Lecture 0

Syllabus

Course: Computational Linguistics 2	**Name:** Thomas Graf
Course#: Lin637	**Email:** lin637@thomasgraf.net
Time: TR 10:00–11:20am	**Office hours:** Tue 11:30–2:30pm
Location: Psychology A 144	**Office:** SBS N249
Course Website: tba	**Personal Website:** http://thomasgraf.net

1 Course Outline

1.1 Contents

This course serves a specific purpose in our program (see Fig. 1 on page III): it acts as the 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, our courses focus on studying the properties of natural language from a computationally informed perspective. The question is not how computers can solve linguistic tasks, but how language can be conceptualized as a computational problem. This emphasis is also reflected in the selection of topics for this course.

- **What this course is not about**
  - Computers as tools for linguistic research
  - Programming
  - Software development for natural language tasks
- **What is not covered but profits from what is covered**
  - Speech recognition
  - OCR
  - Text generation
  - Parsing
  - Semantic analysis
  - Machine translation

Wk	*Classes*	*Formal*	*Linguistics*
1	Jan 27, 29	What is computation?	Marr's Three Levels
2	Feb 3, 5	Formalizing phonology	Why formalize?
3	Feb 10, 12	Strictly local languages	Local dependencies
4	Feb 17, 19	Subregular hierarchy	How powerful is phonology?
5	Feb 24, 26	Regular languages	Abstractness
6	(DGFS)		
7	Mar 10, 12	String transductions	SPE-OT equivalence
8	(Spring Break)		
9	Mar 24, 26	Weak Generative Capacity	Phonology < Syntax
10	Mar 31, Apr 2	Tree languages	Headedness, feature percolation
11	Apr 7, 9	Local tree languages	GPSG
12	(GLOW)		
13	Apr 21, 23	Recognizable tree languages	GB
14	Apr 28, 30	TAG and MGs	Minimalist syntax
15	May 5, 7	Tree transductions	Reinterpreting the T-model

Table 1: Tentative course outline

- **List of topics**
  - *Phonology*
    - The role of formalization
    - String languages
    - Subregular hierarchy
    - Regular languages
    - Generative capacity of phonology
    - String transductions
    - 2-level morphology
    - Equivalence of SPE and OT
  - *Syntax*
    - Tree languages
    - Syntax is more complex than phonology
    - Mildly context-sensitive formalisms (TAG, MGs)
    - Tree transductions
    - Regular representations of MCS formalisms
    - Reinterpreting the T-model

A detailed outline of the course is given in Tab. 1.

1.2 Prerequisites

The only official prerequisite is Computational Linguistics 1 (Lin 537) or comparable programming skills in Python. It is also helpful to have some basic familiarity with

