

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 05.05.2026 16:17:51
Уникальный программный ключ:
ca953a0120d891083f939673078ef1a989dae18a

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

Institute of World Economy and Business

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

COURSE SYLLABUS

QUANTITATIVE LINGUISTICS AND NEW INFORMATION TECHNOLOGIES

(course title)

Recommended by the Didactic Council for the Education Field of:

45.04.02 "LINGUISTICS" (Master's degree)

(field of studies / specialty code and title)

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

«Foreign language of professional communication and specialized translation»

(higher education programme profile/specialisation title)

1. COURSE GOAL

The goal of mastering the course "Quantitative Linguistics and New Information Technologies" is to teach the processing of Russian-language and foreign-language texts for production and practical purposes, the development of educational and methodological materials using modern information resources and technologies, the compilation of databases, dictionaries and methodological recommendations in professionally oriented areas of translation, familiarization with software products of linguistic profile; the organization of information-search activities aimed at improving professional skills.

2. REQUIREMENTS FOR LEARNING OUTCOMES

Mastering the discipline (module) "Quantitative Linguistics and New Information Technologies" is aimed at the development of the following competencies /competences in part:

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

Competence code	Competence description	Competence development indicators (in the framework of this course)
GC-1	Ability to search, critically analyze problem situations based on a systematic approach, and develop a strategy of actions.	GC-1.1 Ability to analyze the problem situation and decomposes it into separate tasks
		GC-1.2 Develops a strategy for solving the task
		GC-1.3 Ability to form possible solutions to the problem
GC-2	Ability to manage the project at all stages of its life cycle	GC-2.1 Participates in project management at all stages of the life cycle
GPC-6	Ability to use modern technologies for collecting, processing and interpreting experimental data, techniques for compiling and processing scientific documentation (dissertation, report, abstract, abstract)	GPC-6.1. Ability to use modern technologies for collecting, processing and interpreting the experimental data obtained in professional activities
		GPC-6.2. Ability to apply various methods of compilation and registration of scientific documentation

3. COURSE IN HIGHER EDUCATION PROGRAMME STRUCTURE

The course "Quantitative Linguistics and New Information Technologies" refers to the core component of (B1.O.01) block of the higher educational programme curriculum.

Within the higher education programme students also master other modules and / or internships that contribute to the achievement of the expected learning outcomes as results of the course "Quantitative Linguistics and New Information Technologies" study.

Table 3.1. The list of the higher education programme components/disciplines that contribute to the achievement of the expected learning outcomes as the course study results

Competence code	Competence description	Previous Disciplines/Modules*	Subsequent Disciplines/Modules*
GC-1	Ability to carry out a critical analysis of problem situations based on a systematic approach, to develop a strategy of actions	History and methodology of science; Pedagogy and psychology of higher school; General Linguistics And The History Of Linguistic Theories ; Methods of teaching foreign languages in a non-linguistic university; Theory and practice of cross-cultural business communication	
GC-2	Able to manage the project at all stages of its life cycle	Pedagogy and psychology of higher school	Pedagogical internship
GPC-6	Owns modern technologies for collecting, processing and interpreting the experimental data obtained, techniques for compiling and processing scientific documentation (dissertation, report, abstract, abstract.)	History and methodology of science; Pedagogy and psychology of higher school; General Linguistics And The History Of Linguistic Theories ;	Educational internship

* - - to be filled in according to the competency matrix of the higher education programme

4. COURSE WORKLOAD AND ACADEMIC ACTIVITIES

The total workload of the course “Quantitative Linguistics and New Information Technologies” is 4 credits.

