fSchedule Spring Semester 2022

Doctoral Programme: Managing Languages, Arguments and Narratives in the Datafied Society

Date	Teacher	Mode	Topic	ECTS
March 3-4	Qi Yu	Online	Python for Linguists	1.5
ca. March 22-25	Cerstin Mahlow Maren Runte	Online	Workshop Annotation	3
April 11 April 25	Tanja Samardzic	Online	Designing empirical studies in linguistics	1.5
April 29 May 06	Roger Flühler Vanessa Klaas Yvonne Klein Nicole Krüger Cara Seitz	Online	Open Science Publication	0.5
May 30 June 01	Marcin Lewinski		Speech acts: From theoretical basics to argumentative polylogues online (course descriptions coming soon)	1.5
May 04-05		Live, Winterthur	Workshop: Managing Languages, Arguments and Narratives in the Datafied Society (course descriptions coming soon)	(1)
May 20-21	Chris Reed		Argumentation Annotation and Argumentation Mining (course descriptions coming soon)	1.5
June 09-12	Eva Kuske	Casa Santo Stefano, Miglieglia	Writing Retreat (sign up directly via kusk@zhaw.ch)	n/a

Phd seminar: “Python for Linguists”

In the context of *Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society*

ECTS: 1.5

Program:

- March 3, 2022, 09:00-17:15 (inclusive of breaks and hands-on sessions)
- March 4, 2022, 09:00-17:15 (inclusive of breaks and hands-on sessions)

- Place: The seminar will take place online via Zoom.
- Lecturer: Qi Yu (Graduate School of the Social and Behavioural Sciences & Department of Linguistics, University of Konstanz)

Topic, focus:
This is an introductory seminar for programming with Python, with a specific focus on its application in processing large-scale text data. The seminar is specifically intended for participants with little to no experience with programming, i.e., no previous knowledge of programming is required.

The first part of the seminar will focus on the basic programming with Python and the concept of object-oriented programming. The second part will introduce the core components of computational analysis of text data and demonstrate the analysis of large-scale corpus with Python using the library NLTK.

Main value added from a theoretical and methodological perspective:
Data-driven approaches for linguistic research has been booming in the last decades owing to the availability of large-scale corpora. Python, along with its numerous handy libraries for natural language processing (NLP), has been one of the most used programming languages in the computational linguistics, NLP and corpus linguistics community. In this seminar, participants will learn the basic Python programming skills and commonly used approaches for processing text data. These approaches can be transferred to data-driven analysis of theoretical linguistic research questions.

Main value added from a practical perspective:
The basic programming skills and the data-driven thinking learned from this seminar can be widely applied to real-life research questions: First, they can be used to automatize the manipulation of large-scale text data for which manual processing would be cumbersome, e.g., corpus cleaning. Second, they can be used for corpora analysis which can provide valuable empirical evidence for linguistic research.

Learning objective: knowledge. Participants will learn 1) the basics of Python programming, 2) the concept of object-oriented programming, and 3) commonly used methods for computational analysis of text data.

Learning objective: skills, practices. Participants will be able to 1) do basic programming with Python, 2) use Python to automatize the processing of large-scale text data, and 3) use Python to conduct basic corpus analysis.

Learning objective: researcher attitudes. Participants will practice their logical thinking skills and design thinking skills.

Evaluation: Assignment which will be assessed with a pass/ fail evaluation.

Bio notes/ Profiles

Qi Yu is a PhD candidate in computational linguistics and research fellow of the Cluster of Excellence “Politics of Inequality”, University of Konstanz. Her main area of interest covers broadly NLP, computational linguistics and machine learning. Qi’s PhD research topic is the automated detection of linguistic cues for framing in political discourses using NLP approaches. In her research, Python has been used as an essential tool, with which she manipulates large-scale corpora, develops rule-based NLP pipelines for automated annotation of linguistic features, and implements machine learning-based approaches for date-driven corpora analysis, such as topic modelling, word embeddings and text classification.

