

---

# ANALYSIS OF THE FIRST FIELD TRIAL

Eva Marcos, Celer Soluciones

November 27th, 2012

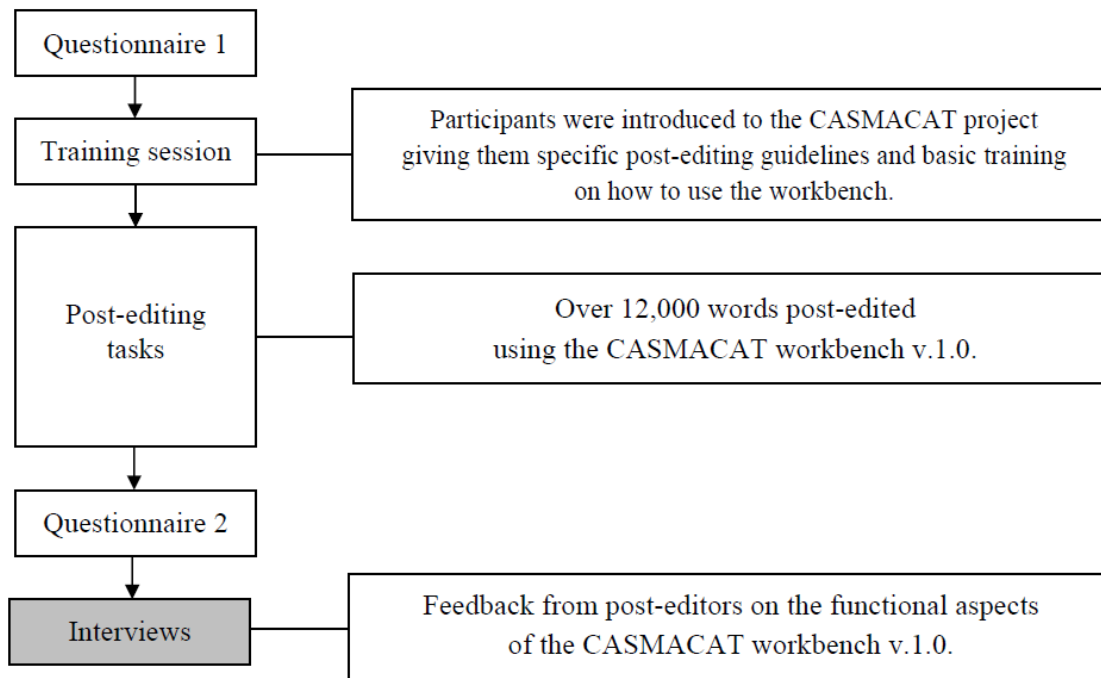


## COMPANY BACKGROUND

- The company has been using CAT tools and MT for more than a decade
- We mainly translate in Life Sciences, Institutional and Technical fields.
- Our most common language pairs are English-Spanish, Spanish-English.
- Participation in many R&D European and Spanish projects (Transtype 2, SMART, CASMACAT, EXPERT) has been a key factor

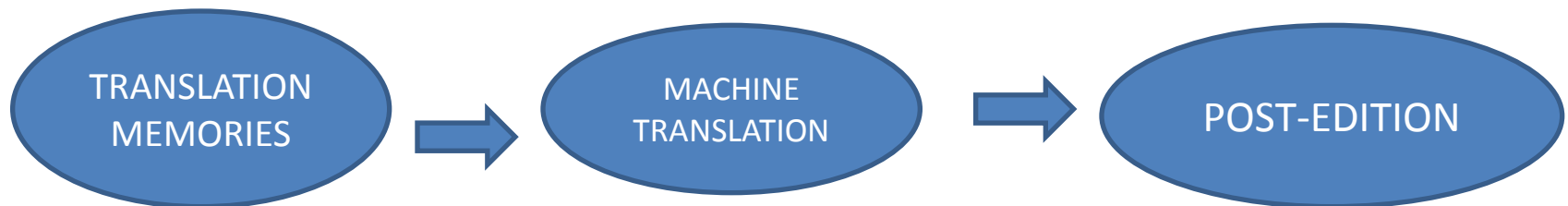
## OVERVIEW OF THE FIRST CASMACAT FIELD TRIAL

