

## Next Generation Translation Tools – project workshop abstracts

Swansea University, 19 July 2016

**Josep Bonet, DGT**

DGT has two main issues related to CAT tools. The most basic one is summarised in this proposal for a research study: Do CAT tools affect the way in which translators think and see the text? (Segmentation in thinking and the target text?); identify the practices that foster best the cohesion and coherence when using CAT tools or any available tools. The second issue is ergonomics or, to be more precise, the interaction between humans and machines allowing the former to be as productive as possible with minimal loss of concentration, quality and consistency.

With a population of over 1500 professional translators who are intensive users of CAT tools, ideas can be brought forward and developments can be tested in a real environment subject to everyday stress.

**Gatis Dilāns, Ventspils University College**

### **Medical terminology translation using the Internet: A look at a combined process of web-search and translation**

The study is a continuation of Dilāns (2015) which found that translation students while doing web search for new medical terminology translations basically relied on phonologic as well as transliteration of borrowings often using Google Autocomplete function for additional clues. By analyzing a combined process of web-search and translation in more detail, we are discovering the strategies and techniques of term search that may be applied in designing an automated web-based search module in e.g., SDL Trados (memoQ already has a basic tool for automated term search, but it can be improved).

The study employs methodological triangulation (screen-recording and verbal reports) to examine strategies in web-mediated translation process of medical terminology. Eight of the participants were medical students (new experts) and the other eight were translation students (new professionals) and they all had to perform the same task. The preliminary results show that even though the performance of the new experts was faster, it almost did not differ from the new translators' performance in terms translated terms and precision. The new experts relied more on their expert knowledge and Latin while the new professionals used their overall linguistic knowledge and search techniques.

Reference: Dilāns, G. (2015). Integrating technology in Latvian translation education: Untranslated medical terminology management practice using online resources and Computer-aided Translation tools. *The Journal of Specialized Translation*, 24, 44-60.

**Maureen Ehrensberger-Dow** & Gary Massey, ZHAW Institute of Translation and Interpreting

### **Why Ergonomics Matters in Translation**

Language technology, which continues to develop faster than researchers can investigate its use, can relieve translators of tedious tasks and increase efficiency. However, many translators do not exploit the potential of their tools, ignore many helpful features, and seem irritated by some of them. More worrying, poor workplace ergonomics may be detrimental to health, job satisfaction, and translation quality. Aspects of recent research done by our team will be presented to illustrate how an ergonomic perspective can contribute to a better understanding of how physical, cognitive, and organizational factors impinge on translators and their work.

**Kevin Flanagan**, SDL International

### **Fine-grained alignment, subsegment recall and fuzzy match repair**

A number of CAT tools offer features that can be categorised as 'fuzzy match repair', but there is a dearth of quantifiable results measuring their effectiveness. This paper will discuss match repair developments based upon doctoral work on fine-grained alignment and subsegment recall, introduce a methodology for measuring match repair effectiveness, then present results obtained through testing on large TMs and some opportunities for further research.

**Nicolas Froeliger**, Université Paris Diderot

### **Attempts at speaking the same language across disciplinary boundaries: the Tralogy conferences so far**

[nf@eila.univ-paris-diderot.fr](mailto:nf@eila.univ-paris-diderot.fr)

This presentation would be a brief outline of the reasons which let players from translation programs, the translation profession, users of translation and ALP specialists to create a forum where they could debate with one another instead of separately: the Tralogy (portmanteau word for translation and technology) conferences. The salient features of the two conferences organized so far (2011 and 2013, in Paris) would be discussed, as well as the necessity to breathe new life into his process, to devise potent research with practical import for the translation professions as well as for society as a whole, and last but not least, to have the whole thing fit into the larger translation-related research system.

**Joanna Gough**, University of Surrey

### **Ergonomics (cognitive and/or physical) of current translation tools: how can they be tested, what problems are there and what improvements are needed?**

It has been argued that translation technology serves as an extension of translator's own memory and contributes to the 'externalisation' of the translation process. This entails a shift in the deployment of cognitive skills - from identifying possible solutions to translation problems using generative cognitive processes to selecting between available solutions using selective cognitive

processes (Pym, 2011, Austermühl, 2013). My recent research into the use of online resources by professional translators (Gough, 2016) showed that half of the sample tended to 'economise' on their research and I hypothesise that this 'economising' tendency also applies to re-using the content of TM and MT. I propose to use the Qualityity plugin for SDL Trados Studio (Hartnett, 2015) to examine the level and nature of such re-used content/suggestions generated by TM/MT by translators in view of the shift in the cognitive processes resulting from using translation technology.

