

# Korpora und Textdatenbanken

- Französische Korpora und Textdatenbanken
  - Actes royaux de Poitou (1302-1464)
  - Altfranzösische Urkunden
  - Anglo-Norman Source Texts
  - Base de Français Médiéval
  - Bibliothèque bleue de Troyes
  - Ciel-F Corpus international écologique de la langue française
  - Corpus d'articles de linguistique issus de la revue "Sciences Humaines"
  - Corpus de français parlé au Québec (CFPQ)
  - Corpus de Français Parlé Parisien des années 2000 (CFPP2000)
  - Corpus de la littérature médiévale: des origines au XVe siècle
  - Corpus de la littérature narrative du moyen âge au XXe siècle
  - Corpus de Langues Parlées en Interaction (CLAPI)
  - Corpus journalistique de l'Est Républicain
  - DEDE: un corpus annoté pour le traitement des DEscriptions DEfinies
  - ESLO: Enquêtes SociLinguistiques à Orléans
  - Corpus FRAN: Français d'Amérique du Nord
  - Français au Canada
  - Français légal ancien de Normandie
  - FRANTEXT
  - FRANTEXT Moyen Français
  - French Learner Language Oral Corpora (filoc)
  - French Treebank - Corpus arboré pour le français
  - Grand Corpus des grammaires françaises, des remarques et des traités sur la langue
  - Guillaume de Digulleville, Pèlerinage de Vie humaine
  - KAsseler MOrgenstern KOrpus (KAMOKO)
  - Korpora des Centre National de Ressources Textuelles et Lexicales (CNRTL)
  - Korpora in den Ressources électroniques de la SIDF (Société Internationale de Diachronie du Français)
  - LangAge: Korpus biographischer Interviews des höheren Lebensalters
  - Les plus anciens documents linguistiques de la France (DocLing)
  - Modéliser le changement: Les voies du français (MCVF)
  - Narrations et Dialogues en français ancien: The Anglo-Norman Year Books Corpus (ANYBC)
  - Nouveau Corpus d'Amsterdam (NCA)
  - Orféo - Corpus d'Etude pour le Français Contemporain (CEFC)
  - PFC – Phonologie du Français Contemporain
  - PERCEO : un Projet d'Etiqueteur Robuste pour l'Ecrit et pour l'Oral
  - SERMO : méthode d'annotation et d'exploitation de corpus paralittéraires pour l'analyse en linguistique diachronique
  - TCOF - Traitement des Corpus Oraux en Français
  - Testaments enregistrés au Parlement de Paris sous le règne de Charles VI
  - Textes de Français Ancien (TFA)
  - Théâtre classique
  - Varitext
- Spanische Korpora und Textdatenbanken
  - Biblioteca digital Arte Lope
  - Clásicos Hispánicos
  - C-Or-DiAL
  - Corpus del Español (Genre/Historical)
  - Corpus del Español (Web/Dialects)
  - Corpus del Español Actual (CEA)
  - Corpus del Español del Siglo XXI (CORPES XXI)
  - Corpus de Referencia del Español Actual (CREA)
  - Corpus Diacrónico del Español (CORDE)
  - Corpus of Spanish Golden-Age Sonnets
  - Corpus of Spanish Novel from 1880-1940
  - Corpus periodístico Hemero
  - epub libre
  - ESCOW14 (Spanish web corpus)
  - IMPACT-es diachronic corpus
  - Moralischen Wochenschriften
  - Proyecto COLA (Corpus Oral de Lenguaje Adolescente)
  - Proyecto para el estudio sociolingüístico del español de España y de América (PRESEEA)
  - Spanish-American Novel (1880-1916)
  - Teatro clásico – Comedias, entremeses, autos y más
- Italienische Korpora und Textdatenbanken
  - "La Repubblica" Corpus
  - APHASIA Corpus
  - BAnca Dati dell'Italiano Parlato (BADIP)
  - Banche dati dell'italiano scritto e parlato
  - CLIPS (Korpora e Lessici dell'Italiano Parlato e Scritto) - sezione "Korpora di Parlato"
  - CoLFIS: Corpus e Lessico di Frequenza dell'Italiano Scritto
  - CORIS/CODIS/DiaCORIS
  - Corpus of Italian Newspapers
  - Corpus parlato di italiano L2
  - Corpus PENELOPE
  - Corpus Taurinense
  - itTenTen16

- ItWac
- KIParla corpus
- LIPS: Lessico Italiano Parlato di Stranieri
- Morfologia dell'italiano in diacronia (MIDIA)
- OVI: Opera del Vocabolario Italiano
- PAISÀ: Corpus italiano
- ParlaTO corpus
- Perugia Corpus (PEC)
- Twita
- VALICO: Varietà di Apprendimento della Lingua Italiana – Corpus Online und andere Lernerkorpora
- Web2Corpus\_IT
- Portugiesische Korpora und Textdatenbanken
  - Colonia Corpus of Historical Portuguese
  - Corpus de Referência do Português Contemporâneo (CRPC)
  - Corpus Histórico do Português Tycho Brahe
  - O Corpus do Português
- Mehrsprachige Korpora und Textdatenbanken
  - BoLC: Bononia Legal Corpus
  - Child Language Data Exchange System (CHILDES)
  - Corpora from the Web (COW)
  - EPIC (European Parliament Interpreting Corpus)
  - Europarl: European Parliament Proceedings Parallel Corpus
  - Hamburg Adult Bilingual Language (HABLA)
  - Korpus-Übersicht des SLDR (Speech and Language Data Repository)
  - MultiUN: Multilingual UN Parallel Text
  - NUNC: Newsgroups UseNet Corpora
  - Patent Translation Resource (PatTR)
  - SLDR: Speech and Language Data Repository
  - Swiss SMS Corpus
  - Web as Corpus (WaCky)
  - WhatsUp, Switzerland?
  - Wikicorpus
- Offline verfügbare Korpora
  - ADIL2: Archivio Digitale di Italiano L2
  - C-ORAL-ROM – Integrated reference corpora for spoken romance languages
  - LABLITA
  - Lessico di frequenza dell'italiano parlato (LIP)
  - LIZ (Letteratura Italiana Zanichelli)
  - VELI (Vocabolario Elettronico della Lingua Italiana)
- Literatur zur Korpuslinguistik
- Internet und Social Media
- Lizenzpflichtige Textkorpora