

## Ninth MT Marathon

Trento, September 8th-13th, 2014

# Interactive(-predictive) Machine Translation

Francisco Casacuberta

fcn@prhlt.upv.es

Pattern Recognition and Human Language Technology

Universitat Politècnica de València



Ninth MT Marathon

Interactive(-predictive) Machine Translation

## Acknowledgment

This talk collects the work carried out by the UPV on IMT, including part of the work developed by the CASMACAT\* consortium.

An special mention to the UPV team in CASMACAT (by alphabetic order): Vicent Alabau, Jesús González-Rubio, Luis Leiva, Daniel Ortiz-Martínez and Germán Sanchis-Trilles. And the collaboration of: José-Miguel Benedí, Mara Chinaea, Jorge González, Pascual Martínez-Gómez and Enrique Vidal. Under the supervision of Francisco Casacuberta.

\* The  project has been co-funded by the European Union under the Seventh Framework Programme Project 287576 (ICT-2011.4.2) 

## Index

- 1 *Machine translation and computer-assisted translation* ▷ 2
- 2 Interactive machine translation ▷ 6
- 3 Adaptation in interactive machine translation ▷ 33
- 4 Multimodal interactive machine translation ▷ 42
- 5 Conclusions ▷ 54
- 6 Bibliography ▷ 56

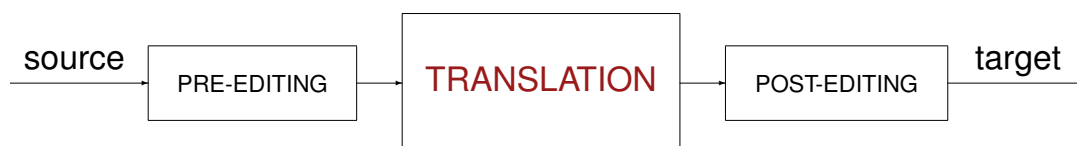


November 7, 2014

2

## Machine translation (MT)

- Existing MT technologies are currently seen as promising approaches to help produce high-quality translations (HQT) cost-effectively.
- However, the current state of the art in machine translation is still very far from allowing fully automatic HQT.
- (Pre-) post-editing



- While the number of errors and bad constructions is high, “post-editing” can make the result useful.
- Many problems could have been avoided by making the source text “simpler”.

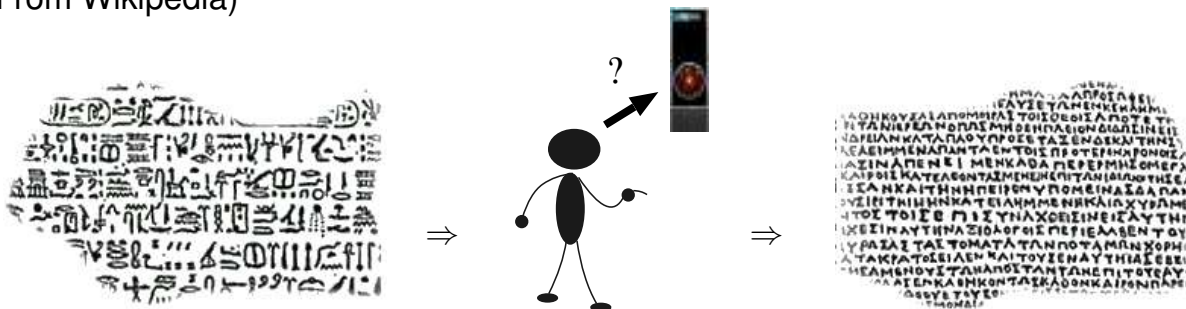


November 7, 2014

3

## HQT: Computer-assisted translation (CAT)

- CAT is a form of translation wherein a human translator translates texts using computer software designed to support and facilitate the translation process. (From Wikipedia)



- Other names: computer-aided translation, computer-assisted translation, machine-aided translation, machine-assisted translation.
- Historically, CAT and MT have been considered as different but close technologies [Kay, MT 1997]. Nowadays, MT is an important part of CAT.



November 7, 2014

## HQT: Human-machine interaction (HMI)

- In classical MT and CAT, the interaction between human translators and machines is very limited.
- At the origin: HMI focused on disambiguation of the source text or for updating user dictionaries or for searching through dictionaries [Slocum, CL 1985] [Whitelock et al., COLING 1986].
- More recently, a MT system is used to produce sentence hypotheses that are post-edited by the human translator.
- **Novel idea:** **interactive-predictive MT** aims to increase the overall (MT + human) productivity by incorporating human correction activities within the translation process itself [Foster et al., MT 1997] [Barrachina et al., CL 2008].
- **Related work:** A method of interactively visualizing and directing the process of translating a sentence [DeNeefe et al. ACL 2005].



November 7, 2014

## Index

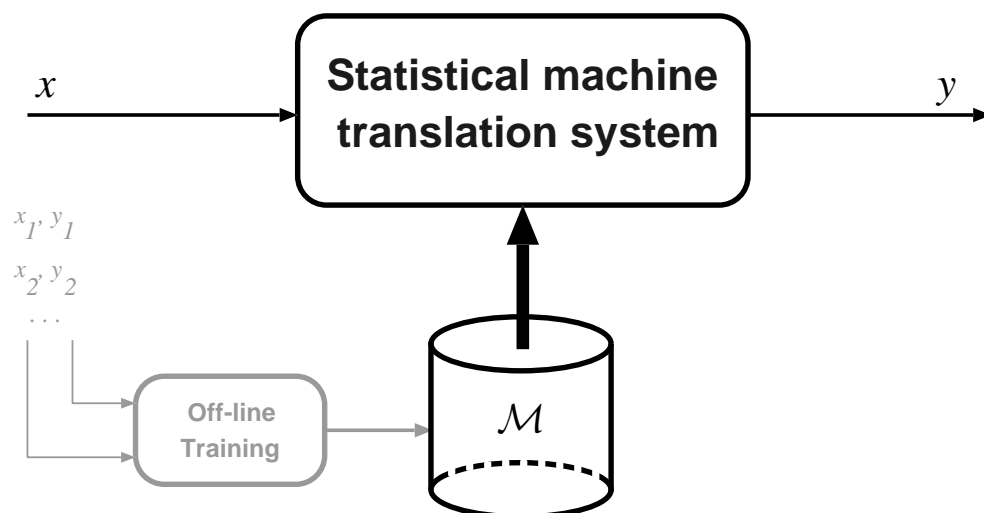
- 1 Machine translation and computer-assisted translation ▷ 2
- 2 *Interactive machine translation* ▷ 6
- 3 Adaptation in interactive machine translation ▷ 33
- 4 Multimodal interactive machine translation ▷ 42
- 5 Conclusions ▷ 54
- 6 Bibliography ▷ 56



November 7, 2014

6

## Diagram of a SMT system



November 7, 2014

7

## Post-editing: example

Translating the source sentence “*Click OK to close the print dialog*” into Spanish (the reference is “*Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión*”):

System: Haga clic para cerrar el diálogo de impresión

User: Haga clic **en** para cerrar el diálogo de impresión

User: Haga clic **en ACEPTAR** para cerrar el diálogo de impresión

User: Haga clic **en ACEPTAR** para cerrar el **cuadro** diálogo de impresión

User: Haga clic **en ACEPTAR** para cerrar el cuadro **de** diálogo de impresión

User: Haga clic **en ACEPTAR** para cerrar el cuadro de diálogo de impresión

Result: Haga clic **en ACEPTAR** para cerrar el **cuadro de** diálogo de impresión

