

Corpus Computer Corpora

English Department, University of Berne

last updated: 22.09.2014

Abbr.	Corpus	type	parsed/tagged	variety	manual
19th Century English	19th-Century English Corpus	historical	no	British	no manual
CEECs	Corpus of Early English Correspondence (CEECs)	historical	no	British	
HELSINKI	The Helsinki Corpus of English Text: Diachronic Part	historical	no	British	manual
INNSBRUCK	Innsbruck Computer-Archive of Machine-Readable English Texts (ICAMET)	historical			
LAMPETER	Lampeter Corpus	historical	no	British	
NEWDIGATE	The Newdigate Newsletters	historical	no	England	
OLD SCOT	The Helsinki Corpus of Older Scots	historical	no	Scotland	
BERN	Bern	spoken			
Blairgowrie, Perthshire	Blairgowrie, Perthshire	spoken			
Bridlington, Yorkshire	Bridlington, Yorkshire	spoken		British	
Bristol	Bristol	spoken			
Coggeshall, Essex	Coggeshall, Essex	spoken		British	
Colchester	Colchester	spoken		British	
COLT	Corpus of London Teenage Language (COLT)	spoken	no	London	manual
COLTTAG	Corpus of London Teenage Language (COLT)	spoken	yes	London	manual
Deepings, Lincolnshire	Deepings, Lincolnshire	spoken			
Falkland Islands	Falkland Islands	spoken	no	Falkland Islands	
Folkestone, Kent	Folkestone, Kent	spoken			
Holbrook, Rural Ipswich	Holbrook, Rural Ipswich	spoken		British	
Ipswich	Ipswich	spoken		British	
LLC	London-Lund Corpus	spoken	no	British	no manual
Lurgan, Northern Ireland	Lurgan, Northern Ireland	spoken			
Mansfield, Nottinghamshire	Mansfield, Nottinghamshire	spoken			
Nailsea, Somerset	Nailsea, Somerset	spoken			
Northhamptonshire	Northamptonshire	spoken			
POW	Polytechnic of Wales Corpus	spoken	yes	South Wales	no manual
Rhode Island	Rhode Island	spoken			
Rhonda Valleys	Rhonda Valleys	spoken			
Romford, Essex	Romford, Essex	spoken		British	
SEC	Lancaster/IBM Spoken English Corpus (SEC)	spoken	no	British	manual
Silver End, Essex	Silver End, Essex	spoken		British	
WSC	The Wellington Corpus of Spoken New Zealand English (WSC)	spoken	no	New Zealand	manual
ACE	Australian Corpus of English (ACE)	written	no	Australia	manual
BROWN1	Brown 1	written	no	American	manual
BROWN2	Brown 2	written			
BROWNTAG	Brown Corpus (tagged)	written	yes	American	manual
FLOB	Freiburg-LOB (FLOB)	written	no	British	manual
FROWN	Freiburg-Brown (FROWN)	written	no	British	manual
KOHLAPUR	Kolhapur Corpus	written	no	India	manual
LOB	LOB Corpus	written	no	British	manual
LOBTAG	LOB Corpus (tagged)	written	yes	British	manual
LPC	Lancaster Parsed Corpus	written	yes		
WC	Wellington Corpus	written	no	New Zealand	manual
WCTAG	Wellington Corpus (tagged)	written	yes	New Zealand	manual
WWC	Wellington Corpus of Written New Zealand English	written		New Zealand	manual
BNC	British National Corpus	written/spoken			
ICE_EA	ICE East-African component	written/spoken		Tanzania/Kenia	

The corpora and manuals are available on the Corpus Computer. Please contact your project supervisor for the student account password.