

Forschungsdaten im Fokus des FID Linguistik

Verzeichnen, Verlinken,
Verfügbarmachen



UNIVERSITÄTS
BIBLIOTHEK
FRANKFURT AM MAIN

Heike Renner-Westermann & Vanya Dimitrova



Fachinformationsdienst (FID) Linguistik

- Linguistik-Portal
- Services im Überblick
- Forschungsdaten in der Linguistik
- Verzeichnen
- Verlinken
- Verfügbarmachen
- Ausblick

Fachinformationsdienst (FID) Linguistik

- Infrastruktur & Services
- Für die allgemeine und vergleichende Sprachwissenschaft und die einzelphilologischen Sprachwissenschaften
- DFG-Programm „Fachinformationsdienste für die Wissenschaft“
- Förderung seit 2017 an der UB Frankfurt/M.
- Lin|gu|is|tik-Portal
- <https://www.linguistik.de>



Katalogsuche

Suchen

Bücher
Datenbanken Bibliothekskataloge
Literaturhinweise
Aufsätze Sprachkorpora
« Open Access Publikationen »
E-Journals Websites Zeitschriften
Bibliografien
Online-Wörterbücher

Recherche | Search

Willkommen beim Lin|gu|is|tik-Portal

Sie finden hier Informationen und Services zu allen Fachgebieten der Sprachwissenschaft.

Das Lin|gu|is|tik-Portal ist die Online-Plattform des Fachinformationsdienstes (FID) Linguistik.

- Katalogsuche
- Link-Verzeichnis
- Zeitschriften-Verzeichnis
- Datenbank-Verzeichnis
- Wörterbuch-Verzeichnis
- Hilfe zur Recherche

Suchen

Bücher
Datenbanken Bibliothekskataloge
Literaturhinweise
Aufsätze Sprachkorpora
« **Open Access Publikationen** »
E-Journals Websites Zeitschriften
Bibliografien
Online-Wörterbücher

Recherche | Search

Willkommen beim Lin|gu|is|tik-Portal

Sie finden hier Informationen und Services zu allen Fachgebieten der Sprachwissenschaft.

Das Lin|gu|is|tik-Portal ist die Online-Plattform des Fachinformationsdienstes (FID) Linguistik.



Annohub-Repository

Annohub RDF
Edition

BLL LOD Edition

LOD-Suche

Suchen

Bücher
Datenbanken Bibliothekskataloge
Literaturhinweise
Aufsätze Sprachkorpora
« Open Access Publikationen »
E-Journals Websites Zeitschriften
Bibliografien
Online-Wörterbücher

Recherche | Search

Willkommen beim Lin|gu|is|tik-Portal

Sie finden hier Informationen und Services zu allen
Fachgebieten der Sprachwissenschaft.

Das Lin|gu|is|tik-Portal ist die Online-Plattform des
Fachinformationsdienstes (FID) Linguistik.

E-Journal-Hosting

Linguistik-Journals

Linguistik-
Repository

Reihe Empirische
Linguistik

Katalogsuche

Suchen

Bücher
Datenbanken Bibliothekskataloge
Literaturhinweise
Aufsätze Sprachkorpora
« Open Access Publikationen »
E-Journals Websites Zeitschriften
Bibliografien
Online-Wörterbücher

Recherche | Search

Willkommen beim Lin|gu|is|tik-Portal

Sie finden hier Informationen und Services zu allen
Fachgebieten der Sprachwissenschaft.

Das Lin|gu|is|tik-Portal ist die Online-Plattform des
Fachinformationsdienstes (FID) Linguistik.

Katalogsuche

Suchen

Neuerwerbungen
Kaufvorschlag
Korpus-Lizenzen
FID-Lizenzen
Web News

Bücher
Datenbanken Bibliothekskataloge
Literaturhinweise
Aufsätze Sprachkorpora
« Open Access Publikationen »
E-Journals Websites Zeitschriften
Bibliografien
Online-Wörterbücher

Recherche | Search

Willkommen beim Lin|gu|is|tik-Portal

Sie finden hier Informationen und Services zu allen Fachgebieten der Sprachwissenschaft.

