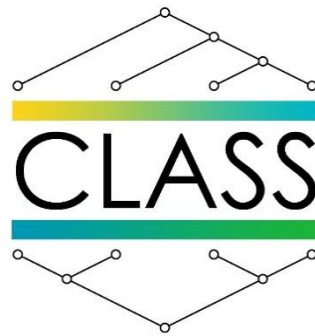




Co-funded by the
Erasmus+ Programme
of the European Union



*Development of the interdisciplinary master program on Computational Linguistics at Central
Asian universities*

585845-EPP-1-2017-1-ES-EPPKA2-CBHE-JP

Developing e-books (methodological books)

Work Package 4 / Task 4.4. / Deliverable 4.2.

March 2021





Content

E-books produced.....	3
ENU.....	3
KazNU	5
SamSIFL.....	8
TSUULL.....	8
ABKRU.....	8
UrSU	8

E-books produced

A total of 12 books have been produced.

The books produced in the CLASS project have been classified into two types:

1. **Books created for the pilot implementation experience of the master.** Some books are considered by the CLASS coordinating institution as a deliverable just accessible for the EACEA because they have been created as internal material for the pilot implementation experience of the master, and not to be accessible for the general public.
2. **Books created and validated for the master.** Some books are considered by the CLASS coordinating institution as a deliverable both accessible for the EACEA and the general public, because they have been produced and, afterwards, validated.

Hereunder the list of the books produced are listed, by the partner institutions that produced them. For each of them it is also indicated the type: **book created for the pilot experience of the master** (in blue) or **book created and validated for the master** (in pink).

ENU

ENU prepared these textbooks:

The textbooks (study guides) are available here:

<https://files.erasmusclass.eu/apps/files/?dir=/CLASS/WP4%20Dissemination%20and%20Sustainability/WP4%20Deliverables/ENU%20e-book&fileid=7661>

- 1) Formal grammar and automata (three versions: in Kazakh, English, and Russian language) – **book created and validated for the master**
- 2) Ontologies, semantic technologies (three versions: in Kazakh, English, and Russian language) - **book created for the pilot experience of the master**
- 3) Speech processing (three versions: in Kazakh and English and Russian language) - **book created for the pilot experience of the master**
- 4) Statistical methods for NLP (three versions: in Kazakh, English, and Russian language) – **book created and validated for the master**

Developing e-books and methodological books on CL

Within the framework of the project, ENU has developed the following electronic textbooks:

The title of the electronic textbook: "Formal grammar and automata". – **book created and validated for the master**

Authors: SHARIPBAY A.A., NIYAZOVA R.S., OMARBKOVA A. S., RAZAKHOVA B. SH.

The electronic textbook is posted on the Moodle.

Volume: 173 pages

The electronic study guide is intended for undergraduates in the educational program "Computational Linguistics", for teachers and all those who independently want to study formal grammars and automata, want to learn more about natural language processing technology and the creation of language processors.

The electronic study guide was developed within the framework of the Erasmus + EAC-A03-2016 CLASS project "Development of an interdisciplinary master's program in computational linguistics at universities in Central Asia".

The title of the electronic textbook: "Ontologies, semantic technologies". - [book created for the pilot experience of the master](#)

Authors: SHARIPBAY A.A., MUKANOVA A.S., NIYAZOVA R.S.,
RAZAKHOVA B. SH., OMARBEKOVA A. S., Abdurakhmonova N. Z.

The electronic textbook is posted on the Moodle.

Volume: 88 pages

The content of the electronic textbook is compiled in accordance with the educational program "Computational Linguistics". The electronic textbook is aimed at mastering the theory and practice, as well as methods and means of ontological engineering, which are an important class of computational linguistics tools. Issues such as the main ways of representing knowledge, ontology and its main components are systematically presented; definition and description of classes; the purpose and functions of the inference machine (resonator); building a class hierarchy; technology for developing a domain ontology; descriptive logics underlying ontologies; OWL - ontology description language; practical application of ontologies. Examples are given; practical exercises performed on a computer; control questions and tasks. The electronic textbook was developed within the framework of the Erasmus + EAC-A03-2016 CLASS project "Development of an interdisciplinary master's program in computational linguistics at universities in Central Asia".

The title of the electronic textbook: "Statistical methods in Natural Language Processing". – [book created and validated for the master](#)

Author: Kudubayeva S.A.

The electronic textbook is posted on the Moodle.

Volume: 165 pages

The electronic study guide is intended for students enrolled in the educational program "Computational Linguistics", for teachers and everyone who wants to study independently the use of statistical methods for analyzing text, audio speech, wants to learn more about natural language processing technology. The electronic study guide was developed within the framework of the Erasmus + EAC-A03-2016 CLASS project "Development of an interdisciplinary master's program in computational linguistics at universities in Central Asia".

