National Research Tomsk State University

приоритет2030^
лидерами становятся

International Research & Practice Seminar

Experimental Studies of Language and Speech: Resource Bases and Technologies (E-SoLaS)

supported by Tomsk State University Development Programme (Priority 2030)

October 16-17, 2023

Experimental Studies of Language and Speech (E-SoLaS-2023)

«Экспериментальные исследования языка и речи: ресурсные базы и технологии» (E-SoLaS-2023)

Проводится при поддержке Программы развития Томского государственного университета (Приоритет-2030)

16-17 октября 2023 года

Hosts

Linguistic Anthropology Laboratory, National Research Tomsk State University, Russia (http://illa.tsu.ru/)

Faculty of Philology, National Research Tomsk State University, Russia (http://www.tsu.ru/)

Organizing Committee

Chair of Organizing Committee

Zoya I. Rezanova, Professor, Head of the Department of General, Computer and Cognitive Linguistics, Deputy Head of the Linguistic Anthropology Laboratory, National Research Tomsk State University

Co-chair of Organizing Committee

Olga V. Nagel, Dean of the Faculty of Foreign Languages; Senior Research Fellow, Linguistic Anthropology Laboratory, National Research Tomsk State University

Members of Organizing Committee

Oksana V. Tsaregorodtseva, Senior Research Fellow, Linguistic Anthropology Laboratory, National Research Tomsk State University; Associate Member of Language and Cognition Group, University of Tübingen

Elena D. Artemenko, Deputy Head of the Laboratory of Social and Cognitive Informatics, National Research University Higher School of Economics; Senior Research Fellow, Laboratory of Linguistic Anthropology, National Research Tomsk State University

Valeriia E. Vladimirova, postgraduate student of the Department of General, Computer and Cognitive Linguistics, Faculty of Philology, Junior Research Fellow, Linguistic Anthropology Laboratory, National Research Tomsk State University

Irina S. Korshunova, postgraduate student of the Department of General, Computer and Cognitive Linguistics, Faculty of Philology, Junior Research Fellow, Linguistic Anthropology Laboratory, National Research Tomsk State University

Working group

Kseniya Kucherova, student of the Department of General, Computer and Cognitive Linguistics, Faculty of Philology, National Research Tomsk State University

Victoriya Tseller, student of the Department of General, Computer and Cognitive Linguistics, Faculty of Philology, National Research Tomsk State University

Executive Secretary

Anastasiia I. Dudareva, secretary of the department of the Department of General, Computer and Cognitive Linguistics, National Research Tomsk State University

Presentation formats:

Plenary Session Minutes:

- Presentation (40 minutes)
- Discussion (5 minutes)

Paper Session Minutes:

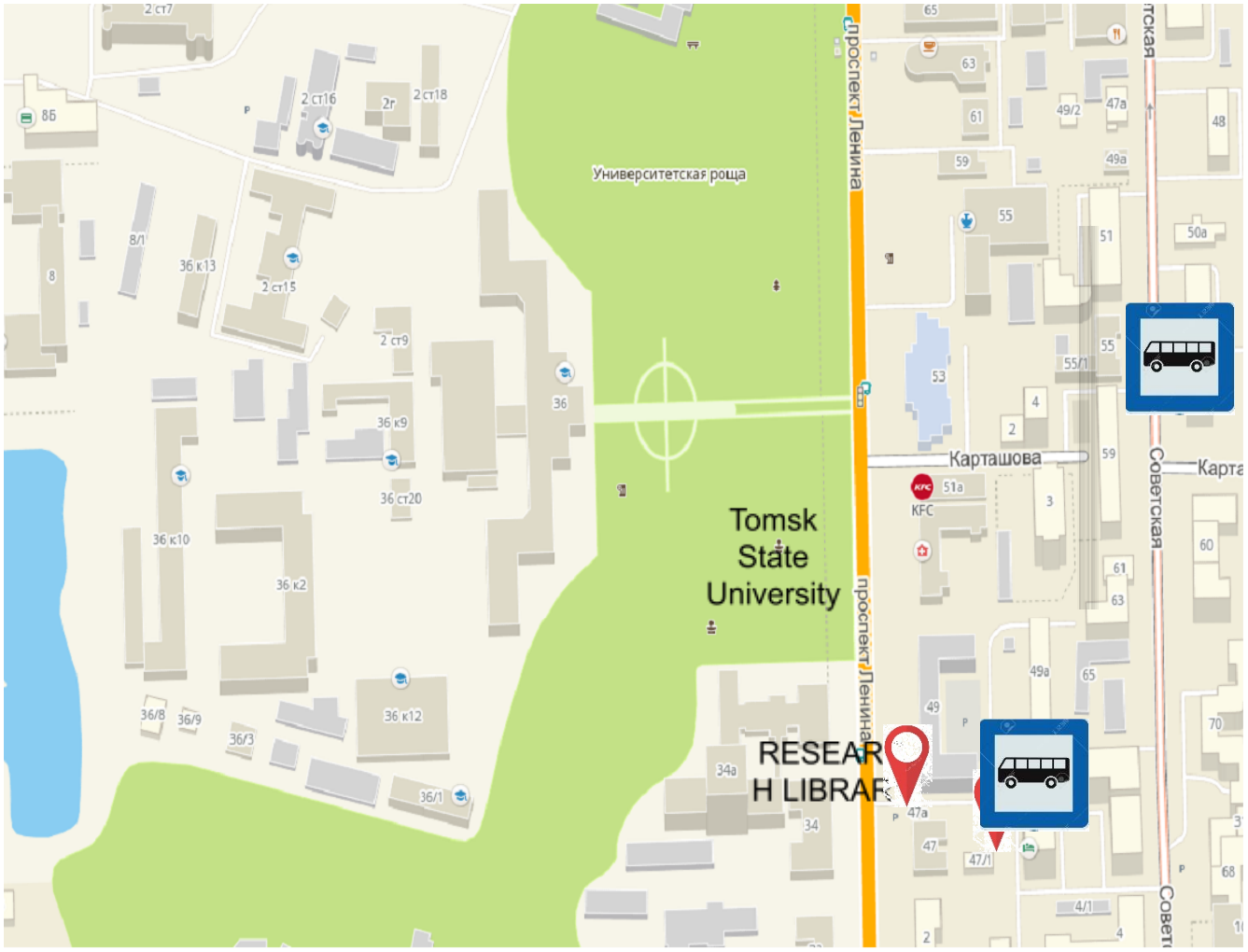
- Presentation (10 minutes)
- Discussion (5 minutes)

