

Linguistica Antverpiensia New Series – Themes in Translation Studies 8/2009. Walter Daelemans & Véronique Hoste: *Evaluation of Translation Technology*. 261 pages. ISBN 9789054876823

The journal “Linguistica Antverpiensia New Series - Themes in Translation Studies” is an annual publication devoted to language and translation studies. The 2009 issue, edited by Walter Daelemans and Veronique Hoste, with the subtitle “Evaluation of Translation Technology” is structured into three parts: Evaluation of Machine Translation, Evaluation of Translation Tools, and it contains in addition a section of six book reviews.

The two first parts of the volume bring together 10 contributions, each of which with around +-20 pages, covering implementation details of machine translation systems and computer assisted translation technology, reports on the usage, usability and reception of the technology in the lab and in field studies, as well as guidelines for evaluation and multidisciplinary analysis of translation technology adoption.

The first part, “Evaluation of Machine Translation”, contains six articles from well-known researchers in their field, covering a wide range of heterogeneous topics.

Andy Way criticizes the attitude of the Statistical Machine Translation (SMT) community and argues that SMT researchers should start communicating with linguists to include syntactic and other linguistically informed resources into their system design. According to him, such an inclusion of statistical and linguistic methods and communities was originally intended by the authors of the seminal IBM papers, but has been forgotten by (many of) the SMT actors today.

Vincent Vandeghinste describes two closely related hybrid MT systems, the METIS-II and the PaCo-MT system, both of which are inspired by SMT, RBMT (rule-based machine translation), and example-based approaches, thus integrating linguistic and corpus-based techniques. The systems differ in the kind of resources they use, such as mono- and bilingual corpora, parsers and transfer rules, and the translation quality of their output.

Two papers in this first part discuss particular MT error analyzes in detail. Bogdan Babych and Anthony Hartley present a method to detect systematically mistranslated multiword expressions (MWE) between French and English. Their method can also be used to highlight correctly translated MWEs. Another error analysis for translating English -ing forms with an RBMT system into various languages is presented by Nora Aranberri-Monasterio and Sharon O’Brien. They find that the translation of -ing forms are not as problematic for RBMT systems as one might think, but suggest using controlled languages at the content authoring stage and a semi-automatic method for post-editing. Both papers come to the same conclusion that BLEU (or other automatic translation evaluation metrics) correlates well with and is a good predictor for human judgements.

A further contribution from Paula Estrella, Andrei Popescu-Belis and Maghi King outlines the principles and resources for a framework of MT evaluation (FEMTI). They go through a number of examples and workflows to explain an interactive web interface which helps generate quality-based MT evaluation plans.

The last contribution in this first part from Lynne Bowker, outlines a detailed recipient evaluation study of machine translation usability for minority language communities in Canada. An English and a French speaking minority community in Quebec and Saskatchewan, respectively, are compared, their background and needs are analyzed and it is convincingly argued when and why (post-edited) MT output can be of help for those groups.

The second part, “Evaluation of Translation Tools”, comprises four contributions.

A paper of Lieve Macken assesses the usefulness of a sentence-based translation memory (SDL Trados) and compares statistics about segments and their matches with a sub-sentential translation memory system (Similis). While some text types (journalistic texts) are not suited for Translation Memories, a sub-sentential system can offer useful help for terminology and MWE translation.

A contribution by Alberto Fernandez Costales describes a software localization experiment using the "Pasolo" software. The paper points to a number of localization problems and concludes that CAT tools are suited for software localization, leading to more coherent and consistent translations in a shorter time.

Miguel A. Jimenez-Crespo investigates quality and consistency of localized web pages in Spanish, and compares them with spontaneously produced Spanish pages. This study shows that localized websites have lower linguistic quality, they have a higher number of lexical, syntactic and typographic inconsistency, and replicate source language structures.

The contribution by Iulia Mihalache presents a view on translation technology adaptation and technology usage from a social, economic, and psychological point of view. The author collects a number of statements from translation communities in order to classify the translators' perception and attitudes on translation technology, with the goal to develop a more complete, multidisciplinary understanding of evaluation criteria.

This broad variety of topics and themes provides a highly interesting insight into the diversity of research activities around machine translation and machine assisted translation, its development, deployment, reception and evaluation, from the translators' as well as from the users' point of view. It is a testimonial of the increasing impact and potentials that machine translation technology today has on the society and other research areas, such as psychology and sociology, inside, as well as outside the labs and specialized companies.

As can be expected in such a heterogeneous compilation of different papers, the methods and findings are not always uncontroversial. While each of the contributions is interesting in itself, giving detailed information about the studied topics, I felt at some places a lack of overall coherence. For instance, where Way's paper calls for more linguistically informed representation in corpus based translation research (and SMT), Vandeghinste actually presents such a translation system, without both taking notice of each other in their writing. Paula Estrella, Andrei Popescu-Belis and Maghi King introduce evaluation guidelines (i.e. the FEMTI framework), but while many contributors discuss evaluation techniques, none of the other chapters take up this framework in their research. Or, as the editors notice in their introduction, whereas Costales states that translation aides help produce more consistent texts, this is disproved in the study of Jimenez-Crespo. It would have been desirable to see more connectivity at this point and allow contributors to refer to each other's work which would have made the lecture of the volume somewhat easier and coherent. Some of the articles would have profited by explicitly pointing out their proximity to related contributions.

As a whole, the volume gives an insight into the diversity of research directions within and around the rapidly growing and changing field of machine translation and their evaluation techniques.

Michael Carl