

Banche dati dell'italiano scritto e parlato

Übersicht der Accademia della Crusca zu diversen, auch weniger bekannten italienischen Korpora und Textdatenbanken

<http://www.accademiadellacrusca.it/it/link-utili/banche-dati-dellitaliano-scritto-parlato>

zum Beispiel:

- CoLFIS (Corpus e Lessico di Frequenza dell'Italiano Scritto) - ISTC
- Corpora e lessici dell'italiano parlato e scritto (CLIPS)
- Corpora LABLITA (Corpus di italiano parlato, corpus Stammerjohann, etc.)
- Corpus Athenaeum
- Corpus Taurinense (22 testi fiorentini della seconda metà del XIII secolo)
- Corpus di parlato cinematografico
- Corpus di parlato telegiornalistico. Anni Sessanta vs. 2005 (CPT)
- Corpus VALICO Varietà Apprendimento Lingua Italiana
- Lessico italiano radiofonico (LIR corpus)
- Lessico italiano televisivo (LIT corpus or LIT 2006 corpus)
- Newsgroup UseNet Corpora (NUNC)
- Corpus della Piattaforma per l'apprendimento dell'italiano su corpora annotati (PAISÀ corpus)
- pvbvarless (Valori di età di acquisizione, immaginabilità e frequenza per 527 parole del Primo Vocabolario del Bambino)
- SMS Monitor Studies
- varless (Variabili lessicali e sub-lessicali: valori per 626 nomi dell'italiano)
- uvm.