

# LIN4770C/5770C INTRODUCTION TO COMPUTATIONAL LINGUISTICS

Spring 2023

---

<b>Instructor:</b>	Prof. Zoey Liu	<b>Office:</b>	Turlington 4016
<b>Email:</b>	<a href="mailto:liu.ying@ufl.edu">liu.ying@ufl.edu</a>	<b>Office Hour:</b>	T 2:50-4pm & R 3:55-5pm Or by appointment
<b>TA:</b>	Raphael Iyamu	<b>Office:</b>	Turlington 4122
<b>Email:</b>	<a href="mailto:r.iyamu@ufl.edu">r.iyamu@ufl.edu</a>	<b>Office Hour:</b>	T 12-2pm & R 12-1pm

---

## Course objectives

This course surveys selected topics and tasks that are fundamental to computational linguistics. Students will learn to gain (1) understanding of different linguistic problems that could be solved with mathematical means; (2) understanding of different computational methods to derive automatic analysis of language structures at different linguistic levels; (3) technical programming skills to model linguistic phenomena in a computational framework.

## Prerequisite

Given the title of the course, where the word *computational* functions as the modifier of *linguistics*, **an interest in Linguistics is required**. In addition, for this class, we will be using the programming language Python fairly frequently. That said, prior coursework in programming is not required.

## Course website

We will be using Canvas as the course website. All lecture notes, code, and supplementary materials will be posted on Canvas. Grades will be posted to the Canvas grade book.

## Textbook

There is no required textbook for the course. At the same time, students are encouraged to use relevant chapters from the following (free!) book as references:

- [Speech and Language Processing \(3rd ed. draft\)](#). Dan Jurafsky & James H. Martin.

## Attendance

Attendance is required. You are responsible for all the material that is covered during class, even when your absence from class is excused. While in attendance, you are expected to actively participate. When you are excused, download the lecture notes from Canvas and ask a classmate what else you might have missed.

## Laptop

Class meetings will include lectures, quizzes, and programming-related activities. Bring your laptop to class, but leave it closed unless you are writing code or taking a quiz.

---

## Course outline

Note: the following course outline is subject to change.

<b>Week</b>	<b>Topic</b>
Week 1	Introduction; Catching up with the Pythonics
Week 2	Text Normalization; Catching up with the Pythonics
Week 3	Regular Expressions with Python; Catching up with the Pythonics
Week 4	Edit Distance; Catching up with the Pythonics
Week 5	<i>N</i> -gram Language Modeling
Week 6	Noisy Channel; Statistical Machine Learning I
Week 7	Statistical Machine Learning II
Week 8	Part-of-Speech Tagging
Week 9	Formal Language
Week 10	<b>Spring Break</b>
Week 11	Context-free Parsing
Week 12	Dependency Parsing
Week 13	Word Senses; Vector Semantics
Week 14	Vector Semantics; Document Similarity; Topic Modeling
Week 15	Topic Modeling; Final Project Presentation
Week 16	Final Project Presentation

## Grade breakdown

<b>Quizzes</b>	10%
<b>In-class exercises</b>	10%
<b>Assignments</b>	40%
<b>Final project</b>	
Three ideas	5%
Summary paragraph	5%
Project proposal	15%
Project poster	10%
Final presentation	5%
Optional extra credit assignment	2%

---

## Grade scale

A	90-100	A-	87-89.9	B+	84-86.9	B	80-83.9
B-	77-79.9	C+	74-76.9	C	70-73.9	C-	67-69.9
D+	64-66.9	D	60-63.9	D-	57-59.9	E	< 57

## Quizzes

Make-up quizzes will be offered only when accompanied by a letter of explanation from your class dean or a medical professional. An unexcused absence on the day of a quiz will result in a score of 0 for that quiz.

## Assignment collaboration

Students are encouraged to collaborate with each other on questions that involve programming in the assignments; if that were the case, the name(s) of the collaborating students are required to be mentioned as a comment on Canvas upon assignment submission.

For questions that involve qualitative data analysis, students are expected to complete them independently.

## Late work

Late assignments will not be accepted. A student's lowest quiz score and lowest assignment grade will be dropped. Late work relating to the term project will be accepted only when accompanied by a letter of explanation from your class dean or a medical professional.

## Optional extra credit assignment

Students who wish to do so may increase their total grade points by 2 (out of 100) by participating 2 hours in psycholinguistic experiments through the LIN/S&P/SLHS participant pool. Deadline is March 31st. A list of experiments can be found [here](#). This site will be updated throughout the semester. Please retain the IRB form with the original signature from the experimenter (no photocopies) as proof that you participated. Experiments that you have previously participated in, experiments that you count towards other classes, or experiments you were paid to participate in cannot count toward this. If you choose not to participate or do not qualify for any of the studies, you can receive the same amount of course credit by means of an alternative assignment. This involves writing a synopsis of a short research article, video, or a podcast. Details will be made available by the instructor.

## Academic Integrity

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](#) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, students are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If students have any questions or concerns, please consult with the Instructor or the TA.

## Classroom Conduct

Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to professional behavioral standards may be subject to discipline. The Instructor pledges

---

to treat each of the students with dignity, respect, and professional courtesy. Students are expected to do the same for the Instructor and for each other.

### **Accommodation Policies**

If a student qualifies for accommodations because of a disability, please submit their accommodation letter from the [Disability Resource Center](#) to the Instructor in a timely manner so that their needs can be addressed.

### **Religious Observances**

A student should inform the Instructor of religious observances that will conflict with class attendance, tests or examinations, or other class activities prior to the class or occurrence of that test or activity. The Instructor is obligated to accommodate students' religious observances. See policy details [here](#).

### **Course Evaluations**

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 [here](#). 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 [here](#). Summaries of course evaluation results are available to students [here](#).

### **Recording lecture content.**

Students are allowed to record class lectures. However, the only allowable purposes of these recordings 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. Specifically, students may not publish recorded lectures without the written consent of the instructor. 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. 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 action and/or discipline the Student Honor Code and Student Conduct Code.