

Language	Period	Acronym	Name	Size	Tagged or parsed	Contact persons	Webpage
Afrikaans	1880-1922		Corpus of Cape Dutch Correspondence	130,000 words		Ana Deumert	
Chinese	960-1911		Sheffield Corpus of Chinese	18,000 words		Xiaoling Hu	http://www.shef.ac.uk/hri/projects/projectpages/sccdls
Chinese, Taiwan			Sinica Historical Corpus				http://www.sinica.edu.tw/ftms-bin/kiwi1/mkiwi.sh?language=1
Czech	13th century-1989	DCNC-DIAKORP	Diachronic section of the Czech National Corpus	1,95 million words		Václav Cvrček, ICNC (Institute of the Czech National Corpus, Charles University, Prague)	http://ucnk.ff.cuni.cz/english/diakorp.php
Dutch	1950/51 & 2002	VU DNC	VU University Diachronic News Text Corpus		YES	Kirsten Vis et al.	http://www.clarin.nl/showcase/direct-quotation/440
Dutch	500-1976		Taalbank				http://gtb.inl.nl/
French	1968	ESLO	L'Enquête Socio-Linguistique à Orléans	c. 10,5 million words			http://eslo.tge-adonis.fr/
French	1998		Fr(leMonde)	1 million		Thierry Selva	
French	1100-1837	MCVF	Corpus MCVF	2,5 million words	YES	France Martineau, Université d' Ottawa	http://www.voies.uottawa.ca/corpus_pg_fr.html
French	1180-2009		FRANTEXT	271,619,272 words		CNRS & Université de Lorraine	http://www.atilf.fr/Les-ressources/Ressources-informatisees/FRANTEXT/
French	9th-13th century		Nouveau Corpus d' Amsterdam	3,3 million words	YES	Piet van Reenen (Free University of Amsterdam)	http://www.uni-stuttgart.de/lingrom/stein/corpus/
French	9th-15th century	BFM	Base de français médiéval	4,7 million words	YES	Laboratoire Analyses de Corpus de l'E.N.S. Lettres et Sciences Humaines, Lyon	http://weblex.ens-lh.fr/wlx/cgi/weblexf?corpus¼bfm
French			Beeching Corpus of Spoken French	150,000 words		Kate Beeching	http://www.uwe.ac.uk/hlss/llas/iclru/corpus.pdf
French		CRFP	Corpus de Référence du Français Parlé	440,000 words			http://www.up.univ-nrs.fr/delic/crfp
Galician	Middle Ages-present		Diachronic Stemmed Corpus and Dictionary of Galician	5,601,290 words		Nieves R. Brisaboa et al. (Universidade da Coruña)	http://gramatica.usc.es/redecorpus/?lang=en
German	1210-1930	TüBa-D/Z	Tübinger Baubank des Deutschen/Diachrones Corpus	252.520.365 words			
German	1650-1900	GerManC	A representative historical corpus of German	c. 800,000 words		Martin Durrell	http://ota.ahds.ac.uk/desc/2544
German	1650-1900	DTA	Kernkorpus des Deutschen Textarchivs	in development			http://www.dwds.de/ressourcen/koorpora/

German	16th-19th century		Register in Diachronic German Science				http://korpling.german.hu-berlin.de/ridges/index_en.html
German	1920-1939		Juilland-D-Korpus	500,000 words			http://www.dwds.de/ressourcen/korpora/
German	1946-2009		Die ZEIT	460 million words			http://www.dwds.de/ressourcen/korpora/
German	1949-1990		DDR-Korpus	9 million words			http://www.dwds.de/ressourcen/korpora/
German	1950-present	DEREKO	German Reference Corpus	4 billion words			http://corpora.ids-mannheim.de/ccdb/
German	1990s		DWDS supplementary corpus	900 million words			
German	1994-2004		Süddeutsche Zeitung	453 million words			http://www.dwds.de/ressourcen/korpora/
