Dialogues in Endangered Language Corpora
Michael Rießler (Universität Bielefeld)

Despite the digital nature of current methodology, fieldwork-based documentation of endangered languages has rarely considered applying computational methods in building and analyzing the resulting corpora more efficiently. In my presentation, I will provide a brief overview of my own projects-in-progress which attempt to work in the paradigm of endangered language documentation and typologically informed language description while systematically applying methods from NLP.

The languages I work most intensively on at present belong to the Permic, Saamic and Samoyedic branches of Uralic and are all spoken in the Barents Sea region of northeastern Europe. Since dialogues (and multilogues) are also recorded and annotated, the resulting spoken language corpora can likely offer relevant comparative data for a project on dialogue structures.