Work Package 1

Michael Carl and Robin Hill
WP1 Periode month 6 – 18
Tasks 1.1, 1.3, 1.4, and 1-5

- Task 1.1 Post-editing (9 PM month 6–18).
  - Preliminary qualitative investigation into post-editing and translation styles

- Task 1.3 Translator Types and Translation Styles (13PM month 1-24).
  - Design of the questionnaire
  - Release of the TPR-DB

- Task 1.4 Text Type (17PM month 6-30).
  - Set up of experimental translation design
  - Collection of experimental data

- Task 1.5 Cognitive Modeling (8PM month 6-36).
  - Edinburgh Eye-tracking Experiment 1
WP1 Task 1.3 Design of the questionnaire

Before Session Q1

QUESTIONNAIRE 1

Name:.................................................................

Sex:  M  F

Wear Glasses:  Yes  No

Years of formal translator training:  .................... Years

Years of translator experience:  .................... Years

Languages L1  ............... L2  ............... L3

How frequently do you use machine translation?

○ Every day
○ Every 2 - 3 weeks
○ Every month
○ Once or twice a year
○ Never

From your previous experience with machine translation outputs, how would you rate your level of satisfaction in relation to machine translation?

○ Highly satisfied
○ Somewhat satisfied
○ Neutral
○ Somewhat dissatisfied
○ Highly dissatisfied

Do you think that you will want to apply machine translation in your future translation tasks?

○ Yes  ○ No  ○ I'm not sure

In general, how feasible do you think it is to apply machine translation to professional translation services?

○ Very likely
○ Somewhat likely
○ Neutral
○ Somewhat unlikely
○ Very unlikely

Have you ever post-edited machine translation?

○ Yes  ○ No

After Session Q2

QUESTIONNAIRE 2

Name:.................................................................

How satisfied are you with the translation you have produced:

through post-editing  through editing?

○ Highly satisfied
○ Somewhat satisfied
○ Neutral
○ Somewhat dissatisfied
○ Highly dissatisfied

Would you have preferred to work on your translation from scratch instead of post-editing machine translation?

○ Yes  ○ No

Do you think that you will want to apply machine translation in your future translation tasks?

○ Yes, at some point  ○ No, never!  ○ I'm not sure yet

Based on the post-editing task you have performed, how much do you rate machine translation outputs on the following attributes?

<table>
<thead>
<tr>
<th></th>
<th>Well Below Average</th>
<th>Below Average</th>
<th>Average</th>
<th>Above Average</th>
<th>Well Above Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammaticality</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Style</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Overall quality</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Based on the post-editing task you have performed, which of these statements will you go for?

○ I had to post-edit ALL the outputs.
○ I had to post-edit about 75% of the outputs.
○ I had to post-edit 25 - 50% outputs.
○ I only had to post-edit VERY FEW outputs.

Based on the post-editing task you have performed, how often would you have preferred to translate from scratch rather than post-editing machine translation?

○ Always.
○ In most of the cases (75% of the outputs or more).
○ In almost half to the cases (approx. 50%).
○ Only in very few cases (less than 25%).
WP1 Task 1.3
Release of TPR-DB V1.0

Collection of Translation Experiments
- Legacy Data (200 Translation Sessions)
- Casmacat collected data (> 250 translation Sessions)
- 3 Source Languages, 7 Target languages

| Study | Text | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | Total |
|-------|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|------|
| ACS08 |      |  |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |   60  |
| BD08  |      |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    | 10   |
| BML12 |      | 9 | 11| 10|   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    | 60   |
| JLG10 |      |   |   |   | 2 | 3 | 2 | 3 | 5 | 5 |    |    |    |    |    |    |    |    |    |    |    | 20   |
| KTHJ08|      | 24| 24| 23|   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    | 71   |
| LWB09 |      |   |   |   | 12| 14| 14|   |   |   |    |    |    |    |    |    |    |    |    |    |    | 40   |
| MS12  |      | 3 | 9 | 7 |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    | 44   |
| NJ12  |      | 15| 19| 14|   |   |   |   |   | 17|    |    |    |    |    |    |    |    |    |    |    | 100  |
| SG12  |      | 6 | 5 | 5 |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    | 32   |
| TPR11 |      | 10| 9 |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    | 19   |
| Total translations | 67 | 69 | 67 | 14 | 16 | 15 | 15 | 53 | 12 | 14 | 14 | 2 | 3 | 2 | 3 | 5 | 5 | 41 | 39 | 456 |
WP1 Task 1.3
Structure of the TPR Database V1.0

- One Folder for each study:
  - Subfolder with Translog data, alignments, annotations
- Meta Data annotation
- Programming tools
WP1 Task 1.4
Translation Experiment Design