German	1994-2005		Berliner Zeitung	252 million words			http://www.dwds.de/ressourcen/korpora/
German	1996-2005		Der Tagesspiegel	170 million words			http://www.dwds.de/ressourcen/korpora/
German	1997-2006 & 1999-2006		BILD und WELT	121 & 240 million words			http://www.dwds.de/ressourcen/korpora/
German	2003-2005		Potsdamer Neueste Nachrichten	15 million words			http://www.dwds.de/ressourcen/korpora/
German	20th century		Kernkorpus des 20	100 million words			http://www.dwds.de/ressourcen/korpora/
German	20th century		Korpus Gesprochene Sprache	2,5 million words			http://www.dwds.de/ressourcen/korpora/
German	21st century		Kernkorpus des 21	in development			http://www.dwds.de/ressourcen/korpora/
German	750-12th century		Referenzkorpus Altdeutsch (Old German Reference Corpus)				http://www.deutschiachrondigital.de/home/?lang=en
German	750-1900	DDD	Deutsch Diachron Digital				http://www.dh.informatik.uni-erlangen.de/IMMD8/Services/sammlung_korpora/DDD.html
German	Old High-Middle High German	Kali	Korpusarbeit Linguistik	213,798 words in development		Gabriele Diewald	www.kali.uni-hannover.de
German		AGD	Archiv für Gesprochenes Deutsch-Datenbank gesprochenes Deutsch				http://dsav-wiss.ids-mannheim.de/
German			Wendekorpus	250,000 words		Norbert Dittmar	http://www.dwds.de/ressourcen/korpora/
Greek, Ancient	8th-5th c. BC	AGDT	Ancient Greek Dependency Treebank	350,000 words	YES	Gregory Crane	http://nlp.perseus.tufts.edu/syntax/treebank/greek.html
Icelandic	12th-20th century		Icelandic Parsed Historical Corpus	1,002,390 words			http://linguist.is/icelandic_treebank/Icelandic_Parsed_Historical_Corpus_%28IcePaHC%29

Italian	1861-1945	DiaCORIS	Diachronic Corpus di Italiano Scritto	in development		University of Bologna	http://corpora.ficlit.unibo.it/coris_eng.html
Italian	1965-present		Il Corpus Stammerjohann	86,000 words			http://lablita.dit.unifi.it/corpora/confro/onto/lablita/
Italian	Old Italian		Padua Corpus				
Italian	origins-1375	TLIO	Tesoro della Lingua Italiana delle Origini	23,178,540 words		Istituto Opera del Vocabolario Italiano, Firenze	http://www.vocabolario.org/
Japanese	10th-19th century		Diachronic Corpus of Japanese	in development		Yasuhiro Kondo (National Institute of Japanese Language and Linguistics)	http://www.ninjal.ac.jp/english/research/project/a/corpus/
Jewish	1887-1938		Korpus jüdischer Periodika	26 million words			http://www.dwds.de/ressourcen/korpora/
Latin	13th century	IT-TB	Index Thomisticus Treebank	230,000 words	YES	Marco Passarotti	http://itreebank.marginalia.it/
Latin	1st c. BC - 4th century AD	LDT	Latin Dependency Treebank	55,000 words	YES	Gregory Crane	http://nlp.perseus.tufts.edu/syntax/treebank/latin.html
Polish	15th century-1791	PoDi	Polish Diachronic Corpus	in development	YES	Arek Danszczyk, Björn Hansen, Thomas Menzel, and Roland Meyer (responsible for corpora)	http://rhssl1.uni-regensburg.de/SlavKo/news/polish-diachronic-corpus-poldi-going-online
Portuguese	1380-1845		Corpus Histórico do Português Tycho Brahe	2,547,503 words	YES		http://www.tycho.iel.unicamp.br/~tycho/corpus/en/index.html