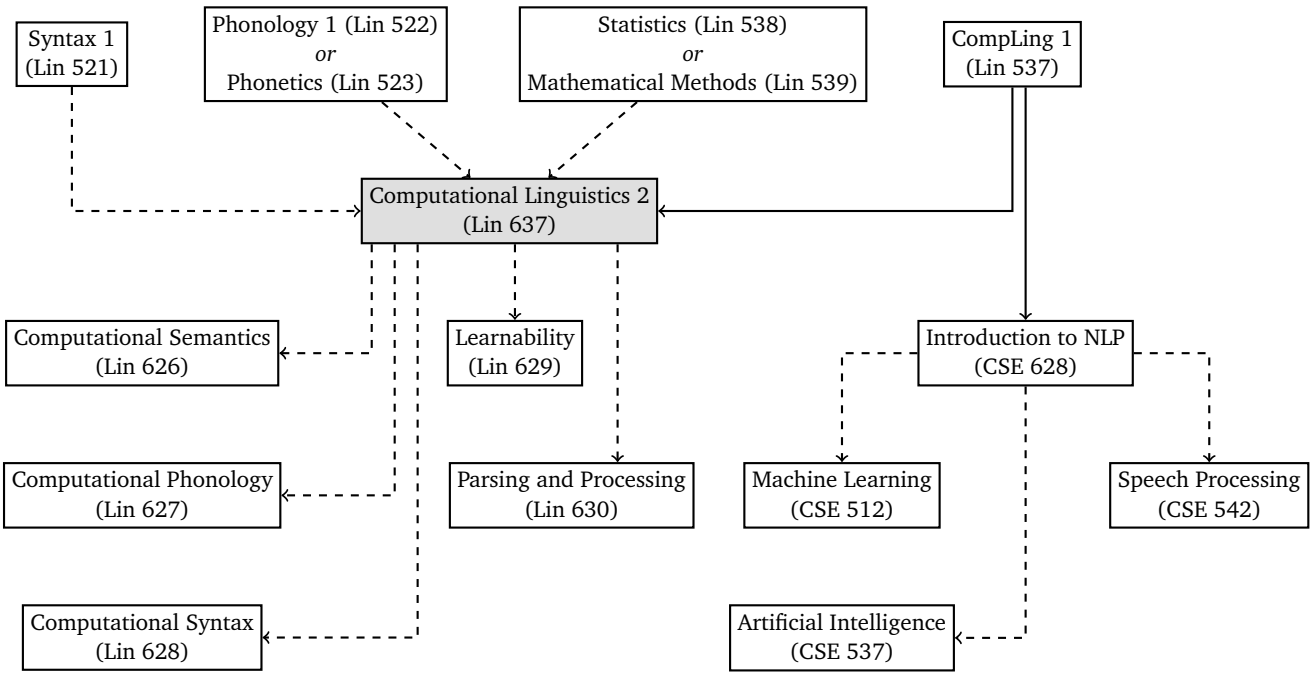

Figure 1: Computational Linguistics 2 in the curriculum (dashed lines indicate recommendations rather than prerequisites)

linguistics (phonemes, phrase structure rules, syntactic trees) and mathematics (sets, functions, relations, and propositional logic as covered in Semantics 1, for instance). You can take an online survey to identify weaknesses, and several introductory readings on these topics are available on the course website.

2 Teaching Goals

- **Practical Skills**
  - conceptualize a problem in mathematical terms
  - optimize your programs through the use of adequate algorithms and data structures (dynamic programming techniques, hash tables, etc.)
  - a more abstract and theoretically informed perspective on current tools and techniques in NLP
  - an understanding for how linguistic insights can be invoked to simplify NLP tasks
- **Research Skills**
  - assess linguistic phenomena from a computational perspective
  - evaluate linguists' claims about computational efficiency
  - basic overview of current research in theoretical computational linguistics
  - use computational concepts to identify new empirical generalizations
  - bring linguistic data to bear on computational claims
  - mathematically informed understanding of linguistic theories

3 Grading

- **Homework**
  - weekly exercises, programming assignments, or critical evaluations of assigned readings
  - a random sample of homeworks will be collected and graded
  - solutions will be made available online after the due date
  - Collaboration on homework problems is encouraged as long as you write up the solutions by yourself, using your own words, examples, notation, and code.
- **Readings**
  - about two readings per week
  - you have to collectively write a summary for each reading in the course wiki
- **Survey Squibs**
  - By the end of the course, the wiki should contain survey articles on a number of topics not covered in this course (or not covered at the same depth).
  - You have to pick a topic and write the corresponding survey article.
  - These articles should be succinct and simple enough that they are comprehensible to a researcher with little exposure to computational linguistics, yet at the same time include enough technical detail that the claims can be verified by somebody with the appropriate background. (Why this weird requirement? Because that's the recipe for writing a computational paper that can be published in a linguistics journal!)
- **Workload per Credits**
  - *3 credits*: homework, readings, squib
  - *2 credits*: homework, readings
  - *1 credit*: readings
  - *0 credits*: none, but I highly recommend that you at least read the assigned papers as they will be important for following the lectures

4 Policies

4.1 Contacting me

- Emails should be sent to lin637@thomasgraf.net to make sure they go to my high priority inbox. Disregarding this policy means late replies and is a sure-fire way to get on my bad side.
- Reply time $<$ 24h in simple cases, possibly more if meddling with bureaucracy is involved.

- If you want to come to my office hours and anticipate a longer meeting, please email me so that we can set apart enough time and avoid collisions with other students.

4.2 Special Needs

If you have any special needs that might impact your class performance (learning disabilities, impaired sight or hearing, etc.), please come to my office hours or contact me via mail so we can make suitable arrangements.