

The next generation translator's workbench: post-editing in CASMACAT v.1.0

Bartolomé Mesa-Lao

bm.ibt@cbs.dk

Copenhagen Business School
Centre for Research and Innovation in Translation
and Translation Technology, Denmark

Aims of the study

- ❑ To identify, analyze and consolidate the needs of translation professionals with regard to post-editing.
- ❑ To uncover areas of improvement to expand the possibilities of the existing CASMACAT prototype.
- ❑ To reveal post-editing processes in need of:
 - greater automation;
 - better machine-based support.

The CASMACAT Project

- ❑ EU 7th framework project.
- ❑ Participants: Celer Soluciones SL, Universitat Politècnica de València, University of Edinburgh, and Copenhagen Business School.
- ❑ Overall goal is to “build the next generation translator’s workbench to improve productivity, quality, and work practices in the translation industry”.

The CASMACAT Workbench

- ❑ Web-based technology.
- ❑ Visual and functional resemblance to commercial CAT tools.
- ❑ Direct integration of state-of-the-art translation technology such as
 - exploitation of machine translation;
 - interactive translation prediction.
- ❑ Integration of new input devices (e-pen).
- ❑ Extensive logging and playback of translator behaviour (key-logging and eye-tracking).

First field trial description

- ❑ June-July 2012.
- ❑ 5 professional translators from Celer Soluciones SL.
- ❑ English to Spanish.
- ❑ Post-editing guidelines.
- ❑ 18 texts were both translated from scratch and post-edited (12,000 words in total).
- ❑ 1 translator monitored with eye-tracking technology.
- ❑ More than 30 hours of user activity data (process).



Proyecto CSMACAT

CELER

Es un proyecto europeo de I+D dentro del programa marco FP7-ICT-2011.4.2 Language Technologies

- El proyecto pretende investigar desde diferentes perspectivas, métodos innovadores de ayuda a la traducción humana que integran la traducción automática dentro del flujo de trabajo. Como resultado de esta investigación se desarrollará una herramienta de trabajo en línea que integre los conocimientos adquiridos y constituya la nueva generación de aplicaciones de traducción asistida.
- El proyecto tiene una duración de tres y comenzó en noviembre de 2011.
- El coste es de 3.216.696 euros, de los que el aporte de la CE es de 2.500.000 euros.

Participan

- THE UNIVERSITY OF EDINBURGH
- COPENHAGEN BUSINESS SCHOOL
- UNIVERSIDAD POLITÉCNICA DE VALENCIA
- CELER SOLUCIONES

CELER ha formado parte de los principales proyectos de investigación en traducción en los últimos 10 años.

Proyecto CSMACAT

Este proyecto europeo de I+D dentro del programa marco FP7-ICT-2011.4.2 Language Technologies pretende investigar desde diferentes perspectivas, métodos innovadores de ayuda a la traducción humana que integran la traducción automática dentro del flujo de trabajo. Como resultado de esta investigación se desarrollará una herramienta de trabajo en línea que integre los conocimientos adquiridos y constituya la nueva generación de aplicaciones de traducción asistida. El proyecto tiene una duración de tres y comenzó en noviembre de 2011. El coste es de 3.216.696 euros, de los que el aporte de la CE es de 2.500.000 euros.