Portuguese	1300-1900		O corpus do Português	45 million words	YES	Mark Davies & Michael Ferreira	http://www.corpusdoportugues.org/
Portuguese	9th-12th century (Latin) & 12th-16th century (Portuguese)	CIPM	Corpus Informatizado do Português Medieval	2,053,582 words		Francisca Xavier (Universidade Nova de Lisboa)	http://cipm.fcsh.unl.pt/
Portuguese, Brazilian	1755-1832	PHPB-Ba	Corpus Compartilhado Diacrônico – Outros manuscritos				http://www.tycho.iel.unicamp.br/ce_dohs/corpora_ot.html
Portuguese, Brazilian	1809-2000	PHPB-Ba	Corpus Compartilhado Diacrônico – Cartas Brasileiras				http://www.tycho.iel.unicamp.br/ce_dohs/corpora.html
Portuguese, Brazilian	18th-20th century		Corpus Compartilhado Diacrônico: cartas pessoais brasileiras			Laboratório de História do Português Brasileiro, Universidade Federal do Rio de Janeiro	http://www.letras.ufrj.br/laborhistorico/
Portuguese, Brazilian	1901-2006	PHPB-Ba	Corpus Compartilhado Diacrônico - Impresso				http://www.tycho.iel.unicamp.br/ce_dohs/corpora_i.html

Portuguese, Brazilian	1990-1999	PHPB-Ba	Corpus Compartilhado Diacrônico - Oral				http://www.tycho.iel.unicamp.br/ce-dohs/corpora_o.html
Portuguese, Brazilian		CE-DOHS	Corpus Eletrônico de Documentos Históricos do Sertão			Zenaide de Oliveira Novais Carneiro & Mariana Fagundes de Oliveira, Universidade Estadual de Feira de Santana/UEFS, Brasil	http://www2.uefs.br/cedohs/
Roviana			Corpus of Early Roviana			Western Province Government of the Solomon Islands (1991)	
Russian			The Regensburg Diachronic Corpus of Russian			Roland Meyer	http://rhssl1.uni-regensburg.de/SlavKo/korpus/rrudi-new
Serbian	1848-1933, 1937 (Lower Sorbian)	DOTKO	DOInoserbski Tekstowy Korpus	12 million words		Sorbian Institute	http://www.dolnoserbski.de/korpus/
Slovene	1750-1918	IMP	The IMP language resources for historical Slovene	15 million words	YES	Tomaž Erjavec	http://nl.ijs.si/imp/index-en.html
Spanish	1975-2004	CREA	Corpus de referencia del español actual	160 million words		Real Academia Española	http://www.rae.es/recursos/banco-de-datos/crea
Spanish	origins-1974	CORDE	Corpus Diacrónico del Español	260 million words		Real Academia Española	http://www.rae.es/recursos/banco-de-datos/corde
Spanish	2001-2012	CORPES XXI	Corpus del español del siglo xxi	300 million words, in development		Real Academia Española	http://www.rae.es/recursos/banco-de-datos/corpes
Spanish			Google books corpus	45 billion words		Mark Davies	http://googlebooks.byu.edu/x.asp
Spanish, Chile	16th-17th century	CorDECh	Corpus Diacrónico del Español de Chile	in development		Manuel Contreras Seitz (Universidad Austral de Chile)	
Sumerian	3200 BC-85 BC		Diachronic Corpus of Sumerian Literature	in development		Jeremy Black (Oxford University)	http://dcs1.orinst.ox.ac.uk/
Welsh	1500-1850		Historical Corpus of Welsh Language			David Willis (University of Cambridge)	http://people.ds.cam.ac.uk/dwew2/hcwl/menu.htm
	Miscellanea						
Old Indo-European		PROIEL	Pragmatic Resources in Old Indo-European Languages			Dag Haug, University of Oslo	http://www.hf.uio.no/ifiikk/english/research/projects/proiel/