

# How to teach machine translation post-editing? Experiences from a post-editing course

Maarit Koponen

University of Helsinki

*maarit.koponen@helsinki.fi*

4th Workshop on Post-Editing Technology and Practice  
MT Summit, Miami, November 3, 2015



# Why should PEMT skills be taught?

- The increasing use of MT and PE workflows leads to a growing need for translators with relevant MT and PE skills.
- Familiarity with translation technology plays an important role in the future employability of translation students.
- PEMT differs from translation in terms of practical and cognitive processes; it likely requires different skills and training.
- Translator training should take these aspects into account.

# What PE skills are needed?

- 1 General skills
  - ▶ language skills (source and target), text linguistic skills
  - ▶ subject area knowledge
  - ▶ cultural and intercultural competence
  - ▶ documentation and research skills
- 2 Technical skills
  - ▶ positive attitude toward technology
  - ▶ software skills, ability to learn new tools quickly
  - ▶ ability to evaluate tools
  - ▶ principles of MT technology
- 3 Specific editing skills
  - ▶ typical MT features and errors
  - ▶ suprasentential errors
- 4 Strategic skills
  - ▶ ability to evaluate data sources
  - ▶ ability to follow and adapt to client specifications
  - ▶ ability to edit efficiently

(O'Brien 2002; Rico and Torrejón 2012; Austermuehl 2013; Pym 2013)



# MT and PEMT in the Finnish context

- The use of MT and PE has not been particularly widespread so far.
  - ▶ MT quality: Finnish is known to be difficult for MT systems due to rich morphology.
  - ▶ MT availability: The small market area means there has been relatively little interest in developing MT systems for Finnish.
- A survey of 238 Finnish translators (Mikhailov 2015) found that they did not consider MT-related skills important.
- Most respondents were only familiar with free online systems; 3 reported using systems other than Google Translate (Mikhailov 2015).
- MT development is being carried out by LSPs and PEMT workflows are used by some translators/LSPs (mainly English to Finnish).



- A Translation Studies module (intermediate level) focusing on PEMT.
- The goal of the course was to familiarize the students with the theoretical and practical aspects of PEMT.
- The course is offered to students of different language and translation programs (Bachelor/Master level).
- Lectures (7 weeks) + 5 hands-on assignments, final reflective essay.
- 15 students took part in the course, 13 handed in the final essay.
- All students except one were native speakers of Finnish, various L2 backgrounds.

# Course description – topics

- Theory and history of MT and PE
- Practical use of MT and PE
- Controlled language and pre-editing for MT
- Post-editing without source text
- Post-editing process research
- Post-editing quality levels and guidelines
- MT quality evaluation and PE effort
- PE competences

# Course description – assignments

- 1 Comparison of 2 MT versions (RBMT and SMT) and post-editing
- 2 Pre-editing
  - ▶ Followed by customizing the RBMT system dictionary
- 3 Post-editing without source text
- 4 Post-editing to different quality levels
  - ▶ Followed by examining own PE time data and edit distance metrics
- 5 MT quality evaluation (fluency/adequacy)

## MT systems and other tools used

- Sunda (RBMT, en-fi-en) <http://www.sunda.fi/kaantaja.html>
- Google Translate <http://translate.google.com>
- MT@EC
- Appraise <https://github.com/cfedermann/Appraise>
- Asiya Online  
[http://asiya.lsi.upc.edu/demo/asiya\\_online.php](http://asiya.lsi.upc.edu/demo/asiya_online.php)

# Students' reflective essays

- At the end of the course, the students were asked to write a reflective essay (1000–1500 words) reflecting on their own experiences and observations against the theoretical background covered.
- They were also asked to comment on the course content, practical organization and potential improvements.

## Questions for the essay

- 1 How did your understanding of the use of MT post-editing and related phenomena develop during the course?
- 2 What benefits and opportunities do you see in the use of MT and PE from the perspective of a translator, an organization requiring translation services, an individual MT user? What about problems or limitations?
- 3 How has the field developed and how do you believe it will develop in the future?



# Themes identified in the reflective essays

- 1 Expectations and attitudes toward MT
  - ▶ Expectations of MT quality were low; only one had a positive attitude toward MT before the course.
- 2 MT quality and trust
  - ▶ The students were positively surprised by the MT quality but found it hard to trust the MT.
- 3 PE process and time
  - ▶ Most found it hard to get used to working with a raw translation.
  - ▶ PE speed – and the students' views on their own speed – varied.
- 4 Adjusting to PE requirements
  - ▶ Most found it difficult to adjust to the idea of “good enough” quality.
- 5 Future visions
  - ▶ All believed MT and PE will play a large part in future translation work – MT as a tool rather than threat.
- 6 PE training
  - ▶ MT and PE related training for translator students was seen as **useful**, even essential.

# Conclusion and on-going work

- The students gained a basic understanding of MT principles and the use of PE workflows in the translation industry.
- The students' essays show a positive attitude toward MT, but also critical evaluation of the suitability of MT tools and PE workflows to specific situations.
- The students gained more awareness of their own work processes and showed ability to evaluate their own work.
- Technical issues and lack of suitable MT texts for the various language pairs posed some challenges during the course.
- The same course is taught again in fall 2015, with some modifications:
  - ▶ A TM system used for most assignments (Memsource).
  - ▶ A wider selection of language pairs for the assignments.
  - ▶ The order of the course topics and assignments has been changed.



# Thanks

Thank you for your attention! Questions and comments?