References:

Austermühl, F. (2013). Future (and not-so-future) trends in the teaching of translation technology, in: *Revista Tradumàtica: Tecnologies de La Traducció*, 11, pp. 326–337

Gough, J. (2016). The patterns of interaction between professional translators and online resources. PhD Thesis, University of Surrey

Hartnett, P (2015). Qualityity. Available at: <http://codificare.net/tools/qualityity/>

Pym, A. (2013). Translation skill sets in a machine translation age, in: *Meta* 58(3), pp. 487-503

**Katell Hernandez Morin**, Fabienne Moreau, Franck Barbin, David Le Roux, Octavia Effraim, Gaëlle Phuez-Favris, Chantal Quéniart, Jean-Marie Le Goff, Samuel Barbier - Université Rennes 2

### **Tackling Multiple Approaches in Adapting Translation Tools to Quality and Productivity Requirements**

The research led by the TRASILT team has taken various directions that may be of use in helping build a bid proposal on translations tools:

- 1) The team has led a three-year study on the translation quality and productivity of final-year master student translations. The experimental design was done in controlled conditions (limited time and sample length, « calibrated » student profiles) on translations done with TM, MT, SR and without translation technology.
- 2) To objectively and evenly assess the quality of translations, a multidimensional quality assessment grid was developed and tuned. The grid rests on professional and functional criteria and is particularly versatile.
- 3) Part of the team has also worked on extracting and structuring fine linguistic knowledge for semantic browsing in text collections; on new generation data treatment; and on combining MT and human translation for dialogue engine training.
- 4) The team is presently working on building a multilingual and multidomain collaborative terminology database system for translators.
- 5) Finally, some team members have a special experience in exploring alignment and memory-sharing functionalities, as well as terminology database and reviewer data integration in CAT tools for professionals (needs of a regional institution in French-Breton translation).

**Hendrik J. Kockaert**<sup>1,2</sup>, Winibert Segers<sup>1</sup>, Bert Wylin<sup>1,3</sup>, Dirk Verbeke<sup>3</sup>

KU Leuven<sup>1</sup>, University of the Free State<sup>2</sup>, Televic Education<sup>3</sup>

**TranslationQ (EvaluationQ & RevisionQ): Automated Translation Process with Real-time Feedback & Evaluation/ Revision with Preselected Items Evaluation (PIE) Method.**

**Can subjectivity be avoided in translation evaluation?**

The Preselected Items Evaluation method is a translation evaluation method which is not based on evaluating the entire target text, but on preselected items in the source text. PIE can help translation evaluators constrain subjectivity and arbitrariness.

RevisionQ is a software tool and is part of a comprehensive translation management tool, called TranslationQ. It works in a "TM" mode where the evaluator inserts comments and suggestions for a better translation and for re-use. Translators can use RevisionQ in an assistant mode when translating.

RevisionQ will benefit businesses, universities and public services on the account of: (1) Time-saving and efficient revision; (2) Transparent and objective evaluation; (3) Visible and justifiable revision statistics; (4) Consistency and overall equality in marking; (5) Statistics allowing future research and adequate training in the area of multilingual and monolingual content management.

**Mikhail Mikhailov**, University of Tampere

**TextHammer corpus tool: not only concordancing**

All the tools which are designed to query parallel corpora seem to concentrate mainly on concordancing. Other features are rarely offered. At the University of Tampere a new online tool is being developed to obtain data of different kinds from parallel corpora.

The application can generate various kinds of output: concordances, collocations, frequency lists, and statistics on texts. The texts are grammatically annotated with open source software. The user can also create subcorpora in order to perform searches on texts representing certain genres/types/authors.

The application can be used to study multilingual experimental data produced by our research group.