In Madrid (offices of Celer Soluciones) months of June and July 2012



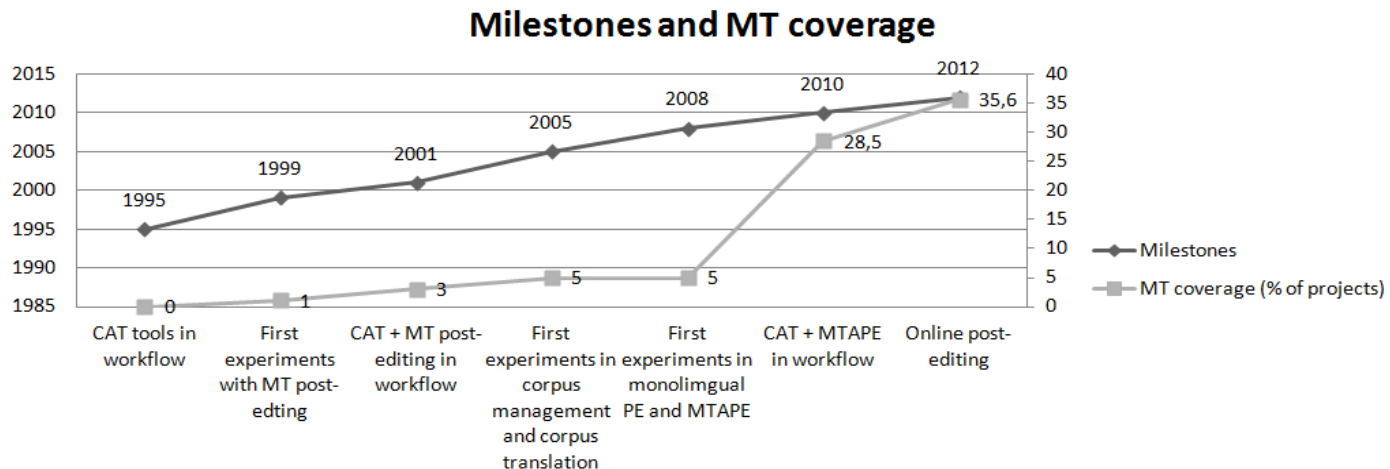
## MT AND POST-EDITING – GENERAL OVERVIEW AND CELER SOLUCIONES EXPERIENCE

- Translation services at greater speed and lower cost initially motivated MT+postedition
- Translation Memory Systems most important tools of translator support technology up until now.
- Correlating Translation Memories to post-editing -> aim for all language service providers.




## THE EXPERIENCE OF CELER SOLUCIONES

- Celer involved in working with MT and post-editing tools for a long time
  - Some figures for 2011: 16.41% projects – 30.72% words



## MAIN GOALS OF THE FIELD TRIAL

- **USER SATISFACTION AND TRANSLATION PRODUCTIVITY.**
- Collect feedback to implement the suggestions of post-editors in the second prototype of the CASMACAT workbench
- Human evaluation  real needs of the post-editing tool
- IDENTIFY, ANALYSE, CONSOLIDATE NEEDS

## SYSTEM AND TASKS

- MT system used based on Moses (UEDIN WMT 2012 evaluation campaign)
- TASKS → Translation of new stories (EN-ES) collected from CNN, Fox News, NY Times...
- System was trained
  - 4 Million word News Commentary
  - 57 million word Europarl
  - 319 million word United Nations parallel Corpora
  - Additional 248 million words of monolingual news language model data on 1,062 million words from the Spanish Gigaword corpus.
- It achieved a BLEU score of 32.9 (case –sensitive) on the 2012 test set
- Analysis by professional translator at Celer Soluciones = mostly useful for post-editing

# METHODOLOGY

Individual interview (30m) with all the post-editors (semi-structured)

User guide

Informed consent

(end of presentation)

Room for additional questions, suggestions, ideas...

## Informed consent

In this experiment your keyboard and mouse activity will be logged. Remember that your participation is confidential and none of the information logged will identify you by name.

I have read the foregoing information. I have had the opportunity to ask questions about it and any questions that I have asked have been answered to my satisfaction. I consent voluntarily to participate as a participant in this research.

– Name:.....

– Date: .....

– Signature: .....



