

Deep networks, transformers, explainability (ASR)



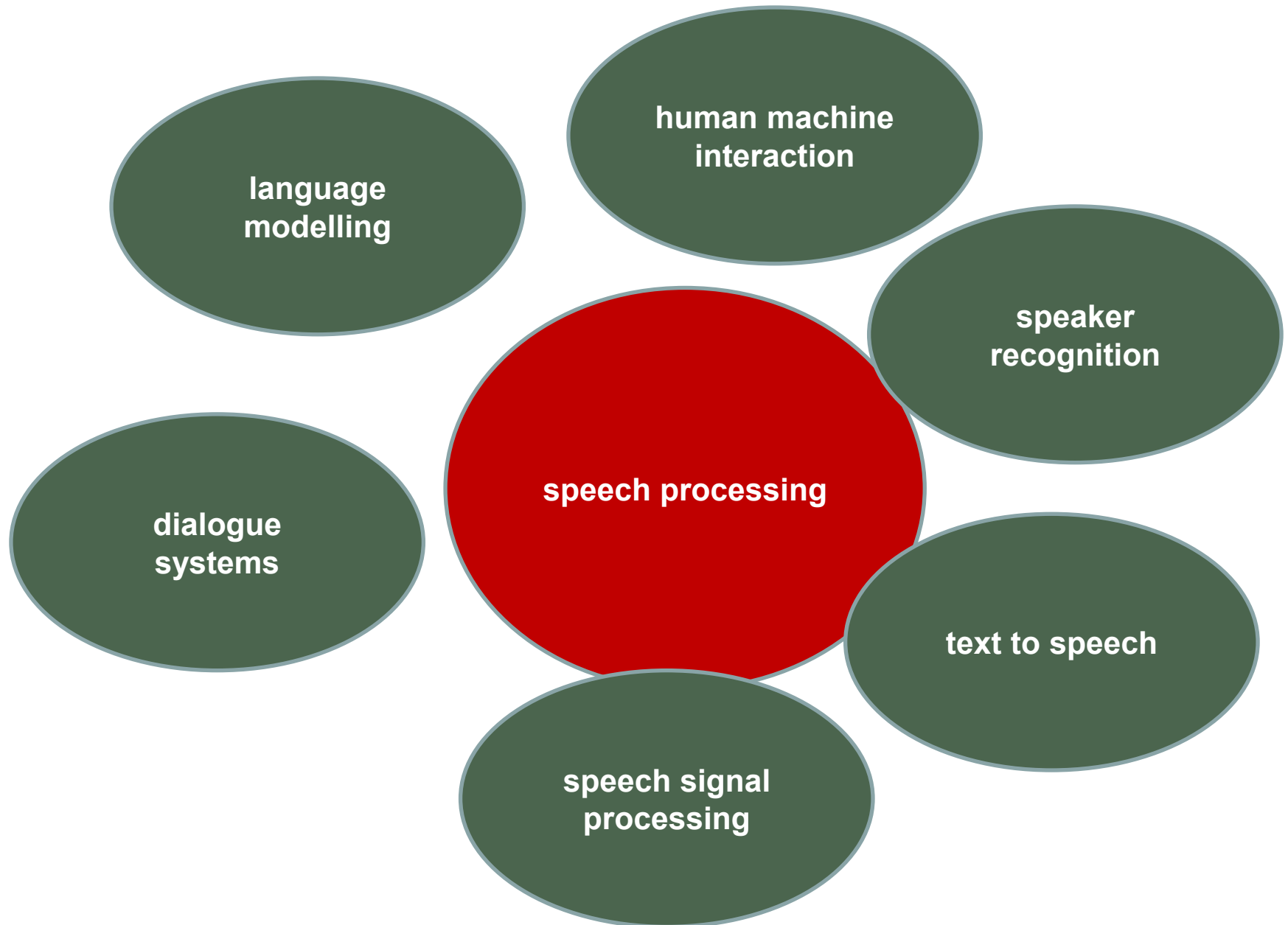
Louis ten Bosch