Working languages: Russian, English

Conference time: Tomsk (GMT+7)

Conference Venue

RESEARCH LIBRARY of Tomsk State University, Lenin Avenue, 34A, room 07



How to get there:

Bus stop Universitet

Buses: 2, 3, 4, 12, 19, 22, 23, 26, 31, 32, 33, 36, 37, 53, 112, 118, 442

Trolleybuses: 1, 3, 4

Bus stop Biblioteka TGU/TSU Library

Buses: 2, 3, 4, 12, 19, 22, 23, 26, 31, 32, 33, 36, 37, 53, 112, 118, 442

Trolleybus: 1, 3

Bus stop Yuridicheskiy Institut/Law Institute

Buses: 8, 9, 29

Programme overview

Research Library of Tomsk State University, Room 07 /

Online (Zoom) via the following link

https://us06web.zoom.us/j/89691067563?pwd=KVPRhbaih5DM7WeXyVKeRmjqtgSbdO.1

Meeting ID: 896 9106 7563

Passcode: 121244

October, 16

9.00-9.30 Registration and greeting

Plenary session

Moderators Zoya Rezanova, Olga Nagel

9.30 – 10.15 Elena Artemenko

The bilingual brain: neuroplasticity

10.15 – 11.00 Yury Bushov

Study of brain activity during pronunciation of emotional and non-emotional words aloud and to oneself

11.00 – 11.15 Coffee break

11.15 – 12.00 Anastasiia Kolmogorova

Specificity of the mental lexicon of Tuvan-Russian bilinguals: neurosemantic characteristics of Russian and Tuvan words

12.00 – 12.45 Oksana Tsaregorodtseva

Sensory Modality Norms for L1 and L2: Heritage Khakass as L1 Compared to Russian as L2 and L1

12.45 – 14.00 Lunch

14.40 – 15.25 Diloro Iskandarova

Self-identification of a bilingual: who am I?

15.25 – 16.10 Elena Oglezneva

Russian-speaking enclaves in China at the late XX - beginning XXI centuries: experience of sociolinguistic modeling

16.10 – 16.55 Igor Itkin, Daniel Buynitskaya, Veronika Dibrova

Experience in creating a universal digital platform for conducting experiments in the field of cognitive linguistics

October, 17

Paper Session

Moderators Elena Artemenko, Oksana Tsaregorodtseva

10.30 – 10.45 Valeriia Vladimirova

Cognitive processing of case forms of Russian nouns: the influence of animacy

10.45 – 11.00 Umidjan Makhmudov

Sociolinguistic Database UzRusSocLing

11.00 – 11.15 Kseniya Ivankova

Cognitive processing of verbal stimulus by men and women: experimental study

11.15 – 11.30 Irina Korshunova

Assessment of the contribution of auditory and visual modalities to the semantics of Russian words by native Russian speakers and Tatar-Russian and Tajik-Russian bilinguals

11.30 – 11.45 Margarita Lutova

The role of priming in the perception of Russian emotional words

11.45 – 12.00 Coffee break

12.00 – 12.15 Dinara Aisheva

Adversarial attacks on Kandinsky 2.0 using macaronic prompts

12.15 – 12.30 Elena Pogodaeva

Thesaurus-guided method of representing topics in dialogical speech (using the example of the corpus of Russian oral speech of Turkic-Russian bilinguals RuTuBiC)

12.30 – 12.45 Mikhail Vlasov

Knowledge in Practice: Development of a multimodal corpus for improving methods of effective work with educational text and developing secondary school students' reading literacy

12.45 – 13.00 Andrey Stepanenko

Topic modeling of the coronavirus disease at the beginning of the pandemic based on Twitter and news texts

13.00 – 13.15 Stanislav Beletskiy

Probabilistic concepts of pandemic discourse

13.15 – 13.30 Stanislav Devi

Review on Machine Translation perception aspects from an end user perspective in current scientific publications

13.30 – 13.45 Kseniia Krapivina

Identification and qualitative and quantitative evaluation of EmEditor interface translation tactics

13.45 – 14.00 Closing Plenary Session

Buffet

Программа семинара

Научная Библиотека Томского государственного университета, 07 ауд. / онлайн по ссылке https://us06web.zoom.us/j/89691067563?pwd=KVPRhbaih5DM7WeXyVKeRmjqtgSbdO.1

Meeting ID: 896 9106 7563
Passcode: 121244

---

16 октября

9.00-9.30 Регистрация и приветствие участникам

Пленарное заседание

Модераторы Зоя Резанова, Ольга Нагель

9.30 – 10.15 Елена Артеменко

Нейропластичность и билингвальный мозг

10.15 – 11.00 Юрий Бушов

Изучение активности мозга при произнесении эмоционального и неэмоционального слов вслух и про себя.

