

D	I	G
H	U	M
L	A	B

Digital
Humanities Lab
Denmark

Language-based
materials and tools



CLARIN-DK

Language data and language tools - the Danish node of an European Infrastructure

The CLARIN-DK archive and tools

The language based data you use for research purposes or want to share with fellow colleagues, can be stored in the CLARIN-DK archive.

The archive contains Danish text corpora, both general and language for special purpose, Danish audio, video and photo collections, as well as lexica, wordnet, and linguistic annotations of texts and videos. Furthermore, CLARIN-DK provides various tools and a help desk.

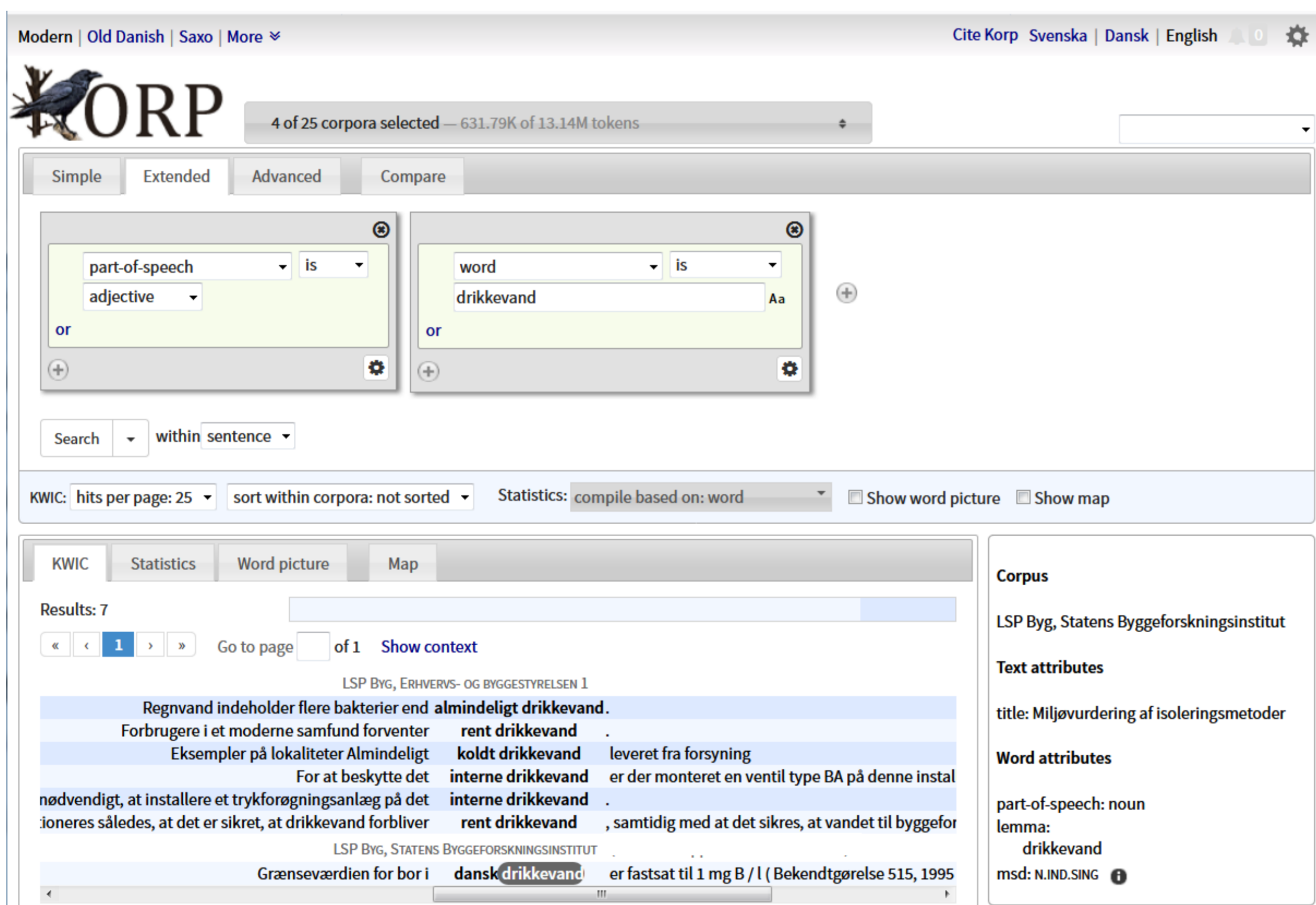
The archive is one of the data archives in the European CLARIN infrastructure.



