

Ocene in poročila – Reviews and Reports

DIHUR Team Members Passing on the Enthusiasm for ParlaMint

In the summer of 2023, the Institute of Contemporary History (INZ) researchers involved with the Digital Humanities: Resources, Tools and Methods Research Programme (DIHUR) participated in two training activities aimed at demonstrating the potential of multilingual parliamentary corpora and their user-friendly querying via concordancers for research purposes in Social Sciences and Humanities.

On 11 July 2023, Dr Darja Fišer, Dr Anna Kryvenko, and Kristina Pahor de Maiti (all INZ), together with Dr Petya Osenova (Bulgarian Academy of Sciences; Sofia University “St. Kl. Ohridski”, Bulgaria), delivered a pre-conference tutorial at the Digital Humanities 2023 conference, organised by the Alliance of Digital Humanities Organizations (ADHO) and hosted in Graz, Austria. The DH 2023 conference theme, “Collaboration as Opportunity”, addressed the issue of transdisciplinary and transnational collaboration with a particular focus on the expanding South-Eastern European Digital Humanities community.¹

The pre-conference tutorial *Put Them In to Get Them Out: the ParlaMint Corpora for Digital Humanities and Social Sciences Research* was designed in response to an ever-increasing interest in records of parliamentary debates across Europe as a fruitful object of study in a variety of disciplines in Social Sciences and Humanities. Its overarching goal was to engage a broad range of DH scholars into exploring the potential of parliamentary corpora as a language resource for studying socio-political phenomena by introducing them to the ParlaMint corpora (Erjavec et al. 2023).²

In this half-day tutorial, the instructors focused on an overview of the ParlaMint project and the description of the corpus structure, including types of metadata and linguistic annotation, the compatibility potential of the corpora in the project, as well as parliament and language-specific characteristics that impact cross-corpus compatibility. The overview was followed by an exploration of the ParlaMint corpora via the Crystal NoSketch Engine concordancer and a presentation of basic corpus analysis techniques. The hands-on part of the tutorial revolved around finding ways to answer

1 Read more about the INZ team activities at the DH 2023: <https://www.inz.si/sl/Dogodki/INZ-na-konferenci-Digital-Humanities-2023/:%20>.

2 Tomaž Erjavec et al., *Linguistically annotated multilingual comparable corpora of parliamentary debates ParlaMint.ana 3.0* (Slovenian language resource repository CLARIN.SI 2023), ISSN 2820-4042, <http://hdl.handle.net/11356/1488>.

research questions related to the discursive construction of socially significant concepts in the British parliamentary corpus (ParlaMint-GB) in the 2015–2022 period. Furthermore, the participants were invited to try the corpus analysis techniques learned on the ParlaMint corpora of their choice; and discuss the comparability potential compared to the ParlaMint-GB corpus both in view of metadata availability as well as content. The participants reported their results, compared and discussed their findings, and provided feedback on the applicability of the knowledge and skills acquired during the tutorial to their own research.

The tutorial was fully booked and well attended by registered participants representing nine countries from three different continents. Since neither programming skills nor prior experience in using language corpora or corpus querying tools were required to take part in this tutorial, it attracted the attention of scholars with various backgrounds ranging from computer science and information modelling to linguistics, history, art, and cultural studies. Extensive comments from the participants during the tutorial and the subsequent discussions between the instructors and the tutorial attendees on the margins of the conference in the following days proved the relevance of FAIR (findable, accessible, interoperable, and reusable) parliamentary corpora to the Digital Humanities community with regard to further research into specific national parliaments or trans-nationally as well as cross-disciplinary cooperation opportunities.

A few weeks later, from 31 July to 4 August 2023, Dr Anna Kryvenko and Kristina Pahor de Maiti conducted a one-week workshop titled *Combining Corpus Linguistics and Discourse Analysis to Explore the Parliamentary Debates across Europe* at the 13th European Summer University in Digital Humanities “Culture and Technology”, organised by the Transylvania Digital Humanities Centre (DigiHUBB) and hosted by the Babeş-Bolyai University, Romania. The target audience of the workshop comprised students and scholars in Digital Humanities or Social Sciences who work with language data. The participants interested in political discourse particularly benefited from the workshop, as the tasks included the topics of the discursive construction of Europe and the European Union by different political groups and actors, political polarisation, and gender representation in national parliaments.

The use cases for this workshop came solely from the ParlaMint corpora, including the parliamentary debates in national or regional languages (Erjavec et al. 2023) and their machine translation into English (Kuzman et al. 2023).³ However, the instructors emphasised that the techniques and tools used during this workshop can be applied to a vast selection of resources and research questions. Among other things, they demonstrated how creating and comparing different data subsets can help refine research questions and improve research design. On the other hand, it was made clear that corpus linguistic techniques are not self-sufficient and that an interpretation of the

3 Taja Kuzman et al., *Linguistically annotated multilingual comparable corpora of parliamentary debates in English ParlaMint-en.ana 3.0* (Slovenian language resource repository CLARIN.SI 2023), ISSN 2820-4042, <http://hdl.handle.net/11356/1810>.

extracted patterns of language use as forms of social interaction requires the utilisation of one of the specific theoretical perspectives in the field of discourse analysis and a set of respective analytical approaches applied to the data extracted from the corpus with the help of corpus analysis techniques.

The instructors used mixed teaching strategies, including elements of personalised learning, inquiry-based learning, and project-based learning, to better tailor the workshop contents to the participants' individual learning goals and interests, where possible, and encourage them to incorporate the newly learned skills in their current and future research projects. Based on the end-course student feedback, the workshop was evaluated as useful and relevant for their research, although only a minority of the participants initially planned to work with parliamentary data. The overall outreach of this workshop exceeded the dedicated group, as there was a short teaser session open to participants of the other workshops at the European Summer University in Digital Humanities, which was also well attended.

Through a critical analysis of the lessons learned, it becomes evident that training courses aimed at empowering scholars and students within the Digital Humanities community on the one hand, and fostering innovative research on the other, should be balanced in terms of: a) being applicable to other similar resources and tools as well as being adaptable to individual research interests and goals; b) exploring the possibilities and addressing the challenges or limitations offered by the resources and tools in focus; c) tackling the issue of finding the right fit among the research questions, data, methods, and theory. Overall, the experience gained this summer by some of the DIHUR team members confirms the need for tailored training courses and underlines the added value of well-constructed comparable resources that can be reused from different scientific angles and explored with user-friendly tools such as concordancers.

*Anna Kryvenko
Kristina Pahor de Maiti Tekavčič*