

# Informationsveranstaltung

## BA Computerlinguistik

### Korpus- und Computerlinguistik

**Prof. Dr. Stephanie Evert**

Lehrstuhl für Korpus- und Computerlinguistik

<https://www.linguistik.phil.fau.de/>

WS 2024/25



FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG

PHILOSOPHISCHE FAKULTÄT  
UND FACHBEREICH THEOLOGIE

# Was ist eigentlich ein Korpus?

Das Korpus  
Die Korpora

- Korpus<sup>1</sup> (im weiten Sinn)  
= Sammlung von Sprachdaten / Texten in maschinenlesbarer Form
  - sehr große Korpora ( $\geq 100$  M Wörter) sind besonders nützlich
  - Auswertung mit statistischen Methoden und maschinellen Lernverfahren
- Korpus<sup>2</sup> (im engen Sinn)  
= Stichprobe authentischer Sprachdaten / Texte, die für eine bestimmte Sprache oder Sprachvarietät repräsentativ ist
  - z.B. literarische Korpora, Dialekte, gesprochene Sprache, IBK, ...
  - Basis für empirisch fundierte sprachwissenschaftliche Studien
  - zentral für Korpuslinguistik (im engen Sinn) und *Digital Humanities*
- **Korpus- und Computerlinguistik**  
= Erstellung, maschinelle Verarbeitung und Auswertung von Korpora<sup>1</sup>

## Was macht eigentlich Computerlinguistik?

- Linguistische Analyse von Texten (Wortarten, Syntax, ...)
- Text Mining, Knowledge Mining, Argument Mining, ...
- Erkennung von Autoren, Meinungen, Sentiment, Emotionen, ...
- Erkennung von Fake News, Hate Speech, Desinformation, ...
- Maschinelle Übersetzung
- Spracherkennung & Sprachsynthese
- Dialogsysteme & virtuelle Assistenten
- Generierung und Überarbeitung von Texten
- Computerlexikographie & Terminologearbeit
- Rechtschreib- und Grammatikkorrektur, stilistische Verbesserungen
- Unterstützung für Korpuslinguistik, Digital Humanities und Computational Social Science

# Künstliche Intelligenz (KI): virtuelle Assistenten

IN 1939's CARTOON  
"THE POINTER", THIS  
GUY GOT A NEW,  
MORE PEAR-SHAPED  
BODY & PUPILS WERE  
ADDED TO HIS EYES



# Künstliche Intelligenz (KI): virtuelle Assistenten

THIS CLAUSE IN A UNION CONTRACT SAYS THAT WAGES WILL RISE OR FALL DEPENDING ON A STANDARD SUCH AS COST OF LIVING



# Methoden der CL: symbolische Verfahren

 **SYSTRANet** Your free online translator

Text Web Page RSS File Dictionary My Dictionary Register for free

From: English To: German Translate Options:   

For instance, on the planet Earth, man had always assumed that he was more intelligent than dolphins because he had achieved so much — the wheel, New York, wars and so on — whilst all the dolphins had ever done was muck about in the water having a good time. But conversely, the dolphins had always believed that they were far more intelligent than man — for precisely the same reasons.

Zum Beispiel auf der Planet Erde, hatte Mann immer angenommen, dass er intelligenter war, als waren Delphine, weil er soviel erzielt hatte — das Rad, New York, Kriege und so weiter — während alle Delphine überhaupt getan hatten, Dreck ungefähr im Wasser, das eine gute Zeit hat. Aber andererseits, hatten die Delphine immer geglaubt, dass sie weit intelligenter als Mann — aus genau den gleichen Gründen waren.

# Methoden der CL: statistische Verfahren

☰ Google Translate



🗨️ Text

📄 Documents

DETECT LANGUAGE

ENGLISH

SPANISH

FRENCH



GERMAN

ENGLISH

SPANISH



For instance, on the planet Earth, man had always assumed that he was more intelligent than dolphins because he had achieved so much—the wheel, New York, wars and so on— whilst all the dolphins had ever done was muck about in the water having a good time. But conversely, the dolphins had always believed that they were far more intelligent than man—for precisely the same reasons. ✕

Zum Beispiel hatte der Mensch auf dem Planeten Erde immer angenommen, er sei intelligenter als Delfine, weil er so viel erreicht hatte - das Rad, New York, Kriege und so weiter -, während alles, was die Delfine jemals getan hatten, im Wasser herumgespielt hatte eine gute Zeit haben. Umgekehrt hatten die Delfine immer geglaubt, sie seien weitaus intelligenter als der Mensch - aus genau den gleichen Gründen. ☆



380 / 5000



# Methoden der CL: Deep Learning



DeepL

Übersetzer ▾

DeepL Pro

Für Unternehmen ▾

Kostenloses Probeabo starten



Anmelden



Text übersetzen

31 Sprachen



Dateien übersetzen

.pdf, .docx, .pptx



DeepL Write

KI-Schreibassistent

Englisch (erkannt) ▾



Deutsch ▾

automatisch ▾

Glossar

For instance, on planet Earth, man had always assumed that he was more intelligent than dolphins because he had achieved so much—the wheel, New York, wars and so on—whilst all the dolphins had ever done was much about in the water having a good time. But conversely, the dolphins had always believed that they were far more intelligent than man—for precisely the same reasons.