## GUIDELINES

*Post-editors should always read the source segment first to understand the meaning of the sentence and apply the following guidelines:*

- *Use as much of the raw MT output as possible.*
- *Aim for a grammatically, syntactically and semantically correct translation.*
- *Don't worry if style is repetitive.*
- *Ensure that key terminology is correctly translated.*
- *Ensure that no information has been accidentally added or omitted.*
- *Basic rules regarding spelling, punctuation and hyphenation apply.*
- *Don't worry about formatting (rules for bold or italics should not be applied).*
- *Make changes only where absolutely necessary, i.e. correct words or phrases that are nonsensical, wrong or ambiguous.*

*These guidelines should NOT be applied if:*

- *Raw MT does not make any sense and it would take longer to post-edit than to translate from scratch.*
- *There are multiple errors that require rearranging most of the text.*
- *In such cases post-editors should proceed to translate from scratch.*

## PROFILE OF PARTICIPANTS

All six participants, professionals working for CS

<i>Participants</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>
<i>Gender</i>	<i>F</i>	<i>F</i>	<i>M</i>	<i>F</i>	<i>F</i>	<i>M</i>
<i>Years of translator training</i>	<i>5</i>	<i>4</i>	<i>1</i>	<i>2</i>	<i>1</i>	<i>3</i>
<i>Years of professional experience as translators</i>	<i>7</i>	<i>7</i>	<i>+2 0</i>	<i>12</i>	<i>15</i>	<i>10</i>
<i>Previous experience in post-editing</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>

## PARTICIPANTS' SUGGESTIONS/COMMENTS

- Fixing of repetitive errors
- 4 out of 6 preferred a TM interface to perform post-editing tasks
- Auto write features
- Search-replace
- Quality control checks

Working environment → variable but in general preferring CASMACAT workbench

I	SOURCE	TARGET
34433	It can be your child or your blood pressure.	Puede ser su hijo o la tensión arterial.
34434	It can certainly help you plan your response.	Sin duda, planificar la respuesta lo hace.
34435	It can occur at any age - even childhood.	Puede aparecer a cualquier edad, incluso en la infancia.
34436	It can only be administered alone.	Sólo puede administrarse en monoterapia.
34437	It can only be administered in combination with a statin.	Sólo puede administrarse en combinación con una estatina.
34438	It can result in paralysis	casi siempre produce parálisis
34439	It can retain the downloadable pdfs.	Puede contener los PDF descargables.
34440	It carries impulses from the medulla to the muscles of the throat, respiratory passages, lungs, esophagus, heart, and many other visceral organs, and conveys sensory input from these sites to the brain (Bear, p. 235).	Transporta impulsos desde el bulbo raquídeo hasta los músculos de la garganta, las vías respiratorias, los pulmones, el esófago, el corazón y muchas otras vísceras, y conduce hasta el encéfalo la información sensitiva procedente de estos puntos (Bear, pág. 235).
34441	It comes in the following dosage strengths	Se presenta en las dosis siguientes.
34442	It concludes with a description of the safety/tolerability profile of beta-blockers.	Concluye con una descripción del perfil de seguridad y tolerabilidad de los betabloqueantes.
34443	It concludes with a discussion of clinical trials related to aldosterone in heart failure.	Termina con una exposición de los ensayos clínicos relacionados con la aldosterona en la insuficiencia cardíaca.
34444	It consists of five statements measured on a 7-point Likert scale (i.e., "completely agree" to "completely disagree").	Consta de cinco frases medidas en una escala de Likert de 7 puntos (es decir, comprendid entre "totalmente de acuerdo" y "en total desacuerdo").
34445	<b>It contains 35 items which assess five domains:</b>	<b>Contiene 35 apartados que evalúan cinco dominios:</b>
34446	It contains a full explanation of the clinical research study your child is being invited to take part in and a consent form that you will be asked to sign if you want your child to participate.	Contiene una explicación completa del estudio de investigación clínica en el que se invita a participar a su hijo y un impreso de consentimiento que deberá firmar si quiere que su hijo participe.
34447	It contains a full explanation of the study you are being invited to take part in and a consent form that you will be asked to sign if you decide to take part.	Contiene una explicación completa del estudio en el que se le invita a participar y un impreso de consentimiento que se le pedirá que firme si decide participar.
34448	It contains a logical flow of information to demonstrate to key stakeholders the clinical and	Ofrece un flujo lógico de información para mostrar a los principales interesados los

## FUNCTIONAL ASPECTS

- Track changes
- Insertion of comments
- Monitor of progress
- Indicators on work performed

## POTENTIAL PRODUCTIVITY IMPROVEMENTS

- SEARCH AND REPLACE
- COPY TEXT FROM SOURCE TO TARGET SEGMENTS.
- SORT OF TEXTS ALPHABETICALLY
- RECORD OF EDITING PATTERNS
- MACRO SUPPORT
- AUTOWRITE SEGMENTS
- TRANSLATION MEMORY MODULE
- GLOSSARY INCLUDING BLACK AND WHITE LISTS
- QUALITY CONTROL