PhD seminar: "Annotation and Modeling of Textual Data: Concepts and Tools"

In the context of *Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society*

ECTS: 3

Program:

- Tue, March 22, 2022, 09:00 – 13:00 (modeling and annotation as concepts, including breaks)
- Wed, March 23, 2022, 09:00 – 12:30 (modeling and annotation as scientific practices, including breaks)
- Thu, March 24, 2022, 08:30 – 12:30 (in-depth introduction to INCEpTION, including breaks)
- Fri, March 25, 2022, 13:30 – 17:30 (hands-on session with your own data, including breaks)
- Thu, April 28, 2022, 09:00 – 12:00 (follow-up, including breaks)

- Place: The seminar will take place online via Zoom.
- Lecturer: Cerstin Mahlow and Maren Runte (Institute of Language Competence, School of Applied Linguistics, ZHAW), Irina Bigoulaeva (INCEpTION)

Topic, focus:
The investigation of specific phenomena in written texts often requires compiling task-specific corpora and creating or extending task-specific knowledge bases. Researchers profit from automatic, semi-automatic, and manual annotation of linguistic structures and semantic or argumentative information. Presently, researchers require a broad range of skills and tools to address such annotation tasks, especially when annotation is an essential part of the research process and used for data-driven investigation. Explorative manual annotation serves as starting point for the development of systematic classifications and annotation schemes. Multi-layer and concurrent annotation of elements and linking between elements and layers help to gain deeper insights and support modeling of complex phenomena.

Manual annotations for a small set of texts can also serve as input for machine learning algorithms to allow for systematic annotation of larger collections to test the validity of the classification system.

The first part of the seminar will focus on the modeling and annotation as concepts and scientific practices. The second part is an in-depth introduction to the annotation tool INCEpTION and participants will then work with INCEpTION using their own data.

Main value added from a theoretical and methodological perspective:
Multi-layer annotation is a prerequisite for comprehensive and hermeneutic modeling of large text collections. The combination of automatic basic linguistic annotation (e.g., Part-of-Speech) and manual data-driven annotation (i.e., marking, labeling, and linking words and phrases), allows researchers to develop scientifically sound perspectives. Explicit machine-readable annotation is also a prerequisite for reproducible research.

Main value added from a practical perspective:
Participants will get to know INCEpTION, a text-annotation environment for various kinds of annotation tasks on written text. INCEpTION is a state-of-the-art open-source web application. It allows several users to work on the same annotation project and it can handle several annotation projects at a time. It provides a recommender system to help you create annotations faster and easier. Participants will be able to work with their own data in INCEpTION with support and help by the developers.

Learning objective: knowledge. Participants will learn about 1) the concept of modeling, 2) annotation as concept and scientific practice, 3) the concept of entities and relations in INCEpTION and understand the general workflow of annotation, 4) the concept of multi-layer annotation.

Learning objective: skills, practices. Participants will be able to 1) use INCEpTION to upload raw and pre-annotated text collections and apply standard automatic linguistic annotation (e.g., POS, NER), 2) use INCEpTION to model the phenomena in their own research using multiple annotation layers, i.e., existing ones and newly created ones

Learning objective: researcher attitudes. Participants will practice their computational thinking skills and reflect on their scientific methods.

Evaluation: Assignment which will be assessed with a pass/ fail evaluation.

Bio notes/ Profiles

Cerstin Mahlow is Professor of Digital Linguistics and Writing Research at the School of Applied Linguistics at ZHAW. She holds a doctorate in computational linguistics from the University of Zurich and a Magistra Artium in Computational Linguistics, Spanish Philology, and Political Sciences from Friedrich-Alexander-Universität Erlangen-Nürnberg (Germany). Her research interests lie at the intersection of writing research, document engineering, natural language processing, pedagogy, and human-computer interaction. As computational linguist, Cerstin Mahlow focuses on systematic modeling of empirical data from writing processes.

Maren Runte is Lecturer at the School of Applied Linguistics at ZHAW. She holds a doctorate in Linguistics from the University Duisburg-Essen (Germany) and a Magistra Artium in German Linguistics from the Ruhr-Universität Bochum (Germany). Maren Runte has many years of research experience in the fields of lexicography and corpus linguistics. In her research on dictionary use, she shows how evaluations of large amounts of language data can be presented in a user-adequate way. In her teaching, she combines corpus linguistic evaluations with discourse linguistic questions.

Irina Bigoulaeva is a member of the INCEpTION team (https://inception-project.github.io) at the UKP Lab at TU Darmstadt. She holds an M.Sc. degree in Computational Linguistics from the Ludwig-Maximilians-Universität München (Germany), and additionally has a background in general theoretical linguistics. As a current PhD student, she is interested in legal document analysis and data annotation tasks.