Das Lin|gu|is|tik-Portal ist die Online-Plattform des Fachinformationsdienstes (FID) Linguistik.

Über uns

Team

Projekt

Publikationen

Lin|gu|is|tik News

Wissenschaftlicher
Beirat

Kontakt

Katalogsuche

Suchen

Bücher
Datenbanken Bibliothekskataloge
Literaturhinweise
Aufsätze Sprachkorpora
« Open Access Publikationen »
E-Journals Websites Zeitschriften
Bibliografien
Online-Wörterbücher

Recherche | Search

Willkommen beim Lin|gu|is|tik-Portal

Sie finden hier Informationen und Services zu allen Fachgebieten der Sprachwissenschaft.

Das Lin|gu|is|tik-Portal ist die Online-Plattform des Fachinformationsdienstes (FID) Linguistik.

Forschungsdaten

- Sprachkorpora (Text, Audio, Video)
- Digitaler Volltext, Mio Wörter
- Keine E-Journals, keine E-Books, kein PDF
- „Rohdaten“ für die maschinelle Verarbeitung
- Linguistisch annotiert
- Tags: Satzgrenze, Satzart, Wortgrenze, N, V, Adj, Singular, Plural, Genitiv, Akkusativ, Wurzel, Suffix etc.

Verzeichnen von Forschungsdaten 1

- im Link-Verzeichnis
- Ressourcentyp
„Korpora /
Forschungsdaten“
- Erschließung:
Objektsprache,
linguistische Spezifika

62. Corpus Paralelo CLUVI/CLUVI Parallel Corpus

<https://repositori.upf.edu/handle/10230/20051>

Thema: Fachsprache; Kontrastive Linguistik; Korpuslinguistik; Übersetzungswissenschaft

Schlagwort: Alignierung; Kookkurrenz; Literatursprache

Sprache: Baskisch; Deutsch; Englisch; Französisch; Galizisch; Katalanisch; Portugiesisch; Spanisch

Ressourcentyp: Korpora / Forschungsdaten

Veröffentlicht durch: Seminario de Lingüística Informática (SLI) = Computational Linguistics Group

Geschrieben in: Englisch; Galizisch

Land: Spanien

Zugang: freier Zugriff

Verzeichnen von Forschungsdaten 2

- Linked Open Data (LOD)
 - BLL-Schlagwort-Thesaurus für LOD veröffentlicht
 - BLL-Schlagwort-Thesaurus mit LLOD-Ressourcen (z.B. OLiA, Lexvo, Glottolog) verlinkt
- ➔ eine Suche im Linguistik-Portal findet Objekte in der LLOD-Cloud



LOD-Suche

Suchen

[Erweiterte Suche](#)

Das Lin|gu|is|tik-Portal ist mit Linked Open Data (LOD) vernetzt. Die LOD-Suche bietet Ihnen die Möglichkeit, gezielt nach Sprachressourcen zu suchen, die über diese Vernetzung in das Portal eingebunden wurden.

Sie finden hier linguistisch annotierte, frei verfügbare Forschungsdaten wie Korpora, Wörterbücher oder Lexika. Für die Integration der Ressourcen wurden im Rahmen des FID Linguistik Routinen entwickelt, die auf Semantic-Web-Technologien und computerlinguistischen Methoden basieren: Die erfassten Ressourcen werden automatisiert hinsichtlich Objektsprache und Annotationsschema analysiert. Die vorliegenden Beschreibungen der Ressourcen werden durch die Analyseergebnisse angereichert und in einem eigens dafür eingerichteten Metadaten-Repository ([Annohub](#)) gespeichert. Diese Annohub-Daten dienen als Basis für die Suchaufbereitung.