Auf dem Planeten Erde zum Beispiel hatte der Mensch immer angenommen, dass er intelligenter sei als die Delphine, weil er so viel erreicht hatte - das Rad, New York, Kriege und so weiter - während alles, was die Delphine je getan hatten, darin bestand, im Wasser herumzutollen und sich zu amüsieren. Umgekehrt hatten die Delfine immer geglaubt, sie seien viel intelligenter als der Mensch - aus genau denselben Gründen.



## Methoden: generative KI



# Digital Humanities: Digitale Editionen

PARATEXTS  
SEARCH | INDEX  
RESULTS  
DIE FACKEL CONTENTS

**DIE FACKEL CONTENTS**

DATE ISSUE PAGES 29 GO SYNC  
5.1918 XX. JAHR, Heft 474 1-160

OVERVIEW

1. Jg. 474-493 (Front cover: contents)  
(Inside front cover: adverts: contents)

2. Jg. 1 Der begabte Czernin  
23 ► Glossen  
23 Niemand geringerer als  
25 Autor und Direktor  
26 Von einem Mann namens Ernst Posse  
26 Ein Kronzeuge für die österreichische Regierung  
26 Ein Ausspruch  
27 Nimmermehr  
27 Auf Deutsches  
28 Österreich 1918  
28 Eine neue Naturgewalt  
28 Fremdwörterchutz  
31 Sauerb  
32 Die Entschädigten  
33 Rassenunterschiede  
33 Unterricht  
34 Gut getroffen  
35 Die überlegene Wirkung unserer Gase  
36 Wir Deutschen begrüßen alle Versuche, dem Völkerrecht und der Menschlichkeit zum Siege zu verhelfen, mit Freude, wir lehnen es aber ab, uns übertölpeln zu lassen  
36 Mit G.  
37 Ein Wunder  
37 Biblisches  
38 Mein Ehrenwort, daß das nicht von mir ist  
38 Die große Kanone oder Beweis gegen Barbaentum  
39 Der Praeceptor Germaniae  
40 Lionardo da Vinci  
41 Das technoromantische Abenteuer  
46 Für Lammasch  
50 ▼ Inschriften  
52 Der darbenende Bürger  
57 ▼ Glossen  
69 ▼ Notizen  
78 Bange Stunde  
81 Halbschlaß  
82 Das zweite Sonett der Louise Labé  
83 An eine Falte  
84 Suchen und Finden  
85 Die Flamme der Epimeleia  
87 ▼ Programme  
95 ▼ Glossen  
113 Ein Staatsstreich  
119 ▼ Inschriften

**DIE FACKEL TEXT**

PAGE TEXT ISSUE VOLUME  
Heft 474-483 23.5.1918 XX. JAHR

NR. 474—483 MAI 1918 XX. JAHR

DIE FACKEL

HERAUSGEBER  
KARL KRAUS

INHALT:

Der begabte Czernin / Glossen / Das technoromantische Abenteuer / Für Lammasch / Inschriften / Der darbenende Bürger / Glossen / Notizen / Bange Stunde / Halbschlaß / Das zweite Sonett / An eine Falte / Suchen und Finden / Die Flamme der Epimeleia / Programme / Glossen / Ein Staatsstreich / Inschriften / Am Sarg Alexander Girardis / Der Weltspiegel / Glossen / Zum ewigen Frieden.

NACHDRUCK VERBOTEN

Preis dieses Heftes:

**3 Kronen = 2 Mark 50 Pf.**

VERLAG ‚DIE FACKEL‘, WIEN  
III/2, HINTERE ZOLLAMTSSTRASSE 3 TELEPHON NR. 187

ERSCHEINT MINDESTENS VIERMAL IM JAHRE.

IMAGE

<https://www.oeaw.ac.at/fackellex/aac/aac.html>



# Team

<b>&lt; Department</b>
Aktuelles
<b>Team</b>
Profil
Kontakt

## Team



**Prof. Dr. Michaela Mahlberg**

Departmentsprecherin



**Christian Sandig, M.A.**

Geschäftsführer



**Andreas Blombach, M. A.**

Studiengangsberatung



**Nathan Dykes, M. A.**

Wissenschaftlicher  
Mitarbeiter



**Prof. Dr. Anastasia Glawion**

Juniorprofessorin für  
Digitale Literatur und  
Methoden



**Dr. Marianna Gracheva**

Wissenschaftliche  
Mitarbeiterin (Postdoc)



**Dr. Dominik Kremer**

Wissenschaftlicher  
Mitarbeiter (Postdoc)



# Sprachwissenschaft: Empirische Sprachbeschreibung

BNCweb Query result

http://localhost/bncweb-cgi/main.pl?program=sort&textOrSpeaker=&theData=%5Bword%3D%22linguistics%22%5D

Your query "linguistics" returned 784 hits in 100 different files on position -1 with tag-restriction any adjective (224 hits)

Sort parameters: Position: 1 Left Tag restriction:

No	Filename	
201	<a href="#">HGR 481</a>	Generative grammar rejects the empirical na
202	<a href="#">CGY 1266</a>	In the semiological phase his work is inspir
203	<a href="#">CGY 1183</a>	The structuralism that Lévi-Strauss applies i
204	<a href="#">H8V 441</a>	But by integrating these theories with their v
205	<a href="#">CGY 1271</a>	But in his second phase the issues of custom
206	<a href="#">G1N 89</a>	CM2: Minds, machines and evolution. Sample containing about 38482 words from a book (domain: belief and thought)
207	<a href="#">CGY 14</a>	Spoken or Written: Written Number of Words (tagged items): 38,616 Average sentence length (w-tags per s-unit): 27.8816 Derived text type: Academic prose Genre: W:ac:humanities_arts Text type: Written books and periodicals Publication date: 1985-1993 Age of Author: unknown Domicile of Author: unknown Sex of Author: unknown Type of Author: Multiple Age of Audience: Adult Text Domain: Informative: Belief and thought Perceived level of difficulty: High Medium of Text: Book Place of publication: UK: South (south of Bristol Channel-Wash line) Text Sample: Whole text Estimated circulation size: High Target audience sex: Mixed
208	<a href="#">CMR 1</a>	
209	<a href="#">CGY 1</a>	
210	<a href="#">KAM 6</a>	
211	<a href="#">HH3 13</a>	
212	<a href="#">J7F 18</a>	
213	<a href="#">CM2 563</a>	Even if my assessment of its implications concerning the relative order of emergence of the intentional ingredient and of syntactic structure were held to be incorrect, the mechanism of that evolution might be of independent interest, and be seen as bearing on other problems besides (especially in developmental psychology and <a href="#">theoretical_AJ0 linguistics</a> ).

**bucket** *noun*

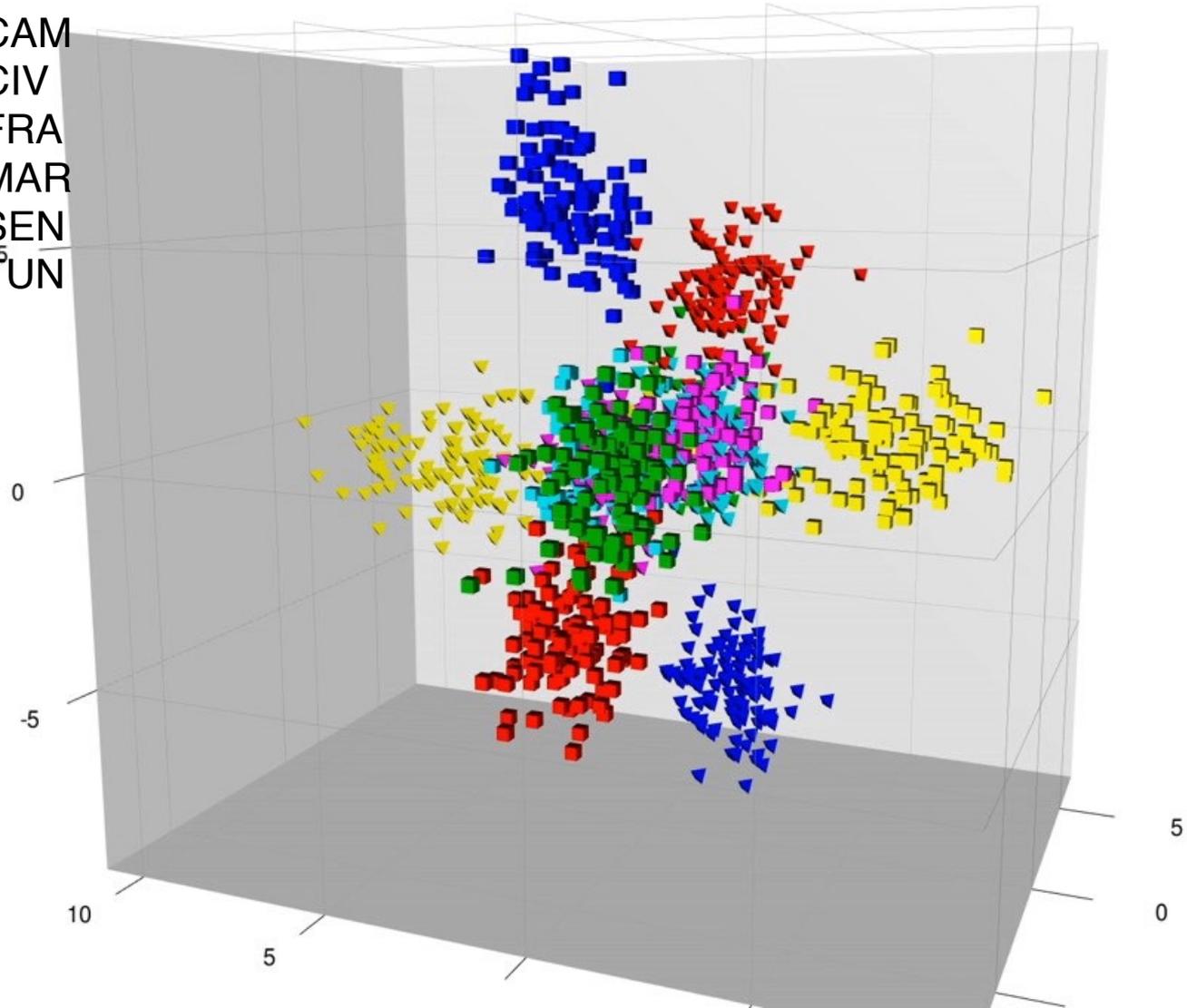
- ADJ. empty, full | galvanized, metal, plastic | leaky | champagne, ice | water | slop, waste *a slop bucket full of scraps of food* | mop | coal | fire *The sand had spilt from the fire bucket.*
- VERB + BUCKET fill *She filled the bucket with fresh water.* | carry | empty, pour, throw *She poured the bucket of dirty water down the drain.*
- BUCKET + VERB be filled with/full of sth, contain sth, hold sth | overflow
- PREP. in a/the ~ | ~ of a bucket of oats for the horses
- PHRASES a bucket and spade *The children ran down to the beach with their buckets and spades.* mop and bucket *The cleaner put down his mop and bucket and sat down.*

