

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 15.06.2025 15:48:07
Уникальный программный ключ:
ca953a0120d891083f939673078ef1a989dae18a

**Federal State Autonomous Educational Institution of Higher Education
PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE
LUMUMBA
RUDN University**

Law Institute

educational division (faculty/institute/academy) as higher education programme developer

COURSE SYLLABUS

Quantitative Linguistics and IT (in English)

course title

Recommended by the Didactic Council for the Education Field of:

45.04.02 Linguistics

field of studies / speciality code and title

**The course instruction is implemented within the professional education programme of
higher education:**

MA in Translation and Interpreting for Public Services and Institutions

higher education programme profile/specialisation title

2025

1. COURSE GOALS

Corpus linguistics is the study of language as expressed in samples (corpora) or "real world" text. This method represents a systematic approach to obtaining a set of abstract rules by which a natural language is governed or else relates to another language. Originally done by hand, corpora are now largely derived by an automated process.

Modern linguistic research increasingly involves quantitative analysis: we count or measure various aspects of linguistic behaviour, and want to use this data to answer questions about the way language works.

This course will guide students through the description, visualization, and basic statistical analysis of both discrete and continuous linguistic data using open-source statistical software environment.

2. REQUIREMENTS FOR LEARNING OUTCOMES

On completion of the course the student:

- knows the conceptual underpinnings of some common statistical tests, and apply them appropriately
- can critically evaluate quantitative analyses in a range of linguistic subfields
- can make appropriate methodological choices when planning future research projects
- masters skills in the management of data, analytical work, and analysis results
- masters skills to summarize, visualize, and otherwise explore data using a variety of methods

Table 2.1. List of competences that students acquire through the course study

Competence code	Competence descriptor	Competence formation indicators (within this course)
GC-2	Can manage a project at all stages of its life cycle.	Indicator 1 Knows relevant approaches and techniques for project management and its implementation Indicator 2 Masters comprehensive tools for the project planning, implementation, quality control and its results application
GC -3	Can create and lead a team, designing a team strategy aimed at achieving the goals	Indicator 1 Knows relevant approaches and techniques for project management and its implementation Indicator 2 Masters comprehensive tools to organize the team work to achieve results
GC-4	Can apply modern communication techniques using the official language of the Russian Federation and foreign language(s) in academic and professional interaction	Indicator 1 Knows the specifics, approaches, tools, and techniques to communicate in Russian and foreign languages within various speech contexts of academic and professional interaction. Indicator 2 Masters comprehensive linguistic and extralinguistic tools to communicate in Russian and foreign languages within various speech contexts of academic and professional interaction.

Competence code	Competence descriptor	Competence formation indicators (within this course)
GC -7	Can find the necessary sources of information and data, comprehend, analyze, memorize information while using digital devices and algorithms; use various types of information from different sources in an effective way in problem-solving; assess the information, its authenticity, infer and deduct based on the input data and information	Indicator 1 Knows the relevant approaches, tools, and techniques to search for information from various sources, its processing and further use Indicator 2 Masters comprehensive tools to evaluate information, integrate, reshape and tailor it in line with context-specific tasks
GPC-2	Knows the specifics of foreign scientific worldview and masters the academic discourse in Russian and foreign languages being learnt	Indicator 1 Knows the specifics of the scientific picture of the world and the components of its discursive modelling in Russian and the foreign languages being learnt Indicator 2 Masters comprehensive tools to produce academic discourse in Russian and foreign languages

3.COURSE IN HIGHER EDUCATION PROGRAMME STRUCTURE

The course refers to the variable component of (B1) block of the higher educational programme curriculum. The course is taught within the general Scientific Disciplines university module.

The course is taught in English, the students are expected to master English at level B2- C1 to follow explanations, take notes, and ask questions, as well as provide the answers for overall course study related activities in English.

Students are expected to know Translation Studies key concepts and theories.

The course provides for theoretic grounds to further training in such fields as legal documents translation, scientific translation in jurisprudence, court interpreting, translation and interpreting for public services, legal translation and human rights.

4. COURSE WORKLOAD AND ACADEMIC ACTIVITIES

The total workload of the course is 2 credits (72 academic hours).

*Table 4.1. Types of academic activities during the periods of higher education programme mastering (**full-time training**)*

Type of academic activities	Total academic hours	Semesters/training modules			
		1	2	3	4
Contact academic hours	18		18		

Type of academic activities	Total academic hours	Semesters/training modules			
		1	2	3	4
Lectures (LC)					
Lab work (LW)					
Seminars (workshops/tutorials) (S)	18		18		
<i>Self-studies</i>	45		45		
<i>Evaluation and assessment (exam/passing/failing grade)</i>	9		9		
Course workload	72		72		
	2		2		

