

ISSN 2500-2902
URAL-ALTAIC STUDIES
УРАЛО-АЛТАЙСКИЕ ИССЛЕДОВАНИЯ

№ 3 (46) 2022

ИНСТИТУТ ЯЗЫКОЗНАНИЯ РАН

МОСКВА

Ural-Altaic Studies
Урало-алтайские исследования

ISSN 2500-2902
ISBN 978-1-4632-0168-5

Ural-Altaic Studies

Scientific Journal

№ 3 (46) 2022

Established in 2009
Published four times a year

Moscow

© Institute of Linguistics, Russian Academy of Sciences, 2022
© Tomsk State University, 2022

ISSN 2500-2902
ISBN 978-1-4632-0168-5

Урало-алтайские исследования

научный журнал

№ 3 (46) 2022

Основан в 2009 г.
Выходит четыре раза в год

Москва

© Институт языкознания Российской академии наук, 2022
© Томский государственный университет, 2022

CONTENTS

No 3 (46) 2022

Amirkul Sh. Amanov, Natalia V. Gagarina. Macrostructural analysis of narrative texts of Turkic-speaking bilingual schoolchildren	7
Anar A. Gadzhieva. Reflection of the peculiarities the dialect of Semirechensk region in “the Tale of Kara-Mergen”	21
Alfiya M. Galieva. Phoneme distribution in a Tatar text: a quantitative approach.....	30
Nikita A. Muravyev. Northern Khanty as a language with the features of accusative and hierarchical alignment	43
Yu. V. Normanskaya, I. V. Fedotova. Phonetic and lexical innovations in Ob-Ugric dialects in the 18 th —21 st centuries: new archival and field data	63
Maria O. Cheremisinova. Comparative-attenuative polysemy in Finno-Ugric languages	75
Style sheet	96

СОДЕРЖАНИЕ

Номер 3 (46) 2022

А. Ш. Аманов, Н. В. Гагарина. Макроструктурный анализ повествовательных текстов тюркоязычных школьников-билингвов.....	7
А. А. Гаджиева. Отражение особенностей говоров жителей Семиреченской области в «Сказке о Кара-Мергене»	21
А. М. Галиева. Распределение фонем в татарском тексте: опыт количественного исследования	30
Н. А. Муравьев. Севернохантыйский как язык с чертами аккузативного и иерархического кодирования	43
Ю. В. Норманская, Ю. В. Федотова. Фонетические и лексические инновации в диалектах обско-угорских языков в XVIII—XXI вв.: новые архивные и полевые материалы.....	63
М. О. Черемисинова. Компаративно-аттенуативная полисемия в финно-угорских языках	75
Требования к оформлению статей.....	96

Главный редактор

А. В. Дыбо

(д.ф.н., чл.-корр. РАН, Институт языкознания РАН)

Заместитель главного редактора

Ю. В. Норманская

(д.ф.н., Институт языкознания РАН)

Редакционная коллегия

В. Ю. Гусев (к.ф.н., Институт языкознания РАН), П. П. Дамбуева (д.ф.н., проф., Институт языкознания РАН), Е. В. Кашкин (к.ф.н., Институт русского языка РАН), О. А. Мудрак (д.ф.н., проф., Российский государственный гуманитарный университет), С. А. Мызников (д.ф.н., проф., Институт лингвистических исследований РАН), И. Николаева (PhD, Школа восточных и африканских исследований, Великобритания), Ф. Ш. Нуриева (д.ф.н., проф., Казанский (Приволжский) федеральный университет), Г. Ц. Пюрбеев (д.ф.н., проф., Институт языкознания РАН), М. Роббеетс (PhD, Лейденский университет, Нидерланды), И. Я. Селютина (д.ф.н., проф., Институт филологии СО РАН), Р. А. Тадинова (д.ф.н., доц., Институт языкознания РАН), З. Н. Экба (к.ф.н., Институт языкознания РАН), А. Б. Шлуинский (к.ф.н., Институт языкознания РАН)

Редакционный совет

В. М. Алпатов (д.ф.н., чл.-корр. РАН, Институт языкознания РАН), А. Е. Аникин (д.ф.н., чл.-корр. РАН, Институт филологии СО РАН), Р. Г. Ахметьянов (д.ф.н., проф., Бирский государственный педагогический институт), М. Бакро-Надь (проф., Университет Сегеда, Венгрия), В. Блажек (проф., Масариков университет, Чехия), Т. М. Гарипов (д.ф.н., чл.-корр. НАН РБ, Башкирский государственный педагогический институт им. М. Акмуллы), Н. И. Егоров (д.ф.н., проф., Чувашский государственный институт гуманитарных наук), И. В. Кормушин (д.ф.н., проф., Институт языкознания РАН), И. Л. Кызласов (д.и.н., проф., Институт археологии РАН), Й. Лааксо (проф., Венский университет, Австрия), К. М. Мусаев (д.ф.н., чл.-корр. РАЕН, Институт языкознания РАН), И. А. Невская (проф., Франкфуртский университет им. Иоганна Вольфганга Гёте, Германия), Т. Ризе (проф., Венский университет, Австрия), Е. К. Скрибник (проф., Мюнхенский университет, Германия), П. А. Слепцов (д.ф.н., проф., акад. НАН РСЯ, Институт гуманитарных исследований и проблем малочисленных народов Севера СО РАН), М. Стаховски (проф., Краковский университет, Польша), Ф. Г. Хисамитдинова (д.ф.н., проф., Институт истории, языка и литературы Уфимского научного центра РАН), Л. Хонти (проф., Будапештский университет, Венгрия), Н. Н. Ширококова (д.ф.н., проф., Институт филологии СО РАН), К. Шониг (проф., Берлинский университет им. Гумбольдта, Германия), М. Эрдал (проф., Франкфуртский университет им. Иоганна Вольфганга Гёте, Германия)

Editor-in-Chief

Anna Dybo

(Institute of Linguistics, Russian Academy of Sciences)

Deputy Editor-in-Chief

Yulia Normanskaya

(Institute of Linguistics, Russian Academy of Sciences)

Editorial Board

Polina Dambueva (Institute of Linguistics, Russian Academy of Sciences), Zarema Ekba (Institute of Linguistics, Russian Academy of Sciences), Valentin Gusev (Institute of Linguistics, Russian Academy of Sciences), Egor Kashkin (Institute of the Russian language, Russian Academy of Sciences), Oleg Mudrak (Russian State University for the Humanities), Sergey Myznikov (Institute for Linguistic Studies, Russian Academy of Sciences), Irina Nikolaeva (School of Oriental and African Studies, Great Britain), Fanuza Nurieva (Kazan (Volga region) Federal University), Grigoriy Pyurbuev (Institute of Linguistics, Russian Academy of Sciences), Martine Robbeets (Leiden University, Center for Linguistics, Institute for Area Studies, the Netherlands), Irina Selyutina (Institute of Philology, the Siberian Branch of Russian Academy of Sciences), Andrey Shluinski (Institute of Linguistics, Russian Academy of Sciences), Roza Tadinova (Institute of Linguistics, Russian Academy of Sciences)

Advisory Board

Rinat Ahmet'yanov (Birk State Pedagogical Institute), Vladimir Alpatov (Institute of Linguistics, Russian Academy of Sciences), Alexandr Anikin (Institute of Philology, the Siberian Branch of Russian Academy of Sciences), Marianne Bakró-Nagy (University of Szeged, Hungary), Václav Blažek (Masaryk University, Czech Republic), Nikolay Egorov (Chuvash State Institute for the Humanities), Marcel Erdal (Goethe University Frankfurt, Germany), Talmas Garipov (M. Akmulla Bashkir State Pedagogical University), Firdaus Hisamitdinova (Institute of History, Language and Literature, Ufa Scientific Centre of Russian Academy of Sciences), László Honti (University of Budapest, Hungary), Igor Kormushin (Institute of Linguistics, Russian Academy of Sciences), Igor Kyzlasov (Institute of Archaeology, Russian Academy of Sciences), Johanna Laakso (University of Vienna, Austria), Kenesbay Musaev (Institute of Linguistics, Russian Academy of Sciences), Irina Nevskaya (Goethe University Frankfurt, Germany), Timothy Riese (University of Vienna, Austria), Klaus Schönig (Humboldt University of Berlin, Germany), Natalia Shirobokova (Institute of Philology, the Siberian Branch of Russian Academy of Sciences), Elena Skribnik (University of Munich, Germany), Piotr Slepsov (Institute of Humanities Research and Indigenous Studies of the North, the Siberian Branch of Russian Academy of Sciences), Marek Stachowski (Krakow University, Poland)

Макроструктурный анализ повествовательных текстов тюркоязычных школьников-билингвов

Аманов Амиркул Шаукатович, Казахский национальный университет имени Аль-Фараби (Алматы); amir_frl@mail.ru
Гагарина Наталья Владимировна, Берлинский университет им. Гумбольдта, Уппсальский университет;
gagarina.natalia@googlegmail.com

Статья посвящена анализу макроструктуры устных повествовательных текстов тюркоязычных учащихся-билингвов. Материалом исследования послужили повествовательные тексты, полученные с помощью ТОН — Теста на Определение Нарративных навыков многоязычных детей — инструмента для оценки повествовательных способностей билингвов. Для анализа были выбраны наиболее распространенные языковые группы среди детей в Казахстане. Результаты исследования существенно отличаются не только соотношением количественно-качественных показателей макроструктурной сложности рассказов билингвов на двух языках, но и сравнительным описанием особенностей близкородственных языков, их связей, сходств и различий в условиях дву- и многоязычия.

Дети начинают создавать повествования примерно в два года, и с этого момента их повествовательная компетентность интенсивно развивается, но не у всех детей с одинаковой скоростью. Более пристальный взгляд на различные макроструктурные компоненты показал, что производство было поразительно схожим в трех языковых группах — как для часто используемых, так и для редко встречающихся компонентов макроструктуры. Тем не менее становится ясно, что при проведении в республике научных исследований по вопросам билингвизма необходимо учитывать исторически сложившиеся на огромной территории страны региональные социально-экономические, демографические и этноязыковые условия.

Ключевые слова: билингвы, макроструктура историй, близкородственные языки, нарративные навыки, рассказ, пересказ

MACROSTRUCTURAL ANALYSIS OF NARRATIVE TEXTS OF TURKIC-SPEAKING BILINGUAL SCHOOLCHILDREN

Amirkul Sh. Amanov, Al-Farabi Kazakh National University (Almaty); amir_frl@mail.ru
Natalia V. Gagarina, Humboldt University of Berlin, Uppsala University; gagarina.natalia@googlegmail.com

The present study is devoted to the analysis of the macrostructure of oral narrative texts of Turkic-speaking bilingual schoolchildren. The research material was obtained using a MAIN — Multilingual Assessment Instrument for Narratives — a tool for assessing the narrative abilities of bilinguals. Subjects for the analysis represent the most common language groups among children in Kazakhstan. The results of the study differ significantly from previous findings regarding the ratio of qualitative and quantitative indicators of the macrostructural complexity of bilingual stories in two languages, as well as in the comparative description of the features of closely related languages, their connections, similarities and differences in the conditions of bilingualism and multilingualism.

Observations show that narrative competence develops depending on the subject's age and propensity for storytelling and becomes fully structured only in the adulthood. Children begin to create narratives at about two years old, and from that moment their narrative competence develops intensively, although at different rates speed. A closer look at the various macrostructural components showed that the production was strikingly similar in the three language groups, both for frequently used and rarely encountered components of the macrostructure. The complexity of the plot was also considered from the other side — by estimating the number of episodes in which participants implemented complete GAO sequences. In all three language groups, on both L1 and L2, most speakers used at least one complete sequence. Nevertheless, it becomes clear that when conducting scientific research on bilingualism in the republic, it is necessary to take into account the regional socio-economic, demographic and ethno-linguistic conditions that have historically developed on the vast territory of the country.

Keywords: bilinguals, macrostructure of stories, closely related languages, narrative skills, story, retelling

Авторы выражают глубокую признательность Элеоноре Дюсеновне Сулейменовой и Аимгуль Каирбековне Казкеновой — за ценные советы при планировании исследования и рекомендации по оформлению статьи.

1. Введение

Способность к повествованию является одной из основных форм использования языка его носителями. Повествовательные навыки личности важны для ее языкового и когнитивного развития, поскольку они обеспечивают соответствующие условия для организации устной и письменной информации. Последние исследования повествовательных навыков подтвердили, что определение способности к повествованию у обучающихся помогает предсказать дальнейшее развитие их навыков чтения, письма и успех в школе в целом [Lindgren 2018: 12].

Повествовательные навыки также важны для социализации личности: хорошие рассказчики привлекают больше внимания и получают больше возможностей практиковать язык в социуме, и, соответственно, они становятся успешными. Кроме того, исследование этих навыков в раннем возрасте представляет большой интерес в клинических условиях, так как помогает выявить нарушения в общении и в использовании других функций языка.

Повествовательная компетентность развивается в зависимости от возраста и склонности к рассказыванию историй и становится полностью структурированной только у взрослого человека. Дети начинают создавать повествования примерно в два года, и с этого момента их нарративные навыки интенсивно растут. Но она развивается не у всех детей с одинаковой скоростью. Учитывая важность ранней нарративной компетентности для последующего становления личности, с одной стороны, и вариативности в развитии, с другой стороны, представляется крайне важным исследовать уровень владения повествовательными навыками у детей в младшем школьном возрасте [Pinto et al. 2018: 4].

Необходимо отметить, что понимание высказываний предшествует развитию общей языковой компетенции, и исследователи в основном изучают понимание лексики и грамматики, следя за формированием соответствующих навыков на протяжении всего детства. Гораздо меньше внимания уделялось пониманию повествований, например рассказов, созданных с помощью изобразительных и/или устных стимулов. Только недавно лингвисты стали проявлять интерес к этому аспекту [Bohnacker et al. 2016: 20].

Развитие нарративных навыков следует за усвоением базовых лексических, синтаксических и морфологических категорий языка [Berman, Slobin 1994: 5]. Следовательно, производство повествования можно оценить начиная со времени активного развития повествования, т. е. примерно с четырех лет; оценка же навыков повествования с помощью заданий на понимание возможна у детей более младшего возраста, которые все еще имеют довольно ограниченные продуктивные способности. Это относится не только к одноязычным, но и к двуязычным детям, которые могут овладевать своими языками в разном возрасте [там же: 8].

Результаты исследований с использованием данных, собранных у двуязычных детей, показывают, что двуязычные дети осваивают макроструктуру таким же образом, как и одноязычные дети. Более того, двуязычные дети, которые последовательно овладевают двумя языками, могут до некоторой степени сохранять свои навыки повествования, переходя на другой язык: «...последовательные двуязычные быстрее усваивают структуру истории, потому что это навык когнитивно-лингвистического уровня, которым можно поделиться с первым языком» [Paradis, Schneider 2008].

Тем не менее необходимо отметить, что при изучении нарративной компетенции у билингвов нужно учитывать ряд дополнительных аспектов по сравнению с монолингвами. Так, ясно одно: анализ повествовательных навыков билингвов — сложный процесс, требующий тщательного планирования. Просить испытуемых рассказать историю — самый распространенный метод для оценки повествовательной компетентности. Кроме этого, для анализа нарративных навыков в мировой практике используются различные задания, в которых преобладают два типа: рассказывание историй и пересказ историй.

Немаловажным является и то, что исследования повествовательных навыков и навыков понимания чаще всего касаются письменных текстов. Гораздо меньше известно о понимании устных повествований.

В настоящем исследовании мы планируем сосредоточиться на анализе макроструктуры устных повествовательных текстов. Материалом исследования стали устные повествовательные тексты, полученные с помощью ТОН — Теста на Определение Нарративных навыков многоязычных детей — инструмента для оценки повествовательных способностей билингвов. Для анализа были выбраны казахско-русская, узбекско-русская и уйгурско-русская языковые группы, поскольку они являются наиболее распространенными среди детей в Казахстане и поэтому представляют большой научный интерес. Выбор данных языков обусловлен и тем, что все они относятся к одной языковой группе — тюркской. Это дает возможность изучить особенности функционирования близкородственных языков, их связей, сходства и различий в условиях дву- и многоязычия.

Статья состоит из шести разделов: в первом разделе обсуждаются инструмент и процедура исследования; во втором разделе описываются участники исследования; в третьем обосновываются критерии

оценки структуры истории; в четвертом проводится количественно-качественный анализ и межъязыковые сравнения историй; в разделе «Обсуждение и заключение» подводятся итоги, определяются перспективы и новые задачи исследования.

2. Инструмент и процедура исследования

Устные повествовательные тексты были получены с помощью ТОН — Теста на Определение Нарративных навыков многоязычных детей — инструмента для оценки повествовательных способностей билингвов. Тест позволяет оценить уровень экспрессивной речи и уровень понимания смысла рассказа с помощью «рассказывания» (создания истории) или «пересказа» и включает набор вопросов для понимания, ориентированных на цели и описание психического состояния. ТОН предусматривает 3 способа создания рассказов: составление рассказа по картинкам, пересказ и составление рассказа по образцу. Тест содержит четыре параллельных истории, каждая из которых тщательно разработана и состоит из шести последовательных картинок. Сюжеты историй были адаптированы для детей из разных этнокультурных сред [Gagarina, Bonhacker, Lindgren 2019: 192].

В настоящем исследовании истории «Козлята» и «Птенцы» использовались для рассказывания историй, а истории «Кошка» и «Собака» — для пересказа. Этому предшествовал этап адаптации тестовых материалов — их перевод на 3 языка (казахский, узбекский и уйгурский). Полный набор файлов для проведения исследования: описание процедуры проведения теста, протоколы тестов и листы ответов, протоколы подсчета очков, анкета для родителей и тексты историй — доступны на 86 языках на официальном сайте Центра общего языкознания Ассоциации Лейбница¹.

В каждой языковой группе эксперименты организовали учителя начальных классов. Тесты проводились в одноязычном режиме с интервалом между языками не менее 4 дней. Порядок проведения, представления историй (пересказ/рассказывание) и языки эксперимента были сбалансированы в соответствии с форматом ТОН.

В начале эксперимента задается набор вопросов для разминки (например, «Расскажите мне о своем друге», «Какие игрушки вам нравятся?», «Какая ваша любимая книга или фильм?» и т. д.). После разминки экспериментатор кладет на стол три цветных конверта и просит ребенка выбрать один: «В каждом конверте есть интересная история. Выбери один, а потом я расскажу тебе эту историю. Ребенок выбирает конверт и достает набор из шести картинок в форме буклета. Экспериментатор разворачивает картинки так, чтобы вся последовательность была видна только ребенку. После того как ребенок просмотрел всю последовательность картинок, экспериментатор спрашивает: «Ты готов? Я расскажу тебе одну историю, а затем задам несколько вопросов». После рассказа истории задаются 10 вопросов для понимания. После первого рассказа экспериментатор переходит ко второй части тестирования. На столе лежат еще три цветных конверта, и исследователь просит ребенка выбрать один: «В каждом конверте отдельная история. Выбери один и расскажи мне одну историю по картинкам». Экспериментатор при необходимости помогает ребенку развернуть буклет. После того как ребенок закончит свой рассказ, экспериментатор задает 10 вопросов на понимание.

По окончании эксперимента тексты, записанные на диктофоны (речь каждого говорящего записывалась отдельно), были затранскрибированы специалистами-носителями соответствующих языков в соответствии с форматом ТОН.

Были проанализированы компоненты макроструктуры истории, описания психологического состояния и ответы на вопросы на понимание.

В соответствии с протоколом оценки производства макроструктурных элементов участникам начисляются баллы за то, что они включают в свои повествования специальные установки (время и место), цели действующих персонажей историй, попытки и результаты, а также внутренние состояния как иницирующие события (т. е. формирующие отправную точку эпизода) и как реакции (т. е. эмоциональные реакции персонажей на исход эпизода) для каждого из трех эпизодов. Помимо этого, испытуемые получали баллы за каждый правильный ответ на понимание текста.

Оценка макроструктуры в ТОН многомерна: ее можно оценить количественно, т. е. через сумму компонентов эпизода во всем тексте (структура сюжета), и качественно через комбинацию компонентов эпизода (сложность сюжета). При суммарной оценке структуры сюжета учитывается наличие следующих компонентов: информация о начале истории (максимум 2 балла: 1 за время, 1 за место) и общее количество компонентов в эпизоде — 5: цели, попытки, результаты, ОПС² в качестве иницирующего со-

¹ См. <https://www.leibniz-zas.de/de/publications/schriftenreihe/zaspil/zaspil-63/main-63-download>.

² ОПС — описание психического состояния.

бытия, и ОПС в качестве реакции. Каждая история содержит три эпизода, таким образом, максимальное количество баллов за повествование составляет 17 (2 за введение, 9 за каждую из трех целей, попыток и результатов и 6 баллов за ОПС как события или реакции).

Таблица 1. Один полный эпизод (Эпизод 1)
для ТОН истории о Козлятах³

Компонент	Козлята
Установка (введение)	однажды / когда-то... (<i>время</i>) ... в лесу / у озера / у дерева / на берегу реки (<i>место</i>)
Эпизод 1	
ОПС как начало события	Мать-коза увидела, что козленок в опасности
Цель	Она хотела спасти его
Попытка	Она бросилась в воду...
Результат	...и вытолкнул ребенка из воды
ОПС как результат	Мать-коза была счастлива

В таблице 2 отмечены 17 компонентов, показывающих максимальный балл по ТОН структуре сюжета (17 баллов). Оценка структуры истории — это *количественная оценка микроструктуры*. Наличие трех основных компонентов (цель, попытка, результат (ЦПР)) в каждом эпизоде отражает максимально возможную сложность сюжета, которая является *качественной оценкой макроструктуры* в ТОН. Поскольку в каждой истории три эпизода, ЦПР может происходить три раза за историю.

Таблица 2. Сценарий рассказа с отмеченными
макроструктурными компонентами истории о Козлятах

	Компоненты ⁴	Оценка структуры истории	Оценка сложности сюжета
Однажды коза-мать <u>увидела</u> , что ее козленок упал в воду и испугался.	Эпизод 1: ОПС как начало события	1	
Она прыгнула в воду	Эпизод 1 попытка	1	
потому что она <u>хотела спасти его</u> .	Эпизод 1 цель	1	Эпизод 1 (цель, попытка, результат (ЦПР))
Голодная лиса увидела, что мать-коза была в воде, и зарычала: «Ммм, вкусно, что я вижу около воды?!».	Эпизод 1 ОПС как начало события	1	
Мать-коза <u>вытолкнула козленка из воды</u> , но лису она не заметила.	Эпизод 1 результат	1	
Она была <i>рада</i> , что ее ребенок не утонул.	Эпизод 1 ОПС как реакция	1	
Тем временем наглая лиса прыгнула вперед	Эпизод 2 попытка	1	Эпизод 2 (цель, попытка, результат (ЦПР))
потому что она <u>хотела поймать</u> другого козленка.	Эпизод 2 цель	1	
Она <u>схватила козленка</u> .	Эпизод 2 результат	1	
Козленок был <i>напуган</i> .	Эпизод 2 ОПС как реакция	1	

³ Все четыре истории ТОН («Козлята», «Птенцы», «Кошка» и «Собака») имеют эквивалентные повествовательные стимульные материалы (параллельные истории). Это дает возможность сравнивать и сопоставлять их результаты.

⁴ Типы компонентов выделены следующим образом: ОПС как начало события, цели, попытки, результаты, **ОПС как реакция**.

	Компоненты	Оценка структуры истории	Оценка сложности сюжета
Храбрая птица, пролетавшая мимо, увидела, что козленок в большой опасности.	Эпизод 3 ОПС как начало события	1	
Она решила остановить лису и спасти козленка.	Эпизод 3 цель	1	Эпизод 3 (цель, попытка, результат (ЦПР))
Птица сказала лисе: «Оставь козленка в покое». А потом она полетела вниз и укусила лису за хвост.	Эпизод 3 попытка	1	
Лиса отпустила козленка, и птица прогнала его.	Эпизод 3 результат	1	
Птица была очень <i>счастлива</i> , что смогла спасти козленка, а лиса все еще была <i>голодна</i> .	Эпизод 3 ОПС как реакция Суммарный балл за структуру	1 17	

Вопросы понимания требуют определения целей, чувств и мыслей главных героев, а не просто воспоминания о содержании истории. Из десяти вопросов, составляющих набор вопросов для понимания в ТОН (D1—D10), три вопроса нацелены на цели эпизодов истории (D1, D4 и D7), шесть — на внутренние состояния как реакции на результаты в разных эпизодах и задаются парами (D2/D3, D5/D6 и D8/D9), и последний вопрос (D10) относится ко всей истории, а не к конкретному эпизоду. Несколько вопросов, касающихся внутренних состояний, а также последний вопрос включают в себя вывод и принятие точки зрения разных главных героев.

Описания психологического состояния (ОПС) включали слова, обозначающие различные внутренние состояния: термины состояния восприятия (например, *видеть*, *слышать*, *чувствовать* и *обонять*), термины физиологического состояния (например, *испытывать жажду*, *голод* и *усталость*), термины сознания (например, *бодрствовать* и *спать*), термины эмоций (например, *грустный*, *счастливый* и *злой*); ментальные глаголы (например, *хотеть*, *думать* и *знать*) и лингвистические глаголы (например, *говорить*, *звонить* и *спрашивать*).

В настоящем исследовании мы сначала проанализировали оценки структуры историй (количественный анализ). Затем более внимательно рассмотрели сложность историй (качественный анализ), сосредоточив внимание на различиях между языками.

3. Описание участников

В исследовании приняли участие тридцать двуязычных детей: в каждой языковой группе (казахско-русская (Kaz⁵), узбекско-русская (Uzb) и уйгурско-русская (Uig)) по 10 участников (5 мальчиков и 5 девочек). Наблюдается сравнительно небольшая разница в возрастном диапазоне между языковыми группами: казахско-русская языковая группа (средний возраст 98 месяцев, возрастной диапазон 7;7—9;5); узбекско-русская языковая группа (средний возраст 89 месяцев, возрастной диапазон 7;6—9;5); уйгурско-русская языковая группа (средний возраст 84 месяца, возрастной диапазон 7;4—9;3).

Все они владеют русским языком в качестве второго языка и одним из тюркских языков в качестве первого. У всех по крайней мере один родитель является носителем русского языка. Все испытуемые — учащиеся-билингвы, живущие на территории Казахстана с рождения. Учащиеся были набраны из классов с русским языком обучения, и все обучаются в третьем классе.

Необходимо отметить, что свободный выбор языка общения, обучения, воспитания закреплен в ст. 6 «О языках в Республике Казахстан» (с изменениями и дополнениями по состоянию на 10.01.2022 г.). В ней отмечается, что «каждый гражданин Республики Казахстан имеет право на пользование родным языком, на свободный выбор языка общения, воспитания, обучения и творчества. Государство заботится о создании условий для изучения и развития языков Казахстана»⁶.

В целом в РК по всем типам общеобразовательных школ в 2013/14 учебном году распределение учащихся по языкам обучения характеризовалось следующими цифрами: на казахском языке обучались 1 660 443 школьника, на русском — 808 576, на узбекском — 77 840, на уйгурском — 14 391, на таджик-

⁵ Общепринятые коды для представления названий языков (ISO 639-2): Kaz — Kazakh, Uig — Uighur; Uyghur, Uzb — Uzbek. [интернет ресурс] https://www.loc.gov/standards/iso639-2/php/code_list.php.

⁶ https://online.zakon.kz/Document/?doc_id=1008034.

ском — 3 970, на английском языке — 2 001 человек⁷. С 2016 года в Казахстане начата постепенная реализация процесса перехода на трехязычную систему образования, по которой место школ с разными языками обучения (казахским, русским, турецким, английским, таджикским, узбекским, уйгурским и др.) постепенно займут школы с трехязычным (казахским, русским и английским) предметным образованием [Сулейменова 2017: 642].

По данным МОН РК за 2021 г., в Казахстане всего имеются 7471 школа, из них с казахским языком обучения — 3916 школ, с русским — 1271 школа, 2243 школы со смешанным языком обучения⁸, 41 школа с обучением на других языках, в том числе на узбекском — 24 школы, и на уйгурском — 17 школ⁹.

Помимо этого, русский язык в РК является обязательной частью единого государственного образовательного стандарта. Функционирование русского языка осуществляется в двух различных по целям, масштабам и структуре секторах:

— русский язык в образовательных учреждениях всех типов и уровней как язык обучения;

— русский язык в образовательных учреждениях всех типов как предмет изучения: второй (неродной) язык для школ с нерусским языком обучения [Туксайтова 2016: 3].

Так, русский язык в Казахстане — это язык обучения в дошкольных учреждениях, школах с русским языком обучения, русских отделений всех факультетов всех вузов, язык обучения в бакалавриате, магистратуре и докторантуре; это обязательный язык преподавания независимо от языка обучения во всех образовательных учреждениях — детских садах, школах (обязательный учебный предмет, входящий в перечень дисциплин, включаемых в документ об образовании), колледжах, бакалавриате, магистратуре, докторантуре PhD; это язык, который большая часть населения страны использует независимо от этнической принадлежности; это язык, коммуникативное пространство которого до сих пор превышает коммуникативное пространство казахского (государственного) языка [Сулейменова 2017: 642].

Эксперименты проводились в городах Туркестан, Шымкент (жилой массив Сайрам) и Алматы в сентябре 2021 г. Все указанные города относятся к южному региону Республики Казахстан. Ярко выраженная полиэтничность данного региона¹⁰ вызывает особый интерес для лингвистического исследования. Все дети посещают средние общеобразовательные двуязычные школы с 5-дневной учебной неделей и 8-часовым учебным днем.

Перед тестированием все родители предоставили письменное согласие на участие в исследовании. Кроме того, родители заполнили анкету, в которой содержалась справочная информация о ребенке.

Таблица 3. Описание участников по возрастным группам

	казахско-русская языковая группа (Kaz)	узбекско-русская языковая группа (Uzb)	уйгурско-русская языковая группа (Uig)
общее количество (N)	(N=10)	(N=10)	(N=10)
Возрастной диапазон в месяцах (г;м)	97—115 (7;75—9;5)	92—115 (7;6—9;5)	89—112 (7;4—9;3)
Возраст в месяцах: средний	101,7	88,9	94
Родной язык (L1) казахский, узбекский, уйгурский / Русский (кол-во)	8/2	4/6	5/5
Социально-экономическое положение (СЭП): средний	35,3	34,8	31,5
Девочки / мальчики (N)	5/5	5/5	5/5

⁷ Бюро национальной статистики Агентства по стратегическому планированию и реформам Республики Казахстан; <https://stat.gov.kz>, 2022.

⁸ Школы, в которых имеются как классы с русским, так и классы с казахским языками обучения.

⁹ Бюро национальной статистики Агентства по стратегическому планированию и реформам Республики Казахстан; <https://stat.gov.kz>, 2022.

¹⁰ При изучении и описании языковой ситуации в Республике Казахстан следует учесть региональный подход. Представленность населения Республики Казахстан по этническому признаку в общем плане выглядит следующим образом: западная, восточная части республики менее полиэтничны, а северная, южная и центральная более полиэтничны. В языковом отношении ситуация в указанных регионах сложилась приблизительно следующим образом: на западе и юге Казахстана преобладает казахский язык, а на севере, востоке и в центральной части страны доминирующим является русский язык. Каждый из регионов имеет свою специфическую языковую особенность. От этих и ряда других факторов зависит языковое развитие регионов [Алтынбекова 2005].

Социально-экономическое положение семьи (низкое, среднее и высокое) определялось в соответствии с методическими рекомендациями Министерства образования и науки Республики Казахстан комбинацией трех переменных: числа книг дома (до 25, от 26 до 100, свыше 100 — это индикатор культурных запросов), уровня образования родителей (полное среднее, среднее профессиональное, высшее) и профессионального статуса родителей (с учетом индекса престижности профессии). Так, для всех языковых групп социально-экономический статус родителей варьировался от низкого до высшего среднего, включая университетских преподавателей. Средний индекс социально-экономического положения семьи (СЭП) существенно не отличался для трех групп: средний индекс СЭП казахско-русской языковой группы = 35,3; средний индекс СЭП узбекско-русской языковой группы = 34,8; средний индекс СЭП уйгурско-русской языковой группы = 31,5.

Исследования, посвященные анализу особенностей билингвизма в Казахстане, в основном сосредоточены на грамматике [например, Утегенова 2016; Казкенова, Рахилина 2019; Булатбаева, Сарсикеева, Айтжанов 2020; Сулейменова, Аканова, Аймагамбетова 2021; Суюнбаева, Шелестюк 2022 и др.]. В проведенных исследованиях охватывается только одна языковая комбинация — казахско-русская и, как правило, изучаются взрослые двуязычные носители.

В лингвистике существуют многочисленные исследования, проведенные в рамках макроструктурного анализа повествовательных текстов билингвов с применением ТОН. Так, например, в нескольких лонгитюдных тематических исследованиях детей, последовательных билингвов, изучалось развитие обоих языков (например, [Bernardini, Schlyter 2004; Granfeldt 2000]). Н. Гагарина, С. Фичман, Е. Галкина и др. провели масштабное исследование с участием 162 русских, финских, немецких, израильских, норвежских и шведских детей-билингвов. В исследовании рассматривались макроструктура повествовательных текстов детей-билингвов, говорящих на русском языке как на родном (L1) и подверженных воздействию разных языков (L2) [Gagarina et al. 2021: 54]. Помимо этого, Дж. Линдгрэн и У. Бонакер провели эксперимент с участием 46 немецко-шведских двуязычных детей в возрасте от 4 до 6 лет. В исследовании рассматривали широкий круг вопросов, касающихся макроструктуры рассказов билингвов [Lindgren, Bohnacker 2020] и др.

Тем не менее, до сих пор не были исследованы региональные особенности билингвов, носителей русского и казахского/узбекского/уйгурского языков Центральной Азии, в том числе в Казахстане. Так, в нашем исследовании делается попытка сопоставить взаимодействие русского и *близкородственных* языков в речи учащихся-билингвов, проживающих с рождения в Казахстане.

4. Оценка структуры истории

В таблице 4 показаны оценки структуры рассказа для трех языковых групп. Баллы по трем группам относительно одинаковы (в среднем 11 или 12 баллов). Хотя по каждому языку есть участники, набравшие близко к максимальному баллу (17 баллов), далеко не у всех участников это так. Средние баллы составляют всего 65—70% от максимального, а некоторые дети набрали относительно низкие баллы.

Таблица 4. Оценка структуры рассказа по языковым группам (максимум 17 баллов)

Три группы двуязычных детей	Kaz	Uzb	Uig
(N=30)	(N=10)	(N=10)	(N=10)
Средний балл по L1 ¹¹ за пересказ	11,2	12,2	10,5
Средний балл по L1 за рассказ	9,6	12,6	9,8
Средний балл по L2 ¹² за пересказ	10,5	10,7	9,6
Средний балл по L2 за рассказ	9,5	9,7	9,8

По графику видно, что участники трех языковых групп набрали приблизительно одинаковые баллы по L1. Видно, что лучшие показатели и за рассказ, и за пересказ историй по L1 у узбекско-русской языковой группы, а относительно низкие — у уйгурско-русской.

Рисунок 2 показывает, что результаты участников на русском языке немного отличаются от результатов на L1. Дети в первых двух языковых группах (Uzb и Kaz) показывают лучшие средние результаты за пересказ историй: Uzb группа — 10,7 баллов, Kaz группа — 10,5 баллов, в то время как дети из группы Uig набрали всего 9,6 баллов. Однако последние набрали максимальный балл за рассказ — 9,8 (Uzb — 9,7; Kaz — 9,5).

¹¹ L1— казахский/узбекский/уйгурский.

¹² L2 — русский язык.

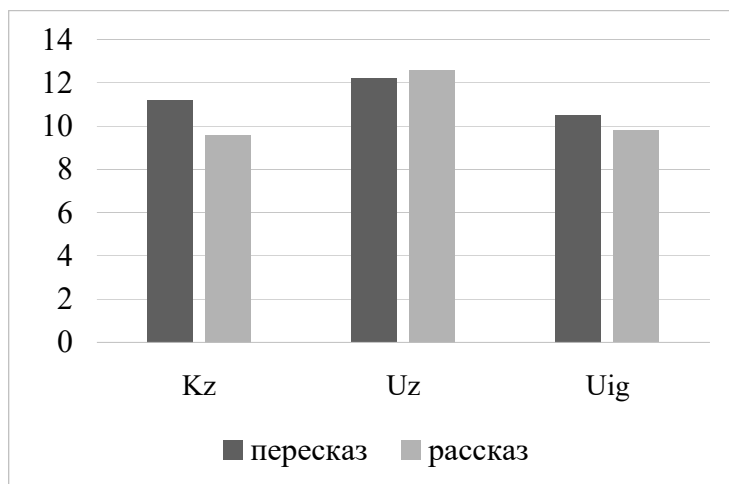


Рис. 1. Оценка структуры рассказа по языковым группам на L1

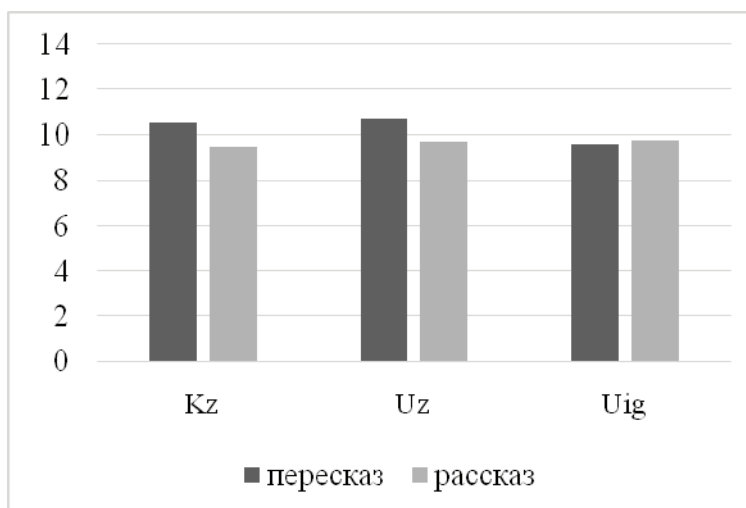


Рис. 2. Оценка структуры рассказа по языковым группам на L2 (русский язык)

Как отмечает Н. В. Гагарина, «нормальный, типичный ребенок производит в среднем от 11 до 12 компонентов на рассказ, а не все 17 компонентов (как в идеализированном сценарии), хотя некоторые дети могут выразить 15 или 16 компонентов» [Gagarina, Bonhacker, Lindgren 2019]. Это важный результат, поскольку мы говорим здесь о цели, т. е. чего мы ожидаем от детей в идеальном случае, этот вывод следует учитывать при оценке успеваемости детей, которые все еще развивают свои навыки повествования.

Важно подчеркнуть, что все дети были набраны из классов с русским языком обучения, но, несмотря на это, в языковых группах Kaz и Uzb баллы за пересказ на русском оказались заметно выше, чем баллы за рассказ. В этой связи необходимо вспомнить, что повествовательная компетентность развивается не у всех детей с одинаковой скоростью. Помимо индивидуальных особенностей развития нарративных навыков у учащихся, на наш взгляд, появлению такой разницы могло способствовать и то, что психологически детям легче пересказывать услышанную историю, нежели выдумывать собственный рассказ по наглядным стимулам (последовательностям картинок). Эта гипотеза подтверждается и тем, что большинство участников часто использовали полные эпизоды (цель, попытка и результат) и элементы описания психологического состояния персонажа именно в пересказах, тем самым достигая максимальных баллов.

