LIN4770C/5770 Introduction to Computational Linguistics

MWF 10:40-11:30, Matherly 102
Spring 2026

Instructor:	Dr. Zoey Liu	**Office:**	Turlington 4109
Email:	liu.ying@ufl.edu	**Office Hour:**	TBD
		Office Phone:	352-294-7457
TA:	Minghao Zheng	**Office:**	Dauer 42
Email:	minghao.zheng@ufl.edu	**Office Hour:**	TBD

Course description and 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, students are expected to have taken LIN4930/6932 Programming for Linguists or the equivalent to ensure they have sufficient background in the programming language, `Python`.

Course website

We will be using Canvas as the course website. All lecture and lab related materials will be posted on Canvas. Instructional materials for this course consist of only those materials specifically reviewed, selected, and assigned by the Instructor. The Instructor is only responsible for these instructional materials. Grades will be posted to the Canvas grade book.

Course communication

The preferred method for any course-related written communications is via **Canvas messaging**. Students are more than encouraged to come to office hours or schedule an appointment with the Instructor for any course-related matter that needs to be discussed in-person.

If students need help resolving technical issues for communication purposes, please visit the helpdesk website or call 352-392-4357.

External references

There is no required textbook for the course. Class content will rely heavily on relevant chapters from the following (free!) book as references:

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

Grade breakdown

Attendance & Participation	5%	**Assignments**	30%	**Final project deliverables**	
Peer engagement	8%			Project pitch	2%
Quizzes	12%			Project proposal	3%
Labs	15%			Final presentation	5%
				Code walk-through	5%
				Final write-up	15%

Grade scale

We will be using the default canvas grading scheme.

		B+	87-89%	C+	77-79%	D+	67-69%	
A	94-100%	B	83-86%	C	73-76%	D	63-66%	F 0-59%
A-	90-93%	B-	80-82%	C-	70-72%	D-	60-62%	

Attendance & Participation

Attendance is required. Acceptable reasons for absence from or failure to engage in class can be found here. In addition, appropriate documentation for absences (except for religious holidays), missed work, or the inability to fully engage in class is required. A warning will be issued after two consecutive weeks of absences. Attendance will be prohibited after four weeks of absences, resulting in a withdrawal recommendation and/or a failing grade for excessive absences.

Students are responsible for all the material that is covered during class, *even when their absence from class is excused*. While in attendance, students are expected to actively participate. When their absence is excused, download the lecture notes from Canvas and ask a classmate what else they might have missed.

Laptop: Class meetings will include lectures, quizzes, and programming-related activities. Make sure to bring laptop to class.

Peer engagement

Engaging in community discussion can constitute a rewarding and beneficial way of learning, because this motivates sharing thoughts with each other and learning from diverse perspectives. In this class, aside from class participation, peer engagement will be operationalized as follows: please choose at least 8 weeks of this semester; for each week, post a question, a comment, a thought, or a response to someone else's question on the week's lecture materials (and/or readings) in Canvas Discussion section. For each week of the student's choice, the deadline for posting on the Discussion board is by 11:59 PM on the Friday of that week.

Quizzes

There will be approximately 8-9 pop quizzes throughout the semester; students are expected to take the quiz in class with their laptop. Quizzes are open-book; usage of AI tools, however, is not allowed. The lowest quiz score will be dropped.

Labs

There will be approximately 6 lab sessions throughout the semester, led by the TA; students are expected to participate in lab sessions with their laptop. Each lab is due by 11:59pm on the Friday of the week with

a lab session. Consultation with AI tools for completing labs after the lab session is not prohibited, but is not recommended. Usage of any AI tools is required to be mentioned as a comment on Canvas upon lab assignment submission. Failing to do so will result in a grade of 0 for the lab.

Assignment

There will be a total of five coding assignments. Students are expected to complete each assignment independently. Consultation with AI tools for completing assignments is not prohibited, but is not recommended. Usage of any AI tools is required to be mentioned as a comment on Canvas upon assignment submission. Failing to do so will result in a grade of 0 for the assignment. Assignments have strict deadlines which will be noted on Canvas.

Final project

Students will form teams to complete their final project; the ideal size of the team is 2 people. The final project consists of five components: (1) project pitch; this requires students to come to office hours and discuss with the Instructor in person for approval of their project idea (2) a two-paragraph project proposal; (3) project presentation in Week 15; (4) a code walk-through of the experiments; (5) a write-up (5-8 pages, excluding references) for the final project. Each of the five components has strict deadlines which will be noted on Canvas; this is to ensure that students are on track making progress towards completing their final project.

The code walk-through is to be performed using jupyter notebooks. Each code chunk in the notebook need to have been executed with output by the time of submission. Written documentation is required in the notebook to link the output in the notebook with specific section/page number in the final project.

Any use of AI for generating code from scratch, or for writing the final project, will result in a grade of 0 for the code walk-through as well as the final write-up.

Late work

Late submissions for this class are not accepted, unless they are accompanied by a letter of explanation from the class dean or a medical professional. Any reasons that fall out of those scenarios should be discussed with the Instructor in person; email inquiries in these cases will not be addressed.

Further Course Policies

The Academic Accommodation Policy, The University's Honesty Policy, The GatorEvals Policy, and The In-class Recording Policy for this course can be found here: UF Academic Policies and Resources.

Academic Resources

A full list of academic resources available to all UF students can be found here: UF Academic Policies and Resources.

Course outline

Note: the following course outline is *subject to change*.

Week	Topic	Lab	Assignment due
Week 1	Introduction; Text Normalization		
Week 2	Regular Expressions with Python; Edit Distance	Lab 1	
Week 3	Formal Language		Assignment 1
Week 4	N-gram Language Modeling	Lab 2	
Week 5	Noisy Channel; Statistical Machine Learning		Assignment 2
Week 6	Statistical Machine Learning, continued	Lab 3	
Week 7	Part-of-Speech (POS) Tagging		Assignment 3
Week 8	Context-free Parsing / TBD	Lab 4	
Week 9	Dependency Parsing		Assignment 4
Week 10	**Spring Break**		
Week 11	Dependency Parsing, continued	Lab 5	Project pitch
Week 12	Word Senses; Vector Semantics		
Week 13	Vector Semantics, Continued	Lab 6	Project proposal
Week 14	Topic Modeling; Neural Networks		
Week 15	Final Project Presentations		Assignment 5