Go to "http://localhost/bncweb-cgi/fileInfo.pl?text=CM2&urlTest=yes"

# Sprachwissenschaft: Quantitative Linguistik

- MUTA
- ▲ TRIB
- FRAT
- ▲ VOIE
- LFI
- ▲ LM
- AJD
- ▲ MAT
- SOL
- ▲ WALFA
- LAPRE
- ▲ TEMPS

- CAM
- CIV
- FRA
- MAR
- SEN
- TUN



## RTG Dimensions of Constructional Space

### Welcome to the DFG-Funded Research Training Group of the Dimensions of Constructional Space at FAU Erlangen-Nürnberg

The Research Training Group will explore a relatively new paradigm in linguistics that has become known under the name of Construction Grammar (CxG). The fundamental premise of this approach is that the totality of speakers' linguistic knowledge is represented in a network of form-meaning pairings called constructions. Constructions differ in size (ranging from morphemes through phrasal and clausal templates to discourse conventions), abstractness (fully elaborated, partially schematic, fully schematic) and entrenchment, as well as in the way they interconnect with other constructions in the network. These properties define a multidimensional space we call 'constructional space'.

### FJUEL 2023

📅 6. October 2023 📁 Category: [PhD Researchers](#), [Talks and Events](#)



On September 29th and 30th, 2023, some of the members of our RTG took part in the annual FJUEL conference, hosted at Friedrich-Alexander Universität Erlangen-Nürnberg. The Forum Junge Englische Linguistik in Bayern (FJUEL) is a platform where doctoral and post-doctoral researchers, as well as stu...

[Continue >](#)

### Contact



**Prof. Dr. Ewa Dabrowska**

**Project Leader**  
Department of English and  
American Studies  
Lehrstuhl für Language und  
Cognition (Alexander von  
Humboldt-Professur)

Bismarckstr. 6  
91054 Erlangen

☎ +49 9131 85-29263

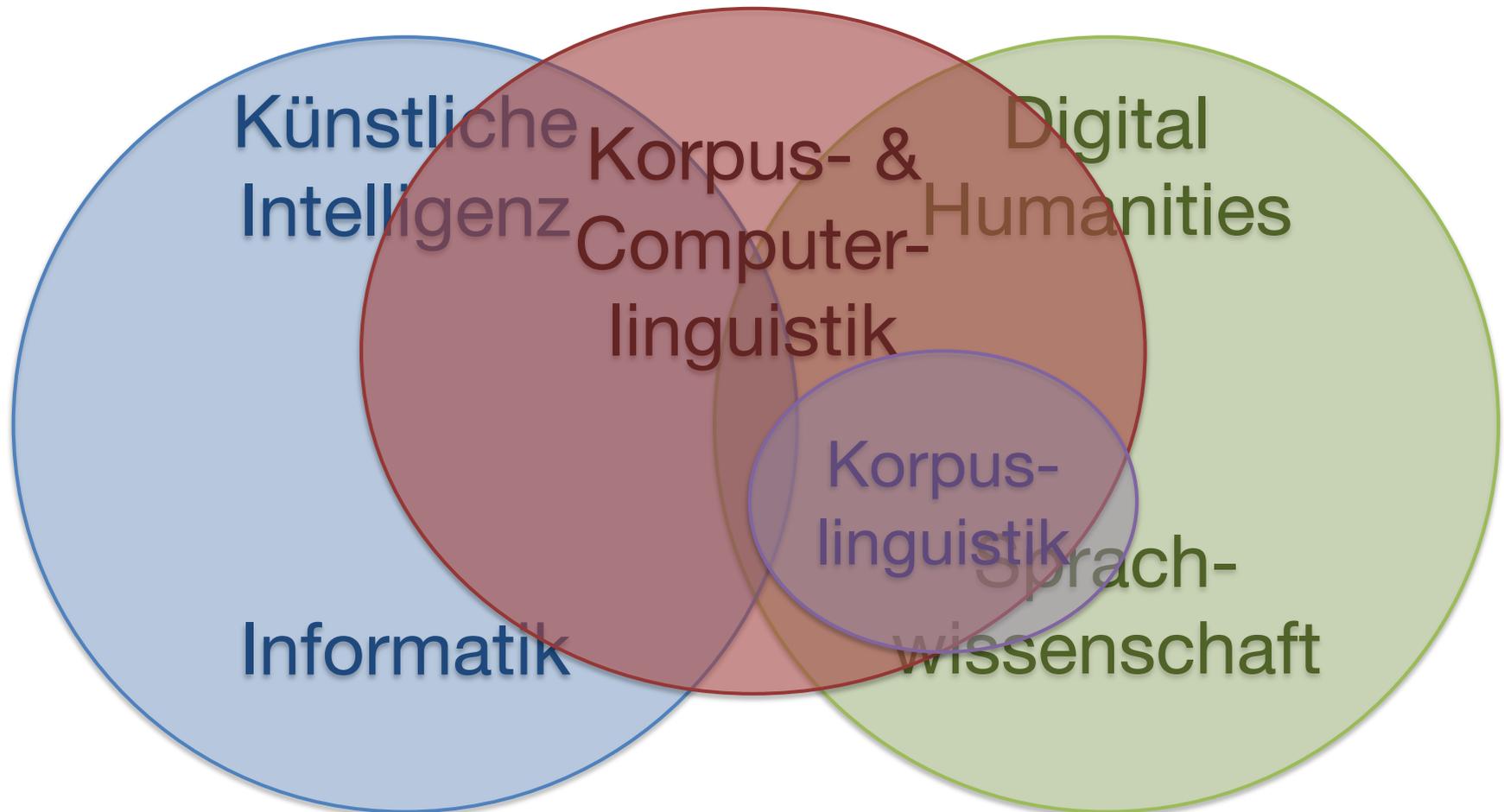
✉  
[ewa.dabrowska@fau.de](mailto:ewa.dabrowska@fau.de)