## NON-FUNCTIONAL ASPECTS

### USABILITY

- *User Interface: menu, tool bar and search an replace areas*
- *Customisability: font size, types...*
- *Learnability: training and help tools*
- *Supportability: documents, help tool..*

### EFFICIENCY: speed

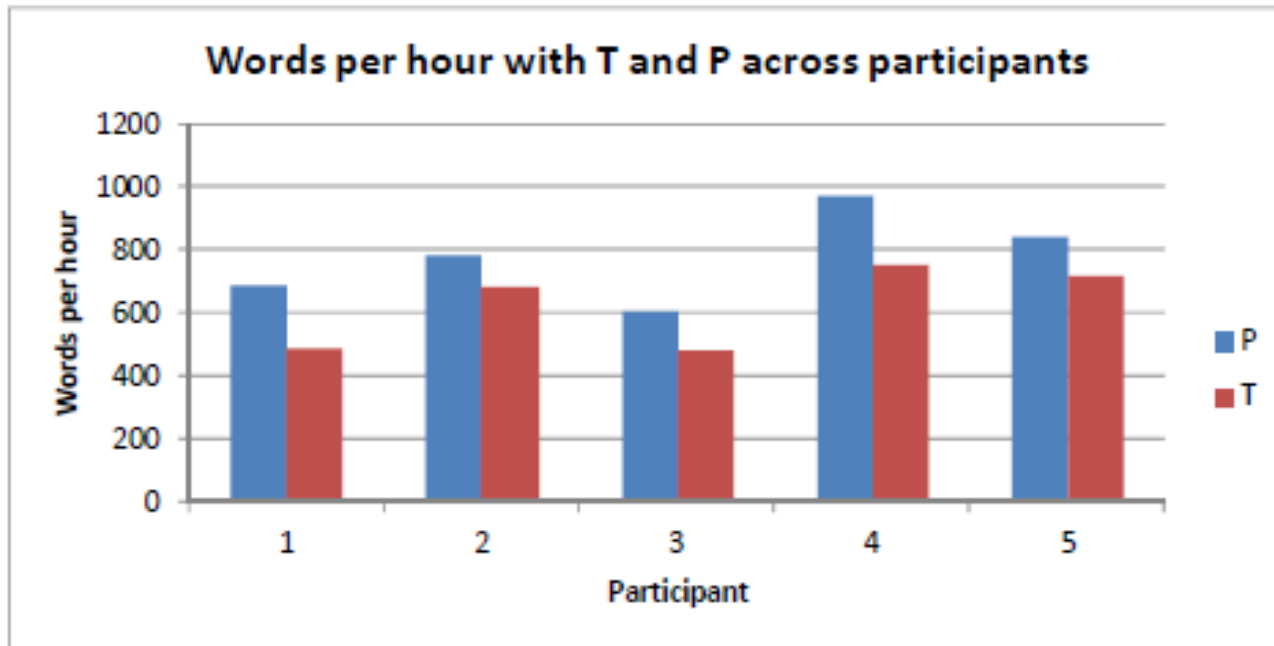
### EXTENSIBILITY: new mechanisms (speech recognition?)