*maarit.koponen@helsinki.fi*

<p><b>Konekäynnön jälkeiditointitehtävä 1</b> (kähtötehti)</p> <p>Tähän: Care of People with the flu The person who is ill with the flu should stay in a room separate from common areas of the home and avoid contact with others as much as possible. The person should stay home for at least 7 days. Konekäynnön 2 Huuletien toimittaja Ransalla Matti, joka sairastaa flunssaa, pitäisi joutua huoneeseen, jossa kodin tavutuksia alueille ja viittä viikoksi muuten niin paljon kuin mahdollista. Huuletien toimittaja Ransalla Matti, joka on saanut flunssaa tulee joutua huoneeseen, jossa kodin tavutuksia alueille ja viittä viikoksi muuten niin paljon kuin mahdollista.</p>	<p><b>Tehtävä 3: Yksikertainen editointi</b> Tässä editointitehtävässä kokelaan yksikertaisesti edustavaa ilmiötä. Käsiteltävä asia on käännöskäsitteiden käyttö, josta on otettu otteita. Tehtävässä on annettu käännöskäsitteiden käyttöä koskevia ohjeita. Tehtävässä on annettu käännöskäsitteiden käyttöä koskevia ohjeita. Tehtävässä on annettu käännöskäsitteiden käyttöä koskevia ohjeita.</p>	<p><b>CKT245 Käännösedittointitehtävä 2</b> Tutustu yllä olevan kontrollin kielien sääntöihin (englanti, heikki moodissa). Valitse tässä jaksossa tehtävissä 1 käytetyt käännöskäsitteet välikäsitteiksi 5 erikseen, jotka olivat erillisen harjoituksen käännöskäsitteitä. Muokkaa nämä välikäsitteet erikseen erikseen käännöskäsitteiksi: tee käännöskäsitteistä artikkelit, joihin käännöskäsitteet joko jätetään välikäsitteiksi tai lisäetään käännöskäsitteiksi. Kirjallisuuden ohjeet ovat näkyvät ensin käännöskäsitteiksi (joka jättäminen tosin...)</p>
<p><b>Appraise</b> Tehtävä 4: Kevyt vs täysi jälkeiditointi</p> <p>otanto</p> <p>Tee tässä "täysi" jälkeiditointi alustavasti on tehty, jonka sisältöä on otettu myös josta käännöskäsitteitä on otettu. 1. Pöytäkirjan käännöskäsitteitä on otettu käännöskäsitteistä. 2. Tehtävässä on annettu josta käännöskäsitteitä on otettu josta käännöskäsitteistä. 3. Käännöskäsitteitä on otettu josta käännöskäsitteistä. 4. Käännöskäsitteitä on otettu josta käännöskäsitteistä. Työskentele tai käännöskäsitteitä on otettu josta käännöskäsitteistä.</p>	<p><b>Tehtävä 5: Konekäynnön editointi</b> Aseta välikäsitteet käännöskäsitteiksi käännöskäsitteiksi "käännöskäsitteiksi" perusteella, eli sen perusteella, mikä välikäsitteistä on otettu käännöskäsitteiksi. Aseta välikäsitteet käännöskäsitteiksi käännöskäsitteiksi "käännöskäsitteiksi" perusteella, eli sen perusteella, mikä välikäsitteistä on otettu käännöskäsitteiksi. Aseta välikäsitteet käännöskäsitteiksi käännöskäsitteiksi "käännöskäsitteiksi" perusteella, eli sen perusteella, mikä välikäsitteistä on otettu käännöskäsitteiksi.</p>	<p><b>Appraise</b> otanto</p> <p>Aseta välikäsitteet käännöskäsitteiksi käännöskäsitteiksi "käännöskäsitteiksi" perusteella, eli sen perusteella, mikä välikäsitteistä on otettu käännöskäsitteiksi. Aseta välikäsitteet käännöskäsitteiksi käännöskäsitteiksi "käännöskäsitteiksi" perusteella, eli sen perusteella, mikä välikäsitteistä on otettu käännöskäsitteiksi.</p>
<p><b>Employment: Commission proposes €1.4 million from Globalisation Fund for former shipbuilding workers in Finland</b></p> <p>Heikki</p> <p>Tämä on käännöskäsitteitä koskeva tehtävä. Käännöskäsitteitä on otettu josta käännöskäsitteistä. Työskentele tai käännöskäsitteitä on otettu josta käännöskäsitteistä.</p>	<p>Is he having years off? Is this wine cold? Is your (SINGULAR) (PLURAL) (SINGULAR) (PLURAL)</p> <p>— Sitä — Rank 1 (0) Rank 2 (0) Rank 3 (0) <b>Is it better to eat?</b> — Translation 1 — Rank 1 (0) Rank 2 (0) Rank 3 (0) <b>Are they the same?</b> — Rank 1 (0) Rank 2 (0) Rank 3 (0) <b>Are they the same?</b> — Rank 1 (0) Rank 2 (0) Rank 3 (0) <b>Are they the same?</b> — Rank 1 (0) Rank 2 (0) Rank 3 (0)</p>	<p>Tehtävä 5: Konekäynnön editointi</p> <p>— Sitä — Rank 1 (0) Rank 2 (0) Rank 3 (0) <b>he eivät tunne häntä.</b> — Translation 1 — Rank 1 (0) Rank 2 (0) Rank 3 (0) <b>he eivät tunne häntä.</b> — Translation 2 — Rank 1 (0) Rank 2 (0) Rank 3 (0) <b>He eivät tunne häntä.</b> — Translation 3</p>

