# DOCTORAL SCHOOL OF LINGUISTICS FACULTY OF HUMANITIES AND SOCIAL SCIENCES PÁZMÁNY PÉTER CATHOLIC UNIVERSITY

## **PROGRAMME PLAN**



Budapest, 2020

### Contents

PLAN OF INSTRUCTION	
A. COURSE WORK	3
B. RESEARCH AND PUBLICATION	4
FACULTY AND STAFF	
COURSE DESCRIPTIONS – EXAMPLES	
GENERAL TRACK (IN HUNGARIAN, SOME COURSES ALSO IN ENGLISH)	8
APPLIED TRACK (IN ENGLISH)	9
EXAMPLES OF COURSES IN THE WORKSHOPS	12
COMPLEX EXAMINATION	
Topics in the first part of the complex examination	14
The second part of the complex examination	15
FURTHER INFORMATION	

#### PLAN OF INSTRUCTION

#### A. COURSE WORK

In each of the first two semesters at least four examination or seminar grades are to be collected, in addition to the consultation seminar grades. All the prescribed course credits are to be collected in the first four semesters.

Compulsory courses are printed bold below.

Elective courses can be taken in any of the first four semesters.

#### **SEMESTER 1**

General track (in Hungarian); grade achieved through	Applied track (in English); grade achieved through	Credits
1. Syntax 1 (generative theory of syntax);	<b>1. Syntax; examination</b> (Semester 1 or 2)	8
examination		
2. Morphology; examination	2. Introduction to applied linguistics;	8
	examination	
3. Phonology; examination	<b>3. Phonology; examination</b> (Semester 1 or 2)	8
4. Logic; examination	4. One of compulsory electives A; seminar grade	8
5. Consultation; seminar grade	5. Consultation; seminar grade	4
6. Elective course	6. Elective course	8

#### **SEMESTER 2**

1. Syntax 2 (Hungarian generative syntax); examination	1. One of compulsory electives A; seminar grade	8
2. Semantics; examination	2. On of compulsory electives B: Semantics OR Morphology; examination (Semester 1 or 2)	8
3. Research design 1; seminar grade	3. Research design 1; seminar grade	8
4. Consultation; seminar grade	4. Consultation; seminar grade	4
5. Elective course	5. Elective course	8
6. Elective course	6. Elective course	8

Compulsory electives A (choose any TWO; 8 credits each):	
Introduction to language pedagogy	
Psycholinguistics 1	
Sociolinguistics 1	
Pragmatics 1	
Introduction to language technology	

#### **SEMESTER 3**

1. Research design 2; seminar grade	1. Research design 2; seminar grade	8
2. Consultation; seminar grade	2. Consultation; seminar grade	4
3. Elective course	3. Elective course	8
4. Elective course	4. Elective course	8
5. Elective course		
6. Elective course		

#### **SEMESTER 4**

1. Consultation; seminar grade	1. Consultation, seminar grade	4
2. Elective course	2. Elective course	8
3. Elective course		
4. Elective course		
Complex examination	Complex examination	16

#### **Elective courses (Semesters 1-4)**

1. Academic English 1; seminar grade	8
2. Academic English 2; seminar grade	8
3. Phonology in first and second language	8
acquisition	
4. First language acquisition	8
5. Second language acquisition	8
6. Psycholinguistics 2	8
7. Sociolinguistics 2; seminar grade	8
8. Phonetics; examination	8
9. Statistics; seminar grade	8
10. Pragmatics 2	8
11. Research on the foreign language learner	8
12. Language policy, language teaching policy;	8
seminar grade	
13. Language testing and assessment; seminar	8
grade	
14. Corpus linguistics; analysing learner text	8
15-20. Workshops 1-6	8
21-24. Electives in Hungarian 1-4	8
	<ol> <li>Academic English 2; seminar grade</li> <li>Phonology in first and second language acquisition</li> <li>First language acquisition</li> <li>Second language acquisition</li> <li>Psycholinguistics 2</li> <li>Sociolinguistics 2; seminar grade</li> <li>Phonetics; examination</li> <li>Statistics; seminar grade</li> <li>Pragmatics 2</li> <li>Research on the foreign language learner</li> <li>Language policy, language teaching policy; seminar grade</li> <li>Language testing and assessment; seminar grade</li> <li>Corpus linguistics; analysing learner text</li> <li>20. Workshops 1-6</li> </ol>

