



Estsemestrigentag 2023: Computerlinguistik und Sprachtechnologie

First Semester Day 2023: Computational Linguistics and Language Technology

Department of Computational Linguistics
Institut für Computerlinguistik www.cl.uzh.ch

14. September 2023

- ▶ Wer seid ihr? Who are you?
- ▶ Wer sind wir? Who are we?
- ▶ Computerlinguistik (CL) an der UZH / Computational Linguistics at UZH
- ▶ BA/MA Studium der Computerlinguistik (CL) an der UZH / BA/MA Study of Computational Linguistics at UZH
- ▶ Fragen? Questions?

Tell us something about your background...

Why is natural language processing interesting for you?

Wer sind wir? Who are we?



- ▶ Prof. Dr. Martin Volk
(Head of Department of
Computational Linguistics)



- ▶ Prof. Dr. Rico Sennrich
(Machine Translation/Machine
Learning)



- ▶ Prof. Dr. Volker Dellwo
(Phonetics & Speech Sciences)



- ▶ Prof. Dr. Stephan Schmid
(Phonetics & Speech Sciences)



Who are we? Digital Linguistics / Computational Neuroscience of Speech & Hearing

- ▶ Prof. Dr. Lena Jäger
(Digital Linguistics)



- ▶ Dr. Natalie Giroud
(Computational Neuroscience of Speech & Hearing)



Who else teaches computational linguistics?

- ▶ Prof. Dr. Gerold Schneider



- ▶ Dr. Manfred Klenner
(Studienfachberatung/Study
Advisor)



Who else teaches computational linguistics?

▶ Dr. Simon Clematide



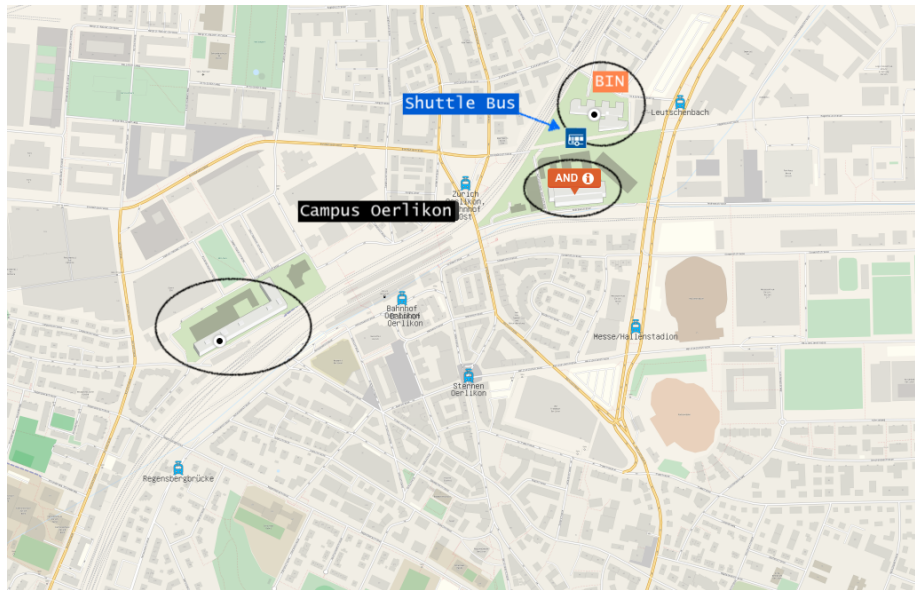
▶ Dr. Sarah Ebling



<http://www.cl.uzh.ch/de/people.html>

- ▶ Post-doctoral students in project positions
- ▶ Doctoral students in project positions
- ▶ Student employees
- ▶ *last but not least*: about 120 students in total

Where to find us? Campus Oerlikon AND 2nd Floor



- ▶ Most lectures are taught at Oerlikon Campus[▲] AND or BIN or City Campus.
- ▶ Shuttlebus City Campus[▲] \longleftrightarrow Oerlikon Campus:
- ▶ Mensa BIN[▲]: friendly service :-)

All important information about your studies, our Department and its research activities.