TOTAL: Four word-strokes



November 7, 2014

8

## Introduction to interactive machine translation (IMT)

### *Main idea in IMT:*

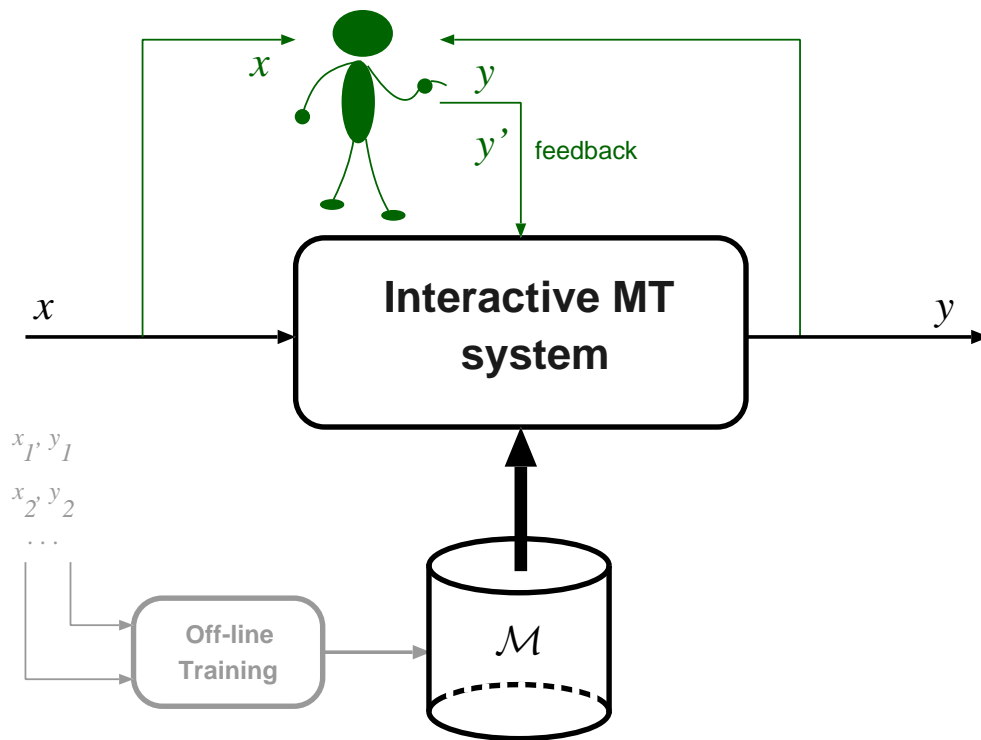
Use a MT system to produce target text segments that can be accepted or amended by a human translator; these correct(ed) segments are then used by the MT system as additional information to achieve further, hopefully improved suggestions [Foster et al., MT 1997] [Barrachina et al., CL 2008][Casacuberta et al., CACM 2008].



November 7, 2014

9

### Diagram of an interactive MT system



### Human-machine (keyboard) interactive process: example

Translating the source sentence ( $x$ ) “Click OK to close the print dialog” into Spanish (the reference is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”):

System ( $\hat{y}_s$ ): Haga clic para cerrar el diálogo de impresión

User ( $y_p$ ): Haga clic en **en**

System ( $\hat{y}_s$ ): **ACEPTAR** para cerrar el diálogo de impresión

User ( $y_p$ ): Haga clic en ACEPTAR para cerrar el **cuadro**

System ( $\hat{y}_s$ ): **de diálogo de impresión**

User ( $y_p$ ): Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión

Result ( $\hat{y}$ ): Haga clic **en** ACEPTAR para cerrar el **cuadro** de diálogo de impresión

TOTAL: Two word-strokes



## Human-machine (keyboard) interactive process

- In each iteration, a correct prefix ( $y_p$ ) of the target sentence is available and the IMT system computes its best (or  $N$ -best) translation suffix hypothesis ( $\hat{y}_s$ ) to complete this prefix.
- Given  $y_p \hat{y}_s$ , the IMT cycle proceeds by letting the user establish a new, longer acceptable prefix.

This prefix is typically formed by  $y_p$ , followed by an initial part of  $\hat{y}_s$  *accepted* by the user, followed by text obtained by means of additional user keystrokes generally aimed to amend remaining incorrect parts of  $\hat{y}_s$ .

This prefix becomes a new  $y_p$ , thereby starting a new IMT prediction cycle

- Ergonomics and user preferences dictate exactly when the system can start its new cycle, but typically, it is started after each user-entered word or even after each new user keystroke.



## Interactive machine translation: the original idea

- These ideas were studied in the *TransType* (TT) project [Foster et al. EMNLP 2002] and have been thoroughly explored in the *TransType-2* (TT2) project [Barrachina et al. CL 2008], in the *MIPRCV* project [Toselli et al. 2011], in the *Caitra* system [Koehn ACL 2009] and in the *CASMACAT* project [Alabau et al. PBML 2013].
- In TT, the IMT system suggests the best target word that follows the given prefix, however in TT2, the IMT system suggests the best complete suffix. In MIPRCV simple features were added to IMT and in CASMACAT, new advanced features have been tested.



## Text prediction for interactive machine translation

- In each iteration and given a source text  $x$  and a “correct” *prefix*  $y_p$  of the target text, search for a *suffix*  $\hat{y}_s$ , that maximizes the posterior probability over all possible suffixes:

$$\hat{y}_s = \operatorname{argmax}_{y_s} p(y_s | x, y_p)$$

- Taking into account that  $p(y_p | x)$  does not depend on  $y_s$ , we can write:

$$\begin{aligned} \hat{y}_s &= \operatorname{argmax}_{y_s} \frac{p(y_p y_s | x)}{p(y_p | x)} \\ &= \operatorname{argmax}_{y_s} p(y_p y_s | x) \end{aligned}$$

- $p(y_p y_s | x)$ : Log-linear models (or SFST if  $p(y_p y_s, x)$  is used)
- Text-input MT is a particular case, where  $y_p = \lambda$
- Main difference of IMT vs. MT: **search over the set of suffixes**



## Interactive machine translation: Search

High speed is needed because typically a new system hypothesis must be produced in real time after each user keystroke [Och et al. EACL 2003] [Barrachina et al. CL 2008].

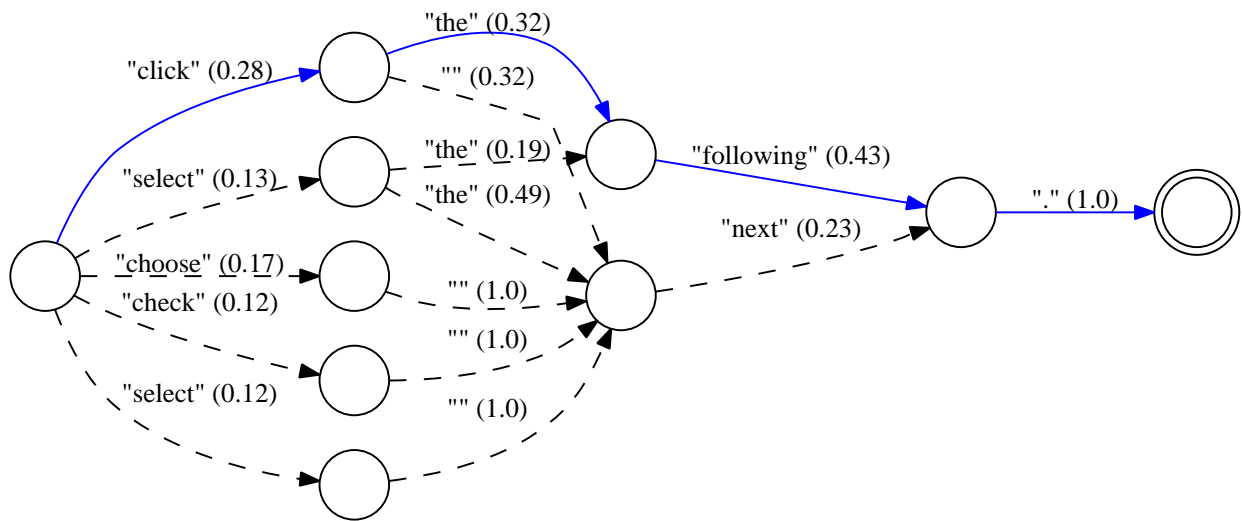
### WORD-GRAPH BASED APPROACH:

- For each source sentence, a *graph representing all its possible translations according to the translation model is generated.*
- *In each IMT iteration, the word-graph is searched for a best path compatible with the prefix given in this iteration.*
- *Error-correcting smoothing (edit distance) is used to allow for user-given prefixes that may not exist in the word-graph.*
- *Computation is carried out in an incremental manner: in each iteration the results from the previous iteration are updated.*