Als Sucheinstiege eignen sich Sprachbezeichner (z.B. Arabisch, Schwedisch, Hindi) sowie linguistische Begriffe für die annotierten Phänomene (z.B. Auxiliärverb, Akkusativ, Dual). Weitere Informationen und Tipps für die Suche finden Sie [hier](#).



Ergebnis eingrenzen:

Schlagwort ▶

Autor / Herausgeber ▶

Jahr ▶

Material ▶

Typ ▶

BLLDB-Zugang ▶

Suche in den Katalogen und Verzeichnissen

Sortierung

Erscheinungsjahr ↓ (absteigend) ▼

Erweiterte Suche

Suchen

📱 Mobile Ansicht

Seite: **1** 2 3 4 5 ... 20

Treffer 1 – 20 von 394

1 Universal Dependencies Treebank Spanish-AnCora

Annohub

[Details anzeigen](#) ▼

2 Universal Dependencies Treebank Hungarian

Annohub

[Details anzeigen](#) ▼

3 DBnary - Wiktionary as Linguistic Linked Open Data (Polish Edition)

Gilles Sérasset

Annohub

[Details anzeigen](#) ▼

4 Blog mix 2005 (2017-02-17); Bloggmix 2005 (2017-02-17)

Annohub

[Details anzeigen](#) ▼

Alle

Kataloge

- [UB Frankfurt Linguistik](#)
- [IDS Mannheim](#)
- [OLC Linguistik](#)
- [UB Frankfurt Retrokatalog](#)
- [DNB Sachgruppe Sprache](#)
- [Institut für Empirische Sprachwissenschaft](#)
- [Leibniz-Zentrum Allgemeine Sprachwissenschaft](#)

Bibliografien

- [BLLDB](#)
- [BDSL](#)
- [IDS Bibliografie zur deutschen Grammatik](#)
- [IDS Bibliografie zur Gesprächsforschung](#)
- [IDS Konnektoren im Deutschen](#)
- [IDS Präpositionen im Deutschen](#)
- [IDS OBELEX meta](#)
- [MPI-SHH Linguistik Sammlung](#)
- [MPI für Psycholinguistik](#)

Linked-Open-Data-Kataloge

- [Annohub](#) 394

Online-Quellen

Verzeichnen von Forschungsdaten 2

- I. Sprachressourcen aus der LOD-Cloud analysiert, um OLiA-Annotationen zu finden
➔ findbar im Linguistik-Portal
- II. Sprachressourcen, deren Metadaten in der LOD-Cloud liegen, anhand von Datenproben analysiert:
 - a) Bestimmung der Objektsprache
 - b) Zuordnung der gefundenen Annotation zu einem vorliegenden Annotationsmodell➔ findbar im Linguistik-Portal

Verzeichnen von Forschungsdaten 2

- III. Sprachressourcen können manuell hochgeladen werden, um die Analyse-Pipeline zu durchlaufen
➔ findbar im Linguistik-Portal

Alle ermittelten Metadaten (Objektsprache, Annotationen) werden in einem dafür eingerichteten Repository gespeichert:

<https://www.linguistik.de/lod/annohub-repository/>

Annohub-Repository

Annohub ist ein Repositorium für Metadaten von Sprachressourcen wie Korpora oder Lexika. Die enthaltenen Ressourcen wurden mittels einer eigens dafür entwickelten Webanwendung hinsichtlich Annotationsschema und Objektsprache analysiert. Der Fokus liegt aktuell auf frei verfügbaren Ressourcen, die im RDF-, CoNLL- oder XML-Format vorliegen. Die erfolgreich ermittelten Metadaten werden – zusammen mit den aggregierten Kern-Metadaten – in das Lin|gu|is|tik-Portal eingebunden und für eine Nachnutzung bereitgestellt (vgl. [Annohub RDF Edition](#)). Eine gezielte Recherche ist über die [LOD-Suche](#) möglich.