11.00 – 11.15 Кофе-брейк

11.15 – 12.00 Анастасия Колмогорова

Специфика ментального лексикона тувинско-русских билингвов: нейросемантические характеристики русских и тувинских слов

12.00 – 12.45 Оксана Царегородцева

Нормы сенсорной модальности для L1 и L2: хакасский как L1 по сравнению с русским как L2 и L1

12.45 – 14.00 Обед

14.40 – 15.25 Дилоро Искандарова

Самоидентификация билингва: кто я?

15.25 – 16.10 Елена Оглезнева

Русскоязычные анклавы в китае в конце XX – начале XXI вв.: опыт социолингвистического моделирования

16.10 – 16.55 Игорь Иткин, Даниэль Буйницкая, Вероника Диброва

Опыт создания универсальной цифровой платформы для проведения экспериментов в области когнитивной лингвистики

17 октября

Секционные доклады

Модераторы Елена Артеменко, Оксана Царегородцева

10.30 – 10.45 Валерия Владимирова

Когнитивная обработка падежных форм русских существительных: влияние фактора одушевленности

10.45 – 11.00 Умиджан Махмудов

Социолингвистическая база данных UzRusSocLing

11.00 – 11.15 Ксения Иванкова

Когнитивная обработка вербального стимула мужчинами и женщинами: экспериментальное исследование

11.15 – 11.30 Ирина Коршунова

Оценка вклада аудиальной и визуальной модальностей в семантику русских слов носителями русского языка, как родного и татарско-русскими и таджикско-русскими билингвами

11.30 – 11.45 Маргарита Лутова

Роль прайминга в восприятии русских эмоциональных слов

11.45 – 12.00 **Кофе-брейк**

12.00 – 12.15 Динара Аишева

Применение мультиязычных вредоносных запросов для генерации изображений нейронной сетью Kandinsky 2.0

12.15 – 12.30 Елена Погодаева

Тезаурусный метод представления тем в диалогической речи (на примере корпуса русской устной речи тюрко-русских билингвов RuTuBiC)

12.30 – 12.45 Михаил Власов

Знания на практике: Разработка мультимодального корпуса для совершенствования методов эффективной работы с учебным текстом и развития читательской грамотности учащихся средних школ

12.45 – 13.00 Андрей Степаненко

Тематическое моделирование пандемии коронавирусной инфекции в период начала пандемии на материалах текстов социальной сети «Твиттер» и средств массовой информации

13.00 – 13.15 Станислав Белецкий

Вероятностные концепты пандемийного дискурса

13.15 – 13.30 Станислав Дэви

Исследование аспектов восприятия машинного перевода со стороны конечных пользователей в актуальных научных публикациях

13.30 – 13.45 Ксения Крапивина

Идентификация и качественная и количественная оценка переводческих тактик в тексте интерфейса EmEditor

13.45 – 14.00 Заключительное пленарное заседание

Фуршет

Abstracts report

Plenary reports

The bilingual brain: neuroplasticity

Elena Artemenko

HSE University, Saint-Petersburg, Russia
National Research Tomsk State University, Tomsk, Russia
nekrasovaed@gmail.com

It is shown that bilingualism requires additional effort from the cognitive system, which may lead to certain changes in the bilingual brain. The need to store and utilise two representational systems corresponding to different languages, and to develop an effective control system that allows the use of language according to context, may lead to functional and structural changes in the brain. Neuroimaging studies show that the neural organisation of language representations in the bilingual brain depends on their type. The paper will discuss the peculiarities of brain structural organisation in processing conceptual, lexical and syntactic level units.

Study of brain activity during pronunciation of emotional and non-emotional words aloud and to oneself

Yury Bushov

National Research Tomsk State University, Tomsk, Russia
bushov@bio.tsu.ru

The study of the mechanisms of thinking is an urgent problem of modern neuroscience. At the same time, the question of the participation of various brain structures and their role in the processes of thinking remains largely open.
The purpose of this study was to study the activity of brain structures during the pronunciation of emotional and non-emotional words aloud and to oneself.

Volunteers participated in the research: 20 men and 20 women - university students aged 18-27 years. During the fMRI scan of the brain, the participsnts uttered the unemotional word "One" and the emotional word "Pain" first aloud, then silently. The results of functional MRI were obtained on a SIEMENS Magnetom Verio 3 Tesla tomograph at the National Research Center "Kurchatov Institute" (Moscow).

It has been established that the pronunciation of emotional and non-emotional words to oneself is accompanied by the activation of only a part of the brain structures that are activated when the same words are pronounced aloud.

Interhemispheric and gender differences were found in the activation of "speech" structures during the pronunciation of emotional and non-emotional words. Gender differences manifested themselves in different levels of activation of the same "speech" structures in men and women. Moreover, most often, these structures are more strongly activated in men.