#### All credits to collect through course work in Semesters 1-4:

Compulsory courses, including the compulsory electives, the consultations and the complex<br/>examination:96 credits<br/>48 credits<br/>48 creditsElective coursesAltogether in Semesters 1-4:144 credits

<b>SEMESTERS 5-8</b> <b>Consultation</b> in each semester; seminar grade	4 credits, altogether:	16 credits
All credits to collect through course work:		160 credits

#### **B. RESEARCH AND PUBLICATION**

The Doctoral School is committed to maintaining the highest ethical standards of academic research and publication as detailed in the recommendations of the British Association for Applied Linguistics and the Committee on Publication Ethics (COPE). Credits are only assigned to academic publications and presentations that conform to these standards.

#### I. Academic publications

### 1. Academic publications based on the doctoral student's own research (min. 26 credits):

Peer-reviewed book in the student's non-first language: 18 credits

Peer-reviewed book in the student's first language: 14 credits

Paper in the student's non-first language, in an international peer-reviewed periodical, 1: 12 credits

Paper in the student's non-first language, in an international peer-reviewed periodical, 2: 12 credits

Paper in the student's non-first language, in an international peer-reviewed periodical, 3: 12 credits

Paper in the student's non-first language, in an international peer-reviewed periodical, 4: 12 credits

Paper in the student's non-first language, in a peer-reviewed periodical or in a peer-reviewed book volume, 1: 9 credits

Paper in the student's non-first language, in a peer-reviewed periodical or in a peer-reviewed book volume, 2: 9 credits

Paper in the student's non-first language, in a peer-reviewed periodical or in a peer-reviewed book volume, 3: 9 credits

Paper in the student's first language, in a peer-reviewed periodical or in a peer-reviewed book volume, 1: 8 credits

Paper in the student's first language, in a peer-reviewed periodical or in a peer-reviewed book volume, 2: 8 credits

Paper in the student's first language, in a peer-reviewed periodical or in a peer-reviewed book volume, 3: 8 credits

Paper in the student's non-first language, in an international academic conference

proceedings, based on competitive anonymous abstract or paper selection, 1: 6 credits Paper in the student's non-first language, in an international academic conference

proceedings, based on competitive anonymous abstract or paper selection, 2: 6 credits Paper in an academic conference proceedings, based on competitive anonymous abstract or paper selection, 1: 4 credits

Paper in an academic conference proceedings, based on competitive anonymous abstract or paper selection, 2: 4 credits

#### 2. Academic reviews:

Review of a book, published in an academic journal and written in the student's non-first language: 6 credits

Review of a book, published in an academic journal and review written in the student's first language: 4 credits

#### 3. Editorial activity (max. 12 credits):

Editing a peer-reviewed publication in the student's non-first language: 7 credits Editing a peer-reviewed publication in the student's first language: 5 credits Editing a publication in the student's non-first language: 4 credits Editing a publication in the student's first language: 3 credits

#### 4. Academic publications based on the research of others (max. 8 credits)

Review article in the student's non-first language, in a peer-reviewed periodical: 5 credits Review article in the student's first language, in a peer-reviewed periodical: 4 credits Peer-reviewed encyclopedia entry, in the student's non-first language: 3 credits Peer-reviewed encyclopedia entry, in the student's first language: 2 credits

#### 5. Publications popularising linguistics (max. 5 credits):

Publication popularising linguistics, min. 10 pages: 3 credits Publication popularising linguistics, 3 to 10 pages, 1: 2 credits Publication popularising linguistics, 3 to 10 pages, 2: 2 credits Publication popularising linguistics, shorter than 3 pages, 1: 1 credit Publication popularising linguistics, shorter than 3 pages, 2: 1 credit The above number of credits are granted for a minimum of 10 pages per paper or review (2,000 characters per page). Count two-thirds of the points prescribed for the category for texts shorter than that, unless the Linguistics Doctoral Board decides otherwise. Round up to a whole number.