Tips for first-year students

<https://www.cl.uzh.ch/en/studies/studies-BA-MA/study-hints.html>

What do we do in computational linguistics?

Analyzing words

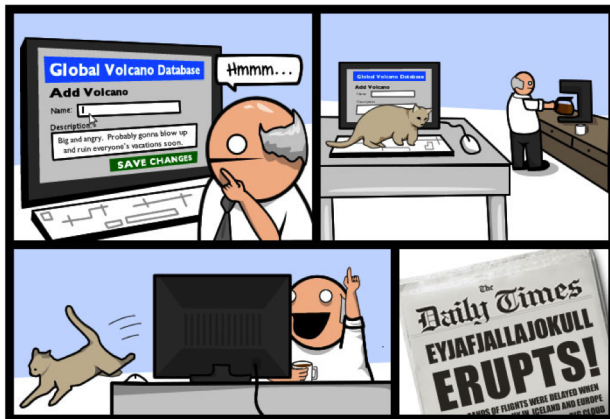


ein Zwei-fels-fall ???

What do we do in computational linguistics?

Recognition of names in texts

HOW TO NAME A VOLCANO



[HTTP://THEOATMEAL.COM](http://theoatmeal.com)

BY THE OATMEAL

What do we do in computational linguistics?

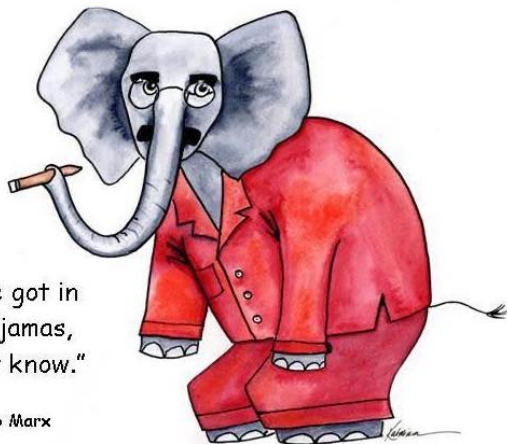
Syntactic analysis: How do words depend on each other?

“One morning I shot an elephant in my pajamas.

What do we do in computational linguistics?

Syntactic analysis: How do words depend on each other?

“One morning I shot an elephant in my pajamas.



How he got in
my pajamas,
I don't know."

Groucho Marx

What do we do in computational linguistics?

Machine translation



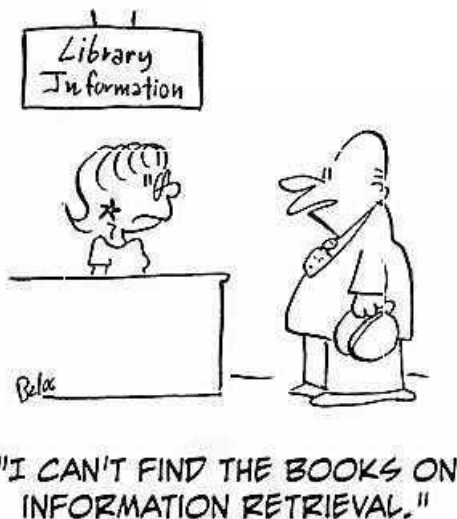
What do we do in computational linguistics?

Computation of word and sentence meaning (in context)



What do we do in computational linguistics?

Automatic extraction of information from texts



Processing spoken language

iPhone-Sprachassistentin

17. August 2015 22:39; Akt: 18.08.2015 10:29 

«Ruf 911 an!» – Siri rettet Sam Ray (18) das Leben

Mit Gehirnerschütterung, gebrochenen Rippen und Verbrennungen war ein 18-Jähriger unter einem Lastwagen eingeklemmt. Hilfe kam dank des Handy-Sprachassistenten.



ChatGPT is built on technology from CL



Hi ChatGPT, is there something in Computational Linguistics that you cannot do?

ChatGPT is built on technology from CL



Hi ChatGPT, is there something in Computational Linguistics that you cannot do?