The obtained results testify to the participation in the provision of sound speech of the basic neural networks: the central executive control network (CEN); operational rest networks (DMN); networks of selection of a relevant stimulus (SN). The Central Executive Control Network (CEN) is partly involved in providing internal speech.

The results of this study are important for understanding the role of various brain structures in the provision of inner speech and verbal thinking.

This study was financially supported by the Russian Foundation for Basic Research (project no. 18-013-00758) and supported by the Ministry of Science and Higher Education of the Russian Federation (project no. 075-015-2020-801).

Self-identification of a bilingual: who am I?

Diloro Iskandarova

Russian-Tajik (Slavic) University

galdobr@peterlink.ru

Historically, on the territory of modern Tajikistan, several areas of bilingualism have developed: Tajik-Turkic (Uzbek, Kyrgyz), Tajik-Shugnan (or other Pamir languages), and in Soviet period, Tajik-Russian. Moreover, with the spread of

spheres of functioning of the Russian language in traditional bilingual areas, multilingualism appeared. Thus, in the north of Tajikistan (in the Sughd region), the majority of the population speaks three languages - Tajik, Uzbek and Russian; in the Gorno-Badakhshan Autonomous Region - local (Shugnan, Rushan, Wakhan, etc.), Tajik and Russian (sometimes English).

In conditions of widespread bilingualism/multilingualism, the question of self-identification is actual. It is known that a person often identifies his native language by nationality, but sometimes the question arises: I equally use two/three languages in different areas, which language can I call my native language? In the Tajik language, the native language is "zaboni modari", i.e. the mother's language, and this somewhat simplifies the answer to the question about the native (maternal) language (although nationality in documents is recorded according to the father). In sociolinguistic literature, there are several terms: native/non-native, first/second, dominant language, etc. But since there is no uniformity in the use of these terms, the question of self-identification of a bilingual/multilingual who is a native speaker of several languages (cultures?), although one of them is dominant (is it always?), remains open.

The basis for the research is experimental methods: sociolinguistic survey and associative experiment (with stimulus words "friend or foe" and ethnonym stimuli). The results obtained allow us to identify the issue of the influence of self-identification of a bilingual/multilingual through language (culture) on possible social tension in a multilingual society, both in the native country and in migration, since one of the acute problems of the modern world is the increasing intensity of migration, which often causes problems social tension and rejection.

The study can also serve as a basis for monitoring ethnic, religious and linguistic tolerance among young people.

Experience in creating a universal digital platform for conducting experiments in the field of cognitive linguistics

Igor Itkin[1], Daniel Buynitskaya*[1,2], Veronika Dibrova[1,2]
[1]Intelligent Profit Solutions Tomsk (IPST)
[2]Tomsk State University, Russia
*correspondence author: buynitskajadn@gmail.com

The report is devoted to the description of a universal digital platform that integrates the functions of preparing, processing, storing and visualizing data for cognitive linguistics tasks.

When working with platforms intended for research in the field of cognitive linguistics, neurobiology, etc., it is difficult for the user to understand the large functionality and structure of the program, and there is often a need to write code in programming languages (mainly Python). As a result, the creation process experiments become much more complicated. Examples of such programs are PsychoPy, E-prime and OpenSesame.

The digital platform features an intuitive experiment designer that allows you to create experiments that include stimulus and response elements, eliminating the need for Python code.

The objective is to present a universal digital platform that includes a user-friendly interface for creating experiments, setting up its parameters, uploading and processing data, aimed at creating a unified experimental base in the field of humanities, as well as automating the process of processing and visualizing the results.

The following stack is used for software development: NestJS, TS, PostgreSQL, MongoDB, Swagger, TypeORM, Mongoose, JWT.

For data storage purposes, the PostgreSQL database management system (DBMS) is utilized, owing to its numerous advantages, including support for databases of unlimited size; powerful and reliable transaction and replication mechanisms; extensible system of built-in programming languages; inheritance; easy expandability.

Currently, the collection of labeled methodological data from experiments has been optimized by improving databases and information collection logic, as well as increasing the number of respondents and reducing the amount of time for creating experiments, collecting and analyzing data from employees of the Laboratory of Linguistic Anthropology at TSU.

Specificity of the mental lexicon of Tuvan-Russian bilinguals: neurosemantic characteristics of Russian and Tuvan words

Anastasia Kolmogorova

Siberian Federal university

nastiakol@mail.ru

The purpose of the study is to describe the neurosemantic characteristics of 73 nouns used in the oral speech of the Tuvans: 55 are words of the Tuvan language, 18 are from Russian. We obtained the word list in a psycholinguistic experiment with 150 Tuvan-Russian bilinguals whose average age was 22.6 years. Five groups of noun stimuli were identified: culturally significant words (names of rituals and their elements, natural objects with a high symbolic load in the traditional worldview, kinship terms); zoonyms; names of household objects in Tuvan, names of household objects in Russian, and vocabulary of the everyday. The research problem lies in the field of joint inquiries of cognitive linguistics, neurosemantics and aphasiology. It is a case of searching for a neural correlate that is called the meaning of a word in traditional linguistics. According to F. Pulvermüller's neurosemantic conception, such a correlate is a distributed structure of neurons. In being activated, they allow a person to get a "feeling" of a word and provide understanding. In order to conduct the experiment, we used a methodology already well-tested in perceptual semantics to identify such psycholinguistic properties of words as subjective frequency, age of acquisition, imageability, familiarity with a concept. The novelty of the work is that for the first time using the Tuvan language material we have found out which groups of words different psycholinguistic properties are essential for and which neurosemantic characteristics of words they correlate with (sensorimotor experience, visual perception, experience of engagement with a word).We hypothesize about factors affecting the identified characteristics of words and specify words that have neural correlates with the strongest connections. It has been found out that in case of culturally significant words the property "familiarity with a concept" turns out to be relatively high in value and the strongest in influence on other properties; in case of zoonyms it is "imageability"; in case of everyday vocabulary it is "subjective frequency"

