

LIN4071/5075 Introduction to Corpus Linguistics (Fall 2025)

Instructor: Stefanie Wulff, Ph.D.

Office hours: Tuesdays 12-1:30pm or by appointment (4131D Turlington Hall)

Email: swulff@ufl.edu

Phone: 352-294-7455

Class Meeting Times

This is a face-to-face class, which means that we will meet regularly on Tuesdays 10:40-11:30am and Thursdays 10:40am-12:35pm in Matherly 0102.

Course Description

This course introduces corpus linguistics as a scientific research method, that is, a way to retrieve language data from corpora and analyze that data qualitatively and quantitatively. Specifically, we will cover how to form hypotheses, and operationalize these hypotheses in a way that allows a corpus-based investigation of our hypotheses; how to retrieve pertinent language data from corpora using a software tool called SketchEngine; how to annotate the data for relevant variables under investigation; and how to quantify our hypotheses and, where applicable, test our hypotheses for statistical significance. We will discuss this research process for a variety of linguistic phenomena and corresponding corpus-linguistic data types, including analyses of lexical, grammatical, and morphological constructions, analyses of different text types, as well as analyses of metaphorical language.

Course Objectives

By the end of this course, you will have acquired the ability to formulate research hypotheses, query corpus data, annotate corpus data, and (statistically) analyze the data. You will have gained competence in using the SketchEngine corpus analysis software. You will have gained an overview of contemporary research in the field of corpus linguistics, including its application in a variety of settings such as foundational linguistic research, sociolinguistics, and language acquisition. Graduate students will additionally gain experience with conducting their own corpus-based research project.

Course Credit Requirements and Grading

Undergraduate students

12 Home Quizzes (6 points each)	72 points
3 Discussion Boards (4 points each)	12 points
Research Summary	10 points
Lecture Attendance	6 points

Graduate students

12 Home Quizzes (6 points each)	72 points
Research Project Proposal	5 points
Research Project Presentation	10 points
Research Project Writeup	13 points

Home Quizzes (undergraduate and graduate students): 72% (6 points each)

Both undergraduate and graduate students will complete 12 quizzes worth 6% of your final grade each. Quizzes are take-home assignments and are open book and notes. Home quizzes are submitted by the due date noted on the schedule below on CANVAS.

Discussion Boards (undergraduate students only): 12%

Undergraduate students will complete three discussion boards that comment on a video lecture. All materials are available on Canvas, and specific instructions for each discussion board are provided on Canvas. Responses are due on the dates shown in the course schedule below and to be submitted on Canvas.

Lecture Attendance (undergraduate students only): 6%

You should attend two language-related lectures, workshops, or conference talks in the Fall 2025 semester (in the Department of Linguistics, World Languages and Literatures, Speech and Hearing Sciences, etc.). Proof of attendance (for example in the form of a screen shot or picture) has to be submitted to Dr. Wulff via email (swulff@ufl.edu) by 9 December 2025.

Research Summary (undergraduate students only): 10%

Undergraduate students only will be responsible for a research summary worth 10 points. Students will select a topic in corpus linguistics of their choosing and summarize no less than three research articles on that topic in writing. The research summary should close with a critical discussion of the research summarized such that students point out any gaps in the existing research and make suggestions for how they would conduct future research to close these gaps. Research summaries should be 5-6 pages long (excluding any cover page and a separate page for references; double-spaced; 12-point Times New Roman; 1-inch margins throughout). Research summaries must be submitted on CANVAS by 9 December 2025.

Research Project Proposal (graduate students only): 5%

Graduate students only will submit a synopsis of a planned research project that is worth 5 points. The proposal should detail (i) the topic area and specific hypothesis you plan to examine, (ii) the corpus data you plan to use, (iii) the corpus-linguistic methods you plan to use, and (iv) the kind of statistical or other form of evaluation you plan to do. The proposal should not exceed 2 pages, excluding title page and references (12-point Times New Roman font, single-spaced, 1-inch margins throughout) and should be submitted on Canvas by 9 October 2025. Graduate students will informally report on their research project proposals in class on 2 October 2025 to receive feedback.

Research Project Presentation (graduate students only): 10%

Graduate students only will give a 15—minute in-class presentation of their research project worth 10 points. Presentations will be scheduled for classes on 13, 18, and 20 November 2025. The presentation should comprise all major sections of a scientific research presentation (Introduction, Literature Review, Methods, Results, and Discussion). If the research project is yet to be completed by the time of the presentation, a section should be added at the end of the presentation that outlines the remaining work and details any challenges yet to be overcome.

Research Project Final Paper (graduate students only): 13%

Graduate students only should submit a final paper of their research project worth 13 points. The paper should comprise all major sections of a scientific research manuscript (Introduction, Literature Review, Methods, Results, and Discussion). The writeup should be between 8 and 12 pages long, excluding title page and references (12-point Times New Roman font, single-spaced, 1-inch margins throughout) and is due on Canvas by 9 December 2025.