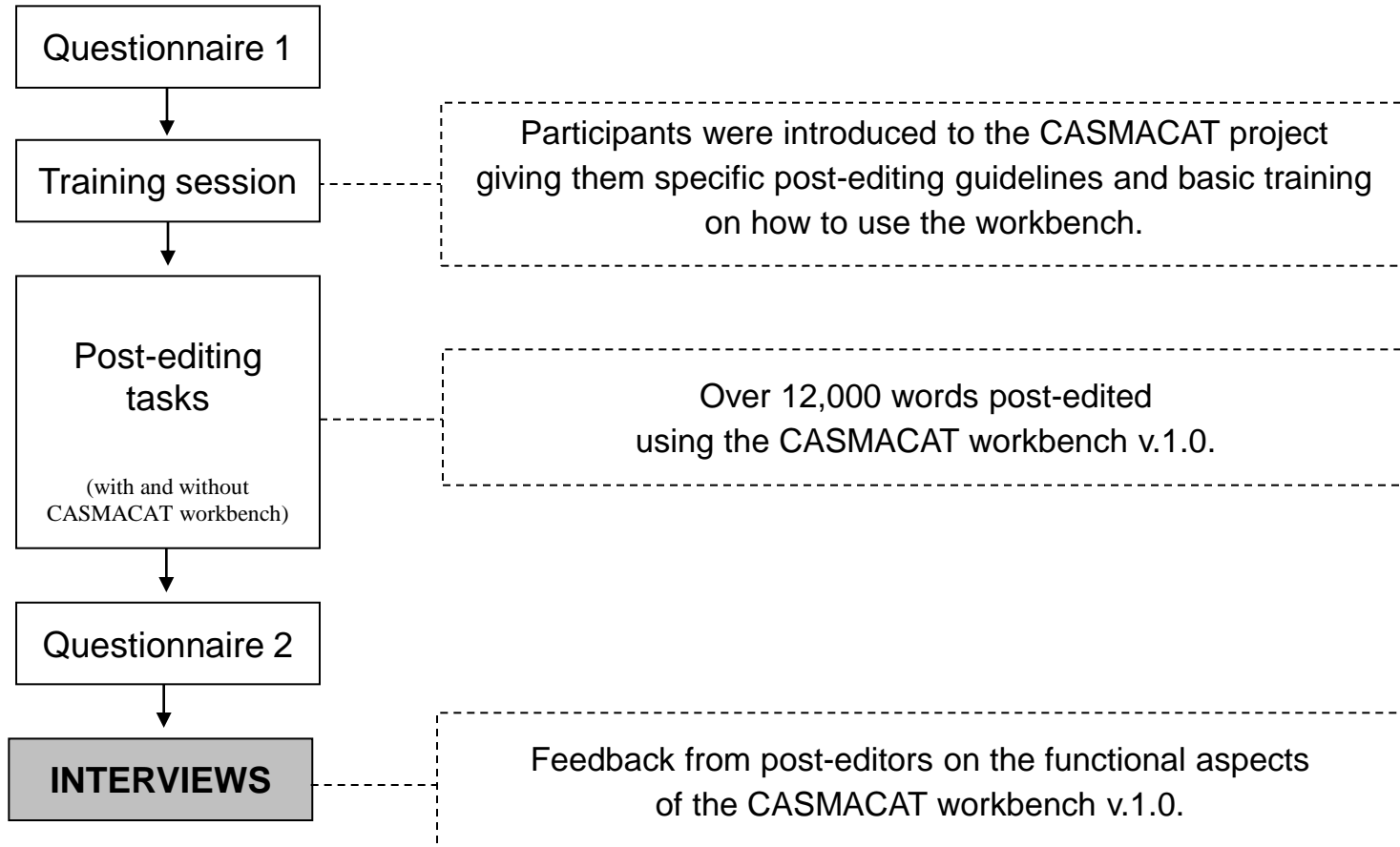
Participan

- THE UNIVERSITY OF EDINBURGH
- COPENHAGEN BUSINESS SCHOOL
- UNIVERSIDAD POLITÉCNICA DE VALENCIA
- CELER SOLUCIONES

CELER ha formado parte de los principales proyectos de investigación en traducción en los últimos 10 años.