Hier finden Sie eine tabellarische Übersicht der Annohub-Ressourcen. Eine Ressource kann aus mehreren Dateien bestehen (siehe *Details*).

10 ▼ Einträge anzeigen

Suchen

Ressource	Sprache	Annotation	URL	Details
3000 Common English words with Māori translations	Maori; English	Ontolex	Download (210,4 KB)	Details
8 SIDOR (2017-10-16); 8 SIDOR (2017-10-16)	Swedish	Suc ; Mamba ; Mamba-Syntax	Download (87,5 MB)	Details
Academic texts – Humanities 1997–2012 (2017-10-16); Akademiska texter – Humaniora 1997–2012 (2017-10-16)	Swedish	Suc ; Mamba ; Mamba-Syntax	Download (208,7 MB)	Details

Verlinken von Forschungsdaten

- Bibliography of Linguistic Literature (BLL)
- Über 1.000 Freitext-Schlagwörter an rund 7.000 Titeln
- Intellektuelle Prüfung
- Online-Präsenz: Homepage, URN, Handle, PURL, DOI, Clarin-VLO
- MediaType: text, audio, video
- ModalityType: written, spoken, sign language
 - ➔ über 450 Normdatensätze verlinkt mit über 4.500 Titeln

COCA

Suchen



Services

Willkommen beim Lin|gu|is|tik-Portal

Sie finden hier Informationen und Services zu allen Fachgebieten der Sprachwissenschaft.

Das Lin|gu|is|tik-Portal ist die Online-Plattform des Fachinformationsdienstes (FID) Linguistik.



Ergebnis eingrenzen:

Schlagwort ▶

Autor /
Herausgeber ▶

Jahr ▶

Material ▶

Typ:

Aufsatz (567)

Buch (76)

Zeitschrift (70)

Korpus (1)

Zeitschrift (1)

Website (1)

BLLDB-
Zugang ▶

Suche in den Katalogen und Verzeichnissen

COCA

Sortierung

Erscheinungsjahr ↓ (absteigend) ▼

Erweiterte Suche

Suchen

Seite: [1](#) [2](#) [3](#) [4](#) [5](#) ... [36](#)

Treffer 1 – 20 von 719

- 1 МЕХАНИЗМЫ ОБРАЗОВАНИЯ И ЛЕКСИКОГРАФИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ НЕОЛОГИЗМОВ-КОЛЛОКАЦИЙ ... : MECHANISMS OF FORMATION AND LEXICOGRAPHIC INTERPRETATION OF NEOLOGISM-COLLOCATIONS ...
Е.И. Архипова. - : Мир науки, культуры, образования, 2022

BASE

Details anzeigen ▼

[ONLINE ANSEHEN](#)

- 2 Purpose of things ...
Rose, David. - : Open Science Framework, 2022

BASE

Details anzeigen ▼

[ONLINE ANSEHEN](#)[subito](#)

- 3 A corpus-based study of the acquisition of the English progressive by L1 Chinese learners: From prototypical activities to marked statives ...

Xiaoyan, Zeng; Xiaoxiang, Chen; Shirai, Yasuhiro. - : Zenodo, 2022

BASE

Details anzeigen ▼

Mobile Ansicht

 Alle

Kataloge

<input checked="" type="checkbox"/> UB Frankfurt Linguistik	15
<input checked="" type="checkbox"/> IDS Mannheim	0
<input checked="" type="checkbox"/> OLC Linguistik	21
<input checked="" type="checkbox"/> UB Frankfurt Retrokatalog	0
<input checked="" type="checkbox"/> DNB Sachgruppe Sprache	7
<input checked="" type="checkbox"/> Institut für Empirische Sprachwissenschaft	0
<input checked="" type="checkbox"/> Leibniz-Zentrum Allgemeine Sprachwissenschaft	0