5. COURSE CONTENTS

Table 5.1. Course contents and academic activities types

Course module title	Course module contents (topics)	Academic activities types
Units	Contents	S
Module 1	Quantitative Linguistics	S
Introduction to the course & Research design	Quantitative methods. Fundamentals of research design: asking answerable questions.	S
Descriptive statistics.	Summarizing data with numbers: measures of central tendency, dispersion, standard error, and confidence intervals. Discrete and continuous distributions.	S
Correlation Analysis and regression	Types of correlation. Exploring the relationship between two measurement variables	S
Software for quantitative linguistics	SPSS, voyant tools, tools for digital content analysis	S
Module 2	Corpus Linguistics	S
Corpus building	Manual creation of a DIV monolingual corpus. Automatic creation of a DIV bilingual comparable corpus.	S
Corpus encoding and annotation	Corpus- based translation studies. Structural information. Text-linguistic information. Tasks Corpus Linguistics for Multidisciplinary Research: Coptic Scriptorium As Case Study https://dh2018.adho.org/corpus-linguistics-for-multidisciplinary-research-coptic-scriptorium-as-case-study/	S
Creating an XML TEI document	Adding a simple header. Adding linguistic annotation. Indexing the corpus.	S
Corpus tools and	Text acquisition. Annotation. Corpus management. Wordlist. Concordances.	S

Course module title	Course module contents (topics)	Academic activities types
corpus analysis		
Creating multilingual corpora	Comparable corpora. Parallel corpora. Alignment. Corpus creation checklist Corpus linguistics – YouTube . https://www.youtube.com/watch?v=YH6RcHzNDmQ Exploring parallel corpus of legal texts	S
Multilingual corpora of legal texts	UN Corpus of legal texts. Projects presentations	S

6. CLASSROOM EQUIPMENT AND TECHNOLOGY SUPPORT REQUIREMENTS

Table 6.1. Classroom equipment and technology support requirements

Type of academic activities	Classroom equipment	Specialised educational / laboratory equipment, software, and materials for course study (if necessary)
Lecture	A lecture hall for lecture-type classes, equipped with a set of specialised furniture; board (screen) and technical means of multimedia presentations.	No course-specific pre-installed software required due to reliance on cloud computing
Seminar	A classroom for conducting seminars, group and individual consultations, current and mid-term assessment; equipped with a set of specialised furniture and technical means for multimedia presentations.	No course-specific pre-installed software required due to reliance on cloud computing
Self-studies	A classroom for independent work of students (can be used for seminars and consultations), equipped with a set of specialised furniture and computers with access to the electronic information and educational environment.	No course-specific pre-installed software required due to reliance on cloud computing

7. RESOURCES RECOMMENDED FOR COURSE STUDY

Main readings:

1. Zufferey, S. (2020). Introduction to Corpus Linguistics. United Kingdom: Wiley.
2. A Practical Handbook of Corpus Linguistics. (2021). Germany: Springer International Publishing.
3. An Introduction to Research Methods in Applied Linguistics: A Practical Guide. (2021). (n.p.): Universitas Brawijaya Press.

Additional readings:

1. Gerlach, M., & Font-Clos, F. (2020). A standardized Project Gutenberg corpus for statistical analysis of natural language and quantitative linguistics. *Entropy*, 22(1), 126.
2. Language and Text: Data, Models, Information and Applications. (2021). Netherlands: John Benjamins Publishing Company.

Internet sources

Corpus Linguistics: The Basics – YouTube

<https://www.youtube.com/watch?v=32RjJ-1A-8Q>

What is CORPUS LINGUISTICS? What does CORPUS LINGUISTICS ...

<https://www.youtube.com/watch?v=d3InPxwZESU>

Increasing the impact of corpus linguistics in disciplinary education ...

<https://www.youtube.com/watch?v=iP9iJ9vHQoM>

A Glossary of Corpus Linguistics - Edinburgh University Press

<https://edinburghuniversitypress.com/book-a-glossary-of-cor> .

A Glossary of Corpus Linguistics

https://books.google.ru/books/about/A_Glossary_of_Corpus_Linguistics.html?id=bQFNaGS9_NkC&redir_esc=y - RUDN Electronic Library System (RUDN ELS)
<http://lib.rudn.ru/MegaPro/Web>

- EL "University Library Online" <http://www.biblioclub.ru>

- EL "Yurayt" <http://www.biblio-online.ru>

- EL "Student Consultant" www.studentlibrary.ru

- EL "Lan" <http://e.lanbook.com/>

2. Databases and search engines:

- electronic foundation of legal and normative-technical documentation

<http://docs.cntd.ru/>

- Yandex search engine <https://www.yandex.ru/>

- Google search engine <https://www.google.ru/>

- Scopus abstract database <http://www.elsevierscience.ru/products/scopus/>

Training toolkit for self-studies to master the course:

1. The set of lectures on the course

8. ASSESSMENT TOOLKIT AND GRADING SYSTEM FOR EVALUATION OF STUDENTS' COMPETENCES LEVEL UPON COURSE COMPLETION

The assessment toolkit and the grading system to evaluate the competences formation level (competences in part) upon the course study completion are specified in the Appendix to the course syllabus.

DEVELOPERS:

Novikov P.N.

position, department

signature

name and surname

HEAD OF EDUCATIONAL DEPARTMENT:

_____	_____	_____
name of department	signature	Atabekova A.A. name and surname

**HEAD
OF HIGHER EDUCATION PROGRAMME:**

_____	_____	_____
position, department	signature	Atabekova A.A. name and surname