For the purposes of the current Programme Plan, *international journal* and *international conference proceedings* are defined as an academic journal or academic conference proceedings volume whose authors (and, in the case of journals, whose editorial board) are sufficiently varied in terms of nationality and their academic affiliation.

To earn credits for a publication, the doctoral student is to inform the Doctoral Office in writing, including all the necessary bibliographic data, possible online access, and the supervisor's confirmation of acceptance for credit. In certain cases the credits can be granted prior to actual appearance of the published work, provided that editorial acceptance is duly confirmed.

#### **II.** Academic presentations

# 1. Academic conference presentation based on the student's own research (min. 14 credits)

Conference talk in the student's non-first language, 1: 6 credits Conference talk in the student's non-first language, 2: 6 credits Conference talk in the student's non-first language, 3: 6 credits Poster presentation in the student's non-first language: 3 credits Conference talk in the student's first language, 1: 4 credits Conference talk in the student's first language, 2: 4 credits Conference talk in the student's first language, 3: 4 credits Poster presentation in the student's first language, 3: 4 credits

#### 2. Talks popularising linguistics

Talk popularising linguistics 1: 2 credits

Talk popularising linguistics, 2: 2 credits

Talk popularising linguistics, 3: 2 credits

To earn credits for a conference presentation and popular talks, the doctoral student is to inform the Doctoral Office in writing, including all the necessary data (the event's name, venue, date, the title and type of the presentation: talk or poster) and the supervisor's signature.

#### 3. Talks at the doctoral school's workshop conference

#### SEMESTERS 1-4 (min. 8 credits to collect)

Talk at the workshop conference of the doctoral school 1: 4 credits Talk at the workshop conference of the doctoral school 2: 4 credits Talk at the workshop conference of the doctoral school 3: 4 credits

#### SEMESTERS 5-8 (min. 8 credits to collect)

Talk at the workshop conference of the doctoral school 4: 4 credits Talk at the workshop conference of the doctoral school 5: 4 credits Talk at the workshop conference of the doctoral school 6: 4 credits Talk at the workshop conference of the doctoral school 7: 4 credits

#### Multi-author publications and presentations:

In the case of multi-author publications and presentations, calculate the credit points for the category divided by the number of co-authors, unless the Linguistics Doctoral Board decides otherwise. Round up for a whole number.

#### III. Other academic and popular linguistic activities (max. 12 credits)

Event (conference) organisation, academic or popular, 1: 3 credits Event (conference) organisation, academic or popular, 2: 3 credits Other academic activity (e.g. research assistance) or activity popularising linguistics, 1: 3 credits (or fewer, depending on the decision by the Linguistics Doctoral Board) Other academic activity (e.g. research assistance) or activity popularising linguistics, 2: 3 credits (or fewer, depending on the decision by the Linguistics Doctoral Board)

The above activities are to be confirmed towards the Doctoral Office by the supervisor.

#### IV. Teaching/ teaching assistance (max. 16 credits)

Teaching a university course, 1: 8 credits Teaching a university course, 2: 8 credits Teaching assistance (organisation, support), 1: 4 credits Teaching assistance (organisation, support), 2: 3 credits Producing teaching material, 1: 4 credits Producing teaching material, 2: 3 credits

This category refers mainly to teaching or teaching assistance at Pazmany Peter Catholic University under the supervision of the supervisor or a member of the Linguistics Doctoral Board, unless the Linguistics Doctoral Board decides otherwise.