**Sharon O'Brien**, Dublin City University

**A Broad Overview of Ergonomics-Related Research at the CTTS in Dublin City University – 2006-2016**

The proposed presentation will provide a high-level overview of collaborative work done in the Centre for Translation and Textual Studies at Dublin City University, along with international collaborators, on the topic of Computer-Aided Translation (CAT) and ergonomics. CAT is broadly defined here to include translation memory tools and machine translation. Ergonomics (or human factors) is the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance ([www.hfes.org](http://www.hfes.org)). The spotlight is on translators as users of the technologies under scrutiny and end users of the text produced by those technologies. Our work to date has focused on cognitive effort

among translators (O'Brien 2006, 2011), UI design (O'Brien et al. 2015, Moorkens & O'Brien 2016, Teixeira & O'Brien 2016), perception (Moorkens et al. 2015), usability (Castilho & O'Brien 2016, Doherty & O'Brien 2014), and finally, the broader ergonomic issues among professionals and within organisations (Ehrensberger-Dow & O'Brien 2015, Cadwell et al. 2016). The objective is to summarise the work to date and to ask what needs to be focused on in the future and why this is important.

## References

- Sheila Castilho & Sharon O'Brien (2016) Evaluating the Impact of Light Post-Editing on Usability. [Poster Presentation (Refereed)], LREC 2016, Slovenia, 23-MAY-16 - 28-MAY-16.
- Moorkens, Joss & Sharon O'Brien (Forthcoming) (2016) 'Assessing User Interface Needs of Post-Editors of Machine Translation' In: Dorothy Kenny (eds). IATIS Yearbook 2015. UK: Routledge.
- Teixeira, Carlos & Sharon O'Brien (Forthcoming) (2016) 'Recent advances in translation tools from a cognitive ergonomics perspective', Conference Paper to be presented at the EST Conference, Aarhus, Denmark, September 2016.
- Cadwell, Pat, Sheila Castilho, Sharon O'Brien (Forthcoming) (2016) 'The Impact of Ergonomic Factors on Machine Translation (Non-)Adoption among Translators in the European Commission's Directorate-General for Translation', Conference Paper to be presented at the EST Conference, Aarhus, Denmark, September 2016.
- Ehrensberger-Dow, Maureen & Sharon O'Brien (2015) 'Ergonomics of the Translation Workplace: Potential for Cognitive Friction'. TRANSLATION SPACES, 4 :98-118
- Moorkens J., O'Brien S., da Silva I., de Lima Fonseca N., Alves F. (2015) 'Correlations of perceived post-editing effort with measurements of actual effort'. Machine Translation, :1-18
- O'Brien, Sharon, J. Moorkens, J. Vreeke & R. Schaefer (2015) Kanjingo: Considering the Ergonomics of Mobile Apps for Post-Editing. Traducteurs à l'œuvre : approches ergonomiques des pratiques professionnelles et des formations de traducteurs, Grenoble, France, 05-MAR-15 - 06-MAR-15.
- Doherty S., O'Brien S. (2014) 'Assessing the Usability of Raw Machine Translated Output: A User-Centered Study Using Eye Tracking'. International Journal of Human-Computer Interaction, 30 (1):40-51
- O'Brien, Sharon (2011) 'Towards predicting post-editing productivity'. Machine Translation, 25 :197-215
- O'Brien, Sharon, O'Hagan M., Flanagan M. (2010) Keeping an eye on the UI design of Translation Memory: How do translators use the Concordance feature? ECCE 2010 - European Conference on Cognitive Ergonomics 2010: The 28th Annual Conference of the European Association of Cognitive Ergonomics, pp.187-190
- O'Brien, Sharon (2006) 'Eye Tracking and Translation Memory Matches'. Perspectives: Studies in Translatology, 14 :185-205

My current research uses observation and interviewing to investigate the workplace practices of translators and project managers in LSPs and translation departments of other organisations. Drawing on practice-theoretical approaches (Schatzki 1996; Reckwitz 2002; Shove et al. 2012), I focus on the interdependencies between materials (including tools), meanings and competences in the practices of translation and project management. This research enables us to learn about the ergonomics of current translation tools and how the technologies shape and are shaped by workplace practices, across spaces and over time.

#### References

Reckwitz, Andreas. 2002. "Toward a Theory of Social Practices A Development in Culturalist Theorizing." *European Journal of Social Theory* 5 (2): 243–63.

Schatzki, Theodore R. 1996. *Social Practices: A Wittgensteinian Approach to Human Activity and the Social*. Cambridge: Cambridge University Press.

Shove, Elizabeth, Mika Pantzar, and Matt Watson. 2012. *The Dynamics of Social Practice*. London: Sage.