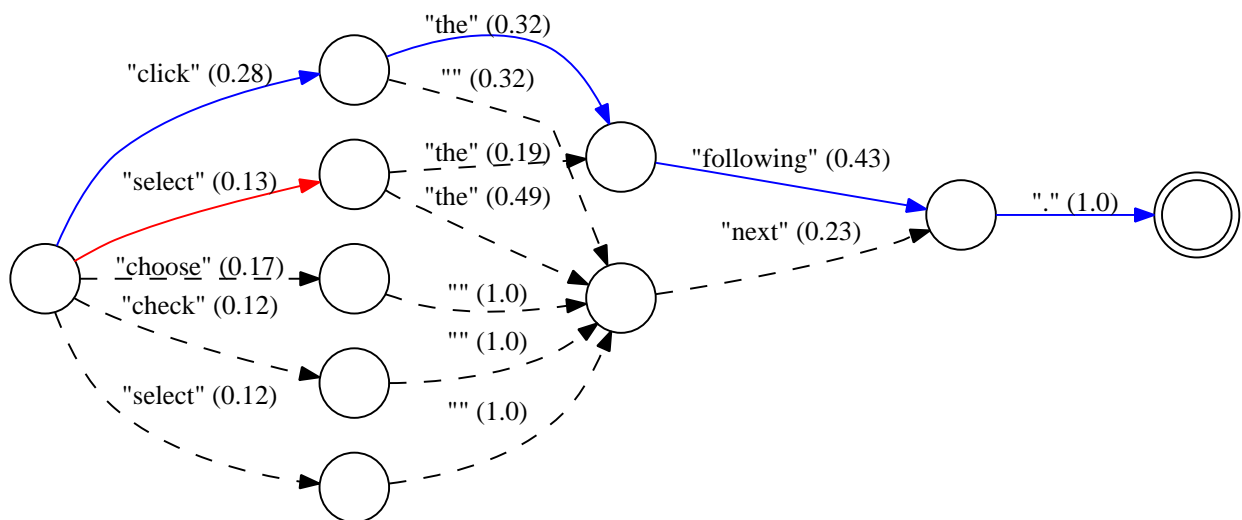
### IMT search



Translation word-graph corresponding to the source “seleccionar el siguiente”



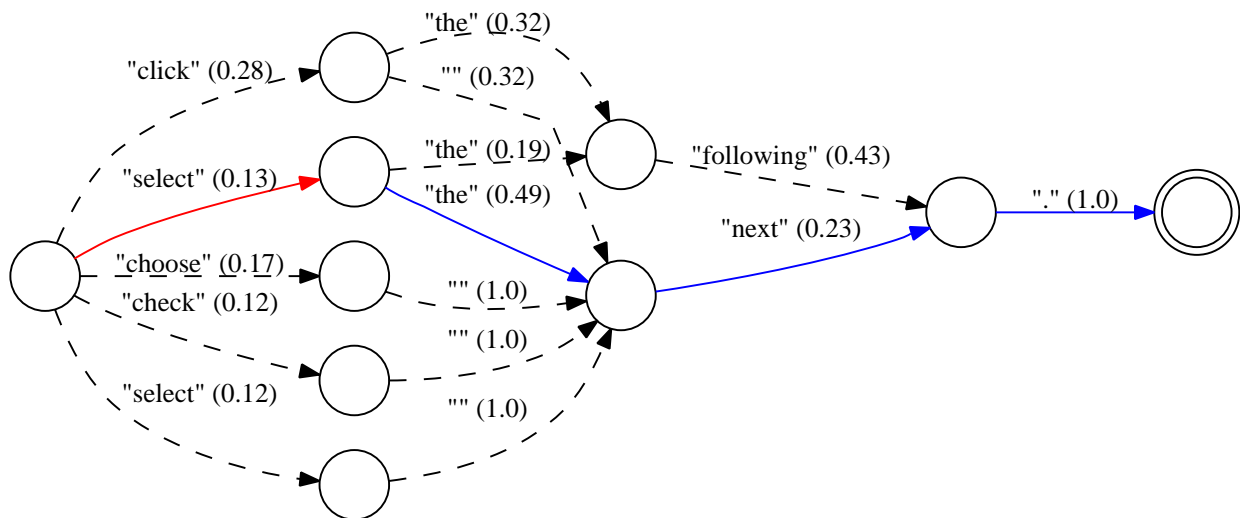
### IMT search



Translation word-graph corresponding to the source “seleccionar el siguiente”



## IMT search



Translation word-graph corresponding to the source “seleccionar el siguiente”



November 7, 2014

18

## Evaluating MT and IMT systems

### FOUR MEASURES:

- **TRANSLATION WORD ERROR RATE (TWER):**  
Minimum number of *word* insertions, deletions and substitutions needed to edit the system output into a (single) target reference
- **WORD-STROKE RATIO (WSR):**  
Number of user interactions that are necessary to achieve the reference target divided by the number of running words. In each interaction only one wrong word is changed
- **TRANSLATION CHARACTER ERROR RATE (CER):**  
Minimum number of *character* insertions, deletions and substitutions needed to edit the system output into a (single) target reference
- **KEY-STROKE RATIO (KSR):**  
Number of key-strokes that are necessary to achieve a (single) target reference divided by the number of running characters. In each interaction only one character is changed



November 7, 2014

19

## Summary of IMT experimental results

[Barrachina et al. CL 2008]

- Benchmark corpora: Xerox printer manuals (XRCE2) (Train: 50K bisentences and test 1K sentences) and EU bulletin (EU) (Train: 200K bisentences and test 800 sentences). English-Spanish, English-German and English-French.
- Results:

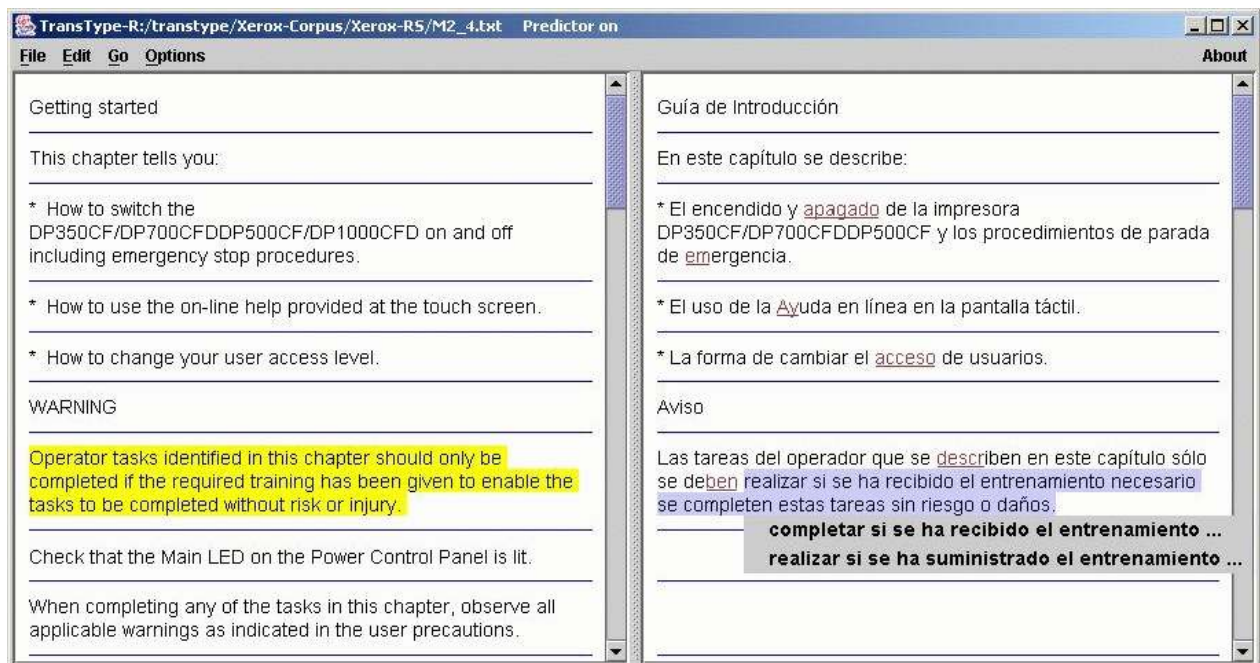
DATA: XRCE2	SFST (1-best)	
	WSR	TWER
English-Spanish	27.4	43.1
Spanish-English	31.7	51.4
English-French	65.1	73.8
French-English	58.5	71.9
English-German	55.4	81.3
German-English	55.0	78.5

DATA: EU	SFST (1-best)	
	WSR	TWER
English-Spanish	52.1	55.8
Spanish-English	48.5	52.5
English-French	62.2	53.9
French-English	60.5	49.2
English-German	49.6	65.5
German-English	44.0	66.6

Similar results were achieved with phrase-based models



## Human evaluation: A typical session in TT2



# Human evaluation in TT2 project

[Casacuberta et al. CACM 2008]