#### All credits to collect through research and publication:

min. 80 credits

#### TOTAL NUMBER OF CREDITS TO COLLECT:

**240 CREDITS** 

#### FACULTY AND STAFF

Head of the Doctoral School of Linguistics: Balázs Surányi, suranyi.balazs (at) btk.ppke.hu Head of the applied linguistics track: Andrea Reményi, remenyi.andrea (at) btk.ppke.hu Head of the Doctoral Office: Kornélia Kiss, kiss.kornelia (at) btk.ppke.hu Coordinator: Katalin Subicz, subicz.katalin (at) btk.ppke.hu

#### **COURSE DESCRIPTIONS – EXAMPLES**

#### GENERAL TRACK (IN HUNGARIAN, SOME COURSES ALSO IN ENGLISH)

#### Syntax 1

#### Balázs Surányi

This course is an introduction to the core concepts and principles of the analysis of sentence structure. It provides an overview of word classes, constituent structure, grammatical functions, the internal structure of phrases, the structure of main and subordinate clauses, unpronounced syntactic elements, thematic relations and case, and the role of displacement in shaping word order. The linguistic material we analyze comes mostly from English, but also from other languages. While aiming to cover the basic notions of both descriptive and formal syntax, the intellectual emphasis of the course is on understanding the reasoning that guides syntactic analysis.

#### Phonology

#### András Cser, Katalin Balogné Bérces

The course is an introduction to current phonological theory. The subjects discussed include the areas of phonology, alternations, phonotactics, prosody; theories of phonological representations, features, elements, suprasegmental organization, autosegmental phonology; theories of phonological derivations, questions of rule ordering; declarative phonologies, and optimality theory.

#### Morphology

#### András Cser

The course aims at acquainting students with the concept of morphological structure, including the properties of inflection, derivation, and compounding, as well as the functions of morphology and the types of morphemes, etc. The course devotes special attention to morphological theory and the interfaces between morphology and phonology on the one hand, and morphology and syntax on the other. Students are also introduced to the chief problems regarding productivity. Morphological systems of different types will also be compared, and students will also learn about the diachronic aspects of the field, paying special attention to lexicalization and grammaticalization. The central topic of the course provides an opportunity to study certain morphological phenomena from a range of theoretical perspectives, comparing and analyzing the theories under scrutiny.

#### Semantics

Tamás Halm

This course offers an informal, intuitive introduction to the logical core of natural semantics (connectives, predicate—argument structure, quantification), followed by an introduction to an array of empirically significant phenomena (genericity, plurals, tense and aspect, modality), and the problem of the division of labour at the semantics—pragmatics interface (entailment versus presupposition versus implicature).

#### **Sociolinguistics 1**

#### Attila Hegedűs, Andrea Reményi

This seminar course will provide an introduction to the basic concepts and current theoretical frameworks of sociolinguistics, the study of language in its social context. It will overview the basic concepts and some current problems of language variation (language/dialect/speech community/community of practice), interactional aspects (spoken interaction, addressing, conversation analysis), cultural aspects (ethnography of communication, first language acquisition and literacy), political aspects (language planning, language policy, language teaching policy, critical approaches) and bilingualism (types, code choice). We will pay some attention to both the social (individual, geographical, global, virtual) and linguistic dimensions (phonology, grammar, conversational and discourse structures). The course will concentrate on studies on English, with some glimpses into Hungarian and the possible other first language(s) of the participants. The understanding of sociolinguistic research methods will be supported by the discussion of problems of data collection and analysis.