**Dr. Susanne Dyka**

**Project Coordinator**  
Lehrstuhl für Language and Cognition  
Institut für Anglistik und Amerikanistik

Bismarckstraße 6  
91054 Erlangen

# Orientierung: Korpus- und Computerlinguistik



# BA Computerlinguistik

---

- Fokus auf maschinelle Lernverfahren (insb. **Deep Learning**)
  - folgt dem Trend der modernen Computerlinguistik (und KI)
  - Schwerpunkt des Studiums: Operationalisierung von computerlinguistischen Aufgaben als Lernprobleme
  - „Baukasten“ geeignet für zahlreiche Anwendungen in Sprachtechnologie, DH, Korpuslinguistik
- Programmierpraxis: **Python**
  - konsequenter Einsatz von Python und entsprechenden Frameworks im ganzen Studienverlauf
  - entsprechende Kenntnisse werden z.B. in praktischen Hauptseminaren vorausgesetzt
- Aber keine Reduktion auf *prompt engineering* für ChatGPT
  - formale und theoretische Grundlagen der Computerlinguistik
  - interdisziplinäre Anwendungen v.a. im Rahmen der Hauptseminare (politische Diskurse, Digital Humanities, Konstruktionsgrammatik, ...)

# Grundstudium

---

- **Grundlagen der Computerlinguistik I – III**
  - GdCL I = klassische symbolische Ansätze, Mengenlehre, Logik
  - GdCL II = statistische Ansätze, maschinelle Lernverfahren, Wahrscheinlichkeitstheorie, lineare Algebra
  - GdCL III = Deep Learning = neuronale Netze, Tensoranalysis, praktische Implementierung
- **Programmierung & Infrastrukturen I + II**
  - Unix-Betriebssystem, Kommandozeile, Editor, reguläre Ausdrücke
  - Einführung in Python
  - Nutzung von Python-Bibliotheken und NLP-Standardwerkzeugen
- **Linguistische Grundkompetenzen**
  - Grundlagen der germ. Linguistik 1 oder Basismodul II: Linguistics (A)
  - bei Germanistik, Anglistik, Romanistik: DH-Modul 1 „Sprache & Text“
- **Proseminar** Computerlinguistik
  - Rezeption von Originalarbeiten, Präsentation, konstruktive Diskussion
  - Schreiben von Hausarbeiten, Literaturrecherche & Zitierung, LaTeX

# Hauptstudium

---

- **3 Hauptseminare** zu wechselnden Themen
  - 1x mit klassischer Hausarbeit
  - 1x mit mündlicher Prüfung
  - 1x mit praktischem Projekt + Projektbericht
- **1 Praxisseminar**
  - angewandtes Gruppenprojekt (z.B. Teilnahme an Shared Task)
  - gemeinsamer Projektbericht in Form eines Konferenzbeitrags
  - Schwerpunkt auf Teamwork, Selbstorganisation, Arbeitsteilung, ...
- Angebot: mind. **2 Seminare / Semester**
- **Oberseminar** (verpflichtend)
  - Besuch der Vorträge über 2 Semester (→ Anmeldung in StudOn)
  - Essay über zwei ausgewählte Vorträge
- **Praktikum** (150h = 1 Monat Vollzeit)
  - bei externer Firma oder internes Forschungspraktikum

# Erstfach ≠ Zweitfach

---

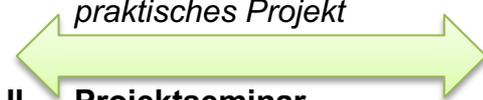
- **Zweitfach**: anwendungsorientierte Ergänzung zu geistes- oder sozialwissenschaftlichem Fach
  - komplett ohne Informatik-Importe
  - Vermittlung der mathematischen und informatischen Grundlagen ganz auf CL zugeschnitten
- **Erstfach**: Interesse an Informatik und NLP-Methoden
  - + 20 ECTS statt Schlüsselqualifikationen
  - 7,5 ECTS **Grundlagen der Informatik** (verpflichtend)
  - 7,5 ECTS + 5,0 ECTS **Wahlpflichtbereich**
    - zahlreiche Module aus dem Bachelor Informatik stehen zur Wahl
    - *Mathematik für Naturwissenschaftler (7,5) und Modellbildung & Statistik (5)*