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
<i>Contact, academic hours</i>	34		34		
Lectures (LC)	34		34		
Laboratory work (LW)					
Practical/seminar classes (S)					

Type of academic activities		TOTAL, academic hours	Semesters/training modules			
			1	2	3	4
<i>Self-studies, academic hours</i>		65		65		
<i>Evaluation and assessment (exam/pass/fail grading), academic hours</i>		45		45		
Course workload	academic hours	144		144		
	credits	4		4		

5. COURSE CONTENTS

Table 5.1. Course contents and academic activities types

Course module title	Course module contents (topics)	Academic activities types*
Module 1. Basic Concepts of the Course "Quantitative Linguistics and New Information Technologies"	Topic 1.1. The general concept of "applied linguistics".	LC
	Topic 1.2. The history and dynamics of the development of ideas and aspects of applied linguistics	LC
	Topic 1.3. The main areas of applied linguistics.	LC
Module 2. Quantitative linguistics	Topic 2.1. Statistical processing of experimental data	LC
	Topic 2.2 General data set and qualitative reliability of the sample	LC
	Topic 2.3 Positional statistics as the main method of studying unknown texts.	LC
	Topic 2.4 The Attribution software package.	LC
Module 3. Computer technologies in linguistic research	Topic 3.1 Computer dictionaries and their classification.	LC
	Topic 3.2 Features of an electronic dictionary that exists in a multidimensional hypertext environment	LC
	Topic 3.3 The tasks and peculiarities of corpus linguistics	LC
	Topic 3.4 Translation problems and their solution using Large Data Corpora created on the basis of various languages.	LC
	Topic 3.5 Solving problems of practical application of information technologies in the study of language phenomena at different levels	LC
Module 4. Automatic text analysis	Topic 4.1 The main tasks of automatic text analysis and its application areas	LC

Course module title	Course module contents (topics)	Academic activities types*
	Topic 4.2 Linguistic support of search and information systems.	LC
	Topic 4.3 Automatic processing of sounding speech and applied phonetics	LC
Module 5. Word processors	Topic 5.1 The linguistic processor as an intermediary between the user and the database that stores the information the user is interested in	LC
	Topic 5.2 Levels of analysis. Lexical analysis: tasks, progress, result. Morphological analysis, its tasks. Syntactic analysis, its tasks.	LC
	Topic 5.3 Semantic analysis, its tasks. Stages of semantic analysis: superficial, deep, pragmatic.	LC
	Topic 5.4 The role of explanatory combinatorial dictionaries and thesauri in the semantic analysis of the text	LC

* - filled in only by **FULL**-time study: LC – lectures; LR – laboratory work; S - seminars.

CLASSROOM EQUIPMENT AND TECHNOLOGY SUPPORT REQUIREMENTS

Table 6.1. Classroom equipment and technology support requirements

Classroom type	Classroom equipment	Specialized educational/laboratory equipment, software and materials for course study (if necessary)
Computer class	An auditorium for conducting classes, group and individual consultations, current and mid-term assessment, equipped with personal computers (in the amount of ___ pcs), a board (screen) and technical means of multimedia presentations.	472 Laptop Asus X751L Intel I5 1700 MHz/8 GB/1000 GB/DVD/audio (15) Multimedia projector Benq MW526 Screen 220*220 MS Windows 8.1 64bit Microsoft Office 2013 SDL Trados Studio 2015 Adobe Reader FastStone Image Viewer
Self-studies	An auditorium 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	324 Casio XJ-F100W Multimedia projector Wall-mounted screen Digis Dsob-1106

* The premises for students' self-studies are subject to **MANDATORY** mention

7. RESOURCES RECOMMENDED FOR COURSE STUDY

Main reading

1. Zubov Aleksandr Vasil'evich. Informacionny`e texnologii v lingvistike [Tekst] : Uchebnik dlya studentov vuzov / A.V. Zubov, I.I. Zubova. - 2-e izd., ispr. - M. : Akademiya, 2022. - 206 s. - (Vy`sshee professional`noe obrazovanie. Bakalavriat. Yazy`koznanie). - ISBN 978-5-7695-9155-6 : 485.10. <http://lib.rudn.ru/MegaPro/Web/SearchResult/ToPage/1>
2. Kurinin Ivan Nikolaevich. Informacionnoe pole, personal`ny`j komp`yuter i rabota v Internete [Tekst/e`lektronny`j resurs]: Uchebnoe posobie po kursu «Informatika»: dlya studentov filologicheskogo fakul`teta i fakul`teta gumanitarny`x i social`ny`x nauk. - e`lektronny`e tekstovy`e danny`e. - M. : Izd-vo RUDN, 2022. - 385 s. <http://lib.rudn.ru/ProtectedView/Book/ViewBook/2734>