#### APPLIED TRACK (IN ENGLISH)

#### **Introduction to applied linguistics**

Andrea Reményi

With the help of this survey course, based on interactive lectures, doctoral students will be introduced to some of the major areas of applied linguistics, defined here not as 'hyphenated linguistics', but as a problem-based approach to language-related issues in life: education, society, etc. Discussions will introduce topics in psycholinguistics, first and second language acquisition, sociolinguistics, bilingualism, language (teaching) policy, pragmatics, language pedagogy, language testing, corpus linguistics, language technology, etc. In each area, theoretical basics will be laid, and theory-practice connections will be emphasised with the help of examples. Some attention will be paid to research design and methodology.

#### Introduction to language pedagogy

#### Andrea Fischer

This course aims to enable the participants to develop a better understanding of the theoretical underpinnings of current principles and practices of language pedagogy. Based on current research different topic areas will be discussed and analysed, including a variety of teaching contexts and learning objectives. As a result, this course also aims to raise an awareness of why language teachers are special agents of change, who should be constantly challenging paradigms and be able to change perspectives in a reflective way.

#### **Psycholinguistics 1**

#### Lilla Pintér

The course provides an introduction into the field of psycholinguistics by summarizing its key findings regarding language perception, comprehension, and production. Special attention will be given to the various methods of psycholinguistic research and their role in testing competing theoretical accounts.

#### **Pragmatics 1**

#### Dániel Zoltán Kádár

The aim of this course is to overview the foundations of pragmatics, with special focus on sociopragmatics – that is, the study of language use in society. The first part of the course lays down the foundations of the field, by overviewing the development and key theories of the field, as well as the most important research methodologies and data types. Following this step, the second part provides insight into a range of sociopragmatic phenomena, spanning linguistic politeness, through humour, to sarcasm.

#### Introduction to language technology

#### Eszter Simon

This course covers the basic concepts and methods of the field of human language technology. The topics we discuss include the following: Introduction, definition, historical background. Character encoding of text files. Detecting, converting, Unicode. Using Unix shells; the main commands. Regular expressions; regular languages and automata. Finite state automata and transducers; the Chomsky-hierarchy of languages; two-level morphology. Corpus building; what is a corpus; corpus typology; corpus design; corpus size. Corpus annotation, annotation tools; evaluation. Annotation levels; tokenization; sentence segmentation; morphological analysis and disambiguation; sequential tagging tasks; syntactic analysis. Machine learning; supervised and unsupervised learning; feature engineering. Specific tasks of HLT: machine translation, computational lexicography, information extraction, ontologies etc.

#### Phonology in first and second language acquisition

#### Katalin Balogné Bérces

A brief overview of the emergence of phonological structures in first language acquisition; phonological markedness and language typology; representations of universals and acquisition in current phonological theory; the Critical Age Hypothesis (CAH) and its role in first and second language acquisition; the "foreign accent"; factors influencing pronunciation in second versus foreign language acquisition; issues in pronunciation teaching: intelligibility, the "target accent", standard vs non-standard varieties, phonetic vs phonological vs lexical components in pronunciation and their teachability in light of the CAH.

#### **Sociolinguistics 2**

#### Andrea Reményi

Based on the overview of the field in the *Sociolinguistics 1* course, here we will study recent trends in sociolinguistics, depending on the students' dissertation topic and interest, for example, cognitive sociolinguistics, language ideologies, interactional sociolinguistics, variation in English as a lingua franca, critical sociolinguistics, language and gender, etc.

#### Academic English 1

#### Andrea Reményi

This is a practical language development course for B2 and above level learners of English, with the following aims: 1. to develop doctoral students' presentation skills, 2. to expand their academic vocabulary, 3. to help them develop their academic writing, including hands-on help with a paper of theirs. Emphasis in the latter will be on text structuring, length and style modification (in terms of both lexis and grammar), plus citation and referencing.