Помимо этого, результаты участников языковых групп Kaz и Uzb на русском языке были относительно ниже, чем на родном языке. Мы полагаем, что это является результатом влияния языковой среды. Эксперименты проводились в разных городах, и, соответственно, языковые ситуации тоже отличались. Так, дети из Kaz и Uzb групп были набраны из русских классов общеобразовательных школ городов Шымкент и Туркестан, где основными языками общения являются казахский и узбекский. В то же время Uig языковая группа показывает приблизительно одинаковые результаты на обоих языках, что тоже можно связать с расположением школ. Участники этой группы были набраны из русских классов

общеобразовательных школ города Алматы. В Алматы языковая ситуация несколько иная: благоприятная русскоязычная языковая среда и большое количество обучающихся в школах на русском языке позволяют практиковать язык не только в стенах школы, но и за ее пределами. И в то же время наибольшее количество уйгуров в Казахстане проживает именно в этом городе¹³, что позволяет сохранять традиции родного языка.

5. Частотность макроструктурных компонентов и сложность историй

Далее мы более подробно рассмотрели, как часто участники включали в свою речь различные компоненты. Можно заметить некую общность в наличии компонентов структуры рассказов. Так, например, во всех рассказах следующие компоненты часто пропускались: описания психологического состояния (А6 — ОПС как реакция) и начало событий эпизодов 2, 3 (А7 и А12 — ОПС как начало события). Эти компоненты упоминались только при пересказах, а в рассказах и вовсе отсутствовали.

Более пристальный взгляд на различные макроструктурные компоненты показал, что их производство было поразительно похоже на обоих языках во всех трех группах, что верно как для часто используемых, так и для редко встречающихся компонентов.

Рисунки 3 и 4 показывают, что не все компоненты создаются участниками одинаково часто: попытки и результаты включены большинством участников (не менее 80%) независимо от эпизода. Компонент, считающийся наиболее важным для хорошо сформированного повествования, цель, создавался гораздо реже, примерно в 50% эпизодов 1 и 2 и только в 30% повествований для эпизода 3.

Это может быть связано с реальными различиями между эпизодами, которые обозначены на картинках. Возможно, детям не так уж важно выразить цель («коза хочет спасти своего козленка») в ситуации, когда показано, что козленок тонет, и необходима быстрая реакция. В повествовании можно было бы отразить эту ситуацию естественным способом, сказав, что коза побежала в воду и спасла своего козленка (или вытолкнула его из воды), обозначая только попытку и результат [Gagarina, Bonhacker, Lindgren 2019: 18].

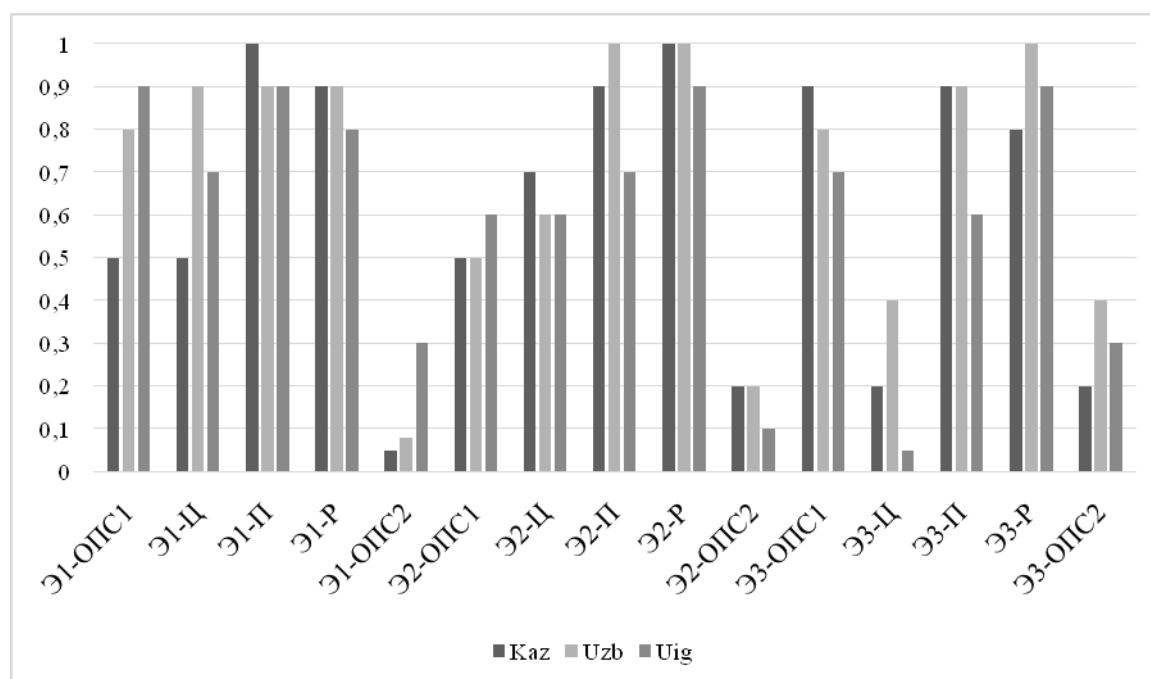


Рис. 3. Сравнение частотности упоминаний макроструктурных компонентов в трех языковых группах на L1

Примечание. Э1 — эпизод 1, Э2 — эпизод 2, Э3 — эпизод 3; ОПС1 — описание психологического состояния как начала события; ОПС2 — описание психологического состояния как реакции; П — попытка; Р — результат; Ц — цель (см. таблицу 2).

¹³ Согласно данным Министерства национальной экономики РК на начало 2020 г., казахи составляли 75,3% (1 531 716 человек), а узбеки 17,6% (344 937 человек) от общей численности населения городов Шымкент и Туркестан. Численность уйгуров в Алматы составляла 160 801 человек, или 57,6% от общей численности уйгуров страны [интернет ресурс] <https://stat.gov.kz/>.

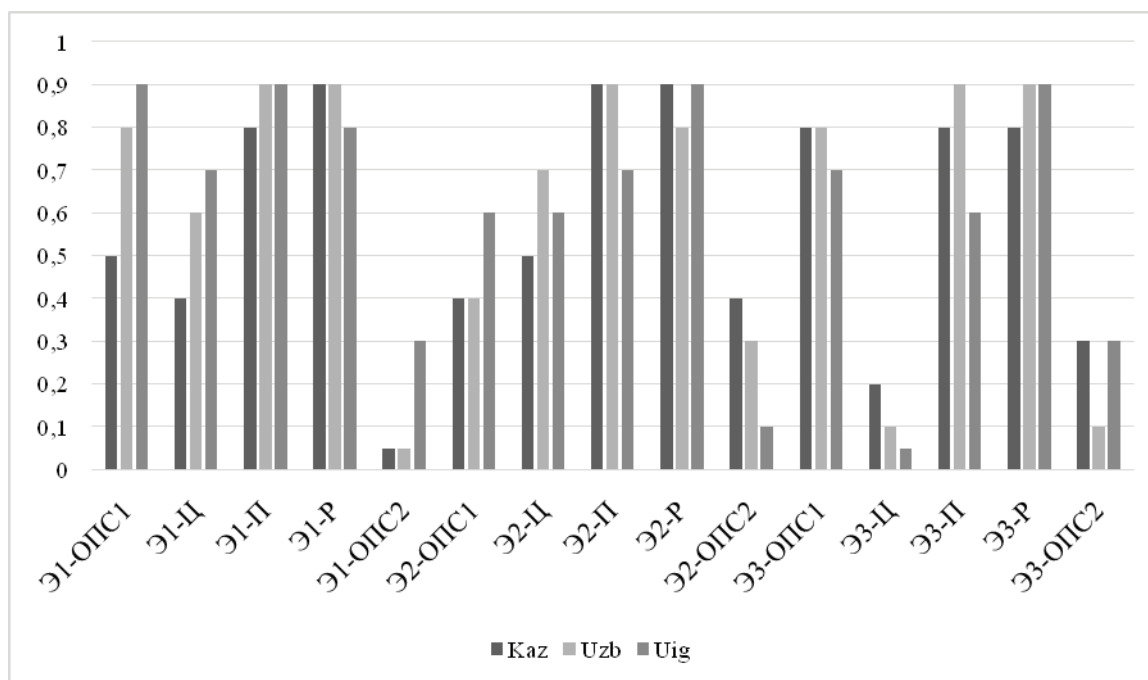


Рис. 4. Сравнение частотности упоминаний макроструктурных компонентов в трех языковых группах на L2 (русский язык)

Таким образом, не все цели одинаково часто включаются в повествование. Напротив, ОПС как начало события были очень частыми в рассказах участников, причем около 70% или более участников создавали их для каждого эпизода. Для ОПС как реакции была заметная разница между эпизодами. В 1-м эпизоде этот компонент производился очень редко (только около 12% участников), а во 2-м и 3-м эпизодах его частотность увеличивается (эпизод 2 — 20%, эпизод 3 — 30%).

Как мы отмечали ранее, наличие трех основных компонентов (цель, попытка, результат (ЦПР)) в каждом эпизоде отражает максимально возможную сложность истории, которая является *качественной оценкой макроструктуры* в ТОН. Так, мы сначала проанализировали долю участников, которые достигли каждого из различных уровней макроструктурной сложности в каждой языковой группе хотя бы один раз в своих рассказах.

Сначала мы проанализировали рассказы на L1 (казахском, узбекском и уйгурском) языке в каждой группе. Так, в двух языковых группах большинство носителей (казахско-русская: 85%, уйгурско-русская: 70%) использовали по крайней мере одну полную последовательность, а в узбекско-русской группе абсолютно все участники отметили хотя бы одну полную структуру (100%), и, таким образом, достигли наивысшего уровня сложности в своем повествовании.

Здесь же можно выделить рассказы, в которых были отмечены все компоненты во всех трех эпизодах и был набран максимальный балл за сложность истории: в казахско-русской группе — 2, в узбекско-русской — 5 историй. Но в то же время были и те, которые не создали ни одного полного эпизода (цель, попытка, результат): казахско-русская: 15%, уйгурско-русская: 30%.

Далее мы проанализировали рассказы на русском языке в каждой языковой группе. Здесь мы увидели относительно схожую картину. Так, во всех трех языковых группах большинство носителей (казахско-русская: 65%, узбекско-русская: 80%, уйгурско-русская: 70%) использовали по крайней мере одну полную последовательность и, таким образом, достигли наивысшего уровня сложности в своем повествовании.

Во всех группах были и те, кто не создал ни одного полного эпизода: казахско-русская — 35%, узбекско-русская — 20%, уйгурско-русская — 30%. Кроме того, ни в одной группе мы не встретили рассказ, в котором присутствуют все компоненты всех трех эпизодов истории.

Сходство между языковыми группами в частотности использования полных компонентов в каждом эпизоде привлекает особое внимание. В рассказах участников эксперимента и на родном, и на русском языках сложность историй значительно различалась между тремя эпизодами. Так, например, наличие всех трех компонентов в основном можно было увидеть в эпизоде 1. В эпизодах 2 и 3 часто пропускались компоненты цель и попытка. Почти во всех историях самая низкая сложность была в эпизоде 3.

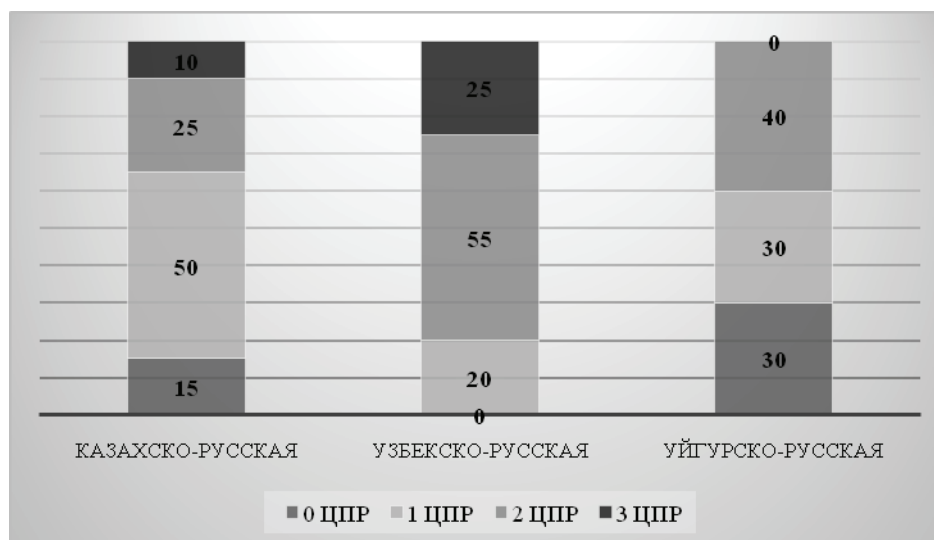


Рис. 5. Процентное соотношение сложности историй в трех языковых группах на L1

Примечание. ЦПР — цель, попытка, результат (основные компоненты истории); 0 ЦПР — ни одного полного эпизода; 1 ЦПР — один полный эпизод; 2 ЦПР — два полных эпизода; 3 ЦПР — три полных эпизода.

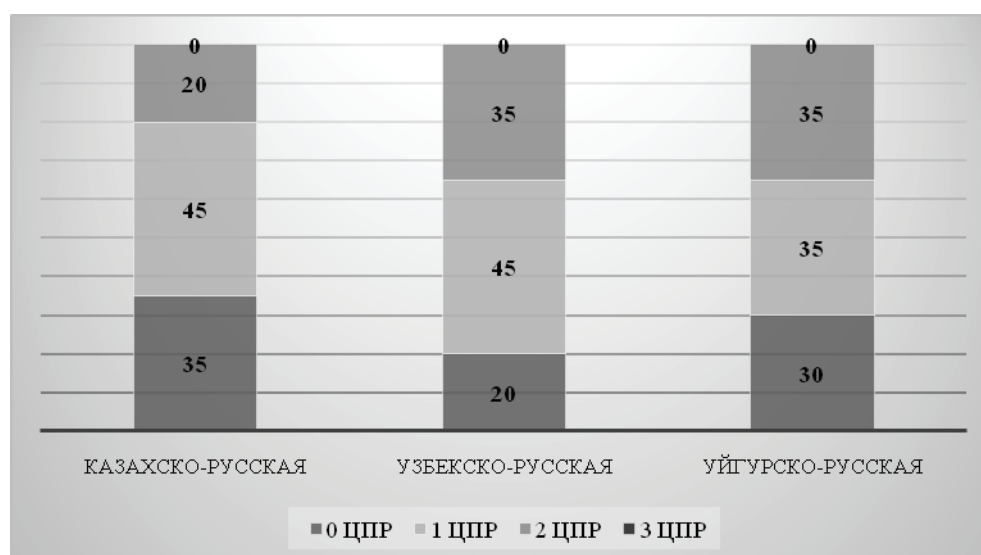


Рис. 6. Процентное соотношение сложности историй в трех языковых группах на L2

6. Обсуждение и заключение

В настоящем исследовании изучалась макроструктура устных повествовательных текстов учащихся билингвов трех языковых групп: казахско-русская, узбекско-русская и уйгурско-русская. Тексты были собраны с использованием историй из Теста на Определение Нарративных навыков многоязычных детей (ТОН). Для оценки структуры сюжета и сложности сюжета была использована многомерная модель макроструктуры. Цель состояла в том, чтобы определить *количественную и качественную оценки микроструктуры* историй и сравнить полученные результаты трех языковых групп. Для этого, во-первых, были проанализированы оценки структуры сюжета по языкам и историям, во-вторых, была исследована частота использования различных компонентов, и, наконец, была изучена сложность сюжета.

Результаты показывают, что структура повествований и сложность сюжета проявляются в трех языковых группах очень похожим образом. Интересно, что средние баллы наших 30 участников-билингвов все еще были довольно далеки от максимума (17 баллов) за структуру рассказа (в среднем 11 или 12 баллов, всего 65—70% от максимального). Это важный результат, поскольку он показывает, что нор-

мальный, типичный рассказчик создает в среднем от 11 до 12 компонентов в рассказе, а не все 17 компонентов (как в идеализированном сценарии в таблице 2), хотя некоторые из наших участников выразили 15 или 16 компонентов.

Помимо этого, результаты участников Kaz и Uzb языковых групп на русском языке были относительно ниже, чем на родном языке. Мы полагаем, что это является результатом влияния языковой среды. Эксперименты проводились в разных городах и, соответственно, языковые ситуации тоже отличаются. Поэтому важно подчеркнуть, что при проведении в республике научных исследований по вопросам билингвизма необходимо учитывать исторически сложившиеся на огромной территории страны региональные социально-экономические, демографические и этноязыковые условия, порой диаметрально противоположные в различных областях Казахстана.

Более пристальный взгляд на различные макроструктурные компоненты показал, что производство было поразительно схожим в трех языковых группах, что верно как для часто используемых, так и для редко встречающихся компонентов макроструктуры. Так, например, во всех рассказах часто пропускались следующие компоненты: описания психологического состояния (A6 — ОПС как реакция) и начало событий эпизодов 2, 3 (A7 и A12 — ОПС как начало события). При этом только при пересказах эти компоненты упоминаются, а в рассказах их и вовсе нет. И более того, не все компоненты создаются участниками одинаково часто: попытки и результаты включены большинством участников (не менее 80%), независимо от эпизода. Компонент, считающийся наиболее важным для хорошо сформированного повествования, цель, создавался гораздо реже, примерно в 50% эпизодов 1 и 2 и только в 30% повествований для эпизода 3. Таким образом, не все цели одинаково часто включаются в повествования. Напротив, ОПС как начало события были очень частыми в рассказах участников, причем около 70% или более участников создавали их для каждого эпизода. Для ОПС как реакции была заметная разница между эпизодами. В 1-м эпизоде этот компонент производился очень редко (только около 12% участников), а во 2-м и 3-м эпизодах частотность увеличивается (эпизод 2 — 20%, эпизод 3 — 30%). Это является хорошим доказательством того, что изобразительные стимулы являются межъязыковыми.

Мы также рассмотрели сложность сюжета с другой стороны, оценив количество эпизодов, в которых участники реализовали полные последовательности ЦПР. Было обнаружено, что во всех трех языковых группах и на L1, и на L2 большинство носителей использовали по крайней мере одну полную последовательность и, таким образом, достигли наивысшего уровня сложности в своем повествовании. Здесь же можно выделить рассказы, в которых были отмечены все компоненты во всех трех эпизодах и был достигнут максимальный уровень за сложность истории.

Так, результаты нашего исследования существенно отличаются не только соотношением качественно-количественных показателей макроструктурной сложности рассказов билингвов на двух языках, но и сравнительным описанием особенностей близкородственных языков, их связей, сходств и различий. В статье впервые делается попытка объединить в рамках одного исследования результаты анализа макроструктуры историй билингвов трех близкородственных языковых групп.

Наблюдения показывают, что повествовательная компетентность развивается в зависимости от возраста и склонности к повествованию историй и становится полностью структурированной только у взрослого человека. Дети начинают создавать повествования примерно в два года, и с этого момента их повествовательная компетентность интенсивно развивается. Но повествовательная компетентность развивается не у всех детей с одинаковой скоростью. Так, анализ повествовательных навыков билингвов — сложный процесс, требующий тщательного планирования.

В данной работе мы не стали детально описывать гендерные различия: во всех языковых группах девочки лучше рассказывали истории и, соответственно, набрали высокие баллы по производствам макроструктурных элементов. Почти во всех рассказах девочек можно обнаружить те или иные элементы описания психического состояния (например, *видеть, слышать, чувствовать, нюхать*), тогда как мальчики больше сосредоточены на сюжете историй и комплексности структуры (*цель + попытка + результат*). Данное явление уже подробно описано в трудах Н. Гагариной, С. Фичмана, Е. Галкиной, Дж. Линдгрена, У. Бонакер и др.

Тем не менее до сих пор не были исследованы региональные особенности билингвов, носителей русского и одного из тюркских языков Центральной Азии, в том числе в Казахстане. В нашем исследовании делается попытка сопоставить взаимодействие русского и *близкородственных* языков в речи учащих-билингвов, проживающих с рождения в Казахстане.

Безусловно, такой поверхностный анализ не дает полной картины нарративных навыков билингвов. Необходимо провести дополнительные исследования: привлечь больше участников, не ограничиваясь только детьми-билингвами; детально анализировать каждый элемент макроструктуры устных повествовательных текстов и сравнить с полученными данными. Кроме того, требуется тщательный лингвисти-

ческий анализ для получения информации о грамматических навыках испытуемого. Ввиду неоднородности национального состава и коммуникативно-языковых предпочтений регионов Казахстана следует расширить географию исследований и проанализировать нарративные навыки на русском и на родном языке в группах с родным языком обучения.

Литература

- Алтынбекова 2005 — О. Б. Алтынбекова. Казахский и русский языки в образовательном пространстве // Вестник КазНУ. Серия филологическая. 2005, 7. С. 14—19.
- Булатбаева, Сарсикеева, Айтжанов 2020 — К. Н. Булатбаева, Г. К. Сарсикеева, К. Б. Айтжанов. Оказиональное в современном языке: новые аспекты анализа // Вестник ЧГПУ им. И. Я. Яковлева. 2020, 1. С. 10—16.
- Казкенова, Рахилина 2019 — А. Казкенова, Е. Рахилина. Русская письменная речь студентов-казахов в зеркале русского учебного корпуса: глагольная префиксация // *Przegląd Wschodnioeuropejski*. 2019, 10 (1). С. 389—398.
- Сулейменова 2017 — Э. Д. Сулейменова. Школы с русским языком обучения в Казахстане: Вызовы трехязычного образования // Текст культуры и культура текста. Алматы, 2017. С. 639—644.
- Сулейменова, Аканова, Аймагамбетова 2021 — Э. Д. Сулейменова, Д. Аканова, М. М. Аймагамбетова. *Biz birgemiz*, или Еще раз о диверсификации русского языка в Казахстане // Вестник Российского университета дружбы народов. 2021, 1. С. 7—22.
- Суюмбаева, Шелестюк 2022 — А. Ж. Суюмбаева, Е. В. Шелестюк. Проблемы переходной диглоссии и функционирование билингвизма в профессиональной коммуникации (на примере языковой ситуации в Республике Казахстан) // Вестник ЧелГУ. 2019, 10. С. 149—164.
- Туксайтова 2016 — П. Туксайтова. Русский язык в мультикультурном пространстве Казахстана: государственная политика и общественные настроения // *Quaestio Rossica*. 2016. Т. 4: 4. С. 94—106.
- Утегенова 2016 — К. Т. Утегенова. Функционирование русского языка в западно-казахстанской области // Русский язык как фактор культурно-образовательной интеграции общества. 2016. С. 21—28.
- Berman, Slobin 1994 — R. A. Berman, D. I. Slobin. Different ways of relating events: Introduction to the study // *Relating Events in Narrative: A Crosslinguistic Developmental Study* / Eds. R. A. Berman, D. I. Slobin. Lawrence Erlbaum Associates, 1994. P. 1—16.
- Bernardini, Schlyter 2004 — P. Bernardini, S. Schlyter. Growing syntactic structure and code-mixing in the weaker language: The Ivy hypothesis // *Bilingualism: Language and Cognition*. 2004, 7 (1). P. 49—69.
- Bohnacker et al. 2016 — U. Bohnacker, J. Lindgren, B. Öztekin. Turkish- and German-speaking bilingual 4-to-6-year-olds living in Sweden: Effects of age, SES and home language input on vocabulary production // *Journal of Home Language Research*. 2016, 1. P. 17—41.
- Gagarina, Bohnacker, Lindgren 2019 — N. V. Gagarina, U. Bohnacker, J. Lindgren. Macrostructural organization of adults' oral narrative texts // *ZAS Papers in Linguistics*. 2019, 62. P. 190—208.
- Gagarina et al. 2021 — N. Gagarina, S. Fichman, E. Galkina, E. Protassova, N. Ringblom, Y. Rodina. How oral texts are organized in monolingual and heritage Russian: Evidence from six countries // *Language Impairment in Multilingual Settings*. Amsterdam, 2021. P. 48—75.
- Granfeldt 2000 — J. Granfeldt. The acquisition of the determiner phrase in bilingual and second language French // *Bilingualism: Language and Cognition*. 2000, 3 (3). P. 263—280.
- Lindgren 2018 — J. Lindgren. Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4—6 (*Acta Universitatis Upsaliensis* 19). Uppsala, 2019.
- Lindgren, Bohnacker 2020 — J. Lindgren, U. Bohnacker. Vocabulary development in closely-related languages: Age, word type and cognate facilitation effects in bilingual Swedish-German preschool children // *Linguistic Approaches to Bilingualism*. 2020, 10 (5). P. 587—622.
- Paradis, Schneider 2008 — J. Paradis, F. Schneider. Distinguishing Bilingual Children from Monolinguals with SLI: Profile Effects on the ENNI. (Symposium on Research in Child Language Disorders, June 2008).
- Pinto et al. 2018 — G. Pinto, Chr. Tarchi, B. Accorti Gamannossi. Kindergarteners' Narrative Competence Across Tasks and Time // *The Journal of Genetic Psychology*. 2018. Vol. 179: 3. P. 143—155.

References

- Altynbekova 2005 — O. B. Altynbekova. Kazahskij i russkij yazyki v obrazovatel'nom prostranstve. In: *Vestnik KazNU. Seriya filologicheskaya*. 2005, 7. S. 14—19. {O. B. Altynbekova. Kazakh and Russian languages in the educational space. In: *Bulletin of KazNU. Philology Series*. 2005, 7. P. 14—19.}

- Bernardini, Schlyter 2004 — P. Bernardini, S. Schlyter. Growing syntactic structure and code-mixing in the weaker language: The Ivy hypothesis. In: *Bilingualism: Language and Cognition*. 2004, 7 (1). P. 49—69.
- Berman, Slobin 1994 — R. A. Berman, D. I. Slobin. Different ways of relating events: Introduction to the study. In: *Relating Events in Narrative: A Crosslinguistic Developmental Study* / Eds. R. A. Berman, D. I. Slobin. Lawrence Erlbaum Associates. 1994, 2. P. 775—790.
- Bohnacker et al. 2016 — U. Bohnacker, J. Lindgren, B. Öztekin. Turkish- and German-speaking bilingual 4-to-6-year-olds living in Sweden: Effects of age, SES and home language input on vocabulary production. In: *Journal of Home Language Research*. 2016, 1. P. 17—41.
- Bulatbaeva, Sarsikeeva, Ajtzhano 2020 — K. N. Bulatbaeva, G. K. Sarsikeeva, K. B. Ajtzhano. Okkazional'noe v sovremennom yazyke: novye aspekty analiza. In: *Vestnik ChGPU im. I. Ya. Yakovleva*. 2020, 1. S. 10—16. {K. N. Bulatbaeva, G. K. Sarsikeeva, K. B. Ajtzhano. Occasional in modern language: new aspects of analysis. In: *Bulletin of the I. Ya. Yakovlev ChSPU*. 2020, 1. P. 10—16.}
- Gagarina, Bohnacker, Lindgren 2019 — N. V. Gagarina, U. Bohnacker, J. Lindgren. Macrostructural organization of adults' oral narrative texts. In: *ZAS Papers in Linguistics*. 2019, 62. P. 190—208.
- Gagarina et al. 2021 — N. Gagarina, S. Fichman, E. Galkina, E. Protassova, N. Ringblom, Y. Rodina. How oral texts are organized in monolingual and heritage Russian: Evidence from six countries. In: *Language Impairment in Multilingual Settings*. Amsterdam, 2021. P. 48—75.
- Granfeldt 2000 — J. Granfeldt. The acquisition of the determiner phrase in bilingual and second language French. In: *Bilingualism: Language and Cognition*. 2000, 3 (3). P. 263—280.
- Kazkenova, Rakhilina 2019 — A. Kozkenova, E. Rakhilina. Russian Written Speech of Kazakh Students in the Reflection of Russian Learner Corpus: Verbal Prefixation. In: *Przegląd wschodnioeuropejski*. 2019, 10(1). P. 389—398.
- Lindgren 2018 — J. Lindgren. Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4—6 (*Acta Universitatis Upsaliensis* 19). Uppsala, 2019.
- Lindgren, Bohnacker 2020 — J. Lindgren, U. Bohnacker. Vocabulary development in closely-related languages: Age, word type and cognate facilitation effects in bilingual Swedish-German preschool children. In: *Linguistic Approaches to Bilingualism*. 2020, 10 (5). P. 587—622.
- Paradis, Schneider 2008 — J. Paradis, F. Schneider. Distinguishing Bilingual Children from Monolinguals with SLI: Profile Effects on the ENNI. (Symposium on Research in Child Language Disorders, June 2008.)
- Pinto et al. 2018 — G. Pinto, Chr. Tarchi, B. Accorti Gamannossi. Kindergarteners' Narrative Competence Across Tasks and Time. In: *The Journal of Genetic Psychology*. 2018. Vol. 179: 3. P. 143—155.
- Sulejmenova 2017 — E. D. Sulejmenova. Shkoly s russkim yazykom obucheniya v Kazahstane: Vyzovy trekh"yazychnogo obrazovaniya. In: *Tekst kul'tury i kul'tura teksta*. 2017. S. 639—644. {E. D. Sulejmenova. Schools with the Russian language of instruction in Kazakhstan: Challenges of trilingual education. In: *The text of culture and the culture of the text*. 2017. P. 639—644.}
- Sulejmenova, Akanova, Ajmagambetova 2021 — E. D. Sulejmenova, D. Akanova, M. M. Ajmagambetova. Biz birgemiz, ili Eshche raz o diversifikacii russkogo yazyka v Kazahstane. In: *Vestnik Rossijskogo universiteta druzhby narodov*. 2021, 1. S. 7—22. {E. D. Sulejmenova, D. Akanova, M. M. Ajmagambetova. Biz birgemiz, or Once again about the diversification of the Russian language in Kazakhstan. In: *Bulletin of the Peoples' Friendship University of Russia*. 2021, 1. P. 7—22.}
- Suyunbaeva, Shelestyuk 2022 — A. Zh. Suyunbaeva, E. V. Shelestyuk. Problemy perekhodnoj diglossii i funkcionirovanie bilingvizma v professional'noj kommunikacii (na primere yazykovoj situacii v Respublike Kazahstan). In: *Vestnik ChelGU*. 2019, 10. S. 149—164. {A. Zh. Suyunbaeva, E. V. Shelestyuk. Problems of transitional diglossia and functioning of bilingualism in professional communication (on the example of the language situation in the Republic of Kazakhstan). In: *Bulletin of ChelSU*. 2019, 10. P. 149—164.}
- Tuksaitova 2016 — R. Tuksaitova. The Russian language in the multicultural space of Kazakhstan: State policy and public mood. In: *Quaestio Rossica*. 2016. Vol. 4: 4. P. 94—106.
- Utegenova 2016 — K. T. Utegenova. Funkcionirovanie russkogo yazyka v zapadno-kazahstanskoj oblasti. In: *Russkij yazyk kak faktor kul'turno-obrazovatel'noj integracii obshchestva*. 2016. S. 21—28. {K. T. Utegenova. Functioning of the Russian language in the West Kazakhstan region. In: *Russian language as a factor of cultural and educational integration of society*. 2016. P. 21—28.}

Отражение особенностей говоров жителей Семиреченской области в «Сказке о Кара-Мергене»

Гаджиева Анар Ахметбековна,
Московский государственный институт (университет) международных отношений МГИМО МИД РФ (Москва),
Институт системного программирования им. В. П. Иванникова РАН; nuanar@yandex.ru

Статья посвящена описанию вокалических, консонантных и морфологических особенностей «Сказки о Кара-Мергене» (Казань, 1901), записанной Н. Н. Пантусовым в Семиречье. Последовательно рассмотрены графико-фонетические и морфологические особенности источника в сопоставлении с данными из ранее исследованных кириллических книг на казахском языке [Премудрости 1891; Букварь 1892; Училище Благочестия 1892; Крещение Руси 1892], изданных в рамках деятельности Переводческой комиссии в конце XIX — начале XX вв.

Обзор примеров, иллюстрирующих тенденции губного сингармонизма, показал, что лабиализация узких гласных второго, третьего и четвертого слога зафиксирована на письме во всех случаях. В результате исследования был сделан вывод, что система сингармонизма в «Кара-Мерген» отличается как от литературного казахского языка, так и от изученных ранее первых книг. Это подтверждает тот факт, что ранее учеными не были изучены первые книги на Семиреченском диалекте.

Анализ аффикса аккузатива, архаичный вариант которого (-ны) sporadически сохранился как в «Кара-Мерген», так и в западных [Премудрости 1891] и северных [Букварь 1892] диалектных книгах, показывает, что его изменение, видимо, еще не завершилось в конце XIX в.

Ключевые слова: казахский язык, семиреченский говор, памятники письменности, графико-орфографические особенности, морфологические особенности, Н. Н. Пантусов

REFLECTION OF THE PECULIARITIES THE DIALECT OF SEMIRECHENSK REGION IN “THE TALE OF KARA-MERGEN”

Anar A. Gadzhieva, Moscow State Institute (University) of International Relations MGIMO
of the Ministry of Foreign Affairs of the Russian Federation (Moscow);
Ivannikov Institute for System Programming, RAS (Moscow); nuanar@yandex.ru

The article describes vowel and consonant features of “The Tale of Kara-Mergen” (Kazan, 1901) from N. N. Pantusov’s collection which was recorded in Semirechye. The author considers in detail the monument’s writing system and analyzes the phonological and morphological features of the text’s language in comparison with the data of previously studied Cyrillic books in the Kazakh language [Wisdom of Jesus the Son of Sirach 1891; Primer for Kyrgyz people 1892; School of Piety 1892; The Baptism of Russia 1892], which were published within the framework of the Kyrgyz mission at the end of 19th and the beginning of the 20th century.

The analysis showed that the vocalism presented in “Kara-Mergen” does not differ significantly from the vocalism of the compared first Cyrillic books in Kazakh. A review of examples illustrating the tendencies of labial vowel harmony showed that the labialization of narrow vowels of the second, third and fourth syllables is fixed in writing in all cases. As a result of the study, it was concluded that the system of vowel harmony in “Kara-Mergen” differs from the literary Kazakh language and from the first books studied earlier. This confirms the fact that scientists have not previously studied the first books in the Semirechensk dialect.

An analysis of the accusative affix, the archaic version of which (-ny) sporadically preserved in “Kara-Mergen” and in Western [Wisdom 1891] and Northern [Primer 1892] dialect books, shows that its change, apparently, was not yet completed at the end of the 19th century.

Keywords: Kazakh language, Semirechensk dialect, written monuments, graphic and orthographic features, morphological features, N. N. Pantusov

1. Введение

Настоящая статья посвящена анализу «Сказки о Кара-Мергене», записанной Н. Н. Пантусовым в Семиречье. С глоссированным текстом и конкордансом сказки можно ознакомиться на следующих страницах портала ЛингвоДок:

- 1) <http://lingvodoc.ispras.ru/dictionary/4105/6/perspective/4105/9/view> (глоссированный корпус);
- 2) <http://lingvodoc.ispras.ru/dictionary/4105/21/perspective/4105/22/view> (конкорданс).

Основное внимание в данной статье направлено на изучение графико-фонетических и морфонологических особенностей источника. Анализируемые формы, зафиксированные в тексте, представлены в оригинальной кириллической транскрипции. При анализе особенностей материала по возможности приводятся соответствия из ранее исследованных кириллических книг на казахском языке [Премудрости 1891; Букварь 1892; Училище 1892 и Крещение 1892].

Текст сказки «Кара-Мерген» изначально был зафиксирован арабицей в 1898 г., а в 1901 г. был издан на кириллице с переводом на русский язык. Н. Н. Пантусов не указывает, с чьих слов была записана сказка. Однако принадлежность говора источника жителю Семиреченской области не вызывает сомнения, так как известно, что ученый долгое время работал чиновником Семиреченского областного управления (с 1885 по 1904 гг.), занимал должность вице-губернатора области. Исполнение государственных задач в качестве чиновника особых поручений, сопровождавшееся частыми поездками по краю, давало ему возможность в полной мере применять приобретенные в университете познания на практике [Горбуров, Горбуров 2016: 119]. Некоторые материалы, записанные и переведенные им, помечались как собранные в Копальском уезде Семиреченской области, например «Киргизские пословицы, записанные в Копальском уезде Семиреченской области» (Казань, 1900) и «Образцы киргизской народной литературы» (Казань, 1909), собранные и переведенные Н. Н. Пантусовым, см. [Пантусов 1909]. В этой связи обращают на себя внимание слова Н. Ф. Катанова: «Образцы киргизской народной литературы, как в киргизском оригинале, так и в русском переводе, изданные в России, были записаны в разных местах, населенных киргизами: в Сырдарьинской, Тургайской, Семипалатинской областях, Оренбургской губернии и Дзунгарской степи. Пословиц же семиреченских киргиз, изданных по-киргизски и с русским переводом, я вовсе не знаю. Ввиду сего собрание неутомимого деятеля на ниве науки Н. Н. Пантусова я считаю делом весьма важным для тюркологии, тем более, что пословицы записаны и переведены им в высшей степени добросовестно. Киргизская транскрипция и русский перевод не потребовали никаких изменений или добавлений, и потому представляются они в том виде, в каком поступили от автора» [Пантусов 1900: V].

2. Графика и орфография памятника

Для записи казахского текста была применена транскрипция, принятая в тюркских изданиях Императорской Академии наук и Императорского Казанского университета. Ниже приводится таблица соответствий графем, использованных Н. Н. Пантусовым в сказке, графемам казахского литературного языка.

Таблица 1. Графические особенности «Кара-Мерген» в сравнении с казахским литературным языком

«Кара-Мерген»	Современный казахский литературный язык
ä	e
ā	ə
e	e
i	i
ö	ө
ÿ	Ү
и	1) жс; 2) дж
б	з
к	қ
у ¹	у губной, заднерядный, краткий гласный 1) у в произношении [w]; 2) ув; үв
у ²	
l	л'
л	л

В целом, если не учитывать отклонений говора-источника от литературной произносительной нормы, графическое оформление исследуемого текста вполне соответствует нормам орфоэпии казахского языка. Описки и опечатки в «Кара-Мерген» встречаются крайне редко.

3. Графико-фонетические особенности

3.1. Вокализм

1. Гласный *a* почти во всех случаях соответствует казахскому литературному *a*, ср.: *арык* ‘тощий’ ~ лит. каз. *арық*; *сақтайды* ‘хранит’ ~ лит. каз. *сақтайды*; *жолдас* ‘товарищ’ ~ лит. каз. *жолдас*.

Исключение составляют несколько слов в позиции непервого слога, например *қауіптанып* ~ лит. каз. *қауіптеніп* ‘опасаясь’; *қасыжатты* ~ лит. каз. *қасиетті* ‘благородный’. В тексте сказки имеется единственный случай с нестандартным соответствием *a* ~ лит. каз. *ә*, ср.: *үана* ~ лит. каз. *және* ‘и’. Аналогичный пример зафиксирован в миссионерской кириллической книге на казахском языке [Премудрости 1891] и в западном диалекте казахского языка [Гаджиева 2019: 62].