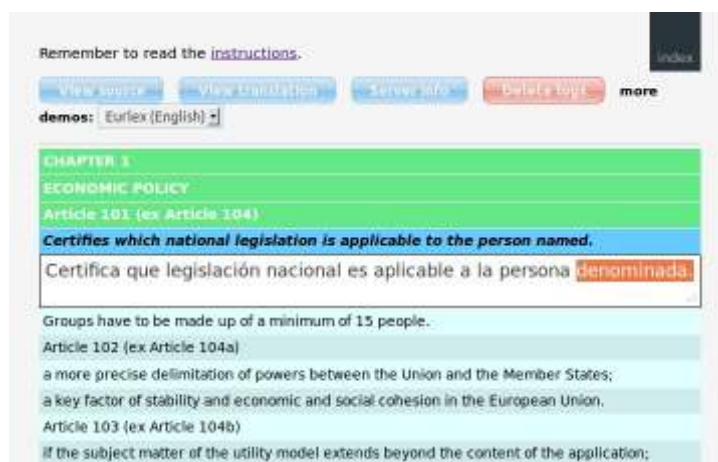
- Assessment: to measure the overall time required to translate a test corpus using the IMT system and without any system.
- Six professional translators, recruited from the two translation agencies. Five rounds during 18 months using Xerox corpus.
- Human evaluation: Overall, an IMT system can allow translators to increase their productivity while maintaining high-quality (about a save of 10-15% of human effort) depending on the task.



November 7, 2014

22

# Snapshot of the MIPRCV prototype



IMT-MIPRCV  
<http://cat.prhlt.upv.es/imt/>



November 7, 2014

23

## Human evaluation in MIPRCV

[Alabau et al. EAMT 2012]

- First field trial
  - 10 users aged 26-43
  - 60 sentences from English-Spanish EU corpus
  - Results:
    - \* the time spent per sentence in the IMT system was higher than in PE.
    - \* the effectiveness was slightly higher for IMT (higher final BLEU)
    - \* IMT failed to succeed in questions regarding the system being easy to use, consistent, and reliable.
- Second field trial
  - 15 users aged 23-24
  - 60 sentences from English-Spanish EU corpus
  - Results:
    - \* IMT was more efficient than with PE (less iterations)
    - \* IMT was perceived as more helpful



November 7, 2014

24



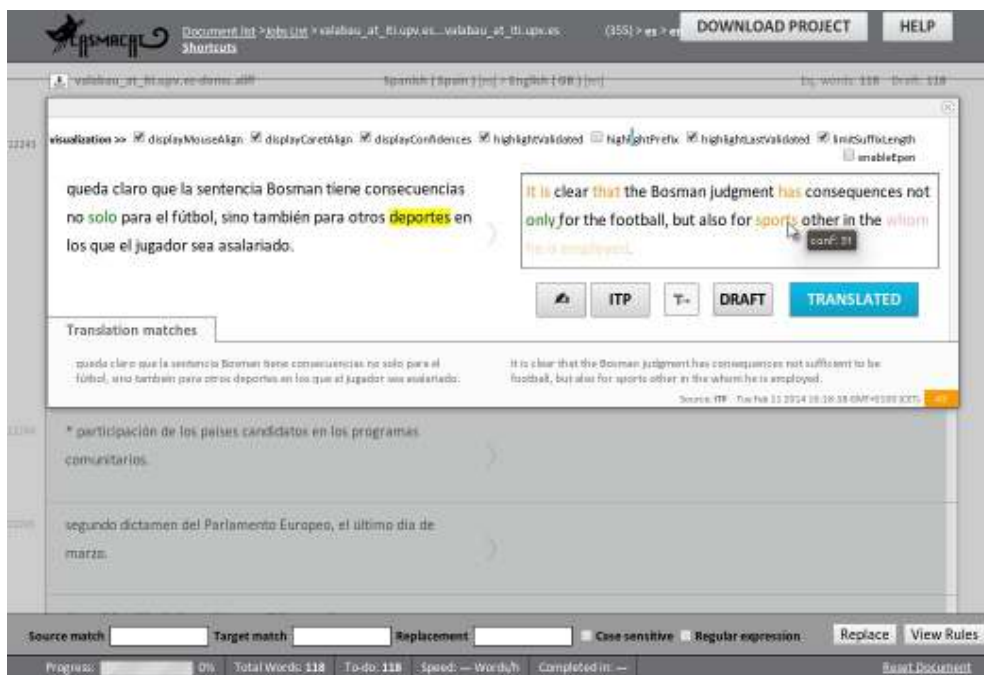
- Co-funded by the European Union under the 7th FP Project 287576
- Partners:
  - [University of Edinburgh](#) (UEDIN), United Kingdom. (Main coordinator).
  - [Universitat Politècnica de València](#) (UPVLC), Spain.
  - [Copenhagen Business School](#) (CBS), Denmark.
  - [Celer Soluciones](#) (CL), Spain.
- The goal is to build the next generation translator's workbench to improve productivity, quality, and work practices in the translation industry by introducing novel types of assistance to human translators.
  - **Interactive translation prediction**
  - **Interactive editing**
  - **Adaptive translation models**
- The workbench was jointly developed with the Matecat project.



November 7, 2014

25

## CASMACAT interface



<http://casmacat.eu/>

November 7, 2014

26

## Human evaluation in CASMACAT project

- First field trial (2012).
  - Post-editing vs. from scratch,
  - Six freelance translators
  - ≈ 12000 words from new stories (en - es) collected from CNN, Fox News, NY Times
  - Results: PE produces a substantial time saving compared to translation
- Second field trial. (2013) [Sanchis et al. MTJ 2014]
  - Post-editing, IMT and advanced IMT.
  - Nine freelance translators and four reviewers
  - ≈ 10000 words in 460 segments from WMT-2012 NC corpus
  - Results: IMT minimizes the number of key strokes that are required to generate the translations but with a little bit higher translation time. The productivity with advanced IMT decreased but the human translators liked the advanced features (IMT requires some human training)
- Third field trial (2014). Under discussion.
  - IMT + On-line learning. Use of the e-pen.



More info in <http://casmacat.eu/>



November 7, 2014

27

## Other tasks



English to Chinese



Japanese to English



November 7, 2014

28

## Other IMT topics

- IMT and hierarchical models [Gonzalez-Rubio et al. EMNLP 2013]
  - Up to now, the translation models used in IMT have been finite-state models and phrase-based models
  - Next step is to use hierarchical models and hyper-graphs instead of word-graphs (from Moses)
  - A decrease of KSMR 1.5 points in the EU and in the TED corpora.
- Other types of interactions
  - Active mouse (in this talk)
  - Fix the correct translated words of the suffix (under development)
  - Passive vs active interaction (in this talk)
  - Other input modalities: speech, e-pen (in this talk)
  - Future input modality: gaze-tracking
  - We need more realistic user models!



November 7, 2014

29



## Using the mouse as an additional information source

- When the user wants to correct the translation hypothesis, he needs to click previously on the position he intends to correct, therefore:
  - he is validating a prefix up to the position where he has set the cursor
  - he is indicating that whatever comes after that cursor position is incorrect and he wants it to be replaced.
- Mouse actions can be considered as a constrained search problem

$$\hat{y}_s = \operatorname{argmax}_{y_s: y_{s1} \neq \hat{y}_{l1}} \Pr(y_s | x, y_p, \hat{y}_l)$$

where  $\hat{y}_l$  is the suffix generated in the previous iteration, already discarded by the user, and  $\hat{y}_{l1}$  is the first word in  $\hat{y}_l$ .

- Lab results on XRCE2 and EU tasks: A decrease of 5-8 WSR points and an increase of the mouse actions.



November 7, 2014

30

## Active interaction in IMT

[Sanchis et al. MTJ 2014]

- **Passive interaction:** In IMT, the user is required to supervise all translations.
- **Active interaction:** The use of confidence measure for **each target word** can help the human to focus his/her attention on doubtful target words/sentences [Ueffing & Ney EAMT 2005].



- Different features have been tested in CASMACAT. Best trade-off between accuracy and computational time: IBM Model 1
- Experimental results with EU task (Spanish to English). For example, allowing a 20% of WSR, the BLEU increases 50 points (from the baseline 30 to 80). If every word is supervised the WSR is 52% (passive interaction).



