Post-Editing Machine Translation

Bob, Daniel, Harald and Xavier
Post-Editing MT vs. TM

- „A good MT match is better than an average fuzzy TM match."

- “There is not a huge technical difference in editing a fuzzy match and editing an MT match. In fact, having experience with TM makes it easier to learn how to post-edit MT.”
Profiles for Post-Editors

- "MT output could also be edited by monolingual speakers of the target language with expertise in the domain."

- "I prefer hiring bilingual linguists for post-editing. They are more flexible."
Training of Post-Editors

“For the success of post-editing, we (i.e. MT software vendors, LSPs, etc.) need to get trained translators on board”

“The topic of training translators for post-editing is overestimated. There is no magic. They just need to learn to be pragmatic.”
Attitude of LSPs

„In the nineties TM was quickly adopted by some smart translators and they make loads of money. Now MT is often requested by customers and LSPs/translators are dragging their feet.“
BLEU, NIST

„Especially SMT vendors are making a huge fuss of these valuation methods. The correlation between the post-editing effort and these scores still needs to be proven.“
PostEditing Guidelines

A tool for sellers, buyers and editors
THE DEFINITION

• “Post editing is the examination and correction of the text resulting from an automatic or semi automatic machine system (machine translation, translation memory) to ensure it complies with the natural laws of grammar, punctuation, spelling and meaning etc.”

*Draft of European Standard for Translation Services (Brussels 2004)*
INITIAL QUESTIONS

• Does your organization understand the different levels of postediting, what is involved for each level, and how that affects cost?

• As a posteditor, have you received conflicting customer acceptance criteria?

• As a client, have you experienced resistance among language specialists to accept post-editing work?

• As a Sales person have you been approached by some customer for MT quick and dirty project and none of you knew how to do it? This is, what to deliver? Or worse… have you tried to convince a hesitating customer to go MT for that reason?
FINAL CRUCIAL QUESTION

• Would standard guidelines provided to posteditors, customers and sales people on how to perform a light and full postedit increase acceptance of the postedited product, increase post-editor productivity and increase customer satisfaction in benefit our overall business?
CRYSTAL CLEAR ANSWER

• IF the answer is **YES** then we clearly see why we need these postediting guidelines and we need them backed by as much of the industry as possible.
TRANSLATORS AND POSTEDITING

• Resistance from translators to handle postediting work
• Ambiguous/unclear demands from customers
• Lack of visibility of quality output from MT system
LSP AND POSTEDITING

• For sales people to rightly balance the traditional human based model versus a faster, and maybe more cost effective MT based process,

  …with some drawbacks.
CUSTOMERS AND POSTEDITING

• Customers willing to get the most out of their budget. And to rightly assess the unrealistic search of perfection and high quality,

… at the lowest possible cost.
KISS

I LOVE KISSING

KEEP IT SIMPLE STUPID!
OUTPUTS WITH KISS IN MIND

- LEVEL 0 RAW OUTPUT
- LEVEL 1 basic, rapid, light, minimalist or… (AKA GOOD ENOUGH)
- LEVEL 2 value, full, publishable, premium (AKA NEARLY AS GOOD AS HumanQ)
- LEVEL 3 (HUMAN QUALITY, HQ)
THE GUIDELINES
THE GUIDELINES

Objectives and scope

• These guidelines are aimed at helping, customers and service providers set clear expectations and can be used as a basis on which to instruct post-editors.

• Certain customization might exist for each company, but it is advisable to stay within definition and keep it simple.

• These guidelines assume bi-lingual postediting

• The guidelines are not for the time being, system or language-specific.
PREVIOUS CONSIDERATIONS

- **Recommendations**

  - **TUNE** your system appropriately, i.e. ensure high level dictionary and linguistic coding for RBMT systems, or **TRAIN** with **clean, high-quality, domain-specific data** for data-driven or hybrid systems.
  - Ensure the **source text** is written well (i.e. correct spelling, punctuation, unambiguous) and, if possible, tuned for translation by MT (i.e. by using specific authoring rules that suit the MT system in question).
  - Integrate **terminology management** across source text authoring, MT and TM systems.
  - **Train** post-editors in advance.
  - Examine the **raw MT output quality** before negotiating throughput and price and set reasonable expectations.
  - **Pilot** editing to either level 1 (Good enough) or 2 (Nearly as good as)
  - Agree a definition for the **final quality** of the information to be post-edited, based on user type and levels of acceptance.
  - Pay post-editors to give **structured feedback** on common MT errors (and, if necessary, guide them in how to do this) so the system can be improved over time.
  - Analyze output from several MT for that language combination and subject matter
  - Evaluate output with language specialist
STANDARD THE GUIDELINES