2. Гласный *ā*, соответствующий литературному *ә*, зафиксирован в основном в анлауте в арабо-персидских заимствованиях: *āp tūrly* ~ лит. каз. *әртүрлі* ‘различный’; *ādāt* ~ лит. каз. *әдет* ‘обычай’; *ākā* ~ лит. каз. *әке* ‘отец’; *ālāmdā* ~ лит. каз. *әлемде* ‘в мире’; *ācīl kылу* ~ лит. каз. *хасіл қылу* ‘осуществлять’; есть несколько примеров употребления и в инлауте, ср. *bārī* ~ лит. каз. *бәрі* ‘все’. Как известно, палатальный звук *ә* в казахском языке почти всегда употребляется в анлауте слов, заимствованных из арабского и персидского языков. По утверждению казахских лингвистов, гласная фонема *ә* в казахском языке — явление позднее, возникшее в результате заимствования из восточных языков [Кенесбаев 1962: 21]. Интересно, что в персидском языке большинство этих слов имеют краткий *a*, ср. *عادت* ‘привычка’ ~ *ādāt* ~ лит. каз. *әдет* ‘обычай’. Можно предполагать, что передача краткого *a* как *ā* или *ā* в «Кара-Мерген» и в других первых кириллических книгах, см. [Дыбо, Норманская 2016], отражает тот факт, что в XIX в. этот звук был более низкого подъема.

3. Гласный *ā* употребляется в соответствии с лит. каз. *e/э*. Представлен чаще всего в инлауте и ауслауте: *мергәнә* N-DAT ~ лит. каз. *мергенге* ‘охотнику’; *дегәндәй* PTCP-COMPAR ~ лит. каз. *дегендей* ‘словно сказал’; *пәрзәнті* N-POSS.3SG ~ лит. каз. *перзенті* ‘дитя его’; *тізәлән* ADV.PTCP ~ лит. каз. *тізелен* ‘сидя на коленях’; *қіләмдәр* N-PL ~ лит. каз. *кілемдер* ‘ковры’; *біткән* PTCP ~ лит. каз. *біткен* ‘произросший’. Лишь в одном случае он употребляется в абсолютном начале условного союза (перс.) *ägär* ~ лит. каз. *егер* ‘если’.

Гласный *e* в анлауте и после согласных в большинстве случаев употребляется в соответствии с лит. каз. *e*, ср.: *ертәгі* ~ лит. каз. *ертегі* ‘сказка’; *екән* ~ лит. каз. *екен* ‘оказывается’; *есәңсіз* ~ лит. каз. *есенсіз* ‘несметный’; *етәгінән* N-POSS.3SG-ABL ~ лит. каз. *етегінен* ‘за полу одежды’; *ертәң ертә* ~ лит. каз. *ертең ерте* ‘рано утром’. В одном корне в «Кара-Мерген» представлено соответствие *e* ~ лит. каз. *i*, ср. *үебәрін* ~ лит. каз. *жіберін* ‘отослав’; *үебәрән* ~ лит. каз. *жіберген* ‘посланный’. При этом интересно отметить, что соответствие *e/э* ~ лит. каз. *i* очень часто встречается в миссионерской кириллической литературе на казахском языке. Так, в [Премудрости 1891] соответствие *e/э* ~ лит. каз. *i* отмечено в 48 лексемах как в корне, так и в аффиксах, например: *белген* ~ лит. каз. *білген* ‘знавший’; *бетип* ~ лит. каз. *бітіп* ‘закончившись’; *езгелек* ~ лит. каз. *ізгілік* ‘благодетельное’. Соответствие *э* ~ лит. каз. *i* встречается в абсолютном начале слов: *эзгі* ~ лит. каз. *ізгі* ‘благий’; *эши* ~ лит. каз. *іші* ‘внутри’; *эсти* ~ лит. каз. *істі* ‘находящийся в беде’. Аналогичное явление представлено также в [Букварь 1892]. Подобного рода переходы наблюдались в разных регионах проживания носителей казахского языка, например в говорах жителей восточно-казахстанской области: *сереңке* ~ лит. каз. *сіріңке* ‘спички’, *ерегесу* ~ лит. каз. *ерегісу* ‘препираться’, *бүлдірге* ~ лит. каз. *бүлдіргі* ‘гречиха’, *керпіш* ~ лит. каз. *кірпіш* ‘кирпич’ [Болатов 1963: 194—195]. В западном диалекте на тенденцию перехода *i > e* указывал Н. И. Ильминский, иллюстрируя это явление примерами *елбеу* ~ лит. каз. *ілбу* ‘брести’, *ілегу* ~ лит. каз. *ілігу* ‘зацепляться’, *кер* ~ лит. каз. *кір* ‘грязь’ [Аманжолов 1997: 400]. На эту же особенность произношения в западной степи указывал В. В. Радлов [Радлов 1870: xxv].

4. Гласный *ы* встречается во всех позициях как в корне (основе) слова, так и в аффиксах: *кыс* ~ лит. каз. *қыс* ‘зима’; *ялғыз* ~ лит. каз. *жалғыз* ‘один’; *алыс* ~ лит. каз. *алыс* ‘далекий’; *сатып* ADV.PTCP ~ лит. каз. *сатып* ‘продавать’; *мылтыбын* N-POSS.3SG-ACC ~ лит. каз. *мылтығын* ‘ружье его’. Звук *ы* употребляется в качестве вставки и в анлауте слова с начальным сонорным *л*: *ылажық* ~ лит. каз. *лайық* ‘достойный’. В двух случаях в интерконсонантной позиции отмечены вставки *ы* и *і*: *адамзат* ~ лит. каз. *адамзат* ‘люди’; *берігі* ~ лит. каз. *бергі* ‘ближний’.

Отметим, что в исследованных нами ранее кириллических книгах на казахском языке — [Премудрости 1891] и [Букварь 1892] — имело место выпадение звуков *ы*, *і* в первом слоге, хотя это явление менее частотно по сравнению со стандартными соответствиями. Интересно в связи с этим отметить, что в «Кара-Мерген» случаи выпадения в первом слоге *ы*, *і* не представлены.

5. В сказке в большинстве случаев *і* в основах и аффиксах слов употребляется в соответствии с лит. каз. *i*: *бір* ~ лит. каз. *бір* ‘один’; *сегіз* ~ лит. каз. *сегіз* ‘восемь’; *ішінән* N-POSS.3SG-ABL ~ лит. каз. *ішінен*

‘изнутри’; *kiciçi* N-POSS.3SG ~ лит. каз. *kiciçi* ‘человек его’; *kemintmi* ADV.PTCP-PST.3SG ~ лит. каз. *kemintmi* ‘ушел’; *nersāniñ* N-GEN ~ лит. каз. *nerseniñ* ‘вещь’; *emti* N-ACC ~ лит. каз. *emti* ‘мясо’.

В нескольких случаях (регулярно в отрицательной частице персидского происхождения *eñ* ‘ни’ и арабском заимствовании *ese* ‘раз’) *i* употребляется в соответствии с лит. каз. *e*: *iñ bir* ~ лит. каз. *eñbir* ‘вовсе, ничего’; *iñ biräüi* ~ лит. каз. *eñbiräüi* ‘никто’; *içcā* ~ лит. каз. *ese* ‘раз’. Такое соответствие имеет место и в современном талдыкорганском говоре: *şikara* ~ лит. каз. *šekara* ‘граница’; *irteñ* ~ лит. каз. *erteñ* ‘завтра’; *žir* ~ лит. каз. *žer* ‘земля’. С полевым материалом талдыкорганского говора можно ознакомиться на сайте ЛингвоДок: <http://lingvodoc.ispras.ru/dictionary/380/4/perspective/380/5/view>.

Обращает на себя внимание регулярное отражение в «Кара-Мерген» *i* в соответствии с лит. каз. *e* в аффиксе инструментального падежа: *ižim-biñän* N-POSS.1SG-INS ~ лит. каз. *ižimmenen* ‘по моим следам’; *kolū-miñän* N-POSS.3SG-INS ~ лит. каз. *qolyminen* ‘руками его’; *myltyk-piñän* N-INS ~ лит. каз. *myltykpenen* ‘ружьем’. Такие морфонологические особенности были зафиксированы в говоре жителей некоторых южных областей Казахстана. Там наряду с литературной формой инструменталиса употребляются формы *-бынан/-мынан/-пынан*: *поезбынан* ~ лит. каз. *пойызбенен* (*пойызбен*) ‘на поезде’; *арба-мынан* ~ лит. каз. *арбаменен* (*арбамен*) ‘на телеге’; *атпынан* ~ лит. каз. *атпенен* (*атпен*) ‘на лошади’ [Қалиев, Сарыбаев 1991: 98].

6. Гласные *y*, *o*, *ы* в основном регулярно соответствуют лит. каз. *у*, *o*, *ы*. Но есть и очень немногочисленные случаи отклонений, приведем все найденные примеры. В исконной лексике 2 примера: *шоңқурда* N-LOC ~ лит. каз. *шұңқырда* ‘в яме’; *бујул* ~ лит. каз. *биыл* ‘в этом году’, остальные случаи отклонений в арабо-персидских заимствованиях: (араб.) *роксат* ~ лит. каз. *рұқсат* ‘разрешение’; (перс.) *суру кайнап* N-POSS.3SG-ADV.PTCP ~ лит. каз. *соры қайнап* ‘попасть в беду’; (перс.) *шунуны* N-ACC ~ лит. каз. *шыныны* ‘чашку’; (араб.) *мырадын* N-POSS.3SG-ACC ~ лит. каз. *мұратын* ‘цель его’; (перс.) *былбыл* ~ лит. каз. *бұлбұл* ‘соловей’.

7. Аналогично и *y*, *ö* стандартно соответствуют лит. каз. *у*, *ө*. Есть несколько исключений, приведем все найденные примеры: *мүнöldү* ~ лит. каз. *жөнелді* ‘отправился’; *өңүр* ~ лит. каз. *үңгір* ‘пещера’; *жетті-сегіз көндө* NUM - N-LOC ~ лит. каз. *жеті-сегіз күнде* ‘за семь-восемь дней’. Интересно отметить, что встречается и стандартное соответствие *y* ~ лит. каз. *у* в различных формах слова *күн*: *бір күндөрдө* NUM - N-PL-LOC ~ лит. каз. *бір күндерде* ‘как-то раз’; *бір күнү* NUM - N-POSS.3SG ~ лит. каз. *бір күні* ‘однажды’; *мас күнүндө* ADJ - N-POSS.3SG-LOC ~ лит. каз. *жас күнінде* ‘в молодости’. Подобные примеры соответствий *y* ~ лит. каз. *ө* и *ö* ~ лит. каз. *у* были зафиксированы также в первых кириллических книгах на казахском языке: [Премудрости 1891; Училище 1892; Букварь 1892]. Приведем лишь некоторые из них:

[Премудрости 1891]: *күрсетип* ~ лит. каз. *көрсетіп* ‘показав’; *күрер* ~ лит. каз. *көрер* ‘увидит’; *сүйлеп* ~ лит. каз. *сөйлеп* ‘говоря’; (араб.) *мөмкен* ~ лит. каз. *мүмкін* ‘возможно’; *жөрген* ~ лит. каз. *жүрген* ‘ходивший’; *өкүт* ~ лит. каз. *үгіт* ‘наставление’;

[Училище 1892]: *сүйтүп* ~ лит. каз. *сөйтіп* ‘таким образом’; *төзу* ~ лит. каз. *түзу* ‘прямая линия’; *жөрүбти* ~ лит. каз. *жүріпті* ‘ходил’; *өлестр* ~ лит. каз. *үлестір* ‘раздай’; *көндө-төндө* ~ лит. каз. *күні-түні* ‘днем и ночью’;

[Букварь 1892]: *сүйтсен* ~ лит. каз. *сөйтсең* ‘если сделаешь так’; *сүйтүб* ~ лит. каз. *сөйтіп* ‘таким образом’; *өйленетн* ~ лит. каз. *үйленетін* ‘тот, кто женится’; *төсүнде* ~ лит. каз. *түсінде* ‘во сне’; (араб.) *дөкен* ~ лит. каз. *дүкен* ‘магазин’.

Здесь хотелось бы отметить, что в [Премудрости 1891] употребление нестандартных соответствий *y* ~ лит. каз. *ө* и *ö* ~ лит. каз. *у* является более частотным. В «Кара-Мерген», [Училище 1892] и [Букварь 1892] указанные отклонения встречаются крайне редко.

3.2. Губная гармония

Что касается губной гармонии гласных непервого слога, то следует отметить, что во всех случаях орфография источника отражает лабиализацию *i*, *ы*, *e* последующих слогов после губных *y*, *у* и *ö* в первом:

Таблица 2. *y* (/ʏ/) во 2-м, 3-м и 4-м слоге вместо *ы* после *y* (/ʏ/) в 1-м

«Кара-Мерген»	Литературный казахский (орфографическая норма)	Перевод
<i>бубу</i>	<i>бұғы</i>	‘марал; олень’
<i>кулун</i>	<i>құлын</i>	‘жеребенок’
<i>бурунбу</i>	<i>бұрынғы</i>	‘прошлый; давний’
<i>цуртунуң</i> N-POSS.3SG-GEN	<i>жұртының</i>	‘народа; людей’
<i>шуркулдады</i> PST.3SG	<i>шұрқылдады</i>	‘закипел; вскипел’
<i>сулудубу</i> N-POSS.3SG	<i>сұлулығы</i>	‘красота ее’

Таблица 3. *y* (/y/), *ö* (/ø/) во 2-м и 3-м слоге вместо *i*, *e* после *ö* (/ø/) в 1-м

«Кара-Мерген»	Литературный казахский (орфографическая норма)	Перевод
<i>өзүнүң</i> PRON.POSS.3SG-GEN	<i>өзінің</i>	‘своего’
<i>өзөңгө</i> N-DAT	<i>өзенге</i>	‘к реке’
<i>шөптү</i> N-ACC	<i>шөпті</i>	‘траву’
<i>өскөңүн</i> PTCP.POSS.3SG-ACC	<i>өскенін</i>	‘что выросла’
<i>көргөн</i> PTCP	<i>көрген</i>	‘видел’
<i>көрүп</i> ADV.PTCP.	<i>көріп</i>	‘увидев’
<i>көрүндү</i> PST.3SG	<i>көрінді</i>	‘показалась’

Таблица 4. *y* (/y/) во 2-м, 3-м и 4-м слоге вместо *ы* после *o* в 1-м

«Кара-Мерген»	Литературный казахский (орфографическая норма)	Перевод
<i>отур</i>	<i>отыр</i>	‘сидит’
<i>отурунту</i> ADV.PTCP-PST.3SG	<i>отырыпты</i>	‘сидел’
<i>қорқуп</i> ADV.PTCP	<i>қорқып</i>	‘испугавшись’
<i>тоңу</i> N-POSS.3SG	<i>тоңы</i>	‘шуба его’
<i>қолунуң</i> N-POSS.3SG-GEN	<i>қолының</i>	‘рук его’
<i>қолтуғанан</i> N-POSS.3SG-ABL	<i>қолтығынан</i>	‘за подмышки его’
<i>мойнунан</i> N-POSS.3SG-ABL	<i>мойнынан</i>	‘за шею его’

Отметим, что в [Премудрости 1891] при *y*, *ü*, *ö* первого слога сингармонизм встречается только в 20—40% словоформ со вторым узким гласным, что сближает его с [Букварь 1892], но, в отличие от «Букваря...», при *o* в первом слоге в [Премудрости 1891] имеется только 8% словоформ со вторым узким гласным, подвергшимся гармонии [Гаджиева 2019: 66].

Анализ текстов на казахском языке, зафиксированных во второй половине XIX в. в различных областях казахской степи, все же дает нам право утверждать, что в речевой реализации аттракция губных гласных распространялась до четвертого и, реже, до пятого слога. В современном казахском языке лабиальная гармония в устной речи распространяется только до третьего слога.

В следующей таблице представлены схемы сингармонизма в рассматриваемом памятнике.

Таблица 5. Сингармонизм в «Кара-Мерген»

		1-й узкий	1-й широкий		1-й узкий	1-й широкий
Задний ряд	Непервые узкие гласные	+	+/- <i>шошуп,</i> <i>шоңқурда</i> но <i>болып, койды</i>	Непервые широкие гласные	-	- <i>мұпар</i> <i>домалак</i>
		+	+		+	+
Передний ряд		<i>мүјүзүн</i>	<i>көзү</i>		<i>мүрөр,</i> <i>мүнөлдү</i>	<i>көргөн</i>

Итак, можно видеть, что тип сингармонизма в «Кара-Мерген» весьма близок к орфоэпическим нормам литературного казахского языка, см. подробнее [Кенесбаев, Мусабаев, Неталиева 1973: 105], но отличается от него несколькими исключениями: после *o* есть примеры без сингармонизма: *болып, койды*. С точки зрения губного сингармонизма к литературному казахскому языку были близки и другие акмолинские книги — [Словарь 1897] и [Букварь 1908], см. [Дыбо, Норманская 2016: 148].

3.3. Консонантизм

1. В большинстве позиций *n* в «Кара-Мерген» соответствует лит. каз. *n*. Лишь в ауслате в позиции перед гласным последующего компонента часто представлен переход *n > б*, что полностью соответствует орфоэпическим нормам казахского языка. Например, *отуруб алды* ~ лит. каз. *отырып алды* ‘сел’; *атыб-еді* ~ лит. каз. *атып еді* ‘выстрелил’; *марылқаб-еді* ~ лит. каз. *жарылқан еді* ‘благодетельство-

вал'; *барыб-еді* ~ лит. каз. *барып еді* 'отправился'; *тартыб-алды* ~ лит. каз. *тартып алды* 'отобрал'; *көб-екән* ~ лит. каз. *көп екен* 'много оказывается'. В слове (перс.) *асман* ~ лит. каз. *аспан* 'небо' на месте ожидаемого *n* употребляется *m*, сходное с произношением персидского аналога. Напомним, что сказка заимствована казахами у персов при посредничестве сартов [Пантусов 1901: 3]. В тексте сказки употребляется еще одно персидское заимствование *дост* ~ лит. каз. *дос* 'друг', сохранившее свое фонетическое содержание.

2. В нескольких лексемах зафиксировано и другое нестандартное употребление *b* в «Кара-Мерген» в соответствии с лит. каз. у в однокоренных словах: *дабыс* ~ лит. каз. *дауыс* 'голос, звук'; *дабыстау* ~ лит. каз. *дауыстау* 'подавать голос'.

3. Согласный *s* обычно соответствует лит. каз. *s*. Исключение составляют слова, где на месте ожидаемых 1) *s* употребляется *ш*: *шалбырап* ~ лит. каз. *салбырап* 'свисая'; 2) *ш* передается *c*: (перс.) *патса* ~ лит. каз. *патша* 'царь'. Отметим, что переход *ш* > *c* наличествует во всех ранее исследованных кириллических памятниках на казахском языке в персидских и арабских заимствованиях.

4. Согласный *y* в тексте сказки стандартно соответствует казахскому литературному *y*, но может иногда встречаться и на месте литературного *ж* в начале второго компонента сложного слова: *kelä-jatyr* ~ лит. каз. *келе жатыр* 'идет'; *kelä-jatca* ~ лит. каз. *келе жатса* 'когда шел'. Такие факты были зафиксированы в говоре жителей южных областей Казахстана. Так, в речи казахов Чимкентской, Туркестанской, Кызылординской областей сохранился *j* в инлауте, так же, как и в «Кара-Мерген», когда предыдущая основа заканчивалась на *-l*: *bujaq* ~ лит. каз. *bul zaq* 'эта сторона'; *sojaq* ~ лит. каз. *sol zaq* 'та сторона'; *sojerde* ~ лит. каз. *sol zerde* 'в том месте'; *kele jatyr* ~ лит. каз. *kele zatyr* 'идет' [Қалиев, Сарыбаев 1991: 48].

5. Согласный *l/l* имеет, видимо, палатально-велярные варианты в зависимости от качества предыдущего гласного. Вариант *l* регулярно встречается после переднерядных гласных: *ā, e, ä, ü, i, ö*. Например: *el* 'народ'; *mižälän* 'став на колени'; *däylämi* 'богатство его'; *är tüplü* 'различный'; *ülkөн* 'большой'; *estildü* 'послышалось'; *kök köйлөкті* 'в синем платье'. В сочетании с заднерядными *a, ы, o, ұ*, у реализуется через *l*: *бала* 'ребенок'; *мылтыбын* 'ружье его'; *болупту* 'был'; *тұбалы* 'с рождения'; *шуркулдады* 'захлопотал'. Приведенные примеры полностью соответствуют произносительным нормам современного казахского языка.

6. Характерной особенностью *n* является то, что перед губным согласным *b* сонорный *n* чередуется с *m*: *көрүмбөйдү* ~ лит. каз. *көрінбейді* 'не видно'; *көрүмбөдү* ~ лит. каз. *көрінбеді* 'не виднелось'. Перед звонкими согласными, среднеязычным *z* и увулярным *ʒ*, альвеолярный сонорный *n* произносится как увулярный сонорный *ŋ*: *буруңбу* ~ лит. каз. *бұрынғы* 'давний'; *өзөңгө* ~ лит. каз. *өзенге* 'к реке'; *күңгө* ~ лит. каз. *күнге* 'на солнце'; *куаңбандай* ~ лит. каз. *куангандай* 'словно обрадовался'; *мергәңгә* ~ лит. каз. *мергенге* 'охотнику'. Указанные факты свойственны и произношению носителей казахского литературного языка.

4. Другие особенности

1. Выпадение согласных звуков: *усап* ~ лит. каз. *ұқсап* 'стать похожим'; *усатып* ~ лит. каз. *ұқсатып* 'уподобляя'; *өңүр тас* ~ лит. каз. *үңгір тас* 'каменная пещера'; (араб.) *екәжат* ~ лит. каз. *хикаят* 'диво'; (араб.) *ацајып* ~ лит. каз. *гажайып* 'дивный'.

2. На стыке слов в составе одной ритмической группы конечный гласный первого компонента поглощается начальным гласным второго: *ат' алмады* ~ лит. каз. *ата алмады* 'не смог выстрелить'; *қан сақтайд' екен* ~ лит. каз. *жан сақтайды екен* 'существует оказывается'; *марылқамас-м' екен?* ~ лит. каз. *жарылқамас па екен?* 'не ниспошлет ли?'; *ос' екен* ~ лит. каз. *осы екен* 'это оказывается'. Такая закономерность свойственна не только казахскому языку, но другим тюркским языкам.

3. В источнике, пусть и в небольшом количестве, встречается диалектная лексика: *шылапшын* 'таз'; *сүйгү* 'полотенце'.

5. Морфологические особенности

В источнике представлены два случая сохранения архаической рефлексии аффикса **аккузатива** *-ны* после сонанта *m* притяжательного аффикса 1 лица ед. ч.: *сыјымны* ~ лит. каз. *сыйымды* 'почтение мое'; *тамабымны* ~ лит. каз. *тамағымды* 'еду мою'. Редкие случаи отклонения в аккумулятиве встречались также в западных [Премудрости 1891] и северных [Букварь 1892] диалектных книгах на казахском языке. Так, формант аккумулятива после сонорного *y* и звонких *z, ж* там был выражен аффиксами *-ны/-ні* вместо *-ды/-ді*. Например: *Кудайны* ~ лит. каз. *Құдайды* 'Бога'; *патсамызны* ~ лит. каз. *патшамызды*

‘царя нашего’. Вероятно, это показывает, что переход $n > d$ в аффиксе аккузатива еще не был окончательно завершен в XIX в.

Наряду с этим в «Кара-Мерген» выявлен один случай особой формы аффикса генитива: в основе на $-ң$ аффикс имеет форму $-дың$ в соответствии с лит. каз. $-ның$: $аңдың \sim$ лит. каз. $аңның$ ‘животного’. В сказке это единственная лексема с основой на $-ң$, поэтому проверить гипотезу о связи формы аффикса с определенным типом основ не представляется возможным. Схожая тенденция употребления аффикса генитива $-дың$ в основах на носовой согласный была отмечена и в книге [Премудрости 1891]: в основах на $-н$ аффикс генитива реализуется как $-дың$ вместо ожидаемого $-ның$. Например, $катындың$ ‘женщины, жены’, $замандың$ ‘времени, эпохи’, $аспандың$ ‘неба’. Наряду с этим встречаются формы, соответствующие литературной норме: $күннің$ \sim лит. каз. $күннің$ ‘солнца’; $катынның$ \sim лит. каз. $қатынның$ ‘женщины’ [Гаджиева 2020: 367].

В двух лексемах имеет место переход $л > d$ в корне слова и анлауте аффикса $-лық$: (араб.) $Алдаба$ N-DAT \sim лит. каз. $Аллаға$ ‘Богу’; $маңдық$ \sim лит. каз. $маңлық$ ‘шелковый’.

6. Выводы

Таким образом, сказка «Кара-Мерген» является образцом фольклора, записанным в юго-восточной части Казахстана, в Семиреченской области. Для отражения орфографии памятника была применена научная транскрипция, принятая в тюркских изданиях Императорской Академии наук и Императорского Казанского университета, в некоторой степени отличающаяся от графики, принятой в миссионерских изданиях Православного миссионерского общества.

Особенности отражения вокализма в «Кара-Мерген» и исследованных ранее кириллических памятниках на казахском языке можно представить в следующей сводной таблице:

Таблица 6. Особенности рефлексии казахских гласных в первых кириллических книгах

Лит. каз.	<i>i</i>	<i>ы</i>	<i>ө</i>	<i>ү</i>	<i>o</i>	<i>ұ</i>	<i>e</i>	<i>ә</i>	<i>a</i>
Кара-Мерген, 1901	<i>i</i>	<i>ы</i>	<i>ö/(j)</i>	<i>ÿ/(ö)</i>	<i>o</i>	<i>y/o</i>	<i>e/ä/i</i>	<i>ä/a</i>	<i>a</i>
[Премудрости 1891]	<i>u/-/e/э</i>	<i>ы/-</i>	<i>ö/ÿ</i>	<i>ÿ/ö</i>	<i>o/(y)/(ö)</i>	<i>y/o/ÿ/(ö)</i>	<i>e/э/u/(ä)</i>	<i>ä/a/e</i>	<i>a/ä</i>
[Училище 1892]	<i>u/-</i>	<i>ы</i>	<i>ö/(j)</i>	<i>ÿ/ö</i>	<i>o</i>	<i>y</i>	<i>e</i>	<i>ä</i>	<i>a</i>
[Букварь 1892]	<i>u/-/e</i>	<i>ы</i>	<i>ö/(j)</i>	<i>ÿ/(ö)</i>	<i>o</i>	<i>y</i>	<i>e</i>	<i>ä</i>	<i>a</i>
[Крещение 1892]	<i>u/-/(ы)</i>	<i>ы</i>	<i>ö/o</i>	<i>ö/ÿ(#)</i>	<i>o</i>	<i>y/o(#)</i>	<i>e/u</i>	<i>ä/a</i>	<i>a</i>
[Букварь 1908]	<i>u</i>	<i>ы</i>	<i>ö</i>	<i>ÿ/ö</i>	<i>o/y</i>	<i>y</i>	<i>e/(a)/(u)</i>	<i>ä/(a)</i>	<i>a/(ä)</i>

Особенностью рассматриваемого источника является позиционная реализация гласного звука $e/э$, транскрибируемого как e в анлауте и $ä$ в инлауте и ауслауте. В других первых книгах $a/ä$ в качестве рефлекса e встречается крайне редко, не более трех раз на весь текст. В абсолютном начале слова этот гласный произносится как дифтонг [je], что не противоречит орфоэпическим нормам литературного казахского языка.

Анализ кратких гласных $ы$, i показал, что в «Кара-Мерген», в отличие от исследованных нами ранее кириллических книг на казахском языке ([Премудрости 1891] и [Букварь 1892]), случаи выпадения $ы$, i в первом слоге не представлены.

Как видно из таблицы, в целом вокализм, представленный в «Кара-Мерген», не отличается значительно от других первых книг. Более интересно оценивать частоту того или иного рефлекса. Например, соответствие $ö \sim$ лит. каз. $ү$ встречается в «Кара-Мерген», [Премудрости 1891; Училище 1892]. Но в «Кара-Мерген» оно представлено лишь в нескольких лексемах, а в [Премудрости 1891] этот тип рефлексии самый частотный.

Обзор примеров, иллюстрирующих тенденции губного сингармонизма в «Кара-Мерген», показал, что при y , $ÿ$, $ö$ первого слога лабиализация узких гласных второго, третьего и четвертого слога зафиксирована на письме во всех случаях. После o в первом слоге встречаются случаи сохранения $ы$. В целом можно отметить, что система сингармонизма «Кара-Мерген» отличается как от литературного казахского языка, так и от изученных ранее первых книг. Это подтверждает тот факт, что ранее учеными не были изучены первые книги на Семиреченском диалекте.

Анализ суффикса аккузатива, архаичный вариант которого ($-ны$) спорадически сохраняется как в «Кара-Мерген», так и в западных [Премудрости 1891] и северных [Букварь 1892] диалектных книгах, показывает, что его изменение, видимо, еще не завершилось в конце XIX в.

Сокращения

Общие

араб. — арабский язык

лит. каз. — литературный казахский язык

перс. — персидский язык

Глоссы

1 — первое лицо

3 — третье лицо

ABL — исходный падеж (аблатив)

ACC — винительный падеж (аккузатив)

ADV.PTCP — деепричастие

ADJ — имя прилагательное

COMPAR — аффикс сравнения и уподобления

GEN — родительный падеж (генитив)

INS — совместный падеж (инструменталис)

LOC — местный падеж (локатив)

N — имя существительное

NUM — имя числительное

PL — множественное число

POSS — притяжательный аффикс

PRON — местоимение

PST — прошедшее время

PTCP — причастие

SG — единственное число

Литература

Аманжолов 1997 — С. А. Аманжолов. Вопросы диалектологии и истории казахского языка. Алматы, 1997.

Болатов 1963 — Ж. Болатов. Шығыс Қазақстан облысы Большенарым селолық ауданында тұратын казактар тіліндегі жергілікті ерекшеліктер туралы (О местных особенностях в языке казахов Большенарымского сельского района Восточно-Казахстанской области) // Вопросы истории и диалектологии казахского языка. Вып. 5. Алма-Ата, 1963. С. 177—230.

Букварь 1892 — Букварь для киргизов. Жазуга үйрететін кнеге. Казань, 1892.

Букварь 1908 — Букварь для киргиз. Букуар. Казань, 1908.

Гаджиева 2019 — А. А. Гаджиева. Анализ системы вокализма и консонантизма в первых кириллических книгах на казахском языке // Урало-алтайские исследования. 2019, 3 (34). С. 53—71.

Гаджиева 2020 — А. А. Гаджиева. Морфонологические особенности первых кириллических книг на казахском языке (на материале книг «Из премудрости Иисуса сына Сирахова» (1891) и «Букварь для киргиз» (1908)) // Вестник Башкирского университета. 2020. Т. 25. № 2. С. 366—371.

Горбуров, Горбуров 2016 — Е. Горбуров, К. Горбуров. Николай Николаевич Пантусов (1849—1909) // Емінак. 2016, 3 (15). Т. 2. С. 118—123.

Дыбо, Норманская 2016 — А. В. Дыбо, Ю. В. Норманская. Первые кириллические книги на казахском языке как источники для изучения истории диалектов и создания литературной нормы // Урало-алтайские исследования. 2016, 4 (23). С. 138—152.

Қалиев, Сарыбаев 1991 — Ф. Қалиев, Ш. Ш. Сарыбаев. Қазақ диалектологиясы (Казахская диалектология). Алматы, 1991.

Кенесбаев 1962 — С. К. Кенесбаев. Современный казахский язык: лексика, фонетика. Алма-Ата, 1962.

Кенесбаев, Мусабаев, Неталиева 1973 — С. К. Кенесбаев, Г. Г. Мусабаев, Х. Неталиева. Орфография казахского языка // Орфография тюркских литературных языков СССР. М., 1973. С. 103—117.

Крещение 1892 — Ауліе княз Владимир тура динге бзи де крип; орыс калкын-да кргизгени. Крещение Руси на киргизском языке. Казань, 1892.

Пантусов 1900 — Н. Н. Пантусов. Киргизские пословицы, записанные в Копальском уезде Семиреченской области. Казань, 1900.

Пантусов 1901 — Н. Н. Пантусов. Материалы к изучению казак-киргизского наречия. Киргизская сказка о Кара-Мергене. Вып. 3. Казань, 1901.

Пантусов 1909 — Н. Н. Пантусов. Образцы киргизской народной литературы, собранные и переведенные Н. Н. Пантусовым // Известия общества археологии, истории и этнографии при Императорском казанском университете. Т. 25. Вып. 1. Казань, 1909.

Премудрости 1891 — Акыл бере тугун кнеге. Из Премудрости Иисуса сына Сирахова. Казань, 1891.

Радлов 1870 — В. В. Радлов. Образцы народной литературы тюркских племен, живущих в Южной Сибири и Дзунгарской степи. Киргизское наречие. Ч. 3. СПб., 1870.

Словарь 1897 — Киргизско-русский словарь. Оренбург, 1897.

Училище 1892 — Езгиликке үйрететін кнеге. Училище благочестия на киргизском языке. Казань, 1892.

References

- Amanzholov 1997 — S. A. Amanzholov. Voprosy` dialektologii i istorii kazaxskogo yazy`ka. Almaty`, 1997. {S. A. Amanzholov. Issues of dialectology and the history of the Kazakh language. Almaty, 1997.}
- Bolatov 1963 — Zh. Bolatov. Shy`gy`s Qazaqstan obly`sy` Bol`shenary`m seloly`q audany`nda turaty`n qazaqtar tilindegi zhergilikti erekshelikter turaly` (O mestny`x osobennostyax v yazy`ke kazaxov Bol`shenary`mskogo sel`skogo rajona Vostochno-Kazaxstanskoj oblasti. In: Voprosy` istorii i dialektologii kazaxskogo yazy`ka. Vy`p. 5. Alma-Ata, 1963. S. 177—230. {Zh. Bolatov. About local features in the Kazakh language of the Bolshenarym rural region of the East Kazakhstan region. In: Issues of dialectology and the history of the Kazakh language. Iss. 5. 1963. P. 177—230.}
- Bukvar` 1892 — Bukvar` dlya kirgizov. Zhazuga yjretetin knege. Kazan`, 1892. {Primer for the Kirghiz people. Zhazuga yjretetin knege. Kazan, 1892.}
- Bukvar` 1908 — Bukvar` dlya kirgiz. Bukuar. Kazan`, 1908. {Primer for the Kirghiz people. Kazan, 1908.}
- Gadzhieva 2019 — A. A. Gadzhieva. Analiz sistemy` vokalizma i konsonantizma v pervy`x kirillicheskih knigax na kazaxskom yazy`ke. In: Uralo-altajskie issledovaniya. 2019, 3 (34). S. 53—71. {A. A. Gadzhieva. Analysis of vowel and consonant system in the first Cyrillic books in Kazakh. In: Ural-Altai studies. 2019, 3 (34). P. 53—71.}
- Gadzhieva 2020 — A. A. Gadzhieva. Morfonologicheskie osobennosti pervy`x kirillicheskih knig na kazaxskom yazy`ke (na materiale knig “Iz premudrosti Iisusa sy`na Siraxova” (1891) i “Bukvar` dlya kirgiz” (1908)). In: Vestnik Bashkirskogo universiteta. T. 25. № 2. 2020, S. 366—371. {A. A. Gadzhieva. Morphophonemic features of the first Cyrillic books in the Kazakh language (on the material of the books “Wisdom of Jesus the son of Sirach” (1891) and “Primer for Kyrgyz People” (1908)). In: Bulletin of Bashkir University. Vol. 25. № 2. 2020. P. 366—371.}
- Gorburov, Gorburov 2016 — E. Gorburov, K. Gorburov. Nikolaj Nikolaevich Pantusov (1849—1909). In: Eminak. 2016, 3 (15). T. 2. S. 118—123. {E. Gorburov, K. Gorburov. Nikolaj Nikolaevich Pantusov (1849-1909). In: Eminak. 2016, 3 (15). Vol. 2. P. 118—123.}
- Dy`bo, Normanskaya 2016 — A. V. Dy`bo, Yu. V. Normanskaya. Pervy`e kirillicheskie knigi na kazaxskom yazy`ke kak istochniki dlya izucheniya istorii dialektov i sozdaniya literaturnoj normy`. In: Uralo-altajskie issledovaniya. 2016, 4 (23). S. 138—152. {A. V. Dybo, Yu. V. Normanskaya. First Cyrillic books in Kazakh as a base for studying the history of dialects and creating literary standard. In: Ural-Altai Studies. 2016, 4 (23). P. 138—152.}
- Qaliev, Sary`baev 1991 — Ğ. Qaliev, Sh. Sh. Sary`baev. Qazaq dialektologiyasy` (Kazaxskaya dialektologiya). Almaty`, 1991. {G. Kaliev, Sh. Sarybaev Kazakh dialectology. Almaty, 1991.}
- Kenesbaev 1962 — S. K. Kenesbaev. Sovremenny`j kazaxskij yazy`k: leksika, fonetika. Alma-Ata, 1962. {S. K. Kenesbaev. Modern Kazakh language: lexicology, phonetics. Alma-Ata, 1962.}
- Kenesbaev, Musabaev, Netalieva 1973 — S. K. Kenesbaev, G. G. Musabaev, X. Netalieva. Orfografiya kazaxskogo yazy`ka. In: Orfografiya tyurkskix literaturny`x yazy`kov SSSR. Moskva, 1973. S. 103—117. {S. K. Kenesbaev, G. G. Musabaev, H. Netalieva. Orthography of the Kazakh language. In: Orthography of the Turkic literary languages of the USSR. Moscow, 1973. P. 103—117.}
- Kreshhenie 1892 — Aulie knyaz Vladimir tura dinge bzi de krip; ory`s kalky`n-da krgizgeni. Kreshhenie Rusi na kirgizskom yazy`ke. Kazan`, 1892. {The Baptism of Russia in the Kirgiz language. Aulie knyaz Vladimir tura dinge bzi de krip; ory`s kalky`n-da krgizgeni. Kazan, 1892.}
- Pantusov 1900 — N. N. Pantusov. Kirgizskie poslovice, zapisanny`e v Kopal`skom uезде Semirechenskoj oblasti. Kazan`, 1900. {N. N. Pantusov. Kyrgyz proverbs collected in the Kopal'sky county of the Semirechensk region. Kazan, 1900.}
- Pantusov 1901 — N. N. Pantusov. Materialy` k izucheniyu kazak-kirgizskogo narechiya. Kirgizskaya skazka o Kara-Mergene. Vy`p. 3. Kazan`, 1901. {N. N. Pantusov. Materials for studying the Kazakh-Kyrgyz dialect. Kyrgyz fairy tale about Kara-Mergen. Iss. 3. Kazan, 1901.}
- Pantusov 1909 — N. N. Pantusov. Obrazcy kirgizskoj narodnoj literatury`, sobranny`e i perevedenny`e N. N. Pantusovy`m. In: Izvestiya obshhestva arxeologii, istorii i e`tnografii. T. 25. Vy`p. 1. Kazan`, 1909. {Samples of Kyrgyz folk literature collected and translated by N. N. Pantusov. In: Proceedings of the Society of Archaeology, History and Ethnography attached to Imperial Kazan University. Vol. 25. Iss. 1. Kazan, 1909.}
- Premudrosti 1891 — Aky`l bere tugun knege. Iz Premudrosti Iisusa sy`na Siraxova. Kazan`, 1891. {Aky`l bere tugun knege. Wisdom of Jesus the Son of Sirach. Kazan, 1891.}
- Radlov 1870 — V. V. Radlov. Obrazcy narodnoj literatury` tyurkskix plemen, zhivushhix v Yuzhnoj Sibiri i Dzungarskoj stepi. Kirgizskoe narechie. Ch. 3. SPb., 1870. {V. V. Radlov. Turkic Tribes of Southern Siberia and Dzungar Steppe: Samples of Folk Literature. Part 3: Kirghiz Dialect. St. Petersburg, 1870.}
- Slovar` 1897 — Kirgizsko-russkij slovar`. Orenburg`, 1897. {Kyrgyz-Russian Dictionary. Orenburg, 1897.}
- Uchilishhe 1892 — Ezgilikke yjretetin knege. Uchilishhe blagochestiya na kirgizskom yazy`ke. Kazan`, 1892. {School of piety in the Kirghiz language. Kazan, 1892.}

Распределение фонем в татарском тексте: опыт квантитативного исследования

Галиева Альфия Макаримовна, ООО «НекстГИС» (Москва);
amgalieva@gmail.com

Наряду с качественным описанием фонем, представленным в татарских грамматиках, требуется их квантитативное исследование, выполненное на репрезентативных выборках текстовых данных. В данной работе анализируется распределение фонем и фонемных сочетаний в повести Амирхана Еники «Әйтелмәгән васыятъ» («Невысказанное завещание»). Приводятся абсолютные и нормализованные частоты (с использованием квантилей стандартного нормального распределения) для каждой из 34 фонем, выделенных в тексте. Использованный подход к нормализации данных (вычисление z -статистик) основан на учете сведений как об общем количестве употреблений каждой фонемы, так и о распределении фонем в зависимости от их позиции в словоформе (в начальной, серединной и финальной позиции), что позволяет не терять информации о позиционных особенностях низкочастотных фонем. Величина отклонения z -статистик от квантилей стандартного нормального распределения, находящихся в интервале $(-1,96, 1,96)$, свидетельствует о степени упорядоченности, характере системной организации фонем в языке.