Certainly! While I'm a product of advancements in computational linguistics and machine learning, there are several things I can't do or have limitations in:

1. **Understanding Context Like Humans:** I analyze the text you provide based on patterns in the data I was trained on. I don't have the ability to understand context in the same way humans do, especially if it's based on real-world experiences or nuanced cultural references.
2. **Emotion and Empathy:** I don't have feelings or emotions. I can recognize and generate text about emotions, but I don't "feel" them or truly understand them in the way humans do.
3. **Creativity and Original Thought:** While I can generate content based on patterns in my training data, I don't have original thoughts or creativity in the same way humans do. I can't come up with genuinely new ideas; I can only remix and recombine what I've been trained on.

Some selected research topics

- ▶ Machine translation of documents
- ▶ Multimodality (Sign language)
- ▶ Creation of multilingual text corpora
- ▶ Working with large language models (like the ones behind ChatGPT)
- ▶ Text mining (information extraction, sentiment analysis, hate speech detection)
- ▶ NLP for accessibility: Assistance systems (sign language)
- ▶ Using computers to understand how production and reception of speech actually work in human beings
- ▶ Using eyetracking devices to understand how we read

Study programs[▲] offered since HS 2019

- ▶ BA: Computerlinguistik und Sprachtechnologie (Major 120/ Minor 60)
- ▶ MA: Computational Linguistics and Language Technology (Major 90/ Minor 30)
- ▶ MA: Digital Linguistics (Minor 30)
- ▶ MA: Focus «Digital Linguistics» and «Phonetics» in the study program «Linguistics» (not suitable for CL students)

Our main goals and topics in study programs

Oriented towards applicable computational linguistics and language technology.

- ▶ Programming skills needed for practical work
- ▶ Machine Learning and statistical methods
- ▶ Machine Translation
- ▶ Text Mining
- ▶ Automatic syntax analysis
- ▶ Use of language technology tools in (corpus) linguistics
- ▶ Courses on analyzing speech and cognitive processing of language[▲] (eye tracking)...

The Rules of Your Study

Look at the pages of the Faculty of Arts

- ▶ Regulations, model curriculum:

<https://www.phil.uzh.ch/de/studium/rechtsgrundlagen.html>

- ▶ Information about deadlines and much more:

<https://www.phil.uzh.ch/de/studium/studentervices.html>

Please also check the prerequisites of each module!

Overview BA and Consecutive MA in Computational Linguistics and Language Technology

BACHELOR

Computerlinguistik &
Sprachtechnologie

Minor (60)/Major(120)

Bachelor Arbeit		15
Weitere Curriculare Module		0
Praxis der CL	0	6
Wissenschaftliche Vertiefung	0	6
Informatik	0	18
Kernbereich CL & SP	18	36
Einführung in die CL	30	30



MASTER

Computational Linguistics &
Language Technology

Minor (30)/Major(90)

Master Thesis		30
Other Curricular Modules		0
Computer Science	0	0
CL & LT in Practice	0	6
Scientific Specialization	6	6
Core Modules	18	24

Mandatory ECTS points per module group

Other MA Programs with “a bit of CL”

Bachelor

Master

Digital Linguistics (Minor)

BA
(no requirements)

MA

You already have a BA in Computational Linguistics or similar field?

Don't choose *Digital Linguistics* (neither minor nor the focus area). Choose *Computational Linguistics and Language Technology*.

Linguistics (Major 120)

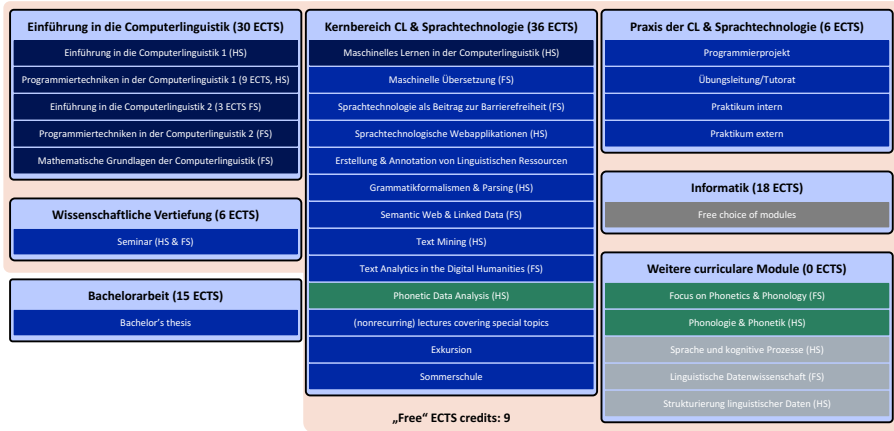
BA
(at least 21 ECTS in linguistics required)