The title of the electronic textbook: "SPEECH PROCESSING". - [book created for the pilot experience of the master](#)

Authors: BEKMANOVA G.T., OMARBEKOVA A.S., KUSSEPOVA L.T.

The electronic textbook is posted on the Moodle.

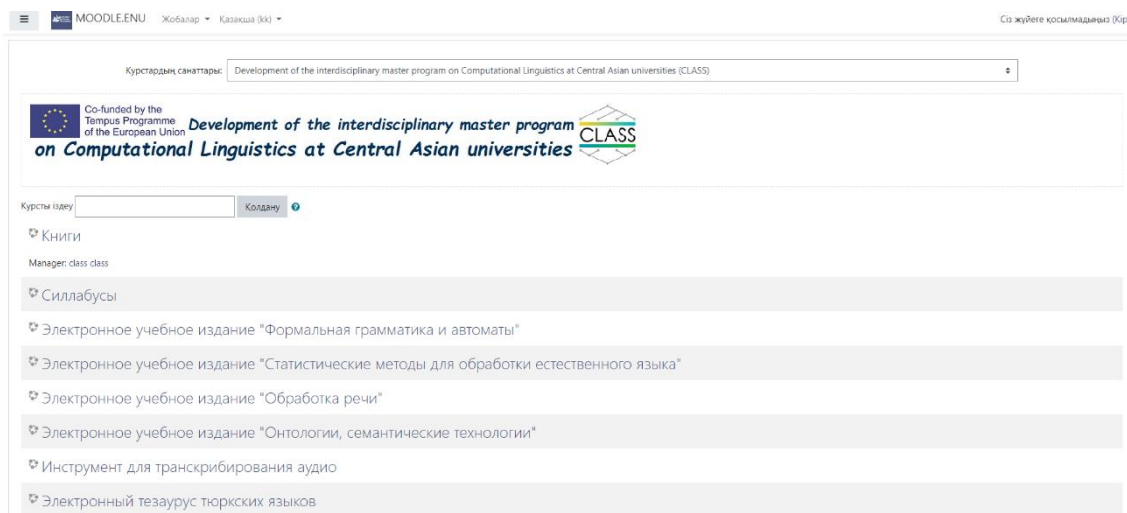
Volume: 105 pages

The electronic textbook is intended for undergraduates and doctoral students in the educational program "Computational Linguistics", teachers, scientists and all those who independently want to study the methods of computer processing of the Kazakh language based on their morphological analysis and synthesis, the current state of the problem of natural language processing, the theoretical foundations of word recognition, morphological and linguistic patterns of the Kazakh language for the formal model of inflection.

The electronic textbook was developed within the framework of the Erasmus + EAC-A03-2016 CLASS project "Development of an interdisciplinary master's program in computational linguistics at universities in Central Asia".

These textbooks are posted on the Moodle:

<https://moodle.enu.kz/course/index.php?categoryid=154>



The textbooks (study guides) are available here:

<https://files.erasmusclass.eu/apps/files/?dir=/CLASS/WP4%20Dissemination%20and%20Sustainability/WP4%20Deliverables/ENU%20e-book&fileid=7661>

Link to the books: <http://comling.enu.kz/index.php/ru/knigi>

KazNU

- *Corpus linguistics: learner's dictionary. Key terms and concepts* (in Kazak language) – **book created and validated for the master**
- *Corporate linguistics: key terms and concepts. Educational dictionary* (bilingual: in Kazak and in English languages) – **book created and validated for the master**
- *Innovative technologies and teaching methods. Dictionary of terms* (in Kazak language) - **book created for the pilot experience of the master**
- *Computer linguistics: language resources. Textbook* (2 versions: in Kazak and in English language) – **books created and validated for the master**

Due to COVID-19 quarantine, all courses in KazNU were done remotely. Professors of the educational program "Computational Linguistics" created new educational material to adapt to this complex situation.

The person responsible for the preparation of the e-book (methodological book) on the discipline "Language Resources" is Madina Mansurova. For the e-book, all the materials are collected, now it remains to carefully check all the material and put everything together. The completion of the task is scheduled for July 5, 2020.

The teachers of KAZNU developed and published the following educational materials for the "Language Resources" discipline:

1. Corpus linguistics: a glossary of basic terms and concepts (in Kazakh language— Almaty: Kazakh University, 2018. – 40 p. (Authors: Madieva G.B., Ismailova N.).) – **book created and validated for the master**. This training dictionary is a collection of basic terms and concepts used in computational and corpus linguistics. Designed for all students interested in the problems of case and computer linguistics.

