Πρόγραμμα «Αριστεία»
«Διαχρονικό σώμα ελληνικών κειμένων του 20ού αιώνα/
Diachronic corpus of Greek of the 20th century»
Κ.Α. Έρευνας: 70/3/11920
Ακρωνύμιο: **Greek Corpus 20**
Κωδικός **«ΑΡΙΣΤΕΙΑ Ι» 2396**
Επιστημονικός υπεύθυνος: Διονύσης Γούτσος

Στο παρόν έγγραφο περιλαμβάνεται η βιβλιογραφική επισκόπηση σχετικά με τη δημιουργία διαχρονικών σωμάτων κειμένων που προβλέπεται στην Ενότητα Εργασίας 1: «Σχεδιασμός προγράμματος». Συγκεκριμένα, παρατίθενται:

Μέρος 1: Βάση βιβλιογραφικών παραπομπών στα αγγλικά, που διακρίνεται σε παραπομπές σε βιβλία, άρθρα σε περιοδικά, άρθρα σε πρακτικά συνεδρίων και άρθρα σε συλλογικούς τόμους

Μέρος 2: Βάση δεδομένων με προγράμματα διαχρονικών σωμάτων κειμένων στα αγγλικά, που διακρίνονται σε προγράμματα για την Αγγλική και προγράμματα για άλλες γλώσσες. (Στη βάση δεδομένων επισημαίνονται τα προγράμματα που είναι συναφέστερα με το σχεδιαζόμενο σώμα κειμένων της Ελληνικής).

Μέρος 1: Βάση βιβλιογραφικών παραπομπών στα αγγλικά

BOOKS

Author(s)	Year	Title	Publisher
Aarts, Bas, Close, Joanne Leech, Geoffrey & Sean Wallis (eds)	2013	*The Verb Phrase In English: Investigating Recent Language Change With Corpora*	Cambridge: Cambridge University Press
Bauer, Laurie	1994	*Watching English Change: An Introduction to the Study of Linguistic Change in Standard Englishes in the Twentieth Century*	London/New York: Longman
Beal, Joan C., Corrigan, Karen P. & Hermann L. Moisl (eds)	2007	*Creating and Digitizing Language Corpora. Volume 1: Synchronic Databases*	Basingstoke: Palgrave Macmillan
Beal, Joan C., Corrigan, Karen P. & Hermann L. Moisl (eds)	2007	*Creating and Digitizing Language Corpora. Volume 2: Diachronic Databases*	Basingstoke: Palgrave Macmillan
Breitbarth, Anne, Lucas, Christopher, Watts, Sheila & David Willis (eds)	2010	*Continuity and Change in Grammar*	Amsterdam/Philadelphia: John Benjamins
Brinton, Laurel J. (ed.)	2001	*Historical Linguistics 1999: Selected Papers from the 14th International Conference On Historical Linguistics, Vancouver, 9–13 August 1999*	Amsterdam/Philadelphia: John Benjamins
Busse, Ulrich & Axel Hóbler (eds)	2012	*Investigations into the Meta-Communicative Lexicon of English: A Contribution to Historical Pragmatics*	Amsterdam/Philadelphia: John Benjamins
Conboy, Martin	2010	*The Language of Newspapers: Socio-Historical Perspectives*	London/New York: Continuum
Detges, Ulrich & Richard Waltereit (eds)	2008	*The Paradox of Grammatical Change: Perspectives from Romance*	Amsterdam/Philadelphia: John Benjamins
Facchinetti, Roberta & Matti Rissanen (eds)	2006	*Corpus-based Studies of Diachronic English*	Bern: Peter Lang
Gries, Stefan Th., Wulff, Stefanie & Mark Davies (eds)	2010	*Corpus-linguistic Applications: Current Studies, New Directions*	Amsterdam/New York: Rodopi
Hernández-Campoy, Juan Manuel & Juan Camilo Conde-Silvestre (eds)	2012	*The Handbook of Historical Sociolinguistics*	Oxford: Blackwell
Jucker, Andreas H. & Irma Taavitsainen (eds)	2008	*Speech Acts in the History of English*	Amsterdam/Philadelphia: John Benjamins

