Draft blog post …

**Reflecting on the CLARIN workshop at the Corpus Linguistics 2023 conference**

**About**

On Sunday 2nd July 2023, members of CLARIN hosted an in person workshop titled [“What can you do with the CLARIN research infrastructure?”](https://www.lancaster.ac.uk/cl2023/pre-conference-workshops/). This took place on the day before the [Corpus Linguistics (CL2023) conference](https://www.lancaster.ac.uk/cl2023/) began at Lancaster University in the UK. The workshop focussed on practical issues in terms of how corpus linguists can benefit from the CLARIN network and infrastructure. It was organised by Darja Fišer, Francesca Frontini, Paul Rayson and Martin Wynne, and we were joined for a special preview session on ParlaMint presented by Dario Del Fante. The workshop attracted 25 participants.

**Practical sessions**

The three hour long workshop began with a welcome from [Paul Rayson](https://www.research.lancs.ac.uk/portal/en/people/paul-rayson/) from the [UCREL research centre](https://ucrel.lancs.ac.uk/) at Lancaster University, and a [CLARIN ambassador](https://www.clarin.eu/content/clarin-ambassadors-programme), who introduced the participants to the aims of the event. Next, [Martin Wynne](https://www.ling-phil.ox.ac.uk/people/martin-wynne) from Oxford University and National Coordinator of [CLARIN-UK](https://www.clarin.ac.uk/), gave a broad overview of the CLARIN infrastructure, covering the distributed network of over 60 centres of expertise, resource families, and key FAIR principles including the interoperability of language resources, corpora, metadata and tools.

Next, [Dario Del Fante](https://docente.unife.it/dario.delfante) from the University of Ferrara and a member of the [Italian node of CLARIN](https://www.clarin-it.it/), presented a flagship CLARIN project, [ParlaMint](https://www.clarin.eu/parlamint), which is creating comparable multilingual corpora of parliamentary debates across Europe. This vast resource forms part of the [Parliamentary corpora resource family](https://www.clarin.eu/resource-families/parliamentary-corpora) of CLARIN, and brings together transcripts of parliamentary proceedings from multiple countries under a common TEI XML schema format, with Universal Dependencies annotation for morphosyntax, named entities and semantic tagging. Workshop participants saw potential use cases in political sciences, digital humanities as well as corpus linguistics, and were assisted in exploring several practical activities including a corpus assisted gender analysis and the parliamentary language around crises. The workshop happened a few days before the full open access release of [ParlaMint 3.0](http://hdl.handle.net/11356/1486).

Finally, in the second half of the workshop, Martin, Dario and Paul supported the participants as they explored further case studies of their choosing which were selected from a wide set of hands-on activities related to CLARIN’s [Virtual Language Observatory](https://vlo.clarin.eu/), [Switchboard](https://switchboard.clarin.eu/), and [Federated Content Search](https://contentsearch.clarin.eu/). These included extended analyses of gender in parliamentary discourse, creating a linguistically annotated corpus of 19th century English novels, how to search for linguistic patterns across corpora, as well as advance access to UPSKILLS learning content prepared by [Iulianna van der Lek](https://www.clarin.eu/person/iulianna-van-der-lek-ciudin), Training and Education Officer at CLARIN-ERIC.

All slides and handouts are available on the [workshop web page](https://www.clarin.eu/event/2023/clarin-corpus-linguistics-2023).