List of corpora and databases

Links accessed in April 2012.

A

ARCHER = A Representative Corpus of Historical English Registers, version 3.1. 1990–93, 2002, 2007, 2010. Compiled under the supervision of Douglas Biber and Edward Finegan at Northern Arizona University, University of Southern California, University of Freiburg, University of Heidelberg, University of Helsinki, Uppsala University, University of Michigan, University of Manchester, Lancaster University, University of Bamberg, University of Zurich, University of Trier, University of Salford, and University of Santiago de Compostela. http://www.llc.manchester.ac.uk/research/projects/archer/.

B


CED = A Corpus of English Dialogues 1560–1760. 2006. Compiled under the supervision of Merja Kyttö (Uppsala University) and Jonathan Culpeper (Lancaster University).


CEEM = Corpus of Early English Medical Writing. In progress. Compiled under the supervision of Irma Taavitsainen and Päivi Pahta. University of Helsinki. See MEMT and EMEMT.


CONCE = A Corpus of Nineteenth-Century English. 2000. Compiled by Merja Kytö (Uppsala University) and Juhani Rudanko (University of Tampere).


E


EMEMT = Early Modern English Medical Texts. 2010. Compiled by Irma Taavitsainen (University of Helsinki), Päivi Pahta (University of Tampere), Martti Mäkinen (Svenska handelshögskolan), Turo Hiltunen, Ville Marttila, Maura Ratia, Carla Suhr, and Jukka Tyrkkö (University of Helsinki).


F

by Christian Mair (Albert Ludwigs-Universität Freiburg) and Geoffrey Leech (University of Lancaster). http://www.helsinki.fi/varieng/CoRD/corpora/FLOB/


G


H


I


LAMSAS = Linguistic Atlas of the Middle and South Atlantic States. 
http://us.english.uga.edu/lamsas/.


Lion = Literature Online. http://lion.chadwyck.co.uk/.


LSWE = The Longman Spoken and Written English Corpus. 
http://www.pearsonlongman.com/dictionaries/corpus/.

http://www.uis.no/research/culture/the_middle_english_grammar_project/meg-c/.

MEMT = Middle English Medical Texts. 2005. Compiled by Irma Taavitsainen (University of Helsinki), Päivi Pahta (University of Tampere), and Martti Mäkinen (University of Stavanger). http://www.helsinki.fi/varieng/CoRD/corpora/CEEM/MEMTindex.html.
NECTE = The Newcastle Electronic Corpus of Tyneside English. 2005. Compiled by Karen Corrigan (Newcastle University), Hermann Moisl (Newcastle University), and Joan Beal (University of Sheffield). http://www.helsinki.fi/varieng/CoRD/corpora/NECTE/.


O


P


S


SCOTS = Scottish Corpus of Texts and Speech. 2007. Compiled by Professor John Corbett (Principal Investigator), Dr. Wendy Anderson (Research Assistant 2004–7), Dr. Fiona Douglas (Research Assistant 2001–3), Dave Beavan (Computing Manager), Professor Christian Kay, Jean Anderson, Dr. Jane Stuart-Smith, Louise Sweeney, Cerwyss O’Hare, and Flora Edmonds. Department of English Language, University of Glasgow. http://www.scottishcorpus.ac.uk.


T


W
WebCorp = The Web as Corpus. Created, operated, and maintained by the Research and Development Unit for English Studies, School of English, Birmingham City University. http://www.webcorp.org.uk.

Y

Z