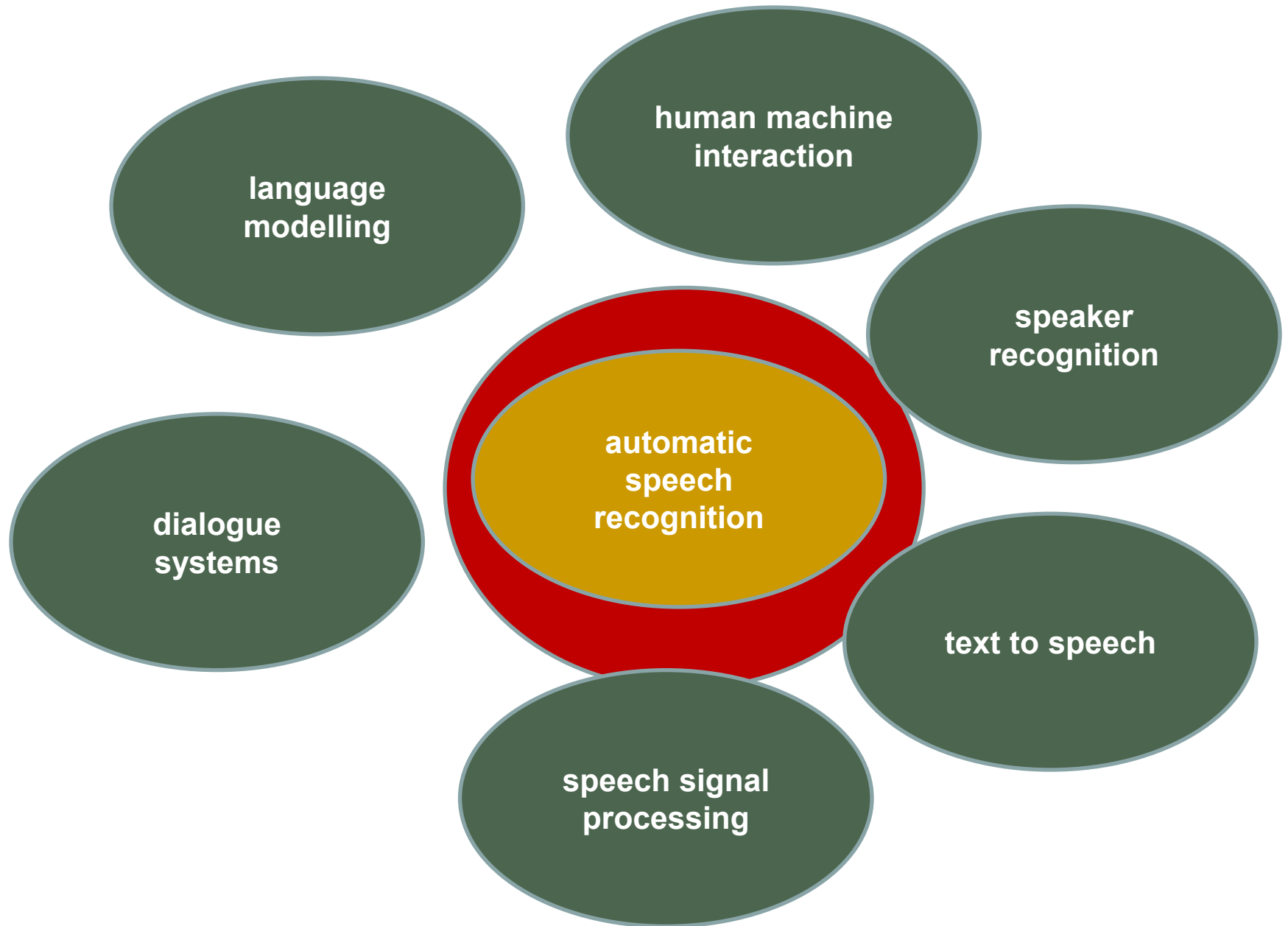
Radboud University
Nijmegen, NL

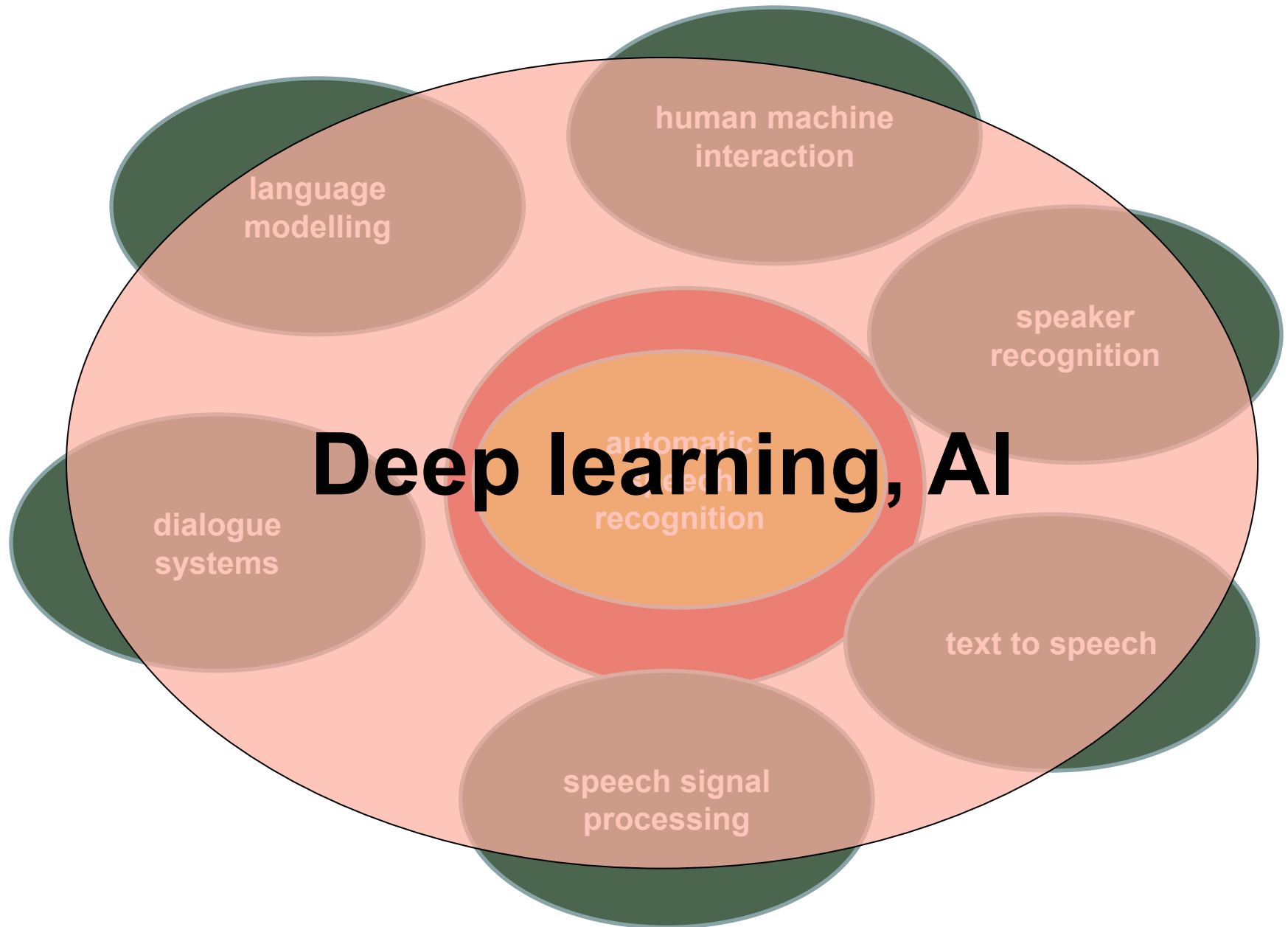
Graz, Nov-Dec 2025

ASR & deep learning is a dynamic field.
Topics and techniques are changing rapidly.









