



Friday, November 24, 2017



Academic coordinators: Laurent Gautier & Will Noonan (uBFC)
Research centre: Centre Interlangues (EA 4182)

**How translators manage generic and contextual constraints:
Annotation, modelling, and information extraction using electronic
corpora.**

Context:

In recent years, the field of translation studies has come to focus not only on the outcomes and applications of traditional academic research, but also, increasingly, on project-based “action research”, in which academics, professionals working in the field, and representatives of the language industry all play an active role. At the same time, projects focused specifically on the role and use of corpora in translator training and professional translation practice (*corpus-based translation studies*) have come to supplement traditional corpus linguistics research. A similar tendency can be observed in the growing role of semantically enriched multilingual text resources in the field of digital humanities, seen, for example, in the dedicated platforms developed by the network of *Maisons des Sciences de l’Homme* in France. In this context, then, this one-day workshop aims to stimulate dialogue between researchers, professional translators, language industry entrepreneurs and students in translation, including from the MA in Multimedia Translation (T2M) course offered at the University of Burgundy.

Focus:

This workshop aims to bring together, and build on, three main observations:

- corpus-based research has tended to focus on microlinguistics, including terms, collocations and recurring syntagms. The results of such research can be observed, for example, in the functioning of termbases and translation memories.
- professional translators have increasingly come to recognize the importance of the “macro” dimension of source texts, including macrolinguistic structures and generic markers. This dimension has long been recognised in the field of technical communication.
- the growing body of research on the textual architecture modelling at different levels of granularity (seen, for example, in the Text Encoding

Initiative) has made it possible to compare convergent and divergent structures within multilingual corpora.

Aims:

This workshop aims to stimulate discussion between professional translators, developers of CAT and MT tools, and researchers in the fields of Applied Linguistics, translation studies and information technology, around the question of how to manage generic criteria:

- for academics: how can research results in text linguistics be modelled?
- for professionals: what software solutions exist? how can new research be incorporated into these tools?
- for translator trainers: how can these innovations be brought to bear when training translators in the use of corpora?

Presentation formats:

Proposals are requested for one of two formats:

- 20-minute academic papers, to be followed by discussion time as part of a thematic panel;
- 15-minute presentations of software solutions, whether at prototype or commercial stage.

The day will end with a round table discussion whose content will be determined by the proposals received. Language industry representatives will be offered the chance to demonstrate their products during the day.

Academic proposals should be written in English or French, and include a title, a 500-word abstract, 5 keywords and a short bio note.

Proposals for software presentations should include a 500-word description of the tool concerned (including its purpose, functions and architecture), along with a brief presentation of the company and of the state of the project.

Proposals are due by September 1st at the latest, and should be sent to laurent.gautier@ubfc.fr and will.noonan@ubfc.fr.