MA Linguistics

Focus areas:

- Digital Linguistics
- Phonetics

BA Major 120 Computational Linguistics and Language Technology



Default: 6 ECTS points per lecture; pink background means overall free minimum ECTS points for covered module groups; dark blue means mandatory introductory modules

Model Curriculum BA Major CL and LT

Sem.	Modulgruppe	Modultitel	Modultyp	ECTS Credits	Σ/Sem.
1	Einführung in die Computerlinguistik	Einführung in die Computerlinguistik 1	P	6	15
		Programmiertechniken in der Computerlinguistik 1	P	9	
		Einführung in die Computerlinguistik 2	P	3	
2	Einführung in die Computerlinguistik	Programmiertechniken in der Computerlinguistik 2	P	6	15
		Mathematische Grundlagen der Computerlinguistik	P	6	
		Maschinelles Lernen in der Computerlinguistik	P	6	
3	Kernbereich Computerlinguistik und Sprachtechnologie Informatik	Module nach freier Wahl	WP/W	6	18
		Module nach freier Wahl	WP/W	6	
		Module nach freier Wahl	WP/W	6	
4	Kernbereich Computerlinguistik und Sprachtechnologie Wissenschaftliche Vertiefung	Module nach freier Wahl	WP/W	12	18
		Modul(e) nach freier Wahl	W	6	
5*	Kernbereich Computerlinguistik und Sprachtechnologie Informatik	Module nach freier Wahl	WP/W	6	24
		Module nach freier Wahl	WP/W	12	
		Module nach freier Wahl	WP	6	
6	Kernbereich Computerlinguistik und Sprachtechnologie Aus allen Modulgruppen des Programms	Module nach freier Wahl	WP	6	30
		Module nach freier Wahl	WP/W	9	
		Bachelorarbeit		15	120

https://www.phil.uzh.ch/de/studium/rechtsgrundlagen/rechtsgrundlagen_bachelor.html

P(flichtmodul) (mandatory modules): You have to pass them (2 attempts only)!

BA Minor 60 Computational Linguistics and Language Technology

BACHELOR (60 ECTS), *Computational Linguistics and Language Technology*

Einführung in die Computerlinguistik (30 ECTS) <ul style="list-style-type: none">Einführung in die Computerlinguistik 1 (HS)Programmiertechniken in der Computerlinguistik 1 (9 ECTS, HS)Einführung in die Computerlinguistik 2 (3 ECTS FS)Programmiertechniken in der Computerlinguistik 2 (FS)Mathematische Grundlagen der Computerlinguistik (FS)	Kernbereich CL & Sprachtechnologie (18 ECTS) <ul style="list-style-type: none">Maschinelles Lernen in der Computerlinguistik (HS)Maschinelle Übersetzung (FS)Sprachtechnologie als Beitrag zur Barrierefreiheit (FS)Sprachtechnologische Webapplikationen (HS)Erstellung & Annotation von Linguistischen RessourcenGrammatikformalismen & Parsing (HS)Semantic Web & Linked Data (FS)Text Mining (HS)Text Analytics in the Digital Humanities (FS)Phonetic Data Analysis (HS)(nonrecurring) lectures covering special topicsExkursionSommerschule <p>„Free“ ECTS credits: 12</p>	Praxis der CL & Sprachtechnologie (0 ECTS) <ul style="list-style-type: none">ProgrammierprojektÜbungsleitung/TutoratPraktikum internPraktikum extern
Wissenschaftliche Vertiefung (0 ECTS) <ul style="list-style-type: none">Seminar (HS & FS)		Informatik (0 ECTS) <ul style="list-style-type: none">Free choice of modules
		Weitere curriculare Module (0 ECTS) <ul style="list-style-type: none">Focus on Phonetics & Phonology (FS)Phonologie & Phonetik (HS)Sprache und kognitive Prozesse (HS)Linguistische Datenwissenschaft (FS)Strukturierung linguistischer Daten (HS)