Audience?



For now:



- AI
- Data Science
- Linguistics
- ...

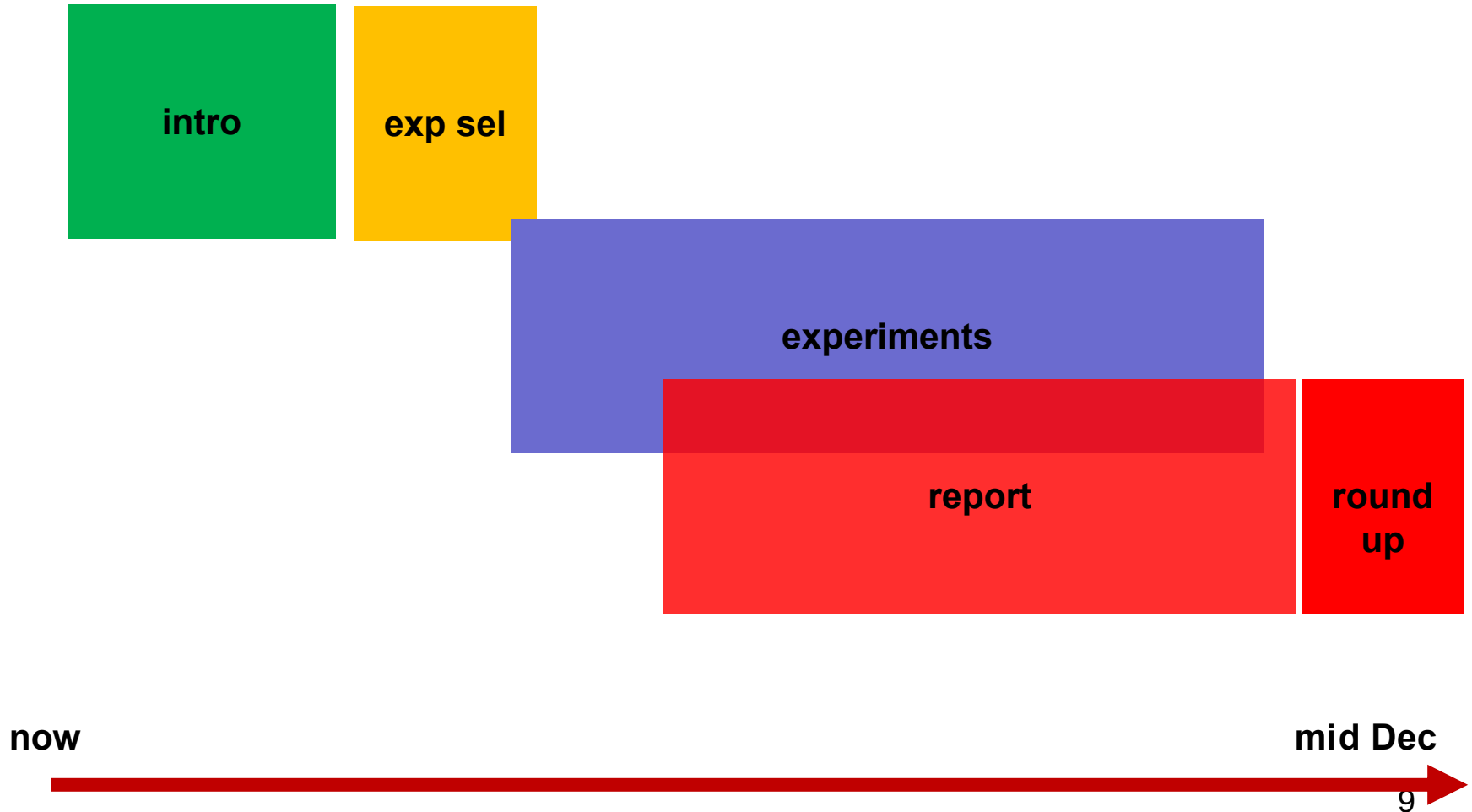
Outline of this course

- Lectures
 - Discussion
- Experiments
 - Working individually or in small groups
 - You may choose your own topic/experiment (from a list)
 - You may use your own database, or one from the web
- Brief report
 - Assessment done by Dec 15/18 (ideally)
- All material on github: [Itenbosch2/Graz_github](https://github.com/Itenbosch2/Graz_github)

