

Historiallisia korpuksia ja korpustyökaluja

Koonnut Tanja Säily

Englannin kielen historiallisia korpuksia:

- Brown Corpus, <https://varieng.helsinki.fi/CoRD/corpora/BROWN/>
- Corpus of Early English Correspondence (CEEC), <https://varieng.helsinki.fi/CoRD/corpora/CEEC/>
- Corpus of Historical American English (COHA), <https://www.english-corpora.org/coha/>
- Helsinki Corpus of English Texts (HC), <https://varieng.helsinki.fi/CoRD/corpora/HelsinkiCorpus/>

Muiden kielten historiallisia korpuksia:

- Corpus del Español, <https://www.corpusdelespanol.org>
- Corpus do Português, <https://www.corpusdoportugues.org>

Digitoituja tekstietokantoja:

- Early English Books Online (EEBO), <https://www.english-corpora.org/eebo/>
- Kansalliskirjaston lehtikokoelma, <https://digi.kansalliskirjasto.fi>

Mistä korpuksia voi löytää:

- Corpus Resource Database (CoRD; vain englantia), <https://varieng.helsinki.fi/CoRD/>
- Kielipankki, <https://www.kielipankki.fi>
- Language Resource Inventory, <https://www.clarin.eu/content/language-resource-inventory>
- Linguistic Data Consortium, <https://catalog.ldc.upenn.edu>
- META-SHARE, <http://www.meta-share.org>

Korpustyökaluja:

- Web-käyttöliittymä
 - Korp, esim. <https://korp.csc.fi> (sisältää useita korpuksia)
 - CQPweb, esim. <https://cqpweb.lancs.ac.uk> (sisältää useita korpuksia)
- Erillinen konkordanssiohjelma
 - AntConc, <https://www.laurenceanthony.net/software/antconc/>
 - WordSmith Tools, <https://lexically.net/wordsmith/>
- Ohjelmointikieli
 - Python, <https://www.python.org>
 - R, <https://www.r-project.org>

Lisää resursseja on lueteltu vapaasti saatavilla olevassa *Kielentutkimuksen menetelmiä* -kirjassa (luvun 9 lopussa ja luvussa 11): <https://doi.org/10.21435/skst.1457>