

French Treebank - Corpus arboré pour le français

Annotiertes Korpus des Französischen (Part of Speech und Phrasenstruktur)

Sprache	Französisch
Varietät	Standard
Sprachliche Realisierung	schriftlich
Umfang	ca. 1 Mio. Wörter
Medium	Artikel der Zeitung <i>Le Monde</i>
Geographischer Ursprung	Frankreich (Paris)
Zeitliche Einordnung	1989-1995
Form der Daten	digitalisierte und zur Analyse aufbereitete Zeitungsartikel
Format	XML
Annotation	lemmatisiert, POS (part of speech)-annotiert, morphosyntaktisch annotiert (Flexionsformen), Annotation der syntaktischen Konstituenten und ihrer Funktion
Mögliche Suchabfragen	aufgrund der umfangreichen Annotation sind detaillierte Analysen in Bezug auf Wortformen, Konkordanzen und syntaktische Strukturen möglich
Quelle/Herausgeber	LLF (Laboratoire de Linguistique Formelle), CNRS, Université Paris Diderot
Nutzungsvoraussetzungen	Nutzung nur im Rahmen einer Forschungsvereinbarung möglich (Kontakt und Antragsformular auf der Homepage)
Link	http://ftb.linguist.univ-paris-diderot.fr/