November 7, 2014

31



## One theoretical issue

[Alabau et al. PRL 2012]

- What function optimizes the rule used for IMT?

$$\hat{y}_s = \operatorname{argmax}_{y_s} p(y_s | x, y_p) \quad (1)$$

and the complete hypothesis in an iteration is  $y = y_p \hat{y}_s$

- For text prediction, this rule minimizes the number of iterations [Oncina PRL 2009]:

$$\hat{y} = \operatorname{argmax}_y p(y | y_p)$$

and the prefix of the hypothesis in an iteration is  $y = y_p \hat{y}$

- For IMT, the following rule minimize the number of iterations:

$$\hat{y} = \operatorname{argmax}_y \sum_{y_s} p(y, y_s | x, y_p) \quad (2)$$

and the prefix of the hypothesis in an iteration is  $y = y_p \hat{y}$

- The rule used in (1) is a max approximation of (2)
- In practice, experimental results show that the improvements are marginal in IMT



November 7, 2014

32

## Index

- 1 Machine translation and computer-assisted translation ▷ 2
- 2 Interactive machine translation ▷ 6
- 3 *Adaptation in interactive machine translation* ▷ 33
- 4 Multimodal interactive machine translation ▷ 42
- 5 Conclusions ▷ 54
- 6 Bibliography ▷ 56



November 7, 2014

33

## Adaptive training

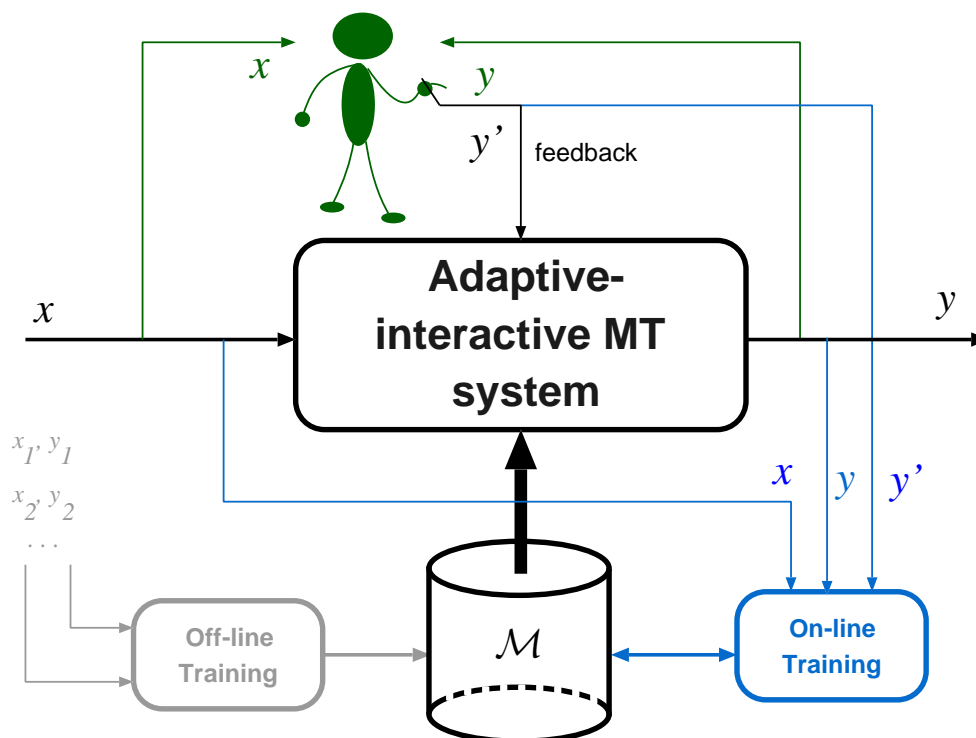
- *Human interaction* offers another unique opportunity to improve IMT system's behavior by tuning the translation models.
- In each iteration, the text obtained by means of additional user keystrokes to correct the suggestion produced by the IMT systems together with the corresponding aligned source segments can generally be converted into new, fresh training data, useful for *adapting* the system to changing environment.



November 7, 2014

34

### Diagram of an interactive MT system: Adaptation



November 7, 2014

35

## Some techniques for adaptation in IMT

- In the TT framework, a *cache technique* both for language models (unigrams, bigrams and trigrams) and translation models (model IBM1 and IBM2) is used [Nepveu et al. EMNLP 2004]
- The output of post-editing has been used for adaptation [Callison-Burch et al. EAMT 2004]
- Another technique is based on *translation memories*, in order to store the data corrected by the human [Biçini & Dymetman C1Cling 08]
- Techniques for topic-adaptation in statistical machine translation: Bayesian adaptation ([Sanchis & Casacuberta COLING 2010]), infrequent n-grams, etc.
- **On-line learning**: incremental EM, Perceptron, passive-aggressive, etc.
- **Active learning** = On-line learning + active interaction (at sentence level)



November 7, 2014

36

## On-line learning

$$\hat{y}_s = \operatorname{argmax}_{y_s} p(y_p y_s | \mathbf{x}; \theta)$$

using a log-linear model:

$$\log p(\mathbf{y} | \mathbf{x}; \theta) \approx \sum_{k=1}^K \lambda_k h_k(\mathbf{x}, \mathbf{y}; \theta_k) + Z(\mathbf{x}; \theta)$$

where  $\theta$  is composed by

- The weights of the log-linear models  $\lambda_k \Rightarrow$  **On line learning the weights**
- The parameters  $\theta_k$  of each feature  $h_k \Rightarrow$  **On line learning the parameters**
  - $h_1(\mathbf{x}, \mathbf{y}; \theta_1)$ : (log) target language model
  - $h_2(\mathbf{x}, \mathbf{y}; \theta_2)$ : (log) phrase-based models
  - $h_3(\mathbf{x}, \mathbf{y}; \theta_3)$ : (log) phrase-based inverse model
  - $h_4(\mathbf{x}, \mathbf{y}; \theta_4)$ : distortion model
  - ...



November 7, 2014

37

## On-line learning the weights

- Different techniques can be used: simplified Perceptron [España & Márquez, EAMT 2008], passive-aggressive [Cesa-Bianchi et al. SMART 2008][Martinez et al. IbPRIA 2011], discriminative ridge regression [Martinez et al. PR 2012], ...
- Discriminative ridge regression:
  - Good hypotheses within a n-best list score higher, bad hypotheses lower
  - Correlation between difference in translation quality and difference in score.
- In a PE scenario (quality=TER): the best results (2 TER points) with discriminative ridge regression (Tra/dev: English-French Europarl v5, test: News-Commentary 2008-09) [Martinez et al. PR 2012]
- In an IMT scenario (quality=KSR): No improvements (Tra/dev: English-Spanish 2013 WSMT, test: News 2008-09) [Chinea et al. LREC 2014]



November 7, 2014

38

## On-line learning the features

[Ortiz et al. NAACL 2010]

- Given a new sentence pair  $(x, y)$  validated by the user, the model of each feature is updated
- To do this, a set of sufficient statistics that can be incrementally updated is maintained for each feature function  $h_k(x, a, y)$
- Standard estimation procedures use word alignment matrices to extract phrase counts (HMM models)
- Estimation of HMM models is based on the incremental EM algorithm
- The Thot toolkit [Ortiz-Martinez and Casacuberta, 2014] implements a fully fledged phrase-based SMT system <https://github.com/daormar/thot/>



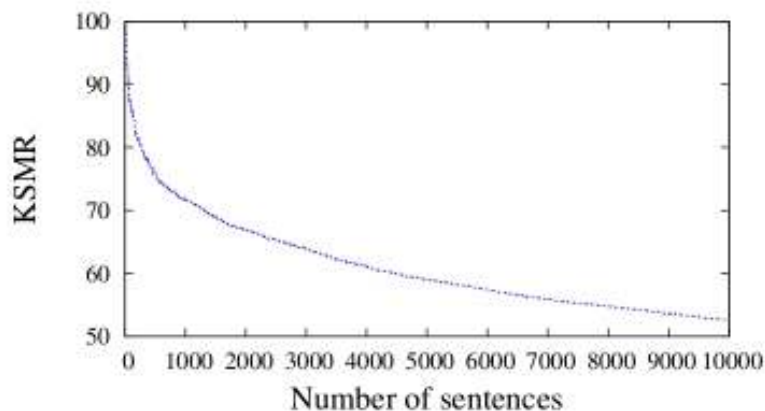
November 7, 2014

39

## Experimental results: features

[Ortiz WFDTR 2012]

- Learning from scratch: 10 000 sentences randomly extracted from the English-Spanish Xerox corpus were interactively translated



- Learning from previous models: English-French Xerox corpus

Mode	BLEU	KSMR	time
batch	33.7	33.9	NA
incremental	44.2	27.9	0.09s



## Active learning in IMT

[Gonzalez et al. PRL 2014]

- **Active learning:** A low confidence measure of a **target translation (at sentence level)** can be used by the user to review it (active interaction) and the models can be adapted.
- If a suggested target sentence presents a low confidence measure of being a correct translation but the human accepts the suggestion, the models can also be adapted.
- Supervising only a subset of the translations using different alternatives.
- Using the on-line learning for IMT to incorporate such subset
- Experimental results, with basic SMT with Europarl, AL with News Commentary, Spanish to English: All the alternatives present similar results, and allowing an effort of 20% of KSMR, the BLEU increase from 40 (only active interaction) to 65 (active interaction plus online learning).