Vasily Kandinsky

First field trial description



The CASMACAT Workbench (Prototype 1)

□ URL: <http://cbs.casmacat.eu/>


Close document

<p>907869 Tech Toys to Keep Your Gift Budget in Balance</p> <p>907870 While the leading candidate for this year's most marvelous toy is the iPad, followed by, perhaps, a Microsoft Kinect video game system, both can put a significant dent in your gift budget.</p> <p>907871 Consider that for the price of one iPad, you can buy five \$99 BMX bicycles, and that one Kinect game costs the same as 10 classic beaded jump ropes, or about 48 cans of Play-Doh.</p> <p>But there are ways to keep your gift budget in balance.</p> <p>907873 I recently visited three national toy sellers - a Wal-Mart, a Toys "R" Us and a BJ's Wholesale Club - to see what electronic toys I could buy for less than \$80.</p> <p>907874 It was immediately clear that Moore's Law has been at work; this year's technology toy is likely to have more power, while costing the same if not a bit less compared with last year's.</p> <p>907875 In addition, while there are generally fewer new toys with</p>	<p>1347826 Tech Toys to Keep Your Gift Presupuesto en equilibrio</p> <p>1347827 Mientras que el principal candidato para este año más maravilloso del juguete es el iPad, seguida de, quizás, un sistema de videojuegos Kinect de Microsoft, ambos pueden poner trabas significativas en su presupuesto de regalos.</p> <p>1347828 Consideran que por el precio de un iPad, usted puede comprar cinco bicicletas BMX 99 dólares, y que un juego Kinect cuesta lo mismo que 10 beaded clásico salto cuerdas, o</p> <p>Pero hay maneras de mantener su regalo de presupuesto en equilibrio.</p> <p>1347830 Recientemente he visitado tres vendedores de juguetes, Wal-Mart, Toys "R" Us y BJ 's Wholesale Club, para ver qué podría comprar juguetes electrónicos a menos de 80 dólares.</p> <p>1347831 Es evidente que la Ley de Moore ha sido en el trabajo; este año el juguete tecnológico es probable que tenga más poder, mientras que los costos de la misma, si no un poco menos en comparación con el del año pasado.</p> <p>1347832 Además, si bien en general es menos juguetes nuevos</p>
---	---

Methodology: semi-structured interviewing

- ❑ Individual retrospective interviews
- ❑ Venue: Celer Soluciones SL (Madrid)
- ❑ Interviewees: 5 professional post-editors
- ❑ Duration: 30 min approx.
- ❑ Grounded Theory approach

Methodology: semi-structured interviewing

□ STRUCTURED PART OF THE INTERVIEW

- Welcome and introduction.
- Signature of informed consent.
- General comments on the first field trial.
- CASMACAT workbench prototype 1:
 - Functional aspects:
 - The editor
 - The process of post-editing in the workbench.
 - Comments on the implemented functionalities:
 - » two-column layout;
 - » keyboard shortcuts;
 - » text formatting.
 - Comments on the futures functionalities to be implemented.

Methodology: semi-structured interviewing

□ UN-STRUCTURED PART OF THE INTERVIEW

- Room for open comments and feedback from the participants in regard to any post-editing issues that may arouse.

Findings: Wish list of functional aspects

POST-EDITING ENVIRONMENT (THE EDITOR)

- Visual track of changes
- Comments in rows
- Real-time word counters and progress bars
- Indicators of work performed

Findings: Wish list of functional aspects

PRODUCTIVITY ENHANCEMENT FEATURES

- Systematic search and replace
- Copy from source to target segments
- Sort segments
- Record of editing patterns / Macro support
- Autowrite/Autocomplete functions
- Automatic segment propagation
- Translation memory module (MT > TM cycle)
- Full glossary and black list integration

Findings: Wish list of functional aspects

QUALITY CONTROL

- QA checks
 - end-of-segment punctuation
 - correct numbers
 - double spaces (op.)
 - length of segments (op.)

Conclusions

- ❑ First field trial successful in uncovering post-editor's needs.
- ❑ Wish list to be implemented in future versions of the CASMACAT workbench (June 2013).
- ❑ Further research in post-editing processes is needed to inform the development of advanced translation assistance tools.
- ❑ Real needs should be the basis for a more successful interaction between the human post-editor and the technology in use.

The next generation translator's workbench: post-editing in CASMACAT v.1.0

Bartolomé Mesa-Lao

bm.abc@cbs.dk

Copenhagen Business School
Centre for Research and Innovation in Translation
and Translation Technology, Denmark