

Ling 317 Introduction to Computational Linguistics

Dr. Nick Danis, nsdanis@wustl.edu

Description

This course introduces the computational tools, both practical and theoretical, to describe and analyze natural language. We will touch upon most of the major aspects of the field of computational linguistics and natural language processing, such as analysis of raw corpora; modeling phonological, morphological, & syntactic patterns; and learnability. We will use Python 3 throughout the course. This course is an introduction to these topics, and therefore no programming experience is required. Students, however, are encouraged to complete optional exercises and readings on basic programming techniques to aid in the completion of assignments and discussions. Prerequisite: L44 Ling 170D and either L44 Ling 258 or CSE 131.

Course Info

Course Number	L44 Ling 317
Semester	Spring 2025
Time	MW 1:00P-2:20P
Location	Eads 003
Office	January 206
Office Hours	Tuesdays 12-1pm
Homepage	https://wustl.instructure.com/courses/149333

Goals

1. Learn practical, computational tools (e.g. Python NLTK) to manipulate natural language data
2. Understand the computational properties of natural language, independent of whether it is being computed by machines or humans (i.e. treating grammars as purely mathematical objects)
3. Apply these techniques to solve problems in phonology, morphology, and syntax

Required Materials

There are no required purchases for this course. Any readings will be made available online. Additionally, all software necessary is free as well. This course assumes you have (or have access to) a working computer with a desktop operating system (macOS, Windows, Linux, etc.). For those

students working primarily off a Chromebook or iOS/Android device, there are cloud options available. More details are given on Canvas.

Readings

A majority of the readings will come from the following source:

- Steven Bird et al. (2014). *Natural Language Processing with Python*. 2nd ed. URL: <http://www.nltk.org/book/>

The entire second edition is available online at the link above; do not purchase a physical copy as the printed edition is the outdated first edition.

Software

All course material that is not a PDF is distributed in a Jupyter notebook for Python 3. If you are unsure what to install, or what this means, we will go over it the first week of class. There is detailed instructions on the Canvas site as well. If you have some experience, and have Python 3 installed, you can get a headstart with `pip install jupyter`.

Grade

The grade breakdown is shown below.

Category	Weight
Participation & Discussion*	15%
Weekly Assignments*	60%
Final Project	25%

*The lowest grade is dropped.

Participation entails completing & uploading the in-class notebooks to Canvas by the end of class time, and regular engagement on the course Piazza & in-class discussions. *There are no make-up opportunities for participation notebooks**, but if you have an excused absence, it will not count against your grade.

The final project consists of a proposal, a presentation during the final weeks of class, and the project (code/notebook) itself. Details are available on Canvas; it is never too early to start thinking about ideas.

Letter grades

Letter grades are assigned based off the following scale. Numerical grades are **not** rounded.

$100 \geq A+ \geq 98$	$80 > C+ \geq 77$
$98 > A \geq 93$	$77 > C \geq 73$
$93 > A- \geq 90$	$73 > C- \geq 70$

90 > B+ \geq 87

70 > D+ \geq 67

87 > B \geq 83

67 > D \geq 63

83 > B- \geq 80

63 > D- \geq 60

If you are taking this class pass/fail, you must receive at least a C- (70%) to pass.

If you believe there has been an error in grading, I am happy to discuss it with you. However, you must bring it up to me within one week of the graded assignment being returned to you. After this, the grade is considered final.

Schedule

The exact schedule is likely to change as the semester progresses. Please see [Canvas](#) for all up-to-date readings and assignment due dates.

date	week	module	comment
1/13/2025	1	introduction	
1/20/2025	2	scripting and control flow	
1/27/2025	3	regex	
2/3/2025	4	corpora and raw text i	
2/10/2025	5	corpora and raw text ii	
2/17/2025	6	vectors, models, and embeddings	
2/24/2025	7	maximum likelihood estimation	
3/3/2025	8	tf-idf and document embeddings	final project proposal due
3/10/2025		spring break	
3/17/2025	9	finite state machines	
3/24/2025	10	finite state transducers	
3/31/2025	11	context free grammars	
4/7/2025	12	machine learning/ai	
4/14/2025	13	presentations	
4/21/2025	14	presentations	presentation material due
5/4/2025			final project due

Academic Integrity

This course adheres to the university's [Academic Integrity Policy](#), and takes cheating and plagiarism very seriously. All work completed online must be done alone unless instructed otherwise, and no resources not approved by the instructor may be used during exams.

Students are welcome and encouraged to use whatever resources they can to aid in understanding the material, though all submitted work must be the work of the student. This means that you cannot submit code and assignments solely authored via AI tools like ChatGPT/Copilot/etc. If you are indeed stuck and even robots cannot help, please ask a human, via email or Piazza.

ADA Compliance

Washington University is committed to providing accommodations and/or services to students with documented disabilities. Students who are seeking support for a disability or a suspected disability should contact Disability Resources at 935-4153. Disability Resources is responsible for approving all disability-related accommodations for WU students, and students are responsible for providing faculty members with formal documentation of their approved accommodations at least two weeks prior to using those accommodations. I will accept Disability Resources VISA forms by email and personal delivery. If you have already been approved for accommodations, I request that you provide me with a copy of your VISA within the first two weeks of the semester. Please see more information at <http://cornerstone.wustl.edu>.

Sexual Assault Resources

The University is committed to offering reasonable academic accommodations (e.g., no contact order, course changes) to students who are victims of relationship or sexual violence, regardless of whether they seek criminal or disciplinary action. If you need to request such accommodations, please contact the Relationship and Sexual Violence Prevention Center (RSVP) at rsvpcenter@wustl.edu or 314-935-3445 to schedule an appointment with an RSVP confidential, licensed counselor. Information shared with counselors is confidential. However, requests for accommodations will be coordinated with the appropriate University administrators and faculty. Please see more information at <https://students.wustl.edu/relationship-sexual-violence-prevention-center>.

Last updated: January 12, 2025