# Syllabus for LIN 637

## Computational Linguistics 2 – Spring 2021
## MWF 09:15-10:10 via ZOOM

Last Updated: January 6, 2021

| | |
|---|---|
| CLASS ZOOM Meeting ID: 936 0394 0635 | Password: 736397 |

| |
|---|
| Course Website: https://github.com/Compling2-Spring2021-SBU/main |

(Boxes are clickable)

# 1    Course Outline

## 1.1    Description

This course is an introduction to the theoretical foundation of computational linguistics. The course emphasizes the importance of algorithms, recursive data structures, formal language theory, and logic for understanding natural language and the development of new tools and software applications. Empirical phenomena in phonology and syntax are sampled from a variety of languages to motivate and illustrate the use of concepts such as strictly local string languages, tree transducers, and semirings. Students will develop familiarity with the literature and tools of the field.

The course also serves as a bridge from introductory courses in linguistics (Syntax 1, Phonology 1, Phonetics) and computational methods (Statistics, Mathematical Methods in Linguistics, Computational Linguistics 1) to advanced courses and seminars in computational/mathematical linguistics. In contrast to the NLP courses offered by the department of computer science, this course focuses on studying the properties of natural language from a computationally informed perspective. The question is not how computers can solve language-related tasks, but how language can be conceptualized as a computational problem. This emphasis is also reflected in the selection of topics for this course.

## 1.2    List of topics

- *Phonology and Morphology*

    - The role of formalization
    - String languages
    - Rational languages
    - String transductions

  – Subregular classes

  – Computational perspectives on phonological theories like SPE and OT

- *Syntax*

  – Tree languages

  – Rational tree languages

  – Syntax is more complex than phonology (but how?)

  – Mildly context-sensitive formalisms (TAG, MGs)

  – Tree transductions

  – Regular representations of MCS formalisms

## 1.3   Prerequisites

The only official prerequisite is Computational Linguistics 1 (Lin 537) or comparable programming skills in Python. Python will be used to illustrate formal concepts, and some of the homeworks will require you to implement an algorithm or procedure in Python. Prior experience with git and markdown is useful for the homeworks but not required.

It is also helpful to have some basic familiarity with linguistics (phonemes, phrase structure rules, syntactic trees) and mathematics (sets, functions, relations, and first-order logic as covered in Semantics 1, for instance).

# 2   Graded Component

- **Homework**

  – Approximately 5-7 exercises, programming assignments, or critical evaluations of assigned readings

  – Homework submission and grading is done via github.

  – Collaboration on homework problems is encouraged as long as you write up the solutions by yourself, using your own words, examples, notation, and code.

  – 2/3 of final grade

- **Project**

  – You will develop a final project in this class.

  – Projects include theoretical research, program/software/tool development, or a software-aided study of language data.

  – A one page project proposal is due Friday March 19.

  – The project itself is due on Friday May 14.

  – 1/3 of final grade

- **Readings**

    - Readings will be given regularly and made available on the course website.
    - It is presupposed in the lectures that you have done the required readings.

- **Workload per Credits**

    - *0 credits*: attend (but I highly recommend that you at least read the assigned papers as they will be important for following the lectures)
    - *1 credit*: attend, participate, readings,
    - *2 credits*: attend, participate, readings, homework
    - *3 credits*: attend, participate, readings, homework, project

# 3   Online Component

This class takes place synchronously over zoom. The only requirements are a computer, laptop, or phone equipped with a typical broswer, microphone and/or camera, and audio output.

I will have office hours Tuesdays, 1-4pm, and by appointment. You are encouraged to drop by even for short questions.

Office hours are held at the "personal zoom meeting room" provided by Stony Brook. The link is here, and is also on the course website.

https://stonybrook.zoom.us/j/4758144071

This class also uses some online tools to facilitate homework collaboration and submission, student discussions, and dynamic lecture evaluation.

- **Homework submission**

    *How it works:* Homeworks will distributed via a github repository. You can fork this repo and upload your own code, or checkout other students' forks to see how they dealt with the problem. In order to submit a homework you upload your solution to your fork and issue a pull request. After the due date, I'll upload my solution to the repository.

    *Why we do it:* This setup mimics the modern workflow in collaborative development projects. Git is one of the best-known version control systems, and github is the biggest online service for hosting git repositories. Familiarity with version control systems is an essential job requirement for computational linguists, and it is also very helpful for academic work. See this discussion on Stackflow for some ideas how git can be used in conjunction with Latex: http://stackoverflow.com/questions/6188780/git-latex-workflow

    *What you'll need:* A github account (the free tier is enough) and a way of uploading your code to a github repository. Linux users will install git via the command line, whereas Windows and Mac users should download and insatall the github app, which comes with a nice GUI.

## University Policies and Services

**Student Accessibility Support Center Statement**  If you have a physical, psychological, medical, or learning disability that may impact your course work, please contact the Student Accessibility Support Center, 128 ECC Building, (631) 632-6748, or at `sasc@ stonybrook.edu`. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential.

**Academic Integrity Statement**  Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty are required to report anya suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty please refer to the academic judiciary website at `http://www.stonybrook.edu/ commcms/academic_integrity/index.html`

**Critical Incident Management**  Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of University Community Standards any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures. Further information about most academic matters can be found in the Undergraduate Bulletin, the Undergraduate Class Schedule, and the Faculty-Employee Handbook.

**Additional resources**  To access mental health services, call Counseling and Psychological Services at 631-632-6720; Counselors are available to speak with 24/7. For updated information on the Academic Success and Tutoring Center please check `www.stonybrook. edu/tutoring` for the most up-to-date information. For IT Support: Students can visit the Keep Learning website at `https://sites.google.com/stonybrook.edu/keeplearning` for information on the tools you need for alternative and online learning. Need help? Report technical issues at `https://it.stonybrook.edu/services/itsm`or call 631-632-2358.

For information on Library services and resources please visit the Continuity of Library Operations guide.