Model Curriculum BA Minor 60 ECTS

Sem.	Modulgruppe	Modultitel	Modultyp	ECTS Credits	Σ/Sem.
1	Einführung in die Computerlinguistik	Einführung in die Computerlinguistik 1	P	6	15
		Programmiertechniken in der Computerlinguistik 1	P	9	
		Einführung in die Computerlinguistik 2	P	3	
2	Einführung in die Computerlinguistik	Programmiertechniken in der Computerlinguistik 2	P	6	15
		Mathematische Grundlagen der Computerlinguistik	P	6	
		Maschinelles Lernen in der Computerlinguistik	P	6	
3	Kernbereich Computerlinguistik und Sprachtechnologie	Module nach freier Wahl	WP/W	6	12
4	Kernbereich Computerlinguistik und Sprachtechnologie	Module nach freier Wahl	WP/W	12	12
5*	Aus allen Modulgruppen des Programms	Module nach freier Wahl	WP/W	6	6
6					0
					60

https://www.phil.uzh.ch/de/studium/rechtsgrundlagen/rechtsgrundlagen_bachelor.html

P(flichtmodul) (mandatory modules): You have to pass them (2 attempts only)!

Mandatory courses:

Introductory lectures “Einführung in die Computerlinguistik”

- ▶ **Programmiertechniken in der CL 1 (PCL 1) (9 ECTS)**
 - ▶ Martin Volk/Simon Clematide, Thu 14.00-15.45h, Room BIN-1-B-01
- ▶ **Einführung in die CL 1 (ECL 1) (6 ECTS)**
 - ▶ Simon Clematide/Martin Volk, Thu 16.15-18.00h, Room BIN-1-B-01

Tutorial sessions

- ▶ **Tutorial for Einführung in die CL 1**
 - ▶ Fri 12.15-13.45h, Room AND-4-06
- ▶ **Tutorial for Programmiertechniken in der CL 1**
 - ▶ Fri 14.00-15.45h, Room AND-4-06

Attending the tutorials is not obligatory, but handing in the written exercises is!

Taking the Informatics I course in parallel: Recommended for boosting your programming skills fast, but quite some work. . .

Informatik I (AINF1166)[▲]

- ▶ Programmierereinführung des Instituts für Informatik (auf Deutsch)
 - ▶ 6 ECTS
 - ▶ Module group “Informatik”
 - ▶ Python introduction with a focus on general software engineering

Retreat for First Semesters in SAC Hut Eseltritt



Two days in the Swiss mountains

Let's get to know each other: from Thursday 5 Oct. 15h to Saturday 7 Oct. 14h.

No tutorials in this week.

If you cannot join, videos of the lectures are available.

Introductory lectures in CL do not have place restrictions (Platzbeschränkung).

- ▶ Nice intro video
- ▶ Important: Different deadlines per faculty and departments!!!
- ▶ Standard module cancellation deadlines 10.10.2023:
https://www.students.uzh.ch/de/booking/fristen_hs.html

MA Major 90 Computational Linguistics and Language Technology ▲

Sem.	Modulgruppe	Modultitel	Modultyp	ECTS Credits	Σ/Sem.
1	Core Modules of Computational Linguistics and Language Technology	Machine Learning for Natural Language Processing 1	WP	6	18
		free choice of modules	WP/W	6	
	All groups of modules of the program	free choice of modules	WP/W	6	
2	Core Modules of Computational Linguistics and Language Technology	Machine Learning for Natural Language Processing 2	WP	6	18
		free choice of modules	WP/W	6	
	Scientific Specialization	free choice of modules	W	6	
3*	Core Modules of Computational Linguistics and Language Technology	free choice of modules	WP/W	6	24
	Computational Linguistics and Language Technology in Practice	free choice of modules	WP/W	6	
	All groups of modules of the program	free choice of modules	WP/W	6	
4		Master's Thesis (part 1)		6/30	30
	All groups of modules of the program	free choice of modules	WP/W	6	
		Master's Thesis (part 2)		24/30	
				90	