Phd seminar: "Designing empirical studies in linguistics"

In the context of *Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society*

ECTS: 1.5

Program:

April 11, 2022, 09.00-12.00 / 13.00-17.00
April 25, 2022, 09.00-12.00 / 13.00-17.00

Dr. **Tanja Samardžić** (University of Zurich)

The seminar will take place online via Ms Teams

Topic, focus:
The goal of this seminar is to provide PhD students in Applied Linguistics with the essential training in the methodology of empirical research. The focus will be on forward-thinking skills, including problem definition, identifying the intermediate and overall goals, planning and preparing empirical studies. These are the key steps that take place at the onset of an empirical study and largely determine its success. While the outcome of a research process cannot be known in advance, a good study design limits the risks, ensures reproducibility and increases the chance of success. It also facilitates scientific communication and reach-out by ensuring clearly formulated research goals and the content of the study. In the seminar, the students will become familiar with the general methodological framework presented against a background of good and bad examples from the linguistic literature. They will have the opportunity to evaluate the design of their own research projects: whether a given goal can be reached with empirical methodology, what data sources can be used, what analyses are suitable. This experience should lead to insights that can be generalised beyond the current projects to all subsequent studies.

Main value added from a theoretical and methodological perspective:
Linguistic research currently encompasses a wide range of methodological frameworks, from introspective inquiry to theory-neutral data descriptions. None of these research frameworks has the status of the standard framework that would be taught explicitly as part of the undergraduate curriculum. As a result, PhD students in linguistics need to acquire methodological skills in an efficient training tailored to their specific needs. This seminar covers only one part of the methodological spectrum, but it will allow the students to distinguish between different options and conduct their own studies with a better awareness of the fundamental methodological steps that are common to many specific approaches.

Main value added from a practical perspective:
The seminar will include practical work with examples from the linguistic literature and the students' own work. The students will perform exercises in formulating study designs leading to new practical skills.

Learning objective: knowledge. Students will know the logical structure of empirical studies in general and applied to linguistic research.

Learning objective: skills, practices. Students will be able to formulate feasible empirical studies, select appropriate data and analyses and will be able to assess and limit research risks.

Learning objective: researcher attitudes. Students will practice their forward thinking skills, design thinking skills, and problem solving skills.

Evaluation: Submitted preparation for the seminar, in-course assignments which will be assessed with a pass/ fail evaluation.

Bio notes/ Profiles

Tanja Samardžić is a lecturer in Computational linguistics and a Group leader at the University of Zurich. Her background combines the theory of language with machine learning. Her research is about developing computational text processing methods and using them to test theoretical hypotheses on how language actually works. Tanja holds a PhD in Computational linguistics from the University of Geneva, where she studied in the group Computational Learning and Computational Linguistics (CLCL). She is committed to promoting and facilitating the use of computational
approaches in the study of language.

Phd seminar: "Open Science in Linguistics: Why is it relevant for me?"

In the context of *Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society*

ECTS: 0.5

Program:

- April 29th, 2022, 09:00-12:00
- May 06th, 2022, 09:00-12:00

- Place: The seminar will take place online via Zoom.
- Lecturers:
  - Dr. Vanessa Klaas (ZHAW Research and Development Unit, Research Data Management),
  - Nicole Krüger (ZHAW University Library, Open Access & OER),
  - Yvonne Klein (ZHAW University Library, Open Access & OER),
  - Roger Flühler (ZHAW University Library, OER)

Topic, focus:
Open Science consists of several movements like open research data, open access for publications, and OER – open educational resources. In its core it aims at:

- making scientific knowledge openly available, accessible and reusable for everyone → for researchers worldwide, to teachers and to the general public;
- sharing of information and data for enhancing the effectiveness and productivity of research and teaching (e.g. in urgent questions of climate change, COVID 19, democracies in times of fake news and conspiracy theories);
- enhancing the quality of research and teaching materials by granting transparency and replicability, as well as peer feedback.

Furthermore: Many funding agencies require the publication of project findings under an open license / as open access or open data.

Main value added from a theoretical and methodological perspective:
Stated by the UNESCO[1], "…increased openness should be promoted in all stages of the scientific endeavour, with the view to reinforcing the strength and rigour of scientific results, enhancing the societal impact of science and increasing the capacity of society as a whole to solve complex interconnected problems.
Increased openness leads to increased transparency and trust in scientific information and reinforces the fundamental feature of science as a distinct form of knowledge based on evidence and tested against reality, logic and the scrutiny of scientific peers."
Learn about the idea of openness in research and which impact it has on research practices and research communication.

Main value added from a practical perspective:
In this workshop, you will learn and discuss how making your research transparent and open can be achieved and managed.
In detail:
Open Research Data: Research Data Management (RDM) is relevant in all phases of the research life cycle, from the research question to publication and archiving. After a short overview on the importance of RDM and its relevance in the context of Open Science, you will learn about details on practical RDM in the context of text corpora and text mining under consideration of possibly existing confidentiality terms.

