



Language Technology Education and Research in Estonia

Mare Koit
mare.koit@ut.ee



PhD Education and Arrangements

3 universities will participate in the Nordic Graduate School of Language Technology

- Tallinn University of Technology
 - Tallinn Pedagogical University
 - University of Tartu
-



PhD Education and Arrangements (2)

Two models of higher education:

- old:
4 years (Bachelor) +2 years (Master) +4 years (PhD)
- from 2002/2003 new (*Bologna declaration*):

3+2+4

- 1 year = 40 credits (AP)
- 1 credit = 40 work hours (=1,5 ECTS)



PhD Education and Arrangements (3)

Tallinn University of Technology:

- no speciality of language technology
- some students of **information technology** are specialising in the field of speech technology, following individual study programmes

<http://eng.ttu.ee/studies/studyprogrammes/?#INFOTECH>



PhD Education and Arrangements (4)

Tallinn Pedagogical University

- no speciality of language technology
 - Computer Science
 - Linguistics

<http://www.tpu.ee/oppeinfo/index.php?act=oppekava>



PhD Education and Arrangements (5)

University of Tartu

- no speciality language technology on the PhD level
- the relevant research training is typically carried out under General Linguistics or Computer Science
- Currently, 6 PhD students are specialising in LT
(3 – general linguistics, 3 – computer science)
 - Individual study plan for every student
 - Obligatory subjects 20 AP
 - Optional subjects related to the field of specialisation 20 AP
 - PhD thesis 120 AP



Education for Master's Degree Studies

University of Tartu

- Computational linguistics at the Faculty of Philosophy (3+2)
=> master of computational linguistics
- Language Technology at the Faculty of Mathematics and Computer Science (3+2)
=> master of informatics
(not of language technology)



Education for Master's Degree Studies (2)

- Some subjects for Computer Science students that are specialising in LT (4th and 5th year):
 - Introduction to general linguistics 3 AP
 - Language technology 3 AP
 - Corpus linguistics 3 AP
 - Computational morphology 2 AP
 - Syntactic theories and models 4 AP
 - Statistical natural language processing 2 AP
 - ...
- Master's thesis 20 AP



Education for Master's Degree Studies (3)

- Some subjects for Computational Linguistics students (3rd, 4th and 5th year):
 - Introduction to computational linguistics 4 AP
 - Language technology 3 AP
 - Corpus linguistics 3 AP
 - Computational morphology 2 AP
 - Data bases 4 AP
 - Statistical natural language processing 2 AP
 - ...
- Master's thesis 20 AP



Research Organisations

- University of Tartu
 - informal group of computational linguistics that involves 2 chairs
<http://www.cl.ut.ee/>
 - Chair of general linguistics at the Faculty of Philosophy (Prof. Haldur Õim) – 2 teachers + 8 researches
 - Chair of language technology at the Faculty of Mathematics and Computer Science (Prof. Mare Koit) – 3 teachers + 1 researcher
- Tallinn University of Technology
 - Institute of Cybernetics
 - Laboratory of Phonetics and Speech Technology (Dr. Einar Meister) – 5 researchers
 - <http://www.phon.ioc.ee/>
- (Research) Institute of Estonian Language
 - 4-5 researchers in language technology (Dr. Ülle Viks, etc.)



Main research fields

- University of Tartu
 - text corpora of Estonian, corpus of spoken Estonian, dialogue corpus
 - computational morphology, syntax of Estonian
 - lexical-semantic data base of Estonian (WordNet)
 - dialogue modelling
 - Tallinn University of Technology
 - speech technology (TTS synthesiser, speech and speaker recognition of Estonian)
 - Institute of Estonian Language
 - computational morphology of Estonian
 - text corpus of Estonian
 - machine-readable lexicons
-