# Index

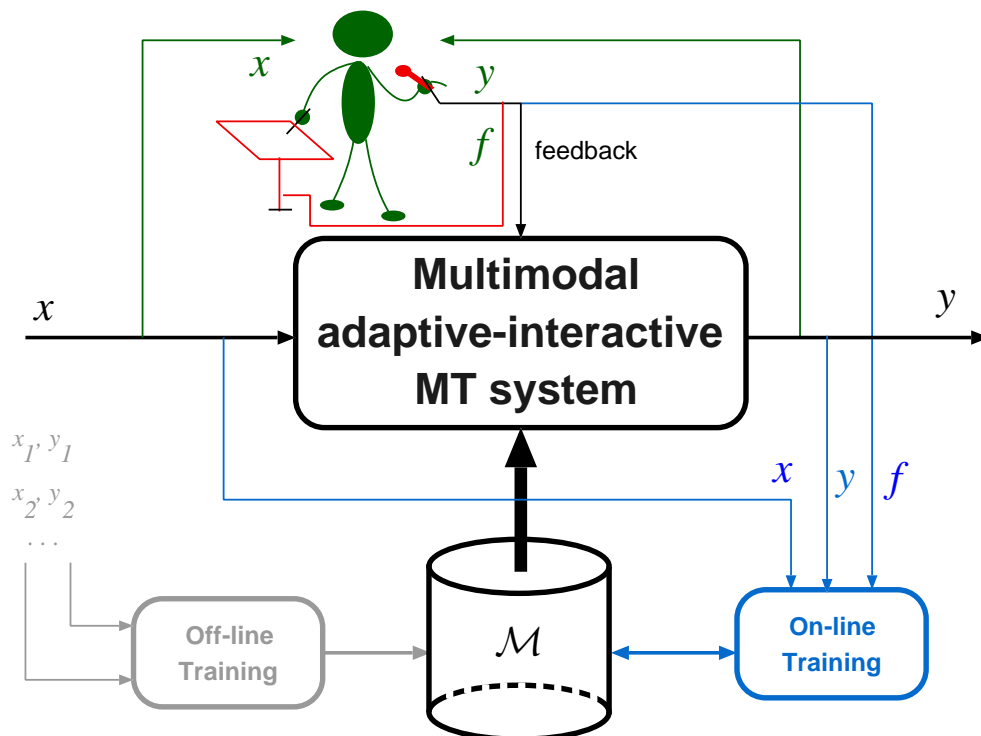
- 1 Machine translation and computer-assisted translation ▷ 2
- 2 Interactive machine translation ▷ 6
- 3 Adaptation in interactive machine translation ▷ 33
- 4 *Multimodal interactive machine translation* ▷ 42
- 5 Conclusions ▷ 54
- 6 Bibliography ▷ 56



November 7, 2014

42

## Diagram of an interactive MT system: Adaptation & multi-modality



November 7, 2014

43

## Other information sources for IMT

- Using speech recognition in IMT (In this section)
- Using e-pen for gesture and handwriting text recognition (In this section)
- Exploiting visual supports (viewgraphs, etc.) (In the future)
- Gaze tracking (In the future)

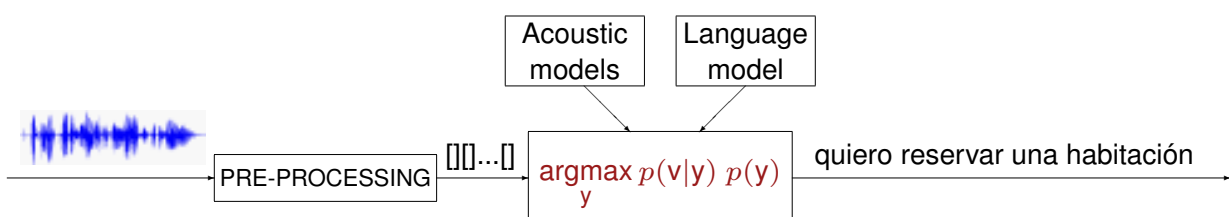


## Speech recognition

Given an input utterance  $v$ , the system should search for a most likely decoding of  $v$ :

$$\begin{aligned} \hat{y} &= \underset{y}{\operatorname{argmax}} p(y | v) \\ &= \underset{y}{\operatorname{argmax}} p(v | y) p(y) \end{aligned}$$

- $p(v | y) \approx$  **acoustic models**: hidden Markov models
- $p(y) \approx$  **language model**: (smoothed) n-grams
- Decoding (maximization): **Viterbi (beam) search**



## Using speech recognition in IMT

- Early idea: a human translator dictates aloud the translation in the *target language*. As the source text is known by the system, this knowledge can be used to reduce recognition errors. (*TransTalk* project) [Dymetman et al. ICSLP 1994] [Brown et al. CSL 1994] [Khadivi et al. COLING-ACL 2006] [Paulik et al. MELECON 2006]).
- Given  $v$  and  $x$ , the system should search for a most likely decoding of  $v$ :

$$\hat{y} = \underset{y}{\operatorname{argmax}} p(y | x, v)$$

By the assumption that  $p(v | x, y)$  does not depend on  $x$ ,

$$\hat{y} = \underset{y}{\operatorname{argmax}} p(v | y) p(x | y) p(y)$$

- $p(v | y) \approx$  (target language) acoustic models
- $p(y) \approx$  target language model
- $p(x | y) \approx$  translation model

- *Similar to plain speech decoding* but additional knowledge is used ( $p(x | y)$ ). In practice, plain speech recognition is used (N-best or word-graphs) and re-scoring.



November 7, 2014

46

## Using speech recognition in IMT

- *Alternative idea within the IMT framework*: the human translator determines acceptable prefixes of the suggestions made by the system by reading (with possible modifications) parts of these suggestions (*TransType 2* project [Vidal et al. IEEE TASLP 2006]).
- Let  $x$  be the source text and  $y_p$  a “correct” prefix of the target sentence. The user is now allowed to *utter some words*,  $v$ , generally aimed at amending parts of  $\hat{y}_s$  and the system has then to obtain a most probable decoding of  $v$  [Vidal et al. IEEE TASLP 2006]:

$$\hat{d} = \underset{d}{\operatorname{argmax}} p(d | x, y_p, \hat{y}_s, v)$$

- Finally, the user can enter additional amendment keystrokes  $k$ , to produce a new consolidated prefix,  $y_p$ , based on the previous  $y_p$ ,  $\hat{d}$ ,  $k$  and parts of  $\hat{y}_s$ .