#### Pragmatics 2

Dániel Zoltán Kádár

In this advanced course students gains both practical and theoretical insight into doing pragmatics on an advanced level. Ultimately, pragmatics is the theory of how language is used to build up and maintain interpersonal relationships, and studying this relational aspect in practice – beyond single utterances – raises a range of noteworthy issues, and also it requires the analyst to develop certain skills, such as self-reflexive thinking. During the course, we overview this deep layer of pragmatic research by combining theoretical training with guided data analysis.

#### Research on the foreign language learner

#### Ágnes Kohlmann

This course equips participants with a basic understanding of the processes of foreign/second language learning. Second language acquisition is its interdisciplinary character and research on this field is concerned with the general question: How are second languages learnt? A range of theoretical models of second language acquisition and multidisciplinary perspectives are examined and evaluated in relation to recent empirical findings, enabling students to develop the ability to think critically about findings in foreign language research.

#### Second language acquisition

#### Ágnes Kohlmann

This course seeks to give an overview about the research on second language acquisition (SLA). In the past 50 years, the field of SLA has developed into an independent and autonomous discipline, with its own research agenda investigating the human capacity to learn languages other than the first language. In this course models of bilingual representation and processing will be reviewed and evaluated in relation to empirical findings, in terms of shared or separated representation of languages, furthermore, issues in comprehension and production in two or more languages.

#### **Psycholinguistics 2**

#### Lilla Pintér

In the advanced course of psycholinguistics, typical and *atypical* developmental paths will be compared to each other. When investigating language disorders, such as Aphasia or Specific Language Impairment, students will also get a chance to learn more sophisticated experimental techniques.

#### First Language Acquisition

#### Lilla Pintér

The course provides a basic overview of the theories of first language learning, as well as of the essential stages of its developmental trajectory. The acquisition of the lexicon, the emergence of syntax, and the problems of disambiguation will be discussed in detail.

#### Phonetics

#### Katalin Mády

The course deals with basic questions of speech research. Main topics are: (1) functions of the articulatory organs and their physiological background (production), (2) acoustic cues of linguistically relevant articulatory differences and their analysis based on speech samples (acoustics), (3) perception of information encoded in the acoustic signal and the relationship between acoustics and human perception (perception). Students get introduced to the freely available Praat software for acoustic analysis. The results of experiments carried out in groups will be interpreted using this software.

#### Methods in phonetic research

#### Katalin Mády

The course provides an introduction to the methodology of experimental research by means of speech data collection and analysis. First, basic concepts of data collection will be discussed and practiced. Then participants will get acquainted with the speech lab, its technical and software tools. In the meantime, they will learn how to use the speech analysis software Praat, along with a brief introduction to acoustic analysis and data presentation. Each participant will carry out an experiment including design, data collection, analysis and presentation throughout the course. This can be done individually or in small groups.

#### Statistics

#### Katalin Mády

Empirical methods have gained considerable importance in today's linguistic research. This course provides an introduction to the methodology of statistical analysis. The course starts with descriptive statistics and distribution types. Based on probability theory, statistical procedures will be discussed for analysing frequency data, scores on a scale and measurable parametric data. For practical work, the freeware software R is used, which is based on a script language but is easy to acquire without programming skills. Statistical methods and the visual presentation of results will be practiced based on empirical data.

#### EXAMPLES OF COURSES IN THE WORKSHOPS

#### **Theoretical linguistics workshop**

#### Information structure and sentence structure

#### Balázs Surányi

This course explores how the form of sentences reflects the properties of the communicative context in which they are uttered—a phenomenon often called 'information packaging' or 'information structure'—from the perspective of syntax. We concentrate on the syntactic marking of topic status (what the sentence is "about"), focus status (contrast with alternatives), and given status (prior presence in the discourse). We look at both English and other languages, highlighting aspects of cross-linguistic differences. We also consider what challenges the syntactic marking of information structure poses both in L1 and in L2 acquisition.