which strongly correlates with the "age of acquisition". It is "familiarity with a concept" in case of names of household items in Tuvan. The subjective frequency has the greatest influence on other properties in Russian.The reflection of the informants has shown that the household sphere is most affected by the competition between words of the native and non-native languages. Those Tuvan words which are most deeply influenced by this competition have the lowest values of psycholinguistic properties.

Russian-speaking enclaves in China at the late XX - beginning XXI centuries: experience of sociolinguistic modeling

Elena Oglezneva

Tomsk State University of Architecture and Civil Engineering (Russia)
eoglezneva@yandex.ru

The research was carried out at the intersection of sociolinguistics, language contactology, the theory of interlingual interaction and linguistic personology.

The focus of the study is on the peculiarities of the functioning of the Russian language in conditions of language contact. Language contact should be considered as a normal state of language in modern conditions, characterized by increased migration flows and multinationality of almost any territory.

The Russian language and its dialects on the territory of Russia interact with the languages and dialects of other peoples inhabiting the Russian Federation (internal contactology). Thus, in the territory of the Russian Far East since the end of the 19th century. There was an interdialectal interaction between the Russian, Ukrainian and Belarusian languages with the subsequent Russification of the linguistic landscape.

Finding itself on the territory of other states as the language of the national diaspora, the Russian language inevitably comes into contact with the languages of the population of these territories (external contactology). Under these conditions, the problem of the vitality of the Russian language in a foreign language environment arises. Thus, the Russian language and its dialects, functioning in different territories of neighboring China, demonstrate

different ways of their existence in a foreign language, foreign national and foreign cultural environment.
Since 2000, several Russian-speaking enclaves in China have been examined: in the city of Harbin, where at the beginning of the 21st century. the last representatives of the Russian post-revolutionary emigration to the eastern abroad were still living; on the right bank of the Amur, in Heilongjiang province, bordering Russia; in the Three Rivers (Sanhe) region in Inner Mongolia, bordering Transbaikalia; in the Xinjiang Uyghur Autonomous Region - in the North-West of the People's Republic of China.
The main research method was the method of speech portraiture of typical linguistic personalities, which illustrated various ways of Russian language existence in a foreign language environment. The degree of vitality of the Russian language in the indicated territories of China was different, and the search for factors that determine different degrees of vitality of the language in seemingly similar conditions is a scientific problem.
The relevance of the undertaken research is related to the need for an exhaustive study of the various variants of the Russian national language and their specificity in a situation of language contact and under the influence of a range of sociolinguistic and linguistic factors themselves.

Sensory Modality Norms for L1 and L2: Heritage Khakass as L1 Compared to Russian as L2 and L1

Oksana Tsaregorodtseva

University of Tübingen, Germany
Tomsk State University, Russia
caregrad@yandex.ru

Research on language and the senses requires access to vast collections of linguistic stimuli and corresponding sensory measurements. While progress has been made in expanding sensory norms across different languages, data on lesser-known languages is still lacking. This study aimed to address this gap by acquiring sensory modality norms for the Khakass language and exploring the role of norming studies in understanding perceptual information in bilingual individuals. Our objectives were to compare sensory modality norms in Khakass (L1) with dominant Russian (L2) and native Russian (L1) to

uncover the impact of heritage and dominant languages on sensory representations. By analyzing ratings and the relationships between modalities, we aimed to gain insights into how individuals' sensory experiences are shaped by their heritage and dominant languages.

Paper Session

Probabilistic concepts of pandemic discourse

Stanislav Beletskiy
National Research University Higher School of Economics (Moscow)
stas.beletskiy@gmail.com

The purpose of this study is to describe the processes of meaning-making of pandemic discourse based on the material of the corpus-collection of students' interviews about life in conditions of self-isolation in the spring of 2020. We measure and describe these relationships using the statistical content analysis program KHCoder, using the functions of collecting associations (probabilities of occurrence of words) and compiling semantic maps. The results of the work consist of description of the center, the near and far periphery of the nodal points of the pandemic discourse, such as terms like covid, corona, self-isolation, quarantine, coronavirus, mask, gloves, remotely, zoom.

Cognitive processing of case forms of Russian nouns: the influence of animacy

Valeriia Vladimirova
Tomsk State University
picture_perfect@mail.ru

The report presents the results of a study dedicated to the cognitive processing of the case system by Turkic-Russian bilinguals (Tatar L1 and Uzbek L1). We hypothesize that bilinguals are less sensitive to Differential Language Sensitivity (DLS) recorded in the corpus than to errors that are atypical for them. This difference is expected to manifest in variations in the speed of processing target words in constructed sentences. To test this hypothesis, an oculographic experiment was designed, incorporating three independent factors: sentence type, DLS type (with three levels: 1) grammatically correct sentence, 2) sentence with registered DLS, 3) sentence with atypical deviation), and the respondent's native language (with two levels: Russian or Turkic). We analyzed oculographic data from 11 Tatar-Russian bilinguals and 6 Uzbek-Russian bilinguals. In Turkic languages, the division of nouns into

animate and inanimate categories lacks grammatical significance, unlike in grammatical gender in other languages. However, the questions asked about people and other nouns differ. This distinction could potentially affect the cognitive processing of objects in bilinguals, as the person to whom the action is directed is conceptually more important than inanimate objects. This importance holds true regardless of the correctness of using these nouns in sentences.