November 7, 2014

47



## Models for speech recognition in IMT

[Vidal et al. IEEE TASLP 2006]

Previous equation (and making some assumptions):

$$\hat{d} = \operatorname{argmax}_d p(d | x, y_p, \hat{y}_s, v) = \operatorname{argmax}_d p(d | x, y_p, \hat{y}_s) p(v | d)$$

- $p(v | d) \approx$  (target language) acoustic models
- $p(d | x, y_p, \hat{y}_s) \approx$  target language model constrained by the source sentence, the prefix and the suffix

Possible scenarios less and more restrictive,

- DEC: Conventional speech recognition of target language utterances
- DEC-PREF: Target speech recognition constrained by the known prefix
- CAT-PREF: Ignore the dependency on the system suggestion  $\hat{y}_s$
- CAT-SEL: Restrict  $d$  to be just a prefix of  $\hat{y}_s$



November 7, 2014

48

## IMT speech recognition results

[Vidal et al. IEEE TASLP 2006]

- SPEECH DATA: Utterances of fragments of target language sentences from the test XEROX CORPUS (485 fragments, 10 speakers, 5,796 utterances)
- MODELS: derived from both source and target sentences of the training XEROX corpus

	DEC	DEC-PREF	CAT-PREF	CAT-SEL
Word Error Rate (%)	18.6	16.1	10.6	1.6
Sentence Error rate (%)	50.2	44.4	30.0	3.6

***Using knowledge about the source sentence is more important than using only user-validated prefixes***



November 7, 2014

49

## CASMACAT interface with e-pen

[Alabau et al. PRJ 2014]

- E-pen based interaction is a promising alternative to keyboard and mouse
- Recognition errors could reduce productivity w.r.t. keyboard, but rule of thumb to user acceptability is 5% ~ 20% acceptable if there is a substantial payoff in terms of achieving task goals
- Aiming at a more comfortable system:
  - Recognition of sub-word units and sequences of multiple words
  - Recognition of e-pen gestures
- Similar to the case of the speech modality

$$\hat{d} = \operatorname{argmax}_d p(d | y_p, x, v) = \operatorname{argmax}_d p(v | d) p(d | y_p, x)$$



November 7, 2014

50

## On-line HTR results

[Alabau et al. PRJ 2014]

- Word and phrase editing

	System	en (CER%)	
		word	phrase
HTR: baseline	HTR	9.9	18.6
4PREF : 4-gram prefix	4PREF	9.5	18.0
M2 : word model 2	M2	7.7	17.4
	M2+4PREF	7.5	15.5

- Using n-best lists for word HTR: 2.8% (en) error rate



November 7, 2014

51

## CASMACAT interface with e-pen



November 7, 2014

52

## Proof-reading gestures

[Alabau et al. PRJ 2014]

<b>deletion</b>	if <sub>1</sub> any <sub>2</sub> feature <sub>3</sub> <del>not</del> <sub>4</sub> is <sub>5</sub> available <sub>6</sub> on <sub>7</sub> your <sub>8</sub> network <sub>9</sub>
<b>insertion</b>	if <sub>1</sub> any <sub>2</sub> feature <sub>3</sub> <i>vis</i> not <sub>4</sub> is <sub>5</sub> available <sub>6</sub> on <sub>7</sub> your <sub>8</sub> network <sub>9</sub>
<b>substitution</b>	if <sub>1</sub> any <sub>2</sub> feature <sub>3</sub> <i>is</i> not <sub>4</sub> is <sub>5</sub> available <sub>6</sub> on <sub>7</sub> your <sub>8</sub> network <sub>9</sub>
<b>shift</b>	if <sub>1</sub> any <sub>2</sub> feature <sub>3</sub> not <sub>4</sub> <i>is</i> <sub>5</sub> available <sub>6</sub> on <sub>7</sub> your <sub>8</sub> network <sub>9</sub>
<hr/>	
<b>transposition</b>	if <sub>1</sub> any <sub>2</sub> feature <sub>3</sub> not <sub>4</sub> <i>is</i> <sub>5</sub> available <sub>6</sub> on <sub>7</sub> your <sub>8</sub> network <sub>9</sub>

Error rate with state-of-the-art gesture recognizer ≈ 10%



November 7, 2014

53

## Index

- 1 Machine translation and computer-assisted translation ▷ 2
- 2 Interactive machine translation ▷ 6
- 3 Adaptation in interactive machine translation ▷ 33
- 4 Multimodal interactive machine translation ▷ 42
- 5 *Conclusions* ▷ 54
- 6 Bibliography ▷ 56



November 7, 2014

54

## Conclusions

- The iterative process where human activity is embedded in the post-editing loop potentially increases the productivity of high quality translations.
- An ergonomic user interface is essential
- The users should be trained in the use of the tool
- Multi-modality can be very important: using vocal correction of transcribed speech, exploiting visual supports (viewgraphs, etc.), gaze tracking, e-pen, touch-screen, etc.
- Open questions:
  - If the MT engine in IMT is very poor or is very good, post-editing is the best choice, but, when an IMT system is useful?
  - What other modalities are useful?
  - Left-to-right is enough? We need more realistic user models



November 7, 2014

55

## Index

- 1 Machine translation and computer-assisted translation ▷ 2
- 2 Interactive machine translation ▷ 6
- 3 Adaptation in interactive machine translation ▷ 33
- 4 Multimodal interactive machine translation ▷ 42
- 5 Conclusions ▷ 54
- 6 *Bibliography* ▷ 56



November 7, 2014

56

## Bibliography

- [Alabau et al. PRL 2012]** V. Alabau, A. Sanchis, F. Casacuberta. *On the Optimal Decision Rule for Sequential Interactive Structured Prediction*. Pattern Recognition Letters, 33(16):2226-2231. 2012.
- [Alabau et al. PBML 2013]** : V. Alabau, R. Bonk, Ch. Buck, M. Carl, F. Casacuberta, M. García-Martínez, J. González-Rubio, P. Koehn, L. A. Leiva, B. Mesa-Lao, D. Ortiz-Martínez, H. Saint-Amand, G. Sanchis-Trilles, Ch. Tsoukala. *An Open Source Workbench for Advanced Computer Aided Translation*. The Prague Bulletin of Mathematical Linguistics, 100:101-112. 2013.
- [Alabau et al. EAMT 2012]** V. Alabau, L. A. Leiva, D. Ortiz-Martínez, F. Casacuberta. *User Evaluation of Interactive Machine Translation Systems*. 16th Conference of the EAMT, 2012.
- [Alabau et al. PRJ 2014]** V. Alabau, A. Sanchis, F. Casacuberta. *Improving on-line handwritten recognition in interactive machine translation*. Pattern Recognition, 47(3):1217-1228. 2014.
- [Barrachina et al. CL 2008]** S. Barrachina, O. Bender, F. Casacuberta, J. Civera, E. Cubel, S. Khadivi, A. Lagarda H. Ney, J. Tomás, and E. Vidal. *Statistical approaches to computer-assisted translation*. Computational Linguistics, 35(1):3-28, 2009.
- [Biçici & Dymetman CICLE 2008]** E. Biçici, M. Dymetman. *Dynamic Translation Memory: Using Statistical Machine Translation to Improve Translation Memory Fuzzy Matches*. 9th CiLing, LNCS, 4919:454–465. 2008
- [Blatz et al. COLING04]** J. Blatz, E. Fitzgerald, G. Foster, S. Gandrabur, C. Goutte, A. Kulesza, A. Sanchis, and N. Ueffing. *Confidence estimation for machine translation*. 20th COLING, 2004.



November 7, 2014

57

## Bibliography

- [**Brown et al. CSL 1994**] P. F. Brown, S. F. Chen, S. A. Della Pietra, V. J. Della Pietra, A. S. Kehler, R. L. Mercer. *Automatic speech recognition in machine-aided translation*. Computer Speech & Language, 8(3):177-187. 1994.
- [**Bulyko et al. ICASSP 2007**] I. Bulyko, S. Matsoukas, R. Schwartz, L. Nguyen, L. Makhoul: *Language model adaptation in machine translation from speech*. ICASSP 2007.
- [**Callison-Burch et al. EAMT 2004**] C. Callison-Burch, C. Bannard and J. Schroeder. *Improving statistical translation through editing*. 9th EAMT Workshop Broadening horizons of machine translation and its applications. 2004,
- [**Casacuberta et al. CACM 2008**] F. Casacuberta, J. Civera, E. Cubel, A.L. Lagarda, G. Lapalme, E. Macklovitch and E. Vidal. *Human interaction for high quality machine translation*. Communications of the ACM, 52(10):135-138, 2009.
- [**Cesa-Bianchi et al. SMART 2008**] N. Cesa-Bianchi, G. Reverberi, S. Szedmak *Online Learning Algorithms for Computer-Assisted Translation D4.2 Smart project*. 2008.
- [**Chinea et al. LREC 2014**] M. Chinea-Rios, Germán Sanchis-Trilles, F. Casacuberta. *Online Optimisation of Log-linear Weights in Interactive Machine Translation*. LREC, 2014.
- [**DeNeefe et al. ACL 2005**] S. DeNeefe, K. Knight, H. H. Chan. *Interactively Exploring a Machine Translation Model*. ACL Interactive Poster and Demonstration Sessions. 97–100, 2005.
- [**Dugast et al. WSMT 2007**] L. Dugast, J. Senellart, P. Koehn. *Statistical post-editing on SYSTRANS's rule-based translation system*. 2nd WSMT, 220-223, 2007