# Studienplan

## Computerlinguistik (Zweifach)



Semester 1 (WiSe)	Semester 2 (SoSe)	Semester 3 (WiSe)	Semester 4 (SoSe)	Semester 5 (WiSe)	Semester 6 (SoSe)
<b>VL Grundlagen der CL 1</b> 2 SWS; 2 ECTS	<b>VL Grundlagen der CL 2</b> 2 SWS; 2 ECTS	<b>VL Grundlagen der CL 3</b> 2 SWS; 3 ECTS	<b>Hauptseminar I</b> 2 SWS; 5 ECTS; <i>Hausarbeit</i>	<b>Hauptseminar III</b> 2 SWS; 5 ECTS; <i>praktisches Projekt</i>	
<b>Ü Grundlagen der CL 1</b> 2 SWS; 3 ECTS	<b>Ü Grundlagen der CL 2</b> 2 SWS; 3 ECTS	<b>Ü Grundlagen der CL 3</b> 2 SWS; 7 ECTS	<b>Hauptseminar II</b> 2 SWS; 5 ECTS; <i>mündl. Prüfung</i>	<b>Projektseminar</b> 2 SWS; 5 ECTS; <i>Teamprojekt</i>	
<b>S Grundkurs Programmierung</b> 2 SWS; 5 ECTS	<b>S Aufbaukurs Programmierung</b> 2 SWS; 5 ECTS	<b>Proseminar Computerling.</b> 2 SWS; 5 ECTS	<b>Oberseminar Computerlinguistik</b> 2×1 SWS; 5 ECTS		
<b>Einführung Linguistik</b> 4 SWS; 5 ECTS; <i>Import Angl./Germ.</i>	<b>DH-Modul 1: Sprache &amp; Text</b> 4 SWS; 5 ECTS; <i>Import DGuS</i>			<b>Praktikum (extern / intern)</b> 150 h = 1 Monat Vollzeit; 5 SWS; unbenotet	
10 SWS 15 ECTS	6 SWS 10 ECTS	6 SWS 15 ECTS	5 SWS 12,5 ECTS	5 SWS 12,5 ECTS	0 SWS 5 ECTS



# Studienplan Computerlinguistik (Erstfach)



Semester 1 (WiSe)	Semester 2 (SoSe)	Semester 3 (WiSe)	Semester 4 (SoSe)	Semester 5 (WiSe)	Semester 6 (SoSe)
<b>VL Grundlagen der CL 1</b> 2 SWS; 2 ECTS	<b>VL Grundlagen der CL 2</b> 2 SWS; 2 ECTS	<b>VL Grundlagen der CL 3</b> 2 SWS; 3 ECTS	<b>Hauptseminar I</b> 2 SWS; 5 ECTS; <i>Hausarbeit</i>	<b>Hauptseminar III</b> 2 SWS; 5 ECTS; <i>praktisches Projekt</i>	
<b>Ü Grundlagen der CL 1</b> 2 SWS; 3 ECTS	<b>Ü Grundlagen der CL 2</b> 2 SWS; 3 ECTS	<b>Ü Grundlagen der CL 3</b> 2 SWS; 7 ECTS	<b>Hauptseminar II</b> 2 SWS; 5 ECTS; <i>mündl. Prüfung</i>	<b>Projektseminar</b> 2 SWS; 5 ECTS; <i>Teamprojekt</i>	
<b>S Grundkurs Programmierung</b> 2 SWS; 5 ECTS	<b>S Aufbaukurs Programmierung</b> 2 SWS; 5 ECTS	<b>Proseminar Computerling.</b> 2 SWS; 5 ECTS	<b>Oberseminar Computerlinguistik</b> 2×1 SWS; 5 ECTS		
<b>Einführung Linguistik</b> 4 SWS; 5 ECTS; <i>Import Angl./Germ.</i>	<b>DH-Modul 1: Sprache &amp; Text</b> 4 SWS; 5 ECTS; <i>Import DGuS</i>			<b>Praktikum (extern / intern)</b> 150 h = 1 Monat Vollzeit; 5 SWS; unbenotet	
	<b>Grundlagen der Informatik (Gdl)</b> 6 SWS; 7,5 ECTS; <i>Import Informatik</i>	<b>Wahlpflichtbereich Informatik</b> <i>ca. 6 SWS; 12,5 ECTS; Import Informatik (+ andere)</i>			
10 SWS 15 ECTS	12 SWS 17,5 ECTS	8 SWS 20 ECTS	9 SWS 20 ECTS	5 SWS 12,5 ECTS	0 SWS 5 ECTS

# Nach dem Studium

---

- Masterstudium / Promotion
  - Computerlinguistik
  - Sprachwissenschaft / Linguistik / Lexikographie (EMLex)
  - Digital Humanities, Computational Social Science
  - Angewandte Informatik
- Tätigkeitsbereiche in der Wirtschaft (→ Sprachtechnologie & KI)
  - Google, Microsoft, Facebook, Amazon, Siemens, ...
  - Text Mining, Information Retrieval, Marktforschung, ...
  - Lexikographie und Terminologie, Legal Tech
  - Spracherkennung und Sprachsynthese, Dialogsysteme
  - Computergestützter Sprachunterricht (CALL)
  - Viele Start-Up-Unternehmen im IT-Bereich und der KI suchen Computerlinguist:innen!