Knowledge in Practice: Development of a multimodal corpus for improving methods of effective work with educational text and developing secondary school students' reading literacy

Mikhail Vlasov
Shukshin Altai State University for Humanities and Pedagogy
vlasov_mikhailo@mail.ru

The research is financed by the Ministry of Education of the Russian Federation within the framework of the state assignment for the implementation of research "Development of a multimodal corpus for improving methods of effective work with educational text and developing secondary school students' reading literacy" (No. PTNI 1022040900277-2-5.3.1)

Diagnosis and development of reading literacy of schoolchildren is one of the most important tasks of modern Russian education system in the context of the development of sovereignty in various fields.

Federal educational programs of secondary school education dictate the need to include tasks on reading literacy in the educational materials. Reading literacy implies students' ability to understand and use written texts, think about them, engage in reading in order to achieve their goals, expand their knowledge and capabilities, and participate in social life (PISA 2009).

PISA 2018 results showed that, compared to 2015, on average there was a significant decline in the level of reading literacy in Russian schoolchildren (Adamovich et al., 2019). This is also due to the fact that standard reading literacy tasks have changed from extracting information from single text to problem solving involving multiple texts (e.g. search, inference and corroborate/conflict) (PISA 2018).

To develop such tasks within the framework of our project "Development of a multimodal corpus for improving methods of effective work with educational text and developing secondary school students' reading literacy" we used a bank of tasks for the formation and assessment of functional literacy of secondary school students (grades 5-9) of the Institute for Educational Development Strategy of the Russian Academy of Education (http://skiv.instrao.ru/bank-zadaniy/), as well as a published PISA 2018 Reading Literacy Framework (PISA 2018). The classification of tasks for educational texts for testing reading literacy is presented in oral report.

Review on Machine Translation perception aspects from an end user perspective in current scientific publications

Stanislav Devi
Tomsk State University
projectheaven@mail.ru

Linguistic technologies application is a rising trend in terms of delivering a workable end-user experience. Such technologies involve machine and automated translation programs, online dictionaries and glossaries, predictive typing systems and voice assistants, etc. Machine translation (MT) systems are one of the most important tools in the work of a modern translator, as they allow translation of large text volumes in relatively short periods of time. MT systems are also widely used in a regular social environment to translate various materials for study, work and everyday life, since the level of MT design provides both quality and usability. Studying MT application aspects requires an interdisciplinary approach, which leads to differences in the object, subject and methodology in scientific research, as a result of which there is a need to systematize this research experience.
This article aims to analyze current trends in the research of end-user interaction with various machine translation systems. In order to do this, the article examines the research aspects of MT perception in the works of various authors currently dealing with this problem.

Cognitive processing of verbal stimulus by men and women: experimental study

Kseniya Ivankova
Tomsk State University
senya.gl@mail.ru

Researchers analyzing the characteristics of male and female speech note that women's speech is characterized by greater expressiveness compared to men's. Gender differences were also discovered in the processing of emotional stimuli, but the data obtained are contradictory and do not allow us to draw clear conclusions. We conducted a series of experiments using E-Prime 2 software to test the hypothesis that the processing of an expressive stimulus during different cognitive tasks will differ in time course depending on the gender of the participant. We also assumed that the response times of the participants may depend on the referent in the sentence, that is, on the gender of the character. As a result of a series of experiments, data were obtained indicating the influence of the degree of expressiveness of a stimulus on its processing, and statistical analysis also showed a high level of significance of the factor of the participants' gender on the response times. Additionally, statistical analysis revealed the significant interaction between the sentence subject and the type of expressiveness. Specifically, when an expressive stimulus was combined with a female referent, the reaction time was longer compared to a neutral stimulus. Thus, our hypothesis that the processing of expressive stimuli is different in men and women is partially confirmed, namely, if a woman acts as the referent in the sentence. To verify these results, it is necessary to conduct an additional series of experiments.

Assessment of the contribution of auditory and visual modalities to the semantics of Russian words by native Russian speakers and Tatar-Russian and Tajik-Russian bilinguals

Irina Korshunova
Tomsk State University
korshunova-61818@mail.ru

The report presents the analysis of the contribution of sensory modalities (hearing, vision, smell, taste, tactile sensations) to the formation of the perceptual component of the semantics of Russian nouns. Based on the assessments of native Russian speakers and Tatar-Russian and Tajik-Russian

bilinguals, the contribution of sensory modalities to the semantics of Russian words is described. At this stage, we received primary data from Tajik-Russian bilinguals. These data were then averaged and the results of descriptive statistics of assessments for 5 sensory modalities of 200 Russian nouns given by three groups of respondents were obtained. It is believed that the visual modality is the most informative for nouns. In our data, the average estimates of the contribution of the visual modality in the L1-R group are 3.773, in the L1-Taj group – 3.529, in the L1-Tat group – 3.112. Which in turn exceeds all average ratings for other sensory modalities. Already based on the data of descriptive statistics, we can judge a significant system of commonality in the relationship under study in the assessments received from the three groups of respondents.