#### Information structure and sentence prosody

#### Balázs Surányi

This course explores how the form of sentences reflects the properties of the communicative context in which they are uttered—a phenomenon often called 'information packaging' or 'information structure'—from the perspective of sentence prosody, including stress and intonation. We concentrate on the prosodic marking of topic status (what the sentence is "about"), focus status (contrast with alternatives), and given status (prior presence in the discourse). We look at both English and other languages, highlighting aspects of cross-linguistic differences. We also consider what challenges the prosodic marking of information structure poses both in L1 and in L2 acquisition.

#### The development of European linguistic thought

#### András Cser

This course provides a survey of the development of linguistic thinking from Antiquity to the twentieth century. The topics covered are: language, thought and reality in ancient Greece;

grammar as an emerging discipline in Rome; reassessing the tradition in the Middle Ages; the study of languages following the Renaissance period; the variety of languages, their description and competition; the beginnings of vernacular grammar; the issue of artificial languages; universal grammar and the relation between language and the mind; the comparative and typological study of languages; the beginnings of Indo-European and Finno-Ugric linguistics; the growth of historical linguistics; the emergence of the Neogrammarians; Saussure's place in the history of linguistics; early structuralism; mature structuralism in Europe and in America.

#### **English linguistics workshop**

#### **Topics in English syntax**

#### Balázs Surányi

Building on Syntax 1, this course provides an overview of key topics in the syntax of presentday English. The topics we consider include the functional structure of clauses, the functional structure of noun phrases, co-reference possibilities of nominal elements, the main classes of displacement operations, logical scope, and the marking of focus. In addition to giving students a detailed picture of English syntax, the course also aims to provide insight into the basic principles of the theoretical framework of generative grammar.

#### Diachronic phonology and morphology of the English language

#### András Cser

The goal of this course is to provide a theoretically oriented survey of the major sound changes and morphological changes in the history of the English language. The typology and modelling of phonological changes is discussed and exemplified through a study of Grimm's Law, Verner's Law, the Great Vowel Shift and various other processes. Theoretical aspects of morphological changes such as reanalysis, analogical change, class reassignment etc. are also given thorough treatment and are exemplified from the history of English.

#### Language technology workshop

#### **Basics of human language technologies**

#### Gábor Prószéky

This course covers the basic concepts and methods of human language technologies. The topics we discuss include the following: Characters, character encoding standards. Research problems in computational representation of natural languages. Applications of formal languages to natural languages. Word analysis and generation. Development methods and application of morphological systems. Structure and representation of sentences. Sentence parsing methods. Knowledge representation, computational semantics, ontologies. Representation of dictionaries, intelligent dictionaries. Computational aids for human translation. Machine translation methods.

#### Tools and resources of human language technologies

#### Gábor Prószéky

This course provides an overview of the main tools and resources that are commonly employed in human language technologies. The course builds on the material covered in the course titled 'Basics of human language technologies'. The topics we discuss include the following: Regular expressions. Computational lexicography and terminology. Ontologies, WordNet. Methods in corpus linguistics. Tokenization, named entity recognition. Word-sense disambiguation. Translation memories, text alignment. Machine translation techniques. Machine learning methods in NLP. Deep learning, neural models in NLP.

#### **COMPLEX EXAMINATION**

The complex examination is to be taken at the end of Semester 4. One re-take is allowed, in the same semester.

To register for the complex examination, the candidate is to submit: 1. his/her *academic CV*, 2. the *Research plan* about the second, research and dissertation writing, phase of his/her doctoral studies via email to the Doctoral Office. The deadline is the earlier one of these dates: (a) either the 15th day preceding the examination,

(b) or the last day of the study period preceding the examination period.

The CV describes the candidate's documented academic achievements: the academic publications (published or accepted for publication), the academic presentations and other academic achievements. The Research plan, in 15-20 pages, outlines the academic background, the initial hypotheses (if any), the research question(s), the objectives and their significance, of the candidate's research for dissertation. The plan details data to be analysed, the methods and the time schedule. Expected results and possible applications are to be presented, if relevant.

