Intro to Computational Linguistics (Tentative Syllabus)

	Course Info
Course#	LING 5300
Semester	Spring 2022
Time	TuTh 9:10AM - 10:30AM
Location	CSC 13
Website	Canvas/Github/Discord
Textbook:	None required. Readings will be available on Canvas
Prerequisites	Review the "Prerequisites" section for more information

	Instructor Info
Instructor	Aniello De Santo
Pronouns	He/His/Him
Email	aniello.desanto@utah.edu
Department	Linguistics
Office	LNCO 2315 or Zoom
Office hours	Review the "Office hours" section for more information

1 Catalog description

An introduction to computational linguistics for students with previous programming experience. This course explores the models, algorithms, and techniques that dominate modern-day language technology, and it evaluates them from a linguistically informed perspective. Topics include corpus-based methods, finite-state approaches, word vectors, computational models of cognitive processes, and model evaluation techniques. Great emphasis is put on discussing the limitations of existing techniques and how they might benefit from linguistic insights. We also discuss Ethics in AI and in Natural Language Processing, and the way fairness considerations should inform data collection and algorithm design. Students will also hone their programming skills and develop familiarity with state-of-the-art software packages for computational linguistics.

2 Goals and Objectives

- familiarity with computational linguistic topics, tools, and resources
- conceptualize linguistics problems in computational terms
- ability to design and carry out a linguistically-informed error analysis of computational linguistic systems
- ability to critically evaluate the societal impact of language technology in scientific contexts as well as in the popular press
- develop learning autonomy and the ability to deepen your programming knowledge through self-study

3 Prerequisites

LING 1069 OR LING 1200 AND LING 3300 OR CS 1410. If you could enroll, you satisfy the official prerequisites (or I gave you permission) Basic familiarity with programming concepts is expected, and familiarity with Python is desirable. If you have no experience in Python, I will make resources available to catch up with the basic syntax. Assignments are designed in a scaffolding fashion, ideally accessible to a variety of linguistics and programming backgrounds.

3.1 Required Materials and Resources

All you need for this course is a device with a recent browser (Windows, OS X, or Linux computer, Chromebook, Android tablet), and an internet connection. Ensure you have these ahead of time as they will be required to complete assignments and activities throughout the course. If you anticipate tech-related issues, please get in touch with me early in the semester, so that I can assist you in finding alternatives via the University.

As a programming language, we will use Python 3, run through *Jupyter Notebooks*, an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. You can install a Jupyter environment locally on your machines, or run them remotely via *CoLab*, an online environment that allows you to write and execute Python in your browser, with no configuration required and free access to GPUs. We will go over the details in the first week of classes.

All class materials will be distributed on Canvas (you can familiarize yourself with Canvas **here**).

No book is required for this course. When needed, weekly readings will be distributed via Canvas. However, if you are interested in pursuing topics in greater depth in this course, you may supplement your reading for class with these:

- Jurafsky and Martin (2008). Speech and Language Processing, Ed. 2. ISBN: 978-0131873216 (Available in the University Library)
- Jurafsky and Martin (—-). Speech and Language Processing, Ed. 3. https://web.stanford.edu/~jurafsky/slp3/

- Eisenstein (2019). Introduction to Natural Language Processing. ISBN: 9780262042840 (free from github: https://github.com/jacobeisenstein/gt-nlp-class/blob/master/notes/eisenstein-nlp-notes.pdf)
- Bender (2013). Linguistic Fundamentals for Natural Language Processing: 100 Essentials from Morphology and Syntax. https://www.morganclaypool.com/doi/abs/10.2200/S00493ED1V01Y201303HLT020

- Bender and Lascarides (2019). Linguistic Fundamentals for Natural Language Processing II: 100 Essentials from Semantics and Pragmatics. https://www.morganclaypoolpublishers.com/catalog_Orig/product_info.php?products_id=1451
- Gorman and Sproat (2021). Finite-State Text Processing. ISBN: 9781636391137 (Online Edition Available with University of Utah Access http://www.morganclaypoolpublishers.com/catalog_Orig/product_info.php?products_id=1636)

