

In search of specialized terminology and phraseology with the help of corpora

Half-day workshop led by [Ana Frankenberg-Garcia](#), [Centre for Translation Studies](#), University of Surrey.

Linguistic proficiency and translation experience are not in themselves enough for translators to deal with the language of specialized texts. Such texts contain terminology and phraseology that are often not accessible or even comprehensible to non-experts. It can take years for translators who have not received any training in a given specialist domain to feel confident enough to translate texts in this field. Bilingual or target language glossaries and term bases – if they are available – can be a good starting point. But translators need more than just equivalents; they also need reassurance from longer stretches of text, and may thus look for this reassurance in comparable or so-called “parallel” texts in the target language. However, it can take ages to find such texts and even longer to read them and locate the exact passages that will help. Could there be a quicker and easier way of becoming acquainted with the terminology and phraseology necessary to work with specialist translation?

In this hands-on workshop, participants will learn to use corpora to help them translate in a domain where they are not (yet) experts. The workshop will begin with an introduction to different ways in which corpora and corpus software can help translators (Frankenberg-Garcia, 2015). No prior knowledge of corpora is assumed. This will be followed by the practical, hands-on part of the workshop, where the participants will be guided through using the state-of-the-art [Sketch Engine](#) tools (Kilgarriff et al. 2004, 2014) developed by Lexical Computing. First they will take a quick look at how pre-loaded corpora that are available from the Sketch Engine can be useful to translators. These pre-loaded corpora include the British National Corpus, the 16 billion-word deTenTen corpus compiled in 2013 for German, and corpora in over sixty languages, including multilingual, parallel corpora. Next, participants will be shown how to build a corpus in a specialist domain by (1) uploading pre-selected files and (2) using WebBootCaT (Baroni et al. 2006), a tool for producing instant web-crawled corpora in any field and language. Participants will learn how to identify specialized terminology in these corpora and how to look up terms in context to acquire also the phraseology needed for their translations. Finally, participants will be shown how to create a specialized corpus to explore terminology and phraseology in any field and language they wish.

By the end of the workshop, participants will be able to (1) understand how corpora can be used as an aid to specialist translation in any language, (2) build a specialized language corpus in a domain and language of their choice, (3) use the corpus to help them with the terminology and phraseology required to translate in a specialized domain and (4) keep the corpus for future reference.

References

- Baroni, M., Kilgarriff, A., Pomikalek, J. & Rychly, P. (2006) [Webbootcat: Instant domain-specific corpora to support human translators](#). In *EAMT-2006 11th Annual Conference of the European Association for Machine Translation*, June 19 & 20, 2006 Oslo University (Norway), 247–252.
- Frankenberg-Garcia, A. (2015) Training translators to use corpora hands-on: challenges and reactions by a group of 13 students at a UK university. *Corpora*, 10/2, 351-380.
- Kilgarriff, A., Rychlý, P., Smrz, P. & Tugwell, D. (2004) [The Sketch Engine](#). *Proceedings of EURALEX 2004*, Lorient, France, 105–116.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014) [The Sketch Engine: ten years on](#). *Lexicography* 1/1, 7-36.