

<b>Course name: Tools and Resources of Language Technologies</b>	<b>Credits: 4</b>
Class type: Lecture+Laboratory hours per week: 2+2	
Type of the exam: presentation, colloquium	
Prerequisites (if exist): Basics of Language Technologies	
<b>Course description:</b>	
<p><b>Introduction.</b> History of computational linguistics. Rule-based CL. Statistical methods in CL. Hybridization.</p> <p><b>Corpus linguistics</b> Encoding corpora. Technical, methodological and practical questions of corpus building. Crawlers and reliable data. Boiler plate algorithms.</p> <p><b>Treating and Analysis of text corpora</b> Corpus building. Corpus operations. Search methods. Methods of corpus analysis. Brown Corpus. Lancaster-.Oslo/Bergen (LOB) Corpus. London-Lund corpus. Bank of English. British National Corpus (BNC). International Corpus of English (ICE). . Gutenberg. Hungarian corpora in practice. Hungarian National Corpus.</p> <p><b>Using corpora in lexicographic applications</b> Dictionary building, planning lexicographic resources for various applications. Dictionary building from text corpora: automatic methods.</p> <p><b>Analysis of conceptual networks</b> Semantic nets. Frames. Semantic relations in computational linguistics,. Case studies: from Quillian nets to CyC, WordNet and FrameNet.</p> <p><b>WordNet</b> Relations, nets. Realization, search. Synsets. Multilingual ontologies. Grammatical categories of WordNet, technical tools for treating WordNet. EuroWordNet. Hungarian WordNet..</p> <p><b>Automatic parsing of monolingual and bilingual sources</b> Dictionary analysis as information extraction. Methodological questions. Parallel corpora and bilingual dictionaries.</p> <p><b>Underspecified structures</b> Unknown proper names in text corpora. Geographical names. Named entity recognition.</p>	
<b>Required reading:</b>	
<p>Daniel Jurafsky &amp; James Martin: <i>Speech and Language Processing</i>. Prentice-Hall, 2000/2008  Gábor Prószték: <i>Computational Linguistics (in Hungarian)</i>. Számalk, 1989  Gábor Prószték &amp; Balázs Kis: <i>Számítógéppel emberi nyelven</i>. SZAK, 1999</p>	
<b>Recommended reading:</b>	

Gábor Prószéky: *Language Technology (and it)s Applications*. Aranykönyv, 2005

**Lecturer** (*name, position, degree*): Gábor Prószéky, full professor; Doctor of HAS

**Additional lecturers**, if exist (*name, position, degree*): Győző Zijian Yang, PhD student,  
Borbála Siklósi PhD student.