[1] UNESCO Recommendation on Open Science, 2021, CC BY-SA 3.0 IGO

Open Access: There is a proven citation advantage for publications in open access. Learn how to fund your OA publication; what options there are to publish for free in open access and how you can find high-quality open access journals and avoid predatory publishers.
OER – Open Educational Resources: In teaching, social media or conference talks it might be useful to use synergy effects and include materials of other experts. However, copyright often prohibits the re-use. Learn how to identify open licensed content and how you can share your own materials for greater visibility under a license like Creative Commons.

Learning objective: knowledge. Participants will learn 1) the advantages of open science in terms of transparency, citation advantages, and synergy effects in research, 2) the importance of planning the data management early in the research process, 3) which data may be published and which is confidential, 4) the benefits of open access, 5) funding options for open access publications and 6) the benefits of open licenses like Creative Commons for the re-use of publications in researcher networks and of teaching materials.

Learning objective: skills, practices. Participants will be able to 1) develop a data management plan, 2) publish their research data for larger transparency and citation advantages, 3) evaluate the quality of open access journals for their publication, 4) acquire funding for their open access publication, 4) find and re-use materials under Creative Commons licenses, and 5) publish materials under a Creative Commons license.

Learning objective: researcher attitudes. Participants will learn how to be part of the culture of sharing and how they can contribute to a development in research as a practice. Open Science allows collaboration, synergy effects, and transparency, it helps to reach a larger audience even beyond your own scientific community.

Evaluation: Assignment which will be assessed with a pass/ fail evaluation.

Bio notes/ Profiles
Dr. Vanessa Christina Klaas received their diploma degree in computer science from the Ludwig-Maximilians-University Munich, Germany, in 2007. After working as a system engineer in the Automotive industry and in the development of medical devices, they pursued their PhD at the Institute for Electronics, ETH Zurich, from where they graduated with the Dr. sc. ETH Zurich degree in 2018. They joined ZHAW in October 2020. As a member of ZHAW Services Research Data, they support researchers in diverse topics concerning research data management, such as data collection (e.g. REDCap), data processing (e.g. Python and R), contribute to community events such as the ZHAW Open Science Cafe, and represent the ZHAW as a “local node” in the Swiss Reproducibility Network.
ZHAW profile and contact

Roger Flühler is a historian and linguist.
After studying history and Latin philology at the University of Zurich, he completed the teaching diploma for Matura schools and taught at various high schools in the cantons of Zurich and Zug. Since July 2021 he is part of the OER Competence Center at the ZHAW University Library.
ZHAW profile & contact

Yvonne Klein has worked in e-learning and as a project manager at various universities in England and Switzerland. She studied at the University of East London, received her postgraduate qualification from the Open University and is a Fellow of the Higher Education Academy (FHEA), UK. Since March 2021, she has joined the ZHAW, Zurich University of Applied Sciences, as a specialist for Open Educational Resources and Open Access.
ZHAW profile & contact

Nicole Krüger is a librarian and MA in philosophy.
After her studies she was a teaching librarian at ZBW – Information Centre for Economics with a teaching assignment at Kiel University. Since August 2020 she is Open Science Specialist in the field of Open Educational Resources (OER) and Open Access at the ZHAW University Library.
ZHAW profile & contact

PhD seminar: "Speech acts: From theoretical basics to argumentative polylogues online"

In the context of *Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society*

ECTS: 1.5

Program:

- May 30, 2022, 09:30-17:30 (inclusive of breaks and hands-on sessions)
- May 31, 2022, 09:30-17:30 (inclusive of breaks and hands-on sessions)
- June 1, 2022, 09:30-17:30 (inclusive of breaks and hands-on sessions)

- Place:
- Lecturer: Marcin Lewiński (ArgLab, Nova Institute of Philosophy, Nova University Lisbon, Portugal)

Topic, focus:
This seminar introduces speech act theory as a powerful framework to understand language across a variety of contexts, including online discussions. Developed within the philosophy of language and widely applied in linguistics, communication studies, and computer science, speech act theory treats uses of language as complex acts fulfilling various functions in broader social activities. Argumentation is one of these functions that will be the focus of the seminar.

The seminar will be divided into three parts. The first part will introduce key concepts of speech act theory, as originally developed by J.L. Austin and J.R. Searle. The second part will focus on various approaches to grasp argumentation as a speech act, notably pragma-dialectics and normative pragmatics within argumentation theory. The third part will discuss recent developments aimed at analyzing and evaluating argumentative speech acts in the context of multi-party online polylogues.