Синтагматическая составляющая фонологической системы языка, так же, как и инвентарь фонем, является уникальной для каждого языка. Поэтому далее исследуется сочетаемость татарских фонем на примере двух- и трехфонемных сочетаний. Квантитативная фонологическая информация дополняется грамматической интерпретацией; в частности, показано, что особенности распределения частотных типов фонемных сочетаний связаны со структурой словоформы и определяются агглютинативным строем татарского языка.

Ключевые слова: фонема, квантитативная лингвистика, татарский язык, сочетаемость фонем, фонология и грамматика

PHONEME DISTRIBUTION IN A TATAR TEXT: A QUANTITATIVE APPROACH

Alfiya M. Galieva, NextGIS Company (Moscow);
amgalieva@gmail.com

Along with the qualitative description of phonemes which is presented in Tatar grammars, their quantitative description is often required. Such a research should be performed on a sufficient volume of text data which would contain lexical items in their natural surroundings and reflect basic regularities of the language system.

This paper analyzes the distribution of phonemes and phoneme combinations in the story by Amirkhan Yeniki “Äytmägän vasiyät” (“The unspoken testament”). The research material was the text converted to a phonologically relevant form. The absolute and relative frequencies (using the quantiles of the standard normal distribution) are provided for each of the 34 phonemes distinguished in the text. The approach to data normalization (z -statistics) is based on information about both the number of phonemes depending on their position (wordinitial, wordfinal or wordmedial), and the total number of occurrences of each phoneme in the text; this allows us to retain information about positional features of low frequency phonemes and to find individual tendencies for each cell in the contingency table. The deviation of the z -statistics from the quantiles of the standard normal distribution $(-1.96, 1.96)$ indicates the character of phoneme sequences in the language. It is shown that for the absolute majority of Tatar phonemes, the z -statistics are far beyond the interval $(-1.96, 1.96)$, with the exception of borrowed phonemes $[ɹ]$ and $[ɹʃ]$, as well as the low-frequency phoneme $[ʒ]$, which indicates the random nature of the distribution of these phonemes in Tatar.

The syntagmatic component of the phonological system, just like the inventory of phonemes, is unique for each language, so without such study the phonological description would be incomplete. That is why a special section examines Tatar phoneme sequences using the example of two and three phoneme combinations. Since there is often a gap in grammar books between the descriptions of the phonological and morphological levels, quantitative information on phonemes is complemented with grammatical interpretation; in particular, the author shows that high frequency combinations of phonemes are associated with affixes and stems, which is determined by the agglutinative structure of the Tatar language.

Key words: phoneme, quantitative linguistics, the Tatar language, combinations of phonemes, phonology and grammar

Введение

В значительной части классических и современных исследований по количественной лингвистике анализируются распределения языковых единиц по длине, структуре слога, распределения единиц текста и корпуса по частям речи и т. п., обосновываются и оттачиваются статистические процедуры анализа языковых данных. При этом эмпирические данные по языкам разных типов часто дополняются теоретическими ожидаемыми значениями, полученными при помощи той или иной модели. Так, для моделирования частот по длине слова используются различные модификации распределения Пуассона [Altmann 2013; Kromer 2001; Best 2008; Best 2011 и др.], закона Ципфа [Altmann 2002; Kromer 2002], закон Мензарата-Альтмана [Altmann 1980]; при моделировании частотности типов слоговых структур хорошие результаты дает использование функции Ципфа-Алексеева [Zörnig et al. 2019], распределения Ципфа-Мандельброта [Mandelbrot 1965; Radojičić et al. 2019; Galieva, Vavilova 2021] и т. п. В частности, журнал «Glottometrics» во многом специализируется на подобном моделировании единиц языка и текста на материале языков разных типов. Количественная лингвистика вносит существенный вклад в развитие теории языка как за счет внедрения новых теоретических конструкций, так и обеспечивая эпистемологический каркас на первый взгляд случайными эмпирическими данными, обнаруживая в них тенденции и закономерности.

Очевидно, что фонологическая система языка, блестяще описанная структуралистами и их последователями, представляющими разные традиции [Jakobson, Halle 1956; Jakobson, Waugh 1979; Chomsky, Halle 1968; Трубецкой 2000 и др.], является хорошим объектом для количественных исследований: текстовый инвентарь фонем поддается различным классификациям и статистическим оценкам, и неслучайно именно исследователи фонологической системы языка высказали мнение о неравномерном распределении фонем в зависимости от их позиции в словоформе [Байчура 1953; Щербак 1970; Jakobson, Waugh 1979; Altmann 2005 и др.].

Тем не менее исследований по количественной фонологии относительно немного, что легко объясняется тем, что фонематическая транскрипция требует большого объема рутинной ручной работы. Обобщенные данные по инвентарю фонем для 2186 языков мира можно получить в базе данных хранилища кросс-лингвистических фонологических данных PHOIBLE, разработанном в Институте эволюционной антропологии Макса Планка [Moran, McCloy 2019]. Инвентарь фонем стандартного татарского языка, представленный в репозитории PHOIBLE, включает 43 единицы.

Данные из репозитория PHOIBLE в дальнейшем могут стать материалом для количественных исследований. В частности, работа С. Брауна показывает, что, согласно данным 451 языка, распределение фонем положительно смещено (*positively skewed*) и может быть аппроксимировано при помощи логнормального распределения [Brown 2017]. При моделировании фонологических данных используются также регрессионные методы; так, в статье Г. Коломы регрессия применена к материалу 100 языков с четырьмя фонологическими переменными (количество согласных, количество гласных, различимость ударения, различимость тона). Предложенный метод основан на принципах синергетической лингвистики: автор говорит о том, что уравнения регрессии могут быть выведены путем решения задачи оптимизации, предполагающей, что уровни фонологических переменных в языке должны быть выбраны таким образом, чтобы максимально упростить декодирование и минимизировать затраты при звукопроизводстве [Coloma 2014].

Б. Кесслер использует количественный подход к исторической фонологии. Исследование включает в себя методы обнаружения и проверки звуковых соответствий, моделирования хода изменения фонем; производится вычисление наиболее вероятного генетического дерева, совместимого с набором нововведений, проверка значимости фонетических свидетельств для определения генетических взаимосвязей между языками и диалектами [Kessler 2015].

Количественные оценки фонологических явлений исследуются также в прикладном аспекте изучения второго языка, результаты таких исследований находят применение в совершенствовании методики преподавания неродных языков [Thomson 2016].

В настоящее время эмпирический материал по фонологии могут обеспечить размеченные речевые корпуса, которые представляют собой специализированные базы данных, состоящие из аудиофайлов и транскрипций текстов [Гришина 2005; Гришина, Савчук 2009; Bogdanova-Beglarian et al. 2017; Bogdanova-Beglarian et al. 2018 и др.]; именно данные из речевых корпусов часто становятся объектом фонетических и фонологических исследований [Turunen, Lipping 2008 и др.]. В статье Н. С. Смирновой и П. Г. Чистикова описаны основные возможности использования статистического анализатора текста TextAnalyser в сфере речевых технологий и приведены некоторые результаты статистической обработки большого текстового корпуса, в частности, представлены ранжированные по частотности списки аллофонов русских фонем и двухфонемных сочетаний [Смирнова, Чистиков 2011].

В татарских грамматиках [Татарская грамматика 1993; Татар грамматикасы 2015] фонемный уровень языка представлен без количественных данных, что легко объяснить тем, что на момент разработки грамматик отсутствовали достаточные массивы подготовленных фонологических данных. Тем не менее для татарского языка достаточно качественный материал для квантитативных исследований могут дать предварительно обработанные письменные тексты.

В данной работе анализируется распределение фонем в татарском тексте на материале повести Амирхана Еники «Эйтелмәгән васыять» («Невысказанное завещание»). Поскольку часто в грамматике языков имеется зазор между описаниями фонологического и морфологического уровней, квантитативные данные дополняются комментариями из области грамматики.

Краткие сведения о татарской фонологии

В академических грамматиках татарского языка говорится о том, что в современном татарском языке имеется 12 гласных фонем, 9 из которых — исконные: $|a|$, $|\partial|$, $|y|$, $|y'|$, $|yl|$, $|e|$, $|o|$, $|\theta|$, $|u|$. При этом выделяются 3 дополнительные гласные фонемы, встречающиеся в русских и европейских заимствованиях: $|\bar{o}|$, $|\bar{e}|$, $|\bar{e}'|$ [Татарская грамматика 1993: 88, Татар грамматикасы 2015: 43]. 9 исконных фонем и 3 отдельные фонемы для заимствованных слов выделяются и в работах ряда западных исследователей (например, [Berta 2015]). Долгота гласных не является дифференциальным признаком татарских фонем. Следует отметить, что существуют и другие подходы к выделению фонем в татарском языке: например, в работах В. А. Богородицкого [Богородицкий 1953], Н. А. Баскакова [Баскаков 1988], Г. Бурбиела [Burbiel 2018] выделяются 10 фонем; в частности, В. А. Богородицкий (а за ним и другие исследователи) дополнительно выделяет дифтонгоидный $|y\bar{y}'|$ в качестве отдельной фонемы [Богородицкий 1953: 28].

Систему консонантизма образуют 25 исконных согласных фонем и 3 добавочные фонемы, употребляющиеся в заимствованиях: $|u|$, $|ul|$, $|\beta|$ [Татарская грамматика 1993: 91; Татар грамматикасы 2015: 46].

Относительно позиционных вариантов фонем нужно отметить следующее: $|a|$ в инициальных слогах произносится как $|a^o|$; в грамматиках для фонем $|y|$, $|y'|$, $|u|$ в абсолютном конце слова приводятся позиционные варианты $|y^m|$, $|y'^m|$, $|u^l|$. Составителями грамматик отмечается, что бифонемность $|y|$, $|y'|$ в финальной позиции «доказана фонологическими, фонетическими и особенно экспериментально-фонетическими исследованиями» [Татарская грамматика 1993: 91; Татар грамматикасы 2015: 45].

В трудах В. А. Богородицкого, основавшего в 1884 году при Казанском университете первую в мире лабораторию экспериментальной фонетики, большое внимание уделялось вопросам изменения вокализма в зависимости от позиции гласных в словоформе, исследователь вскрывал тесную связь фонетики и фонологии с морфологией; к сожалению, многие идеи Богородицкого не находят отражения в современных академических грамматиках татарского языка, ориентированных на русские грамматики, в частности на «Русскую грамматику» 1980 года [Татар грамматикасы 2015: 43].

Бифонемная природа $|y|$ и $|y'|$ в позиции непервых слогов находит отражение и в исследованиях, вышедших за рубежом. Так, Н. Поппе говорит о том, что $|y|$, $|y'|$, $|u|$ в исконных словах не встречаются в финальных и неинициальных слогах, будучи заменены на соответствующие алломорфы бифонемного происхождения [Porre 1963: 9]. На различия $|y|$, $|y'|$, $|u|$ в позиции начального и неначальных слогов указывают и другие исследователи, например Г. Бурбиел [Burbiel 2018: 3].

Тезис о неравномерном распределении фонем в зависимости от их позиции в словоформе, высказанный в научной литературе, не только требует эмпирического подтверждения на текстовых данных, но и оценки, каковы эти смещения в ту или иную сторону, с использованием строгих статистических методов, поэтому основная часть данного исследования посвящена анализу распределения фонем в татарском тексте, причем за основу будут браться нормализованные значения и квантили нормального распределения.

Отбор языкового материала и подготовка данных

Часто в качестве материала для исследования фонологической или слоговой структуры слов того или иного языка берутся данные из лексикографических источников ([Altmann 2005; Мороз 2019]; и др.), что объясняется в первую очередь доступностью таких данных.

В данном исследовании мы не используем материалы татарских словарей, так как специфика отбора и представления форм слов в словарях во многом не согласуется с их реальным использованием в речи, что для татарского языка с его агглютинативной морфологией совершенно критично. Например, в словарях фиксируется большое количество заимствованных слов с нетипичной для татарского языка фо-

немной организацией, среди которых может оказаться очень большой процент редко используемых в текстах слов, что существенно искажает данные о распределении фонем. Кроме того, в словарях лексемы даются в основной форме (например, глаголы фиксируются в форме инфинитива или имени действия с их стандартными аффиксами), поэтому данные из лексикографических источников не отражают аффиксальных цепочек, которые связывают слова в предложениях, при этом они могут содержать непропорционально большое количество аффиксов инфинитива или имени действия. Все это приводит к тому, что распределения фонем, выделенных из единиц словаря, оказываются существенно смещены в ту или иную сторону (например, могут быть переоценены данные по относительно редким для текстов фонемам и фонемным сочетаниям и недооценены фонемы, типичные для аффиксов и аффиксальных цепочек). В данном исследовании нас интересует распределение фонем в словоформах в реальном употреблении, поэтому текстовый материал, представляющий образец классической татарской прозы XX века, мы считаем оптимальным. Этим обусловлен выбор эмпирической базы исследования — текста повести А. Еники «Невысказанное завещание».

Важным этапом в подготовке данных стало приведение кириллического текста к форме, близкой к фонологической: 1 символ — 1 фонема. Считается, что татарское письмо главным образом основано на фонетическом принципе, согласно которому слова пишутся так, как произносятся [Татарская грамматика 1993: 108]. Тем не менее имеется достаточно много разных случаев нарушения принципа 1 буква — 1 звук. Здесь нужно отметить 3 основных случая:

- непроизносимые символы (1 буква — 0 звуков), влияющие на произношение соседних букв: ь и ъ;
- буквы *e*, *ю*, *я* в определенных позициях (1 буква — 2 звука, причем ряд гласного определяется сингармонистическим классом слова, что приводит к вариантам);
- буквы *у* и *ү* после некоторых гласных (когда они произносятся как сонорный согласный (неслоговой гласный) *w*).

Кроме того, буква *в* может читаться как шумный согласный $|v|$ (в словах, заимствованных из русского или европейских языков) или сонант $|w|$ (в словах тюркского происхождения или в восточных заимствованиях) в зависимости от происхождения слов. В таблице 1 представлены примеры преобразования татарских словоформ из кириллической в фонологически релевантную форму.

Таблица 1. Примеры приведения словоформ к фонологическому виду

Графический облик словоформы	Транскрипция
<i>урман</i> 'лес'	урман
<i>ямьле</i> 'красивый, приятный'	јәмле
<i>аулау</i> 'охотиться'	аулау
<i>кияу</i> 'жених, зять'	кијәу
<i>егет</i> 'молодой человек, юноша'	јегет
<i>ел</i> 'год'	јыл

Важная оговорка: выделение фонем в данной работе основано на письменном тексте и выполнено машиной (без последующей ручной проверки). Как следствие, в классе гласных фонем отдельно не выделены 3 гласные фонемы, встречающиеся лишь в русских и европейских заимствованиях: $|\bar{o}|$, $|\bar{i}|$, $|\bar{e}|$. В составе согласных фонем не различались следующие классы фонем: увулярные и неувулярные варианты для $|z|$ и $|k|$. Впрочем, следует отметить, что количество фонологически неассимилированных заимствований в анализируемом тексте относительно небольшое, что не должно, на наш взгляд, существенно исказить данные. С учетом этого были выделены 34 фонемы (см. таблицу 2), позиционные варианты фонем не рассматривались. Таким образом, наш материал не претендует на исчерпывающее описание фонологической системы татарского языка, а лишь представляет ее приближенное подобие, условную модель. Тем не менее эта модель позволяет получить вполне адекватные сведения о распределении и комбинаторных свойствах татарских фонем.

Распределение фонем в зависимости от позиции в словоформе

Приведем предварительные сведения об анализируемом тексте. Изначально (после токенизации) в тексте повести А. Еники содержалось 12063 слова (4283 уникальных словоформы), которые были записаны в фонологической транскрипции. На рисунке 1 представлено количество словоформ в зависимости от длины. Самыми частотными (2450 единиц) являются слова, состоящие из 5 фонем.

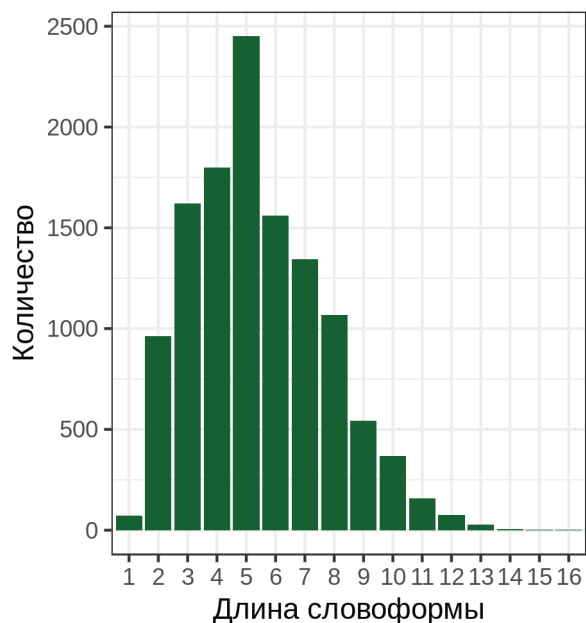


Рисунок 1. Количество словоформ в зависимости от их длины

Так как процедура анализа требовала подсчета числа фонем в начальной и финальной позиции и позиции в середине слова (сюда отнесены все фонемы за исключением крайних слева и справа), были удалены слова, состоящие из одной и двух фонем. В результате в анализируемом файле осталось 10944 словоформы (4202 уникальных словоформы), состоящие из 63102 фонем. Их распределение (абсолютные значения и z -статистики) в зависимости от позиции в словоформе представлены в таблице 2.

Таблица 2. Распределение фонем в зависимости от позиции в словоформе

Фонема	Значения в зависимости от позиции						Всего употреблений
	в начале слова		в конце слова		в середине слова		
	кол-во	z	кол-во	z	кол-во	z	
а	881	-11,43	1396	5,76	4778	4,51	7055
э	374	-15,64	980	9,59	2969	4,81	4323
б	1510	69,89	0	-20,80	487	-39,05	1997
в	54	8,80	3	-4,05	53	-3,78	110
w	19	-5,33	36	-2,80	261	6,47	316
г	331	1,33	0	-19,65	1457	14,57	1788
д	265	-1,23	3	-18,57	1367	15,75	1635
е	89	-29,11	1152	13,42	3468	12,48	4709
ж	0	-1,02	0	-1,02	5	1,63	5
з	86	-7,89	255	6,01	708	1,49	1049
и	569	4,99	210	-13,58	1946	6,84	2725
ј	601	14,22	356	-0,23	1118	-11,13	2075
к	2059	57,91	589	-5,01	1423	-42,08	4071
л	70	-25,63	183	-20,58	3457	36,75	3710
м	329	5,12	285	2,06	859	-5,71	1473
н	205	-24,47	1593	31,06	2911	-5,24	4709
ң	0	-12,66	470	32,79	285	-16,01	755
о	125	-2,13	10	-12,58	721	11,71	856
ө	146	7,11	0	-10,25	451	11,96	497
п	59	-11,63	1022	61,96	137	-40,03	1218
р	64	-26,83	990	13,49	2868	10,61	3922
с	626	23,59	41	-15,71	918	-6,27	1585
т	862	18,70	217	-13,99	1764	-3,75	2843
у	249	1,03	58	-12,83	1047	9,39	1354

Фонема	Значения в зависимости от позиции						Всего употреблений
	в начале слова		в конце слова		в середине слова		
	кол-во	z	кол-во	z	кол-во	z	
Ү	210	8,50	20	-10,35	487	1,47	717
Ф	18	-0,94	6	-3,76	103	3,74	127
х	188	20,13	0	-8,10	123	-9,57	311
ц	1	-1,32	0	-1,94	17	2,60	18
ч	275	8,13	99	-6,53	649	-1,27	1023
ш	321	9,52	72	-10,04	759	0,41	1152
щ	1	0,73	1	0,73	1	-1,16	3
ы	14	-31,15	894	4,96	3550	20,83	4458
ж	209	27,15	0	-7,36	48	-15,74	257
һ	133	22,43	2	-5,30	21	-13,62	156
Всего	10943	-	10943	-	41216	-	63102

Нормализованные значения z для фонем вычислены по методике, использованной в ряде исследований по квантитативной лингвистике [Altmann 2005: 193—194; Popescu et al. 2015: 32], когда значения для отдельной ячейки в таблице сопряженности вычисляются по формуле:

$$z = \frac{n_{ij} - E_{ij}}{\sqrt{\frac{n_{i\cdot} \cdot n_{\cdot j} (N - n_{i\cdot})(N - n_{\cdot j})}{N^2(N - 1)}}}$$

где z — квантиль стандартного нормального распределения $N(0,1)$, n_{ij} — количество фонем в ячейке (i, j) ; $n_{i\cdot}$ — сумма в строке i ; $n_{\cdot j}$ — сумма в колонке j , а N — общее количество фонем в тексте. Выражение в знаменателе — стандартное отклонение для значения в ячейке (i, j) . Ожидаемое значение E для ячейки (i, j) вычисляется по формуле:

$$E_{ij} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{N}$$

Z-статистика, вычисленная по формуле, приведенной выше, удобна тем, что при оценке частотности фонемы в той или иной позиции в слове учитывается как суммарное количество употреблений данной фонемы в тексте, так и суммарное количество фонем в каждой позиции; таким образом, мы можем оценить индивидуальные тенденции для каждой ячейки в таблице сопряженности. Если бы фонемы в тексте были распределены случайным образом, z-статистики для них должны были бы располагаться в интервале квантилей стандартного нормального распределения $(-1,96, 1,96)$; следовательно, величина отклонения реальных z-статистик от данного интервала свидетельствует о степени упорядоченности, о характере системной организации фонем в языке.

В частности, для фонемы $|\delta|$ мы имеем 1510 употреблений в начальной позиции, что соответствует значению 69,89 после нормализации. Так как 69,89 намного больше, чем $|z_{0,05}| = 1,96$, то можем заключить, что фонема $|\delta|$ имеет очень сильную тенденцию к использованию в начале слова. Для фонемы $|\delta|$ в конце слова мы имеем абсолютное значение 0 и нормализованное значение $-20,8$, то есть очень сильную отрицательную тенденцию. Более того, можно даже говорить о том, что фонема $|\delta|$ не может быть использована в абсолютном конце слова в словоформах татарского языка. В анализируемом тексте не встречаются арабские заимствования, оканчивающиеся на букву δ , типа *табиб* ‘врач’. Впрочем, о нетипичности таких слов, об оглушении $|\delta|$ в речи свидетельствует присоединение к ним глухих вариантов словоизменяемых аффиксов:

- табибка* (врач-DIR) ‘к врачу’,
- табибта* (врач-LOC) ‘у врача’,
- табибтан* (врач-ABL) ‘от врача’.

Именно такие варианты словоизменения мы обнаруживаем в Татарском национальном корпусе «Туган тел» [Татарский национальный корпус 2022].

Также очень высокую z-статистику для начала слова имеет фонема $|\kappa|$, а именно 57,91 при 2059 употреблениях в абсолютном начале слова.

Для фонемы $|\zeta|$ в начальной позиции нормализованное значение равно 1,33, что располагается в интервале $(-1,96, 1,96)$ и свидетельствует о том, что данная величина может быть случайной. Возможно,

низкая частотность данной фонемы в исконно татарских словах восполняется ее частотностью в начальной позиции в арабских заимствованиях, и влияние этих двух противоположных тенденций обеспечивает нейтральное значение *z*-статистики.

Анализ таблицы показывает, что лишь в достаточно редких случаях *z*-статистика располагается в интервале (–1,96, 1,96), в частности, это заимствованные фонемы |*y*| и |*ɯ*|, а также фонема |*ʒ*|, которая характеризуется очень низкой частотностью.

Таким образом, текстовый материал свидетельствует о том, что фонемы имеют явную тенденцию использоваться в той или иной позиции в тексте, отражая специфику организации фонологической системы языка.

Если выбор морфологических категорий зависит от способа выражения мысли, то фонологическая структура языка тесно связана с морфологическим строем слова [Баскаков 1988: 3]. Поэтому далее дадим небольшой грамматический комментарий к таблице 2.

Самой частотной фонемой является фонема |*a*|; фонемы |*a*| и |*ə*| используются в аффиксах множественного числа, в аффиксах падежей с пространственным значением (директив, аблатив, локатив), в аффиксах многих глагольных форм (инфинитива, будущего времени, перфекта, прошедшего времени и причастия на *-ган* и др.).

Фонемы |*ɨ*| и |*e*| используются в формах принадлежности у существительных, в аффиксах генитива и аккузатива, категорического прошедшего времени и др.

Фонемы |*l*| и |*p*| используются в аффиксе множественного числа, фонема |*l*| также часто используется во многих словообразующих аффиксах у слов разных частей речи (*-лы*, *-лык*, *-ла*, *-лан* и пр.).

Фонема |*n*| используется в аффиксах генитива и аккузатива, в пассивных и рефлексивных формах глаголов и т. п.

Фонема |*k*| — самая частотная в позиции в начале слова.

Очень большое значение *z*-статистики для фонемы |*n*| в финальной позиции, а именно 61,96, можно объяснить тем, что данная фонема используется в конвербах на *-n*, когда словоформа оканчивается аффиксом конверба.

Далее для наиболее частотных 8 фонем (4 гласных и 4 согласных) приведем относительные частоты употреблений в зависимости от позиции в слове (см. таблицу 3).

Таблица 3. Самые частотные фонемы в тексте

Фоне- ма	В начале		В конце		В середине	
	Кол-во	%	Кол-во	%	Кол-во	%
а	881	12,5	1396	19,8	4778	67,7
э	374	8,7	980	22,7	2969	68,7
ы	14	0,31	894	20,1	3550	79,6
е	89	1,9	1152	24,5	3468	73,6
л	70	1,9	183	4,9	3457	93,2
р	64	1,6	990	25,2	2868	73,1
н	205	4,4	1593	33,8	2911	61,8
к	2059	50,6	589	14,5	1423	35,0

То, что татарские фонемы по-разному распределены в абсолютном начале и абсолютном конце словоформ (см. абсолютные и относительные значения в таблицах 2 и 3), свидетельствуют о том, что в татарском языке фонемы, наряду с другими важными факторами (в частности, ударением и гармонией гласных), важны в когнитивных процессах определения границ словоформ. Любая словоформа представляет собой целостность, структуру, а опознавание структур предполагает их различия, а это возможно лишь в случае, если отдельные структуры различаются друг от друга известными признаками, ключ к которым дают именно фонемы [Трубецкой 2000: 41—42]. Любая словоформа существует в языке и «как структура всегда представляет собой нечто большее, нежели сумму его членов (= фонем)», но «в противоположность отдельным фонемам эта целостность не может быть локализована в звуковой оболочке слова» [Трубецкой 2000: 42], именно поэтому те стохастические, на первый взгляд случайные распределения фонем, которые раскрывает квантитативный анализ, возникают под действием внутренних законов, управляющих структурами языка.

Статистические характеристики фонем в процессах маркирования границ словоформ на массиве данных — один из интереснейших вопросов, возникших при исследовании распределения фонем в зависимости от позиции в словоформе, который требует специального исследования.

Анализ двухчленных и трехчленных сочетаний фонем

Системность всех уровней в тюркских языках во многом определяется тем, что «слово в тюркских языках по своему строению изоморфно словосочетанию, а фонологический состав слова изоморфен структуре самого слова» [Баскаков 1988: 3]. Важный аспект при квантитативном анализе фонемного состава языка — сочетаемость фонем. Синтагматическая составляющая фонологической системы языка — комбинаторные свойства фонем — является уникальной для каждого языка. В своей классической работе Н. С. Трубецкой говорит о том, что «...сочетания фонем подчиняются в любом языке своим особым законам или правилам, которые имеют значение только для данного языка и которые необходимо устанавливать для каждого языка отдельно» [Трубецкой 2000: 265]. Языки из всех логически возможных комбинаций фонем допускают лишь относительно небольшое количество двух-, трех- или n -членных сочетаний, образующих геометрию характерных признаков [Halle 1995].

В современной компьютерной лингвистике n -граммы являются важным статистическим инструментом, а информация о типах и частотности сочетаний фонем необходима для полноты описания фонологической системы языка.

Таблица 4 содержит сведения о 22 наиболее частотных двух- и трехфонемных сочетаниях из повести А. Еники. Далее в работе такие сочетания могут называться биграммами и триграммами соответственно.

Таблица 4. Биграммы и триграммы на материале фонем

Ранг	Биграммы	Количество	Триграммы	Количество
1	ар	1206	лар	453
2	ла	982	ала	325
3	ан	916	енә	308
4	ын	870	ләр	302
5	ен	869	ары	279
6	ка	837	ына	278
7	лә	782	кар	268
8	ән	709	елә	255
9	ал	674	иде	255
10	ер	640	бер	250
11	ул	607	аны	231
12	эр	591	ның	228
13	га	591	ган	212
14	на	586	лән	182
15	ак	582	бул	178
16	бе	582	гән	176
17	ны	575	бел	174
18	да	551	нда	172
19	ки	522	гын	169
20	ел	521	эре	168
21	ры	505	рга	167
22	нә	501	ара	166

Всего в тексте обнаружено 61065 бифонемных сочетаний, при этом количество уникальных биграмм относительно невелико — 685 (из логически возможных 1156 вариантов). Из них 574 бифонемные последовательности встречаются 3 и более раз. 50 наиболее частотных биграмм составляют свыше 43% всех бифонемных сочетаний.

Согласно таблице 4, большая часть высокочастотных биграмм имеет вид гласный + сонорный согласный или сонорный согласный + гласный. Между тем в общей таблице биграмм их удельный вес не настолько большой: биграммы со структурой гласный + сонорный согласный составляют 22,7 процентов, биграммы со структурой сонорный согласный + гласный — 16,3% от общего количества биграмм. Такое неравномерное распределение высокочастотных биграмм с сонорными можно объяснить тем, что сонорные согласные очень частотны в составе словоизменяемых и регулярных словообразующих аффиксов.

В таблице 5 приведены 15 наиболее частотных биграмм, встречающихся в начале слова, то есть это фонемные сочетания, которыми начинаются корни (основы) слов.

Таблица 5. Биграммы, частотные в начале слова

Ранг	Биграммы	Количество
1	ка	588
2	бе	465
3	ки	435
4	ба	346
5	бу	315
6	ја	272
7	кү	251
8	кы	239
9	би	233
10	да	219
11	ал	217
12	ке	212
13	шу	210
14	дә	208
15	ди	206
16	та	199
17	ул	189
18	ид	188
19	то	180

Приведем также краткий лингвистический комментарий к сочетаниям из трех фонем в анализируемом тексте. Выделено всего 42 334 трехфонемных сочетания, из них уникальных — 3038 (при логически возможных 39 304). 1902 триграммы встречаются 3 и более раз. 50 наиболее частотных триграмм составляют свыше 20% от их общего числа.

Интересно, что самое частотное сочетание из трех фонем совпадает с основным — заднерядным — алломорфом множественного числа *-лар* (триграмма, совпадающая с переднерядным вариантом этого аффикса *-лар*, имеет ранг 4). Поскольку, как отмечалось выше, фонемы $|\delta|$ и $|k|$ имеют очень сильную тенденцию употребляться в абсолютном начале слова, можно предполагать, что частотные трехфонемные сочетания *кар*, *бер*, *бул*, *бел* в большинстве случаев соответствуют основам слов, а не являются частями аффиксальных цепочек. Это отчасти подтверждается данными из таблицы 6, представляющей 12 самых частотных словоформ из текста.

Таблица 6. Частотные словоформы в тексте А. Еники

Ранг	Словоформа	Перевод	Количество	Длина
1	иде	служеб. гл.	165	3
2	бер	‘один’	124	3
3	белән	‘с’	122	5
4	бик	‘очень’	119	3
5	генә	‘только, лишь’	88	4
6	һәм	‘и’	85	3
7	аның	‘его/ее’	81	4
8	гына	‘только, лишь’	77	4
9	дип	цитатив	71	3
10	инде	‘уже, уж, же’	67	4
11	карчык	‘старушка’	65	6
12	кына	‘только, лишь’	58	4

Таким образом, сочетания фонем, представленные в татарском тексте, весьма разнообразны как по своей структуре, так и по возможности их реализации в разных позициях в пределах конкретных словоформ.

Вопрос о том, какие количественные характеристики фонем обусловлены языковой системой, а какие — выбором индивидуального текста, его стилиевой принадлежностью и пр., требует дальнейшего изучения на специально подготовленных выборках текстов.

Заключение

Для полноты картины, изображающей фонологическую систему татарского языка, наряду с качественным описанием фонем, представленным в татарских грамматиках, требуется их количественное описание на достаточно больших объемах лингвистических данных.

Данное исследование строилось на графическом текстовом материале, приведенном (с рядом оговорок, отмеченных выше) к фонологически релевантной форме 1 символ — 1 фонема. Несмотря на то, что использованный подход к фонологической транскрипции не дает исчерпывающего представления о фонемах татарского языка, он, тем не менее, позволяет получить вполне адекватные сведения о распределении и комбинаторных свойствах татарских фонем.

Проведенный количественный анализ фонемного состава татарского текста на материале повести А. Еники «Невысказанное завещание» подтвердил тезис о неравномерном распределении фонем в зависимости от их положения в словоформе — в начале, середине или конце слова. Для оценки смещения частот фонем были использованы нормализованные значения (z -статистики) и квантили нормального распределения.

Было показано, что татарские фонемы очень неравномерно распределены в зависимости от позиции в слове: для нормализованных величин, вычисленных при помощи z -теста, наблюдаются как очень большие, так и очень маленькие значения z -статистики, при этом лишь в отдельных случаях нормализованные значения лежат в интервале $(-1,96, 1,96)$, который требуется законами нормального распределения. Величина отклонения реальных z -статистик от данного интервала свидетельствует о степени неслучайности, о характере упорядоченности употребления фонем в языке.

Значения z находятся в интервале $(-1,96, 1,96)$ лишь у заимствованных фонем $|ц|$ и $|щ|$, а также у очень низкочастотной фонемы $|ж|$. Таким образом, возможные случайные распределения затрагивают лишь очень небольшое количество низкочастотных фонем, что не столько ломает глубоко упорядоченный строй татарской фонологической системы, сколько свидетельствует о ее относительной гибкости.

Правила сочетания фонем, так же, как и их инвентарь, являются уникальными для каждого языка. Анализ двух- и трехфонемных сочетаний показал, что сочетания с высоким рангом соответствуют частям аффиксов или основ слов. Таким образом, распределения фонем и фонемных сочетаний во многом определяются агглютинативным строем татарского языка.

Отметим, что данное исследование и его выводы во многом носят предварительный характер. Вопрос о статистических различиях в исконно тюркских и заимствованных словах (как из восточных, так и из западных языков), а также варьирование, обусловленное выбором текста, его стилистикой, тематикой и т. п., требуют дальнейшего изучения. В перспективе было бы также целесообразно провести аналогичное исследование на данных звукового корпуса.

Сокращения

ABL — аблатив

DIR — директив

LOC — локатив

Литература

- Байчура 1953 — У. Ш. *Байчура*. Гласные татарского языка в свете экспериментальных данных: автореф. дис. ... канд. филол. наук. М.; Л., 1953.
- Баскаков 1988 — Н. А. *Баскаков*. Историко-типологическая фонология тюркских языков / Отв. ред. Э. Р. Тенишев. М., 1988.
- Богородицкий 1953 — В. А. *Богородицкий*. Введение в татарское языкознание в связи с другими тюркскими языками. Казань, 1953.
- Гришина 2005 — Е. А. *Гришина*. Устная речь в Национальном корпусе русского языка // Национальный корпус русского языка: 2003—2005. М., 2005. С. 94—110.
- Гришина, Савчук 2009 — Е. А. *Гришина*, С. О. *Савчук*. Корпус устных текстов в НКРЯ: состав и структура // Национальный корпус русского языка: 2006—2008. Новые результаты и перспективы. СПб., 2009. С. 129—149.
- Мороз 2019 — Г. А. *Мороз*. Слоговая структура адыгейского языка: от данных к обобщениям // Вопросы языкознания. 2019, 2. С. 82—95.
- Смирнова, Чистиков 2011 — Н. С. *Смирнова*, П. Г. *Чистиков*. Программа анализа фонетических статистик в текстах на русском языке и ее использование для решения прикладных задач в области речевых технологий //

Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 25—29 мая 2011 г.). Вып. 10. М., 2011. С. 632—643.

Татарская грамматика 1993 — Татарская грамматика / Гл. ред. М. З. Закиев. Т. 1. Казань, 1993.

Татар грамматикасы 2015 — Татар грамматикасы / Ред. Ф. М. Хисамова. Т. 1. Казань, 2015.

Татарский национальный корпус 2022 — Татарский национальный корпус «Туган тел» // <http://tugantel.tatar/?lang=ru>, 2022.

Трубецкой 2000 — *Н. С. Трубецкой*. Основы фонологии / Пер. с нем. А. А. Холодовича; под ред. С. Д. Кацнельсона. М., 2000.

Щербак 1970 — *А. М. Щербак*. Сравнительная фонетика тюркских языков. Л., 1970.

Altmann 2002 — *G. Altmann*. Zipfian linguistics // *Glottometrics*. 2002, 3. P. 19—26.

Altmann 2013 — *G. Altmann*. Aspects of Word Length // *Issues in Quantitative Linguistics* / Eds. R. Köhler, G. Altmann. Vol. 3. Lüdenscheid, 2013. P. 23—38.

