Preliminary program:

- 20.1. instructions, election of first 2 speakers
- 27.1. introductory lecture, enrollment of the rest of the speakers
- 3.2. No meeting
- 10.2. the first 2 presentations
- maximally 12 additional sessions

Presentations

- about 35 min
- Dig up the essentials in the material and present them in an understandable way.
- Present the results of a small-scale research project.
- Give out copies of your slides in the beginning of the presentation.

Research project

- Study some essential aspect of the material. For example, test the best algorithm by programming it and running it on a suitable test material.
- Report on your topic before you start, at least 1 week before your presentation!
- Write a report on the research project before 31.5. (more instructions later)

The grade:

Each of the three parts: the presentation, the summary of it, and the report of the project work is given a grade. The whole course is passed with distinction if all parts are passed with distinction.

Material

Christopher D. Manning and Hinrich Schütze: Foundations of Statistical Natural Language Processing. MIT Press, 1999.

Available in electronic bookstores (example: amazon.com $60). A copy is available for inspection in room B326.

It is assumed that everybody masters Part I (pp. 1-147): Some (simple) mathematical and linguistic foundation.

First 2 topics (to be presented 10.2.)

1. Collocations (Ch. 5)
2. N-gram models (Ch. 6)

Suggestion for research projects: Try a couple of the best approaches presented in your chapter.