

Statistical Natural Language Processing

Organizational Information

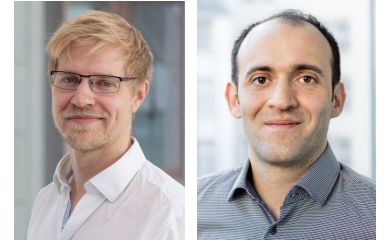
Henning Wachsmuth

<https://ai.uni-hannover.de>

Organizational Information

Course

- **Lectures.** Henning Wachsmuth
- **Tutorials.** Yamen Ajjour
- **Language.** English



Information

- **Web.** <https://www.ai.uni-hannover.de/en/teaching/courses/snlp>
- **Stud.IP.** <https://studip.uni-hannover.de/dispatch.php/course/overview?cid=b1e86956151676abbf4ad0c96d3fe952>

Time and location

- **Lectures.** Thursday 11:00–12:30, as of October 17, Appelstr. 11, A145
- **Tutorials.** Wednesday 13:15–14:45, as of October 23, Appelstr. 9A, MZ2

Notice: The lab “Human Language Technology” is mostly aligned with this course.

Consultation?

- Set up appointment via e-mail: h.wachsmuth@ai.uni-hannover.de

Organizational Information

Teaching at NLP Group, <https://ai.uni-hannover.de/en/studies/courses>

Courses

- Introduction to Natural Language Processing (bachelor, summer). NLP fundamentals, from rule-based methods to statistical methods
- **Statistical Natural Language Processing (master, winter)**. Core NLP, from statistical methods to neural methods
- Computational Argumentation (master, summer). State-of-the-art NLP, advanced methods in a specific research context

Seminars

- Natural Language Processing (bachelor, winter)
- Natural Language Generation (master, summer)

Labs and projects

- Argumentation Technology (master, summer)
- Ethical Artificial Intelligence (master, winter)

Organizational Information

This Course

Overall goals

- Learn major skills needed to tackle typical natural language processing (NLP) tasks.
- Get to know the main NLP techniques used nowadays.

Contents

- Basics of corpus construction, machine learning, and data mining
- Several statistical NLP techniques, mainly based on machine learning
- Several NLP tasks and applications

Competences

- Understanding of theory and practice of NLP
- Development of NLP methods for given tasks
- Scientific evaluation of NLP methods

Organizational Information

Outline of the Course

Introduction

1. Overview
2. Basics of NLP (recap of fundamentals, presented via video only)
3. Basics of Statistical NLP

Techniques

4. Representation Learning
5. NLP using Clustering
6. NLP using Classification and Regression
7. NLP using Sequence Labeling
8. NLP using Neural Networks
9. NLP using Transformers

Application

10. Practical Issues

Organizational Information

Course Elements

Teaching

- **Lectures.** Presentation of course content and organizational info
- **Tutorials.** Presentation of assignments and solutions, Q&A

Assignment sheets (details in first tutorial)

- **Amount.** 6 in total, bi-weekly (all pencil-and-paper)
First sheet published on Oct 21; to be submitted by Nov 4, 23:59 (UTC+1)
- **Group work.** You need to submit with 3–4 people
- **Bonus.** (a) Min. 60% of all points: exam grade + 1/3, (b) Min. 85%: + 2/3
Example for (b): grade of 2.7 is changed to 2.0; only grades < 5.0 can be improved.

Exams (details in some weeks)

- **Oral,** ~30 minutes, questions on all lecture parts, in English
- **Registration.** November 15–30, 2024
- **Dates.** First exams tentatively in first half of February 2024
A list of example questions will be provided early enough.

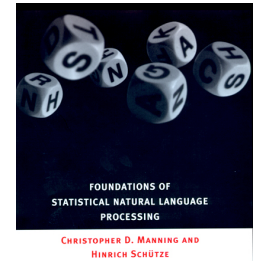
Organizational Information

Textbooks (Not Obligatory)

Foundations of Statistical Natural Language Processing

(Manning and Schütze, 1999)

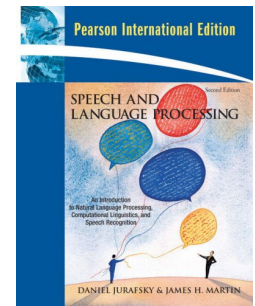
- More oriented towards computer science
- Profound basics, outdated techniques



Speech and Language Processing, 2nd edition

(Jurafsky and Martin, 2009)

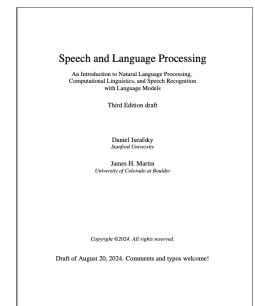
- More oriented towards computational linguistics
- Comprehensive, neural techniques not covered



Speech and Language Processing, draft of 3rd edition

(Jurafsky and Martin, 2024) → **draft freely available online**

- More oriented towards computational linguistics
- Comprehensive, up-to-date, excellently written



References

- **Jurafsky and Martin (2009)**. Daniel Jurafsky and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics. 2nd edition, Prentice-Hall, 2009.
- **Jurafsky and Martin (2024)**. Daniel Jurafsky and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd edition. Online manuscript released August 20, 2024.
<https://web.stanford.edu/jurafsky/slp3/>
- **Manning and Schütze (2009)**. Christopher D. Manning and Hinrich Schütze. Foundations of Statistical Natural Language Processing. MIT Press, 1999.