Bibliografien

<input checked="" type="checkbox"/> BLLDB	207
<input checked="" type="checkbox"/> BDSL	0
<input checked="" type="checkbox"/> IDS Bibliografie zur deutschen Grammatik	0
<input checked="" type="checkbox"/> IDS Bibliografie zur Gesprächsforschung	0
<input checked="" type="checkbox"/> IDS Konnektoren im Deutschen	0
<input checked="" type="checkbox"/> IDS Präpositionen im Deutschen	0
<input checked="" type="checkbox"/> IDS OBELEX meta	1
<input checked="" type="checkbox"/> MPI-SHH Linguistik Sammlung	0
<input checked="" type="checkbox"/> MPI für Psycholinguistik	1

Linked-Open-Data-Kataloge

<input checked="" type="checkbox"/> Annohub	0
---	---

Online-Quellen

<input checked="" type="checkbox"/> Link-Verzeichnis	1
--	---



Ergebnis eingrenzen:

Schlagwort ▶

Autor / Herausgeber ▶

Jahr ▶

Material ▶

Typ:
Korpus (1)

BLLDB-Zugang ▶

Suche in den Katalogen und Verzeichnissen

COCA

Sortierung Erscheinungsjahr ↓ (absteigend) ▼

Erweiterte Suche Suchen

Treffer 1 – 1 von 1

- 1** Corpus of Contemporary American English [COCA]
<https://www.english-corpora.org/coca/>
Schlagwort: Amerikanisches Englisch
Homepage: <https://www.english-corpora.org/coca/>
[Zugehörige Publikationen \(199\)](#)
BLLDB
[Details verbergen ↕](#)



Alle Mobile Ansicht

Kataloge

<input checked="" type="checkbox"/> i UB Frankfurt Linguistik	0
<input checked="" type="checkbox"/> i IDS Mannheim	0
<input checked="" type="checkbox"/> i OLC Linguistik	0
<input checked="" type="checkbox"/> i UB Frankfurt Retrokatalog	0
<input checked="" type="checkbox"/> i DNB Sachgruppe Sprache	0
<input checked="" type="checkbox"/> i Institut für Empirische Sprachwissenschaft	0
<input checked="" type="checkbox"/> i Leibniz-Zentrum Allgemeine Sprachwissenschaft	0

Bibliografien

<input checked="" type="checkbox"/> i BLLDB	1
<input checked="" type="checkbox"/> i BDSL	0
<input checked="" type="checkbox"/> i IDS Bibliografie zur deutschen Grammatik	0
<input checked="" type="checkbox"/> i IDS Bibliografie zur Gesprächsforschung	0
<input checked="" type="checkbox"/> i IDS Konnektoren im Deutschen	0
<input checked="" type="checkbox"/> i IDS Präpositionen im Deutschen	0
<input checked="" type="checkbox"/> i IDS OBELEX meta	0
<input checked="" type="checkbox"/> i MPI-SHH Linguistik Sammlung	0
<input checked="" type="checkbox"/> i MPI für Psycholinguistik	0

Linked-Open-Data-Kataloge

<input checked="" type="checkbox"/> i Annohub	0
---	---



Ergebnis eingrenzen:

Schlagwort ▶

Autor / Herausgeber ▶

Jahr ▶

Material ▶

Typ:

Aufsatz (185)

Buch (14)

BLLDB-Zugang ▶

Suche in den Katalogen und Verzeichnissen

Sortierung

Erscheinungsjahr ↓ (absteigend) ▾

Erweiterte Suche

Suchen

Seite: [1](#) [2](#) [3](#) [4](#) [5](#) ... [10](#)

