

ELMIRA MUSTAKIMOVA

Moscow, Russia

Mobile:

+7 (926) 908 06 56

Email:

egmustakimova_2@edu.hse.ru

Webpage:

www.hse.ru/en/staff/egmustakimova

View this resume online at www.elmiram.github.io

EDUCATION

BACHELORS STUDIES

2012 – present

National Research University «Higher School of Economics» (NRU HSE), Moscow, Russia

Currently – senior student

Program: Fundamental and Computational Linguistics

Modules

- Theory of Language (Phonetics, Morphology, Syntax, Semantics, Discourse)
- Experimental Linguistics and Neurolinguistics
- Language Diversity and Typology
- Programming (Python, R)
- Algorithms
- Databases
- Natural Language Processing

SUMMER SCHOOLS

July 2014 – Yiddish summer school "Yiddish on the Continent"

Université Libre de Bruxelles, Institute for Jewish Studies, Brussels, Belgium

September 2014 – Summer School "Introduction to Contemporary Neurolinguistics"

NRU Higher School of Economics, Moscow

June 2015 – Summer School "Pronouns: syntax, semantics, processing"

NRU Higher School of Economics, Moscow

WORK EXPERIENCE

NOVEMBER 2013 – PRESENT TEACHING ASSISTANT

School of Linguistics, National Research University “Higher School of Economics”

Courses: *Programming for Linguists* in 2013, *Programming (Python)* from 2014 to present

Duties:

- Collecting and assessing programming assignments and tests
- Taking a small group of students for additional consultations

- Regularly giving feedback to the teacher on students' progress and major mistakes

FEBRUARY 2014 – PRESENT LABORATORY ASSISTANT

Linguistic Laboratory of Corpus Technologies, National Research University "Higher School of Economics"

Duties:

- Developing a modern platform for two corpora developed at NRU HSE Corpus Lab (including search engine and annotation system, Python+Django framework used)
- Updating the corpus webpages
- Providing technical support to the users of the corpora

PROJECTS

2014 – PRESENT «RESEARCH IN LEXICAL TYPOLOGY»

A team of lexical typologists studies the verbs denoting 'falling' in the world's languages. In this project I collect and analyze data of 4 languages (Indonesian, Yiddish, Tatar and Japanese) and develop a web-interface for the database. The interface will be accessible by link www.web-corpora.net/falling, the work is now in progress.

2014 – PRESENT «CORPUS OF RUSSIAN STUDENT TEXTS» AND «RUSSIAN LEARNER CORPUS»

The Corpus of Russian Texts (www.web-corpora.net/learner_corpus) contains texts written by students, native speakers of Russian, in academic environment (usually project writing samples from various courses). The Russian Learner Corpus (www.web-corpora.net/heritage_corpus) contains texts written by learners of Russian with various backgrounds (different native languages, heritage speakers of L2 learners). I develop the online platform for these two corpora. My goal is to create a complex system that would enable users to add and annotate texts, create subcorpora and to perform exact and lexico-grammatical search.

NOVEMBER – MARCH 2015 STUDENT SEMINAR ON LINGUISTICS AND RELATED FIELDS

At some point, I realized that I am not aware of the research that some of my classmates or other students in our department are doing. Therefore, I suggested a monthly student seminar, where everyone can share his/her research and receive comments from each other. Quite many students supported the idea, and eventually I began chairing our monthly sessions.

2014 «AUTOMATIC IDENTIFICATION OF SEMANTIC AUREOLE OF METER»

"Automatic identification of semantic aureole of meter" is a final project at a course "Meter in Russian Poetry". The work is inspired by the idea that a particular poetry meter in the poet's mind is subconsciously connected to a particular set of topics. This idea was quite popular among some Russian researchers in the XXth century. I tried to use various machine learning techniques to cluster Russian poems by topics and see whether the clusters correspond to meter clusters. Though the answer was negative, the work drew interest of philologists and I gave a talk about the project at an International Conference of Young Philologists in Tartu.

2014 «AUTOMATIC IDENTIFICATION OF PARADIGM GAPS IN CORPORA»

"Automatic identification of paradigm gaps in corpora" is the topic of my third year final paper. At that time, articles of Adam Albright on paradigmatic gaps fascinated me; I also became interested in the potential of machine learning for linguistic research. These two interests resulted in a project where I used decision trees, KNN classifier and a Naïve Bayes classifier to automatically decide on the basis of accessible Russian corpus data whether a particular word has a defective paradigm or not.

2013 – 2014 «CREATION OF CORPUS INSTRUMENTS FOR YIDDISH»

In this project, a team of students and professors created instruments for the Yiddish language, such as a transliterator, automatic alignment of Yiddish texts with corresponding audio files, comprising a corpus out of a mere collection of texts. My part in the project was automatic disambiguation of Yiddish. We already obtained a morphological parser for Yiddish, but its output left much to be desired, so I took disambiguation upon myself. I used several disambiguation strategies (rule-based, data-driven Brill algorithm, HMM and Viterbi decoding, simple bigram model). Later, this work became my second year final paper.

