

Syllabus for LING 334
Introduction to Computational Linguistics
Winter 2009

Course:

Class time: TuTh 3:30–4:50
Class location: University Library Mac Lab, B182
ROOM CHANGE FOR WEEK 2:
Plan to meet in **Cresap101** on **Tuesday January 13th**
and **Thursday January 15th**

Instructor: Hannah Rohde

Office: 2016 Sheridan Road, room 103
Phone: 491-4860
Email: hannah@northwestern.edu
Office Hours: ThTh 1:00-2:00 and by appointment

What's this course all about?

This course aims to provide an introduction to the fundamental concepts of Computational Linguistics through the window of experimental Psycholinguistics. Topics will be introduced in the context of a well-known psycholinguistic model or result and then will be further explored with computational tools and approaches. The course will cover methodology for empirical linguistic analysis and natural language processing involving the use of text corpora, language modeling, part-of-speech tagging, syntactic analysis, and semantic interpretation.

The central goal is to familiarize students with core questions and approaches in the field while providing a toolbox of practical programming skills. Theoretical material on such topics as n-gram models, neighborhood density, context-free grammars, etc. will be supplemented with laboratory exercises and mini-projects to give students hands-on experience in the use of corpora and the implementation of algorithms.

What am I going to do in this class?

The course will be taught through:

Readings and lectures to provide an overview of natural language processing. We'll look at computational models of language, comparing those with psycholinguistic results on human processing.

Hands-on programming experience. Homeworks and in-class labs will pose questions about language processing and the properties of our language input. Trying to answer those questions will provide an opportunity to gain programming skills and to learn about computational tools.

Formulating your own research questions. Working alone or in a group, you'll design a project which explores a psycholinguistic question that using one or more computational techniques.

No programming experience will be assumed. An interest in and some exposure to Linguistics is definitely a plus.

Website: Blackboard. You should be able to access it by going to `courses.northwestern.edu`.
Let me know if you have trouble.

Resources

There is no required textbook but some useful resources may include:

- Manning & Schütze. Foundations of Statistical Natural Language Processing. MIT Press.
- Jurafsky & Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice-Hall.
- <http://perldoc.perl.org/>

Assessment

Course Component	% of Final Grade
Participation in class meetings	20%
Homework assignments (4)	60%
Final Project	20%
$\left(\begin{array}{l} \text{Initial proposals} \\ \text{Project outline} \\ \text{Paper} \end{array} \right)$	$\left(\begin{array}{l} 2\% \\ 6\% \\ 12\% \end{array} \right)$

1. Participation in class meetings

Plan to attend all classes. Your presence, questions, discussion, and general involvement will define this part of your grade.

In-Class Exercises: On some days, especially early in the quarter, part of class time will consist of lab exercises. On those days, participation will include work (usually with a partner) on a task to write a short program or modify an existing one.

2. Homework assignments

These programming tasks and problem sets will give you practice in computational methodology as well as asking you to relate corpus data to behavioral data.

Collaboration: You are encouraged to work together in pairs or small groups when discussing the assignments. However, all write-ups must be your own. I also ask that you note on each assignment who you have worked with.

3. Final Project

The course focuses on psycholinguistic questions that can be addressed with computational tools. The final project will focus on a question that is of interest to you. You will work on pilot research projects in groups or on your own. These are not intended to be complete projects; rather, they are opportunities to identify a question that interests you and to step through the initial planning of a project that could answer that question (including, if possible, preliminary data collection).

During Week 5 we will brainstorm about project ideas. I'll give you some suggested topics, but you are welcome to bring in ideas that are of interest to you and your own research. Based on your **initial proposal** and your interest in working in a group, I will assign you a partner with similar interests who you will work with to create a **project outline** and a joint final **paper**.

Initial proposals: In Week 6, you will submit a one-page write-up that lists: (i) 2 papers on a topic of interest to you, (ii) a description of the issues raised in these papers, and (iii) one or more questions that you have regarding these papers and a proposal for how you might start to address the questions with the tools we have covered in this class

Project Outline: In Week 8, you will submit a 1-2 page outline covering: (i) the question that you (and partner) are investigating, (ii) what theoretical implications an answer to this question would have, (iii) the tool(s) you will use to address this question, (iv) the resources (e.g., corpora) that you would use, (v) examples of relevant data, (vi) a plan for acquiring more data if needed, and (vii) hypotheses about the types of results you may find.

Paper: Incorporating the feedback from the outline, you will conduct preliminary research on your topic. You will then review the work in a final paper (5-10 pages), describing the questions and research design and any preliminary results. The paper is due at the beginning of finals week.

Course Policies

1. Keep up with class. The syllabus (topics, assignments, due dates) may change. These changes will be announced in class and posted on the website. It is your responsibility to keep up with them.
2. Hand in your work on time. Due dates are firm. The highest possible grade is a C for work received in the first 24 hours after the deadline; after that, I will give comments and suggestions on work turned in, but you will not receive credit for the assignment.
3. If you have a problem, meet with me early! If you've tried to complete an assignment but are stuck, come to me well before it's due. If many students have similar questions, we may use the end of a class period to work on the assignment (especially if it involves programming and lab time would be of use).
4. Working together is encouraged but final write-ups must be your own. Working together will be a big part of in-class work and is permissible on homework assignments; however, assignment write-ups and code hand-ins must be your own. If you work with someone (brainstorming, comparing program output, debugging), please note that in your write-up.