Additional reading:

1. Zalevskaya A.A. «Kvantitativnaya lingvistika i novy`e informacionny`e texnologii»: praktikum/dlya studentov 1 kursa magistratury`: 035700 Lingvistika. Ch.1: «Osnovny`e ponyatiya kursa “Kvantitativnaya lingvistika”», «Komp`yuterny`e texnologii v lingvisticheskix issledovaniyax». – Tver`: Tver. gos. un-t. 2019 – 58 s..
2. Baranov A.N. Vvedenie v prikladnuyu lingvistiku : ucheb. posobie. – 2-e izd., ispr. – M.: Editorial URSS, 2018. – 360 s.
3. Belonogov G.G. Komp`yuternaya lingvistika i perspektivny`e informacionny`e texnologii. – M.: Russkij mir, 2018. – 248 s.
4. Belyaeva L.N., Gerd A.S., Ubin I.I. Avtomatizaciya v leksikografii // Prikladnoe yazy`koznanie. – SPb: SPU, 2016. – S. 318–334.
5. Bovtenko M.A. Komp`yuternaya lingvodidaktika : ucheb. posobie. – M.: Flinta; Nauka, 2015. – 216 s.
6. Gorodeczkij B.Yu. Komp`yuternaya lingvistika: modelirovanie yazy`kovogo obshheniya // Novoe v zarubezhnoj lingvistike. – M.: Progress, 2020. – Vy`pusk XXIV: komp`yuternaya lingvistika. – S. 5–29.
7. Daczyuk S.G. Absolyutny`j gipertekst. – 2019 [E`lektronny`j resurs] URL: <http://xyz.org.ua> (data obrashheniya 16.08.2017).
8. Ivanova E.P. E`lektronny`e slovari kak komponent mul`timedijnogo informacionnogo polya // VI Stepanovskie chteniya. Yazy`k i kul`tura. Na materiale romano-germanskix i vostochny`x yazy`kov : mat-ly` dokl. i soob. mezhdunar. konf. M.: RUDN, 2019. – S. 244–246.
9. Kazakevich O.A. Avtomatizaciya leksikograficheskix rabot. Avtomaticheskie slovari (Obzor zarubezhny`x publikacij) // NTI. – Ser. 2. – 2018. – № 9. – 25–29.
10. Karpova O.M., Menagarishvili O.V. E`lektronny`e slovari anglijskogo yazy`ka // Slovar` v sovremennom mire : mat-ly` 3 mezhdunar. shk.–sem. – Ivanovo: Yunona, 2015. – S. 121–124.
11. Kobrin R.Yu. Opy`t semanticheskij klassifikacii slovarej // Teoreticheskie problemy` nauchno-texnicheskij terminologii i praktiki perevoda. – Omsk, 2019. – S. 70–85.
12. Komp`yuterny`j korpus gazetny`x tekstov russkogo yazy`ka konca XX-go veka <http://www.philol.msu.ru/~lex/korpus.html>
13. Kuper I.R. Gipertekst kak sposob kommunikacii. – 2022 [E`lektronny`j resurs] URL: <http://http:www.nir.ru/socio/scipubl/sj1-2-00kuper.html>xyz.org.ua (data obrashheniya 10.11.2016).
14. Marchuk Yu.N. Vy`chislitel`naya leksikografiya. – M.: Nauka, 2018. – 183 s.