Treffer 1 – 20 von 199

1 From 'movement into action' to 'manner of causation': changes in argument mapping in the "into"-causative

Flach, Susanne

In: Linguistics. - Berlin [u.a.] : Mouton de Gruyter 59 (2021) 1, 247-283

Verwendetes Korpus: Corpus of Historical American English

Verwendetes Korpus: Corpus of Late Modern English Texts

Verwendetes Korpus: Corpus of Contemporary American English

Schlagwort: Amerikanisches Englisch; Art und Weise (Semant.); Bewegungsverben; Corpus of Contemporary American English; Corpus of Historical American English; Corpus of Late Modern English Texts; Einzelsprachliche Elemente; Entrenchment-and-Conventionalization Model; Expansion (Sprachwandel); Frühneuenglisch; Gerundium; Grammatikalisierung; into - (engl. Konstr.); Kategorie (ethnoling.); Konstruktionsgrammatik; Korpuslinguistik; Raumpräposition; Semantische Felder; Semantische Syntax; Textkorpus (engl.) Early English Books Online (EEBO); Thematische Rolle; Verba causativa; Verbvalenz

BLLDB

Details verbergen ✎

ONLINE ANSEHEN

subito

ZDB

Mobile Ansicht

Alle

Kataloge

<input checked="" type="checkbox"/> i UB Frankfurt Linguistik	14
<input checked="" type="checkbox"/> i IDS Mannheim	0
<input checked="" type="checkbox"/> i OLC Linguistik	16
<input checked="" type="checkbox"/> i UB Frankfurt Retrokatalog	0
<input checked="" type="checkbox"/> i DNB Sachgruppe Sprache	0
<input checked="" type="checkbox"/> i Institut für Empirische Sprachwissenschaft	0
<input checked="" type="checkbox"/> i Leibniz-Zentrum Allgemeine Sprachwissenschaft	0

Bibliografien

<input checked="" type="checkbox"/> i BLLDB	199
<input checked="" type="checkbox"/> i BDSL	0
<input checked="" type="checkbox"/> i IDS Bibliografie zur deutschen Grammatik	0
<input checked="" type="checkbox"/> i IDS Bibliografie zur Gesprächsforschung	0
<input checked="" type="checkbox"/> i IDS Konnektoren im Deutschen	0
<input checked="" type="checkbox"/> i IDS Präpositionen im Deutschen	0
<input checked="" type="checkbox"/> i IDS OBELEX meta	0
<input checked="" type="checkbox"/> i MPI-SHH Linguistik Sammlung	0
<input checked="" type="checkbox"/> i MPI für Psycholinguistik	0

Linked-Open-Data-Kataloge

<input checked="" type="checkbox"/> i Annohub	0
---	---

Verfügbarmachen von Forschungsdaten

- Überregionale FID-Lizenzen
- Untypische Produkte
- Nutzung: maschinelle Auswertung, statistische Verfahren, Text- and-Data-Mining etc.
- Download des gesamten Korpus
oder
spezifische Webschnittstelle zur Analyse des gesamten Korpus

Lizenzierung von Sprachkorpora 1

- European Language Resources Association (ELRA)
- Liste mit angebotenen Korpora (anfangs 86, heute 304)
- individuelle Lizenzbeantragung
- Einzelpreise zwischen 0,- € und 10.000,- €
- Korpus-Fonds
- Services / **Korpus-Lizenzen**
- <https://www.linguistik.de/services/licences-corpora/>

Lizenzierung von Sprachkorpora 2

- English-Corpora.org / Mark Davies
(ehem. BYU-Corpora)
- 56 Campuslizenzen
- Services / **FID-Lizenzen**
- <https://www.linguistik.de/services/fid-licences/>

Ausblick

- Ausbau des Nachweises von Forschungsdaten
- Thematisches Modul „Bedrohte Sprachen“
- Quantitativer Ausbau der Katalog-Nachweise
- Vernetzung mit weiteren LOD-Ressourcen
- Datenmodell und Technologie des Linguistik-Portals auf LOD umstellen
- Weiterentwicklung der Recherche-Funktionalität
- Volltextsuche über Open-Access-Publikationen
- Kooperation mit NFDI-Konsortium Text+
- Usability-Evaluation

Vielen Dank für Ihre Aufmerksamkeit

Kontakt:

Fachinformationsdienst Linguistik

info@linguistik.de