Jucker, Andreas H., Schreier, Daniel & Marianne Hundt (eds)	2009	*Corpora: Pragmatics and Discourse. Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME 29). Ascona, Switzerland, 14-18 May 2008*	Amsterdam/New York: Rodopi
Kranich, Svenja	2010	*The Progressive in Modern English: A Corpus-Based Study of Grammaticalization and Related Changes*	Amsterdam/New York: Rodopi
Leech, Geoffrey, Hundt, Marianne, Mair, Christian & Nicholas Smith	2009	*Change in Contemporary English: A Grammatical Study*	Cambridge: Cambridge University Press
Lenker, Ursula & Anneli Meurman-Solin (eds)	2007	*Connectives in the History of English*	Amsterdam/Philadelphia: John Benjamins
Lindquist, Hans & Christian Mair (eds)	2004	*Corpus Approaches to Grammaticalization in English*	Amsterdam/Philadelphia: John Benjamins
Mair, Christian	2006	*Twentieth Century English: History, Variation and Standardization*	Cambridge: Cambridge University Press
McEnery, Tony & Andrew Hardie	2012	*Corpus Linguistics: Method, Theory and Practice*	Cambridge: Cambridge University Press
Nevalainen, Terttu & Elizabeth Closs Traugott (eds)	2012	*The Oxford Handbook of the History of English*	Oxford: Oxford University Press
Nevalainen, Terttu & Sanna-Kaisa Tanskanen (eds)	2007	*Letter Writing*	Amsterdam/Philadelphia: John Benjamins
Nevalainen, Terttu, Klemola, Juhani & Mikko Laitinen (eds)	2006	*Types of Variation: Diachronic, Dialectal and Typological Interfaces*	Amsterdam/Philadelphia: John Benjamins
Nevalainen, Terttu, Taavitsainen, Irma, Pahta, Päivi & Minna Korhonen (eds)	2008	*The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*	Amsterdam/Philadelphia: John Benjamins
Renouf, Antoinette & Andrew Kehoe	2009	*Corpus Linguistics: Refinements and Reassessments*	Amsterdam/New York: Rodopi
Rudanko, Juhani	2011	*Changes in Complementation in British and American English: Corpus-Based Studies on Non-Finite Complements in Recent English*	Basingstoke: Palgrave Macmillan
Trips, Carola	2009	*Lexical Semantics and diachronic morphology:The Development of -hood, -dom and -ship in the History of English*	Tübingen: Niemeyer
Wilson, Andrew, Archer, Dawn &	2006	*Corpus Linguistics Around the World*	Amsterdam/New York: Rodopi

Paul Rayson

JOURNAL ARTICLES

Author(s)	Year	Title	Journal	Volume-Issue	Pages
Baker, Paul	2011	Times may change, but we will always have money: Diachronic variation in recent British English	*Journal of English Linguistics*	39 (1)	65-88
Baker, Paul	2009	The BE06 Corpus of British English and recent language change	*International Journal of Corpus Linguistics*	14 (3)	312–337
Baker, Paul	2010	Will Ms ever be as frequent as Mr? A corpus-based comparison of gendered terms across four diachronic corpora of British English	*Gender and Language*	4 (1)	125–149
Clark, Caroline	2010	Evidence of evidentiality in the quality press 1993 and 2005	*Corpora*	5 (2)	139–160
Collins, Peter	2004	Let-imperatives in English	*International Journal of Corpus Linguistics*	9 (2)	299-319
Contreras Seitz, Manuel	2009	Hacia la Constitución de un Corpus Diacrónico del Español de Chile (Towards Chilean Spanish Language Diachronic Corpus)	*RLA. Revista De Lingüística Teórica y Aplicada*	47 (2)	111-134
Davies, Mark	2012	Expanding horizons in historical linguistics with the 400-million word Corpus of Historical American English	*Corpora*	7 (2)	121–157
Duguid, Alison	2010	Newspaper discourse informalisation: a diachronic comparison from keywords	*Corpora*	5 (2)	109–138
Duguid, Alison	2010	Investigating anti and some reflections on Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS)	*Corpora*	5 (2)	191–220
Dury, Pascaline	2004	Building a bilingual diachronic corpus of ecology: The long road to completion	*ICAME*	28	5-16
Eshkol-Taravella, Iris, Baude, Olivier, Maurel, Denis, Hriba, Linda, Dugua, Céline & Isabelle Tellier	2011	Un grand corpus oral «disponible»: le corpus d'Orléans 1968-2012	*TAL.*	53 (2)	17-46
Gabrielatos, Costas, Mc Enery	2012	The peaks and troughs of corpus-based contextual	*International Journal of Corpus*	37 (2)	151-175

Tony, Diggle, Peter J. & Paul Baker		analysis	*Linguistics*		
Gries, Stefan Th. & Martin Hilpert	2008	The identification of stages in diachronic data: Variability-based neighbour clustering	*Corpora*	3 (1)	59–81
Hardie, Andrew	2007	Part-of-speech ratios in English corpora	*International Journal of Corpus Linguistics*	12 (1)	55–81
Heremans, Katrien & Hubert Cuyckens	2012	DIACHRONEX: Corpus-based exercises for English diachronic linguistics	*ICAME*	36	67-93
Hilpert, Martin	2011	Dynamic visualizations of language change: Motion charts on the basis of bivariate and multivariate data from diachronic corpora	*International Journal of Corpus Linguistics*	16 (4)	435-461
Hilpert, Martin & Stefan Th. Gries	2009	Assessing frequency changes in multistage diachronic corpora: Applications for historical corpus linguistics and the study of language acquisition	*Literary and Linguistic Computing*	24 (4)	385-401
Hinrichs, Erhard & Thomas Zastrow	2012	Linguistic Annotations for a Diachronic Corpus of German	*Linguistic Issues in Language Technology*	7 (7)	1-16
Hu, Xiaoling, Williamson, Nigel & Jamie McLaughlin	2005	Sheffield Corpus of Chinese for Diachronic Linguistic Study	*Literary and Linguistic Computing*	20 (3)	281-293
Jacobson, Michel & Oliver Baude	2011	Corpus de la parole: Collecte, catalogage, conservation et diffusion des ressources orales sur le franηais et les langues de France	*TAL.*	52 (3)	47-69
Jenset, Gard B.	2013	Mapping meaning with distributional methods: A diachronic corpus-based study of existential there	submitted to *Journal of Historical Linguistics*		
Koteyko, Nelya	2007	A diachronic approach to meaning: English loanwords in Russian opposition discourse	*Corpora*	2 (1)	65–95
Kwary, Deny Arnos & Arum, Kirana Wuri Anjar	2011	Lincoln's vs. Obama's presidencies: A diachronic corpus-based analysis of the adjectival collocates of [man] and [woman] in the American English	*Revista Virtual de Estudos da Linguagem - ReVEL*	9 (17)	211-225
Leech, Geoffrey & Nicholas Smith	2005	Extending the possibilities of corpus-based research on English in the twentieth century: A prequel to LOB and FLOB1	*ICAME*	29	83-98