15. Marchuk Yu.N. Metody` modelirovaniya perevoda. – M.: Nauka, 2019. – 201 s.
16. Marchuk Yu.N. Osnovy` komp`yuternoj lingvistiki. – M.: MGOU, 2022. – 234 s.
17. Nelyubin L.L. Komp`yuternaya lingvistika i mashinny`j perevod. – M.: VCzP, 2017. – 151 s.
18. Novikov L.A. Uchebny`e slovari, ix specifika i tipy` // Novikov L.A. Izbranny`e trudy`. T.2. – M.: Izd-vo RUDN, 2014. – S. 688–705.
19. Potapova R.K. Novy`e informacionny`e texnologii i lingvistika : ucheb. posobie dlya vuzov. – 4-e izd., stereotip. – Editorial URSS KOMKNIGA, 2015. – 368 s.
20. Selegej V. E`lektronny`e slovari i komp`yuternaya leksikografiya [E`lektronny`j resurs]. Sivakova N.A. Leksikograficheskoe opisanie anglijskix i russkix fitonimov v e`lektronnom glossarii : dis. ... kand. filol. nauk. – Tyumen`, 2014. – 162 s.
21. Tuzlukova V.I. Tipologiya pedagogicheskix leksikograficheskix istochnikov v mezhdunarodnoj pedagogicheskij leksikografii http://rspu.edu.ru/science/conferences/01_04_09/tuzlukova_pr.html [21.05.17]
22. Ubin I.I. Avtomaticheskij slovar` kak sredstvo avtomatizacii leksikograficheskix rabot // Teoriya i praktika nauchno-texnicheskij leksikografii : sb. st. – M.: Russkij yazy`k, 2014. – S. 234–240.
23. Cherkasova G.A. Komp`yuterny`j asociativny`j tezaurus: baza nauchny`x issledova-nij // Yazy`kovoje soznanie: formirovanie i funkcionirovanie / Pod red. N.V. Ufimcevoj. – M., 2015. – S. 129–134.
24. E`pshtejn V.L. Vvedenie v gipertekst i gipertekstovy`e sistemy`, 2014 [E`lektronny`j resurs] URL: <http://www.ipu.rssi.ru/publ/epstn.htm> (data obrashheniya 27.01.2016).
25. Kennedy G. An Introduction to Corpus Linguistics. – Addison Wesley: Longman Ltd., 2015 – xii 315 p.

Internet sources:

1. Electronic libraries (EL) of RUDN University and other institutions, to which university students have access on the basis of concluded agreements
<http://lib.rudn.ru/MegaPro/Web> - RUDN Electronic Library System (RUDN ELS)
<http://lib.rudn.ru/MegaPro/Web>
<http://www.biblioclub.ru/> EL "University Library Online" <http://www.biblioclub.ru>
<http://www.biblio-online.ru/> EL "Yurayt" <http://www.biblio-online.ru>
<http://www.studentlibrary.ru/> EL "Student Consultant" www.studentlibrary.ru
<http://e.lanbook.com/> EL "Lan" <http://e.lanbook.com/>
- EL "Trinity Bridge"
2. Databases and search engines:
<http://docs.cntd.ru/> 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. Lectures on the course of «Quantitative Linguistics and New Information Technologies».

Section 1. Basic Concepts of the Course "Quantitative Linguistics and New Information Technologies"

The general concept of "applied linguistics". Differences in the understanding of the tasks and specifics of applied linguistics in the West and in Russia. The connection of applied linguistics with other sciences.

The history and dynamics of the development of ideas and aspects of applied linguistics

The main areas of applied linguistics. Applied linguistics, quantitative linguistics, computational linguistics. Leading methods used in various areas of applied linguistics.

The general concept of "information technology".

"Electronic resource" as a general concept. Types of electronic resources and principles of their classification. "Database" as a set of ordered information about certain objects in a certain way. Different interpretations of the concept of "data corpus". Various interpretations of the concept of "hypertext".

Tasks of using the capabilities of quantitative linguistics and new information technologies in the professional activity of a graduate of a master's degree in the field of "Theory of teaching foreign languages and cross-cultural communication".

Section 2. Quantitative linguistics.