2. Corpus Linguistics: Educational Dictionary of Key Terms and Concepts. Learner's Dictionary (in Russian and English languages)/ Compiled by G.B. Madieva, S. Bektemirova, N. Ismayilova. – Almaty: Kazakh University, 2019. – 52 p. – **book created and validated for the master**. This dictionary is an experience of creating an educational dictionary of corpus linguistics in Russian, prepared for educational purposes for undergraduate, masters and Ph.D. doctoral students and for anyone interested in corpus linguistics problems. The dictionary contains basic terms and concepts used to study, study the problems of corpus linguistics, the creation of cases, as well as to perform independent work by students, undergraduates, and doctoral students in relevant disciplines. The study of key terms and concepts in courses devoted to the problems of corpus linguistics makes it possible to form competencies in the possession and use of conceptual and terminological apparatus of the new for

Kazakhstani universities direction of linguistics, dealing with theoretical and applied issues of compiling national corpuses of various languages.

3. Innovative technologies and teaching methods. Dictionary of terms. Learner's Dictionary/ Compiled by G. Madiyeva, S. Bektemirova, N. Ismayilova, Madiyeva D. – Almaty: Kazakh University, 2019. – 31 p. - **book created for the pilot experience of the master**. This dictionary is

an experience of creating an educational dictionary of corpus linguistics in Russian, prepared for educational purposes for undergraduate, undergraduate and doctoral students Ph.D. and for anyone interested in corpus linguistics problems. The dictionary contains basic terms and concepts used to study, study the problems of corpus linguistics, the creation of cases, as well as to perform independent work by students, undergraduates, and doctoral students in relevant disciplines. The study of key terms and concepts in courses devoted to the problems of corpus linguistics makes it possible to form competencies in the possession and use of conceptual and terminological apparatus of the new for Kazakhstani universities direction of linguistics, dealing with theoretical and applied issues of compiling national corpuses of various languages.

4. Computational Linguistics: Language Resources

The name of the textbook: "Computational Linguistics: Language Resources".). – **book created and validated for the master**

Authors: Madiyeva G., Mansurova M., Dossanova A., Kadyrbek N.

Recommended for publication: by the Scientific Council of the Faculty of Philology and World Languages and the Scientific Council of the Faculty of Information Technologies and the Editorial and Publishing Council of al-Farabi KazNU.

Reviewers: Doctor of Technical Sciences, Professor Tukeyev U.A. and Doctor of Philology Ibrayeva Zh.K.

House and year of publication: Almaty: Kazakh University, 2021.

Volume: 151 pp.

The textbook "Computational Linguistics: Language Resources" is intended for undergraduates enrolled in the Master's program in the direction of training "Information and Communication Technologies". This tutorial will improve students' in-depth understanding of the structure, purpose and function of various language resources, as well as develop the ability to use existing language resources and create new ones for natural language processing. As a result of studying this course, students will be able to describe the theoretical and organizational-methodological issues of the construction and operation of the linguistic support subsystem, develop new tools for collecting, storing and managing data, reuse existing and develop new language resources, and adapt language resources for certain NLP applications.

SamSIFL

- *Dictionary of terms of computational linguistics and NLP in English, Russian and Uzbek (trilingual)* - [book created for the pilot experience of the master.](#)

TSUULL

- *Linguistic support for machine translation (monograph)* (in Uzbek language) - – [book created and validated for the master](#)

ABKRU

- Satmaganbetova Zh. Z. *Basics of Computer linguistics*. Handbook. - Kostanay: A. Baitursynov KSU, 2020. - 64 pages (in Kazakh language) - [book created for the pilot experience of the master.](#)

The Handbook covers the following topics: Introduction to computer linguistics Tools of computer linguistics. Types of text processing applications. Word processing and analysis programs. The concept of machine translation. Phonetic computer program. Linguistic Corpus. Trends of Modern Corpus Linguistics Development. Overview of Text Corpora in the World. Kazakhstan Corpora. Almaty Kazakh Language corpus. New Corpus of the Kazakh language "Kazcorpus"; Types of modern linguistic buildings; Stages of the body assembly process; The Handbook information is for training the students of Technology and Linguistics major and minor courses with the Kazakh language of instruction.

Approved by the Educational and Methodological Council of A.Baitursynov KSU, Recommended for publishing. Protocol dated 4.06. 2020

Responsible for the task: Zhanar Satmagambetova

<https://ksu.edu.kz/cooperation/erasmusplus/class/materials/>

UrSU

- *Natural Language understanding* (in English language) – [book created and validated for the master](#)