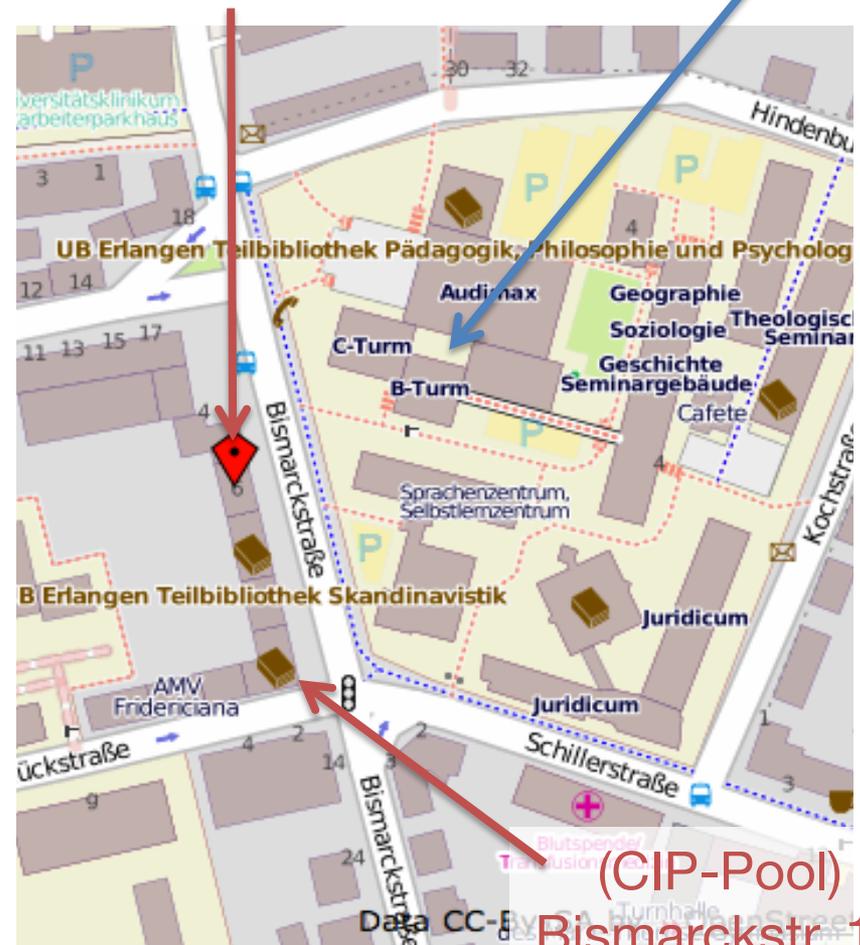
# Über uns

- Prof. Dr. Stephanie Evert  
[stephanie.evert@fau.de](mailto:stephanie.evert@fau.de)
- Dr. Besim Kabashi  
[besim.kabashi@fau.de](mailto:besim.kabashi@fau.de)
- Philipp Heinrich, M.Sc.
- Nathan Dykes, M.A.
- Bao Minh Doan Dang, M.A.  
[baominh.d.dang@fau.de](mailto:baominh.d.dang@fau.de)
- Dr. Alexander Piperski
- Mahdi Mantash, M.Sc.
- Naveed Unjum, M.Sc.
- Steffen Bothe, M.Sc.
- Anke Lutz  
[ccl-sekretariat@fau.de](mailto:ccl-sekretariat@fau.de)

[www.linguistik.phil.fau.de/team/](http://www.linguistik.phil.fau.de/team/)

Bismarckstr. 6  
Raum 4.000  
(Lehrstuhlbüro)

Philologien

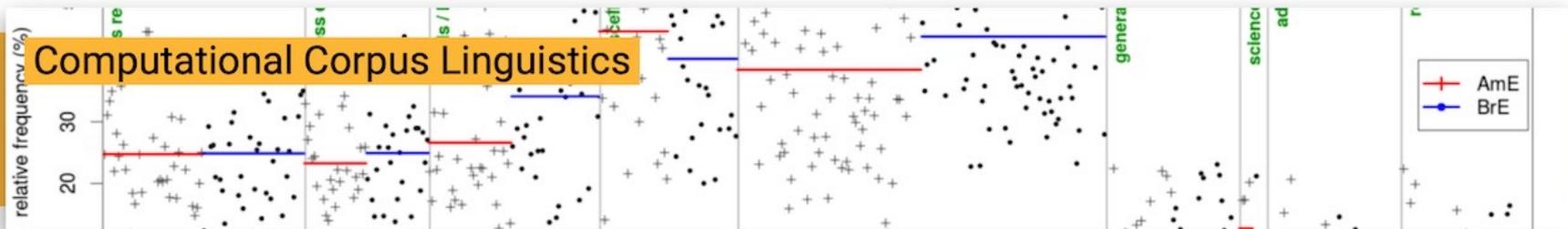


(CIP-Pool)

Bismarckstr. 12

Raum 0.320

Ganz wichtig: <https://www.linguistik.fau.de/>



## Welcome!

The [computational corpus linguistics group](#) carries out foundational methodological research on the quantitative analysis of large text corpora. Our [research](#) lies at the intersection of corpus linguistics, computational linguistics, and the digital humanities.

We offer a Bachelor's degree program in computational linguistics ("Linguistische Informatik"). Teaching usually takes place in German. We currently offer courses in the [teaching portal](#).