Statistical processing of experimental data: the study of varying features (quantitative, qualitative, branching). General data set and qualitative reliability of the sample Decryption of messages or texts to detect information presented in a way unknown to the researcher. The concepts of "cipher" and "code". Differences between cryptography and decryption tasks. "Machine decryption". Positional statistics as the main method of studying unknown texts. Stages of application of the method of positional statistics: splitting of continuous text into separate blocks, analysis of the morphology of the word; reception of the "environment" of the word. Combinatorial method as decryption "from the inside" and "from the outside". Attribution (authorization) of a message or text; areas of application for various purposes. Formal-quantitative methods and identification of the features of the author's language, realized at the subconscious level. The Attribution software package: the principles of its operation, the results obtained.

Section 3. Computer technologies in linguistic research

Computer dictionaries and their classification. Features of an electronic dictionary that exists in a multidimensional hypertext environment Electronic libraries. Electronic encyclopedias. Concordances. Associative thesauri and the possibilities of their use for different purposes. The possibilities of using monolingual and multilingual electronic dictionaries to form a data corpus when conducting research for various purposes.

Tasks and specifics of corpus linguistics as a science engaged in the development of general principles for the construction and use of language data corpora using computer technology.

The method of selecting texts (sampling). Representativeness of certain phenomena in a particular corpus. Types of markings. The main types of buildings and the possibilities of their use for solving educational, methodological and research tasks. Translation problems and their solution using Large Data Corpora created on the basis of various languages. The importance of the corpus of national languages for solving problems of cross-cultural communication; use for the purpose of teaching languages (native, second / foreign).

Issues of forming a database based on electronic dictionaries and corpora of national languages. The possibilities of quantitative and qualitative analysis of the selected data

corpus. Presentation of the results of statistical processing and products of qualitative analysis of materials using drawings, diagrams, etc.

Solving problems of practical application of information technologies in the study of language phenomena at different levels Ways of using information technologies in solving translation problems. The specifics of the use of information technologies in solving problems of cross-cultural communication. Information technologies in solving problems of language teaching.

Section 4. Automatic text analysis

The main tasks of automatic text analysis and its application areas Linguistic support of search and information systems: automatic categorization of documents; automatic classification (splitting texts into groups according to a given classifier) and clustering of texts (based on thematically similar content); automatic referencing (generation of texts from the most significant sentences of a document or group of documents). Tasks of extracting information from the text. Deeper analysis of the extracted information – extraction of knowledge (data mining): names of entities (persons, geographical names, etc.); extraction of facts; extraction of opinions; automatic abstracting. Systems that simulate the language interaction of a computer with a person. Automatic processing of sounding speech and applied phonetics Linguistic components of automatic text analysis: tokenization (selection of words and sentence boundaries), morphological analysis (reduction of all forms of a word to one word form), syntactic analysis module, semantic analysis module, anaphora resolution module.

Section 5. Word processors

The linguistic processor as an intermediary between the user and the database in which the information of interest to him is stored; the goals and tasks performed by him. Operations performed by the linguistic processor. Levels of analysis. Lexical analysis: tasks, progress, result. Morphological analysis, its tasks. Methods of morphological analysis: declarative, procedural, combined, probabilistic and statistical; their advantages and disadvantages. Syntactic analysis, its tasks. Formal-grammatical and probabilistic-statistical approaches to syntactic analysis. Stages of syntactic analysis: pre-syntactic, proper syntactic, post-syntactic analysis – the sentence is transformed into a sequence of basic text units that correspond to individual words and punctuation marks. Semantic analysis, its tasks. Stages of semantic analysis: superficial, deep, pragmatic. The concept of a semantic node. Sources of information about links: data from the parser and thesaurus dictionaries. Systems: syntagmas and paradigms of relations between lexemes as the basis of semantic analysis. The role of explanatory combinatorial dictionaries and thesauri in the semantic analysis of the text Areas of application of word processors as internal translators. Internal translators (from some indexing languages to others and to translation languages).

