

Visualisation of terminology in a corpus on climate change

Abstract

Recent evolutions in software and practices enable different modalities to be included in corpus analysis, not only from the point of view of the object analyzed (corpus with multimodal data) but also from the point of view of the tools that are now accessible to identify significant patterns in the data. Thanks to visualisation methods, precious synthetic views of the corpus become accessible (Anthony 2018, Picton, Drouin & Humbert-Droz 2020, Chambó & León-Araúz 2021). It may be useful to question the impact of these tools on the researcher's intellectual process when analyzing a specialized corpus. Should we consider this simply as a different presentation of data, with no fundamental change in the object being analyzed? Or do these tools contribute to radically changing the linguist's perception of his corpus by shedding light on processes that remained invisible otherwise? To contribute to the reflection on these questions, we study the specialized terminology in a corpus on climate change including two discourse communities (UN experts and the general press) over three distinct time periods (years around the COP15, COP21 and COP26). Drawing on methods inspired by Lavrentiev, Sherstinova, Chepovskiy & Pincemin (2021) a correspondance analysis tool is used alongside a specificity graph from the software TXM to consider diachronic variation in the two sub-corpora. These tools allow a quick comparison of keywords in each sub-corpora and thus offer a relevant starting point for the analysis of the specialized terminology as it is used in discourse. They seem to offer great potential for the analysis of specialized languages.

References

- Anthony L., 2018, "Visualisation in Corpus-Based Discourse Studies", in Taylor C. & A. Marchi (Eds), *Corpus Approaches to Discourse*, Routledge: Oxon, 197-224.
- Chambó S. & P. León Araúz, 2021, "Visualising Lexical Data for a Corpus-Driven Encyclopedia", *Electronic lexicography in the 21st century. Proceedings of the eLex 2021 conference*, 5–7 July 2021, virtual. Brno: Lexical Computing CZ, s.r.o., 29-55.
- Lavrentiev A. M., Sherstinova T. Y., Chepovskiy A. M., Pincemin B., 2021, "Using TXM Platform for Research on Language Changes Over Time: The dynamics of vocabulary and punctuation in Russian literary texts", *Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya*, (70), 69-89, <https://doi.org/10.17223/19986645/70/5>.
- Picton, A., Drouin, P., & Humbert-Droz, J., 2020, « Exploration et visualisation de la variation terminologique en corpus spécialisés complexes : réflexions et propositions méthodologiques » [Chapitre de livre]. Dans *Lexique(s) et genre(s) textuel(s) : approches sur corpus*. Actes de la conférence 11e Journées du réseau "Lexicologie, Terminologie, Traduction" (Vol. , p. 99-116), Éditions des archives contemporaines <https://doi.org/10.17184/eac.9782813003454>.