



Language Technology: Research and Development

Introduction

Sara Stymne

Uppsala University
Department of Linguistics and Philology
sara.stymne@lingfil.uu.se



Teaching Team

- ▶ Course coordinator, examiner and lectures:
 - ▶ Sara Stymne
- ▶ Seminars
 - ▶ Ali Basirat
 - ▶ Daniel Dakota
 - ▶ Sara Stymne



Course Content

Theory

Philosophy of science

Research methods in LT

Scientific writing

Practice

Survey a research field

Plan and implement a project

Write and review scientific papers

- ▶ Lectures covering theory (large group)
- ▶ Seminars devoted to practice (small group)
- ▶ Individual projects on a common theme (small group)



Research Themes

- ▶ Cross-lingual NLP [[Sara](#)]
 - ▶ Cross-lingual systems involve more than one language
 - ▶ Useful especially for low-resource languages
- ▶ Word embeddings [[Ali](#)]
 - ▶ A meaningful representation of words
 - ▶ Beneficial to many NLP tasks and linguistic studies
- ▶ Sentiment analysis classification tasks [[Daniel](#)]
 - ▶ Used for automated opinion mining extraction
 - ▶ Difficulty across language and domain in sentiment tasks types



Course Structure

1. Background part:

- ▶ Philosophy of science and research methods [lectures]
- ▶ Survey of the state of the art in research theme [seminars]
- ▶ Planning an R&D project [lecture, seminar]

2. Project part:

- ▶ Implementing an R&D project [seminars]
- ▶ Writing a scientific paper [lecture, seminars]
- ▶ Reviewing scientific papers [lecture]



Reading List

- ▶ Textbooks:
 - ▶ Okasha, S. (2002) *Philosophy of Science: A Very Short Introduction*. Oxford University Press.
 - ▶ Zobel, J. (2004) *Writing for Computer Science*. Second Edition. Springer.
- ▶ Papers:
 - ▶ Available online from the course home page
<https://cl.lingfil.uu.se/kurs/rd20/>



Assignments and Examination

1. Take home exam on philosophy of science (15%) [written]
2. Research paper presentation and discussion (15%) [oral]
3. Project proposal (15%) [written, oral]
4. Term paper (40%) [written, oral]
5. Review of term papers (15%) [written]



Assignments and Examination

1. Take home exam on philosophy of science (15%) [written]
 2. Research paper presentation and discussion (15%) [oral]
 3. Project proposal (15%) [written, oral]
 4. Term paper (40%) [written, oral]
 5. Review of term papers (15%) [written]
-
- ▶ Pass (G) = all assignments passed
 - ▶ Distinction (VG) = at least 50% of 1, 3–5 with distinction



Deadlines

Choose your preferred topic	September 4, 13.00
Hand in take home exam	September 17
Project proposal	October 2
Present project proposal	October 7
First version of term paper	December 11
Peer review of (other) term papers	December 21
Final seminar	January 13
Final term paper	January 15



Deadlines

Choose your preferred topic	September 4, 13.00
Hand in take home exam	September 17
Project proposal	October 2
Present project proposal	October 7
First version of term paper	December 11
Peer review of (other) term papers	December 21
Final seminar	January 13
Final term paper	January 15

Backup deadlines available on course web page, but important to try to respect original deadlines! (This course is a prerequisite for the master thesis course.)



Seminars

- ▶ All seminars are obligatory!
- ▶ Group seminars:
 - ▶ Research papers
 - ▶ Project proposal (presentations with slides)
 - ▶ Progress reports (and ethics)
- ▶ Final seminar in full group
 - ▶ Full day "mini workshop"
 - ▶ First-year master students also invited
 - ▶ Social event (if possible)



Seminars

- ▶ All seminars are obligatory!
- ▶ Group seminars:
 - ▶ Research papers
 - ▶ Project proposal (presentations with slides)
 - ▶ Progress reports (and ethics)
- ▶ Final seminar in full group
 - ▶ Full day "mini workshop"
 - ▶ First-year master students also invited
 - ▶ Social event (if possible)
- ▶ If you miss a seminar, there will be a compensation task



Going for the Real Thing

- ▶ The goal is to do **real** research resulting in **real** publications
- ▶ Guidelines for submission and reviews:
 - ▶ Transactions of the Association for Computational Linguistics
- ▶ Term papers may be revised and submitted for publication
- ▶ Actual submission is **not** a course requirement