Overview of the course

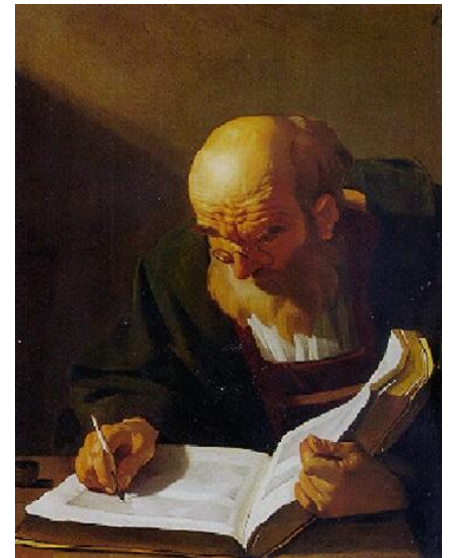
Date	Topic
Nov 10, 14	Lecture: Intro, overview course Literature
Nov 17, 20	Literature, brainstorm topics for experiment
Fr Nov 21	Literature presentations, selection of experimental topic Set-up (small) experiments
Nov 24 (2x)	Discussion ongoing experiments
Dec 1	In groups or individual: short presentations of prelim experimental results
Tue Dec 9	Finalization of experiments
Mon Dec 15	Assessments and closing

Time line of this course



Report

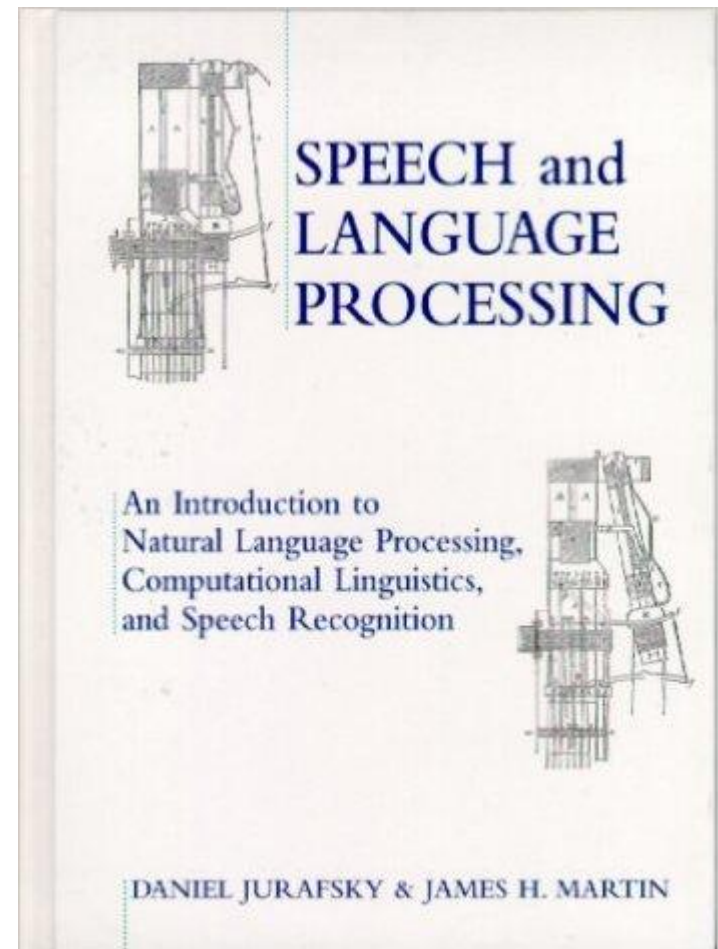
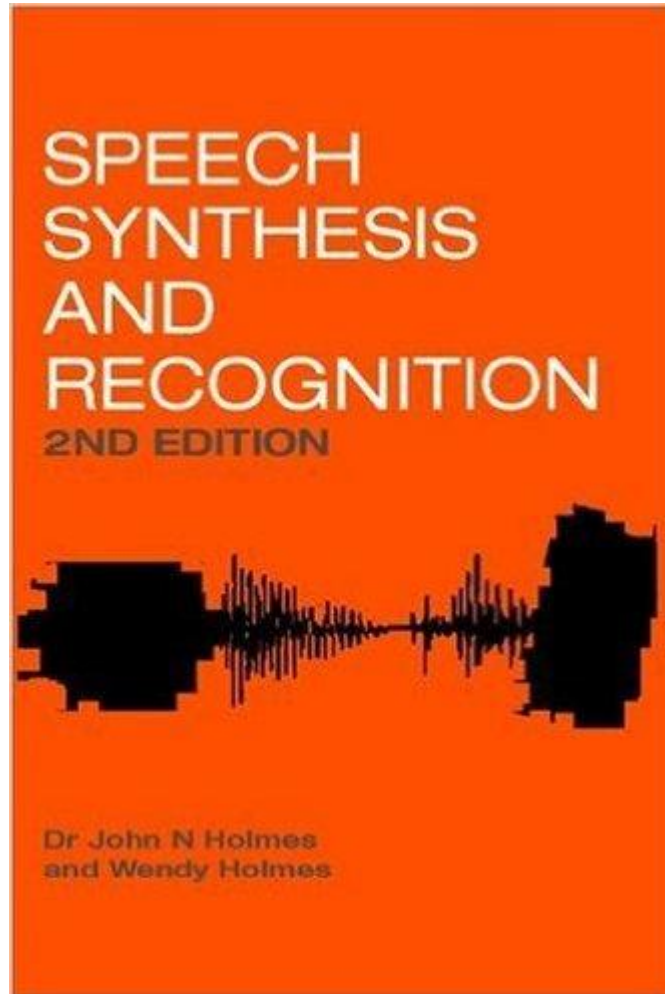
- Topic:
 - ASR/AI is way too broad to be dealt with in a few weeks, so start small
 - based on experimental work
 - (but may also be more theoretical)
 - Topics: will be discussed later
- Experiment: **individual or in a group**
- Writing: always individual
- Assessment: individual



Reading material

- Classical books
 - Holmes & Holmes
 - Jurafsky and Martin: <https://web.stanford.edu/~jurafsky/slp3/>
- Papers
 - Github
 - the “classics” (e.g., Vaswani et al., Baevski et al.)
 - Interspeech
 - https://www.isca-archive.org/interspeech_2025/index.html
 - ICASSP
- A number of key papers are given in the course
 - Additional material may vary – usually depends on your experimental project

Books – the classical ones



May be helpful: Interspeech 2025 tutorial presentations

- 1. Tutorials - Automatic Quality Assessment for Speech and Beyond https://voicemos-challenge-2023.github.io/speech-synthesis-evaluation/IS2025_tutorial.pdf
- 2. Nvidia NeMo toolkit for speech applications https://drive.google.com/file/d/1hlfFnVNrdHQ_FrvgVJRw4XNszl-aXVUz/view?usp=drivesdk
- 3. Speech Technology meets early Language acquisition: How interdisciplinary efforts benefit both the fields <https://zenodo.org/records/17018214>

As simple as it can be ...

Your three lines of python code!

```
1 from transformers import pipeline
2 classifier = pipeline("sentiment-analysis")
3 classifier(["I've been waiting for a HuggingFace course my whole life.", "I hate this so much!"])
```

```
[{'label': 'POSITIVE', 'score': 0.9598047137260437},
 {'label': 'NEGATIVE', 'score': 0.9994558095932007}]
```