Source: https://www.phil.uzh.ch/de/studium/rechtsgrundlagen/rechtsgrundlagen_master.html

MA Minor 30 Computational Linguistics and Language Technology ▲

Sem.	Modulgruppe	Modultitel	Modultyp	ECTS Credits	Σ/Sem.
1	Core Modules of Computational Linguistics and Language Technology	Machine Learning for Natural Language Processing 1	WP	6	12
		free choice of modules	WP/W	6	
2	Core Modules of Computational Linguistics and Language Technology	Machine Learning for Natural Language Processing 2	WP	6	12
		free choice of modules	WP/W	6	
3	Scientific Specialization	free choice of modules	W	6	6
4*					0
					30

Source: https://www.phil.uzh.ch/de/studium/rechtsgrundlagen/rechtsgrundlagen_master.html

Mandatory course:

- ▶ Machine Learning for Natural Language Processing (6 ECTS)
 - ▶ Simon Clematide, Mon 14.00-15.45h
- ▶ Tutorial session
 - ▶ Mon 16.00-17.45h

Attending the tutorials is not obligatory, but handing in the written exercises is!

Laptop

- ▶ You need a laptop for attending the exercises (in programming courses also in the lecture).
- ▶ MacOS, Windows, Linux is fine. For BA students, more infos and software installation in the first week's tutorial session on Friday.
- ▶ In case you need a new machine: <https://www.projektneptun.ch>

All modules/lectures are there <https://lms.uzh.ch>

- ▶ Detailed program
- ▶ Slides, books
- ▶ Handing out and submitting exercises
- ▶ Forums and contact via e-mail...

As soon as you book a course module, you have access to the course on OLAT.

Do not use a very strange user name on OLAT on the first login.... No change possible!

Documentation for students[▲]

Manfred Klenner (Office: AND-2-30): contact him by e-mail:
`study@cl.uzh.ch`

- ▶ Follow us on: X https://twitter.com/cl_uzh ; Instagram https://www.instagram.com/uzh_cl
- ▶ Our Department maintains a moderated low-traffic mailing list that serves as a forum for computational linguistics in Zurich. Information about local events, announcements of talks and job opportunities are posted there as well.
cl-list@lists.ifi.uzh.ch
- ▶ Web interface for registering:
<https://lists.ifi.uzh.ch/listinfo/cl-list>
- ▶ Registering/following now is a good idea:-)

Zentrale Informatik – Studierende

Passwörter/Account • Arbeitsplatz/Kollaboration • Software/E-Learning • IT-Kurse • Studium • IT-Sicherheit • Support



Passwörter + Account →

- Passwort vergessen
- Passwörter verwalten
- Konto-Sperrung lösen
- UZH-E-Mail-Adresse ändern
- Getting Started



Arbeitsplatz + Kollaboration →

- Office 365 / Teams
- E-Mail
- Daten verwalten
- Online zusammenarbeiten
- UZH Print Plus
- VPN
- Zugang zum Netz
- mehr...



Software + E-Learning →

- OLAT
- Software fürs Studium
- Angebote von Drittanbietern
- Microsoft
- Mobile Apps



IT-Kurse + Weiterbildung →

- Kursprogramm
- Lunchveranstaltung



Studium + Organisation →

- Modulbuchung
- Studierendenportal
- UZH-App
- Semestereinschreibung
- Leistungsübersicht
- Vorlesungsverzeichnis



IT-Sicherheit →

- Informationssicherheit und Datenschutz
- IT-Sicherheit Management
- Guidelines + Sicherheitsregeln

<https://www.zi.uzh.ch/de/students.html>

Fragen? Questions?

Do not hesitate, ask your questions . . .

Klein, aber oho! / Small, but multifaceted

- ▶ Gatherings
- ▶ Spielabende
- ▶ Organisation der Teilnahme an TaCoS - Tagung der Computerlinguistik-Studierenden
- ▶ Staffel an der SOLA-Stafette
- ▶ Volley-Night
- ▶ Kurse (LaTeX, ...)
- ▶ Erstsemestrigenzvieri

Please join:-) We need you