Controlled Languages and Machine Translation

Introduction

This course introduces the concept of *controlled languages* as a key aspect in the new working environment for the professional translator, where *Machine Translation* is the centre. In this context, the course presents how such scenario can be tailor-made to answer the needs of the translation industry. Namely, the course concentrates on the introduction of controlled languages techniques to students, on training in pre-editing skills for machine translation use, and on post-editing practice.

The translation industry has dramatically changed over the last ten years as the gradual process of introducing translation memory (TM) and machine translation (MT) systems in the documentation and translation workflow has been gaining momentum. Also, the burgeoning development of the software localization business in an increasingly globalized world presents a significant challenge to the translation industry by multiplying the number of languages that need to be translated and by expecting shorter turnaround time of translations. Online documentation and web delivery call for a responsive translation industry ready to deliver in a highly competitive environment.

In this scenario, the use of MT systems in the industry has spawned two additional fields that require careful consideration by the professional translator: the use of controlled languages in specific industries (aeronautics, automotive industry, heavy-equipment manufacturers, software companies, etc.) and the post-editing needs of documents translated with MT systems. Future professional translators need to be aware of issues and skills such as what minimum post-editing entails, which standard metrics are starting to be developed to score machine translation output, which controlled languages are used in the industry, how to write documents in controlled language to get better results when using MT systems, etc. They also need to have a thorough exposure to different professional and online MT systems, not only from the user´s viewpoint but also from the developer´s since only by a good understanding of the complexity of these systems can the future translator judge the quality of the output and foresee what errors will be frequent.

Course / module Description

The course uses the automotive industry as an example of a translation scenario. Therefore, the student will learn *controlled languages and machine translation* by working on the following aspects:

- Understand the **multilingual requirements and needs of the automotive industry** by reviewing the different document types, the complexity of technical terminology, and the targeted audience.
- Study existing **Controlled Language rules** from different industries such as Caterpillar Technical English and AECMA Simplified English, in order to adapt them and come up with suitable rules for the technical documentation used in the automotive industry.

- **Compile specialized automotive bilingual glossaries** to feed ENGSPAN´s English-to-Spanish dictionaries placing special emphasis on the disambiguation of multiple noun compounds describing procedures and parts.
- Experiment with the advanced features of ENGSPAN, which allows users to **customize the system** by modifying the grammatical coding of lexical entries according to parsing information.
- Learn how to **post-edit the raw MT output** generated by ENGSPAN following recommendations for minimal post-editing principles and practice these principles extensively.
- Perform **evaluations of the translation quality using SAE J2450.**

Training topics

- Introduction to controlled languages for Machine Translation
- Pre-editing: definition, rules, applications
- Post-editing: definition, rules, applications
- Controlled languages and their impact in translation quality control