3.2 Discord Server

To facilitate remote discussion/collaboration among peers, this course has an associated Discord server. Discord is a collaborative chatroom environment extensively used for Team synchronization in the tech industry. You can create a free account and join the course channel with an invite I will share on Canvas the first week of classes. **Make sure to join the channel before the second week of classes.**

4 Grading

Please try not to take grades as a war between you and me. My aim is to help you be as successful in this course as possible. Grading is meant as a way for both you and the instructors to check on how well the course is helping you internalize concepts, and keep track if your progress. Student grades are determined by the following components:

1. **Weekly homework (60%) and readings**
   Weekly homework assignments will be a mix of theoretical exercises, coding practices in Python, and reading write-ups. One homework may take the form of

   - several minor practice exercises, or
   - a single, more open ended task (e.g. implementing or evaluating a chatbot for English).

   Sometimes HW assignments will be used to explore concepts and tools we did not have time to look at in class. Therefore, my recommendation is always to at least glance at an homewrok as soon as it is out, to assess how much time you will need to dedicate to it.

   Additionally, one or two readings might be associated to the homework. The content of the reading might be referenced in an assignment and/or in class discussions, and general familiarity with the content will be expected.

   Assignments are unlocked each Tuesday and, unless stated otherwise, are due the following

Monday at 11:59pm. Details on the requirements for each assignment will be shared with the students in advance. When computing the final grade, the lowest homework grade will be dropped.

Late Submissions: Life is messy (especially these days)! While I encourage you to do assignments on time due to the incremental nature of the materials, I generally adopt a flexible deadline approach: if you think you are going to be late on an assignment, just let me know and we can figure things out. I do not need doctor's notes or complicated reasons (in fact, I do not need to know why at all)! Just let me know you need an extension. I trust you to not abuse this policy.

2. **Class participation (20%)**
Class participation is highly encouraged and can take various forms:

   - asking questions in class
   - participating during in-class discussions (in person or by posting questions in the Zoom chat)
   - participating in office hours
   - pointing out problems with the lecture notes (e.g. typos, confusing wording, broken code)
   - frequent use of the course Discord channel (this includes both asking questions of your own and answering fellow students' questions)
   - posting links to relevant online material (e.g. tutorials or news paper articles)
   - participating in the Linguistics Department research events (e.g. colloquia)

3. **Final Project (20%)**
A final project will be due by the end of the semester. I encourage students to work in groups with different backgrounds (e.g. a linguist and a CS/DS major), but I won't enforce it and I will also allow individual projects. Final projects can be of three types:

   - Type 1: Implement your own software targeting an NLP problem (max. 3 students per group)
   - Type 2: Conduct an extensive, linguistically informed evaluation (comprehensive of error analysis) of an existing NLP package (max. 2 students per group)
   - Type 3: Formulate a research proposal for a computationally informed linguistic question, and conduct an extensive (critical) literature review on the topic (minimum 10 pages, max. 1 student per group)

Student projects (topic and group composition) must be approved by the instructor. Students are encouraged to attend office hours during the first half of the semester to brainstorm ideas. An official proposal (max. 500 words) will be due at the end of week 8. Rubrics for each projects will be shared with the students following the approval of their proposals.

4. **Extra-credits**

It is possible for you to earn up to 5% of positive increase to your final grade in the form of extra credits. This can be accomplished by

- attending a Department of Linguistics colloquium talk
- participating in office hours
- pointing out problems with the lecture notes (e.g. typos, confusing wording, broken code)
- frequent use of the course Discord channel (this includes both asking questions of your own and answering fellow students' questions)
- suggesting extra topics for lectures
- posting links to relevant online material (e.g. tutorials or news paper articles)
- participating in one (or more) experimental Linguistic stud(ies) (if available)
- other extra-credit opportunities as arise during the semester

To sign up for a linguistic study, please visit: http://speechlab.utah.edu/participate. php. **Note that you are strongly encouraged to sign up early, rather than wait to the end of the semester to participate.** If for any reason you are not comfortable participating in Linguistics studies for credit in this course, an alternative assignment will be made available. If you need accommodation of any kind, have a question about accommodation, and/or prefer to complete an alternative assignment for any reason, please create a Sona account and then email at speechacquisitionlab@gmail.com. Arrangements for alternative assignments must be made before the current semester's add/drop deadline. If no study appointments are available, check back later, as new appointments are added periodically. If you have any questions or concerns regarding the Linguistics Study Pool or the Sona sign-up system, please contact the Linguistic Study Pool coordinators at speechacquisitionlab@gmail.com.