The complex examination consists of a theoretical part followed by a dissertation part, both orally.

#### Topics in the first part of the complex examination<sup>1</sup>

In the theoretical part the candidate is examined in three topics. Two of them are chosen by the Linguistics Doctoral Board, based on the recommendation of the supervisor – at least one of them are Topics A in the list below. A third topic (A or B) is chosen by the candidate.

#### **Topics A:**

- 1. Syntax
- 2. Morphology
- 3. Phonology
- 4. Semantics
- 5. Pragmatics
- 6. Psycholinguistics
- 7. Sociolinguistics
- 8. Historical linguistics

#### **Topics B:**

- 1. Hungarian syntax or the syntax of the candidate's non-first language
- 2. Hungarian phonology or the phonology of the candidate's non-first language
- 3. Hungarian morphology or the morphology of the candidate's non-first language
- 4. Phonetics
- 5. Text linguistics
- 6. Rhetoric
- 7. Corpus linguistics
- 8. Computational linguistics

<sup>&</sup>lt;sup>1</sup> Based on the "Principles and rules of the complex examination" ("A komplex vizsga általános elvei és szabályai") by the Hungarian Doctoral Board, accepted on February 12, 2016, modified on April 15, 2016.

#### 9. Lexicography

- 10. Research on bilingualism
- 11. First language acquisition
- 12. Second language acquisition
- 13. Discourse analysis
- 14. Prosody
- 15. Dialectology
- 16. Theory of translation
- 17. The history of a chosen language
- 18. The philology of a chosen language

#### The second part of the complex examination

In the second part of the examination the candidate reports on his/her familiarity in the research field and the steps already taken. On the basis of the Research plan, s/he presents the plans and the schedule to follow in the second phase of his/her doctoral studies. The committee and the candidate discuss that plan, and the committee may suggest modifications and corrections.

The committee evaluates the text, the presentation and the discussion of the Research plan, and either accepts it (with or without suggested modifications) or rejects it if it is found an unsuitable starting point for the research and dissertation work in the following two years. The committee prepares a short written evaluation about the second part of the examination: if the Research plan is accepted, the committee may suggest modifications; if the Research plan is rejected, the committee explains the decision.

#### FURTHER INFORMATION

General information for international students at Pázmány: <u>http://btk.ppke.hu/en/international-students</u>

How to use the inter-university wifi-pass 'EduRoam': <u>https://info.ppke.hu/index\_en.html</u>

How to use the studies administration site 'Neptun': http://btk.ppke.hu/uploads/articles/1558340/file/NEPTUN\_step\_by\_step\_2018\_end.pdf

How to use the faculty library: <u>https://btk.ppke.hu/konyvtar/library</u> Leaflet: <u>https://btk.ppke.hu/uploads/articles/1016581/file/2019\_09\_broschure.pdf</u>

On the Doctoral School of Linguistics (DSL): http://btk.ppke.hu/en/academics/degree-programmes#lingphd

On the Doctoral Office, including contact data (in Hungarian): http://btk.ppke.hu/oktatas/doktori-kepzesek-ph-d-es-habilitacio/doktori-es-habilitacios-iroda

Application for admission into the DSL:

http://btk.ppke.hu/uploads/articles/3782/file/Application%20for%20Admission%20to%20the %20PhD\_2019.docx

On the Applied Linguistics track within the DSL: <u>http://btk.ppke.hu/uploads/articles/3782/file/English%20language%20PhD%20program%20ip n%20Applied%20Linguistics\_revised\_formatted-RA(1).pdf</u>

Committee on Publication Ethics (COPE): https://publicationethics.org)

Recommendations of the British Association for Applied Linguistics: https://www.baal.org.uk/wp-content/uploads/2017/08/goodpractice\_stud.pdf (an abridged version)

https://baalweb.files.wordpress.com/2017/08/goodpractice\_full.pdf (the full version)