2013 – 2014 «DATABASE AS AN INSTRUMENT FOR LEXICO-TYPOLOGICAL RESEARCH: VERBS OF SOUNDS AND THE MODELS OF THEIR SEMANTIC DERIVATION»

The aim of this project was to collect a database of verbs of animal sounds and then to analyze the linguistic categorization of animals and the metaphorical uses of their sound verbs. The work was accomplished in a research group consisting of students and professors of School of Linguistics NRU HSE. My contribution to the project is the following: a) collecting and analyzing the linguistic data in Tatar and Indonesian languages, b) developing a visualizing interface for the collected database (www.web-corpora.net/zvukimu).

PUBLICATIONS

Kuzmenko, Elizaveta, Mustakimova, Elmira. Syntactic Phrase of the Type X-t' ne pereX-t' // In [Proceedings of the 6th International Conference on Meaning-Text Theory](#)

Kuzmenko, Elizaveta, Mustakimova, Elmira. [Automatic Disambiguation in the Corpora of Modern Greek and Yiddish](#) // In Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue" (2015)

Mustakimova, Elmira, Orekhov, Boris. Verbs of Animal Sounds in Bashkir and Tatar // In «И хрюкотали зелюки: verbs of animal sounds – typology of metaphors» / Eds: T.Reznikova, A.Vyrenkova, B.Orekhov, E.Rakhilina. Yazyki slavyanskih kultur, 2015. (in print)

Mustakimova, Elmira. Verbs of Animal Sounds in the Indonesian Language // In «И хрюкотали зелюки: verbs of animal sounds – typology of metaphors» / Eds: T.Reznikova, A.Vyrenkova, B.Orekhov, E.Rakhilina. Yazyki slavyanskih kultur, 2015. (in print)

CONFERENCES

2015

October 14 – 16. 3rd Conference in General, Nordic and Slavic linguistics, GeNSLing 2015 (Moscow, Russia). Talk: *Verbs of Falling in Typological Perspective*

October 10 – 11. ConCorT Junior. Corpus Technologies and Computational Methods in Modern Humanities (Moscow, Russia). Talk: *Database Visualization: Verbs of Animal Sounds*

July 20 – 24. Corpus Linguistics 2015 (Lancaster, UK). Talk: *Corpus of Russian Student Texts: goals, annotation, and perspectives*

June 16 – 19. Pronouns: Syntax, Semantics, Processing (Moscow, Russia). Talk: *Automatic Identification of Paradigm Gaps in Corpora*

May 27 – 30. International Conference "Dialogue 2015": Computational Linguistics and Intellectual Technologies (Moscow, Russia). Talk: *Automatic Disambiguation in the Corpora of Modern Greek and Yiddish*

April 24 – 27. International Conference of Young Philologists (Tartu, Estonia). Talk: *Automatic Identification of Semantic Aureole of Meter*

April 14. Student Conference of School of Linguistics NRU HSE and MCTTU (Moscow, Russia). Talks: *Automatic Identification of Paradigm Gaps Based on the Corpus Data (Project Overview)* and *Corpus of Russian Student Texts*

2014

December 1 – 3. Seminar on Cognitive Semantics and Lexical Typology (Helsinki, Finland). Talk: *A New Typological Project: Typology of Verbs of Falling*

September 26 – 27. RES ET VERBA - 4 (Moscow, Russia). Talk: *What Sounds Animals Make in the World's Languages*

2013

August 29 – 30. 6-th International conference on Meaning-Text theory (Prague, Czech Republic). Talk: *Syntactic Phraseme of the Type X-t' ne pereX-t'*

SKILLS

Linguistics research

- linguistic data collection: work with native speakers, corpus research and internet language data
- linguistic data analysis: frequency, collocations, semantic analysis
- design of linguistics experiments and questionnaires

Programming

- Languages: Python, R, JavaScript
- Frameworks: Django, Flask
- Databases: SQL (MySQL, SQLite), NoSQL (ElasticSearch, Neo4j)
- IDE: PyCharm
- LaTeX
- Git

Language usage:

- Russian – native
- English – advanced
- German, French, Yiddish – intermediate
- Japanese, Indonesian – beginner

HOBBIES

- Papercraft. I love all sorts of paper art: origami, applique, paper sculptures, and also cardmaking.
- Bookbinding. Bookbinding is a long process of sewing signatures from pieces of paper, gluing flyleaves, designing and making book-covers from cardstock and cloth. Usually I create blank notebooks and give it as presents to friends, but sometimes I print out texts and compile a real book. I love it because you can clearly see how your work converts into a resulting object, and I love the feeling when you hold a finished book after several evenings of work on it.
- Smashbooking. The idea is simple: to collect photos, tickets, short notes about the day and smash it all into a book. Thus, you have sort of a mixture of visual, textual and tangible memories. It helps me express my feelings, analyze emotions, how I spend time, and at the same time keep memories, and do it in an unusual fun way.
- Rollerskating. Both roller blades and quad skates are equally good for me. I'm not a professional but can do some cool-looking tricks.
- Sketching. I have learned about sketchnoting from Mike Rohde's "The Sketchnote Handbook" about 2 years ago. Since then I enjoy drawing sketches and presenting ideas visually instead of merely writing them in plain text. It helps at lectures, conferences and when planning a new project!
- Japanese language. For a long time I've been fascinated by the Japanese language and culture. In my free time I try to learn another couple of kanji and new words, or write a couple of sentences.