Identification and qualitative and quantitative evaluation of EmEditor interface translation tactics

Kseniia Krapivina
Tomsk State University
krapivinaxena@gmail.com

Up to 2010s, software user interfaces were considered solely graphical design products. However, a verbal component of user interface for an individual desktop software application could amount to thousands of separate textual elements linked into complex semantic structure by hyperlinks and visual means, which makes it a complex creolized text. The goal of current research is to identify, qualify, quantify and compare tactics and techniques the UX-writers and translators utilize to create and adapt such creolized cohesion for EN and RU languages respectively. To achieve that, the bilingual database of software strings from EmEditor text processor was created completed with identifying and structural tag system. The differences between the EN and RU texts, such as changes in capitalization, parts of speech, word order, elaborations and truncations etc., were detected and marked up. The resulting base was subjected to exploratory statistical analysis to identify patterns and establish baseline for following studies. In this seminar, the preliminary qualitative and quantitative results will be presented.

The role of priming in the perception of Russian emotional words

Margarita Lutova
Tomsk State University
keinkk265@gmail.com

The phenomenon of implicit memory, including priming, has garnered increasing interest among researchers in recent years. Evidence suggests that depending on the type of prime, individuals can process and absorb information at different speeds. Emotionality and priming are closely linked, as emotions can influence decision-making and task completion.

Our research analyzes the cognitive processing of Russian adjectives by typical Russian speakers, aiming to explore the impact of primes on the processing of emotional words. The primary method employed in this study was experimentation. Initially, a psycholinguistic survey was conducted using the Likert scaling technique to select appropriate stimulus materials. Subsequently, a behavioral experiment was carried out using E-prime software, allowing us to record reaction speeds. Visual and auditory primes were utilized, involving a total of 48 participants.

The study results indicate a tendency for priming to influence respondents' reaction speeds, although not significantly. In the first experiment, a notable difference was observed in participants' reaction speeds to congruent and incongruent pairs of stimulus words and primes. In the second experiment involving auditory primes, such a difference was not observed, but reaction speeds varied based on the type of prime used.

We recognize the study's potential in exploring diverse factors such as respondents' age and gender, as well as utilizing various stimuli, including different parts of speech. This approach would enable us to discern the factors influencing respondents' processing of emotional words and the strength or weakening of the priming effect.

Sociolinguistic database UzRusSocLing

Umidjan Makhmudov
Tomsk State University
abduvaliumid@yandex.com

In our study, which aims to investigate the impact of Uzbek-Russian bilingualism on the perception of sensory semantics in Uzbek and Russian words, it is essential to explore how sociolinguistic factors influence the subjective language assessment of bilingual individuals.
The database (DB) includes assessments of language experience and social aspects of the use of native (maternal) Uzbek and second (acquired) Russian languages, given by 125 speakers of Uzbek-Russian bilingualism. The data was collected on the basis of questionnaires developed at the Laboratory of Linguistic Anthropology of TSU and adapted for collecting material on another type of bilingualism using the sociolinguistic questionnaire of O.A. Kazakevich and the language questionnaire of the bilingual Mariam, Blumenfeld and Kaushanskaya.
Expected results: The conducted analyses can provide insights into the sociolinguistic factors that influence bilinguals' subjective assessment of their maternal language proficiency.
This database can be used as a source of data for studying language situations in the Republic of Uzbekistan, as well as as material for studying the psycholinguistic and social aspects of bilingualism in the system of studying Russian as a foreign language.

Thesaurus-guided method of representing topics in dialogical speech (using the example of the corpus of Russian oral speech of Turkic-Russian bilinguals RuTuBiC)

Elena Pogodaeva
Tomsk State University
elenanikolaevnapogodaeva@gmail.com

The report presents the use of thesaurus description of topics in the RuTuBiC corpus of Russian oral speech of Turkic-Russian bilinguals. To date, approaches have been taken to create universal recommendations for annotating corpora of various types. However, existing recommendations

are approximate, and the final version of the annotation system depends on the purpose of creating the corpus and the characteristics of the language material. The main specificity of the analyzed text material is the polythematic nature of the dialogues and the hierarchical structure of topics with the genre-based thematic limitations of sociolinguistic interviews. The report describes an analyzed list of theme sections from two questionnaires that were used for interviewing respondents, as well as the coordinated decisions of corpus developers regarding the list of topics and thematic markup tags. The resulting hierarchical structure of thematic classes was supplemented with words and expressions from the Russian language thesaurus RuWordNet. The supplemented list of topics was analyzed to detect intersections of thematic classes, and the resulting examples were compared with examples of manual annotation of text material. The resulting terminological description of the thematic content of dialogues can be used as a list of labels for applying an unsupervised machine learning algorithm, as well as for clarifying instructions for marking the corpus.

Topic modeling of the coronavirus disease in the the beginning of the pandemic based on the Twitter and the news texts

Andrey Stepanenko
Tomsk State University
stepanenkone@mail.ru

The investigation reveals topic disparities between media and social network (Twitter) genres during the onset of the coronavirus pandemic in 2020. Topic modeling methods, including Latent Dirichlet Allocation (LDA) and BERT Topic Modelling, were employed to address this issue. The results indicate common themes across genres in 2020, such as research and pandemic-related restrictions. However, despite these commonalities, distinct topics were identified in each genre. These variances are attributed to the socially determined nature of the text and its focus, influenced by genre-specific characteristics.This study highlights significant topic differences between genres and could pave the way for further sociological research.

