poster and demo)

Vicent Alabau, Ragnar Bonk, Christian Buck, Michael Carl, Francisco Casacuberta, Mercedes Garcia-Martinez, Philipp Koehn, Luis Leiva, Bartolome Mesa-Lao, Herve Saint-Amand, Chara Tsoukala,
German Sanchis, Daniel Ortiz and Jesus Gonzalez

Quality as the sum of its parts: A two-step approach for the identification of translation problems and translation quality assessment for HT and MT+PE (Poster)
Joke Daems and Lieve Macken

Transferring Source Tags to the Target Text in Statistical Machine Translation: A Two-Stream Approach (Poster)
Eric Joanis, Darlene Stewart, Samuel Larkin and Roland Kuhn

Assessing Post-Editing Efficiency in a Realistic Translation Environment (Poster)
Samuel Läubli, Mark Fishel, Gary Massey,
Maureen Ehrensberger-Dow and Martin Volk

Integrated Post-Editing and Translation Management for Lay User Communities (Demo)
Adrian Laurenzi, Megumu Brownstein, Anne M. Turner and Katrin Kirchhoff

VistaTec (Demo)
Phil Ritchie

The ACCEPT Post-Editing environment: a flexible and customisable online tool to perform and analyse machine translation post-editing (Demo)
Johann Roturier, Linda Mitchell and David Silva

PET: A tool for evaluating translation quality through post-editing (Demo)
Lucia Specia

An Evaluation of Tools for Post-Editing Research: The Current Picture and Further Needs (Poster)
Lucas Nunes Vieira

Online production of HQ parallel corpora and permanent task-based evaluation of multiple MT systems: both can be obtained through iMAGs with no added cost (Poster and demo)
Lingxiao Wang and Christian Boitet

Reverso (Demo)

16:25 – 17:30: Session 3

16:25 *Issues in Incremental Adaptation of Statistical MT from Human Post-edits*
Mauro Cettolo, Nicola Bertoldi, Marcello Federico,
Christophe Servan, Loïc Barrault and Holger Schwenk.

16:50 *The ACCEPT Post-Editing environment: a flexible and customisable online tool to perform and analyse machine translation post-editing*
Johann Roturier, Linda Mitchell and David Silva

17:15 Conclusion

Table of Contents

This translation is not too bad: An analysis of post-editor choices in a machine translation post-editing task
Maarit Koponen 1

What can we learn about selection mechanism for post-editing?
Maja Popovic, Aljoscha Burchardt, Eleftherios Avramidis and David Vilar 11

User Attitudes to the Post-Editing Interface
Joss Moorkens and Sharon O'Brien 19

Integrated Post-Editing and Translation Management for Lay User Communities
Adrian Laurenzi, Megumu Brownstein, Anne M. Turner and Katrin Kirchhoff 27

Community-based post-editing of machine-translated content: monolingual vs. bilingual
Linda Mitchell, Johann Roturier and Sharon O'Brien 35

Combining pre-editing and post-editing to improve SMT of user-generated content
Johanna Gerlach, Victoria Porro, Pierrette Bouillon and Sabine Lehmann 45

Advanced Computer Aided Translation with a Web-Based Workbench
Vicent Alabau, Ragnar Bonk, Christian Buck, Michael Carl, Francisco Casacuberta, Mercedes Garcia-Martinez, Philipp Koehn, Luis Leiva, Bartolome Mesa-Lao, Herve Saint-Amand, Chara Tsoukala, German Sanchis, Daniel Ortiz and Jesus Gonzalez 55

Quality as the sum of its parts: A two-step approach for the identification of translation problems and translation quality assessment for HT and MT+PE
Joke Daems and Lieve Macken 63

Transferring Source Tags to the Target Text in Statistical Machine Translation: A Two-Stream Approach
Eric Joanis, Darlene Stewart, Samuel Larkin and Roland Kuhn 73

Assessing Post-Editing Efficiency in a Realistic Translation Environment
Samuel Läubli, Mark Fishel, Gary Massey, Maureen Ehrensberger-Dow and Martin Volk 83

An Evaluation of Tools for Post-Editing Research: The Current Picture and Further Needs

Lucas Nunes Vieira 93

Online production of HQ parallel corpora and permanent task-based evaluation of multiple MT systems: both can be obtained through iMAGs with no added cost
Lingxiao Wang and Christian Boitet 103

Issues in Incremental Adaptation of Statistical MT from Human Post-edits
Mauro Cettolo, Nicola Bertoldi, Marcello Federico,
Christophe Servan, Loïc Barrault and Holger Schwenk. 111

The ACCEPT Post-Editing environment: a flexible and customisable online tool to perform and analyse machine translation post-editing
Johann Roturier, Linda Mitchell and David Silva 119

Welcome from MT Summit 2013 Workshop Chair

On behalf of the hosts of the 14th Machine Translation Summit I am delighted to welcome our participants to Nice for the 2nd Workshop on Post-Editing Technologies and Practice.

In recent years we have witnessed a significant increase in popularity of machine translation post-editing. It has a great potential in developing methodologies and computer tools for correcting the raw MT output to improve the quality of translation. MT post-editing is inherent to MT as such and it is very important that a special full-day workshop devoted to this issue runs in conjunction with he current MT Summit to further promote up-to-date achievements in this field.

I should like to take this opportunity to thank Sharon O'Brien, Michel Simard and Lucia Specia, the organisers of the Workshop, whose enthusiasm and commitment made this event a high standard scientific gathering with a wide range of activities like oral presentations, discussions, poster and demo sessions.

I am sure you will find the Workshop a stimulating opportunity for sharing knowledge and expertise, meeting colleagues and friends.

Enjoy the Workshop and the gorgeous city of Nice!

Svetlana Sheremetyeva.