- 6 English source texts
  - News, sciences
- Various target languages
- Three tasks:
  - Translation
  - Post-editing
  - Monolingual editing
- Keep order of tasks
- Permutation of texts

<table>
<thead>
<tr>
<th>Participant</th>
<th>Date</th>
<th>Time</th>
<th>From Scratch</th>
<th>Post-editing</th>
<th>Editing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 01</td>
<td>T1</td>
<td>T2</td>
<td>P3</td>
<td>P4</td>
<td>E5</td>
</tr>
<tr>
<td>Participant 02</td>
<td>T3</td>
<td>T4</td>
<td>P5</td>
<td>P6</td>
<td>E1</td>
</tr>
<tr>
<td>Participant 03</td>
<td>T5</td>
<td>T6</td>
<td>P1</td>
<td>P2</td>
<td>E3</td>
</tr>
<tr>
<td>Participant 04</td>
<td>T2</td>
<td>T1</td>
<td>P4</td>
<td>P3</td>
<td>E6</td>
</tr>
<tr>
<td>Participant 05</td>
<td>T4</td>
<td>T3</td>
<td>P6</td>
<td>P5</td>
<td>E2</td>
</tr>
<tr>
<td>Participant 06</td>
<td>T6</td>
<td>T5</td>
<td>P2</td>
<td>P1</td>
<td>E4</td>
</tr>
<tr>
<td>Participant 07</td>
<td>T1</td>
<td>T3</td>
<td>P2</td>
<td>P4</td>
<td>E5</td>
</tr>
<tr>
<td>Participant 08</td>
<td>T3</td>
<td>T5</td>
<td>P4</td>
<td>P6</td>
<td>E1</td>
</tr>
<tr>
<td>Participant 09</td>
<td>T5</td>
<td>T1</td>
<td>P6</td>
<td>P2</td>
<td>E3</td>
</tr>
<tr>
<td>Participant 10</td>
<td>T2</td>
<td>T4</td>
<td>P1</td>
<td>P3</td>
<td>E6</td>
</tr>
<tr>
<td>Participant 11</td>
<td>T4</td>
<td>T6</td>
<td>P3</td>
<td>P5</td>
<td>E2</td>
</tr>
<tr>
<td>Participant 12</td>
<td>T6</td>
<td>T2</td>
<td>P5</td>
<td>P1</td>
<td>E4</td>
</tr>
<tr>
<td>Participant 13</td>
<td>T1</td>
<td>T3</td>
<td>P2</td>
<td>P5</td>
<td>E4</td>
</tr>
<tr>
<td>Participant 14</td>
<td>T3</td>
<td>T5</td>
<td>P4</td>
<td>P1</td>
<td>E6</td>
</tr>
<tr>
<td>Participant 15</td>
<td>T5</td>
<td>T1</td>
<td>P6</td>
<td>P2</td>
<td>E2</td>
</tr>
<tr>
<td>Participant 16</td>
<td>T2</td>
<td>T4</td>
<td>P1</td>
<td>P6</td>
<td>E3</td>
</tr>
<tr>
<td>Participant 17</td>
<td>T4</td>
<td>T6</td>
<td>P3</td>
<td>P2</td>
<td>E5</td>
</tr>
<tr>
<td>Participant 18</td>
<td>T6</td>
<td>T2</td>
<td>P5</td>
<td>P4</td>
<td>E1</td>
</tr>
<tr>
<td>Participant 19</td>
<td>T6</td>
<td>T3</td>
<td>P2</td>
<td>P5</td>
<td>E4</td>
</tr>
<tr>
<td>Participant 20</td>
<td>T2</td>
<td>T5</td>
<td>P4</td>
<td>P1</td>
<td>E6</td>
</tr>
<tr>
<td>Participant 21</td>
<td>T4</td>
<td>T1</td>
<td>P6</td>
<td>P3</td>
<td>E2</td>
</tr>
<tr>
<td>Participant 22</td>
<td>T5</td>
<td>T4</td>
<td>P1</td>
<td>P6</td>
<td>E3</td>
</tr>
<tr>
<td>Participant 23</td>
<td>T1</td>
<td>T6</td>
<td>P3</td>
<td>P2</td>
<td>E5</td>
</tr>
<tr>
<td>Participant 24</td>
<td>T3</td>
<td>T2</td>
<td>P5</td>
<td>P4</td>
<td>E1</td>
</tr>
</tbody>
</table>
WP1 Task 1.4
Collection of Experimental Translation Data

Collected translation sessions for 6 English source texts:

- **BML12**: en->es (Mesa-Lao, 2012)
  - 20 translation, 20 post-editing and 20 editing

- **MS12**: en->zh (Schmalz, 2012)
  - 15 translation, 19 post-editing and 10 editing

