

LingTUtorial: Corpus Analysis

List of corpora



British National Corpus (BNC): <https://corpus.byu.edu/bnc/>

Corpus of American Soap Operas (SOAP): <https://corpus.byu.edu/soap/>

Corpus of Contemporary American English (COCA): <https://corpus.byu.edu/coca/>

Corpus of Global Web-based English (GloWbE): <https://corpus.byu.edu/glowbe/>

Corpus of Historical American English (COHA): <https://corpus.byu.edu/coha/>

COSMAS II (DWDS corpus, IDS): <http://www.ids-mannheim.de/cosmas2/>

International Corpus of English (ICE): <http://ice-corpora.net/ice/>

International Corpus of Learner English (ICLE): not available online, but in the English department of the TU Braunschweig

Michigan Corpus of Academic Spoken English (MiCASE):

<https://quod.lib.umich.edu/cgi/c/corpus/corpus?c=micase;page=simple>

Projekt Deutscher Wortschatz: <http://wortschatz.uni-leipzig.de/de>

Subtitle Frequency Corpora (SUBTLEX): <http://crr.ugent.be/programs-data/subtitle-frequencies>

Talkbank (CHILDES, SLA Bank etc.): <https://talkbank.org/>

TIME Magazine Corpus: <https://corpus.byu.edu/time/>

WebCELEX: <http://celex.mpi.nl/>