

LIN4071/5075 Introduction to Corpus Linguistics (Fall 2021)

Instructor: Stefanie Wulff, Ph.D.

Office hours: Wednesdays 11am-12pm or by appointment (face-to-face: 4131D Turlington Hall; or see zoom meeting ID and passcode provided on Canvas)

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 Wednesdays 1:55-4:55pm in Matherly Hall (MAT) 0011.

Students are allowed to record video or audio of class lectures. However, the purposes for which these recordings may be used are strictly controlled. The only allowable purposes are (1) for personal educational use, (2) in connection with a complaint to the university, or (3) as evidence in, or in preparation for, a criminal or civil proceeding. All other purposes are prohibited. Specifically, students may not publish recorded lectures without the written consent of the instructor.

A “class lecture” is an educational presentation intended to inform or teach enrolled students about a particular subject, including any instructor-led discussions that form part of the presentation, and delivered by any instructor hired or appointed by the University, or by a guest instructor, as part of a University of Florida course. A class lecture does not include lab sessions, student presentations, clinical presentations such as patient history, academic exercises involving solely student participation, assessments (quizzes, tests, exams), field trips, private conversations between students in the class or between a student and the faculty or lecturer during a class session.

Publication without permission of the instructor is prohibited. To “publish” means to share, transmit, circulate, distribute, or provide access to a recording, regardless of format or medium, to another person (or persons), including but not limited to another student within the same class section. Additionally, a recording, or transcript of a recording, is considered published if it is posted on or uploaded to, in whole or in part, any media platform, including but not limited to social media, book, magazine, newspaper, leaflet, or third party note/tutoring services. A student who publishes a recording without written consent may be subject to a civil cause of action instituted by a person injured by the publication and/or discipline under UF Regulation 4.040 Student Honor Code and Student Conduct Code.

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 AntConc; 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 competency in using the AntConc 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

| | |
|---|-----------|
| 8 home quizzes (7.5 points each) | 60 points |
| Research summary | 20 points |
| 3 discussion boards (5 points each) | 15 points |
| Research participation/lecture attendance | 5 points |

Graduate students

| | |
|-------------------------------------|-----------|
| 8 home quizzes (7.5 points each) | 60 points |
| 3 discussion boards (5 points each) | 15 points |
| Research project proposal | 5 points |
| Research project presentation | 10 points |
| Research project writeup | 10 points |

Discussion Boards (undergraduate and graduate students): 15% (5% each)

Undergraduate and graduate students will complete three discussion boards that analyze a given data set or 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.

Home Quizzes (undergraduate and graduate students): 60% (7.5% each)

Both undergraduate and graduate students will complete 8 quizzes worth 7.5% 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.

Research Participation/Lecture Attendance (undergraduate students only): 5%

You will choose one of the following options to get credit of 5% towards your final grade:

1. participation in a research experiment: you can participate in any research experiment conducted on campus during the fall 2021 semester. The faculty member in charge of the research experiment should confirm your participation in an email to Dr. Wulff

(swulff@ufl.edu). Proof of participation has to be submitted to Dr. Wulff by 13 December 2021.

2. lecture attendance: you can attend a language-related lecture (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 by 13 December 2021.

Research Summary (undergraduate students only): 20%

Undergraduate students only will be responsible for a research summary worth 20 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 8 December 2021.

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 13 October 2021.

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 1 and 8 December 2021. 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 writeup (graduate students only): 10%

Graduate students only should submit a writeup of their research project worth 10 points. The writeup 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).

Attendance and Participation

Attendance is mandatory for all class meetings for undergraduate and graduate students. Students can miss up to three 50-minute class periods without any excuse. Each further unexcused absence will lower the final grade by one level (e.g. from an A to A-). Excused absences must be consistent with university policies in:

<<https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx>>.

Extra Credit and Make-up Assignments

There is one extra credit opportunity in this class: you can complete the Kognito training to earn 3 extra points. This assignment is available on Canvas like all other assignments. If you have three or more excused absences, please contact Dr. Wulff to discuss the possibility of a make-up exam or equivalent. You cannot make up for unexcused absences.

Grading Scale

Your total score determines your final grade on this grading scale:

100-93=A; 92.9-90=A-; 89.9-87=B+; 86.9-83=B; 82.9-80=B-; 79.9-77=C+; 76.9-73=C(S); 72.9-70=C-(U); 69.9-67=D+; 66.9-63=D; 62.9-60=D-; 59.9-0=E. For information on current UF grading policies for assigning grade points, please go to <<https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies>>.

I aim to grade and give feedback on any course assignments as promptly as possible, within a week at the latest. For more information on how exactly each assignment will be graded, please read the grading rubric page on Canvas.

Classroom Etiquette

Please switch your cell phone to silent during class. If we are meeting on Zoom, mute your microphone unless you are speaking to minimize background noise for everyone. Please make every effort to arrive to class on time (even if it is on Zoom). Being more than 10 minutes late to the start of the course will count as an absence on each third occasion of such tardiness. We will familiarize ourselves with other rules of “netiquette” in the first week of classes; a summary of (n)etiquette rules is on Canvas.