Main value added from a theoretical and methodological perspective:
Speech act theory, which is enjoying nothing short of a revival today, provides a particularly useful framework for studying argumentation in its communicative complexity. It allows to understand intricate communicative strategies arguers are bound to follow in the context of new, digital forms of interaction. Given its normative underpinnings, the theory also affords a setting for studying the variety and dynamics of norms governing today's argumentative practices. The qualitative pragmatic methodology of speech act theory is capable of capturing fine details of communicative exchanges and, as such, provides a particularly useful complement to the quantitative methods of corpus linguistics and argument mining.

Main value added from a practical perspective:
One important objective of the seminar is to demonstrate that speech act theory, apart from its theoretical sophistication, offers a practical toolkit for understanding what is at stake in today's online-driven communication. Attention to the various elements of speech acts – notably their illocutionary forces, perlocutionary effects, and conditions of felicity – enhances critical skills of communication and argumentation analysis that can be readily applied to real-life research questions.

Learning objective: knowledge. Participants will learn: 1) the basic concepts of speech act theory and their broader philosophical context; 2) various ways to understand argumentation as a complex, situated speech act; and 3) the most recent results of research on speech acts in the context of polylogues, especially in online contexts.

Learning objective: skills, practices. Participants will be able to: 1) understand interactions in terms exchanges of speech acts; 2) apply the framework of speech act theory to the study of argumentation;

and 3) use qualitative pragmatic methods to analyze and evaluate instances of complex online communication.

Learning objective: researcher attitudes. Participants will enhance their logical thinking skills and practice skills of philosophically based, rigorous analysis of language-in-use.

Evaluation: Work in groups and individual assignment which will be assessed with a pass / fail evaluation.

Bio notes/ Profiles

Marcin Lewiński is an Assistant Professor and the Chair of the Reasoning and Argumentation Lab (ArgLab) at the Nova Institute of Philosophy, Nova University Lisbon, Portugal. He completed his PhD at the University of Amsterdam (2010), and from 2010 to 2016 worked as a post-doctoral fellow of the Portuguese Foundation for Science and Technology (FCT), before becoming an assistant professor in 2016. His work focuses on the basic issues in the philosophy of language and argumentation theory such as rationality of everyday conversations, social and strategic aspects of speech acts, practical reasoning, pragmatic meaning, fallacies. He is currently leading the Horizon 2020 COST Action project CA17132*, European Network for Argumentation and Public Policy Analysis (APPLY)* (2018-2023), where philosophical and linguistic concepts and methods are applied to a critical study of public argument.

Writing Retreat

In the context of *Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society*

ECTS: n/a

Program:

- June 9, 2022, afternoon - June 12, 2022, afternoon
- Place: Casa Santo Stefano, Miglieglia
- Coach: Dr. Eva Kuske (Zurich University of Applied Sciences)

Topic, focus:
The writing retreat at Casa Santo Stefano offers you the chance to spend a few days focusing solely on your academic writing, whether you are working on your PhD monograph or an article. Sign up directly via kusk@zhaw.ch, stating preferred room option.

Rooms: (space limited for approx. 10 people, available on first come, first serve basis)

a) Shared room (2 people) with shower and restroom
b) Single room with shower and restroom in the room
c) Single room with shower and restroom on the corridor

Amenities:

- Breakfast including homemade bread
- Cappuccino, espresso, latte macchiato and teas available during breaks
- Seasonal fruits avaiable during the day, homemade cake available in the afternoons
- Free parking available behind Monte Lema gondola
- Lunches and dinners at nearby restaurants

Preliminary Schedule:

Thursday

- Arrival
- Afternoon Writing session
- Dinner

Friday, Saturday

- Optional early bird writing session: A writing prompt and time to write before breakfast
- Breakfast
- Morning Writing session (including writing prompts and optional coaching)
- Lunch
- Afternoon Writing session (including writing prompts and optional coaching)
- Dinner
- Optional night owl writing session

Sunday

- Checkout
- Optional Writing session in the seminar room

Bio notes/ Profiles

Dr. Eva Kuske is a research associate at the Zurich University of Applied Sciences and manages the Doctoral Programme in Applied Linguistics *Managing Languages, Arguments and Narratives in the Datafied Society.* Since completing her MA and PhD in English languages and literature, she has continuously worked in university management, and she has been teaching courses on academic English writing for Bachelor and PhD students at the University of Bern and at the University of Applied Sciences in Winterthur.