If you attend a colloquium talk, to receive credit you have to write a 1 page (typed double-spaced) summary of what the talk was about, including your general opinions on the presentation. Comments do not have to pertain to the scientific content of the talk (although that is obviously welcome), but can be about presentation style/clarity/etc. This report is officially due on the last day of class, but may be turned in on Canvas as soon as it is complete. For information about the time and location of Linguistics talks by department faculty and visiting faculty from other institutions visit the Department webpage.

4.1 University of Utah Grading Scale

Letter	Scoring
A	100% - 94%
A-	93.9% - 90%
B+	89.9%–87%
B	86.9%–84%
B-	83.9% - 80%
C+	79.9%–77%
C	76.9%–74%
C-	73.9% - 70%
D+	69.9%–67%

Letter	Scoring
D	66.9%–64%
D-	63.9% - 60%
E	59.9%–0%

5 (Tentative) Schedule by week (Subject to change)

As a coding introduction, this course relies heavily on active-learning strategies. Students are strongly encouraged to bring their own laptops to class, and will be prompted to take part to the lecture development. If you need help procuring affordable technology (or happen to have tech related issues — e.g., lack of stable internet) through the semester, let me know! The University can help!

Week	Date	Topic	HW # Due (Sunday)
1	01/11	Syllabus; Python Primer	
	01/13	Big Picture	
2	01/18	Societal Impact of NLP	
	01/20	Fairness in NLP/AI	HW 0 Due
3	01/25	Formal Languages and Regular Expressions	
	01/27	Dialogue Systems	HW 1 Due
4	02/01	Word Prediction	
	02/03	Language Models with n-grams	
5	02/08	Spell-checkers and model evaluation	
	02/10	Spell-checkers: representing words	HW 2 Due
6	02/15	Finite State Phonology	
	02/17	Finite State Phonology; Intro to Pynini	
7	02/22	Finite State Morphology	
	02/24	Finite State Morphology	HW 3 Due
8	02/22	Beyond words: Syntax	
	02/24	Syntax	
9	03/01	Final project/proposal planning	
	03/04	Final project/proposal planning	HW 4 Due
10	03/08	Spring Break	
	03/10	Spring Break	
11	03/15	Syntax [cont.]	
	03/17	Syntax [cont.]	Project Proposal Due
12	03/22	Meaning	
	03/24	Meaning	HW 5 Due
13	03/29	Meaning	
	03/31	Meaning	
14	04/05	Comp. Psycholinguistics: A Primer	HW Due
	04/07	Final project presentations	
15	04/26	Wrap Up	

Week	Date	Topic	HW # Due (Sunday)
	04/28	No classes	
16	05/03	Finals week	
	05/05	Finals week	Project due

Topics and deadlines are tentative and might be changed as the semester progresses.

This class has no Final exam. The final project and an optional class survey will be due for Finals week. HW assignments are unlocked each Tuesday morning and, unless stated otherwise, are due the following **Monday at 11:59pm**.

6 General Policies

6.1 Contacting me

- Emails should be sent to aniello.desanto@utah.edu
- Please put LING 5300 as part of the subject of the email, and *never* leave the subject empty!
- Reply time < 24h in simple cases, possibly more if meddling with bureaucracy is involved.
- Do not expect quick replies to emails sent overnight and/or during the weekend.

6.2 Office Hours

- Office hours will take place twice per week either in my office in LNCO, or in a dedicated Zoom room (link available on Canvas).
- You do not need to schedule an appointment for these, you can just show up (but an heads up is welcome if you anticipate a long discussion).
- Extra office hours are also always available by appointment.