Assuming the recommendations before they have been implemented, the effort involved in postediting will be determined by two main criteria:

• The quality of the MT raw output.
• The expected end quality of the content.
THE GUIDELINES - STANDARD

- THE GUIDELINES - GOOD ENOUGH – Edit as little as possible

- Comprehensible
- Accurate
- but as not being stylistically compelling. The text may sound like it was generated by a computer, syntax might be somewhat unusual, grammar may not be perfect but the message is accurate.
- Aim for semantically correct translation.
- Ensure that no information has been accidentally added or omitted.
- Edit any offensive, inappropriate or culturally unacceptable content.
- Use as much of the raw MT output as possible.
- Basic rules regarding spelling apply.
- No need to implement corrections that are of a stylistic nature only.
- No need to restructure sentences solely to improve the natural flow of the text.
THE GUIDELINES - STANDARD

• THE GUIDELINES - NEARLY AS GOOD AS – Edit as much as necessary

• Comprehensible
• Accurate
• **stylistically fine**, though the style may not be as good as that achieved by a native-speaker human translator. Syntax is **normal**, grammar and punctuation are **correct**.
• Aim for grammatically, syntactically and semantically **correct** translation.
• Ensure that key terminology is correctly translated and that untranslated terms belong to the client’s list of “Do Not Translate” terms”.
• Ensure that no information has been accidentally added or omitted.
• Edit any offensive, inappropriate or culturally unacceptable content.
• Use as much of the raw MT output as possible. But re-write if needed.
• Basic rules regarding spelling, punctuation and hyphenation apply.
• Ensure that **formatting** is correct.
Recommendations:

• Contact a few MT engine companies and compare outputs for your language and subject matter.
• Pilot the process with a few thousand words. Ideally with different engines.
• Calculate the cost/time it takes to reach either to good enough or nearly as good quality level
• GO FOR IT!
THANKS!!!

IDISC INFORMATION TECHNOLOGIES
XAVIER MAZA
xmaza@idisc.es

LINKEDIN - Post Editing Service Level Group
MTPE – What is holding us back?

Bob Donaldson
Carson Strategy Group
LSP adoption of MT technology is limited by both the current MT deployment paradigm and the MT vendor business model
Where we are (or seem to be):

- Many “studies” and “pilot projects” indicate cost savings
- Many large customers are pushing for more MT
- Many (most?) LSP’s do not yet see the benefit
- There is still a lot of talk about MT “quality”
How should an LSP think about quality?

- BLEU Score?
- Utility Score?
- Translator quality evaluation?
- ??
- Productivity!!
  - Do more for less money
  - Do more in less time
Doing more in less time:

• This is (or should be) a no-brainer
  – Easy to measure
  – Depends on workflow integration
  – Requires translator/editor training
  – Editing standards vary with customer and usage

• But there are hidden costs in the above
Doing more for less money:

- **Project Margin = Price – Cost**
  - Price is set by the market
  - Cost = Translation Cost + PM Overhead

- **Profit = Sum (Project Margin) – Fixed Costs**
  - Fixed Costs include facility, senior management, IT infrastructure, etc.
  - Note: MT deployment can increase fixed costs

- **Volume is the key!**
Potential MT Deployment Options

• License or Public Domain
  – Increase fixed costs
  – Increase Project Margin
  – Productivity gain + volume must exceed fixed cost increase

• Software as a Service
  – MT as a component of project cost
  – Productivity gain must exceed MT cost
Hidden MT Costs

• Training of Post-editors
  – Not all translators are candidates
• Customization
  – Expertise is a concern
• Configuration
  – Integration of MT into workflow
  – MT infrastructure
• Customers X Languages X Cost
Hidden MT Complexity for LSP

Becomes …

Becomes … (with multiple projects)
Summary of Complexity

- Multiple Language Pairs
- Multiple Vendors
  - No vendor has best solution for all languages
- Multiple Customers
- Multiple Customizations
  - 1 per language pair per customer
- Multiple MT Instances
- Multiple Cost Components
Is there a way out of the conundrum?

?? Automate (some of) the Post-Editing
   – Limit MT Instances to 1 per language pair
   – Move customization effort to post-editing
   – Bundling options
     • Built in by MT vendors (hybrid solution)
     • Separate “APE” vendor (MT agnostic)

?? Align Business Models
   – Is MT a service or a solution?
   – Can an LSP “afford” the MT productivity gain?

?? Volume as the Solution
   – Another wave of industry consolidation?
   – Expanding addressable market using MT?