### PORTABILITY: Mozilla Firefox only, mobile devices.

### RELIABILITY: in 2 cases the system crashed while working.

## DATA ON PRODUCTIVITY

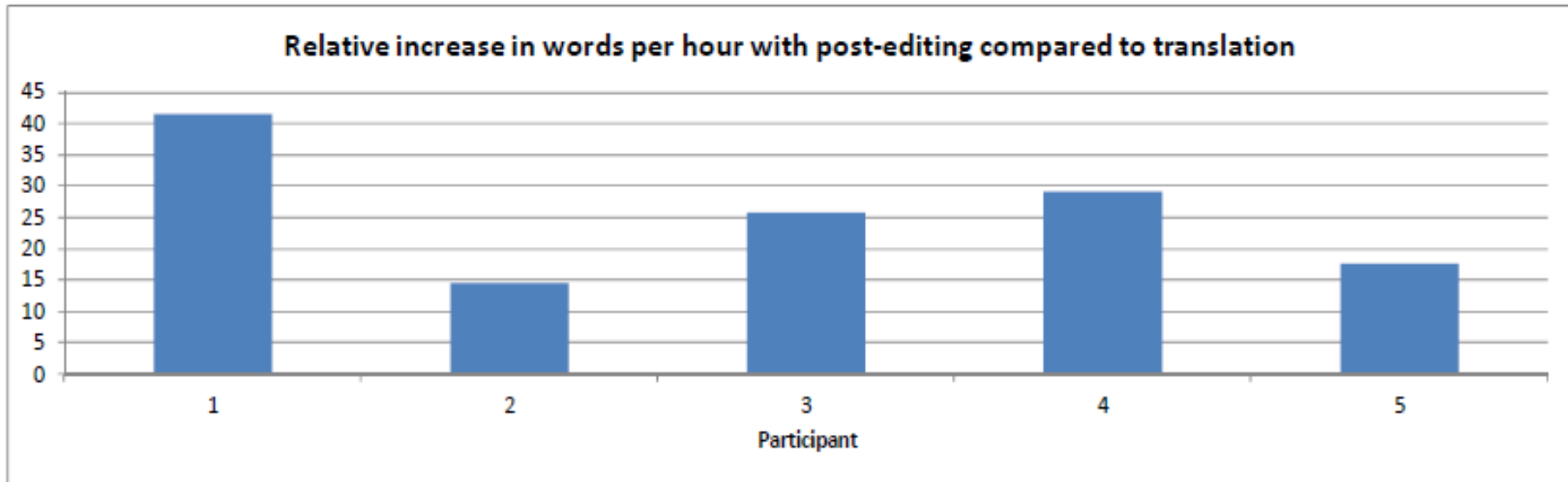
1) subjects that translate faster also post-edit faster





## DATA ON PRODUCTIVITY

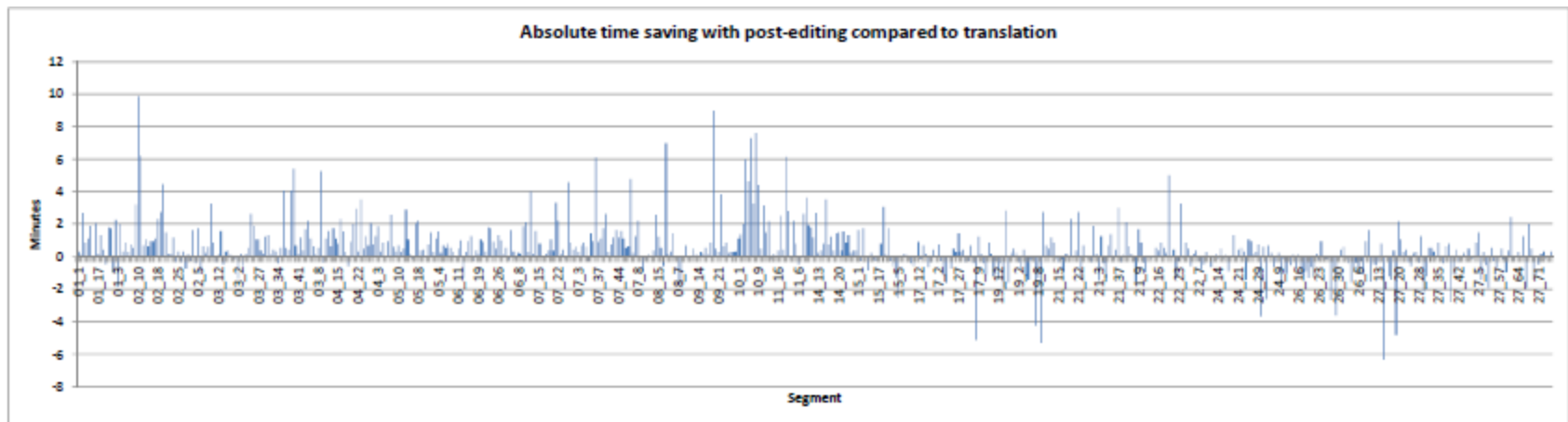
2) some translators however exhibit larger speed-ups than others (15 > 40 % speed-ups)



- Findings based on 1 post-editing text (30 segments) and 3 translation texts (115 segments)

