Course Description and goals:

This course is an introduction to computational linguistics. Our goals include the following:

- Learn basic Python programming and core natural language processing (NLP) tools using the Natural Language Toolkit (NLTK)
- Understand how analysis of language can be done computationally
- Understand why automated analysis of language is so challenging
- Learn core methods in computational linguistics, including both rule-based and statistical methods

Textbooks:

1. Required: Python Crash Course (PCC), Eric Matthes (https://nostarch.com/pythoncrashcourse)

2. Useful resources:
   a. Natural Language Processing with Python (NLTK), Steven Bird, Ewan Klein & Edward Loper:
      A complete new edition of this text is available for free online at http://www.nltk.org/book
      The online version is up-to-date and works with the latest edition of the NLTK. A printed version of the text is only available for the First Edition.

   Important: The printed version of the text is the First Edition. The online version – the third edition - is up-to-date and works with the latest edition of the NLTK. Either version of the book is okay. If you buy the hard copy, you’ll need to occasionally check the online version for small changes. Less than 10% of the book has changed across editions, but some of the changes are important to know about in order to get your code to run.

3. Supplemental text (optional, and freely available online):

4. Suggested text for those interested in learning more: Speech and Language Processing – Daniel Jurafsky & James H. Martin. Second edition, 2009. Pearson Prentice Hall. This is one of the foundational texts for computational linguistics. We will read only selected chapters from this text, and most of those can be found online. HOWEVER, if you anticipate a future in computational linguistics, you will most certainly want to own this book.
Class communications – Canvas, Canvas, Canvas

1. All class communications will happen via Canvas. IT IS YOUR RESPONSIBILITY TO STAY INFORMED – check Canvas at least every weekday. You can also use Canvas to ask questions, discuss the readings, stay in touch with your classmates, share interesting things you find online, etc.

2. All assignments will be posted on Canvas, with all the details you need to complete them. Unless otherwise specified, please submit all assignments via Canvas.

3. Course schedule: The course schedule will live in a shared Google spreadsheet, which you can access through the Canvas site. This schedule may change from time to time – IT IS YOUR RESPONSIBILITY TO MONITOR THE COURSE SCHEDULE AS WELL.

Attendance & Participation
Attendance is REQUIRED, but each student can have two free absences, no need to explain, just let us know in advance. If a situation develops such that you will need to miss more class than this, please make an appointment to discuss this with me in person, or email me if in-person isn’t possible.

Course credit requirements and grading
1. 32% - Frequent exercises (almost every week, 4 points each, lowest 2 scores dropped)
2. 15% - First pair programming project [due October 13]
3. 20% - Midterm [November 4]
4. 25% - Final small group programming project: code, report & presentation [presentations December 2 and 4, code and report due December 9]
5. 3% - Cheat sheet assignment [due October 28]
6. 5% - Participation: you-trys, reading, discussion, etc.

Schedule:
This course is divided up into three units:
1. Python Bootcamp (for linguists): August 26 – October 4
2. Working with language data in Python: October 7 – November 4
3. Computational models for linguistic analysis: November 6 – December 6

Exercises:
Learning to program is all about practice, practice, practice! The point of the weekly exercises is to give you plenty of opportunities to practice, and for you to get thoroughly comfortable with the process of turning an idea into a bit of Python. Please submit these, even if you know you got things wrong, even if you couldn’t finish everything! You’ll get plenty of points for the parts you could complete.

Late assignments:
Late assignments will not be accepted, unless you’ve made prior arrangements with me.

Linguistics Colloquia and COI Discovery Series Talks: The Department of Linguistics and the College of Information each offer a series of talks from experts in the field. Please plan to attend these if at all possible – I’ll be sure to announce the talks, and they will also appear on the course schedule.
**Class behavior (University policy)**

Student behavior that interferes with an instructor’s ability to conduct a class or other students' opportunity to learn is unacceptable and disruptive and will not be tolerated in any instructional forum at UNT. Students engaging in unacceptable behavior will be directed to leave the classroom and the instructor may refer the student to the Center for Student Rights and Responsibilities to consider whether the student's conduct violated the Code of Student Conduct. The university's expectations for student conduct apply to all instructional forums, including university and electronic classroom, labs, discussion groups, field trips, etc. The Code of Student Conduct can be found at [www.unt.edu/csrr](http://www.unt.edu/csrr).

**Academic Integrity**

Please make sure that all the work you submit in this course is your own!!!

It is your responsibility to know the definitions and consequences of plagiarism, cheating etc. as outlined in the Code of Student Conduct and Discipline, available in the Undergraduate Catalog and online at [http://catalog.unt.edu/content.php?catoid=5&navoid=244#Student_Standards_of_Academic_Integrity](http://catalog.unt.edu/content.php?catoid=5&navoid=244#Student_Standards_of_Academic_Integrity)

**Disability Accommodation Statement**

“The University of North Texas makes reasonable academic accommodation for students with disabilities. Students seeking accommodation must first register with the Office of Disability Accommodation (ODA) to verify their eligibility. If a disability is verified, the ODA will provide you with an accommodation letter to be delivered to faculty to begin a private discussion regarding your specific needs in a course. You may request accommodations at any time, however, ODA notices of accommodation should be provided as early as possible in the semester to avoid any delay in implementation. Note that students must obtain a new letter of accommodation for every semester and must meet with each faculty member prior to implementation in each class. For additional information see the Office of Disability Accommodation website at [http://www.unt.edu/oda](http://www.unt.edu/oda). You may also contact them by phone at 940.565.4323.”

Students with any special needs or disabilities (including learning disabilities) should inform me of their needs as soon as possible so that proper accommodations can be made.

**SEXUAL DISCRIMINATION, HARASSMENT, & ASSAULT**

UNT is committed to providing an environment free of all forms of discrimination and sexual harassment, including sexual assault, domestic violence, dating violence, and stalking. If you (or someone you know) has experienced or experiences any of these acts of aggression, please know that you are not alone. The federal Title IX law makes it clear that violence and harassment based on sex and gender are Civil Rights offenses. UNT has staff members trained to support you in navigating campus life, accessing health and counseling services, providing academic and housing accommodations, helping with legal protective orders, and more.

UNT’s Dean of Students’ website offers a range of on-campus and off-campus resources to help support survivors, depending on their unique needs: [http://deanofstudents.unt.edu/resources_0](http://deanofstudents.unt.edu/resources_0). Renee LeClaire McNamara is UNT’s Student Advocate and she can be reached through e-mail at SurvivorAdvocate@unt.edu or by calling the Dean of Students’ office at 940-565-2648. You are not alone. We are here to help.