- Sehr kleiner Studiengang (~ 30 Studierende)
- Aktive Fachschaftsinitiative:

<https://www.linguistik.phil.fau.de/teaching/fsi-computerlinguistik/>

- FSI-Stammtisch

- Grundlagen- und Orientierungsstudium:

<https://www.ziwis.fau.de/grundlagen-und-orientierungsstudium/>

### Computational Corpus Linguistics

Lehrstuhl für Korpus- und Computerlinguistik  
Department Germanistik und Komparatistik

Bismarckstr. 6  
91054 Erlangen

+49913185-22426

info@linguistik.uni-erlangen.de

http://linguistik.fau.de

OCT 17 4:15 PM - 6:00 PM  
Einführungsveranstaltung BA  
Computerlinguistik

Subscription

# Stundenplan

siehe <https://www.linguistik.fau.de/teaching/lehrveranstaltungen/>

	Montag	Dienstag	Mittwoch	Donnerstag	Freitag
08–10				09:30–10:15 <b>Fragestunde zu GdCL 1</b> (S. Evert) Bismarckstr. 12, 0.320	
10–12			<b>Sprechstunde Prof. Evert</b> Bismarckstr. 6, 4.006 / Zoom (Anmeldung über StudOn)	<b>Übung GdCL 1: Traditionelle Verfahren</b> (S. Evert) Bismarckstr. 12, 0.320	<b>Vorlesung GdCL 3: Deep Learning</b> (S. Evert) Bismarckstr. 12, 0.320
12–14					<b>Übung GdCL 3: Deep Learning</b> (S. Evert) Bismarckstr. 12, 0.320
14–16		<b>Proseminar Computerlinguistik</b> (B. Kabashi) Bismarckstr. 12, 0.320	<b>HS Mehrwortausdrücke und Kollokationen aus NLP-Sicht</b> (B. Kabashi) Bismarckstr. 12, 0.320		<b>Praxisseminar Building the Söd★mizer</b> (S. Evert, F. Schäfer) Bismarckstr. 12, 0.320 + Sondertermine gem. StudOn
16–18	<b>Programmierung &amp; Infrastrukturen I</b> (B. M. Doan Dang) Bismarckstr. 12, 0.320		<b>Oberseminar Computerlinguistik</b> (B. Kabashi) Bismarckstr. 12, 0.320		

# Lehrangebot im WS 2024/25

- VL + Ü Grundlagen der Computerlinguistik 1
  - PS Programmierung & Infrastrukturen 1
  - Linguistische Grundkompetenzen (teilweise)
  - Modul Grundlagen der Informatik (Erstfach, optional)
- } **1. Sem.**
- VL + Ü Grundlagen der Computerlinguistik 3
  - PS Computerlinguistik
- } **3. Sem.**
- HS: Kollokationen & Mehrwortausdrücke
  - PraxisS: Building the Söd★mizer (mit Jap. + Germ.)
  - ggf. weitere HS als Importe (→ selbst anfragen)
- } **5. Sem.**
- Oberseminar Computerlinguistik
  - Wahlpflichtbereich Informatik
  - Schlüsselqualifikationen
- } **allgemein**

# Informationen für Erstsemester

## Grundlagen der Computerlinguistik I

- **Vorlesung:** Mo 14:00 – Do 9:00 (Zeitfenster)
    - Bereitstellung als Screencasts von 45–60 Minuten Dauer
    - Dozent: Prof. Dr. Stefan Evert 😊
  - **Fragestunde:** Do 9:30–10:15 (CIP-Pool)
    - Fragen und Diskussion zur Vorlesung
  - **Übung:** Do 10:15–11:45 (CIP-Pool)
    - Veranschaulichung mit Beispielen und interaktiven Übungen
    - wöchentliche Übungsaufgaben zur schriftlichen Bearbeitung
    - Dozentin: Prof. Dr. Stephanie Evert
  - **Tutorium:** n.V. (Tutorin: Anta Dzene)
- } **Beginn: Do 24.10.**

---

# Informationen für Erstsemester

## Programmierung & Infrastrukturen I

- **Grundkurs Programmierung:** Mo 16:15–17:45 (CIP-Pool)
  - erste Sitzung am Montag 21.10. (→ weitere Organisation)
  - Dozent: Bao Minh Doan Dang, M.Sc.
  
- **Tutorium:** n.V. (Tutorin: Laura Tatlow)

# Informationen für Erstsemester

## Importmodule

- **Grundlagen der germanistischen Linguistik 1 oder Basismodul II: Linguistics (A)**
  - falls zweites Fach nicht Germanistik, Anglistik oder Romanistik
  - jeweils mehrere Gruppen in verschiedenen Zeitfenstern
- **Grundlagen der Informatik (nur Erstfach)**
  - Vorlesung: Do 8:30–12:00 / „asynchron“ als Videoaufzeichnung
  - Fragestunde zu GdI: Di 10:15–11:45 (online)
  - Sprechstunden zu GdI: verschiedene Termine (Chat / virtuell)
  - alternativ: im 2. Semester belegen

## Nächste Schritte

- Alle LVen werden durch **StudOn-Kurse** begleitet
  - separater Kurs für jede Lehrveranstaltung in Campo
  - dort aktuelle Informationen, Materialien, Zugang Videos + Zoom
  - gilt unseres Wissens als offizielle Anmeldung
- Lehrveranstaltungen für 1. Semester in dieser Woche
  - Vorlesung GdCL1: ersten Screencast schauen → Diskussion am 24.10.
  - Fragestunde + Übung GdCL 1: beginnt erst **am Do 24.10.**
  - Linguistische Grundkompetenzen / GdI nicht vergessen!
- Prüfungsanmeldung über **Campo** nicht vergessen!
  - Anmeldezeitraum: **18.11.–08.12.2024**
- **Accounts** für CIP-Pool = IdM-Kennung + Passwort
- **Sprechstunde** Prof. Dr. Stephanie Evert: **hybrid**
  - Mi 10:00–12:00 (nur mit Voranmeldung per StudOn)
  - bis 04.11. Mo 9:00–11:00 nur per Zoom (mit Voranmeldung)