Altmann 1980 — *G. Altmann*. Prolegomena to Menzerath's Law // *Glottometrika*. 1980, 2. P. 1—10.

Altmann 2005 — *G. Altmann*. Phonic Word Structure // *Quantitative Linguistik. Ein internationales Handbuch* / Eds. R. Köhler, G. Altmann, R. Piotrowski. Berlin; New York, 2005. S. 191—208.

Berta 2015 — *A. Berta*. Tatar and Bashkir // Eds. L. Johanson, É. Á. C. Johanson. *The Turkic Languages*. London; New York, 2015. P. 283—300.

Best 2008 — *K.-H. Best*. Word Length in Persian // *Glottometrics*. 2008, 16. P. 27—30.

Best 2011 — *K.-H. Best*. Word Length Distribution in French // *Glottometrics*. 2011, 22. P. 57—61.

Bogdanova-Beglarian et al. 2017 — *N. Bogdanova-Beglarian, T. Sherstinova, O. Blinova, G. Martynenko*. Linguistic Features and Sociolinguistic Variability in Everyday Spoken Russian // *International Conference on Speech and Computer*. Springer, Cham, 2017. P. 503—511.

Bogdanova-Beglarian et al. 2018 — *N. Bogdanova-Beglarian, T. Sherstinova, O. Blinova, G. Martynenko, E. Baeva*. Towards a Description of Pragmatic Markers in Russian Everyday Speech // *International Conference on Speech and Computer*. Springer, Cham, 2018. P. 42—48.

Brown 2017 — *S. Brown*. The Distribution of Phoneme Inventory and Language Evolution // *CAES*. 2017. Vol. 3.1. P. 22—34.

Burbiel 2018 — *G. Burbiel*. Tatar grammar. A Grammar of the Contemporary Tatar Literary Language. Stockholm; Moscow, 2018.

Chomsky, Halle 1968 — *N. Chomsky, M. Halle*. The Sound Pattern of English. Cambridge, Massachusetts, 1968.

Coloma 2014 — *G. Coloma*. Towards a Synergetic Statistical Model of Language Phonology // *Journal of Quantitative Linguistics*. 2014. P. 100—122.

Galieva, Vavilova 2021 — *A. Galieva, Z. Vavilova*. Initial and Final Syllables in Tatar: from Phonotactics to Morphology // *Glottometrics*. 2021, 50. P. 57—75.

Halle 1995 — *M. Halle*. Feature Geometry and Feature Spreading // *Linguistic Inquiry*. 1995, 26 (1). P. 1—46.

Jakobson, Halle 1956 — *R. Jakobson, M. Halle*. Fundamentals of Language. S.-Gravenhage, 1956.

Jakobson, Waugh 1979 — *R. Jakobson, L. Waugh*. The Sound Shape of Language. Brighton, 1979.

Kessler 2015 — *B. Kessler*. Computational and Quantitative Approaches to Historical phonology // *Oxford handbook of historical phonology* / Eds. P. Honeybone, J. Salmons. Oxford, 2015. P. 171—188.

Kromer 2002 — *V. Kromer*. Zipf's law and its modification possibilities // *Glottometrics*. 2002, 5. P. 1—13.

Kromer 2001 — *V. Kromer*. Word length model based on the one-displaced Poisson-uniform distribution // *Glottometrics*. 2001, 1. P. 87—96.

Mandelbrot 1965 — *B. B. Mandelbrot*. Information Theory and Psycholinguistics // *Scientific Psychology* / Eds. B. Wolman, E. Nagel. New York, 1965. P. 550—562.

Moran, McCloy 2019 — *S. Moran, D. McCloy*. PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History, 2019 // <http://phoible.org>, 2019.

Popescu et al. 2015 — *I. Popescu, M. Lupea, D. Tatar, G. Altmann*. Quantitative Analysis of Poetic Texts // *Quantitative Linguistics*. Vol. 67. Berlin, 2015.

Poppe 1963 — *N. Poppe*. Tatar manual: descriptive grammar and texts with a Tatar-English glossary. Bloomington: Indiana University; The Hague: Mouton and Co., 1963.

Radojičić et al. 2019 — *M. Radojičić, B. Lazić, S. Kaplar, R. Stanković, I. Obradović, J. Mačutek, L. Leššová*. Frequency and Length of Syllables in Serbian // *Glottometrics*. 2019, 45. P. 114—123.

Thomson 2016 — *R. I. Thomson*. Does training to perceive L2 English vowels in one phonetic context transfer to other phonetic contexts? // *Canadian Acoustics*. 2016, 44 (3). P. 198—199.

Turunen, Lipping 2008 — *J. Turunen, T. Lipping*. Phoneme analysis based on quantitative and qualitative entropy measurement // *Computer Speech & Language*. 2008, 22 (4). P. 313—329.

Zörnig et al. 2019 — P. Zörnig, K. Stachowski, A. Ráková, Y. Qu, M. Mistecký, K. Ma, M. Lupea, E. Kelih, V. Gröller, H. Gnatchuk, A. Galieva, S. Andreev, G. Altmann. Quantitative Insights into Syllabic Structures // *Studies in Quantitative Linguistics* 30. Lüdenscheld, 2019.

References

- Altmann 2002 — G. Altmann. Zipfian linguistics In: *Glottometrics*. 2002. 3. P. 19—26.
- Altmann 2013 — G. Altmann. Aspects of Word Length. In: *Issues in Quantitative Linguistics* / Eds. R. Köhler, G. Altmann. Vol. 3. Lüdenscheld, 2013. P. 23—38.
- Altmann 1980 — G. Altmann. Prolegomena to Menzerath's Law. In: *Glottometrika*. 1980. 2. P. 1—10.
- Altmann 2005 — G. Altmann. Phonic Word Structure. In: *Quantitative Linguistik. Ein internationales Handbuch* / Eds. R. Köhler, G. Altmann, R. Piotrowski. Berlin; New York, 2005. P. 191—208.
- Baychura 1953 — U. Sh. Baychura. Glasnye tatarskogo jazyka v svete eksperimental'nyh dannyh: avtoref. dis. ... kand. filol. nauk. M.; L., 1953. {U. Sh. Baychura. Vowels of the Tatar language in the light of experimental data: abstract of dis. ... cand. philol. sciences. Moscow, Leningrad, 1953.}
- Baskakov 1988 — N. A. Baskakov. Istoriko-tipologicheskaja fonologija tjurkskih jazykov / Ed. E. R. Tenishev. M., 1988. {N. A. Baskakov. Historical and typological phonology of Turkic languages / Ed. E. R. Tenishev. Moscow, 1988.}
- Berta 2015 — A. Berta. Tatar and Bashkir. In: L. Johanson, É. Á. C. Johanson (eds.). *The Turkic Languages*. London; New York, 2015. P. 283—300.
- Best 2008 — K.-H. Best. Word Length in Persian. In: *Glottometrics*. 2008. 16. P. 27—30.
- Best 2011 — K.-H. Best. Word Length Distribution in French. In: *Glottometrics*. 2011. 22. P. 57—61.
- Bogdanova-Beglarian et al. 2017 — N. Bogdanova-Beglarian, T. Sherstinova, O. Blinova, G. Martynenko. Linguistic Features and Sociolinguistic Variability in Everyday Spoken Russian. In: *International Conference on Speech and Computer*. Springer, Cham, 2017. P. 503—511.
- Bogdanova-Beglarian et al. 2018 — N. Bogdanova-Beglarian, T. Sherstinova, O. Blinova, G. Martynenko, E. Baeva. Towards a Description of Pragmatic Markers in Russian Everyday Speech. In: *International Conference on Speech and Computer*. Springer, Cham, 2018. P. 42—48.
- Bogorodickij 1953 — V. A. Bogorodickij. Vvedenie v tatarskoe jazykoznanie v svjazi s drugimi tjurkskimi jazykami. Kazan, 1953. {V. A. Bogorodickij. Introduction to Tatar linguistics in connection with other Turkic languages. Kazan, 1953.}
- Brown 2017 — S. Brown. The Distribution of Phoneme Inventory and Language Evolution In: *CAES*. 2017. Vol. 3. 1. P. 22—34.
- Burbiel 2018 — G. Burbiel. *Tatar grammar. A Grammar of the Contemporary Tatar Literary Language*. Stockholm — Moscow, 2018.
- Chomsky, Halle 1968 — N. Chomsky, M. Halle. *The Sound Pattern of English*. Cambridge, Massachusetts, 1968.
- Coloma 2014 — G. Coloma. Towards a Synergetic Statistical Model of Language Phonology In: *Journal of Quantitative Linguistics*. 2014. P. 100—122.
- Galieva, Vavilova 2021 — A. Galieva, Z. Vavilova. Initial and Final Syllables in Tatar: from Phonotactics to Morphology. In: *Glottometrics*. 2021. 50. P. 57—75.
- Grishina 2005 — E. A. Grishina. Ustnaja rech' v Nacional'nom korpuse russkogo jazyka. In: *Nacional'nyj korpus russkogo jazyka: 2003—2005*. M., 2005. S. 94—110. {E. A. Grishina. Oral speech in the National Corpus of the Russian Language. In: *National Corpus of the Russian Language: 2003—2005*. Moscow, 2005. P. 94—110.}
- Grishina, Savchuk 2009 — E. A. Grishina, S. O. Savchuk. Korpus ustnyh tekstov v NKRJa: sostav i struktura. In: *Nacional'nyj korpus russkogo jazyka: 2006—2008. Novye rezul'taty i perspektivy*. St. Petersburg, 2009. S. 129—149. {E. A. Grishina, S. O. Savchuk. A Corpus of Oral Texts in the NCRL: Composition and Structure. In: *National Corpus of the Russian Language: 2006—2008. New results and prospects*. St. Petersburg, 2009. P. 129—149.}
- Halle 1995 — M. Halle. Feature Geometry and Feature Spreading. In: *Linguistic Inquiry*. 1995. 26 (1). P. 1—46.
- Jakobson, Halle 1956 — R. Jakobson, M. Halle. *Fundamentals of Language*. S.-Gravenhage, 1956.
- Jakobson, Waugh 1979 — R. Jakobson, L. Waugh. *The Sound Shape of Language*. Brighton, 1979.
- Kessler 2015 — B. Kessler. Computational and Quantitative Approaches to Historical phonology. In: *Oxford handbook of historical phonology* / Eds. P. Honeybone, J. Salmons. Oxford, 2015. P. 171—188.
- Kromer 2002 — V. Kromer. Zipf's law and its modification possibilities. In: *Glottometrics*. 2002. 5. P. 1—13.
- Kromer 2001 — V. Kromer. Word length model based on the one-displaced Poisson-uniform distribution. In: *Glottometrics*. 2001. 1. P. 87—96.
- Mandelbrot 1965 — B. B. [Mandelbrot]. *Information Theory and Psycholinguistics*. In: *Scientific Psychology* / Eds. B. Wolman, E. Nagel. New York, 1965. P. 550—562.

- Moran, McCloy 2019 — S. Moran, D. McCloy. PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History, 2019. In: <http://phoible.org>, 2019.
- Moroz 2019 — G. A. Moroz. Slogovaja struktura adygjejskogo jazyka: ot dannyh k obobshhenijam. In: *Voprosy jazykoznanija*. 2019. 2. S. 82—95. {G. A. Moroz. The syllable structure of the Adyghe language: from data to generalizations. In: *Voprosy jazykoznanija*. 2019. 2. P. 82—95.}
- Popescu et al. 2015 — I. Popescu, M. Lupea, D. Tatar, G. Altmann. *Quantitative Analysis of Poetic Texts // Quantitative Linguistics*. Vol. 67. Berlin, 2015.
- Poppe 1963 — N. Poppe. *Tatar manual: descriptive grammar and texts with a Tatar-English glossary*. Bloomington: Indiana University; The Hague: Mouton and Co., 1963.
- Radojičić et al. 2019 — M. Radojičić, B. Lazić, S. Kaplar, R. Stanković, I. Obradović, J. Mačutek, L. Leššová. Frequency and Length of Syllables in Serbian. In: *Glottometrics*. 2019. 45. P. 114—123.
- Shcherbak 1970 — A. M. Shcherbak. *Sravnitel'naja fonetika tjurkskih jazykov*. L.: Nauka, 1970. {A. M. Shcherbak. *Comparative phonetics of Turkic languages*. Leningrad: Nauka, 1970.}
- Smirnova N., Chistikov 2011 — N. Smirnova, P. Chistikov. Programma analiza foneticheskikh statistik v tekstah na russkom jazyke i eje ispol'zovanie dlja reshenija prikladnyh zadachah v oblasti rechevyh tehnologij. In: *Komp'juternaja lingvistika i intellektual'nye tehnologii: Po materialam ezhegodnoj Mezhdunarodnoj konferencii «Dialog»* (Bekasovo, 25—29 maja 2011 g.). Vol. 10. M., 2011. S. 632—643. {N. Smirnova, P. Chistikov. A program for the analysis of phonetic statistics in texts in Russian and its use for applied solutions in the field of speech technologies. In: *Computational Linguistics and Intelligent Technologies: Based on the materials of the “Dialogue” annual International Conference* (Bekasovo, May 25—29, 2011). Vol. 10. Moscow, 2011. P. 632—643.}
- Tatarskaja grammatika 1993 — *Tatarskaja grammatika* / Ed. M. Z. Zakiev. Vol. 1. Kazan, 1993. {Tatar grammar / Ed. M. Z. Zakiev. Vol. 1. Kazan, 1993.}
- Tatar grammatikası 2015 — *Tatar grammatikası* / Ed. F. M. Khisamova. Vol. 1. Kazan, 2015. {Tatar grammar / Ed. F. M. Khisamova. Vol. 1. Kazan, 2015.}
- Tatar National Corpus 2022 — “Tugan Tel” Tatar National Corpus. In: <http://tugantel.tatar/?lang=en>, 2022.
- Trubeckoj 2000 — N. S. Trubeckoj. *Osnovy fonologii* / Ed. S. D. Kacnel'son. M., 2000. {N. S. Trubeckoj. *Fundamentals of phonology* / Ed. S. D. Kacnel'son. Moscow, 2000.}
- Thomson 2016 — R. I. Thomson. Does training to perceive L2 English vowels in one phonetic context transfer to other phonetic contexts? In: *Canadian Acoustics*. 2016. 44 (3). P. 198—199.
- Turunen, Lipping 2008 — J. Turunen, T. Lipping. Phoneme analysis based on quantitative and qualitative entropy measurement. In: *Computer Speech & Language*. 2008. 22 (4). P. 313—329.
- Zörnig et al. 2019 — P. Zörnig, K. Stachowski, A. Ráková, Y. Qu, I. Místecký, K. Ma, M. Lupea, E. Kelih, V. Gröller, H. Gnatchuk, A. Galieva, S. Andreev, G. Altmann. Quantitative Insights into Syllabic Structures. *Studies in Quantitative Linguistics* 30. Lüdenscheid, 2019.

Севернохантыйский как язык с чертами аккузативного и иерархического кодирования

Муравьев Никита Алексеевич, Институт языкознания РАН (Москва), Университет Гамбурга; nikita.muraviev@gmail.com

В статье рассматривается стратегия ролевого кодирования участников в казымском диалекте севернохантыйского языка. Некоторые факты, такие как аккузативное маркирование личных местоимений, глагольное согласование с S/A-участником, преимущественно начальная позиция S/A-участника в клаузе и пассивный залог с продвижением O-участника, указывают на принадлежность казымского хантыйского к языкам с аккузативной стратегией. Вместе с тем, анализ дистрибуции активного и пассивного залога, субъектного и субъектно-объектного спряжений на полевом и текстовом материале обнаруживает ряд иерархических эффектов кодирования. В частности, выбор между активным и пассивным залогом диктуется относительной позицией A- и O-участника в иерархиях одушевленности и референциального статуса, а необходимым условием использования субъектно-объектных форм является определенность обоих участников. Совокупность данных фактов побуждает нас рассматривать казымский хантыйский как промежуточный случай, имеющий некоторые внешние признаки аккузативной стратегии кодирования, но в остальном функционирующий по принципам языков с иерархической стратегией, таких как алгонкинские или тибето-бирманские.

Ключевые слова: уральские языки, севернохантыйский язык, морфосинтаксис, ролевое кодирование, залог, иерархии, референциальный статус, одушевленность

NORTHERN KHANTY AS A LANGUAGE WITH THE FEATURES OF ACCUSATIVE AND HIERARCHICAL ALIGNMENT

Nikita A. Muravyev, Institute of Linguistics, RAS (Moscow), University of Hamburg; nikita.muraviev@gmail.com

This study features an alignment system in Kazym Khanty. Several facts including accusative marking of personal pronouns, verbal agreement with S/A and a strong tendency towards the clause-initial position of S/A suggest treating Kazym Khanty alignment as accusative. Meanwhile field and text data on the distribution of the active and passive voice as well as subject and subject-object agreement forms reveal several hierarchical coding effects. These include sensitivity of active/passive distinction to animacy and referentiality of the core participants and definiteness requirement of both participants for subject-object agreement. In sum, these facts suggest that Kazym Khanty can be considered an intermediate case of an alignment system combining certain surface properties of an accusative alignment with underlying hierarchical distribution patterns resembling some hierarchical coding systems of Algonquian and Tibeto-Burman languages.

Keywords: Uralic languages, Northern Khanty, morphosyntax, alignment, voice, hierarchies, referentiality, animacy

Введение

Севернохантыйский язык (обско-угорские < финно-угорские < уральские), наряду с восточным и южным хантыйским, распространен на территории Ямало-ненецкого и Ханты-мансийского автономных округов Российской Федерации. Хантыйский язык представляет собой диалектный континуум, и его диалектное членение на данный момент остается дискуссионным. В частности, в [Терешкин 1981] хантыйский язык подразделяется на западное и восточное наречия, или диалектные группы, а в [Штейниц 1937] — на северное, южное и восточное, некоторыми исследователями также выделяются промежуточные говоры. Северное наречие хантыгского языка, далее — севернохантыйский язык, на первый взгляд мало чем отличается от генетически и ареально близких ему языков с базовой аккумулятивной стратегией ролевого кодирования¹, распространенных на территории Западной Сибири. Об аккумулятивности и частичной нейтральности кодирования участников в севернохантыйском говорится, в частности, в грамматике [Nikolaeva 1999: 38]. Это следует из устройства падежного маркирования именных групп и глагольного согласования: S/A-участник выступает в немаркированной форме и вызывает первичное согласование, тогда как P-участник маркируется аккумулятивом либо остается немаркированным и иногда вызывает вторичное согласование. Однако такой взгляд на севернохантыйское кодирование основан только на устройстве активного залога и не учитывает тот факт, что активный залог в значительной мере функционально ограничен. Например, он не может использоваться в случае, если A-участник является неодушевленным или неопределенным, подробнее см. раздел 2.1. Если же рассматривать активный и пассивный залог в совокупности как единую систему кодирования участников, то такие функциональные лакуны обоих залогов складываются в иерархические принципы, свойственные языкам с базовой иерархической стратегией, где распределение синтаксических позиций между участниками зависит не от семантических ролей, а от взаимного расположения участников на иерархиях лиц и одушевленности. К примеру, одушевленный участник имеет приоритет над неодушевленным: если A одушевленный, а P неодушевленный, то A становится подлежащим, используется активный залог, если наоборот, то уже P становится подлежащим, используется пассивный залог. С учетом этого категория залога в севернохантыйском обнаруживает сходство с характерной для иерархических языков категорией направления, различающей значения директива и инверсива, которые отражают, соответственно, прямое и обратное соответствие морфосинтаксических ролей и иерархических позиций. Цель исследования — установить, какое место занимают иерархические принципы кодирования участников в системе севернохантыйского языка. Материалом для исследования послужили данные дистанционного опроса носителей с. Казым Ханты-Мансийского автономного округа — Югры (2020—2021 гг.). Для иллюстрации отдельных положений работы также используется текстовый материал из севернохантыйского корпуса на тегинском, казымском и шурышкарском диалектах².

Статья состоит из трех разделов. В разделе 1 представлен краткий обзор типологической литературы по ролевому кодированию, а также даны основные сведения о морфологии и кодировании участников в севернохантыйском языке. Раздел 2 посвящен анализу иерархических принципов организации залога и спряжения в севернохантыйском языке на материале казымского диалекта. В разделе 3 подводятся итоги исследования и обсуждается статус кодирования участников в севернохантыйском языке в типологии стратегий ролевого кодирования.

1. Ролевое кодирование в хантыйском языке: постановка задачи

1.1. Типологические исследования ролевого кодирования

Важнейшим периодом в типологических исследованиях кодирования участников стали 70-е годы прошлого столетия, в которые на смену традиционным евроцентричным понятиям субъекта и объекта пришла общепринятая ныне система SAO (SAP), впервые предложенная Р. Диксоном [Dixon 1972] и получившая развитие в работах Б. Комри, Б. Бикеля, Дж. Николс, А. Е. Кибрика и многих других [Comrie 1981; Nichols 1992; Kibrik 1997; Bickel, Nichols 2009], ср. обзор теоретических нюансов модели в [Haspelmath 2011]. В этой системе S соответствует единственному участнику непереходного глагола, A — первому участнику (как правило, агенту и субъекту) переходного глагола, O или P — второму участнику

¹ Стратегия кодирования участников в предложении является одним из основных параметров, определяющих типологический профиль языка. Выделяется аккумулятивная стратегия, при которой единственный участник кодируется так же, как и субъект переходного глагола, эргативная стратегия, при которой единственный участник кодируется так же, как и объект переходного глагола, и ряд других стратегий (подробнее см. раздел 1.1).

² Корпус текстов на хантыйском языке общим объемом в 42174 словоупотребления, разработанный Е. В. Кашкиным в рамках программы фундаментальных исследований президиума РАН «Корпусная лингвистика» (2012—2014 гг.).

(как правило, пациенсу и прямому объекту) переходного глагола. Данная система позволяет единообразно описывать данные языков с разными базовыми стратегиями падежного кодирования аргументов. Исходя из простого совпадения или несовпадения кодирования трех участников, выделяется пять логически возможных стратегий: нейтральная ($S=A=P$), аккузативная: ($S=A\neq P$), эргативная: ($S=P\neq A$), переходная, или горизонтальная ($S\neq A=P$) и трехчастная ($S\neq A\neq P$). Вместе с тем существуют различные отклонения от единообразного кодирования участников, как, например, активно-стативные системы, в которых S кодируется по-разному в зависимости от агентивности/пациентивности глагола. Кроме того, широко известны случаи расщепленной эргативности, когда выбор между аккузативной и эргативной стратегиями зависит от времени глагола, референциального выражения (местоимения vs. полные именные группы) и других параметров. Крайним случаем такого расщепленного кодирования являются иерархические языки Северной и Южной Америки, а также австронезийские языки с особой разновидностью залоговой системы, называемой залогом филиппинского типа. В этих языках выбор стратегии кодирования напрямую зависит от свойств участников: их одушевленности, топикальности, определенности и/или прагматической выделенности, что побудило некоторых исследователей постулировать особые типы: иерархическую стратегию [Nichols 1992] и австронезийскую стратегию [Blust 2013].

Отдельной типологической проблемой, затронутой в работе [Siewierska 2003], являются стратегии глагольного согласования. В некоторых случаях они могут совпадать с падежным кодированием, но довольно часто устроены сложнее, ср. обзор в [Bickel et al. 2013], и зависят от следующих факторов: разнообразия категорий, по которым согласуется глагол (например, число, лицо, род), наличия/отсутствия показателя согласования на глаголе, (не)совпадения показателей для разных участников, позиции согласовательного показателя в словоформе, условий наличия согласования. Достаточно вспомнить пример английского языка, в котором аккузативность глагольного согласования проявляется только в третьем лице единственного числа настоящего времени при помощи показателя *-s*, тогда как в остальных лицах и числах согласование морфологически не выражается и, тем самым, следует нейтральной стратегии, ср. *They like sailing* ‘Они любят плавать’ и *He like-s sailing* ‘Он любит плавать’ [Bickel et al. 2013: 17].

Наиболее интересной для нас в свете дальнейшего обсуждения представляется иерархическая стратегия кодирования, при которой кодирование участников определяется одной или несколькими семантическими, синтаксическими и/или дискурсивными иерархиями, ср. [Thompson 1994; Dixon, Aikhenvald 1997; Zuñiga 2006]. Это означает, что в качестве подлежащего выбирается не фиксированный A или P -участник, а тот из них, который стоит выше по лицу, одушевленности или, например, референциальному статусу. Внешне это может проявляться как в вариативном падежном кодировании, так и в вариативном согласовании. Иерархические системы кодирования фиксируются преимущественно в языках Северной и Южной Америки, тибето-бирманских языках и на севере Австралии, см. [Siewierska 2013].

Рассмотрим пример южноамериканского языка пикурис с базовой иерархической стратегией кодирования участников, кодирование в котором, как будет показано далее, имеет много общего с севернохантыйским. Характерный признак большинства иерархических глагольных систем — дифференцированное кодирование участников в зависимости от их соотношения по лицу, а именно приоритетное кодирование локуторов (участников первого или второго лица) по сравнению с нелокуторами (участниками третьего лица), как в примерах (1, 2) ниже. В (1) A -участник первого лица производит действие над P -участником третьего лица, тогда как в (2), напротив, A -участник третьего лица производит действие над P -участником первого лица. Поскольку первое лицо иерархически выше третьего, глагол в обоих случаях согласуется с участником первого лица, какую бы роль в ситуации он ни играл. Избежать омонимии в данном случае помогает (а) выбор парадигмы согласования: в первом случае для участника первого лица используется субъектный ряд $ПА$, во втором — объектный ряд I , (б) именной показатель косвенного объекта для третьего лица во втором случае и (в) показатель инверсива на глаголе, указывающий на то, что действие совершается иерархически более низким участником над более высоким.

ПИКУРИС (кайова-таноанские)

(1) *sənene ti-mɔn-ʔan*

человек 1SG:ПА-видеть-PST

‘Я увидел человека’ [Klaiman 1993: 359]³.

(2) *ta-mɔn-mia-ʔan sənene-pa*

1SG:I-видеть-INV-PST человек-OBL

‘Человек увидел меня’ [там же].

³ Здесь и далее англоязычные переводы в иностранных источниках даны в русском варианте, перевод на русский язык — автора настоящей статьи.

Однако гораздо более типична для языков с иерархическим кодированием ситуация, при которой показатель директива/инверсива является единственным средством установить соответствие между иерархией участников и их ролями. Так, в примерах (3, 4) из равнинного кри [Zuñiga 2006: 2] мы видим аналогичное сочетание участников первого и третьего лица. Если действие направлено от более иерархически высокого участника первого лица к более низкому участнику третьего лица, используется показатель директива (3), в противном случае — показатель инверсива (4).

РАВНИННЫЙ КРИ (алгонкинские)

- (3) *ni-wāpat-ā-w*
1-видеть-DIR-3
'Я вижу его/ее' [Zuñiga 2006: 2].
- (4) *ni-wāpat-ikw-w*
1-видеть-INV-3
'Он(а) видит меня' [там же].

Отдельные иерархические закономерности в глагольной морфологии встречаются и в уральских языках. Так, в венгерском и в северносамодийских языках действует так называемый запрет на инверсивное согласование (англ. *inverse agreement constraint*), ср. [É. Kiss 2013; Bárányi 2015]. В языках с этим ограничением субъектно-объектное спряжение невозможно при Р-участнике первого или второго лица, который более топикален, чем А-участник третьего лица. Так, венгерский язык имеет парадигмы субъектного (5) и субъектно-объектного спряжения (6), последнее из которых используется при определенном прямом объекте.

ВЕНГЕРСКИЙ (уральские)

- (5) *Lát-sz egy kutyá-t.*
видеть-2SG.SUBJ INDEF собака-ACC
'Ты видишь собаку (неопределенную)' [Bárányi 2015: 2].
- (6) *Lát-od a kutyá-t.*
видеть-2SG.OBJ DEF собака-ACC
'Ты видишь собаку (определенную)' [там же].

Дополнительно существуют субъектно-объектные формы для первого лица А-участника и второго лица Р-участника (7), ср. (8).

- (7) *Lát-lak téged.*
видеть-1SG>2 ты.ACC
'Я вижу тебя' [Bárányi 2015: 5].
- (8) **Lát-ok téged.*
видеть-1SG.OBJ ты.ACC
Ожид.: 'Я вижу тебя' [там же].

При этом для случая третьего лица А-участника и первого либо второго лица Р-участника отдельных форм не существует, ср. формы субъектного спряжения в (9, 11), а использование основных форм субъектно-объектного спряжения оказывается неграмматичным (10, 12).

- (9) *Lát-om őt.*
видеть-1SG.SUBJ он.ACC
'Я вижу его/ее' [Bárányi 2015: 3].
- (10) **Lát-ok őt.*
видеть-1SG.OBJ он.ACC
Ожид.: 'Я вижу его/ее' [там же].
- (11) *Lát-Ø engem.*
видеть-3SG.SUBJ я.ACC
'Он/она видит меня' [Bárányi 2015: 3].
- (12) **Lát-j-a engem.*
видеть-3SG.OBJ я.ACC
Ожид.: 'Он/она видит меня' [там же].

Случай венгерского языка показывает, что даже заведомо аккумулятивные языки могут обнаруживать закономерности иерархического кодирования. Поэтому для полного понимания закономерностей кодирования в родственных венгерскому обско-угорских языках и, в частности, в севернохантыйском необходим комплексный подход, учитывающий кодирование участников при разнообразных, в том числе непрототипических сочетаниях двух основных участников.

1.2. Ролевое кодирование в севернохантыйском языке

Рассмотрим подробнее релевантные аспекты морфологии севернохантыйского языка. Согласно описаниям [Steinitz 1950; Николаева 1995; Каксин 2010] и др., именная система севернохантыйского языка

различает три падежа: немаркированный падеж, локатив, транслатив (обдорский диалект) или латив (другие северные диалекты), ср. таблицу 1 ниже. Во всех севернохантыйских говорах немаркированная форма используется в позициях субъекта, прямого объекта и приименного зависимого, фактически сочетая в себе функции номинатива, аккузатива и генитива. У личных местоимений падежная система отличается наличием морфологически выраженных аккузатива и датива и отсутствием форм локативных падежей.

Таблица 1. Образец склонения существительного и местоимения в обдорском и казымском диалектах севернохантыйского языка, по [Nikolaeva 1999: 13, 16; Каксин 2010: 83, 91].

	Существительное <i>χot</i> ‘дом’		Местоимение <i>luw, lūw</i> ‘он/она’	
	обдорский	казымский	обдорский	казымский
Номинатив			<i>luw</i>	<i>lūw</i>
Генитив	<i>χo:t</i>	<i>χot</i>		
Аккузатив			<i>luwe:l</i>	<i>lūwti / lūweλ</i>
Датив	–	–		<i>lūwela</i>
Локатив	<i>χo:t-na</i>	<i>χot-ən</i>	<i>luwe:l-na</i>	–
Латив	–	<i>χot-a</i>	–	–
Транслатив	<i>χo:t-ti</i>	–		–

Помимо падежа, именная морфология хантыйского языка различает категории числа (единственное, двойственное и множественное) и посессивности. Посессивность используется, во-первых, в собственной функции для указания на лицо и число обладателя, ср. *ma χot-em* ‘мой дом’ (я дом-POSS.1SG) и, во-вторых, в дискурсивной функции для указания на топикальность или определенность именной группы [Михайлов 2018], ср. *an-en mij-e* ‘дай чашку (опред.)’ (чашка-POSS.2SG дать-IMP.SG).

Глагольная система севернохантыйского языка различает активный и пассивный залог. Глагольное спряжение в активном залоге маркирует согласование глагола по лицу и числу с S/A-участником и по числу с P-участником и различает две парадигмы: субъектную и субъектно-объектную, ср. таблицы 2 и 3 ниже.

Таблица 2. Образец субъектного спряжения глагола *mānti* ‘идти’ в непрошедшем времени в казымском диалекте севернохантыйского языка [Каксин 2010: 101].

Л.ч. S	Форма	Л.ч. S	Форма	Л.ч. S	Форма
1SG	<i>mān-λ-əm</i>	1DU	<i>mān-λ-əmn</i>	1PL	<i>mān-λ-ūw</i>
2SG	<i>mān-λ-ən</i>	2DU	<i>mān-λ-ətn</i>	2PL	<i>mān-λ-əti</i>
3SG	<i>mān-λ-∅</i>	3DU	<i>mān-λ-əjn</i>	3PL	<i>mān-λ-ət</i>

Таблица 3. Образец субъектно-объектного спряжения глагола *təχərti* ‘запереть’ в непрошедшем времени в казымском диалекте севернохантыйского языка [Каксин 2010: 107—108].

Л.ч. S, ч. O	Форма	Л.ч. S, ч. O	Форма	Л.ч. S, ч. O	Форма
1SG>SG	<i>təχər-λ-em</i>	1DU>SG	<i>təχər-λ-emən</i>	1PL>SG	<i>təχər-λ-ew</i>
2SG>SG	<i>təχər-λ-en</i>	2DU>SG	<i>təχər-λ-əln</i>	2PL>SG	<i>təχər-λ-əln</i>
3SG>SG	<i>təχər-λ-əlle</i>	3DU>SG	<i>təχər-λ-əln</i>	3PL>SG	<i>təχər-λ-el</i>
1SG>NSG	<i>təχər-λ-əlləm</i>	1DU>NSG	<i>təχər-λ-əlləmn</i>	1PL>NSG	<i>təχər-λ-əllūw</i>
2SG>NSG	<i>təχər-λ-əllən</i>	2DU>NSG	<i>təχər-λ-əllən</i>	2PL>NSG	<i>təχər-λ-əllən</i>
3SG>NSG	<i>təχər-λ-əlle</i>	3DU>NSG	<i>təχər-λ-əllən</i>	3PL>NSG	<i>təχər-λ-əlləl</i>

Вопрос о дистрибуции субъектного и субъектно-объектного спряжений остается предметом дискуссий, однако на данный момент принято считать, что при переходных глаголах, где наблюдается вариативность между данными двумя типами форм, выбор между ними зависит от информационной структуры (см. более подробное обсуждение далее в 2.2). При непереходных глаголах единственно возможным является субъектное спряжение.

Пассивный залог в севернохантыйском языке маркируется специальным показателем *-a(j)/-i(j)-*, который располагается после показателя времени и перед показателем лично-числового согласования

субъектной парадигмы спряжения. Алломорф с конечным йотом используется, если далее следует лично-числовой показатель, тогда как алломорф без йота фиксируется в форме 3-го л. ед. ч. с нулевым показателем. Алломорф *-a(j)-* выбирается при глаголах с основой на согласный, *-i(j)-* — с основой на гласный [Каксин 2010: 112—114], см. таблицу 4 ниже.

Таблица 4. Образование форм пассивного залога в казымском диалекте севернохантыйского языка

	<i>wɛr-ti</i> ‘делать’, основа на согласный	<i>tɔ-ti</i> ‘нести’, основа на гласный
Актив, 3SG	<i>wɛr-əs</i>	<i>tɔ-s</i>
Актив, 3PL	<i>wɛr-s-ət</i>	<i>tɔ-s-ət</i>
Пассив, 3SG	<i>wɛr-s-a</i>	<i>tɔ-s-i</i>
Пассив, 3PL	<i>wɛr-s-aj-ət</i>	<i>tɔ-s-ij-ət</i>

А-участник (исходное подлежащее) при пассивном залоге занимает позицию косвенного дополнения и оформляется падежным показателем локатива, в то время как позицию подлежащего занимает другой аргумент, чаще всего Р-участник (исходное прямое дополнение), с которым согласуется глагол в лице и числе.

Устройство именной и глагольной систем севернохантыйского языка в целом свидетельствует в пользу базовой аккумулятивной стратегии кодирования участников. Она проявляется, во-первых, в аккумулятивном оформлении местоимений, ср. аккумулятивную форму *mǎnǎtti* ‘меня’ в (14) ниже, и, во-вторых, в согласовании в первую очередь с S/A-участником и лишь во вторую — с Р-участником, ср. формы субъектного спряжения глаголов *šǔkǎtti* ‘разбить’ и *χǎtsǎti* ‘ударить’ в (13) и (14) и форму субъектно-объектного спряжения глагола *χǎtsǎti* ‘ударить’ в (14). Свою роль играет и порядок слов: даже при формальной нейтральности падежного кодирования полных именных групп субъект отчетливо тяготеет к начальной позиции, тогда как прямой объект располагается вслед за субъектом, ср. расположение участников ‘Вася’ и ‘чашка’ в (13), ‘Вася’ и ‘я’ в (14).

- (13) *waša-j-en an šǔkǎt-əs*
Вася-EP-POSS.2SG чашка разбить-PST[3SG]
‘Вася разбил чашку’.
- (14) *waša-j-en mǎnǎtti χǎtsǎ-s / χǎtsǎ-s-λe*
Вася-EP-POSS.2SG я.ACC ударить-PST[3SG] / ударить-PST-3SG>SG
‘Вася ударил меня’.

Вместе с тем обнаруживаются случаи, в которых А-участник не может занимать позицию подлежащего, а Р-участник — позицию прямого дополнения. Например, если А-участник оказывается в фокусе частного вопроса, как ‘Иван’ в примере (15) из обдорского диалекта, или является неодушевленным, как ‘ветер’⁴ в примере (16) из казымского диалекта, единственно возможный вариант — использовать пассивную конструкцию.

- (15) *kalan χoj-na we:l-s-a? Juwan-na we:l-s-a*
олень кто-LOC убить-PST-PASS[3SG] Иван-LOC убить-PST-PASS[3SG]
‘Кто убил оленя? Иван убил’ [Nikolaeva 2001: 25].
- (16) *wət-ən maša-j-en iλ pǎwət-s-a*
ветер-LOC Маша-EP-POSS.2SG вниз уронить-PST-PASS[3SG]
‘Порыв ветра сбил Машу’.

Такая обязательность для пассива, в целом, типологически нехарактерна [Klaiman 1993: 348], поскольку, как правило, в языках с противопоставлением актива и пассива любому предложению в активе можно подобрать пассивный аналог и наоборот, ср. *Маша была сбита порывом ветра* и *Порыв ветра сбил Машу* или в безличном варианте *Машу сбilo порывом ветра*. Альтернативный подход состоит в том, чтобы рассматривать оппозицию актива и пассива в севернохантыйском не как залоговую альтернативу, а как вариативную стратегию кодирования участников, при которой морфосинтаксические роли

⁴ В данном случае А-участник ‘порыв ветра’ является не агенсом, а каузаторм, ср. [Падучева 2001: 57].

А и Р соотносятся с синтаксическими позициями не единственным образом, а в зависимости от взаимного иерархического положения этих участников, а глагол посредством залоговой морфологии указывает на соответствие (актив) или несоответствие (пассив) А-участника наиболее высокой иерархической позиции среди участников в предложении. Возвращаясь к примеру (2) из языка пикурис, стоит отметить формальное сходство севернохантыйского пассива с инверсивным кодированием в этом языке: в обоих случаях имеется специальный показатель на глаголе (пассив в хантыйском и инверсив в пикурис) и косвенный именной показатель (локатив в хантыйском и обликвус в пикурис), а глагол согласуется по лицу и числу с Р-участником.