6.2.1 A Note on Office hours

You might feel like you need to have a specific question (e.g., about a homework) in order to come to office hours. This is *not* true. **Office hours are a space for us to discuss any issue/concern you might be having with the class, and you are always welcome regardless of whether you have a precise question or not.** Thus, you also do not need an appointment to show up to office hours. However, if you anticipate a longer meeting, it might be helpful for you to email me in advance so that we can set apart enough time and avoid collisions with other students.

6.3 Names/Pronouns

Class rosters are provided to the instructor with the student's legal name as well as *Preferred first name* (if previously entered by you in the Student Profile section of your CIS account, which managed can be managed at any time). While CIS refers to this as merely a preference,

I will honor you by referring to you with the name and pronoun that feels best for you in class or on assignments. Please advise me of any name or pronoun changes so I can help create a learning environment in which you, your name, and your pronoun are respected. If you need any assistance or support, please reach out to the LGBT Resource Center: https://lgbt.utah.edu/campus/faculty_resources.php

7 Covid Policies & Online component

Generally, I trust your judgment. These are weird times, and I hope we can all collaborate to create an environment that is safe for everyone. I welcome suggestions that can help maximize what you get out of this course while minimizing risks to your (and your families!) mental and physical health.\

- Masks
    - As per CDC guidance, and the new Salt Lake County order, **mask are required** to attend class in person.
    - Please be respectful of me and your fellow classmates, and wear a mask without me having to call your attention to it.
    - I will keep masks in class in case you need one.
- Get Vaccinated and Boosted!
    - Getting a vaccine is not only an act of protection for yourself, but it is a kindness towards vulnerable members of our community
    - You can get free vaccination on campus! More information here: https://alert.utah.edu/covid/vaccine/
- Testing and Staying Home
    - If you feel sick, please stay home! I will work with you to arrange ways to attend the class virtually and catch up on missed materials.
    - Students must self-report if they test positive for COVID-19 via this website: https://coronavirus.utah.edu/
    - UofU offers free, asymptomatic testing, you can find information about that and vaccine availability here: https://healthcare.utah.edu/coronavirus/reserve-test/
- Virtual Attendance (to be updated first week of classes)}
    - This class is scheduled for in-person attendance. If classroom technology allows, Zoom Meetings will also take place during normal class time and **synchronously**. You are welcome to join class via these online meetings instead of in person.
    - Meetings will be also recorded and also made available on Canvas, but these are not a substitute for classroom participation.
    - If you need asynchronous accommodations, please discuss them with me explicitly.

8 Accommodations

Accommodations will be considered on an individual basis and may require documentation.

8.1 Americans with Disabilities Act (ADA)

The University of Utah seeks to provide equal access to its programs, services and activities for people with disabilities. If you will need accommodations in the class, reasonable prior notice needs to be given to the Center for Disability & Access, 162 Olpin Union Building, 801-581-5020. CDA will work with you and the instructor to make arrangements for accommodations. All written information in this course can be made available in alternative format with prior notification to the Center for Disability & Access. The website for the Center for Disability and Access: https://disability.utah.edu

8.1.1 Extreme personal circumstances

Please contact your instructor as soon as possible if an extreme personal circumstance (hospitalization, death of a close relative, natural disaster, etc.) is interfering with your ability to complete your work.

8.1.2 Religious Practice

To request an accommodation for religious practices, contact your instructor at the beginning of the semester.

8.1.3 Active Duty Military

If you are student on active duty with the military and experience issues that prevent you from participating in the course because of deployment or service responsibilities, contact your instructor as soon as possible to discuss appropriate accommodations.

8.1.4 An additional note on diversity and accessibility

It is my intent that students from all diverse backgrounds and perspectives be well served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that students bring to this class be viewed as a resource, strength and benefit. It is my intent to present materials and activities that are respectful of diversity: gender, sexuality, age, socioeconomic status, ethnicity, race, and culture.

I also aim to maximize accessibility for all my lecture materials. However, it is probable that I miss or underestimate some issues. Your suggestions are encouraged and appreciated.
Please let me know ways to improve the effectiveness of the course for you personally or for other students or student groups. In addition, if any of our class meetings conflict with your religious events, please let me know so that we can make arrangements for you.