Search in Danish Corpora

Danish Corpora are available in the Korp concordance search tool.

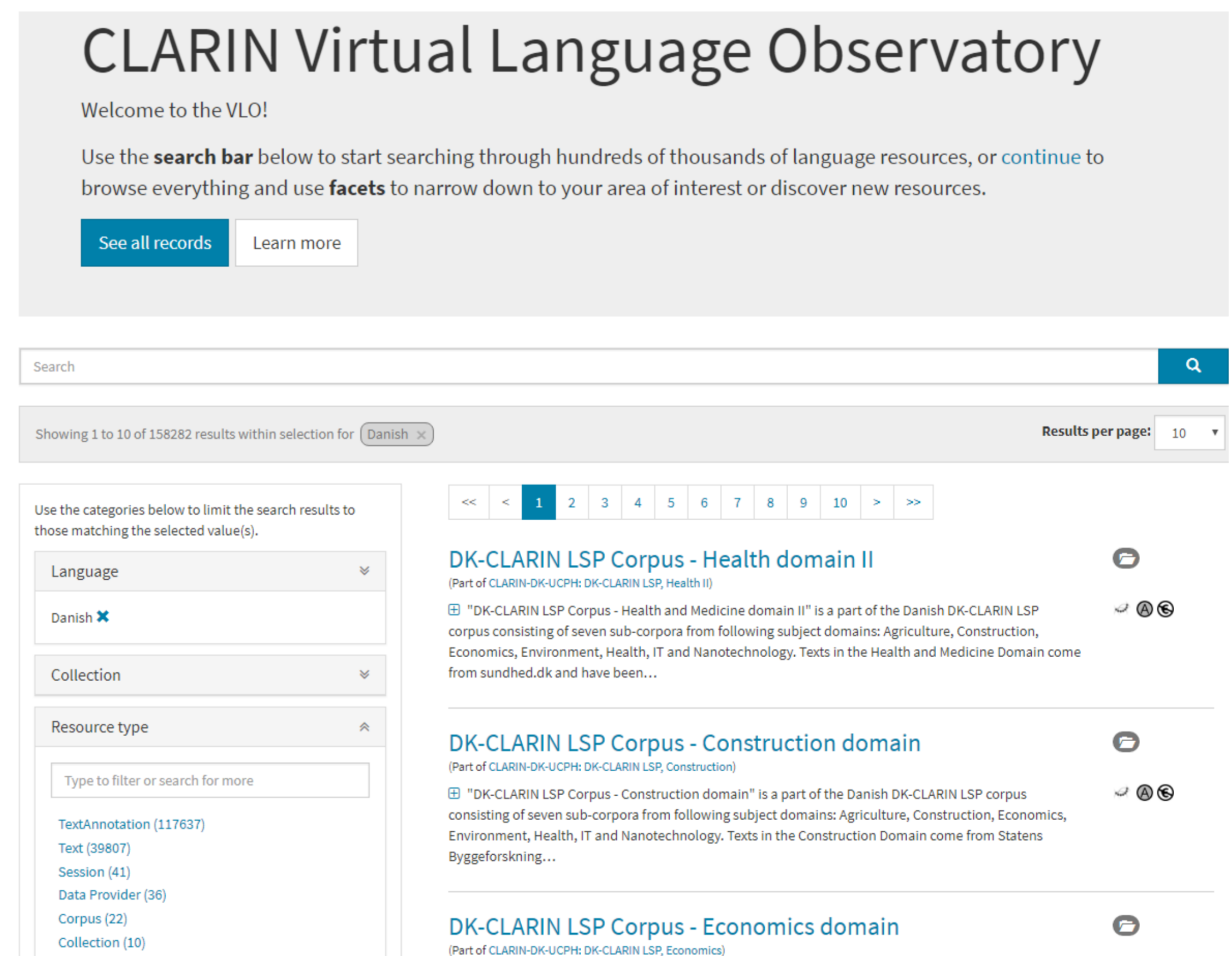
Below a search in the Danish Corpus “Byggeri og anlæg” for adjectives followed by the word “drikkevand”:



Discovering European Language Resources

Clarín.eu offers a search interface – the VLO - to 1.6m language resources of different types.