Для того чтобы адекватно оценить, в какой пропорции севернохантыйский язык сочетает в себе черты аккузативного и иерархического кодирования, необходимо последовательно обсудить основные аспекты кодирования участников в севернохантыйском языке, ответив на следующие вопросы:

- 1) **От каких параметров зависит вариативное кодирование участников в севернохантыйском языке?**
- 2) **Носит ли кодирование участников иерархический характер?**
- 3) **Если да, то какое место занимает иерархическое кодирование в системе севернохантыйского языка?**

Вопросы 1 и 2 будут далее рассмотрены во второй главе, вопрос 3 — в третьей главе.

2. Параметры кодирования участников в севернохантыйском языке

Как мы упомянули в разделе 1, севернохантыйский язык формально относится к языкам с базовой аккузативной стратегией кодирования участников, но в то же время обнаруживает ряд черт, сближающих его с иерархическими языками, что требует более детального взгляда на параметры кодирования участников. Далее мы отдельно рассмотрим основные параметры кодирования участников, влияющие на выбор ключевых для севернохантыйской системы глагольных форм: активного и пассивного залога (2.1), а также субъектного и субъектно-объектного спряжения (2.2).

2.1. Активный и пассивный залог

Прототипический пассив характеризуется смещением А-участника из позиции подлежащего и продвижением на его место Р-участника [Keenan 1985; Shibatani 1988]. Изначально предполагалось, что основным свойством пассива является синтаксическое повышение Р-участника [Givón 1979: 186], но в последнее время ключевая роль отводится понижению А-участника [Shibatani 1985: 830; Thompson 1994: 47], в том числе из-за существования непродвигающих безличных пассивов. Специфичной чертой обско-угорских пассивов и, в частности, пассивного залога в севернохантыйском языке является его связь с информационной структурой. Его основной функцией является поддержание топика [Кошкарёва 2002: 35] в случаях, когда участник с семантической ролью А-участника не соответствует данной информационно-структурной роли [Kulonen 1989: 71; Nikolaeva 2001: 25; Virtanen, Sosa 2018: 235]. К примеру, описание залога в обдорском диалекте в [Nikolaeva 2001] строится с опорой на модель информационной структуры высказывания в [Lambrecht 1994], в которой на основании вопросно-ответных тестов выделяется а) аргументный фокус, как в {Кого Вася ударил?} Он ударил ПЕТЮ, б) предикатный фокус {Что Вася сделал с Петей?} Он УДАРИЛ Петю и в) сентенциальный фокус, как в {Что случилось?} Вася ударил Петю. Так, в вопросно-ответной паре (15), повторенной в (17), топиком является Р-участник 'олень', тогда как А-участник 'Иван' оказывается в фокусе и поэтому не может стоять в позиции подлежащего при активном залоге. По этой причине используется пассивный залог, при котором 'олень' занимает позицию подлежащего, а 'Иван' помещен в позицию косвенного дополнения в локативе.

- (17) *kalan xoj-na we:l-s-a? Juwan-na we:l-s-a*
 олень кто-LOC убить-PST-PASS[3SG] Иван-LOC убить-PST-PASS[3SG]
 'Кто убил оленя? Иван убил' [Nikolaeva 2001: 25].

Залоговое противопоставление в севернохантыйском можно рассматривать и в терминах дискурсивной топикальности, или активированности референтов в дискурсе, по [Givón 1983]. Так, в статье [Кошкарёва 2002] приводится нарративная последовательность (18), в которой 'свекровь' и 'сноха' являются наиболее топикальными участниками, тогда как А-участник последнего высказывания 'нити выделяющая женщина' упоминается впервые и также не может занимать позицию подлежащего, поэтому и в этом случае используется пассив с А-участником в позиции косвенного дополнения в локативе.

(18) {В одно время люди жили, сноха у них тоже была. Это было в наших краях. Каждый вечер эта сноха нити делает. Свекровь ей говорит: «Невестка, нельзя по вечерам нити делать...»}

p̄in wert̄i imi-j-ən j̄öχət-λ-aj-ən
нити делать женщина-EP-LOC прийти-NPST-PASS-2SG

‘Нити выделяющая женщина к тебе придет’ [Кошкарева 2002: 35].

Как можно видеть из этого примера, а также из примеров (19—21) ниже, в севернохантыйском пассивизация не ограничивается продвижением топикального Р-участника, а охватывает широкий спектр семантических ролей, ср. реципиент (19), бенефициант (20) и место (21).

(19) *wen-jəχ-λ-аλ-ən m̄ört-a m̄əjλ-s-ij-ət*
зять-люди-PL-POSS.3SG-LOC мера-LAT дарить-PST-PASS-3PL

‘Они были щедро одарены семьей зятя’ [Kulonen 1989: 229].

(20) *maw-əŋ p̄asan-ən, s̄ör-əŋ p̄asan-ən wer-s-aj-ət*
мед-ADJ стол-LOC пиво-ADJ стол-LOC делать-PST-PASS-3PL

‘Стол с медом и пивом был накрыт для них’ [там же: 220].

(21) *m̄ow šiw-ən əm̄əs-λ-a*
земля туман-LOC сесть-NPST-PASS[3SG]

‘Погода туманная (букв. на землю сел туман)’ [там же: 185].

Однако сравнение севернохантыйского языка с материалом языков с базовой иерархической стратегией кодирования предполагает обзор поведения исследуемой категории с точки зрения не только информационной структуры, но и других параметров, таких как лицо, одушевленность и референциальный статус, ср. [Zúñiga 2006: 48].

Рассмотрим для начала параметр лица и конкретнее контексты А-участника нелокутора и Р-участника локутора, А-участника локутора и Р-участника нелокутора и обоих участников локуторов. Особенностью севернохантыйских диалектов, за исключением обдорского, является отсутствие у местоимений падежной формы локатива, ср. обзор в [Вальгамова и др. 2011: 168—171] и связанный с этим повсеместный запрет (в том числе и в обдорском) на использование личных местоимений для выражения А-участника при пассиве. Поэтому в контекстах обоих участников первого/второго лица используется исключительно активный залог, ср. (22a) и неграмматичность (22b).

(22) a. *n̄əŋ m̄ənətti χ̄äts̄ə-s-ən*
ты я.ACC ударить-PST[3SG]

‘Ты меня ударил’.

b. **m̄a <n̄əŋenən>⁵ χ̄äts̄ə-s-ij-əm*
я ты.LOC ударить-PST-PASS-1SG

Ожид.: ‘Ты меня ударил’.

То же касается и конфигурации А-участника локутора и Р-участника нелокутора, в которой также может использоваться только активный залог (23a), ср. невозможность пассива в (23b).

(23) a. *m̄a w̄aša-j-en χ̄äts̄ə-s-əm*
я Вася-EP-POSS.2SG ударить-PST-1SG

‘Я ударил Васю’.

b. **w̄aša-j-en <m̄ənemən> χ̄äts̄ə-s-i*
Вася-EP-POSS.2SG я.LOC ударить-PST-PASS[3SG]

Ожид.: ‘Я ударил Васю’.

В свою очередь, в обратной конфигурации А-участника нелокутора и Р-участника локутора может использоваться как активный (24a), так и пассивный залог (24b).

(24) a. *m̄a w̄aša-j-en-ən χ̄äts̄ə-s-ij-əm*
я Вася-EP-POSS.2SG-LOC ударить-PST-PASS-1SG

‘Вася меня ударил’.

⁵ Здесь и далее в угловых скобках приведена искусственно построенная по аналогии с дативной формой *m̄ənem̄a* форма локатива местоимения. Данная форма, хотя и признается приемлемой отдельными носителями в специальных контекстах, запрещается большинством других носителей, не фиксируется в текстах, в том числе в позиции агентивного дополнения при пассиве.

- b. *waša-j-en mānatti χūtšə-s*
 Вася-EP-POSS.2SG я.АСС ударить-PST[3SG]
 ‘Вася меня ударил’.

Если примеры (23а, б) укладываются в логику иерархических языков, в которых подлежащим обязан быть участник, более высокий в иерархии локутор > нелокутор, то примеры (24а, б), на первый взгляд, отходят от этой логики. Однако, если обратиться к материалу текстов, ср. таблицу 5 ниже, становится очевидно, что активный залог при А-участнике нелокуторе и Р-участнике локуторе встречается заметно реже, чем пассивный, и картина получается практически зеркальной к контекстам с А-участником локутором и Р-участником нелокутором, ср. приведенные в начале примеры (1, 2) из языка пикурис с аналогичным, хотя и более строгим распределением директива и инверсива.

Таблица 5. Использование актива и пассива в зависимости от лиц участников в казымском диалекте севернохантыйского языка: корпусные данные

	А локутор, Р нелокутор	А нелокутор, Р локутор	Всего
Актив	43	4	47
Пассив	—	42	42
Всего	43	46	89

Теперь рассмотрим параметры выбора залога в контекстах обоих участников третьего лица. Одним из значимых параметров является одушевленность участников. Так, при неодушевленном А-участнике и одушевленном Р-участнике активный залог попадает под запрет (25а), и единственно возможным оказывается пассивный залог (25б).

- (25) a. **wət-en maša-j-en ił pāwt-əs*
 ветер-POSS.2SG Маша-EP-POSS.2SG вниз уронить-PST[3SG]
 Ожид.: ‘Порыв ветра сбил Машу’.
 б. *wət-ən⁶ maša-j-en ił pāwət-s-a*
 ветер-LOC Маша-EP-POSS.2SG вниз уронить-PST-PASS[3SG]
 ‘Порыв ветра сбил Машу’.

Случаи, когда А-участник третьего лица одушевленный, но уступает в одушевленности Р-участнику, как ‘собака’ в (26), попадают под менее категоричный запрет.

- (26) a. *?amp-en pet'a-j-en tɔχm-əs*
 собака-POSS.2SG Петя-EP-POSS.2SG кусать-PST[3SG]
 ‘Собака укусила Петю’.
 б. *pet'a-j-en amp-ən tɔχəm-s-a*
 Петя-EP-POSS.2SG собака-LOC кусать-PST-PASS[3SG]
 ‘Собака укусила Петю’.

При этом более значимую роль, по-видимому, играет именно одушевленность А-участника, поскольку даже при неодушевленном Р-участнике, как в (27), использование активного залога оказывается затруднено⁷.

⁶ Стоит отметить, что в (26б) участник ‘ветер’ в позиции косвенного объекта не маркирован дискурсивным показателем посессива второго лица, который, как уже упоминалось выше, используется практически исключительно в субъектной позиции (кроме имен собственных), ср. [Mikhailov 2021: 43–48].

⁷ На данный момент мы не можем утверждать, что в таких контекстах использование активного залога полностью исключено, однако построить соответствующий естественный контекст крайне сложно. Так, например, предложение (i), по комментариям носителей, возможно в том случае, если ветер в предшествующем контексте был главным действующим лицом, совершившим ряд других действий.

(i) *wət-en tām sɔχ-ət ił pawt-əs*
 ветер-POSS2SG этот шкура-PL вниз ронять-PST[3SG]
 {‘Что натворил сегодня ветер?’} ‘Ветер уронил эту одежду’.

Данный контекст можно рассматривать как результат коэрсии изначально неподходящего для позиции подлежащего неодушевленного участника посредством его персонификации.

- (27) a. *wət-en lipt-ət ar-a tɔ-s
ветер-POSS.2SG лист-PL много-DAT нести-PST[3SG]
Ожид.: ‘Порыв ветра разбросал листья’.
- b. lipt-ət wət-ən ar-a tɔ-s-ij-ət
лист-PL ветер-LOC много-DAT нести-PST-PASS-3PL
‘Порыв ветра разбросал листья’.

Помимо одушевленности, на выбор залога влияет также референциальный статус участников. При референтном неопределенном А-участнике и определенном Р-участнике допустимо использование только пассивного (28b, 29b), но не активного залога (28a, 29a).

- (28) a. *mǔlsər aj_iki waša-j-en χǎtsə-s
что.INDEF мальчик Вася-EP-POSS.2SG ударить-PST[3SG]
Ожид.: ‘Какой-то мальчик ударил Васю’.
- b. waša-j-en mǔlsər aj_iki-j-ən χǎtsə-s-i
Вася-EP-POSS.2SG что.INDEF мальчик-EP-LOC ударить-PST-PASS[3SG]
‘Какой-то мальчик ударил Васю’.
- (29) a. *mǔlsər aj_iki tām an-en šūkət-əs
что.INDEF мальчик этот чашка-POSS.2SG ломать-PST[3SG]
Ожид.: ‘Какой-то мальчик разбил чашку’.
- b. tām an-en mǔlsər aj_iki-j-ən šūkət-s-a
этот чашка-POSS.2SG что.INDEF мальчик-EP-LOC ломать-PST-PASS[3SG]
‘Какой-то мальчик ударил разбил чашку’.

Если же оба участника оказываются референтными неопределенными, как в (30), ограничение на активный залог снимается.

- (30) a. aj_iki amp χǎtsə-s
мальчик собака ударить-PST[3SG]
{Что случилось?} ‘Мальчик ударил собаку’.
- b. amp aj_iki-j-ən χǎtsə-s-i
собака мальчик-EP-LOC ударить-PST[3SG]
{Что случилось?} ‘Мальчик ударил собаку’.

Однако в безличных переходных контекстах, где референтный неопределенный (31) или генерический А-участник (32) не выражен, используется исключительно пассивный залог.

- (31) an šūkət-s-a / *šūkət-s-ət
чашка ломать-PST-PASS[3SG] / ломать-PST-3PL
{Что случилось?} ‘Чашку разбили’.
- (32) tāta išne-t at keša tɔxər-λ-aj-ət / *tɔxər-λ-ət
здесь окно-PL ночь для запырять-PST-PASS-3PL / запырять-PST-3PL
‘Здесь окна на ночь закрывают’.

Кроме того, существует целый ряд глаголов, у которых пассивный залог является основным или единственным вариантом, ср. [Соловар 2010: 91—92]. Такие глаголы описывают изменение психического, физического, физиологического или метеорологического состояния в результате воздействия нереперентной и неодушевленной внешней силы, которая, как и ветер, не может стоять в позиции подлежащего переходного глагола. Так, пассив образуется от одноместных глаголов *mǎreməti* ‘соскучиться’ (33), не имеющего активной формы, *pǎtləti* ‘стемнеть’ (34), имеющего параллельную активную форму для безличного контекста *na ulice stemnelo*, и *šəŋχiti* ‘гореть, тлеть’ с синонимичными активной и пассивной формой (35).

- (33) apəlne-ŋ-en mǎremə-s-i
младшая.сестра-EP-POSS.2SG соскучиться-PST-PASS[3SG]
‘Сестра твоя соскучилась’ [Соловар 2010: 92].
- (34) lǎw šāta pǎtlə-s-i
она там стемнеть-PST-PASS[3SG]
‘Ее там темнота застала’ [там же: 92].

- (35) *неpek śəŋχi-s-a* / *śəŋχij-əs*
 бумага тлеть-PST-PASS[3SG] / тлеть-PST[3SG]
 ‘Бумага истлела’.

В последнем примере семантическая разница между глагольными формами, по-видимому, состоит в том, что при активе событие тления концептуализуется как внутренний естественный процесс, тогда как при пассиве — это результат внешнего воздействия, ср. примеры ниже с глаголом *pətti* ‘мерзнуть’ и неодушевленным участником (36) и с одушевленным участником (37), которого событие замерзания в нормальном случае может затронуть только извне, ср. возможность выразить каузатора ‘холод’ при пассиве, но не при активе (38).

- (36) *jəχan-ew pət-s-a* / *pət-əs*
 река-POSS.1PL замерзать-PST-PASS[3SG] / замерзать-PST[3SG]
 ‘Наша река замерзла’.

- (37) *amp-en pət-s-a* / *??pət-əs*
 собака-POSS.2SG замерзать-PST-PASS[3SG] / замерзать-PST[3SG]
 ‘Собака замерзла’.

- (38) *amp-en iški-j-ən pət-s-a* / **pət-əs*
 собака-POSS.2SG холод-EP-LOC замерзать-PST-PASS[3SG] / замерзать-PST[3SG]
 ‘Собака замерзла (букв.: холодом замерзла)’.

Приведенные примеры можно рассматривать как свидетельство того, что пассивный залог сохраняет переходность глагола, не устраняя А-участника полностью, а только не выражая или ограниченно выражая его в силу его крайне низкой позиции на референциальной иерархии.

Таким образом, залог в севернохантыйском обнаруживает признаки иерархической организации: преимущество А-участника по параметрам ведет к использованию активного залога, преимущество Р-участника по тем же параметрам — к использованию пассивного залога, см. таблицу 6 ниже. В этом отношении сравнение противопоставления актив-пассив с оппозицией директив-инверсив оказывается вполне уместным.

Таблица 6. Использование актива и пассива
 в зависимости от исследуемых параметров
 в казымском диалекте севернохантыйского языка

А-участник	Р-участник	Актив	Пассив
топик	в фокусе		
определенный	неопределенный	+	—
локутор	(не)локутор		
личный	безличный		
равные по информационной структуре			
равные по референциальному статусу			
равные по одушевленности		+	+
одушевленный	неодушевленный		
нелокутор	(не)локутор		
одинаково личные			
в фокусе	топик		
неопределенный	определенный	—	+
неодушевленный	одушевленный		
безличный	личный		

Что касается конкретных параметров, как можно видеть из таблицы, локуторы обнаруживают явный, хотя и не совсем последовательный приоритет над нелокуторами, как и в классических иерархических языках. В случае же двух нелокуторов выбор между активом и пассивом зависит не только от топикальности А-участника и Р-участника, как в обдорском, но и от их взаимного расположения в референциальной иерархии и иерархии одушевленности.

2.2. Субъектное и субъектно-объектное спряжение

Вариативное глагольное согласование лишь относительно недавно стало предметом широкого обсуждения [Iemmolo 2011; Iemmolo, Klumpp 2014; Haig 2018]. В частности, в работе [Haig 2018: 789] вводится понятие «дифференцированное выражение объекта в глаголе» (англ. Differential Object Indexing, DOI), которое определяется как система выражения прямого объекта, в которой показатель числа на предикате зависит от прагматических и семантических факторов, не связанных с самим предикатом. В числе факторов упоминается а) наличие/отсутствие выраженного кореферентного прямого объекта в данной клаузе, б) лицо, одушевленность, информационный статус прямого объекта, в) информационная структура всей клаузы и г) наличие дополнительного объекта, более высокого по одушевленности. В севернохантыйском языке такую систему образуют парадигмы субъектного и субъектно-объектного спряжения. Субъектное спряжение используется всегда с непереходными глаголами, а также в пассивном залоге, тогда как у переходных глаголов оно употребляется наравне с субъектно-объектным спряжением. Дистрибуция двух парадигм у переходных глаголов нетривиальна и является предметом дискуссий. Изначально считалось, что, как и в родственном венгерском, субъектно-объектное спряжение в хантыйском используется при определенном прямом объекте, а субъектное — при неопределенном [Honti 1986]. Однако в недавних работах использование спряжений анализируется в основном в терминах информационной структуры. Так, в работе [Nikolaeva 2001] на материале обдорского диалекта хантыйского утверждается, что субъектное спряжение используется в случае, если топику/теме соответствует только субъект (39), тогда как для использования субъектно-объектного спряжения необходимо, чтобы топику/теме соответствовал как субъект, так и прямой объект (40), ср. аналогичные выводы на материале восточных диалектов [Filchenko 2012: 120—122].

(39) *luw / Juwan Pe:tra re:sk-əs / *re:sk-əs-li*
он / Иван Петр ударить-PST[3SG] / ударить-PST-3SG>SG
{Кого ударил Иван?} ‘Он/Иван ударил Петра’ [Nikolaeva 2001: 16].

(40) *luw Pe:tra / luw-e:l re:sk-əs-li / *re:sk-əs*
он Петр / он-ACC ударить-PST-3SG>SG / ударить-PST[3SG]
{Что Иван сделал с Петром?} ‘Он ударил Петра/его’ [там же].

В свою очередь, в работе [Кошкарева 2002: 32] на материале казымского диалекта акцент делается на позиции фокуса/ремы, но условие использования субъектно-объектного спряжения оказывается примерно тем же: оно используется, если ремой является одно сказуемое, что соотносится с требованием тематичности субъекта и прямого объекта в [Nikolaeva 2001], ср. (41) с глаголами *nõx ałeməti* ‘поднять’ и *temətti* ‘выбросить’ в субъектно-объектном спряжении.

(41) *in ime-λ kim et-əm jūpijən*
этот женщина-POSS.3SG наружу выйти-PST.PTCP после
‘После того, как эта женщина наружу вышла’.
sorni an-ije-λ nõx ałəm-əs-λe
золотой чашка-DIM-POSS.3SG вверх взять-PST-3SG>SG
‘(Та женщина) взяла золотую чашку’.
pūn-λ-aλ jira temət-s-əλe
нитка-PL-3SG в.сторону выбросить-PST-3SG>SG
‘Все нитки выбросила’ [Кошкарева 2002: 33].

Выводы И. А. Николаевой и Н. Б. Кошкаревой в целом подтверждаются на имеющемся текстовом и полевом материале казымского диалекта, однако нуждаются в ряде дополнений и уточнений. Для этого рассмотрим более подробно различные типы информационно-структурных контекстов, опираясь на работу [Lambrecht 1994]. Поскольку информационная структура в большей степени зависима от контекста высказывания, чем другие обсуждаемые нами параметры, рассмотрим примеры из естественных текстов, а конкретнее — диалогические реплики, в которых информационно-структурные функции наиболее очевидным образом заданы. В контекстах предикатного фокуса, при котором субъект и объект являются топиками, как в ответной реплике⁸ в (42), в соответствии с уже известными данными используется субъектно-объектное спряжение.

⁸ В самом вопросе, однако, при таком же предикатном фокусе используется субъектное спряжение, что может быть связано с неассертивностью высказывания или какими-то другими особенностями вопросительных предло-

- (42) — *xil-en wŷ-s-ən?* — *xɔn wŷ-s-əm*
внук-POSS.2SG брать-PST-2SG NEG брать-PST-1SG>SG
‘Внука твоего ты взяла? — Я не брала’ [Корпус: «Женщина, выделяющая жилы», 39—40].

В контекстах сентенциального фокуса, при котором в фокус попадает все предложение, включая субъект, объект и глагол, как можно видеть в ответе в (43), используется субъектное спряжение, что также вписывается в существующие представления о рематичности/фокусности объекта при субъектном спряжении.

- (43) — *toja mānət lɛpəl-s-ən?*
зачем я.ACC обмануть-PST-2SG
‘Ты почему меня обманул?’
mitxɔ lɔp-əs:
слуга говорить-PST[3SG]
‘Слуга говорит.’
— *nāŋ kānši part-s-ən n̄ar ox-əp petr*
ты искать:NFIN.NPST велеть-PST-2SG лысый голова-ATTR Петр
‘Ты приказал мне найти лысого Петра’ [Корпус: «Находчивый слуга царя», 37—39].

Наконец, в контекстах аргументного фокуса, а конкретнее фокуса на прямом дополнении, наблюдается вариативность в выборе спряжения. Так, в одном фрагменте (44) в ответ на эксплицитно поставленный частный вопрос используется субъектное спряжение.

- (44) — *mōj-λ-an nāŋ lipsi-λ-ən?*
что-PL-POSS.2SG ты лизать-NPST-2SG
‘Что ты там слизываешь?’
— *pa sɔλ-λ-am kim karalə-s-əm, sɔλ-λ-am lɛ-λ-əm*
ADD кишка-PL-1SG наружу ковырять-PST-1SG кишка-PL-1SG есть-NPST-1SG
‘Я вытащил свои кишки, ем кишки’ [Корпус: «Лживый лис», 13, 14].

В другом фрагменте (45) в ответе на похожий вопрос используется уже субъектно-объектное спряжение.

- (45) — *χōjt-en nāŋ in lawəl-λ-en pa?*
человек-POSS.2SG ты сейчас ждать-NPST-2SG>SG ADD
‘Или ждешь кого-то?’
— *sa sa — lōp-əs balda lūw-el*
PRTC PRTC говорить-PST[3SG] Балда он-3SG
‘Да-да, — сказал Балда ему’
— *lawəl-λ-əm ma tōm aj pāləŋ*
ждать-NPST-1SG>SG я тот маленький туча
‘Жду я ту маленькую тучку’ [Корпус: «Сказка о попе и его работнике Балде», 186—188].

При внимательном взгляде на контрасты (44, 45) обнаруживается, что разницу в форме глагола можно объяснить различным референциальным статусом Р-участника. Так, выражение ‘та маленькая тучка’ в (45) описывает уникальный объект в текущем дискурсе, поэтому является определенным, тогда как выражение ‘кишки’ в (44) не описывает уникальный объект, поэтому является неопределенным.

Влияние разницы в референциальном статусе подтверждается и на элицитированных севернохантыйских примерах. В контекстах с аргументным фокусом на прямом дополнении, в которых дефолтно используется субъектное спряжение (ср. примеры 41 выше из обдорского и 46 из казымского диалекта), субъектно-объектное спряжение также может использоваться, если позицию прямого объекта занимает определенный участник (46b), и неграмматично при референтном неопределенном участнике (46a).

- (46) a. *pet'a-j-en mōlsər aj_iki χātsə-s / *χātsə-s-λe*
Петя-EP-POSS.2SG какой.INDEF мальчик ударить-PST[3SG] / ударить-PST-3SG>SG
{Кого ударил Петя?} ‘Петя ударил КАКОГО-ТО МАЛЬЧИКА’.
b. *pet'a-j-en tām aj_ike-n χātsə-s / χātsə-s-λe*
Петя-EP-POSS.2SG этот мальчик-POSS.2SG ударить-PST[3SG] / ударить-PST-3SG>SG
{Кого ударил Петя?} ‘Петя ударил ЭТОГО МАЛЬЧИКА’.

жений, которые на данный момент неизвестны. Мы ограничимся упоминанием этого факта, оставив его обсуждение за рамками данной работы.

Контексты определенных, но не упоминавшихся ранее участников в позиции прямого дополнения, определенность которых следует только из их уникальности в текущем дискурсе (ср. работу [Gundel et al. 1993], а также анализ маркирования прямого дополнения в терминах уникальности в бесермянском удмуртском [Сердобольская 2017]), устроены менее регулярно. Так, в приведенном ниже фрагменте (47) объектное спряжение используется только с Р-участником ‘печь’, но не с другими упоминаемыми участниками, хотя ни один из этих участников не был упомянут ранее. Вероятное объяснение состоит в том, что печь является наиболее значимой частью традиционного жилища, поэтому здесь она в наибольшей степени активирована, ср. похожее рассуждение в [Nikolaeva 2001: 36—39] о влиянии семантического отношения ‘часть-целое’ на использование объектного спряжения. Однако окончательные выводы могут быть сделаны только по итогам отдельного исследования подобных примеров на материале текстов.

- (47) *kort-a imōltijn si jōχat-s-əm, tāl χātł-ət sora pātłə-λ-ət,*
 деревня-DAT наконец FOC прийти-PST-1SG зима день-PL быстро темнеть-NPST-3PL
jett-a si jūw-əm-əs, lampa-j-ew nōχ wūšit-s-əw,
 вечер-DAT FOC стать-PNCT-PST лампа-EP-POSS.1PL вверх разжигать-PST-1PL
kor-ew āł-s-ew, mōłs-əm nir-a nōχ iχat-s-əm
 печка-POSS.1PL разжечь-PST-1PL>SG малица-POSS.1SG жердь-DAT вверх вешать-PST-1SG
 ‘Наконец я пришел на стойбище. Зимой быстро темнеет. Наступил вечер. Зажгли лампу. Затопили печь. Малицу на жердь повесил’ [Корпус: «Вечером», 2—7].

Для использования субъектно-объектных форм необходимо также, чтобы и А-участник занимал высокую позицию в референциальной иерархии, ср. референтный неопределенный А-участник при определенном (48) и неопределенном Р-участнике (49). Следовательно, для использования субъектно-объектного спряжения необходимым условием является определенность обоих участников.

- (48) *aj_iki pet'a-j-en χātśa-s / *χātśa-s-λe*
 мальчик Петя-EP-POSS.2SG ударить-PST[3SG] / ударить-PST-3SG>SG
 {Кого ударил мальчик?} ‘Мальчик ударил ПЕТЮ’.
- (49) *aj_iki amp χātśa-s / *χātśa-s-λe*
 мальчик собака ударить-PST[3SG] / ударить-PST-3SG>SG
 {Кого ударил мальчик?} ‘Мальчик ударил СОБАКУ’.

Значима также позиция А-участника в иерархии одушевленности: если А-участник — человек, допустимо использование обоих спряжений (50). Если же А-участник — животное, единственным вариантом остается субъектное спряжение (51) с тем условием, что Р-участник уступает А-участнику по одушевленности, в противном случае используется пассив (52).

- (50) *pet'a-j-en tām an-en šūkat-əs / šūkat-s-əλλe*
 Петя-EP-POSS.2SG этот чашка-POSS.2SG разбить-PST[3SG] / разбить-PST-3SG>SG
 {Что Петя разбил?} ‘Петя разбил ЭТУ ЧАШКУ’.
- (51) *amp-en tām an-en šūkat-əs / *šūkat-s-əλλe*
 собака-POSS.2SG этот чашка-POSS.2SG разбить-PST[3SG] / разбить-PST-3SG>SG
 {Что собака разбила?} ‘Собака разбила ЭТУ ЧАШКУ’.
- (52) *pet'a-j-en amp-ən toχəm-s-a*
 Петя-EP-POSS.2SG собака-LOC кусать-PST-PASS[3SG]
 {Что собака сделала с Петей?} ‘Собака Петю УКУСИЛА’.

Исключение составляют контексты типа (53), в которых животное приобретает топикальный статус в текущем фрагменте дискурса за счет того, что выполняет последовательность действий.

- (53) *χōłta aj wəj-lejki mān-əs, amōj jiləp χət wēr-əs,*
 куда маленький зверь идти-PST[3SG] или новый дом делать-PST[3SG]
amōj katra χət-əλ nōχ leśət-s-əλλe
 или старый дом-POSS.3SG вверх готовить-PST-3SG>SG
 ‘Куда хоть мышка делась: или новую норку она сделала, или старую норку починила’ [Корпус: «Мышка», 15].

Таким образом, выбор между субъектным и субъектно-объектным спряжением действительно опирается главным образом на информационную структуру в соответствии с наблюдениями [Nikolaeva

2001; Filchenko 2012; Кошкарева 2002] и других авторов. Субъектно-объектное спряжение используется в контекстах фокуса на глаголе с топикальными субъектом и объектом, тогда как в большинстве других случаев используется субъектное спряжение. Единственный контекст в казымском хантыйском, где наблюдается вариативность в выборе спряжения — контекст с аргументным фокусом на прямом дополнении. В этом случае выбор спряжения, как и в случае выбора между активом и пассивом, зависит от референциального статуса и одушевленности и также носит иерархический характер, ср. таблицу 7 ниже.

Таблица 7. Использование субъектного и субъектно-объектного спряжений в зависимости от исследуемых параметров в казымском диалекте севернохантыйского языка

А-участник	Р-участник	Субъектное спряжение	Субъектно-объектное спряжение
топик одушевленный определенный	топик одушевленный определенный	–	+
топик одушевленный	в фокусе (не)одушевленный	+	+
равные по референциальному статусу			
в фокусе неодушевленный	в фокусе (не)одушевленный	+	–
разные по референциальному статусу			

В целом, сопоставляя выбор между субъектным и субъектно-объектным спряжением с выбором между активным и пассивным залогом, можно сделать вывод о том, что для залога определяющей является прежде всего иерархическая позиция А-участника: при иерархически высоком А-участнике по обсуждаемым параметрам используется активный залог, в противном случае — пассив. Для спряжения же значима иерархическая позиция обоих участников: если оба они занимают высокую позицию на одной или нескольких иерархиях, используется субъектно-объектное спряжение, в противном случае — субъектное спряжение. Тем самым, мы получили ответ на первые два заявленных вопроса: 1) вариативное кодирование участников в севернохантыйском зависит от информационной структуры, лица, одушевленности и референциального статуса и 2) кодирование носит иерархический характер, так как зависит от положения участников на шкале, задаваемой данными параметрами.

4. Севернохантыйский язык: аккузативная и иерархическая трактовка

Рассмотрев параметры кодирования участников в севернохантыйском языке, мы можем обратиться к более общему вопросу 3 о том, какое место занимает иерархическое кодирование в системе севернохантыйского языка и как это соотносится с представлением о севернохантыйском как о языке с аккузативной стратегией ролевого кодирования. Как мы уже упомянули ранее в разделе 1.2, если рассматривать кодирование участников в строго формальном ключе, базовой стратегией кодирования в севернохантыйском языке является аккузативная стратегия, что следует из аккузативного оформления местоимений, первичного согласования глагола с S/A-участником и тяготения А-участника (переходного субъекта) к начальной позиции в предложении. При таком подходе севернохантыйская категория пассива подобно индоевропейским пассивам считается производным залогом и не рассматривается как часть системы кодирования. Однако, исключая пассив из рассмотрения, мы не только упускаем из вида общую иерархизированность кодирования, но также лишаемся возможности интерпретировать контексты, в которых пассивная конструкция является единственным способом выразить некоторую переходную пропозицию, ср. примеры (15, 16), повторенные ниже в (54, 55). В результате либо обязательное требование пассива в подобных контекстах приходится рассматривать как своего рода исключение из аккузативного анализа, либо такие контексты стоит выделять в самостоятельную, параллельную стратегию ролевого кодирования.

(54) *kalan ɣoj-na we:l-s-a? Juwan-na we:l-s-a*
 олень кто-LOC убить-PST-PASS[3SG] Иван-LOC убить-PST-PASS[3SG]
 ‘Кто убил оленя? Иван убил’ [Nikolaeva 2001: 25].

(55) *wət-ən maša-j-en ił pāwət-s-a*
 ветер-LOC Маша-EP-POSS.2SG вниз уронить-PST-PASS[3SG]
 ‘Порыв ветра сбил Машу’.

Альтернативный подход состоит в том, чтобы считать пассивный залог частью базовой стратегии кодирования, применяя, тем самым, к севернохантыйскому языку инструментарий для описания иерархических языков и приписывая пассиву функцию инверсива. Такой подход отходит от дифференциации стратегий по формальному принципу в сторону более функционального описания, а также позволяет охватить гораздо больше переходных контекстов, включая вышеупомянутые примеры. Тем не менее, и его применение имеет свои недостатки. Во-первых, в прототипических иерархических языках, таких как алгонкинские, ср. примеры (3, 4) из равнинного кри, сочетание иерархически низкого А-участника с высоким Р-участником маркируется только за счет инверсива (тип «*remapping inverse*», по [Zuñiga 2006: 64]), тогда как дополнительное именное маркирование А-участника косвенным падежом, как в языке пикурис (1, 2) и в севернохантыйском, является признаком более периферийной иерархической системы (тип «*passive inverse*», там же), для которой преимущества инверсивного анализа над пассивным менее очевидны. Во-вторых, инверсив в отличие от пассива как правило сохраняет переходность клаузы, ср. [Klaiman 1993: 348], тогда как севернохантыйская пассивная конструкция выглядит как непереходная: А-участник понижен до косвенного дополнения, и отсутствует выраженное прямое дополнение. В некотором смысле локативного А-участника в этой конструкции можно было бы рассматривать как неканонического субъекта, но дополнительных формальных свидетельств в пользу его подлежащей позиции не обнаруживается. В этом отношении примечателен пример восточнохантыйского языка, в котором, помимо базовой переходной конструкции с глаголом в субъектно-объектном спряжении (56), существует и другой вариант конструкции, в которой глагол также стоит в активном субъектно-объектном спряжении, но А-участник, как и при пассиве, маркируется локативом (57), поэтому такая конструкция характеризуется как эргативная переходная [Filchenko 2005: 31—33].

(56) *ma ämp-äm tiyl-a kari-mta-s-im*
я собака-POSS.1SG DEM-LAT тянуть-INTNS-PRF-1SG>SG
'Я притянул собаку ближе' [Filchenko 2005: 30].

(57) *alla sart män-nə löyöli-s-im*
большой щука я-LOC готовить-PTST-1SG>SG
'Я большую щуку приготовил' [Filchenko 2012: 122].

Не отдавая предпочтения какой-либо из двух трактовок, мы лишь констатируем, что севернохантыйский язык нельзя отнести ни к прототипическим аккузативным, ни к прототипическим иерархическим языкам. Формально он тяготеет к языкам с базовой аккузативной стратегией, что неудивительно, учитывая генетический и ареальный фон. Однако в своем функционировании он явно обнаруживает черты иерархической организации, которые частью укладываются в аккузативную модель, частью существуют в виде ряда исключений. В конечном итоге вопрос о том, которая из двух стратегий является для хантыйского языка базовой, представляется в большей степени дескриптивным, чем теоретическим, и ответ на него будет зависеть от целей конкретного исследователя.

Стоит также упомянуть возможные дальнейшие направления исследования.

Во-первых, что касается диалектной специфики севернохантыйской системы, то на данный момент известно об ориентации ролевого кодирования на топикальность участников в обдорском диалекте [Nikolaeva 2001] и в восточных диалектах [Filchenko 2012: 118—120], а также о дифференцированном маркировании А-участника в восточных диалектах в зависимости от агентивности/контроля [Filchenko 2005: 37—39]. Все это при определенном ракурсе может рассматриваться в ряду иерархических принципов кодирования участников и пополняться другими подобными принципами. Исследование данных явлений может способствовать пониманию общей диалектной и ареальной картины ролевого кодирования.

Во-вторых, интересным направлением представляется диахроническая эволюция севернохантыйской системы кодирования, которая привела к наблюдаемому в настоящий момент диалектному разнообразию. Поскольку мы располагаем ранними текстами только на обдорском диалекте, помимо их детального изучения и сопоставления с современными текстами, для этой цели необходимо углубленное сопоставление современных текстов на разных севернохантыйских диалектах с привлечением материала ареально и генетически близких языков.

В-третьих, остается открытым вопрос о кодировании участников в менее прототипических с точки зрения (морфо)синтаксиса и семантики контекстах, например, в неканонических переходных и дитранзитивных конструкциях, а также в эвиденциальных и императивных формах, имеющих собственные парадигмы спряжения. В этих контекстах могут обнаруживаться отклонения от наблюдаемого кодирования в силу особенностей морфологии, синтаксиса, семантики и диахронического развития используемых конструкций. Рассмотрение этих контекстов позволит существенно расширить наши представления об общих механизмах кодирования участников.

Заключение

В статье рассмотрен вопрос о кодировании участников в севернохантыйском языке на материале полевых и текстовых данных казымского диалекта. Севернохантыйский язык принято описывать как аккумулятивный язык, в котором ролевое кодирование в значительной мере определяется информационной структурой. В ходе исследования было установлено, что, помимо информационной структуры, важную роль играют также параметры лица, одушевленности и референциального статуса и что выбор залога и глагольного согласования во многом носит иерархический характер. По итогам анализа был сделан вывод о том, что базовая стратегия кодирования в севернохантыйском формально тяготеет к аккумулятивной, тогда как функционально она близка к системам иерархических языков.

Сокращения

1, 2, 3 — лицо	LAT — латив
I, II — согласовательные парадигмы	LOC — локатив
A — агенс	NEG — отрицание
ACC — аккумулятив	NFIN — нефинитная форма
ADD — аддитив	NPST — непрошедшее время
ADJ — адъективизатор	NSG — неединственное число
ATTR — атрибутивизатор	OBJ — объектное спряжение
DAT — датив	OBL — косвенная форма
DEF — определенность	P — пациенс
DEM — демонстратив	PASS — пассив
DIM — диминутив	PL — множественное число
DIR — директив	PNCT — пунктив
DU — двойственное число	POSS — посессив
EP — эпентеза	PRF — перфект
FOC — фокус	PRTC — частица
IMP — императив	PST — прошедшее время
INDEF — неопределенность	PTCP — причастие
INTNS — интенсив	S — единственный участник
INV — инверсив	SG — единственное число
	SUBJ — субъектное спряжение

Литература

- Вальгамова и др. 2011 — Вальгамова С. И., Кошкарева Н. Б., Онина С. В., Шиянова А. А. Диалектологический словарь хантыйского языка (шурышкарский и приуральский диалекты) / Ред. Н. Б. Кошкарева. Екатеринбург, 2011.
- Каксин 2010 — Каксин А. Д. Казымский диалект хантыйского языка. Изд. 2-е, доп. Ханты-Мансийск, 2010.
- Кошкарева 2002 — Кошкарева Н. Б. Коммуникативная парадигма хантыйского предложения // Языки коренных народов Сибири. Вып. 12. Новосибирск, 2002. С. 29—44.
- Михайлов 2018 — Михайлов С. К. Почему хантыйские имена нуждаются в определенном артикле? // Пятнадцатая Конференция по типологии и грамматике для молодых исследователей. Тезисы докладов (Санкт-Петербург, 22—24 ноября 2018 г.) / Ред. Д. Ф. Мищенко. СПб, 2018. С. 79—81.
- Николаева 1995 — Николаева И. А. Обдорский диалект хантыйского языка. М.; Гамбург, 1995.
- Падучева 2001 — Падучева Е. В. Каузативный глагол и декаузатив в русском языке // Русский язык в научном освещении. 2001, 1. С. 52—79.
- Сердобольская 2017 — Сердобольская Н. В. Инвариант определенности и дифференцированное маркирование прямого дополнения в бесермянском удмуртском // Acta Linguistica Petropolitana. Труды института лингвистических исследований. 2017, 3 (13). С. 76—122.
- Соловар 2010 — Соловар В. Н. Статальные элементарные простые предложения с пассивными формами глагольного сказуемого в казымском диалекте хантыйского языка // Вестник НГУ. Серия: История, филология. 2010, 2 (9). С. 91—98.
- Терешкин 1981 — Терешкин Н. И. Словарь восточнохантыйских диалектов. Л., 1981.
- Штейниц 1937 — Штейниц В. К. Хантыйский (остяцкий) язык // Языки и письменность народов Севера: Ч. 1. Языки и письменность самоедских и финно-угорских народов / Ред. Г. Н. Прокофьева. М.; Л., 1937. С. 193—228.

Bárány 2015 — *Bárány A.* Differential object marking in Hungarian and the morphosyntax of case and agreement. PhD thesis. Cambridge, 2015.

Bickel, Nichols 2009 — *Bickel B., Nichols J.* Case marking and alignment // *The Oxford handbook of case* / Eds. A. Malchukov, A. Spencer. Oxford, 2009. P. 304—321.

Bickel et al. 2013 — *Bickel B., Iemmolo G., Zakharko T., Witzlack-Makarevich A., Bakker D., Haspelmath M.* Patterns of alignment in verb agreement // *Languages across boundaries: Studies in memory of Anna Siewierska* / Eds. D. Bakker, M. Haspelmath. Berlin, 2013. P. 15—36.

Blust 2013 — *Blust R.* The Austronesian languages. Revised edition. Canberra, 2013.

Comrie 1981 — *Comrie B.* Language universals and linguistic typology: Syntax and morphology. Oxford, 1981.

Dixon 1972 — *Dixon R. M. W.* The Dyirbal language of north Queensland. Cambridge, 1972.

Dixon, Aikhenvald 1997 — *Dixon R. M. W., Aikhenvald A.* A typology of argument-determined constructions // *Essays on Language Function and Language Type. Dedicated to T. Givón* / Eds. J. Bybee, J. Haiman, S. Thompson. Amsterdam, 1997. P. 71—113.

Filchenko 2005 — *Filchenko A.* Non-canonical agent-marking in Eastern Khanty: A functional-pragmatic perspective // *University of Pennsylvania Working Papers in Linguistics*. 2005, 1 (11). P. 29—40.

Filchenko 2012 — *Filchenko A.* Continuity of information structuring strategies in Eastern Khanty // *Argument Structure and Grammatical Relations. A crosslinguistic typology* / Eds. P. Suihkonen, B. Comrie, V. Solovyev. Amsterdam, 2012. P. 115—131.

Givón 1979 — *Givón T.* On understanding grammar. New York, 1979.

Givón 1983 — *Givón T.* Topic continuity in discourse: A quantitative cross-language study. Amsterdam, 1983.

Gundel et al. 1993 — *Gundel J. K., Hedberg N., Zacharski R.* Cognitive status and the form of referring expressions in discourse // *Language*. 1993, 2 (69). P. 274—307.

Haig 2018 — *Haig G.* The grammaticalization of object pronouns: Why differential object indexing is an attractor state // *Linguistics*. 2018, 4 (56). P. 781—818.

Haspelmath 2011 — *Haspelmath M.* On S, A, P, T, and R as comparative concepts for alignment typology // *Linguistic Typology*. 2011, 3 (15). P. 535—567.

Honti 1986 — *Honti L.* *Chrestomathia ostiatica*. Budapest, 1986.

Iemmolo 2011 — *Iemmolo G.* Towards a typological study of differential object marking and differential object indexing. Doctoral dissertation. Pavia, 2011.

Iemmolo, Klumpp 2014 — *Iemmolo G., Klumpp G.* Introduction to the special issue “Differential Object Marking: theoretical and empirical issues” // *Linguistics*. 2014, 2 (52). P. 271—279.

Keenan 1985 — *Keenan E. L.* Passive in the world’s languages // *Language Typology and Syntactic Description*, vol. 1: *Clause Structure* / Ed. *T. Shopen*. Cambridge, 1985. P. 243—281.

É. Kiss 2013 — *É. Kiss K.* The inverse agreement constraint in Uralic languages // *Finno-Ugric Languages and Linguistics*. 2013, 1 (2). P. 2—21.

Kibrik 1997 — *Kibrik A. E.* Beyond subject and object: Toward a comprehensive relational typology // *Linguistic Typology*. 1997, 1. P. 279—346.

Klaiman 1993 — *Klaiman M. H.* The relationship of inverse voice and head-marking in Arizona Tewa and other Tanoan languages. *Studies in Language* // *International Journal sponsored by the Foundation “Foundations of Language”*. 1993, 2 (17). P. 343—370.

Kulonen 1989 — *Kulonen U. M.* The Passive in Ob-Ugrian. Helsinki, 1989.

Lambrecht 1994 — *Lambrecht K.* Information structure and sentence form: topic focus and the mental representation of discourse referents. Cambridge, 1994.

Mikhailov 2021 — *Mikhailov S.* Semantics of the Northern Khanty 2sg possessive and of related and competing markers. MA dissertation. Moscow, 2021.

Nichols 1992 — *Nichols J.* Linguistic diversity in space and time. Chicago, 1992.

Nikolaeva 1999 — *Nikolaeva I.* *Ostyak*. Munich, 1999.

Nikolaeva 2001 — *Nikolaeva I.*, 2001. Secondary topic as a relation in information structure // *Linguistics*. 2001, 1 (39). P. 1—50.

Shibatani 1985 — *Shibatani M.* Passives and related constructions: A prototype analysis // *Language*. 1985, 4 (61). P. 821—848.

Shibatani 1988 — *Passive and voice* / Ed. M. Shibatani. Amsterdam; Philadelphia, 1988.

Siewierska 2003 — *Siewierska A.* Person agreement and the determination of alignment // *Transactions of the Philological Society*. 2003, 2 (101). P. 339—370.

Siewierska 2013 — *Siewierska A.* Alignment of Verbal Person Marking // *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology / Eds. Dryer M. S., Haspelmath M. 2013. Vol. 406. URL: <http://wals.info/chapter/100>.

- Steinitz 1950 — *Steinitz W.* Ostjakische Grammatik und Chrestomatie mit Wörterverzeichnis. 2. Verb. Leipzig, 1950.
- Thompson 1994 — *Thompson C.* Passive and inverse constructions // *Voice and inversion* / Ed. T. Givón. 1994. Vol. 28. P. 47—63.
- Virtanen, Sosa 2018 — *Virtanen S., Sosa S.* Remarks on areal linguistics in the information structure of the Ob-Ugrian languages // *Folia Uralica Debreceniensia*. 2018, 25. P. 233—260.
- Watters 2002 — *Watters D.* A grammar of Kham. Cambridge, 2002.
- Zuñiga 2006 — *Zuñiga F.* Deixis and alignment: Inverse systems in indigenous languages of the Americas. Amsterdam, 2006.

References

- Bárány 2015 — Bárány A. Differential object marking in Hungarian and the morphosyntax of case and agreement. PhD thesis. Cambridge, 2015.
- Bickel, Nichols 2009 — Bickel B., Nichols J. Case marking and alignment. In: *The Oxford handbook of case* / Eds. A. Malchukov, A. Spencer. Oxford, 2009. P. 304—321.
- Bickel et al. 2013 — Bickel B., Iemmolo G., Zakharko T., Witzlack-Makarevich A., Bakker D., Haspelmath M. Patterns of alignment in verb agreement. In: *Languages across boundaries: Studies in memory of Anna Siewierska* / Eds. D. Bakker, M. Haspelmath. Berlin, 2013. P. 15—36.
- Blust 2013 — Blust R. *The austronesian languages*. Revised edition. Canberra, 2013.
- Comrie 1981 — Comrie B. *Language universals and linguistic typology: Syntax and morphology*. Oxford, 1981.
- Dixon 1972 — Dixon R. M. W. *The Dyirbal language of north Queensland*. Cambridge, 1972.
- Dixon, Aikhenvald 1997 — Dixon R. M. W., Aikhenvald A. A typology of argument-determined constructions. In: *Essays on Language Function and Language Type. Dedicated to T. Givón* / Eds. J. Bybee, J. Haiman, S. Thompson. Amsterdam, 1997. P. 71—113.
- Filchenko 2005 — Filchenko A. Non-canonical agent-marking in Eastern Khanty: A functional-pragmatic perspective. In: *University of Pennsylvania Working Papers in Linguistics*. 2005, 1 (11). P. 29—40.
- Filchenko 2012 — Filchenko A. Continuity of information structuring strategies in Eastern Khanty. In: *Argument Structure and Grammatical Relations. A crosslinguistic typology* / Eds. P. Suihkonen, B. Comrie, V. Solovyev. Amsterdam, 2012. P. 115—131.
- Givón 1919 — Givón T. *On understanding grammar*. New York, 1979.
- Givón 1983 — Givón T. *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam, 1983.
- Gundel 1993 — Gundel J. K., Hedberg N., Zacharski R. Cognitive status and the form of referring expressions in discourse. In: *Language*. 1993, 2 (69). P. 274—307.
- Haig 2018 — Haig G. The grammaticalization of object pronouns: Why differential object indexing is an attractor state. In: *Linguistics*. 2018, 4 (56). P. 781—818.
- Haspelmath 2011 — Haspelmath M. On S, A, P, T, and R as comparative concepts for alignment typology. In: *Linguistic Typology*. 2011, 3 (15). P. 535—567.
- Honti 1986 — Honti L. *Chrestomathia ostiacica*. Budapest, 1986.
- Iemmolo 2011 — Iemmolo G. *Towards a typological study of differential object marking and differential object indexing*. Doctoral dissertation. Pavia, 2011.
- Iemmolo, Klumpp 2014 — Iemmolo G., Klumpp G. Introduction to the special issue “Differential Object Marking: theoretical and empirical issues”. In: *Linguistics*. 2014, 2 (52). P. 271—279.
- Kaksin 2010 — Kaksin A. D. *Kazym dialect of Khanty*. 2nd edition. Hanty-Mansijsk, 2010.
- Keenan 1985 — Keenan E. L. Passive in the world’s languages. In: *Language Typology and Syntactic Description*. Vol. 1: *Clause Structure* / Ed. T. Shopen. Cambridge, 1985. P. 243—281.
- É. Kiss 2013 — É. Kiss K. The inverse agreement constraint in Uralic languages. In: *Finno-Ugric Languages and Linguistics*. 2013, 1 (2). P. 2—21.
- Kibrik 1997 — Kibrik A. E. Beyond subject and object: Toward a comprehensive relational typology. In: *Linguistic Typology*. 1997, 1. P. 279—346.
- Klaiman 1993 — Klaiman M. H. The relationship of inverse voice and head-marking in Arizona Tewa and other Tanoan languages. Studies in Language. In: *International Journal sponsored by the Foundation “Foundations of Language”*. 1993, 2 (17). P. 343—370.
- Koshkareva 2002 — Koshkareva N. B. Communicative paradigm of Khanty sentence. In: *Languages of Indigenous Peoples of Siberia*. Issue 12. Novosibirsk, 2002. P. 29—44.
- Kulonen 1994 — Kulonen U. M. *The Passive in Ob-Ugrian*. Helsinki, 1989.
- Lambrecht 1994 — Lambrecht K. *Information structure and sentence form: topic focus and the mental representation of discourse referents*. Cambridge, 1994.

- Mihajlov 2018 — Mihajlov S. K. Why do Khanty nouns need a definite article? In: 15th Conference on Typology and Grammar of Young Researchers. Book of Abstracts. (Saint Petersburg, 22—24 November 2018) / Ed. D. F. Mishhenko. Saint Petersburg, 2018. P. 79—81.
- Mikhailov 2021 — Mikhailov S. Semantics of the Northern Khanty 2sg possessive and of related and competing markers. MA dissertation, Moscow, 2021.
- Nichols 1992 — Nichols J. Linguistic diversity in space and time. Chicago, 1992.
- Nikolaeva 1995 — Nikolaeva I. A. Obdorsk dialect of Khanty. Moscow, Hamburg, 1995.
- Nikolaeva 1999 — Nikolaeva I. Ostyak. Munich, 1999.
- Nikolaeva 2001 — Nikolaeva I., 2001. Secondary topic as a relation in information structure. In: *Linguistics*. 2001, 1 (39). P. 1—50.
- Paducheva 2001 — Paducheva E. V. Causative verb and decausative in Russian. In: *Russian Language and Linguistic Theory*. 2001, 1. P. 52—79.
- Serdobol'skaja 2017 — Serdobol'skaja N. V. Invariant of definiteness and differential object marking in Beserman Udmurt. In: *Acta Linguistica Petropolitana*. 2017, 3 (13). P. 76—122.
- Shibatani 1985 — Shibatani M. Passives and related constructions: A prototype analysis. In: *Language*. 1985, 4 (61). P. 821—848.
- Shibatani 1988 — *Passive and voice* / Ed. M. Shibatani. Amsterdam, 1988.
- Siewierska 2003 — Siewierska A. Person agreement and the determination of alignment. In: *Transactions of the Philological Society*. 2003, 2 (101). P. 339—370.
- Siewierska 2013 — Siewierska A. Alignment of Verbal Person Marking. In: *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology / Eds. Dryer M. S., Haspelmath M. 2013. Vol. 406. (URL: <http://wals.info/chapter/100>.)
- Steinitz 1937 — Steinitz W. Khanty (Ostjak) language. In: *Languages and writings of the peoples of the North: Part 1. Languages and writings of the Samoyed and Finno-Ugric peoples* / Ed. G. N. Prokof'eva. Moscow; Leningrad, 1937. P. 193—228.
- Solovar 2010 — Solovar V. N. Statal elementary simple sentences with passive forms of the verbal predicate in Kazym Khanty. In: *Bulletin of the NSU. Series: History, Philology*. 2010, 2 (9). P. 91—98.
- Steinitz 1950 — Steinitz W. *Ostjakische Grammatik und Chrestomatie mit Wörterverzeichnis*. 2. Verb. Leipzig: Harrassowitz, 1950.
- Tereshkin 1981 — Tereshkin N. I. *Dictionary of Eastern Khanty dialects*. Leningrad, 1981.
- Thompson 1994 — Thompson C. Passive and inverse constructions. In: *Voice and inversion* / Ed. T. Givón. 1994. Vol. 28. P. 47—63.
- Val'gamova i dr. 2011 — Val'gamova S. I., Koshkareva N. B., Onina S. V., Shijanova A. A. *Dialectological dictionary of Khanty (Shuryshkary and Priuralsk dialects)* / Ed. N. B. Koshkareva. Ekaterinburg, 2011.
- Virtanen, Sosa 2018 — Virtanen S., Sosa S. Remarks on areal linguistics in the information structure of the Ob-Ugrian languages. In: *Folia Uralica Debreceniensia*. 2018, 25. P. 233—260.
- Watters 2002 — Watters D. *A grammar of Kham*. Cambridge, 2002.
- Zuñiga 2006 — Zuñiga F. *Deixis and alignment: Inverse systems in indigenous languages of the Americas*. Amsterdam, 2006.

Phonetic and lexical innovations in Ob-Ugric dialects in the 18th—21st centuries: new archival and field data

Yulia V. Normanskaya, Institute of Linguistics, RAS, Ivannikov Institute for System Programming, RAS (Moscow); julianor@mail.ru
Idalia V. Fedotova, HSE University, Ivannikov Institute for System Programming, RAS (Moscow); idal.fedotova@gmail.com

This paper offers a comprehensive analysis of proximity between the dialects within Khanty and Mansi languages, which form the Ob-Ugric group within the Uralic languages. Even though they are thought to be closely related, research has shown the discrepancy between the phonological and lexicostatistical distance between Khanty varieties vs. Mansi varieties.

In the article, innovations in phonology, morphology and basic vocabulary are brought together and compared. New data from archives (18th century) and latest fieldwork shed the light on chronology of phonetic, morphological and lexical changes. Using comparative method, it is shown that in the late 18th century Mansi dialects were still not very distant from each other. Moreover, phonetic differences which hindered mutual understanding between speakers of eastern and northern dialects in the 20th century became distinctive not earlier than 250 years ago. On the contrary, Khanty dialects were as distant as separate languages back in the late 18th century. Judging by glottochronological calculations, the time of divergence of contemporary Kazym and Vakh Khanty can be dated back as early as the beginning of the first millennium AD. These Khanty varieties have more differences than any Slavic and even Turkic (except Chuvash) languages between each other. Therefore, they should be counted as different languages, not dialects.

Overall, from the extensive analysis of phonetics, morphology and basic vocabulary, it is proved that Mansi varieties were still dialects (until all but one became extinct) and Khanty varieties were separate languages when they were first recorded in the late 1700s.

Keywords: Ob-Ugric languages, Khanty, Mansi, phonetic innovations, lexicostatistics, glottochronology

ФОНЕТИЧЕСКИЕ И ЛЕКСИЧЕСКИЕ ИННОВАЦИИ В ДИАЛЕКТАХ ОБСКО-УГОРСКИХ ЯЗЫКОВ В XVIII—XXI ВВ.: НОВЫЕ АРХИВНЫЕ И ПОЛЕВЫЕ МАТЕРИАЛЫ

Норманская Юлия Викторовна, Институт языкознания РАН,
Институт системного программирования им. В. П. Иванникова РАН (Москва); julianor@mail.ru
Федотова Идалия Вячеславовна, Институт системного программирования им. В. П. Иванникова РАН,
Национальный исследовательский университет «Высшая школа экономики» (Москва); idal.fedotova@gmail.com

В статье проводится комплексный анализ близости между диалектами хантыйского и мансийского языков — обско-угорской ветви уральской семьи. Несмотря на то, что эти языки традиционно считаются близкородственными, исследования показывают неравномерность в близости хантыйских диалектов между собой по сравнению с мансийскими по фонетическим изоглоссам и процентам совпадения в базисной лексике.

Исследование сводит воедино и сопоставляет инновации в фонологии, морфологии и базисной лексике обско-угорских языков. Новые материалы из архивов (XVIII в.) и новейшие полевые исследования проливают свет на хронологию фонетических, морфологических и лексических изменений. При помощи сравнительно-исторического метода показано, что в конце XVIII в. мансийские диалекты все еще были достаточно близки, а фонетические различия, которые мешали взаимопониманию между носителями восточных и северных диалектов в XX в., стали появляться не ранее 250 лет назад. Напротив, еще в конце XVIII в. хантыйские диалекты были далеки друг от друга, как отдельные языки. По глоттохронологическим расчетам, время распада предка современных казымских и ваховских ханты можно датировать еще началом первого тысячелетия нашей эры. Эти хантыйские идиомы имеют больше различий, чем любые славянские и даже тюркские (кроме чувашского) языки между собой.

Из результатов комплексного анализа фонетики, морфологии и базисной лексики следует, что мансийские идиомы (до исчезновения восточных диалектов в начале XXI в.) все еще были диалектами, хантыйские идиомы были уже отдельными языками, когда они были впервые засвидетельствованы в конце XVIII в.

Ключевые слова: обско-угорские языки, хантыйский язык, мансийский язык, фонетические инновации, лексикостатистика, глоттохронология

The paper was prepared with the aid of Russian Scientific Fund, grant № 20-18-00403 “Digital description of the dialects the Uralic languages based on the analysis of big data”.

One of the most prominent issues in comparative historical linguistics is the possibility to date language divergence. Such significance is connected with the fact that datings are used by other researchers: geneticists, archaeologists, historians. As it is so important to get the most precise datings, comparative linguists should assess the time of language divergence not only using lexicostatistical calculations (suggested by M. Swadesh and modified and refined by S. A. Starostin), but also taking into account innovation processes on other levels of language system: phonetics and morphology.

At present, a great number of the first dictionaries and books in Uralic and Altaic languages created in the 18th—19th centuries have been found in archives of Saint-Petersburg and Kazan, as well as in libraries outside Russia. Some of them have been digitally processed. Based on these data, online etymological dictionaries were created that connect materials of the archival books with contemporary dialects and glossed corpora at the platform LingvoDoc (www.lingvodoc.ispras.ru). Having substantial number of archival sources of 18th—19th centuries in almost all divisions of the Uralic and Altaic languages at our disposal allows us to assess change rate in dialects for the last 250 years. Using computational methods for the full available language material (on average 400 to 1500 lexemes), it is possible to calculate the number of phonetic and morphological innovations and make correspondences between these data and proximity rates gained by lexicostatistics for each of the languages.

Such analysis should answer the question if it is possible to develop a speed rate of change not only for the lexicon, but also for morphology and phonetics. Are there instances when phonetic and morphological changes happen very fast and they should not be taken into account like influx of lexical borrowings during intensive contacts with speakers of other languages? Should all the phonetic innovations be equally assessed? Or instead should we distinguish between the instances when sounds are close to each other by place and manner of articulation, e. g. **k > x*, and when are very different, e. g. **l > t*, because the latter undergo a series of intermediary changes?

A final answer to these questions needs further work, as the sufficient statistics for changes should be compiled to answer them. In this paper we are going to show a piece of analysis of Ob-Ugric material from this perspective. All the available archival materials and field sound recordings have been presented to the scholarly community. At present, 36 Khanty and 21 Mansi dictionaries and concordances are available on the LingvoDoc platform. As has been mentioned, the average volume of the dictionaries is 800 lexemes.

M. A. Zhivlov has conducted lexicostatistical calculations mostly based on the dictionaries that were compiled from the materials collected in the 19th century: Mansi [Munkácsi, Kálmán 1986] and Khanty [DEWOs] with addition of contemporary data on the dialects [Tereshkin 1961; Solovar 2006]. For Mansi the minimal lexicostatistical distance between northern and eastern dialects is 87%, for Khanty it is 73%.

Overall these results are quite unusual, as 65—73% distance between Khanty dialects noted by researchers [Zhivlov 2011; Fedotova in print] are typical not for dialects, but for different languages of Slavic, Germanic and Turkic families, according to [Burlak, Starostin 2005]. Consequently, M. A. Zhivlov dates the divergence of Mansi dialects 300—400 AD and Khanty dialects 100—200 AD (see [Zhivlov 2011]).

Surprisingly, morphological differences in Mansi dialects “are not significant or substantial for classification of dialects”, as is shown in [Koshelyuk 2021: 279], the vast majority of morphological markers in Mansi dialects are the same, whereas the differences between them are caused by phonetic processes that have taken place in these dialects. On the contrary, western and northern Khanty dialects have big differences in morphology, for example the number of nominal cases in the western are 3—4, whereas there are 8—9 in the eastern, according to [Solovar, Nakhracheva, Shiyanova 2016]. Is it possible that the time of divergence of Mansi and Khanty was only two centuries apart, the speakers lived close to each other and in the similar language environment, but in Mansi dialects morphological markers have survived without change, while in Khanty dialects they differ more than in Permic, Fennic or Slavic languages?

The first 18th century dictionaries that we found in Saint-Petersburg Archive of the Russian Academy of Sciences and uploaded to the platform LingvoDoc allow us to solve the issue of such discrepancy of the rate of morphological differences in dialects of Mansi and Khanty.

I. Differences in graphics and phonetics in manuscripts of Mansi dialects of the 18th century

As J. Gulya shows based on unpublished archival data, and Yu. Normanskaya based on the online dictionaries on LingvoDoc (see further [Gulya 1958, 1963; Normanskaya 2022]), **Mansi dialects almost did not have phonetical differences in the 18th century**. The ample differences recorded since the early 19th century only started to emerge in the 18th as duplicate forms.

See below the table for distinguishing features of Mansi dialects based on the following dictionaries¹:

- northern dialects compiled by P. S. Pallas in village Beryozovo:
<http://lingvodoc.ispras.ru/dictionary/3066/11/perspective/3066/14/view>
- unlocated, found in A. Sjögren's archive in Saint Petersburg, having both northern and western features, see [Normanskaya, Kosheliuk 2020] for analysis:
<http://lingvodoc.ispras.ru/dictionary/1393/29131/perspective/1393/29132/view>
- western dialects, published by P. S. Pallas:
Solikamsk: <http://lingvodoc.ispras.ru/dictionary/867/9/perspective/867/10/view>
Cherdyn: <http://lingvodoc.ispras.ru/dictionary/2685/846/perspective/2685/847/view>
Kungur: <http://lingvodoc.ispras.ru/dictionary/2685/6/perspective/2685/7/view>
Verkhoturys: <http://lingvodoc.ispras.ru/dictionary/2685/1653/perspective/2685/1654/view>
Ust'-Uls, recorded by Major Karpinsky in the M. A. Castrén's archive:
<http://lingvodoc.ispras.ru/dictionary/3096/54761/perspective/3096/54762/view>

Table 1. Reflexes of Proto-Mansi phonemes in the first dictionaries of 18th century

Proto-Mansi, Honti	Proto-Mansi [Normanskaya 2022]	Beryozovo Pallas	Unlocated Pallas	Solikamsk Cherkalov	Cherdyn Pallas	Kungur Pallas	Verkhoturys Pallas	Ust'-Uls Castrén
* <i>ā</i>	* <i>ā</i>	<i>o/a</i>	<i>o</i>	<i>o</i>	<i>o/a</i>	<i>o/a</i>	<i>o/a</i>	<i>o/a/oa</i>
* <i>ī</i>	* <i>a</i>	<i>a/e</i>	<i>a/aa</i>	<i>e/a</i>	<i>a/i</i>	<i>a/e</i>	<i>o/a</i>	<i>o/e/ea/aa/u</i>
* <i>ȳ</i>	* <i>ȳ</i>	<i>0/-y/-j</i>	<i>0/-y/-u/-j</i>	<i>0/-ȳ</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0/-ȳ/-χ/-j</i>
* <i>č</i>	* <i>č</i>	<i>č/-z-</i>	<i>č/-z-</i>	<i>č/s</i>	<i>č/š</i>	<i>s</i>	<i>č</i>	<i>č/s</i>
* <i>k</i> _Vback		<i>χ</i>	<i>k/χ</i>	<i>k</i>	<i>k</i>	<i>k</i>	<i>k</i>	<i>k</i>
* <i>š</i>		<i>š</i>	<i>š</i>	<i>š</i>	<i>š</i>	<i>š</i>	<i>š</i>	<i>š</i>

As table 1 demonstrates, Proto-Mansi phonemes that are relevant for classifying Mansi dialects in the late 19th century (the first column)² did not have the distinct reflexes in the western and northern dialects. The Proto-Mansi shift **ā* > *o*, which occurred in the 19th century in all dialects except southern, according to L. Honti, in the 19th century was not completed in Beryozovo (northern), Kungur, Verkhoturys, Cherdyn and Ust'-Uls (western) dialects.

II. Differences in graphics and phonetics in manuscripts of Khanty dialects of the 18th century

On the contrary, **Khanty dialects were significantly different in phonetics**, see below table 2 based on the analysis of Khanty 18th century materials from A. Sjögren's fund № 94 "Linguistic materials collected by academician P. S. Pallas":

- 1) Narym district (Vasjogan dialect):
<http://lingvodoc.ispras.ru/dictionary/2639/4152/perspective/2639/4156/view>;
- 2) Vasjogan dialect: <http://lingvodoc.ispras.ru/dictionary/2059/5674/perspective/2059/5675/view>;
- 3) Surgut district (Lumpokolskoye dialect):
<http://lingvodoc.ispras.ru/dictionary/2639/77/perspective/2639/81/view>;
- 4) Salym dialect:
<http://lingvodoc.ispras.ru/dictionary/1827/6/perspective/1827/7/view>;
- 5) Irtysh dialect: <http://lingvodoc.ispras.ru/dictionary/1674/4209/perspective/1674/4210/view>;
- 6) Beryozovo dialect: <http://lingvodoc.ispras.ru/dictionary/2639/4221/perspective/2639/4225/view>;
- 7) dialect of Tobolsk district:
<http://lingvodoc.ispras.ru/dictionary/2516/17279/perspective/2516/17283/view>.

The table shows that the majority of changes occurred in Tobolsk dialect of the 18th century. It saw the following innovations: Proto-Khanty **a* > *o*, **ɔ* > *u*, **l* > *t/tl*, **č* > *t*, **w* > *0*_u and a number of processes that had started but were not completed: **č* > *š*, **λ* > *tl*, **kV* > *x*.

¹ For more information about the places of residence of Mansi in the 17th—20th centuries, see [Koryakov 2022].

² More information about Proto-Mansi phonemes that are relevant for classifying Mansi dialects see in [Honti 1988].

Table 2. Main phonetic differences in 18th century Khanty dialects by differentiating features suggested by L. Honti

Proto-Khanty	eastern			southern			northern
	Vasjugan-Narym	Vasjugan	Surgut	Salym	Irtysk	Tobolsk	Beryozovo
* <i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>o</i>	<i>o</i>	<i>o</i>	<i>o/a</i>
* <i>ɔ</i>	<i>o</i>	<i>o</i>	<i>o</i>	<i>u/(o)</i> ³	<i>u/(o)</i>	<i>u</i>	<i>o</i>
* <i>ʌ</i> -	<i>o/j</i>		<i>l</i>	<i>l</i>	<i>l</i>	<i>l/tl</i>	<i>l</i>
* <i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>l</i>	<i>t/tl</i>	<i>l</i>
* <i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
* <i>č</i>	<i>č/t'</i>	<i>č</i>	<i>č</i>	<i>č</i>	<i>š/č</i>	<i>š/č</i>	<i>š</i>
* <i>ć</i>						<i>t</i>	<i>š</i>
* <i>k(V)</i>	<i>k</i>	<i>k</i>	<i>k</i>	<i>x</i>	<i>x</i>	<i>x/k</i>	<i>x</i>
* <i>k(V̇)</i>	<i>k</i>	<i>k</i>	<i>k</i>	<i>k</i>	<i>k</i>	<i>k</i>	<i>k</i>
* <i>w</i> -	<i>v</i>	<i>v</i>	<i>v</i>	<i>u</i>	<i>v</i>	<i>v/0(u)</i>	<i>v</i>

Irtysk and Salym dialects saw only three out of these innovations: **kV* > *x*, **a* > *o*, **ɔ* > *u*. The shift **č* > *š* also started in Irtysk. Beryozovo dialect underwent three innovative changes *kV* > *x*, **č* > *š*, **ć* > *š*, and **a* > *o* started to emerge. Vasjugan dialect of that time underwent only one innovation **ʌ* > *o/j*. Surgut dialect preserved the Proto-Khanty phonemes in the most archaic state.

Thus, it can be concluded that the western (Beryozovo, Irtysk, Salym) and eastern (Vasjugan, Surgut) dialects of the 18th century, which remained until the 19th century, were considerably closer to each other than to the Tobolsk dialect, which disappeared quite early and in the 18th century was undoubtedly a separate dialectal group. It was distinguished from the known Khanty dialects of that time by the following processes: Proto-Khanty **l* > *t/tl*, **ʌ* > *tl*, **ć* > *t*, **w* > *0|u*.

Overall, Mansi dialects in the 18th century did not have distinct isoglosses that would have divided them. Meanwhile, Tobol dialect of Khanty was significantly different from the other dialects. This correlates with the data of considerable morphological differences between Khanty dialects and almost none in Mansi.

As these results contradict M. A. Zhivlov's account of glottochronology (performed mainly on the materials of 19th century dictionaries) stating that the time of divergence of Khanty and Mansi dialects was only two centuries apart, we decided to analyze lexicostatistical distances in Swadesh lists of contemporary Khanty and Mansi dialects. Native speakers were surveyed using the questionnaire published in [Kassian et al. 2010]. As a result, we have compiled the Swadesh lists of Kazym, Nizyam and Vakh Khanty as well as Sosva and Yukonda Mansi.

III. Differences in the Swadesh basic vocabulary lists of Mansi dialects

Table 3 below contains Swadesh lists of the words chosen by native speakers, with etymological references. The survey on the hundred-word list of Swadesh was conducted using a questionnaire described in [Kassian et al. 2010]. The last native speaker of the Konda eastern dialect of the Mansi language was interviewed. Two Sosva northern dialect native speakers were also interviewed. All their answers are included in the table. Words with the same meaning in modern dictionaries were not considered, since the methodology of the work is aimed precisely at those words that a native speaker uses in the contexts collected in the article [Kassian et al. 2010]. In the cases when we elicited more than one variant of Khanty or Mansi lexemes, we noted them all. The words that are relevant for calculating lexicostatistical distances — that is, they are not well-known borrowings⁴, but have different roots in dialects — are in bold and their meanings are underlined.

Overall, from the elicitation of native speakers for Sosva and Yukonda Mansi we detected 89 shared words, out of which 4 lexemes are borrowings in one or both dialects. From the remaining 85 words 74 lexemes are etymologically connected and 11 have different roots. Thus, the lexicostatistical distance between Yukonda and Sosva dialects is 87%.

This percentage is quite high, more than between Selkup (85%) dialects according to [Koryakov 2018].

The relatively low number of differences in Mansi Swadesh lists (11 words) correlates with almost full coincidence of morphological markers and a lack of distinct graphic and phonetic innovations in the 18th century manuscripts.

³ The reflexes found in duplicate forms are in brackets.

⁴ According to S. A. Starostin's method, which was used by M. A. Zhivlov too, borrowings are not calculated while measuring the lexicostatistical distances, as they usually emerge not because of evolution of a language (the average speed of change is 5 words in a millenium), but as a consequence of intense language contacts, see further [Burlak, Starostin 2005].

Table 3. Swadesh lists with etymological references for Sosva and Yukonda dialects of Mansi

Concept	Proto-form	eastern Yukonda	northern Sosva
1. all, все			нуссын
ashes, зола	FU * <i>kuḍmV</i> [UEW: 194] ⁵	kivert- <i>χu:ləm</i>	<i>хүлюм</i>
bark, кора	A Komi borrowing [Munkácsi, Kálmán 1986: 571]	<i>χaχ</i>	<i>сүл</i>
belly, живот	FU * <i>kirkV</i> [UEW: 161] A Tungusic borrowing, cf. Proto-Tungus * <i>pugi-</i> / <i>*puki-</i> 'the inside' [TMS 2: 339]	kaχr-	пуки
big, большой	PU * <i>enä</i> [UEW: 74—75]	<i>jañi-</i>	<i>яныг</i>
bird, птица	PU * <i>tulka</i> [UEW: 535] + * <i>wajV</i> (* <i>wojV</i>) [UEW: 553] A Tungusic borrowing, cf. Proto-Tungus * <i>šVču-</i> 'little bird' [TMS 2: 401, 422] ?	εauε	товлыу йй
bite, кусать	PU * <i>pure</i> [UEW: 405] PU * <i>soske</i> [UEW: 448]	<i>porán</i>	<i>пуруңкве</i> <i>тöвмуңкве</i>
black, черный	A Khanty or Komi borrowing [UEW: 758]	<i>εemel'</i>	<i>сэмьл</i>
blood, кровь	PU * <i>kälV</i> [UEW: 134]	<i>kel'p-, k'al'p</i>	<i>кэлп</i>
bone, кость	PU * <i>luwe</i> [UEW: 254]	<i>losum</i>	<i>лув</i>
breast, грудь	FU * <i>mälke</i> (* <i>mäl̥ye</i>) [UEW: 267]	<i>maul'-</i>	<i>мāгьл</i>
burn, жечь	FU * <i>kiče</i> (* <i>küče</i>) [UEW: 153]	<i>ka:si</i>	<i>хүнсуңкве</i> <i>сяритаңкве</i> <i>нāйн</i> <i>тыттуңкве</i>
claw, nail коготь, ноготь	PU * <i>kinče</i> (* <i>künče</i>) (~ ? * <i>kiče</i> (* <i>küče</i>)) [UEW: 157]	<i>k'ant's'</i>	<i>костер</i>
cloud, облако	FP * <i>tule</i> 'ветер' [UEW: 800]	<i>tol-</i>	<i>тул</i>
cold, холодный	FU * <i>pała</i> 'мерзнуть' [UEW: 352]	<i>pul-</i>	<i>полям</i>
come, приходиться	PUg. * <i>juktV</i> [UEW: 851]	<i>juo-</i>	<i>ёхтуңкве</i>
2. die, умирать	PU * <i>šur(e)-ma</i> [UEW: 489]	el'εekrem	сорумн патуңкве āтимьг
dog, собака	PUg. * <i>āmpV</i> (* <i>empV</i>) [UEW: 836]	<i>omr</i>	<i>āмп</i>
drink, пить	FU * <i>juye-</i> (* <i>juke-</i>) [UEW: 103]	<i>ajaχ-</i>	<i>аюңкве</i>
dry, сухой	PUg. * <i>sasV</i> [UEW: 844]	<i>tos-</i>	<i>тосам</i>
ear, ухо	FU * <i>peljä</i> [UEW: 370]	<i>pa:l'</i>	<i>паль</i>
earth, земля	PU * <i>maγe</i> [UEW: 263]	<i>ma:</i>	<i>мā</i>
eat, есть	FU * <i>sewe</i> (* <i>seye</i>) [UEW: 440]	<i>teχ</i>	<i>тэуңкве,</i> <i>тэпъялаңкве</i>
egg, яйцо	PU * <i>muna</i> [UEW: 285]	<i>mon</i>	<i>муни</i>
eye, глаз	PU * <i>šilmä</i> [UEW: 479]	<i>s'a:m-</i>	<i>сам</i>
fat, жир	FU * <i>woje</i> [UEW: 578]	<i>woj</i>	<i>вой</i>
feather, перо	FU * <i>puna</i> [UEW: 402]	<i>pon</i>	<i>пун</i>
3. fire, огонь	PUg. * <i>tüyV-tV</i> ~ * <i>tüwV-tV</i> [UEW: 895] POUg. * <i>nājä</i> [Honti 1982: 421]	towt	най
fish, рыба	PU * <i>kala</i> [UEW: 119]	<i>χul'</i>	<i>хул</i>
fly, летать	FU * <i>šilkV</i> (* <i>šülkV</i>) [UEW: 500]	<i>teul'i</i>	<i>тыламланкве</i>
foot, нога	PUg. * <i>lolkV</i> [UEW: 865]	<i>loaile</i>	<i>лагьл</i>

⁵ Here and further, when quoting forms from [UEW], we, like S. A. Starostin and A. V. Dybo, who created the [UEW] database available at <https://starling.rinet.ru/cgi-bin/response.cgi?root=config&morpho=0&basename=/data/uralic/uralet&first=1>, accept the following transcription changes that make it more convenient for the modern reader to perceive: 3 replaced by V, 8 by O, 8̄ by E.