Lüdeling, Anke	2005	Heterogeneity and standardization in data use and annotation: A Diachronic Corpus of German	*Interdisciplinary Studies on Information Structure*	2	43-54
Macalister, John	2006	The Maori lexical presence in New Zealand English: Constructing a corpus for diachronic change	*Corpora*	1 (1)	85-98
Mair, Christian, Hundt, Marianne, Leech, Geoffrey & Nicholas Smith	2002	Short-term diachronic shifts in part-of-speech frequencies: A comparison of the tagged LOB and F-LOB corpora	*International Journal of Corpus Linguistics*	7 (2)	245-264

Marchi, Anna	2010	'The moral in the story': A diachronic investigation of lexicalised morality in the UK press	*Corpora*	5 (2)	161–189
Millar, Neil	2009	Modal verbs in TIME: Frequency changes 1923–2006	*International Journal of Corpus Linguistics*	14 (2)	191–220
Nokkonen, Soili	2006	The semantic variation of NEED TO in four recent British English corpora	*International Journal of Corpus Linguistics*	11 (1)	29–71
Partington, Alan	2010	Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK newspapers: An overview of the project	*Corpora*	5 (2)	83–108
Partington, Alan	2012	The changing discourses on antisemitism in the UK press from 1993 to 2009: A modern diachronic corpus-assisted discourse study	*Journal of Language and Politics*	11 (1)	51–76
Petrova, Svetlana, Solf, Michael, Ritz, Julia, Chiarcos, Christian & Amir Zeldes	2009	Building and using a richly annotated interlinear diachronic corpus: The case of Old High German Tatian	*TAL.*	50 (2)	47-71
Picton, Aurélie	2011	Picturing short-period diachronic phenomena in specialised corpora: A textual terminology description of the dynamics of knowledge in space technologies	*Terminology*	17 (1)	134–156
Rudanko, Juhani	2006	Watching English grammar change: A case study on complement selection in British and American English	*English Language and Linguistics*	10 (1)	31-48
Rudanko, Juhani	2010	Explaining grammatical variation and change: A case study of complementation in American English over three decades	*Journal of English Linguistics*	38 (1)	4–24
Taylor, Charlotte	2010	Science in the news: A diachronic perspective	*Corpora*	5 (2)	221–250
Vázquez, Nila, Esteban-Segura, Laura & Teresa Marqués-Aguado	2011	A descriptive approach to computerised English historical corpora in the 21st century	*International Journal of English Studies (IJES)*	11 (2)	119-139

ARTICLES IN PROCEEDINGS

Author(s)	Year	Title	Editor(s)	Volume title	Pages
Aguilar-Guevara, Ana, Aloni, Maria, Port, Angelika, Šimík, Radek, de Vos, Machteld & Zeijlstra, Hedde	2011	Semantics and pragmatics of indefinites: Methodology for a synchronic and diachronic corpus study	Stefanie Dipper & Heike Zinsmeister (eds)	*Bochumer Linguistische Arbeitsberichte*	
Calle-Martín, Javier, Marqués-Aguado, Teresa, Esteban-Segura, Laura & Miranda-García, Antonio	2012	The Reference Corpus of Late Middle English Scientific Prose		*Proceedings of KONVENS 2012 (LThist 2012 workshop), Vienna, September 21, 2012*	424-432
Degaetano-Ortlieb, Stefania, Lapshinova-Koltunski, Ekaterina & Elke Teich	2012	Feature discovery for diachronic register analysis: A semi-automatic approach		*LREC 2012*	
Dipper, Stefanie, Faulstich, Lukas, Leser, Ulf & Anke Lüdeling	2004	Challenges in modelling a richly annotated diachronic corpus of German		*Workshop on XML-based richly annotated corpora XBrac-Workshop, 29th May 2004 https://informatik.hu-berlin.de/forschung/gebiete/wbi/research/publications/2004/xbrac04_final.pdf*	
Gerow, Aaron & Khurshid, Ahmad	2012	Diachronic variation in grammatical relations		*Proceedings of COLING 2012: Posters Volume 1. COLING 2012, Mumbai, December 2012*	381–390
Hinrichs, Erhard & Zastrow, Thomas	2012	Automatic Annotation and Manual Evaluation of the Diachronic German Corpus TüBa-D/DC		*Proceedings of Language Resources Evaluation Conference (LREC) 21-27 May 2012 Istanbul*	1622-1627
Krause, Thomas, Lüdeling, Anke, Odebrecht, Carolin & Zeldes, Amir	2012	Multiple tokenizations in a diachronic corpus		*Paper presented at the conference Exploring Ancient Languages through Corpora, Oslo, 14-16/6/2012*	
Kučera, Karel	2007	Mapping the Time Continuum: A Major Raison D'être for Diachronic Corpora		*Proceedings of Corpus Linguistics 2007. Birmingham, UK.*	
Mayoral Hernández, Roberto & Alcázar, Asier	2008	A Diachronic Analysis of Frequency Adverbials: Variation in Peninsular and Latin American Spanish	Westmoreland, Maurice & Thomas, Juan Antonio (eds)	*Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics* Somerville, MA: Cascadilla	81-90
Onelli, C. Proietti, D. Seidenari C. & F. Tamburini	2006	The DiaCORIS project: a diachronic corpus of written Italian		*Proceedings of the 5th International Conference on Language Resources and Evaluation - LREC 2006. Genoa*	1212-1215

Sánchez-Marco, Cristina, Boleda, Gemma Fontana, Josep Maria & Domingo, Judith	2010	Annotation and Representation of a Diachronic Corpus of Spanish	*Proceedings of Language Resources and Evaluation (LREC), Malta, May 2010*	2713-2718
Sánchez-Marco, Cristina, Boleda, Gemma & Lluís Padró	2011	Extending the tool, or how to annotate historical language varieties	*Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage,Social Sciences, and Humanities, Portland, OR, USA, 24th June*	1-9
Scarpa, Federica	2003	Using An Italian Diachronic Corpus To Investigate The "Core" Patterns Of The Language Of Science	*Paper presented at the 14th European Symposium on Languages for Special Purposes "Communication, Culture, Knowledge". University of Surrey, 18-22 August 2003*	
Štajner, Sanja	2011	Towards a Better Exploitation of the Brown 'Family' Corpora in Diachronic Studies of British and American English Language Varieties	*Proceedings of the Student Research Workshop associated with RANLP 2011. Hissar, Bulgaria, 13 September 2011*	17-24
Štajner, Sanja & Mitkov, Ruslan	2012	Using comparable corpora to track diachronic and synchronic changes in lexical density and lexical richness	*The 5th Workshop on Building and Using Comparable Corpora. Vol. 17. No. 17*	88-97
Štajner, Sanja & Mitkov, Ruslan	2012	Diachronic Changes in Text Complexity in 20th Century English Language: An NLP Approach	*Proceedings of Language Resources Evaluation Conference (LREC) 21-27 May 2012 Istanbul*	1577-1584

ARTICLES IN COLLECTED VOLUMES

Author(s)	Year	Τίτλος	Editor(s)	Book title	Publisher	Pages
Asmussen, Jørg	2006	Towards a methodology for corpus-based studies of linguistic change: Contrastive observations and their possiblediachronic interpretations in the Korpus 2000 and Korpus 90 General Corpora of Danish	Wilson, Andrew, Archer, Dawn & Paul Rayson (eds)	*Corpus Linguistics Around the World*	Amsterdam/ New York: Rodopi	33-48
Davies, Mark	2011	Synchronic and diachronic uses of corpora	Viana, Vander, Zyngier, Sonia & Geoff Barnbrook (eds)	*Perspectives on Corpus Linguistics*	Amsterdam/ Philadelphia: Benjamins	63–80
Davies, Mark	2012	Some methodological issues related to corpus-based investigations of recent syntactic changes in English	Nevalainen, Terttu & Traugott, Elizabeth Closs (eds)	*The Oxford Handbook of the History of English*	Oxford: Oxford University Press	
Degani, Marta	2009	Re-analysing the semi-modal ought to: An investigation of its use in the LOB, FLOB, Brown and Frown corpora	Renouf, Antoinette & Andrew Kehoe (eds)	*Corpus Linguistics: Refinements and Reassessments*	Amsterdam/ New York: Rodopi	327-346
Eckkrammer, Eva Martha	2004	A diachronic genre corpus: Problems and findings from the DIALAYMED-Corpus.	Barnbrook, Geoff, Danielsson, Pernilla, Mahlberg, Michaela (eds)	*Meaningful Texts: The Extraction of Semantic Information from Monolingual and Multilingual Corpora*	London/New York: Continuum	17-30
Facchinetti, Roberta & Rissanen, Matti	2006	Introduction	Facchinetti, Roberta & Rissanen, Matti (eds)	*Corpus-based Studies of Diachronic English*	Peter Lang	7-14
Hundt, Marianne & Geoffrey Leech	2012	"Small is beautiful": On the value of standard reference corpora for observing recent grammatical change	Nevalainen, Terttu & Traugott, Elizabeth Closs (eds)	*The Oxford Handbook of the History of English*		
Kehoe, Andrew	2006	Diachronic linguistic analysis on the web with WebCorp	Renouf, Antoinette & Kehoe, Andrew (eds)	*The Changing Face of Corpus Linguistics*	Amsterdam/ New York: Rodopi	297-307

Kehoe, Andrew & Gee, Matt	2009	Weaving web data into a diachronic corpus patchwork	Renouf, Antoinette & Andrew Kehoe (eds)	*Corpus Linguistics: Refinements and Reassessments*	Amsterdam/ New York: Rodopi	255-279
Kohnen, Thomas	2007	From Helsinki through the centuries: The design and development of English diachronic corpora	Päivi Pahta, Irma Taavitsainen, Terttu Nevalainen & Jukka Tyrkkö	*Studies in Variation, Contacts and Change in English 2: Towards Multimedia in Corpus Studies*	VARIENG	
Krishnamurthy, Ramesh	2003	Freeze-frame pictures: Micro-diachronic variations in synchronic corpora	Andor, Jozsef, Horvath, Jozsef and Marianne Nikolov (eds)	*Studies in English Theoretical and Applied Linguistics.*	Lingua Franca Csoport, Pecs, Hungary	15-31
Leech, Geoffrey & Smith, Nicholas	2006	Recent grammatical change in written English 1961-1992: Some preliminary findings of a comparison of American with British English	Renouf, Antoinette & Kehoe, Andrew (eds)	*The Changing Face of Corpus Linguistics*	Amsterdam/ New York: Rodopi	185-204
Mair, Christian	2008	Corpora and the study of recent change in language	Anke Lüdeling & Kytö, Merja	*Corpus Linguistics. An International Handbook. Volume II*	Berlin: Mouton de Gruyter	
Marchi, Anna & Taylor, Charlotte	2009	Establishing the EU: The representation of Europe in the press in 1993 and 2005	Jucker, Andreas H., Schreier, Daniel & Marianne Hundt (eds)	*Corpora: Pragmatics and Discourse, Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME 29).*	Amsterdam/ New York: Rodopi	203-226
Meyer, Roland	2005	The Regensburg Diachronic Corpus of Russian: A new source for linguistic research (not only) on modality	Hansen, Björn & Karlík, Petr (eds)	*Modality in Slavonic Languages. New Perspectives.*	München: Sagner	315-336
Nesselhauf, Nadja	2007	Diachronic analysis with the internet? Will and shall in ARCHER and in a corpus of e-texts from the web	Hundt, Marianne, Nesselhauf, Nadja & Carolin Biewer (eds)	*Corpus Linguistics and the Web*	Amsterdam/ New York: Rodopi	287-305
Scarano, Antonietta & Signorini, Sabrina	2005	Corpus linguistics and diachronic variability. A study on Italian spoken language corpora from the 1960s until nowadays	Pusch, Claus D., Kabatek, Johannes & Raible, Wolfgang (eds)	*Romanistische Korpuslingustik II: Korpora und diachrone Sprachwissenschaft/Romance Corpus Linguistics II: Corpora and*	Tübingen: Gunter Narr Verlag	191-202

Μέρος 2: Βάση δεδομένων με προγράμματα διαχρονικών σωμάτων κειμένων

ENGLISH

Language	Period	Acronym	Name	Size (in words)	Tagged or parsed	Contact persons	Webpage
English	Old English		Electronic Beowulf	17,327			http://www.uky.edu/~kiernan/BL/kportico.html
English	Old English		The Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English	106,210	YES	Susan Pintzuk (University of York) et al.	http://www-users.york.ac.uk/~sp20/corpus.html
English	Old English		The Canterbury Tales Project			Oxford, Sheffield, Brigham Young	http://www.shef.ac.uk/uni/projects/ctp/
English	Old English		The Dictionary of Old English Web Corpus				http://www.doe.utoronto.ca/pub/webcorpus.html
English	Old English		The York-Helsinki Parsed Corpus of Old English Poetry	71,490	YES	Susan Pintzuk et al.	http://www-users.york.ac.uk/~lang18/pcorpus.html
English	Old English	YCOE	The York-Toronto-Helsinki Parsed Corpus of Old English Prose		YES	Ann Taylor, Anthony Warner, Susan Pintzuk, Frank Beths	http://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm
English & Latin	450-1100		The Dictionary of Old English Corpus in Electronic Form			Antonette di Paolo Healey	http://www.press.umich.edu/webhome/healey/siteform.html
English	Old English (-1300)	CoNE	A Corpus of Narrative Etymologies from primitive Old English to early Middle English			Heinz Giegerich (University of Edinburgh)	http://www-test.research.ed.ac.uk/portal/en/projects/a-corpus-of-narrative-etymologies-from-primitve-old-english-to-early-middle-english%284c764637-08a7-40be-a2fd-853ddd378784%29.html
English, Northern	600-1590	SCONE	Seville Corpus of Northern English			Julia Fernández-Cuesta & Gabriel Amores (University of Sevilla)	http://www.helsinki.fi/varieng/CoRD/corpora/SCONE/index.html
English	730-1710		The Helsinki Corpus of English Texts	1,572,800		Matti Rissanen	http://www.helsinki.fi/varieng/CoRD/corpora/Helsinki Corpus/
English	Old & Middle (950-1991?)	LEON	Leuven English Old to New	in preparation (12,6 million)		Peter Petré	http://wwwling.arts.kuleuven.be/fll/ppetre/index.htm
English	Middle English		Corpus of Early Middle English Tagged Texts and Maps			M.Laing@ed.ac.uk (Institute for Historical Dialectology, Edinburgh)	
English	Middle English		Middle English Collection			University of Virginia	http://etext.virginia.edu/mideng.browse.html
English	Middle English		Middle English Compendium			University of Michigan Digital Library Production Service	http://www.hti.umich.edu/mec/

English	1027-1613		**Corpus of MED quotations**	12 million		Peter Petré	https://perswww.kuleuven.be/~u0050685/Corpus_of_MED_quotations.htm
English	1100-1500		**Corpus of Middle English Prose and Verse**			University of Michigan	http://quod.lib.umich.edu/c/cme/
English	1100-1500		**Innsbruck Computer Archive of Middle English Texts**	5,5 million		Manfred Markus	
English	1150-1500	PPCME2	**Penn-Helsinki Parsed Corpus of Middle English**	c. 1,2 million	YES	Anthony Kroch	http://www.ling.upenn.edu/hist-corpora
English	1150-18th century	COREP	**Corpus of English Religious Prose**			Thomas Kohnen	http://www.helsinki.fi/varieng/CoRD/corpora/COERP/index.html
English	1272-1509		**The Parliament Rolls of Medieval England**	8 million			http://www.sd-editions.com/PROME/home.html
English	1300-1500		**Reference Corpus of Late Middle English Scientific Prose**	1,2 million	YES	Antonio Miranda-García	http://prezi.com/-m1rrwyrxvky/malaga-corpus-of-late-middle-english-scientific-prose/
English	1350-1500		**The Middle English Grammar Corpus**			Merja Stenroos (University of Stavanger)	http://www.helsinki.fi/varieng/CoRD/corpora/MEG-C/index.html
English	1350-1500		**The Middle English Scribal Texts Programme - The Middle English Grammar Corpus (MEG-C)**	664,543		Merja Stenroos (University of Stavanger)	www.uis.no/mest
English	1375-1800	CEEM	**Corpus of Early English Medical Writing**	c. 3,75 million		Irma Taavitsainen & Paivi Pahta	http://www.helsinki.fi/varieng/CoRD/corpora/CEEM/
English	1375-1850	CoER	**Corpus of Early English Recipes**	1,5 million		Francisco Alonso Almeida et al.	http://www.gi.ulpgc.es/tell/page2/coer/coer.html
English, Scottish	1380-1500		**Edinburgh Corpus of Older Scots**	in preparation		I.K.Williamson@ed.ac.uk	http://www.lel.ed.ac.uk/research/ihd/
English	1386-1698	ICAMET	**Innsbruck Corpus of Archive of Machine-Readable English Texts**	7,8 million & 182,000		Manfred Markus (Universitat Innsbruck)	http://www.uibk.ac.at/anglistik/projects/icamet/
English	1402-1663	CEECSu	**Corpus of Early English Correspondence Supplement**	c. 0,44 million		Terttu Nevalainen	http://www.helsinki.fi/varieng/domains/CEEC.html
English	1403-1681	CEEC	**Corpus of Early English Correspondence**	2,7 million		Terttu Nevalainen	http://www.helsinki.fi/varieng/domains/CEEC.html
English	1450-1700	HCOS	**Helsinki Corpus of Older Scots**	c. 834,200		Anneli Meurman-Solin	http://www.helsinki.fi/varieng/CoRD/corpora/HCOS/index.html
English, Irish	14th century-present		**The Corpus of Irish English**	in preparation		Raymond Hickey	http://www.uni-due.de/IERC/CIE.htm
English	Early Modern	KEMPE	**Corpus of Early Modern Playtexts in English**	8,9 million	YES	Lene Petersen (University of the West of England)	http://corp.hum.sdu.dk
English	Early Modern		**Early Modern English Dictionaries**			Toronto	http://www.chass.utoronto.ca/cgi/english/emed

			Database				
English	Early Modern		**Michigan Early Modern English Materials**			Richard W. Bailey, Jay L. Robinson, James W. Downer, Patricia V. Lehman	http://etext.lib.virginia.edu/memem.query.html
English	1500-1710		**Penn-Helsinki Parsed Corpus of Early Modern English**				http://www.ling.upenn.edu/histcorpora/PPCEME-RELEASE-2/index.html
English, Scottish	1500-1715	CSC	**Corpus of Scottish Correspondence**	256,300		Anneli Meurman-Solin	http://www.helsinki.fi/varieng/CoRD/corpora/CSC/index.html
English, dialects	1500-1950		**The Salamanca Corpus**	6,748,873		García-Bermejo Giner	http://salamancacorpus.usal.es/SC/index.html
English, British, American	1500-2000		**Google books corpus**	155 billion & 34 billion	YES	Mark Davies	http://googlebooks.byu.edu/x.asp
English	1500-present		**Modern English Collection**			University of Michigan	http://www.hti.umich.edu/english/pd-modeng/
English	1545-2010		**The Small Corpus of Political Speeches**	845,176		Jukka Tyrkkö	http://www.helsinki.fi/varieng/CoRD/corpora/SCPS/index.html
English	1560-1760		**A Corpus of English Dialogues**	1,183,690		Merja Kytö & Jonathan Culpeper	http://www.helsinki.fi/varieng/CoRD/corpora/CED/index.html
English, American	1600-1700		**The Corpus of Early American English**			Merja Kytö	
English, British, American	1600-1999	ARCHER	**A Representative Corpus of Historical English Registers**	3,298,080		Douglas Biber & Edward Finegan	http://www.anglistik.uni-freiburg.de/seminar/abteilungen/sprachwissenschaft/ls_mair/research/projects/archer
English	1640-1717	CEMET	**The Corpus of Early Modern English Texts**	2 million		Hendrik De Smet	
English	1640-1740		**Lampeter Corpus of Early Modern English Tracts**	1,193,385		Josef Schmied & Eva Hertel	http://khnt.hit.uib.no/icame/manuals/LAMPETER/LAMPHOME.HTM
English	1641-1661	FEEN	**Florence Early English Newspapers**			Nicholas Brownlees & Francesca Benucci	
English	1661-1791		**Zurich English Newspaper Corpus**	1,6 million		Udo Fries & Hans Martin Lehmann	http://es-zen.unizh.ch
English	1673-1715		**Newdigate Newsletters**	c. 750,000		Philip Hines	http://khnt.hit.uib.no/icame/manuals/NEWDIGAT/INDEX.HTM
English	1680-1780		**Century of English Prose**			Louis Milic (Cleveland State University)	r0097@vmcms.csuohio.edu
English	1680-1780		**The Century of Prose Corpus**			Louis Milic (Cleveland State University)	
English	1681-1800	CEECE	**Corpus of Early English Correspondence Extension**	c. 2,2 million		Terttu Nevalainen	http://www.helsinki.fi/varieng/domains/CEEC.html

English	1700-1879	COLMOBA ENG	**A Corpus of Late Modern British and American English Prose**	1,17 million		Teresa Fanego	http://www.usc-teresafanego.es/COLMOBAENG.pdf
English	1700-1900	CC	**Coruña Corpus of English Scientific Writing**	in preparation		MuStE, University of A Coruña	http://www.udc.es/grupos/muste/
English	1700-1900	CETA	**Corpus of English Texts on Astronomy** (subcorpus of CC)	10,000 words		MuStE, University of A Coruña	
English	1700-1900	CEPhiT	**Corpus of English Philosophy Texts** (subcorpus of CC)	under publication		MuStE, University of A Coruña	
English	1700-1900	CELisT	**Corpus of English Life Sciences Texts** (subcorpus of CC)	in preparation		MuStE, University of A Coruña	
English	1700-1900	CHET	**Corpus of Historical English Texts** (subcorpus of CC)	in preparation		MuStE, University of A Coruña	
English	1700-1900	CECheT	**Corpus of English Chemistry Texts** (subcorpus of CC)	In preparation		MuStE, University of A Coruña	

Language	Period	Abbreviation	Name	Size	Parsed	Compiler	URL
English	1700-1930		**Women Scientists**	in preparation		MuStE, University of A Coruña	http://womenscientistsudc.com/
English	1700-1914	PPCMBE	**The Penn-Helsinki Parsed Corpus of Modern British English**	948,895	YES	Anthony Kroch, Beatrice Santorini & Ariel Diertani	http://www.ling.upenn.edu/hist-corpora/PPCMBE-RELEASE-1/index.html
English, Scottish	1700-1945	CMSW	**Corpus of Modern Scottish Writing**	5,5 million		John Corbett	http://www.scottishcorpus.ac.uk/cmsw/
English	1700-2000	RNC	**Rostock Corpus of English Newspapers**			Kristina Schneider, birte.boes@uni-rostock.de	
English	1710-1920	CLMETEV	**Corpus of Late Modern English Texts (extended version)**	14,970,622		Hendrik De Smet	https://perswww.kuleuven.be/~u0044428/clmetev.htm
English	1710-1920	CLMET 3.0	**Corpus of Late Modern English Texts, version 3.0**	34,386,225		Hendrik De Smet	https://perswww.kuleuven.be/~u0044428/clmet3_0.htm
English	1720-1913		**The Old Bailey Corpus**	c. 14 million		Magnus Huber (Universität Giessen)	http://www.uni-giessen.de/oldbaileycorpus/
English	1750-1850	LModE-OBC	**Late Modern English-Old Bailey Corpus**	1,008,234 words	YES	Ljubica Leone (University of Salerno)	
English	1761-1790		**A Corpus of late 18c Prose**	300,000		David Denison	http://personalpages.manchester.ac.uk/staff/david.denison/late18c
English, Canadian	1776-1899	CONTE	**Corpus of Early Ontario English**				http://tlioweb.ovi.cnr.it/(S(0lum44i02zylaz55f0ro5ev2))/CatForm01.aspx
English	1788-1900	COOEE	**Corpus of Oz Early English**	2 million		Clemens Fritz	http://www.helsinki.fi/varieng/CoRD/corpora/COOEE/index.html
English	18th century		**18th Century English letter corpus**				http://www.helsinki.fi/varieng/journal/volumes/10/voutilainen/
English	18th century	NEET	**Network of Eighteenth century English Texts**			Susan Fitzmaurice	
English	1810-2009	COHA	**Corpus of Historical American English**	400 million	YES	Mark Davies	http://corpus.byu.edu/coha/
English	1861-1919		**A Corpus of late Modern English Prose**	100,000		David Denison	http://personalpages.manchester.ac.uk/staff/david.denison/LModE_Prose.html
English	1861-1919		**Corpus of Late Modern English Prose**	100,000		David Denison	
English	1881-1922	CEN	**Corpus of English Novels**	26,227,428		Hendrik De Smet	https://perswww.kuleuven.be/~u0044428/cen.htm
English, Scottish	19th century		**A Corpus of Nineteenth-century Scottish Correspondence**	500,000 (in preparation)		Marina Dossena & Richard Dury	
English	19th century	CONCE	**The Corpus of Nineteenth Century English**	1,030,409		Merja Kytö, Juhani Rudanko, Erik Smitterberg (Uppsala University & University of Tampere)	
English	19th century	CNNE	**The Corpus of Nineteenth-Century Newspaper English**	in preparation		Erik Smitterberg	http://www.engelska.uu.se/Research/English_Language/Research_Areas/Electronic_Resource_Projects/