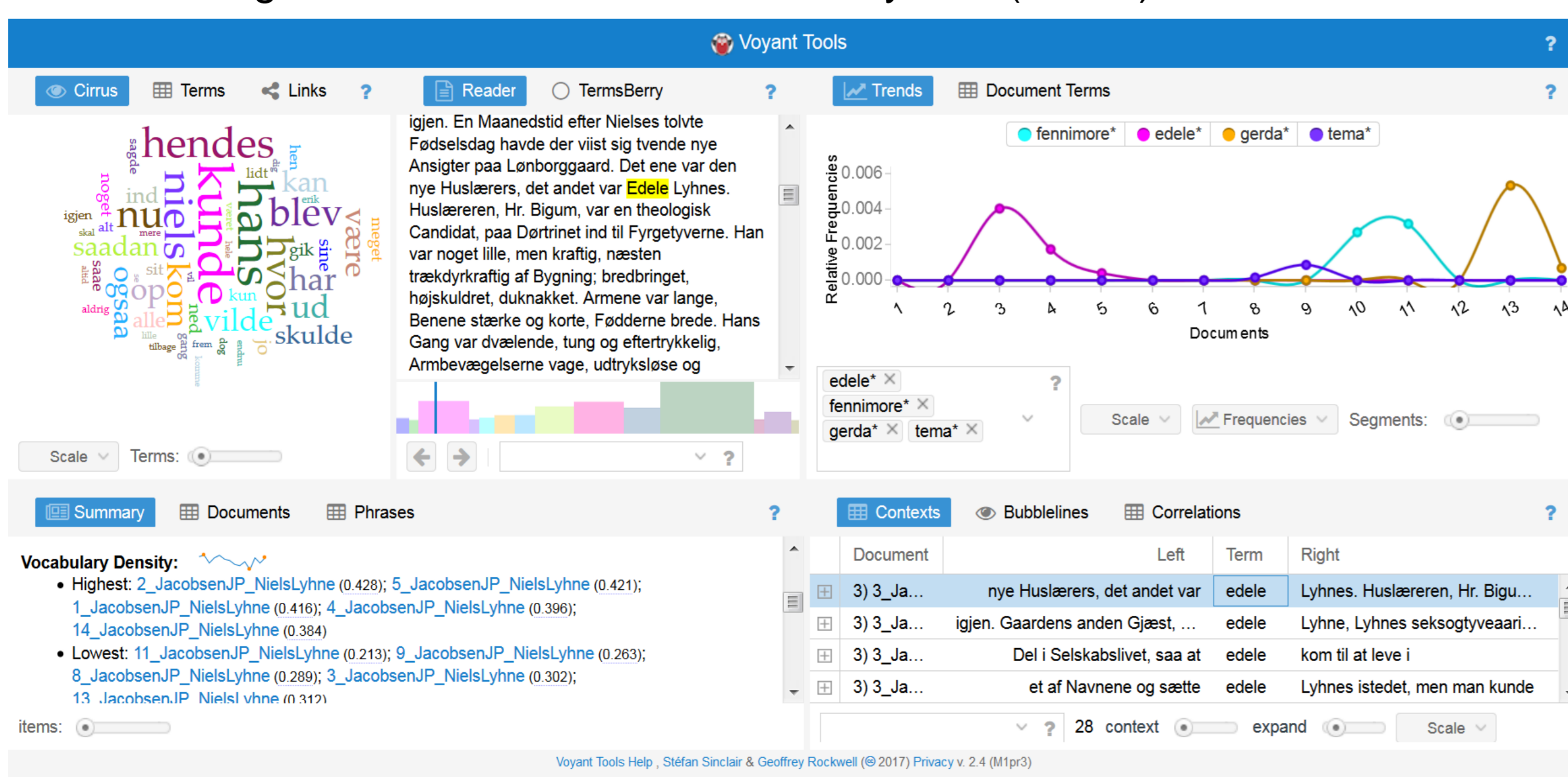
Below a search for resources in Danish:



Distant Reading of Danish Literature

The Voyant Tools allows distant reading of your own or pre-loaded Danish novels.

Below the upper right corner shows the frequencies of the names of four female characters though J.P. Jacobsens novel “Niels Lyhne” (adl.dk):



The CLARIN Toolbox

From the toolbow you can do linguistic annotation of texts, create metadata and format your text into TEI .

Below is the tools for linquistic annotation:

