



i. Name of the course Automatic Speech Recognition/Forced Alignment
ii. Level of the course BA Linguistics, but can be scaled up towards MA, research MA, PhD
iii. Workload 6 ECTS
iv. Institution Radboud University, Nijmegen, NL
v. Course instructor(s) Louis ten Bosch c.s.
vi. Brief course description <p>This ASR/FA course is based on research-based teaching (RBT), i.e. will be based on the integration of ongoing research into ongoing teaching. The course's topic, ASR and FA, is quite technical in nature and progresses rapidly over recent years. It is therefore not at all straightforward to have bachelor students involved in ongoing research projects or even have them apply the same research tools and infrastructures that the researcher uses. However, it is possible for them to acquire and actively apply problem-solving and analytical skills using methods based on actual research if their background matches sufficiently with the background of the course. In this way, students gain experience from actual real-life research as part of their training. Since ASR has matured into a broad research domain with over 10 specialized subdomains, short-term student involvement in projects is necessarily very specific, both in terms of the topic and in terms of method. In order to guarantee that students get a sufficiently wide perspective on ASR/FA research, this course provides a general introduction to ASR/FA principles.</p> <p>The course is built around the following structure</p> <ul style="list-style-type: none">(a) general framework and theoretical design(b) analytic skills, tools, use of infrastructure, and use of literature(c) thesis/research report <p>The course consists of 10 modules. Each of these modules addresses a specific topic within ASR/FA. The last module (number 10) consists in the writing of a thesis. The topic of this thesis is up to the student (e.g., can be chosen from a predefined list).</p>

<p>vii. Research related subject Speech decoding, signal analysis</p>
<p>viii. Data the students work with Open source audio data (mainly English)</p>
<p>ix. Topics</p>

A: Research design
A1: General research design
<p>[TEACHING MATERIAL] https://docs.google.com/document/d/1nSP-QeJJHWCfjDO_0IUmAfF61kw7D2kf/edit?usp=sharing&oid=111755079485368776002&rtpof=true&sd=true https://web.stanford.edu/~jurafsky</p>
A2: Adapting the general research design to the specific topic of interest
A2: Adapting the general research design to the specific topic of interest
A2.1: Formulation of a specific theoretical or practical issue in the general framework of signal-to-word ASR pipeline
A2.2: Formulation of detailed question
<p>A2.3: Selection of appropriate research techniques, selection and creation of corresponding data sources</p> <ul style="list-style-type: none"> • Experimental paradigms (e.g., feature selection, acoustic modelling, interpretation of non-semantic ASR output) • AM/LM-based architectures versus end-to-end • Developing and exploiting databases and corpora for improving ASR performance
A2.4: Identifying an ASR performance error analysis method
A2.5: Inferring theoretical and practical consequences of chosen approach
A3: Adapting the research design to the available research infrastructures
<p>A3.1 Selection of optimal research techniques, selection and creation of corresponding data sources (see also A2.3)</p> <ul style="list-style-type: none"> • literature search in e-journals • connection literature findings to own results

A4: Research reporting
A4.1 Presentation modes for research reporting: thesis (3EC)
<p>A4.2 Established procedures and conventions in research reporting, such as:</p> <ul style="list-style-type: none"> • the ordering of thematic units in the thesis, • organization and clarity of themes addressed, • balance text and graphical items in thesis • clarity terminology, • citing conventions
<ul style="list-style-type: none"> • B: Infrastructures & techniques
B1: For obtaining literature
[GENERAL-PURPOSE REPOSITORY] ResearchGate, Google Scholar, Academia.edu
B2: For obtaining, sharing and managing data
B2.1: Definition of research infrastructures are, and the main concepts around data interoperability , such as data , metadata and standards .
<p>B2.2: Platforms and repositories.</p> <ul style="list-style-type: none"> • General-purpose repositories and disciplinary repositories
<p>B2.3: Identifying, collecting, creating and/or using relevant data for research projects</p> <ul style="list-style-type: none"> • Searching and selecting relevant speech corpora • Citing data sets as appropriate. • Awareness about research data, certified repository and selection of an appropriate licence for sharing data
<ul style="list-style-type: none"> • B3: For analysing data
B3.1: Software for investigating audio signals
B3.2: Software platform for computations (python)
B3.3: Familiarity with existing ASR platforms

x. Learning outcomes

A: Research design

A1: Students will be able to make an overview of the general research design.

A2: Students will be able to create a suitable research design for the specific topic of interest.

A2.1: Students will be able to formulate general questions and hypothesis in terms of ASR/FA pipeline.

A2.2: Students will be able to formulate a specific question about an aspect of the ASR/FA pipeline.

A2.3: Students will be able to select optimal research techniques, and search/create corresponding data sources.

A2.4: Students will be able to select and implement the optimal speech decoding method.

A2.5: Students will be able to infer (speech-related, linguistic, theoretical) consequences from the specific ASR/FA outcomes.

A3: Students will be able to adapt a research design to the available research infrastructures.

A3.1 Students will be able to select of optimal research techniques, select and create corresponding data sources (see also A2.3)

- students will be able to link research outcomes to data.

- A4: Students will be able to report on their performed research in accordance with standards and conventions in the field.

A4.1 Students will be able to cast research outcomes in terms of a thesis.

A4.2 Students will be able to implement established procedures and conventions in research reporting, such as:

- the ordering of thematic units
- organization of the presentation,
- balance text/graphical items
- proper citation of references
- citing conventions.

B: Infrastructures & techniques

B1: Students will be able to identify and apply suitable infrastructures & techniques for obtaining literature.
B2: Students will be able to identify and apply suitable infrastructures & techniques for obtaining, sharing and managing data.
B2.1: Students will understand what ASR/FA pipelines are.
B2.2: Students will be able to identify suitable platforms ASR/FA. <ul style="list-style-type: none"> • Understand the difference between ASR and FA.
B2.3: Students will be able to identify, collect, create and/or use relevant data for their research projects <ul style="list-style-type: none"> • Cite audio data sets as appropriate.
B2.4: Students will be aware of the relevance of data management plan <ul style="list-style-type: none"> • Understand the data lifecycle • Understand how to generate data, analyse and handle it • Understand the legal and ethical issues around speech data generation and use.
B3: Students will be able to identify and apply suitable infrastructures & techniques for analysing data.
B3.1: Students will be able to select and use software for speech analyses.
B3.2: Students will be able to select and use software for ASr/FA.
B3.3: Students will be able to select and use software for interpreting ASR output.

xi. Overview of evaluation	
Rubric	Weighing
Participation incl. homework (initiative, forward-thinking, problem solving, critical thinking, organisation, time management)	30%
On (individual) written final research report in line with the local university standards	70%