- **NJ12**: en->hi (Jaiswal et al. 2012)
  - 39 translation and 61 post-editing (2 MT systems)

- **SG12**: en->de (Hansen and Gutermuth, forthcoming)
  - 12 translation, 10 post-editing and 10 editing
WP1 Task 1.4
Qualitative investigation into translation styles

<table>
<thead>
<tr>
<th>Orientation</th>
<th>Drafting</th>
<th>Revision</th>
</tr>
</thead>
</table>

[Graph showing data points across Orientation, Drafting, and Revision phases.]
Translation Progression Graph

Michael Carl 1st year review meeting November 27, 2012
Large Context Planner

Reading sentence-wise ahead into the source text

Michael Carl 1st year review meeting November 27, 2012
Translation Styles

- Initial orientation:
  - systematic orientation vs. skimming
  - quick planning vs. head start
- Translation drafting:
  - large-context vs. small-context planning
  - backtracking vs. non-backtracking
- Translation Revision:
  - online revision vs. end revision
  - constant revision
Professional vs. Novice Translators

- Professional translators:
  - head-starting
  - small-context planning
  - end-revising

- Student translators:
  - systematic initial orientation
  - large-context planning
  - online revising
WP1 Task 1.1
Unexperienced Post-editor
WP1 Task 1.1
Post-editing: reading ST; checking TT
WP1 Task 1.1
Post-editing: reading TT; checking ST
Cognitive Analysis

Initial Objective: to explore human error-checking behaviour in a simulated post-editing environment.
These are more extreme examples but it is always worth having a human with knowledge of the target language check the MT output. (Mistakes are costly and embarrassing.)
But how are errors or “clunky” translations identified and corrected?

• A lot of translation and post-editing is highly subjective
  • $x$ translators will result in $\geq x$ versions.
• Are some scanning techniques / styles more effective and efficient than others?
• What types of error are easier to spot (and why)?
  • Immediate vs only at end of phrase or sentence;
  • Identified in isolation or broader contextual knowledge required.

Research programme: start with the basics and establish a clear baseline.
Two Stages of Post-Editing

1) Checking and identification  
2) Revision

Error Classification

Obviously some overlap, but:

i. Non-language specific – analogous to more conventional proof-reading;

ii. Translation specific – more subtle language differences between source and target.

Are different “anomaly detection” processes involved?
Methodology: Eye-tracking

1. Precise indication of where attention is focused.

2. Gaze duration is an indicator of processing effort.
Majority of experimental items based on the German-to-English WTM12 shared task (EuroMatrixPlus project) so original German source exists.

1. Easy letter transposition (typo): first two letters of word switched to produce a non-word.

   An unknown person shot an employee at the mimigration office of the city of Gelsenkirchen.
   An unknown person shot an employee at the immigration office of the city of Gelsenkirchen.

2. Hard letter transposition (typo): two internal letters switched to produce a real but inappropriate word (spell-checker would fail to identify).

   Learning how to remain clam in times of stress will help things go more smoothly.
   Learning how to remain calm in times of stress will help things go more smoothly.

The other inmates on the tram were unhurt.
The other passengers on the tram were unhurt.
[Die übrigen Insassen der Tram blieben unverletzt.]

4. Mistranslation: tense or agreement.

Earnings of both Italian and Spanish bonds is growing towards critical limits.
Earnings of both Italian and Spanish bonds are growing towards critical limits.

5. Word order

The fans of the saga vampire were not disappointed.
The fans of the vampire saga were not disappointed.
[Die Vampirsaga-Fans wurden nicht enttäuscht.]
Exp 1 (continued)

• Only one error per sentence.

• Not all sentences contain errors.

Task

• After reading a sentence, the participant clicks the left mouse button if they identified an error or the right button if they did not.

• If yes, the sentence is redisplayed and the task is then to click on the first word of any error.
Phase 1: monolingual (native English speakers). Baseline condition.

Phase 2: bilinguals and professional translators. Contrast condition.

Convention for investigations into bilingual (dis)advantages to contrast mono- and bilinguals on the same tasks.

Aim to validate the procedure by replicating existing monolingual error-detection findings and task effects while also enabling comparisons with language-critical (i.e. translation) errors.
The fans of the saga vampire were not disappointed.
Example 2

The other inmates on the tram were unhurt.

mistranslation

<table>
<thead>
<tr>
<th>error_type</th>
<th>WORD_CLICK_POSITION</th>
<th>WORDS_LIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>mistranslation</td>
<td>[0, 0, 1, 0, 0, 0, 0, 0]</td>
<td>'The', 'other', 'inmates', 'on', 'the', 'tram', 'were', 'unhurt.'</td>
</tr>
</tbody>
</table>