

**English-Corpora.org**[Corpora Overview](#) [PDF](#) [Videos](#) [Resources](#) [Help / FAQ](#) [My account](#)**English-Corpora.org**[corpora](#)[Overview](#) [Language learning](#) [Insight into variation](#) [Queries](#) [Variable length queries \(new\)](#) [Representativity](#) [Compare COCA and BNC](#) [Size](#) [Speed](#) [Collocates+](#) [Training / workshops](#) [History / updates](#) [↑ corpus.byu.edu](#)[PDF guides](#)[Overview / guided tour](#) [Architecture](#) [Association measures](#) [Collocates \(cf Sketch Engine\)](#) [Topics \(and collocates\)](#) [Word sketches](#) [Browsing words](#) [Analyzing texts](#) [KWIC -> analyze text](#) [Saved words and phrases](#) [Saving KWIC entries](#) [Customized word lists](#) [Search history](#) [External resources](#) [Monitor corpus](#) [Virtual Corpora VC: quick overview](#)[videos](#)[Overview](#) [Language learning and teaching](#) [Word sketches](#) [Browsing words](#) [Analyze texts](#) [Search history](#) [Customized word lists](#) [Saved words \(favorites\)](#) [KWIC lines: limiting and sorting](#) [Saved KWIC lines](#) [Analyze KWIC lines](#) [External resources](#) [Virtual Corpora](#) [Examining recent change](#) [Unable to access YouTube?](#)[related resources](#)[Overview](#) [Full-text data](#) [Word frequency](#) [Collocates](#) [N-grams](#)[users](#)[Overview](#) [Number of users](#) [Researchers](#)[my account](#)[Register / profile](#) [Log in](#) [Log out](#) [Name of university](#) [Reset password](#) [Delete account](#)[upgrade](#)[Premium \(individual\) license](#) [Academic \(group\) license](#)[help](#)[FAQs](#) [Citing / screenshots](#) [Problems](#) [Workshops](#) [Third party materials](#) [Contact us](#)


<a href="#">Corpus</a>	<a href="#">Size</a>	<a href="#">Countries</a>	<a href="#">Time</a>	<a href="#">Genre</a>
<a href="#">IWEB</a>	13.9 <b>b</b>	6	2017	Web
<a href="#">NOW</a>	16.2 <b>b</b>	20	2010-now	Web: News
<a href="#">CORONA</a>	1.58 <b>b</b>	20	2020-now	Web: News
<a href="#">GLOWBE</a>	1.9 <b>b</b>	20	2012-13	Web/blogs
<a href="#">WIKI</a>	1.9 <b>b</b>	(+)	2014	Wikipedia
<a href="#">COCA</a>	1.0 <b>b</b>	Am	1990-2019	Balanced
<a href="#">COHA</a>	400m	Am	1810-2009	Balanced
<a href="#">TV</a>	325m	6	1950-2018	TV shows
<a href="#">MOVIES</a>	200m	6	1930-2018	Movies
<a href="#">SOAP</a>	100m	Am	2001-2012	TV shows
<a href="#">HANSARD</a>	1.6 <b>b</b>	Br	1803-2005	Parliament
<a href="#">EEBO</a>	755m	Br	1470s-1690s	Various
<a href="#">SUP.CRT</a>	130m	Am	1790s-2010s	Legal
<a href="#">TIME</a>	100m	Am	1923-2006	Magazine
<a href="#">BNC</a>	100m	Br	1980s-1993	Balanced
<a href="#">CAN</a>	50m	Can	1970s-2000s	Balanced
<a href="#">CORE</a>	50m	6	2014	Web

**Overview:** [brief](#) | [detailed](#)

These are the most [widely used](#) online corpora, and they are used for [many different purposes](#) by teachers and [researchers](#) at [universities](#) throughout the world. In addition, the corpus data (e.g. [full-text](#), [word frequency](#)) has been used by a [wide range of companies](#) in many different fields, especially technology and [language learning](#).

The links below are for the free online interface. You can also purchase and download the corpora for use on your own computer.

Corpus	Download	# words	Dialect	Time period	Genre(s)
<a href="#">News on the Web (NOW)</a>		19.2 <b>billion+</b>	20 countries	2010- <b>yesterday</b>	Web: News
<a href="#">iWeb: The Intelligent Web-based Corpus</a>		14 <b>billion</b>	6 countries	2017	Web
<a href="#">Global Web-Based English (GloWbE)</a>		1.9 <b>billion</b>	20 countries	2012-13	Web (incl blogs)
<a href="#">Wikipedia Corpus</a>		1.9 <b>billion</b>	(Various)	2014	Wikipedia

<a href="#">Coronavirus Corpus</a>		1.5 <b>billion</b>	20 countries	2020-2023	Web: News
<a href="#">Corpus of Contemporary American English (COCA)</a>		1.0 <b>billion</b>	American	1990-2019	Balanced
<a href="#">Corpus of Historical American English (COHA)</a>		475 million	American	1820-2019	Balanced
<a href="#">The TV Corpus</a>		325 million	6 countries	1950-2018	TV shows
<a href="#">The Movie Corpus</a>		200 million	6 countries	1930-2018	Movies
<a href="#">Corpus of American Soap Operas</a>		100 million	American	2001-2012	TV shows
<a href="#">Hansard Corpus</a>		1.6 <b>billion</b>	British	1803-2005	Parliament
<a href="#">Early English Books Online</a>		755 million	British	1470s-1690s	(Various)
<a href="#">Corpus of US Supreme Court Opinions</a>		130 million	American	1790s-present	Legal opinions
<a href="#">TIME Magazine Corpus</a>		100 million	American	1923-2006	Magazine
<a href="#">British National Corpus (BNC)*</a>		100 million	British	1980s-1993	Balanced
<a href="#">Strathy Corpus (Canada)</a>		50 million	Canadian	1920s-2000s	Balanced
<a href="#">CORE Corpus</a>		50 million	6 countries	2014	Web
From <a href="#">Google Books n-grams</a> ( <a href="#">compare</a> )					
<a href="#">American English</a>		155 billion	American	1500s-2000s	(Various)
<a href="#">British English</a>		34 billion	British	1500s-2000	(Various)