2. Methodological recommendations for the preparation of a report on the discipline "Quantitative linguistics and new information technologies"

One of the forms of independent work of students is the preparation of a scientific report to discuss it at a practical (seminar) lesson.

The purpose of the scientific report is to develop students' skills of analytical work with scientific literature, analysis of debatable scientific positions, argumentation of their own views. Preparation of scientific reports also develops the creative potential of students.

One of the forms of independent work of students is the preparation of a scientific report to discuss it at a practical (seminar) lesson.

The purpose of the scientific report is to develop students' skills of analytical work with scientific literature, analysis of debatable scientific positions, argumentation of their own views. Preparation of scientific reports also develops the creative potential of students. The scientific report is prepared under the guidance of a teacher who conducts practical (seminar) classes

Recommendations to the student:

- before starting work on writing a scientific report, coordinate the topic, structure, literature with the teacher, as well as discuss key issues that should be disclosed in the report;
- submit a report to the supervisor in writing;-to speak at the seminar with a 10-minute presentation of his scientific report, to answer the questions of the students of the group.

Requirements for the design of a scientific report:

font -TimesNewRoman, font size -14, line spacing -1.5, margin size - 2.5 cm, indent at the beginning of the paragraph -1.25 cm, formatting by width); the report sheets are foldered

The title page indicates the name of the educational institution, the name of the department, the name of the discipline, the topic of the report, the full name of the student;-to the structure of the report -the table of contents, introduction (relevance, purpose and objectives are indicated), the main part, the conclusions of the author, the list of references (at least 5 positions). The volume is agreed with the teacher. At the end of the work, the date of its completion and the signature of the student who completed the work are put.

The overall assessment for the report takes into account the content of the report, its presentation, as well as the answers to the questions.

3. Methodological recommendations on working with literature of the course of «Quantitative Linguistics and New Information Technologies»:

Any form of independent work of a student (preparation for a seminar, writing an essay, term paper, report, etc.) begins with the study of relevant literature both in the library and at home. The main and additional reading is selected for each topic of the discipline.

Main reading:

- these are textbooks and teaching aids.

Additional reading:

- these are monographs, collections of scientific papers, magazine and newspaper articles, various reference books, encyclopedias, Internet resources.

Recommendations to the students:

it is advisable to carefully review the selected monograph or article. In the books, you should look through the table of contents and the scientific reference apparatus, read the annotation and the preface. It is advisable to look through illustrations, tables, diagrams, appendices. It will allow you to find out which chapters should be read and which ones should be skipped;- in a book or magazine owned by the student, key positions can be highlighted with a marker or noted. When working with an online source, it is also advisable to highlight important information;-if a book or magazine is not the property of a student, then it is advisable to write down the page numbers. Later, you should return to them, re-read or rewrite the necessary information. The physical act of recording helps information get into the "memory bank".

The following types of records are distinguished when working with literature:

Synopsis is a brief schematic record of the main content of the scientific work. The goal is not to rewrite the work, but to identify its logic, the system of evidence, and the main conclusions. A good synopsis should combine completeness with brevity.

A quote is an exact reproduction of the text. Enclosed in quotation marks indicates the source page.

Abstracts are a concentrated presentation of the main provisions of the material read.

Abstract - a very brief summary of the content of the work read.

Summary - the most general conclusions and provisions of the work, its conceptual results.

Writing in one form or another not only contributes to the understanding and assimilation of the material being studied, but also helps to develop the skills of clear presentation in writing of certain theoretical issues.

* - The training toolkit for self- studies to master the course is placed on the course page in the university telecommunication training and information system under the set procedure

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. <https://esystem.rudn.ru/course/view.php?id=641>

The assessment toolkit and the grading system are formed on the basis of the requirements of the relevant local normative act of RUDN University (regulations / order).

DEVELOPERS:

Ass. Professor FLD EF

Position, Educational Department

Ivanova A.G.

Signature

Name and surname

HEAD OF THE HIGHER EDUCATION PROGRAMME:

FLD EF

Educational Department

Malyuga E.N.

Signature

Name and surname