Extra Credit and Make-up Assignments

There are two different extra credit opportunities in this class, from which you can choose one:

1. You can complete the Kognito training to earn 2 extra points. This assignment is available on Canvas like all other assignments.
2. You can participate in an experiment on campus. You can find information about ongoing experiments here: <https://ufl-slhs.sona-systems.com>. Proof of participation should be sent to Dr. Wulff via email (swulff@ufl.edu) no later than 9 December 2025.

Technology and Materials Requirements

To participate fully in this course, you will need to:

- 1) Get a subscription to SketchEngine, an online software tool that allows you to access and analyze corpora in more than 90 languages, and lets you build your own corpora. A subscription to SketchEngine is ~\$9/month. You can sign up here: <https://www.sketchengine.eu/Links to an external site.>
- 2) Install R, a free statistics and programming environment (<https://www.r-project.org/Links to an external site.>). Instructions for how to install R are on Canvas.
- 3) Install RStudio, a free integrated development environment (IDE) for R and Python (<https://posit.co/products/open-source/rstudio/Links to an external site.>). Instructions for how to install RStudio are on Canvas.

There is one required textbook for this class: Anatol Stefanowitsch (2020). *Corpus Linguistics: A guide to the methodology*. Berlin: Language Science Press [AS]. This book is open access and will be made available as a pdf on Canvas. All other course materials (research papers, worksheets, etc.) will also be available on CANVAS. To participate fully in this course, you will need to be able to access the Canvas course site.

UF Academic Policies

This course complies with all UF academic policies. For information about those policies and for resources for students, please see <https://syllabus.ufl.edu/syllabus-policy/uf-syllabus-policy-links/>.

Course Schedule

Date	Weekday/Time	Topic/Assignments Due
8/21	R 10:40-12:35	Organizational matters
8/26	T 10:40-11:30	Lecture: What is corpus linguistics?
8/28	R 10:40-12:35	Lecture: Corpus linguistics as a scientific research method
9/2	T 10:40-11:30	Lecture: Data retrieval and annotation
9/4	R 10:40-12:35	SketchEngine: Word Sketch and Word Sketch Difference Home Quiz 1 due (all students)
9/9	T 10:40-11:30	Lecture: Quantifying research questions
9/11	R 10:40-12:35	SketchEngine: Thesaurus and Word List Home Quiz 2 due (all students)
9/16	T 10:40-11:30	Lecture: Quantifying research questions <i>ctd.</i>
9/18	R 10:40-12:35	SketchEngine: Regular Expressions Home Quiz 3 due (all students)
9/23	T 10:40-11:30	Lecture: Significance testing
9/25	R 10:40-12:35	SketchEngine: Regular Expressions <i>ctd.</i> Home Quiz 4 due (all students)
9/30	T 10:40-11:30	Lecture: Significance testing <i>ctd.</i>
10/2	R 10:40-12:35	SketchEngine: Keywords / N-grams Graduate students pitch research project proposals Home Quiz 5 due (all students)
10/7	T 10:40-11:30	Lecture: Collocation
10/9	R 10:40-12:35	SketchEngine: Concordance Research project proposals due (graduate students) Home Quiz 6 due (all students)
10/14	T 10:40-11:30	Lecture: Collocation <i>ctd.</i>
10/16	R 10:40-12:35	SketchEngine: Corpus Query Language (CQL) Home Quiz 7 due (all students)
10/21	T 10:40-11:30	Video lecture: Corpus methods in language learning
10/23	R 10:40-12:35	SketchEngine: Corpus Query language(CQL) <i>ctd.</i> Discussion Board 1 due (undergraduate students) Home Quiz 8 due (all students)
10/28	T 10:40-11:30	Video lecture: Language change and corpus methods
10/30	R 10:40-12:35	SketchEngine: Text Types / Subcorpora / Corpus Building Discussion Board 2 due (undergraduate students) Home Quiz 9 due (all students)
11/4	T 10:40-11:30	Video lecture: Corpus linguistics and legal interpretation
11/6	R 10:40-12:35	Introduction to basic text processing Discussion Board 3 due (undergraduate students) Home Quiz 10 due (all students)
11/11		Holiday (no class)
11/13	R 10:40-12:35	Research project presentations (graduate students) Home Quiz 11 due (all students)
11/18	T 10:40-11:30	Research project presentations (graduate students)
11/20	R 10:40-12:35	Research project presentations (graduate students) Home Quiz 12 due (all students)
11/23-11/30		Holiday (no class)
12/2	T 10:40-11:30	Wrapping Up
12/9		Research project final paper due (graduate students) Research summary due (undergraduate students) Extra credit documentation due (all students) Lecture attendance documentation due (undergraduate students)