Technology and Materials Requirements

There is one required text books for this class: Anatol Stefanowitsch (2020). *Corpus Linguistics: A guide to the methodology*. Berlin: Language Science Press. 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, and log on to Zoom when you want to attend our course meetings online. You will have to download the following software (all free) to your laptop:

AntConc: <https://www.laurenceanthony.net/software/antconc/> (Links to an external site.)

R: <https://cran.r-project.org/> (Links to an external site.)

NotePad++: <https://notepad-plus-plus.org/> (Links to an external site.)

Furthermore, you will have to access google drive to download corpus data (<https://drive.google.com/drive/folders/1Yo7F3uaRUPwmFH3X33BHmWq7EIYZKTLH?usp=sharing>).

Student Accommodations

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting <<https://disability.ufl.edu/students/get-started>>. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.

Course Evaluation

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at <<https://gatorevals.aa.ufl.edu/students>>. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via <<https://ufl.bluera.com/ufl>>. Summaries of course evaluation results are available to students at <<https://gatorevals.aa.ufl.edu/public-results>>.

University Honesty Policy

UF students are bound by The Honor Pledge which states, “We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: “On my honor, I have neither given nor received unauthorized aid in doing this assignment.” The Honor Code (<<https://www.dso.ufl.edu/sccr/process/student-conduct-honor-code>>) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

Software Use

All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

Student Privacy

Federal laws protect your privacy with regards to grades earned in courses and on individual assignments. See <http://registrar.ufl.edu/catalog0910/policies/regulationferpa.html> for more information.

Campus Resources: Health and Wellness

- U Matter, We Care: if you or someone you know is in distress, please contact umatter@ufl.edu, 352-392-1575, or visit umatter.ufl.edu to refer or report a concern and a team member will reach out to the student in distress.
- Counseling and Wellness Center: visit counseling.ufl.edu or call 352-392-1575 for information on crisis services as well as non-crisis services.
- Student Health Care Center: call 352-392-1161 for 24/7 information to help you find the care you need, or visit shcc.ufl.edu. University Police Department: visit police.ufl.edu or call 352-392-1111 (or 9-1-1 for emergencies).
- UF Health Shands Emergency Room / Trauma Center: for immediate medical care, call 352-733-0111 or go to the emergency room at 1515 SW Archer Road, Gainesville, FL 32608; ufhealth.org/emergency-room-trauma-center.

Campus Resources: Academic Resources

- E-learning technical support: contact the UF Computing Help Desk at 352-392-4357 or via e-mail (helpdesk@ufl.edu).
- Career Connections Center: career assistance and counseling services career.ufl.edu, Reitz Union Suite 1300, 352-392-1601.
- Library Support: cms.uflib.ufl.edu/ask various ways to receive assistance with respect to using the libraries or finding resources.
- Teaching Center: Broward Hall, 352-392-2010 or to make an appointment 352-392-6420. General study skills and tutoring teachingcenter.ufl.edu.
- Writing Studio: 2215 Turlington Hall, 352-846-1138. Help brainstorming, formatting, and writing papers writing.ufl.edu/writing-studio.
- Student Complaints On-Campus: sccr.dso.ufl.edu/policies/student-honor-code-student-conduct-code
- On-Line Students Complaints: distance.ufl.edu/student-complaint-process.

Course Schedule

| Date | Topic/Assignments Due | Read Before Class |
|------|--|--------------------|
| 8/25 | Organizational matters Lecture: What is corpus linguistics? Practice: Introduction to AntConc software | [AS] chapter 2 |
| 9/1 | Lecture: Corpus linguistics as a scientific research method Introduction to AntConc software <i>contd.</i> | [AS] chapter 3 |
| 9/8 | Lecture: Data retrieval and annotation // Quantifying research questions Practice: Concordances Practice: Quantifying research questions Home Quiz 1 due | [AS] chapters 4, 5 |
| 9/15 | Lecture: Significance testing Practice: Significance testing Home Quiz 2 due | [AS] chapter 6 |
| 9/22 | Lecture: Collocation | [AS] chapter 7 |

| | | |
|-------|---|------------------------------|
| | Practice: Collocation Home Quiz 3 due | |
| 9/29 | Practice: Collostructional Analysis Home Quiz 4 due | Stefanowitsch & Gries (2003) |
| 10/6 | Video lecture: Constructions and their collocates (Martin Hilpert) https://www.youtube.com/watch?v=lKWc8pIDO_w | |
| 10/13 | Lecture: Collocation at the interfaces of words and grammar/morphology Practice: TTRs and HTRs Discussion Board 1 due | [AS] chapter 8, 9 |
| 10/20 | Lecture: Collocation at the interface of words and the texts they occur in Practice: Word and keyword lists Brainstorming session for graduate student research proposals Home Quiz 5 due | [AS] chapter 10 |
| 10/27 | Practice: Working with tagged data Home Quiz 6 due Research Project Proposal due (grad students only) | |
| 11/3 | Practice: Regular expressions Discussion Board 2 due | |
| 11/10 | Practice: Basics of text processing Home Quiz 7 due | |
| 11/17 | Practice: Corpus compilation and annotation Home Quiz 8 due | Aedel (2020) |
| 11/24 | No class (Thanksgiving holiday) | |
| 12/1 | Graduate student project presentations Discussion Board 3 due | |
| 12/8 | Graduate student project presentations Research Summary due (undergrad students only) | |
| 12/13 | Research Project Writeup due (grad students only) | |