**Emmanuel Planas**, Université de Nantes, **Amelie Josselin-Leray**, Aurélie Picton, Université de Toulouse 2

#### **A Useful Feature for Translators: Knowledge Rich Contexts**

While comparable corpora have proved to be useful to translators (the CULT vogue: Baker 1993, Beeby 2012, Bowker 2004, Kübler 2004), there is a continuum of possible uses and techniques for exploiting them (e.g. for building monolingual or bilingual terminologies). In the CRISTAL project, we focused on providing translators with "good contexts" taken from comparable corpora, i.e. contexts selected as Knowledge Rich Contexts (KRCs, Meyer 2001) that provide useful conceptual and usage information. Several experiments based on 64 testers (trainee translators and professional translators) show that these KRCs do help translators during the translation process and in what respect.

Alice Carré, Cécile Frérot, Elisabeth Lavault-Olléon & **Caroline Rossi**, Université Grenoble Alpes (ILCEA4-GREMUTS)

#### **Ergonomics as a framework for translation teaching and translator training**

Through a series of conferences launched at the University of Grenoble and a partnership developed with colleagues at ZHAW, ergonomics emerged as a new paradigm for translation studies. We discuss the reasons which led colleagues to emphasise ergonomics (Lavault-Olléon, 2011) before outlining how this paradigm-shift fostered research impacting translation teaching and translator training in Grenoble. A first line of research in our team seeks to better characterise CAT tools with a view to matching professional and academic requirements (Frérot, C. Karagouch, L. to appear), and investigates the use of corpora and corpus technology and its impact on translation quality (Frérot, 2016). A second line of research identifies metacognition as a crucial methodological tool for translator training, and one which takes into account the evolution of professional environments (Lavault-Olléon, E. Carré, A. 2012). Finally, a more recent line of research is linked with the design

and implementation of a new statistical machine translation (SMT) class for M2 students: in line with Doherty and Kenny's qualitative evaluation (Doherty, S., Kenny, D. 2014) we designed pre and post questionnaires aimed at assessing evolutions in student's perceptions of SMT –or lack thereof.

Frérot, C. (2016). Corpora and corpus technology for translation purposes in professional and academic environments. Major achievements and new perspectives. *Cadernos de Tradução* [online], 36:2.

Frérot, C., Karagouch, L. (to appear). Outils d'aide à la traduction et formations de traducteurs : vers une adéquation des contenus pédagogiques avec la réalité technologique des traducteurs. *ILCEA* [online], 27.

Lavault-Olléon, E. (2011). L'ergonomie, nouveau paradigme pour la traductologie. *ILCEA* [online], 14.

Lavault-Olléon, E., Carré, A. (2012). Traduction spécialisée : l'ergonomie cognitive au service de la formation. *Asp* [online], 62.

Doherty, S., Kenny, D. (2014). The design and evaluation of a Statistical Machine Translation syllabus for translation students. *The Interpreter and Translator Trainer*, 8:2.

**Tomáš Svoboda**, ITS Prague

### **Enhancing PEMT performance in a synthetic/fusional (flective) language**

tomas.svoboda@ff.cuni.cz

I have used generic SMT since 2009 in my own translation practice (EN-CS and DE-CS). This way, I was able to develop a substantial understanding of the challenges involved in PEMT with a Slavonic language such as Czech.

In terms of Ergonomics, PEMT in a synthetic/fusional (flective) language is different from analytic languages. It actually involves occupational health hazards posed by the elevated frequency of monotonous keyboard strokes (speaking of my own experience here).

To avoid such risks and enhance PEMT performance, I have devised a set of commands (macros), including user-friendly shortcuts. These include ending convertors, word order switching operations, truncating commands, caps convertors, oppositors, and declinators (including across grammatical numbers). The main issue today is the fact that these features cannot be easily implemented a CAT environment, so I am working in MS Word and have to reimport the translated text into – e.g. – SDL Trados Studio each time.

I have a collection of MT error types and challenges for PEMT in DE-CS and EN-CS, which I could quickly refer to as well.

**Gareth Watkins**, Cardiff University.

### **Introducing a novel Translation Memory Tool Metric**

Abstract: Much research suggests that the usefulness of Translation Memory (TM) is dependent on the text being translated and/or the content of the TM database. These factors alone are not sufficient to measure or evaluate TM tool usefulness. In this paper, a highly customisable TM metric

is introduced, one which not only considers the text type and TM content, but also the skills of the individual translator who will be using the tool. The metric is discussed, along with potential future metric based research and applications.