## Bibliography

- [**Dymetman et al. ICSLP 1994**] M. Dymetman, J. Brousseau, G. Foster, P. Isabelle, Y. Normandin, P. Plamondon, Pierre. *Towards an automatic dictation system for translators: the transtalk project*, ICSLP. 691-694. 1994.
- [**España & Márquez, EAMT 2008**] C. España-Bonet, Ll. Màrquez. *Robust Estimation of Feature Weights in Statistical Machine Translation*. EAMT 2010.
- [**Foster et al. MT 1997**] G. Foster, G., P. Isabelle, and P. Plamondon. *Target-text mediated interactive machine translation*. Machine Translation, 12(1-2):175-194. 1997.
- [**Foster et al. EMNLP 2002**] G. Foster, P. Langlais and , G. Lapalme. *User-friendly text prediction for translators*. EMNLP. 2002
- [**Gonzalez et al. EAMT 2010**] J. González-Rubio, D. Ortiz-Martínez, J.M. Benedí, F. Casacuberta. *On the Use of Confidence Measures within an Interactive-predictive Machine Translation System*. EAMT. 2010.
- [**Gonzalez et al. EMNLP 2013**] J. González-Rubio, D. Ortiz-Martínez, J.M. Benedí, F. Casacuberta. *Interactive Machine Translation using Hierarchical Translation Models*. EMNLP. 2014
- [**Khadivi et al EUROSPEECH 2005**] S. Khadivi, A. Zolnay and H. Ney *Automatic text dictation in computer-assisted translation*. Interspeech, 2005.
- [**Koehn ACI 2009**] P. Koehn. *A Web-Based Interactive Computer Aided Translation Tool*, ACL-IJCNLP, 2009.7
- [**Koehn & Schroeder WSMT 2007**] P. Koehn, J. Schroeder. *Experiments in Domain Adaptation for Statistical Machine Translation*, Second WSMT, 2007



## Bibliography

- [Isabelle et al. **MT-Summit 2007**] P. Isabelle, G. Goutter, M. Simard. *Domain adaptation of MT systems through automatic port-editing*. MT Summit XI, Sept. 2007.
- [Langlais et al. **NAACL 2000**] P. Langlais, P., G. Foster, and G. Lapalme. *TransType: a computer-aided translation typing system*. NAACL Workshop on Embedded Machine Translation Systems, 2000.
- [Langlais et al. **MT 2000**] P. Langlais, P., G. Lapalme, and M. Loranger. *Transtype: Development-evaluation cycles to boost translator's productivity*. Machine Translation, 15(4):77-98. 2002.
- [Macklovitch **LREC 2006**] E. Macklovitch, *TransType2: The last word*. 5th LREC, 2006.
- [Martinez et al. **PR 2012**] P. Martínez, G. Sanchis, F. Casacuberta. *Online adaptation strategies for statistical machine translation in post-editing scenarios*. Pattern Recognition, 45(9):3193-3203, 2012.
- [Martinez et al. **IbPRIA 2011**] P. Martínez-Gómez, G. Sanchis-Trilles, F. Casacuberta. *Passive-Aggressive for On-line Learning in Statistical Machine Translation*. IbPRIA 2012,
- [Nepveu et al. **EMNLP 2004**] L. Nepveu, G. Lapalme, P. Langlais, G. Foster: *Adaptive language and translation models for interactive machine translation*. EMNLP, 2004.
- [Omcima **PRL 2009**] J. Oncina, *Optimum Algorithm to Minimize Human Interactions in Sequential Computer Assisted Pattern Recognition* Pattern Recognition Letters, 30:558-563. 2009.
- [Ortiz et al. **NAACL 2010**] D. Ortiz-Martínez, I. García-Varea, F. Casacuberta. *Online Learning for Interactive Statistical Machine Translation*. NAACL 2010.
- [Ortiz & Casacuberta, **2014**] D. Ortiz-Martínez, F. Casacuberta. *The New Thot Toolkit for Fully-Automatic and Interactive Statistical Machine Translation* EACL 2014.



November 7, 2014

60

## Bibliography

- [Paulik et al. **MELECON 2006**] M. Paulik, S. Stüker, C. Fügen. *Speech Recognition in Human Mediated Translation Scenarios*. IEEE MELECON, 2006.
- [Sanchis et al. **MLMI 2008**] G. Sanchis-Trilles, M.-T. González, F. Casacuberta, E. Vidal and J. Civera. *Introducing Additional Input Information into Interactive Machine Translation Systems*. 5th Joint Workshop on MLMI, 2008
- [Sanchis & Casacuberta **COLING 2010**] G. Sanchis-Trilles, F. Casacuberta. *Log-linear weight optimisation via Bayesian Adaptation in Statistical Machine Translation*. COLING, 2010.
- [Sanchis et al. **MTJ 2014**] G. Sanchis-Trilles, V. Alabau, Ch. Buck, M. Carl, F. Casacuberta, M. García-Martínez, U. Germann, J. González-Rubio, R. L. Hill, P. Koehn, L. A. Leiva, B. Mesa-Lao, D. Ortiz-Martínez, H. Saint-Amand, Ch. Tsoukala. *Interactive Translation Prediction vs. Conventional Post-editing in Practice: A Study with the CasMaCat Workbench*. Machine Translation, under review, 2014
- [Simard et al. **WSMT 2007**] M. Simard, N. Ueffing, P. Isabelle, R. Kuhn. *Rule-based translation with statistical phrase-based post-editing*. 2nd WSMT, 2007.
- [Simard et al. **NAACL 2007**] M. Simard, G. Goutter, P. Isabelle. *Statistical phrase-based post-editing*. NAACL-HLT, 2007
- [Toselli et al. **2011**] A. Toselli, E. Vidal, F. Casacuberta. *Multimodal Interactive Pattern Recognition and Applications*. Springer. 2011.



November 7, 2014

61



## Bibliography

- [Ueffing & Ney EAMT 2005]** N. Ueffing and H. Ney. *Application of Word-Level Confidence Measures in Interactive Statistical Machine Translation*. EAMT 2005.
- [Ueffing & Ney CL 2007]** N. Ueffing and H. Ney. *Word-Level Confidence Estimation for Machine Translation*. *Computational Linguistics*, 33(1):9–40. 2007
- [Vidal et al. IEEE TASLP 2006]** E. Vidal, F. Casacuberta, L. Rodríguez, J. Civera and C. Martínez. *Computer-assisted translation using speech recognition*. *IEEE Transaction on Audio, Speech and Language Processing*, 14(3), 2006.
- [Vidal et al. MLMI 2007]** E. Vidal, L. Rodríguez, F. Casacuberta, and I. García-Varea. *Interactive pattern recognition*. *Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms*, LNCS 4892:60-71. 2007.



November 7, 2014

62

## Projects

- **TransType (TT)**. Natural Sciences and Engineering Research Council of Canada (NSERC). 1997-2000. RALI LAB.
- **TransType-2 (TT2)**. Information Society Technologies Programme. IST-2001-32091, 2002-2004. SEMA Group, Spain (coordinator); Universitat Politècnica de València, Spain; Aachen University, Germany; RALI LAB, Canada; Celer Soluciones, Spain; Xerox RCE, France.
- **Multimodal Interaction in Pattern Recognition and Computer Vision (MIPRCV)**. Consolider Ingenio 2010. 2007-2012. Universitat Politècnica de València, Spain (coordinator); Centre de Visió per Computador; Universitat jaume I; Universidad de Granada; Universitat d'Alacant; Institut de Robòtica i Informàtica Industrial; Universitat de València; Universidad Politécnica de Madrid.
- **Casmacat: Cognitive Anlysis and Statistical Methods for Advanced Computer Aided Translation**. Information and Communication Technologies, 7th Framework Programme. 2011-2014. The University of Edinburgh, UK (coordinador); Universitat Politècnica de València, Spain; Copenhagen Business School, Denmark; Celer Soluciones SL, Spain.  
<http://www.casmacat.eu/>
- Other related projects:
  - **Matecat** <http://www.matecat.com/>
  - **Faust** <http://divf.eng.cam.ac.uk/faust>
  - **transLectures** <https://translectures.eu//>
  - **tranScriptorium** <http://transcriptorium.eu/>



November 7, 2014

63



## Demos

- *TransType*
- *IMT-MIPRCV*
- *Caitra*
- *Casmacat*

<http://rali.iro.umontreal.ca/rali/?q=en/node/1282>  
<http://cat.prhlt.upv.es/imt/>  
<http://www.caitra.org/>  
<http://casmacat.eu/>



November 7, 2014

64

# Thank you!



November 7, 2014

65