## DATA ON PRODUCTIVITY

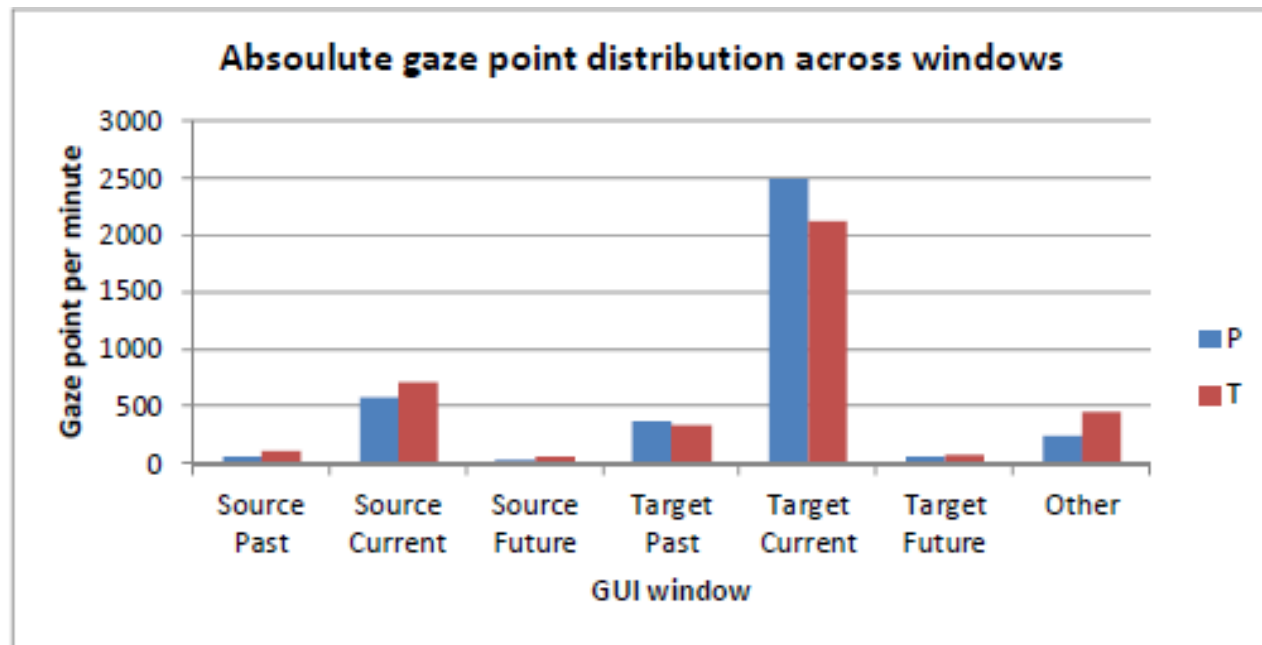
3) Post-editing produces a substantial time saving compared to translation (15-40%)



- Findings based on the 590 segments that were both post-edited and translated by a total of 2 to 4 participants each.

## EYE-TRACK

3) Translation and post-editing show about the same number of gaze-points per minute on average but post-editing spends more time on the target segment being created while translation spends more time on the source and off-screen



- Findings based on only 1 participant but a fairly large amount of data: 8 post-edited files (178 segments, 1.2 million gaze points), 9 translated files (298 segments, 3.7 million gaze points)

## FINAL CONCLUSIONS

- The field trial provided quality information on needs and possible future features (functional and non-functional)
- KEY CONCEPTS:
  - 1 - Flexibility
  - 2 – Simplicity
  - 3 – Accesibility

All data and translation system freely available for other researchers

## SOME COMMENTS...

*"...I would really appreciate it if a visible cursor could be included, as well as format marks. A search tool allowing to perform searches throughout the text is also needed. It would be recommended to include a search and replace option..."*

*"...I would like to have an overview of the full text while translating, to know how long it is and the number of words I have translated/have to translate in every moment..."*

*"...I would like to have the possibility to scroll up and down, without having to go through each segment..."*

*"..Sometimes it was difficult to translate without having access to the whole document, just to get the gist of it. Also, the segment window was too small for the text, which appeared in large-size letters making it difficult to see the whole segment in the window..."*

# Script of the interview

## [STRUCTURED PART OF THE INTERVIEW]

- Welcome and introduction.
- Signature of informed consent in order to be able to record the interview.
- General comments on the first field trial performed with the CASMACAT workbench prior to the interviews.
- 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 (font size, etc.).
- Comments on the desired functionalities to be implemented in future versions: *Departing from previous experiences in any TM system, which specific functions would you like to see implemented in a post-editing tool?*

### Non-functional aspects:

- Report on the usability, customisability, learnability and supportability of the GUI.

## [UNSTRUCTURED PART OF THE INTERVIEW]

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

## BIBLIOGRAPHY

- Mesa-Lao, Bartolomé (2012). "The next generation translator's workbench: post-editing in CASMACAT v.1.0" . Proceedings of the 34th Translating and the Computer Conference. 29 & 30 November 2012. ASLIB - The Association for Information Management, London.



Thank you!