

## **Analyzing Language in Context: Practical Tools and Methods in Corpus Linguistics**

**In the context of: [Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society \(LAND\)](#)**

### **Details**

- Dates: 10-11 April 2025 (from 9am to 4.45pm, including breaks)
- Place: USI Università della Svizzera Italiana, Lugano, Room TBD
- Course lecturer:  
Dr. Émeline Pierre (SNSF Postdoctoral researcher at Università della Svizzera Italiana)
- 1,5 ECTS

### **Description**

Over the past decades, linguistics has experienced a fundamental methodological turning point and has become more data-driven than ever (Zufferey 2020). Therefore, mastering corpus linguistics – i.e. the study of language through large collections of naturally occurring spoken and/or written productions – is now essential for every researcher working with big textual datasets. This includes the ability to collect, analyze, and interpret these large-scale datasets using quantitative approaches.

Corpus methodologies help scholars to move beyond intuition and provide empirically grounded and statistically robust insights, whether the researcher works in discourse analysis, argumentation, language acquisition, translation studies, across any text genres and domains.

### **In this PhD course, students will:**

- Master the theoretical and practical aspects of corpus compilation and analysis;
- Critically evaluate the ethical and methodological issues inherent in corpus-based studies;
- Start to develop proficiency in using statistical methods and computational tools to conduct large-scale, reproducible linguistic research (e.g., AntConc, SketchEngine);
- Be able to apply corpus methodologies to their own research

This course focuses on advanced corpus linguistics techniques and guides PhD students through best practices in corpus data collection, analysis, and interpretation to examine language patterns and linguistic phenomena across various contexts. Hands-on sessions

with software tools will offer the participants the opportunity to analyze real-world linguistic data and directly apply the newly acquired research skills and methodologies to their own project. Prior to the course, students will be given some mandatory preparatory readings. After the course, students will be required to hand in a short study summarizing a sample quantitative analysis of data of their choice.

### Provisional timetable

April, 10		April, 11	
9-10.30	<i>Introduction: Corpus linguistics</i>	9-10.30	<i>The role of statistical analysis and inter-rater agreement</i>
10.30-11	Coffee break	10.30-11	Coffee break
11-12.30	- <i>Corpus compilation/ ethical aspects</i> - <i>Research designs and annotation frameworks</i>	11-12.30	<i>Data session</i>
12.30-13.30	Lunch break	12.30-13.30	Lunch break
13.30-15	<i>Software tools and applications (AntConc, SketchEngine)</i>	13.30-15	<i>Final discussion: Implications and challenges</i>
15.15-16.45	<i>Data session</i>		

### ECTS recognition and evaluation

These are the requirements for earning the 1,5 ECTS:

- Preparatory reading(s)
- Presence (no less than 80%)
- Active participation throughout the course
- Short study case, to hand in after the course. Evaluation: pass/fail.

More detailed information (e.g. preparatory materials, final schedule, short study case) will be provided at a later time.

### Enrolment

The deadline for enrolment is 31<sup>st</sup> of March 2025. PhD students from USI can enroll via the dedicated platform. PhD students from other institutions can enroll by sending an e-mail to: narjes.sheikh.asadi@usi.ch.

Participation in the course is free and open to all people interested. The LAND Doctoral Programme provides financial support to PhD students participating in the Programme and to other PhD students based in Switzerland.

For any questions please contact emeline.pierre@usi.ch or narjes.sheikh.asadi@usi.ch.