9 University Policies

9.1 Drop/Withdraw Policies

Students may drop a course within the first two weeks of a given semester without any penalties.

Students may officially withdraw (W) from a class or all classes after the drop deadline through the midpoint of a course. A "W" grade is recorded on the transcript and appropriate tuition/fees are assessed. The grade "W" is not used in calculating the student's GPA.

For deadlines to withdraw from full-term, first, and second session classes, see the U's Academic Calendar.

10 Plagiarism and Cheating

It is assumed that all work submitted to your instructor is your own work. When you have used ideas of others, you must properly indicate that you have done so.

Plagiarism and cheating are serious offenses and may be punished by failure on an individual assignment, and/or failure in the course. Academic misconduct, according to the University of Utah Student Code,

"...Includes, but is not limited to, cheating, misrepresenting one's work, inappropriately collaborating, plagiarism, and fabrication or falsification of information...It also includes facilitating academic misconduct by intentionally helping or attempting to help another to commit an act of academic misconduct."

For details on plagiarism and other important course conduct issues, see the U's Code of Student Rights and Responsibilities.

11 Course Materials Copyright

The Content is made available only for your personal, noncommercial educational and scholarly use. You may not use the Content for any other purpose, or distribute, post or make the Content available to others, unless you obtain any required permission from the copyright holder. Some Content may be provided via streaming or other means that restrict copying; you may not circumvent those restrictions. You may not alter or remove any copyright or other proprietary notices included in the Content.

Please see the Code of Student Rights and Responsibilities, Section III.A.5 regarding use and distribution of class Content and materials. https://regulations.utah.edu/academics/6-400.php Section III.A.5. prohibits the following: Sale or distribution of information representing the work product of a faculty member to a commercial entity for financial gain without the express written permission of the faculty member responsible for the course. ("Work product" means original works of authorship that have been fixed in a tangible medium and any works based upon and derived from the original work of authorship.)

11.1 Safety at the U

The University of Utah values the safety of all campus community members. You will receive important emergency alerts and safety messages regarding campus safety via text message.

For more information regarding safety and to view available training resources, including helpful videos, visit safeu.utah.edu.

To report suspicious activity or to request a courtesy escort, contact: Campus Police & Department of Public Safety, 801-585-COPS (801-585-2677), dps.utah.edu

11.2   Wellness at the U

Your personal health and wellness are essential to your success as a student. Personal concerns like stress, anxiety, relationship difficulties, depression, or cross-cultural differences can interfere with a student's ability to succeed and thrive in this course and at the University of Utah.

Please feel welcome to reach out to your instructor or TA's to handle issues regarding your coursework.

For helpful resources to manage your personal wellness and counseling options, contact: wellness.utah.edu or womenscenter.utah.edu

11.3   Addressing Sexual Misconduct

Title IX makes it clear that violence and harassment based on sex and gender (which Includes sexual orientation and gender identity/expression) is a civil rights offense subject to the same kinds of accountability and the same kinds of support applied to offenses against other protected categories such as race, national origin, color, religion, age, status as a person with a disability, veteran's status or genetic information. If you or someone you know has been harassed or assaulted, you are encouraged to report it to the Title IX Coordinator in the Office of Equal Opportunity and Affirmative Action, 135 Park Building, 801-581-8365, or the Office of the Dean of Students, 270 Union Building, 801-581-7066. For support and confidential consultation, contact the Center for Student Wellness, 426 SSB, 801-581-7776. To report to the police, contact the Department of Public Safety, 801- 585-2677(COPS).

11.4   Campus Safety

The University of Utah values the safety of all campus community members. To report suspicious activity, call campus police at 801-585-COPS (801-585-2677). You will receive important emergency alerts and safety messages regarding campus safety via text message. For more information regarding safety and to view available training resources, including helpful videos, visit safeu.utah.edu.

12   Changes to the Syllabus

This syllabus is not a contract. It is meant to serve as an outline and guide for your course. Please note that your instructor may modify it to accommodate the needs of your class.

You will be notified of any changes to the Syllabus.