Our members

Aisheva Dinara, Tomsk State University
Artemenko Elena, Tomsk State University, Higher School of Economics
Beletskiy Stanislav, Higher School of Economics
Bushov Yury, Tomsk State University
Buynitskaya Daniel, Tomsk State University, Intelligent Profit Solutions Tomsk
Devi Stanislav, Tomsk State University, Intelligent Profit Solutions Tomsk
Dibrova Veronika, Tomsk State University
Iskandarova Diloro, Russian-Tajik (Slavic) University
Itkin Igor, Intelligent Profit Solutions Tomsk
Ivankova Kseniya, Tomsk State University
Kolmogorova Anastasia, Siberian Federal University
Korshunova Irina, Tomsk State University
Krapivina Kseniia , Tomsk State University
Lutova Margarita, Tomsk State University
Makhmudov Umidjan, Tomsk State University
Oglezneva Elena, Tomsk State University of Architecture and Civil Engineering
Pogodaeva Elena, Tomsk State University
Rezanova Zoya, Tomsk State University
Stepanenko Andrei, Tomsk State University
Tsaregorodtseva Oksana, Tomsk State University, University of Tübingen
Vladimirova Valeriia, Tomsk State University
Vlasov Mikhail, Shukshin Altai State University for Humanities and Pedagogy

Наши участники

Аишева Динара Арматовна, аспирант кафедры общей, компьютерной и когнитивной лингвистики, младший научный сотрудник лаборатории лингвистической антропологии ТГУ, Томск

Артеменко Елена Дмитриевна, кандидат филологических наук, сотрудник лаборатории социальной и когнитивной информатики НИУ ВШЭ, Санкт-Петербург; старший научный сотрудник лаборатории лингвистической антропологии филологического факультета НИ ТГУ, Томск

Белецкий Станислав Борисович, кандидат филологических наук, научный сотрудник Школы филологических наук НИУ ВШЭ (Москва)

Буйницкая Даниэль Николаевна, магистрант 2 года обучения по направлению «Фундаментальная и прикладная лингвистика», старший программист ООО «ИНТЕЛЛИДЖЕНТ ПРОФИТ СОЛЮШНС ТОМСК»

Бушов Юрий Валентинович, доктор биологических наук, профессор Томского государственного университета, психофизиолог

Владимирова Валерия Евгеньевна, аспирант кафедры общей, компьютерной и когнитивной лингвистики; младший научный сотрудник лаборатории лингвистической антропологии НИ ТГУ, Томск

Власов Михаил Сергеевич, кандидат филологических наук, доцент, Алтайский государственный гуманитарно-педагогический университет имени В. М. Шукшина, Бийск

Диброва Вероника Сергеевна, аспирант кафедры общей, компьютерной и когнитивной лингвистики; младший научный сотрудник лаборатории лингвистической антропологии НИ ТГУ, Томск; программист ООО «ИНТЕЛЛИДЖЕНТ ПРОФИТ СОЛЮШНС ТОМСК»

Искандарова Дилоро, доктор филологических наук, профессор, зав. лабораторией «Междисциплинарные прикладные исследования», профессор кафедры теоретического и прикладного языкознания, ведущий научный сотрудник лаборатории лингвистической антропологии ТГУ.

Дэви Станислав Николаевич, аспирант кафедры общей, компьютерной и когнитивной лингвистики

Иванкова Ксения Сергеевна, аспирант кафедры общей, компьютерной и когнитивной лингвистики

Иткин Игорь Иосифович, кандидат физико-математических наук; ООО «ИНТЕЛЛИДЖЕНТ ПРОФИТ СОЛЮШНС ТОМСК», консультант по ИТ

Колмогорова Анастасия Владимировна, доктор филологических наук, профессор, заведующая кафедрой романских языков и прикладной лингвистики ИФ и ЯК Сибирского федерального университета, Красноярск

Коршунова Ирина Сергеевна, аспирант кафедры общей, компьютерной и когнитивной лингвистики; младший научный сотрудник лаборатории лингвистической антропологии НИ ТГУ, Томск

Лутова Маргарита Артемовна, магистрант 1 года обучения по направлению «Фундаментальная и прикладная лингвистика»

Махмудов Умиджан Реимбаевич, аспирант кафедры общей, компьютерной и когнитивной лингвистики НИ ТГУ, Томск

Оглезнева Елена Александровна, доктор филологических наук, директор Института международных связей и интернационализации образования Томского государственного архитектурно-строительного университета, Томск

Погодаева Елена Николаевна, аспирант кафедры общей, компьютерной и когнитивной лингвистики НИ ТГУ, Томск

Резанова Зоя Ивановна, доктор филологических наук, профессор, заведующая кафедрой общей, компьютерной и когнитивной лингвистики, зам. зав. лаборатории лингвистической антропологии НИ ТГУ, Томск

Степаненко Андрей Александрович, младший научный сотрудник лаборатории лингвистической антропологии НИ ТГУ, Томск

Царегородцева Оксана Викторовна, кандидат филологических наук, сотрудник лаборатории 'Язык и познание' Тюбингенского университета, старший научный сотрудник лаборатории лингвистической антропологии НИ ТГУ, Томск