Going for the Real Thing

- ▶ The goal is to do **real** research resulting in **real** publications
- ▶ Guidelines for submission and reviews:
 - ▶ Transactions of the Association for Computational Linguistics
- ▶ Term papers may be revised and submitted for publication
- ▶ Actual submission is **not** a course requirement

- ▶ You are meant to function as a **real** research group
- ▶ Projects are individual, but you should support each other



Publications from Recent Years (1)

Arra'Di Nur Rizal and Sara Stymne. Evaluating Word Embeddings for Indonesian–English Code-Mixed Text Based on Synthetic Data. *The 4th Workshop on Computational Approaches to Code Switching*.

2 students from the 2019 course currently have papers under review

Artur Kulmizev, Miryam de Lhoneux, **Johannes Gontrum**, **Elena Fano** and Joakim Nivre. Deep Contextualized Word Embeddings in Transition-Based and Graph-Based Dependency Parsing – A Tale of Two Parsers Revisited. *The Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing*.

Ailsa Meechan-Maddon and Joakim Nivre. How to Parse Low-Resource Languages: Cross-Lingual Parsing, Target Language Annotation, or Both? *International Conference on Dependency Linguistics (Depling 2019)*.



Learning Outcomes 1

The student should at least be able to do the following, in relation to a scientifically organised language technology project:

- ▶ explain the basic principles of scientific work and research methodology in general and in relation to a current project
- ▶ make an overview of earlier research and the state of the art within the field that the project treats and identify its most urgent research issues,
- ▶ show an ability to identify and formulate research questions in a critical, independent, and creative way



Learning Outcomes 2

The student should at least be able to do the following, in relation to a scientifically organised language technology project:

- ▶ plan and carry out research tasks based on sound methodological principles and within given time limits,
- ▶ evaluate results and partial results with current validation methods,
- ▶ present the purpose of the project and its results in a professional manner, both for scientists and for the general public, orally and in writing, taking the target audience into consideration.



Student Feedback

- ▶ 2019 students were very happy with the course: (4.7/5)
- ▶ Some comments:
 - ▶ Seminars were very useful
 - ▶ Doing a full project from proposal to reviewing and final workshop was very useful
 - ▶ Shaping and running our own projects were fun
 - ▶ Good pace
 - ▶ Good to have three topics to choose from
 - ▶ The course was very useful in preparation for the master's thesis
 - ▶ Intense workload, but it was worth it
 - ▶ Philosophy of science and take-home exam felt a bit out-of-place



Changes in the course

- ▶ Two new teachers and topics
- ▶ Redistributed work load last year, this worked well, so will be kept
- ▶ Changes due to the current situation
 - ▶ Lectures mainly mixed mode (Zoom+campus)
 - ▶ Seminars either Zoom or Campus: decided by each group
 - ▶ Final seminars: TBD based on current recommendations



Campus / Zoom

- ▶ Please follow current regulations!
 - ▶ Do not come to Campus if you do not feel well!
 - ▶ Try to maintain social distancing
 - ▶ You can always attend online if needed!
- ▶ Campus activities may be cancelled on short notice
Check your email before going to Campus!
- ▶ Class rooms may become full (unlikely)
Students gets places on a first come, first served basis
- ▶ Please use your camera and real name during Zoom seminars



Coming up

- ▶ Now: introduction to the topics
- ▶ Choose your preferred topics
 - ▶ Rank your preference for the 3 topics, and indicate your preference for Campus/online seminars
 - ▶ By email to Sara: deadline Friday September 4, 13.00
- ▶ Lecture on science, research and NLP: Thursday
- ▶ Lecture: debates on philosophy of science and NLP: next week (Zoom only)
- ▶ First research paper seminars: Monday 14



Research Paper Seminars

- ▶ Obligatory attendance
- ▶ Each student is responsible for introducing one article each
 - ▶ briefly summarize the paper (2 min)
 - ▶ discuss the main points being made
 - ▶ bring up difficult to understand parts
 - ▶ initiate a discussion by proposing themes to discuss
- ▶ All students are supposed to have read all articles, to bring discussion points, and actively discuss the articles
- ▶ Bring the articles to the seminars (on paper or electronically)
- ▶ The list of articles and presenters will be available on the web page early next week



Questions?