Instructions for the Pre-main Post-editing Task

Profile of the Translator

- Any professional translator profile with proficiency in Italian and English is suitable for this post-editing assignment. If possible, we would prefer translators with 3+ years of professional translation experience and 1+ years of post-editing experience.

- All translators will each post-edit the same set of English sentences that we previously machine-translated into Italian. The translations should be carried out independently by the translators, and any discussion between translators that could influence the post-edited outcomes should be avoided.

Contents of the Task

- We provide **a folder** containing 6 **.txt files** named doc1.txt, doc2.txt, … containing the original English sentences and the machine translation outputs in Italian. These files are in a format accepted by our custom translation interface (GroTE) that will be used for this study.

- This pre-task objective is for all translators to post-edit all texts in all provided files using our interface. For additional information on using the interface, refer to the interface guide we provided.

- Each file represents a piece of a document (either a scientific abstract or a social media post). Files contain between 3 and 8 segments each, with 38 editable segments in total. All segments in the same file are contiguous sentences sourced from the same text, and translators are expected to use the available context information during the translation when appropriate.

- The total number of English words to translate for every translator is 887. The machine-translated Italian outputs needing post-editing contain 902 words.

Editing Settings

- Translators must post-edit the provided machine translations using the interface provided. The interface records editing actions, so copy-pasting to external editors is not allowed.
- Translators are asked **to fill in a brief questionnaire _before_** starting the task. Each translator identifies themselves using their name in "Firstname Lastname" format.

- The final outcomes of the post-editing are expected to be **publishable, professional-quality** translations. We **advise against fully deleting the machine-translated segments and starting a translation from scratch** unless necessary.
- Translators can use any source on the Internet to identify correct terminology, but **no other machine translation systems** (such as Google Translate, DeepL, or Microsoft Translator) **should be used to provide alternative translations.**
- Segments presented to the translators in the interface for post-editing may contain highlights denoting two severity levels for potential mistakes in the automatic translation. If deemed useful, these can be used to support the editing process, but it is important to remember that **highlighted spans are not guaranteed to be wrong, and non-highlighted spans are not guaranteed to be correct**, hence simply editing highlights might not be sufficient to produce publishable quality translations.
- Highlights on a span are removed automatically after the span is edited. The green checkmark on the top right corner of every segment can be clicked to remove highlights for that segment. These can be done, for example, after finishing the editing of a segment to avoid leftover highlights on non-edited spans. **Manually removing all highlights is not mandatory**, however, and the final output can be downloaded even if some highlights are still present. In this case, they will just be automatically removed from the output.
- The editing flow for each file is the following:
  1. Open the webpage of the GroTE interface
  2. Insert the provided login code
  3. Load one of the provided files
  4. Press "📝 Start"
  5. Perform the editing following the guidelines above
  6. When all segments for the file are finished, click "✅ Done"
  7. A message *"Saving trial information. Don't close the tab until the download button is available!"* will appear. **Do not close the tab.**
  8. When the message *"Saving complete! Download the output file by clicking the 'Download translations' button below."* appears, click "📥 Download translations" to download the edited files. The file will have the name "<LOGIN CODE>_<FILENAME>_output.txt"
  9. Click "⬅ Back to data loading" to return to the file loading page.
  10. **If needed, pause and take a break**

Repeat steps 2 to 9 for each one of the files.

Delivery

- The deliverable should contain a folder named "Firstname Lastname" for each translator, containing the output files produced using our interface.

Queries

- We are aware that you will solve all translation issues internally to ensure the project's consistency. However, if you have any questions, you can direct them to g.sarti@rug.nl.