	20th century						
English	1923-2006		**TIME Magazine Corpus**	100 million	YES	Mark Davies	http://corpus.byu.edu/time/
English	1928-1934	BLOB-1931	**BLOB-1931**	c. 1 million		Geoffrey Leech & Paul Rayson	http://www.helsinki.fi/varieng/CoRD/corpora/BLOB-1931/index.html
English, Scottish	1945-2007	SCOTS	**Scottish Corpus of Texts & Speech**	4 million		John Corbett	www.scottishcorpus.ac.uk.
English	1950-1990	DCPSE	**The Diachronic Corpus of Present-Day Spoken English**	c. 800,000		Bas Aarts	http://www.ucl.ac.uk/english-usage/projects/dcpse/
English	1953-1987	LLC	**The London-Lund Corpus of Spoken English**	500,000		Jan Svartvik	http://www.helsinki.fi/varieng/CoRD/corpora/LLC/index.html
English, Australian	1960-1986	ACE	**Australian Corpus of English**	1 million		Pam Peters, Peter Collins & David Blair	http://www.ausnc.org.au/corpora/ace
English, Northern	1960-present		**The Diachronic Electronic Corpus of Tyneside English**	804,266		Karen Corrigan	http://research.ncl.ac.uk/decte/
English, Northern	late 1960s & 1994	NECTE	**Newcastle Corpus of Tyneside English**		YES		http://www.ncl.ac.uk/necte/
English	1968-2000	FRED	**Freiburg Corpus of English Dialects**	2,493,645		Bernd Kortmann	http://www2.anglistik.uni-freiburg.de/institut/lskortmann/FRED/
English	1970/1980	HARES-CAM	**Helsinki Archive of Regional English Speech–Cambridgeshire sampler**			Anna-Liisa Vasko	http://www.helsinki.fi/varieng/CoRD/corpora/HARES/index.html
English	1970/1980	HD	**Helsinki Corpus of British English Dialects**	1,008,641		Ossi Ihalainen, Kirsti Peitsara & Anna-Liisa Vasko	http://www.helsinki.fi/varieng/CoRD/corpora/Dialects/index.html
English	1990-present	COCA	**The Corpus of Contemporary American English**	450 million	YES	Mark Davies	http://corpus.byu.edu/coca/
English	20th century	BNC	**British National Corpus**	100 million			http://CORPUS.BYU.EDU/BNC
English, American	20th century	Frown	**The Freiburg-Brown Corpus of American English**	1 million		Christian Mair	http://www.helsinki.fi/varieng/CoRD/corpora/FROWN/index.html
English, British	20th century	F-LOB	**The Freiburg-LOB Corpus of British English**	1 million		Christian Mair	http://www.helsinki.fi/varieng/CoRD/corpora/FLOB/index.html
English, Indian	20th century		**The Kolhapur Corpus of Indian English**			S. V. Shastri	http://khnt.hit.uib.no/icame/manuals/kolhapur/INDEX.HTM#sourc
	21st century						
English	2003-2008	BE06	**BE06 Corpus**	1 million		Paul Baker	http://www.ling.lancs.ac.uk/profiles/Paul-Baker/

	Miscellanea						
English		DIA	DIA View				http://www.scottishcorpus.ac.uk/corpus/diaview/
English, American			Early American English Corpus	in preparation		Merja Kytö	
English			Leverhulme Corpus Project (Lancaster 1901 corpus)				
English, American			Santa Barbara Corpus of American Spoken English	in preparation (249,000)			http://www.linguistics.ucsb.edu/research/santa-barbara-corpus#Intro
English			The Lancaster Newsbooks Corpus				

OTHER LANGUAGES

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, Universite 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?corpus ¼ bfm
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

Language	Period	Abbreviation	Name	Size		Author	Link
German	1210-1930	TüBa-D/Z	**Tübinger Baumbank 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/korpora/
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.deutschdiachrondigital.de/home/?lang=en
German	750-1900	DDD	**Deutsch Diachron Digital**				http://wwwdh.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

Language	Period	Abbr.	Name	Size	Parsed	Author/Institution	URL
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/confronto/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	PolDi	**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/cedohs/corpora_ot.html
Portuguese, Brazilian	1809-2000	PHPB-Ba	Corpus Compartilhado Diacrônico – Cartas Brasileiras				http://www.tycho.iel.unicamp.br/cedohs/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/cedohs/corpora_i.html

Language	Period	Abbr.	Title	Size	Free	Author/Institution	URL
Portuguese, Brazilian	1990-1999	PHPB-Ba	**Corpus Compartilhado Diacrônico - Oral**				http://www.tycho.iel.unicamp.br/cedohs/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	**DOlnoserbski 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://dcsl.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/ifikk/english/research/projects/proiel/