Concept	Proto-form	eastern Yukonda	northern Sosva
full, полный	FU * <i>tāwδe</i> (* <i>tālkV</i>) [UEW: 518]	<i>tajle</i>	<i>таглэкв</i>
give, давать	PU * <i>miye</i> [UEW: 275]	<i>me-</i>	<i>минункве</i>
good, хороший	POUg. * <i>jomV</i> (* <i>jamV</i>) [UEW: 850]	<i>jomas</i>	<i>ёмас</i>
	A Turkic borrowing * <i>bañ</i> ‘holy’ [EDT: 385]?	<i>wan’χor</i>	
green, зеленый	POUg. * <i>ñErV</i> (* <i>ñOrV</i> , * <i>ñOrkV</i>) [UEW: 331]		<i>нярппум оспа</i>
hair, волосы	PU * <i>apte</i> [UEW: 14]	<i>a:t</i>	<i>ат</i>
hand, рука	FU * <i>kāte</i> [UEW: 140]	<i>kuete-</i>	<i>кат</i>
head, голова	PU * <i>pāñe</i> [UEW: 365]	<i>p’añk</i>	<i>пунк</i>
hear, слышать	FU * <i>kule</i> [UEW: 197]	<i>χ’ontl’-</i>	<i>хунтлункве</i>
heart, сердце	PU * <i>siδā</i> (- <i>mV</i>) (* <i>süδā</i> (- <i>mV</i>)) [UEW: 477]	<i>səm</i>	<i>сьм</i>
horn, рог	PU * <i>añtV</i> (* <i>oñtV</i>) [UEW: 12]	<i>on’t-</i>	<i>анът</i>
I, я	PU * <i>mE</i> [UEW: 294]	<i>om</i>	<i>ам</i>
kill, убивать	FU * <i>wedV</i> [UEW: 566]	<i>el-</i>	<i>алункве</i>
knee, колено	PU * <i>sāncV</i> [UEW: 471]	<i>s’ans</i>	<i>санспуңк</i>
know, знать	PU * <i>wOjV</i> ‘видеть’ [UEW: 588]	<i>ui-</i>	<i>вāнкве</i>
leaf, лист	FU * <i>lEpV</i> (* <i>lepV</i>) [UEW: 259]	<i>lopta</i>	<i>лүпта</i>
lie, лежать	FU * <i>kujV</i> [UEW: 197]	<i>χojen</i>	<i>хуюнкве</i>
liver, печень	PU * <i>maksa</i> [UEW: 264]	<i>ma:t, iʃ mait-</i>	<i>майт</i>
long, длинный	PU * <i>koñćV</i> ~ * <i>koćV</i> [UEW: 180]	<i>χoεə</i>	<i>хоса</i>
louse, вошь	FU * <i>tāje</i> [UEW: 515]	<i>taχət</i>	<i>такум</i>
man, мужчина	PU * <i>koj(e)-mV</i> [UEW: 168]	<i>χət</i>	<i>хум</i>
many, много	FU * <i>čukkV</i> (* <i>čokkV</i>) [UEW: 62]	<i>s’uow</i>	<i>сав</i>
meat, мясо	POUg. * <i>ñāyčəl</i> [Honti 1982: 169]	<i>n’oul’</i>	<i>нёвьль</i>
moon, луна	PU * <i>jitV</i> (* <i>jütV</i>)- [UEW: 99]		<i>этнос</i>
mountain, гора	A Tinguistic borrowing? cf. Proto-Tungus *(<i>x</i>) <i>uKu-</i> ‘hill’ [TMS 2: 256] ⁶	<i>a:χ</i>	
			<i>нёр</i>
mouth, рот	PU * <i>śuwe</i> [UEW: 492]		<i>суп</i>
	A Komi borrowing [Rédei 1970: 167—168]	<i>tuwsən</i>	
name, имя	PU * <i>nime</i> (? * <i>lime</i>) (* <i>nime</i> (? * <i>lime</i>)/* <i>ñime</i>) [UEW: 305]	<i>n’a:mə-</i>	<i>нам</i>
neck, шея	FU * <i>šepä</i> [UEW: 473] (+ * <i>luwe</i>)	<i>sople</i>	<i>сыплув</i>
new, новый	POUg. * <i>iləp, il</i> [Honti 1982: 142] / ПУ * <i>elä-</i> ‘live’ [UEW: 73]		<i>йильпи</i>
4. night, ночь	PU * <i>jitV</i> (* <i>jütV</i>) [UEW: 99]		<i>эт</i>
	FU * <i>eje</i> (* <i>üje</i>) [UEW: 72]	<i>je</i>	
nose, нос	POUg. * <i>ñäl</i> [Honti 1982: 451]	<i>n’oul</i>	<i>нёл</i>
not, не	FU/PU * <i>e~ä~a</i> [UEW: 68-69]	<i>at</i>	<i>ат</i>
one, один	FU * <i>ikte</i> (* <i>ükte</i>) [UEW: 81]	<i>a:χ</i>	<i>аква</i>
person, человек	POUg. * <i>iləm</i> [Honti 1982: 189]	<i>χols</i>	<i>элумхолас</i>
		<i>uetan</i>	
rain, дождь		<i>raχ</i>	<i>ракв</i>
red, красный	FU * <i>wire</i> [UEW: 576]		<i>вбгыр</i>
	PU * <i>kälV</i> [UEW: 134]	<i>kel’p-</i>	<i>кэлп</i>
road, дорога	POUg. * <i>lāñk</i> (- <i>a</i>) [Honti 1982: 189]	<i>l’onχ</i>	<i>лѣңх</i>
root, корень	FU * <i>särV</i> [UEW: 437]	<i>t’or</i>	<i>тār</i>
round, круглый			<i>пувыр</i>
sand, песок			<i>сэй</i>
say, говорить	POUg. * <i>lāw-</i> [Honti 1982: 201]	<i>l’a-</i>	<i>лāвуункве</i>
see, видеть	A Komi borrowing <i>kažal-</i> ‘see’ [UEW: 640]		<i>касалаңкве</i>
	PU * <i>wOjV</i> [UEW: 588]	<i>u:-</i>	

⁶ Reflexes of this word have unmatching correspondences in vocalism in Mansi dialects, cf. Middle Ob *uχ* ‘mountain’ in field notes, other reflexes in [Kannisto et al. 2013: 21]. Such unusual vocalism tends to point to a borrowing. Besides, A. Kannisto does not make any etymological references on this word.

Concept	Proto-form	eastern Yukonda	northern Sosva
seed, семя	PU * <i>šilmä</i> [UEW: 479] ‘eye’		<i>сам</i>
sit, сидеть	PU * <i>amV-</i> [UEW: 8]	<i>wondlen</i>	<i>унлуңкве</i>
5. skin, кожа	FU * <i>šuka</i> [UEW: 488]		<i>сов</i>
	PU * <i>tulka</i> ‘feather, wing’ [UEW: 535]	<i>to:wəl’</i>	
sleep, спать	FU * <i>kujV</i> [UEW: 197]	<i>χojen</i>	<i>хуюңкве</i>
small, маленький	FU * <i>ičä</i> (* <i>üčä</i>) / a Komi borrowing / PUg. <i>wänčV</i> ~ <i>wäčV</i> ‘thin’ [UEW: 899]	<i>wie</i>	
			<i>мāнь</i>
smoke, дым	PUg. * <i>pičV</i> (* <i>pičV</i>) [UEW: 879]		<i>посым</i>
stand, стоять	FU * <i>saŋća</i> [UEW: 431]	<i>tuncən-</i>	
	A Komi borrowing ? [DEWOs: 759] ⁷		<i>лōлюңкве</i>
star, звезда		<i>sov</i>	<i>сов</i>
stone, камень	A Komi borrowing [Rédei 1970: 110]	<i>kerəmc</i>	
			<i>āxətas</i>
sun, солнце	PU * <i>koje</i> [UEW: 167]	<i>χotəl’</i>	<i>хōтал</i>
swim, плавать	PU * <i>uje</i> ~ * <i>oje</i> [UEW: 543]	<i>wujitəχ</i>	<i>уюңкве,</i> <i>уигтаңкве</i>
	FU * <i>pilkV</i> (* <i>pülkV</i>) [UEW: 380]		<i>пувлуңкве</i>
tail, хвост		<i>l’ewuət</i>	<i>лōг</i>
that, тот	PU * <i>tä/te/ti</i> [UEW: 513—514]	<i>tot</i>	<i>таи</i>
thin, тонкий			<i>вовта</i>
this, этот	PU * <i>tä/te/ti</i> [UEW: 513—514]	<i>tet</i>	<i>тыи</i>
thou, ты	POUg. * <i>näy</i> /* <i>něj</i> [Honti 1982: 167]	<i>nan</i>	<i>нан</i>
tongue, язык	FU * <i>hālmä</i> [UEW: 313]	<i>n’elm’en</i>	<i>нēлм</i>
tooth, зуб	FU * <i>pije</i> [UEW: 382]	<i>tus’pank</i>	<i>пуңк</i>
tree, дерево	PU * <i>juwV</i> [UEW: 107]	<i>jiw</i>	<i>йив</i>
two, два	PU * <i>kakta</i> ~ * <i>käktä</i> [UEW: 118]	<i>kit’a</i>	<i>кит/китыг</i>
walk (go), идти	PU * <i>mene</i> [UEW: 272]	<i>te:nen</i>	<i>минункве</i>
	PU * <i>jomV-</i> [UEW: 100]	<i>juomen</i>	
warm, теплый	FU * <i>lońća</i> [UEW: 250]	<i>lanəən</i>	<i>лоньциң</i>
water, вода	PU * <i>wete</i> [UEW: 570]	<i>wit</i>	<i>вит</i>
we, мы	PU * <i>mE</i> [UEW: 29]	<i>man</i>	<i>ман</i>
what, что	PU * <i>mV</i> [UEW: 296]	<i>manə</i>	<i>маныр</i>
white, белый			<i>войкан</i>
		<i>sairən</i>	
	FU * <i>jäje</i> ‘лед’ [UEW: 93]		<i>яңк</i>
who, кто	PU * <i>ku-</i> ~ * <i>ko</i> [UEW: 191]	<i>χənér</i>	<i>хотьют,</i> <i>хōңха</i>
woman, женщина	PU * <i>niñä</i> (* <i>niñä</i> / * <i>hiñä</i>) [UEW: 305]	<i>nia:</i>	<i>нэ</i>
	FU * <i>ewkkV</i> [UEW: 76]		<i>эква</i>
yellow, желтый			<i>хāль лўнта</i> <i>холом оспа</i>

IV. Differences in the Swadesh basic vocabulary lists of Khanty dialects

See below the information about differences between Khanty dialects. The survey was also conducted using a questionnaire developed in [Kassian et al. 2010]. The last native speaker of Nizyam dialect, who lives in Khanty-Mansiysk and speaks the dialect fluently, was interviewed. Also, two native speakers of northern Kazym northern dialect of Khanty, who use their mother tongue in everyday life, were interviewed. All their answers are included into the table. The words that are relevant for calculating lexicostatistical distances between

⁷ A borrowing hypothesis from Komi *люллявны* ‘stagger’ is suggested, because this word is attested only in North Mansi, according to [Munkácsi, Kálmán 1986: 286], and L. Honti does not reconstruct it for Proto-Ob-Ugric,

Kazym (northern) and Vakh (eastern) dialects — that is, they are not well-known borrowings, but have different roots — are in bold and their meanings are underlined. Khanty lexemes that distinguish between Nizyam (transitional) and Kazym (northern) dialects are also underlined. These words show that differences between those western dialects is much less than between the western and eastern ones.

Table 4. Swadesh lists with etymological references for Vakh, Nizyam and Kazym dialects of Khanty

Word	Protoform	Vakh	Nizyam	Kazym
all, все	PU *ku- ~ *ko [UEW: 191]		<i>χudi jewa</i>	<i>χölljewa / χöll</i>
ashes, зола	FU *kuđmV [UEW: 194]		<i>χəjəm / χujəm</i>	<i>χəjəm</i>
bark, кора	FU *kere [UEW: 148—149]	<i>kär</i>	<i>kar</i>	<i>kar</i>
<u>1. belly, живот</u>	FU *kunV [UEW: 408]		<i>χun</i>	<i>χön</i>
		<i>póyji</i>		
<u>2. big, большой</u>	FU/ PU *enā [UEW: 74—75; DEWOs: 115]	<i>állä</i>		
	PU *anV (*onV) [UEW: 9]		<i>wəŋ</i>	<i>wəŋ</i>
bird, птица	FU *wajV (*wojV) [UEW: 553] + *tulka [UEW: 535]	<i>wájäy / wájəy</i>		<i>töχləŋ wəj</i>
bite, кусать	PU *pure [UEW: 405]	<i>porté</i>		<i>pörti</i>
	PU *soske [UEW: 448—449]			<i>təχəmti</i>
black, черный	PUg. *pEkKV-ttV [UEW: 882]	<i>pəyätə</i>	<i>piti</i>	<i>pití</i>
<u>3. blood, кровь</u>	FU *wire [UEW: 576]	<i>wär / wär</i>		
	PU *kälV [UEW: 134]		<i>kədi</i>	<i>kälí</i>
bone, кость	PU *luwe [UEW: 254]	<i>löy</i>	<i>θuw</i>	<i>lüw</i>
breast, грудь	FU *mälke (*mälye) [UEW: 267]	<i>möyäl</i>	<i>mewəd</i>	<i>mewəl</i>
burn tr., жечь	FU *äsV- [UEW: 27]		<i>ə:ijiti</i>	
	PUg. *tüyV-tV ~ *tüwV-tV [UEW: 895]			<i>tüta lapətti</i>
claw, ноготь	PU *kinče (*künče) [UEW: 157]	<i>könteə, künte</i>	<i>kunf</i>	<i>küşkar</i>
cloud, облако	FU *pilwe (*pilhe) [UEW: 381]	<i>pəliěŋ</i>	<i>pədəŋ</i>	<i>päləŋ</i>
<u>4. cold, холодный</u>	PU *acV ‘cold’ [Thesaurus 2010: 55]	<i>ätəy</i>		
	PUg. *pOtV ‘ice crust; to freeze’ [UEW: 882]			<i>pətam</i>
come, прийти	PUg. *juktV [UEW: 851]		<i>juχtjiti</i>	<i>jöχətti</i>
	PU *kola [UEW: 173]	<i>il qöl-iyən</i>	<i>χədti</i>	
	[DEWOs: 1458—1459] connected with Khanty <i>tewər</i> ‘litter, waste’	<i>téwərsäy.jə-yäl</i>		
<u>5. die, умирать</u>	FU *śur(e)-ma ‘death’ [UEW: 489]			<i>sörma jiti</i>
	[DEWOs: 259—260] connected with <i>čäyi-</i> / <i>šawi-</i> ‘save, hide’		<i>faweməti</i>	
dog, собака	PUg. *ämpV (*empV) [UEW: 836]	<i>ämp</i>	<i>amp</i>	<i>amp</i>
	FU *sewe (*seye) [UEW: 440]	<i>li-tä</i>		
<u>6. drink, пить</u>	FU *juye- (*juke-) [UEW: 103]		<i>jəŋeti / jaŋeti</i>	<i>jaŋsi</i>
dry, сухой	FU *sarV ‘сухой’		<i>səram</i>	<i>səram</i>
ear, ухо	FU *peljä [UEW: 370]	<i>pəl</i>	<i>pəd</i>	<i>päl</i>
earth, земля	PU *maye [UEW: 263]	<i>məy</i>	<i>məw</i>	<i>müw</i>
eat, есть	FU *sewe (*seye) [UEW: 440]	<i>litä</i>	<i>θewa / teti</i>	<i>letí</i>
egg, яйцо		<i>kärəŋ moq</i>		<i>kar pöšəχ</i>
eye, глаз	PU *šilmä [UEW: 479]	<i>sem</i>	<i>sem</i>	<i>sem</i>
fat, жир	FU *woje [UEW: 578]	<i>woj</i>	<i>wij</i>	<i>wöj</i>
feather, перо	FU *puna [UEW: 402]	<i>pun</i>		<i>pün</i>
fire, огонь	PUg. *tüyV-tV ~ *tüwV-tV [UEW: 895]	<i>tüyö</i>	<i>tüt</i>	<i>tüt</i>
	POUg. *näj [Honti 1982: 167]	<i>näj</i>	<i>naj</i>	
fish, рыба	PU *kala [UEW: 119]	<i>qu(l)</i>	<i>χud</i>	<i>χül</i>
	FU *šilkV (*šülkV) [UEW: 500]	<i>läyäl-tä</i>		
<u>7. fly, летать</u>	POUg. *jōŋk- ~ *jäŋkā [Honti 1982: 143]		<i>jəŋχtjiti</i>	<i>jäŋχti</i>

Word	Protoform	Vakh	Nizyam	Kazym
foot, нога	[DEWos: 664—667] ⁸	<i>kūr</i>	<i>kur</i>	<i>kūr</i>
full, полный	FU * <i>tāwde</i> (* <i>tālkV</i>) [UEW: 518]	<i>téleltä / télel</i>		<i>teł</i>
give, давать	PU * <i>miye</i> [UEW: 275]	<i>mi-</i>	<i>mäti</i>	<i>māti</i>
good, хороший	PUg. * <i>jomV</i> (* <i>jamV</i>) [UEW: 850]	<i>jäm / jāmāṅ</i>	<i>jəm</i>	<i>jām</i>
green, зеленый	[DEWos: 1637]			<i>wōstī</i>
hair, волосы	PU * <i>apte</i> [UEW: 14]	<i>āwtä / āwat</i>	<i>əpət</i>	<i>əpət</i>
8. hand, рука	FU * <i>käte</i> [UEW: 140]	<i>köt</i>		
	POUg. * <i>jīčz</i> [Honti 1982: 140]		<i>jəf</i>	<i>jəš</i>
head, голова	PU * <i>ukV</i> (* <i>okV</i>) [UEW: 542]	<i>oγ / óγ</i>	<i>uγ</i>	<i>ōγ</i>
hear, слушать	FU * <i>kule</i> [UEW: 197]	<i>kólme-</i>	<i>χədti</i>	<i>χəłti</i>
heart, сердце	PU * <i>šidä</i> (- <i>mV</i>) (* <i>šüdä</i> (- <i>mV</i>)) [UEW: 477]	<i>säm</i>	<i>səm</i>	<i>sām</i>
horn, рог	PU * <i>aṅtV</i> (* <i>oṅtV</i>) [UEW: 12]	<i>āṅət / āṅət</i>		<i>əṅət</i>
I, я	PU * <i>mE</i> [UEW: 294]	<i>mä</i>	<i>ma</i>	<i>ma</i>
kill, убивать	FU * <i>weδV</i> [UEW: 566]	<i>wel-</i>	<i>wedti / wejiti</i>	<i>welti</i>
кnee, колено	FU * <i>šānčV</i> [UEW: 471]	<i>teänte</i>	<i>ʃaf</i>	<i>šanš</i>
9. know, знать	PU * <i>wOjV</i> [UEW: 588]		<i>wədti / wəjti</i>	<i>wəti</i>
	[DEWos: 1448—1449] connected with Khanty <i>tōṅ</i> ‘straight, true’	<i>tōṅəmt-tü</i>		
leaf, лист	FU * <i>IEpV</i> (* <i>lepV</i>) [UEW: 259]	<i>ʔəwət / ʔiwət</i>	<i>łipət</i>	<i>łipət</i>
lie, лежать	FU * <i>oda-mV</i> [UEW: 335]	<i>älä-ta</i>		<i>əłti</i>
liver, печень	PU * <i>maksa</i> [UEW: 264]	<i>müyöl</i>	<i>muxəd</i>	<i>mōχəl</i>
long, длинный	FU * <i>kawka</i> [UEW: 132]	<i>qoγ</i>		<i>χiw</i>
louse, вошь	FU * <i>täje</i> [UEW: 515]	<i>tōytäm</i>		<i>tewtəm</i>
man, мужчина	PU * <i>ekä</i> [UEW: 72]	<i>iki / iki / ik</i>		<i>iki</i>
	PU * <i>koje</i> [UEW: 166]	<i>ku</i>	<i>χə</i>	<i>χə</i>
many, много	PU * <i>erV</i> [UEW: 75]	<i>āryi / ārki</i>	<i>ar / arʃək</i>	<i>ar</i>
meat, мясо	POUg. * <i>nāyäl</i> [Honti 1982: 169]	<i>n'ōyi</i>	<i>nyxi</i>	<i>nōxi</i>
moon, месяц, луна	PU * <i>ekä</i> [UEW: 72]	<i>iki / iki / ik</i>		
	A Komi borrowing [DEWos: 1430]		<i>tüdəe</i>	<i>tüläs</i>
10. mountain, гора	[DEWos: 330]	<i>jöγ</i>		
	[DEWos: 1278]			<i>rep</i>
11. mouth, рот	FU * <i>šule</i> [UEW: 903]	<i>lul</i>		
	PU * <i>aṅe</i> [UEW: 11]		<i>uṅid</i>	<i>uṅəl</i>
name, имя	PU * <i>nime</i> (? * <i>lime</i>) [UEW: 305]	<i>nem</i>	<i>nem</i>	<i>nem</i>
neck, шея	FU * <i>šepä</i> [UEW: 473]		<i>sapəd</i>	<i>sapəl</i>
new, новый	PU * <i>elä</i> - ‘жить’ [UEW: 73]	<i>jälw / jäläw</i>	<i>jidup</i>	<i>jiləp</i>
12. night, ночь	[DEWos: 345]	<i>jöγön</i>		
	PU * <i>jitV</i> (* <i>jütV</i>) [UEW: 99]		<i>at</i>	<i>at</i>
nose, нос	POUg. * <i>näl</i> [Honti 1982: 451]	<i>nöl</i>		<i>nöl</i>
not, не	[Honti 1982: 198]	<i>an</i>	<i>an / ənt</i>	<i>än, änt</i>
one, один	PU * <i>e</i> [UEW: 67]	<i>əj</i>	<i>i / it</i>	<i>it / i</i>
person, человек	PU * <i>koje</i> [UEW: 166] ‘man’	<i>qu</i>	<i>χanti χə / χujat</i>	<i>χänneχə</i>
	A Siberian Tatar or Samoyed borrowing	<i>qäsí</i>		
	FU * <i>kunta</i> [UEW: 206]		<i>χanti</i>	
13. rain, дождь	[DEWos: 852]	<i>ʔəwət</i>		
	POUg. * <i>jirt</i> [Honti 1982: 144]		<i>jert</i>	<i>jert</i>
red, красный	FU * <i>wire</i> [UEW: 576]	<i>wärtä</i>	<i>wirti</i>	<i>würti</i>
14. road, дорога	[Honti 1982: 202]	<i>lək</i>		
	FU * <i>junča</i> ~ * <i>juča</i> [UEW: 104]		<i>jəf</i>	<i>jəš</i>

⁸ Here and further, when a simple reference to the source is given, it shows that the word has no external etymology, but the reader can find forms from other dialects in the specified dictionary.

Word	Protoform	Vakh	Nizyam	Kazym
15. root, корень	FU * <i>särV</i> [UEW: 437]		<i>θer / teramət:i</i>	<i>λer</i>
	POUg. * <i>θōr</i> [Honti 1982: 139]	<i>lor</i>		
	[DEWOs: 330]	<i>joγ</i>		
16. round, круглый	[DEWOs: 394]	<i>kómľäk</i>	<i>jeŋta</i>	<i>jeŋta</i>
	[DEWOs: 1351]	<i>səŋq / səŋqí / səŋqí</i>		
17. sand, песок	FU * <i>kOčV</i> [UEW: 226]		<i>χif</i>	<i>χiš</i>
	FU * <i>kuđmV</i> [UEW: 194]		<i>χujəm</i> ‘песок, зола’	
	FU * <i>wäntV</i> [UEW: 564]			<i>wantti</i>
18. see, видеть	FU * <i>wOjV</i> [UEW: 588]	<i>ú-ta</i>		
			<i>fiwadəti</i>	<i>šiwələti</i>
19. say, говорить	[DEWOs: 415]		<i>jastijti / jastəti</i>	
	[DEWOs: 1436]	<i>tölöy-ta</i>		
	[DEWOs: 1248]		<i>putərti</i>	
	[DEWOs: 848]		<i>lupti</i>	<i>löpti</i>
seed, семя	PU * <i>šilmä</i> [UEW: 479]			<i>sem</i>
sit, сидеть	PU * <i>amV-</i> [UEW: 8]	<i>áməs-ta</i>	<i>əməsti</i>	<i>əməsti</i>
skin, кожа	FU * <i>šuka</i> [UEW: 488]	<i>söγ</i>	<i>suχ</i>	<i>el söχ</i>
20. sleep, спать	POUg. * <i>wōj-</i> ~ * <i>wājā</i> [Honti 1982: 192]	<i>wójä-ta</i>		
	FU * <i>oda-mV</i> [UEW: 335]		<i>udti / wudti</i>	<i>əlti</i>
21. small, маленький	PUg. * <i>äjV</i> [UEW: 835]		<i>aj</i>	<i>aj</i>
	[DEWOs: 43]	<i>üküm / úküm / úkümáli / úküm</i>		
22. smoke, дым	POUg. * <i>pörkzj</i> [Honti 1982: 179]	<i>pörqí</i>		
	PUg. * <i>pičV</i> (* <i>pičV</i>) [UEW: 879]		<i>pəsəŋ</i>	<i>pəsəŋ</i>
stand, стоять	FU * <i>saŋća</i> [UEW: 431]	<i>lál'-ta</i>		<i>λəl'ti</i>
	[DEWOs: 721]		<i>tujci / təcici</i>	
star, звезда	PU * <i>kuńćV</i> ~ * <i>kućV</i> [UEW: 210—211]	<i>qos</i>	<i>χəs</i>	<i>χəs</i>
stone, камень	FU * <i>kiwe</i> [UEW: 163]	<i>köγ</i>	<i>kew</i>	<i>kew</i>
	[DEWOs: 1351]	<i>süŋk</i>		
23. sun, солнце	PU * <i>koje</i> [UEW: 167]		<i>χət</i>	<i>χətəl</i>
	POUg. * <i>nāj</i> [Honti 1982: 167]		<i>naj</i>	
			<i>χəwiji</i>	
24. swim, плавать	FU * <i>pilkV</i> (* <i>pülkV</i>) [UEW: 380]		<i>pewədti</i>	<i>pewəłti</i>
	PU * <i>uje</i> ~ * <i>oje</i> [UEW: 543]		<i>wəcti</i>	<i>wəł'si</i>
tail, хвост	[DEWOs: 727]	<i>lamt-</i>		
	A Komi borrowing [DEWOs: 246]	<i>läγ</i>		
that, тот	ПУ * <i>tä / te / ti</i> [UEW: 513—514]		<i>fəjt / fəjət</i>	<i>šəjət</i>
	ПУ * <i>tä / te / ti</i> [UEW: 513—514]	<i>tom / tūt</i>	<i>tun</i>	<i>töm</i>
this, этот		<i>tü / teu</i>		<i>täm</i>
			<i>ei / eit</i>	
thou, ты	POUg. * <i>näγ / *něj</i> [Honti 1982: 167]	<i>nuy</i>	<i>nəŋ</i>	<i>näj</i>
tongue, язык	FU * <i>ńälmä</i> [UEW: 313]	<i>ńäləm</i>	<i>ńadəm</i>	<i>ńaləm</i>
tooth, зуб	FU * <i>piŋe</i> [UEW: 382]	<i>püŋkä</i>	<i>peŋk</i>	<i>peŋk</i>
tree, дерево	PU * <i>juwV</i> [UEW: 107] ‘сосна’?	<i>juγ</i>	<i>juχ</i>	<i>jüχ</i>
two, два	PU * <i>kakta</i> ~ * <i>käktä</i> [UEW: 118]	<i>kät / kä</i>	<i>kət:em</i>	<i>kät</i>
	PU * <i>mene</i> [UEW: 272]			<i>mänti</i>
25. walk (go), идти	PU * <i>kulke</i> [UEW: 198]	<i>kówəłta</i>		
	PU * <i>čančV</i> ~ * <i>čačV</i> [UEW: 53]		<i>fəfijti / fəfti</i>	
26. warm (hot),	PUg. * <i>mälV</i> [UEW: 868]	<i>méləγ</i>		

Word	Protoform	Vakh	Nizyam	Kazym
теплый (горячий)	[DEWOs: 510]	<i>qónǒq</i>		
	FU * <i>lońca</i> [UEW: 250]		<i>лунсаη</i>	<i>лǒńsi</i>
	[DEWOs: 427]		<i>χǒfǎm</i>	<i>χǒšǎm</i>
water, вода	FU * <i>jǎne</i> [UEW: 93]	<i>jǎŋki / jink / jǎnk / jǎŋk</i>	<i>jɪŋk</i>	<i>jɪŋk</i>
we, мы	PU * <i>mE</i> [UEW: 29]	<i>maŋ</i>	<i>mɪŋ</i>	<i>min</i> (мы двое)
what, что	PU * <i>mV</i> [UEW: 296]	<i>mǔyǐ / mǔyǔ / mǔy / mǔyǔli</i>	<i>muj</i>	<i>mōj</i>
white, белый	[DEWOs: 990]	<i>nǎyi</i>		<i>nōwi</i>
	FU * <i>jǎne</i> [UEW: 93]			<i>jeŋk</i>
who, кто	PU * <i>ku-</i> ~ * <i>ko</i> [UEW: 191]	<i>qóji</i>	<i>χǒj</i>	<i>χǒj</i>
	PU * <i>niŋǎ</i> (* <i>niŋǎ</i> / * <i>niŋǎ</i>) [UEW: 305]	<i>ni</i>	<i>ne</i>	<i>ne</i>
woman, женщина	PU * <i>imV</i> [UEW: 83]			<i>imi</i>
			<i>ɔwʃek</i>	
yellow, желтый				<i>wǔślǎm</i>

Overall, from the elicitation of native speakers for Kazym and Vakh dialects of Khanty 94 shared words were detected, out of which 6 lexemes are borrowings in one or both dialects. From the remaining 88 words 62 lexemes are etymologically connected and 26 have different roots. Thus, the lexicostatistical distance between Kazym and Vakh dialects is 70%.

This percentage is unusually low not only for dialects of one language, but also for closely related languages. In particular, Swadesh lists of contemporary Slavic languages compiled by M. N. Saenko [Saenko 2015, 2017] demonstrate more than 70% of coincidence in basic vocabulary. M. N. Saenko, after S. A. Starostin, dates divergence of Proto-Slavic back to 130 AD. As [Dyachok 2001] shows, the percentage of coincidence in basic vocabulary between any Turkic languages except Chuvash is also more than 70%. The traditional dating of divergence of Common Turkic is the early first centuries AD.

Given that Kazym and Vakh Khanty share the same percentage of coincidence in basic vocabulary, on the grounds of glottochronology it could be suggested that they diverged in the beginning of the first millennium AD.

These data are in agreement with substantial differences in morphology between contemporary Khanty dialects and a number of innovations that are present in Khanty dictionaries as early as the 18th century.

Conclusion

The analysis of phonetic differences in Ob-Ugric languages correlates with the results of glottochronology calculations conducted on the materials of a survey of native speakers of contemporary dialects.

They show that Mansi dialects neither have substantial differences in basic vocabulary, nor did they differ much in morphology and in graphics-phonetics in the 18th century. Phonetic innovations that hindered mutual understanding between speakers of the eastern and northern dialects in the 20th century occurred not earlier than 250 years ago.

On the contrary, Kazym and Vakh Khanty have more differences than any Slavic and even Turkic (except Chuvash) languages between each other. Undoubtedly, they should be accounted for different languages, not dialects. The time of their divergence can be dated as early as the beginning of the first millennium AD.

Abbreviations

FP — Finno-Permic

FU — Proto-Finno-Ugric

POUg. — Proto-Ob-Ugric

PU — Proto-Uralic

PUg. — Proto-Ugric

References

- Burlak, Starostin 2005 — *Burlak S. A., Starostin S. A. Comparative-historical linguistics*. Moscow, 2005.
 DEWOs — *Steinitz W. Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin, 1966—1989.
 Dyachok 2001 — *Dyachok M. T. Glottochronology of the Turkic languages (a preliminary analysis)*. In: *Proceeding of the Second Scientific conference*. Novosibirsk, 2001. P.14—16.

- EDT — *Clauson G.* An Etymological Dictionary of Pre-Thirteenth-Century Turkish. Szeged, 1981.
- Fedotova in print — *Fedotova I. V.* Dialectal classification of Khanty in the light of basic vocabulary. In: Ural-Altai studies. 2022. In print.
- Gulya 1958 — *Gulya J.* Egy 1736-ból származó manysi nyelvemlék (Előzetes közlés). In: Nyelvtudományi Közlemények. 1958, 60. P. 41—45.
- Gulya 1960 — *Gulya J.* Old Mansi dialects. In: Congressus Internationalis Fenno-Ugristarum Budapestini habitus 20—24. IX. 1960. Budapest, 1963. P. 172—175.
- Honti 1982 — *Honti L.* Geschichte des obugrischen Vokalismus der ersten Silbe. Budapest, 1982.
- Honti 1988 — *Honti L.* Die wogulische Sprache. In: The Uralic Languages. Description, History and Foreign Influences. Handbuch der Orientalistik / Ed. Sinor Denis. Leiden; New York; København; Köln, 1988. P. 147—171.
- Kannisto 2013 — *Kannisto A., Eiras V., Moisio A.* Wogulisches Wörterbuch. Helsinki: Société Finno-Ougrienne: Kotimaisten Kielten Keskus, 2013.
- Kassian et al. 2010 — *Kassian A., Starostin G., Dybo A., Chernov V.* The Swadesh wordlist. An attempt at semantic specification. In: Journal of Language Relationship. 2010, 4. P. 46—89.
- Koryakov 2018 — *Koryakov Yu. B.* Language vs dialect question and the Samoyedic languages. In: Ural-Altai studies. 2018, 4 (31). P. 156—217.
- Koryakov 2022 — *Koryakov Yu. B.* Changes in the ethno-linguistic territory of the Western and Southern Mansi in the 17th—20th centuries. Part I. Cis-Urals and the Tura basin. In: Ural-Altai studies. 2022, 2 (45). P. 58—78.
- Koshelyuk 2021 — *Koshelyuk N. A.* General innovation processes in Mansi dialects: nominal morphology. In: Bulletin of Ugric Studies. 2021, 11 (2). P. 272—292.
- Munkácsi, Kálmán 1986 — Wogulisches Wörterbuch. Gesammelt von B. Munkácsi. Geordnet, bearbeitet und herausgegeben von B. Kálmán. Budapest, 1986.
- Normanskaya 2022 — *Normanskaya Yu. V.* How the classification of Mansi dialects was changed (on the material of the first Cyrillic books and dictionaries of the 18th and 19th centuries). In: Siberian Journal of Philology. 2022, 1. P. 127—143.
- Normanskaya, Kosheliuk 2020 — *Normanskaya Yu. V., Kosheliuk N. A.* The unpublished Mansi dictionary of P. S. Pallas — an earlier unknown Mansi dialect? In: Ural-Altai Studies. 2020, 1 (36). P. 92—100.
- Rédei 1970 — *Rédei K.* Die syrjänischen Lehnwörter im Wogulischen. Budapest, 1970.
- Saenko 2015 — *Saenko M. N.* Shared innovations in basic vocabulary as an argument in the discussion about Balto-Slavic unity. PhD thesis. Moscow, 2015.
- Saenko 2017 — *Saenko M. N.* Annotated Swadesh wordlists for the Slavic group (Indo-European family). 2017. In: The Global Lexicostatistical Database / Ed. Starostin G. Moscow; Santa Fe, 2011—2020. URL: <https://starling.rinet.ru/cgi-bin/response.cgi?root=new100&morpho=0&basename=new100\ier\slv&limit=-1>
- Solovar 2006 — *Solovar V. N.* Khanty-Russian dictionary. Saint Petersburg, 2006.
- Solovar, Nakhracheva, Shiyanova 2016 — *Solovar V. N., Nakhracheva G. L., Shiyanova A. A.* Dialects of the Khanty language. Khanty-Mansiysk; Izhevsk, 2016.
- Tereshkin 1961 — *Tereshkin N. I.* A survey of Khanty dialects. Part One. Vakh dialect. Moscow, Leningrad, 1961.
- Thesaurus 2010 — *Normanskaya Yu. V., Dybo A. V.* Thesaurus. Lexicon of natural surroundings in the Uralic languages. Moscow, 2010.
- TMS — Comparative dictionary of Manchu-Tungus languages / Ed. Tsintsius V. I. Leningrad, 1975—1977.
- UEW — Uralisches etymologisches Wörterbuch / Hrsg. von K. Rédei. Budapest, 1986—1991.
- Zhivlov 2011 — *Zhivlov M. A.* Annotated Swadesh wordlists for the Ob-Ugrian group (Uralic family). 2011. In: The Global Lexicostatistical Database / Ed. Starostin G. Moscow; Santa Fe, 2011—2020. URL: <https://starling.rinet.ru/cgi-bin/response.cgi?root=new100&morpho=0&basename=new100\ura\oug&limit=-1>

Компаративно-аттенуативная полисемия в финно-угорских языках

Черемисинова Мария Олеговна, Национальный исследовательский университет «Высшая школа экономики» (Москва); masha.cherem@gmail.com

Работа посвящена компаративно-аттенуативной полисемии в финно-угорских языках. Компаративные показатели некоторых финно-угорских языков иногда обладают еще и аттенуативным значением ('немного'). В статье рассматриваются три показателя, демонстрирующих такую полисемию: *-ges* в бесермянском, *-šək* в казымском хантыйском и *-rak* в горномарийском. Эти показатели употребляются в компаративной конструкции, обладают аттенуативной функцией и сочетаются с показателями отрицания. Показатели кросскатегориальны, и в этой работе мы также описываем сочетаемость показателей в различных функциях. Во всех рассмотренных нами языках показатели в компаративных конструкциях сочетаются с аспектом сравнения, выраженным прилагательным и наречием; кроме того, в бесермянском возможно также сочетание показателя *-ges* с глаголами и существительными. В аттенуативной функции показатели могут сочетаться с прилагательными, наречиями и реляционными именами / послелогами. В бесермянском и горномарийском языках показатели могут сочетаться еще и с глаголами. Также все рассмотренные показатели могут сочетаться с отрицательным глаголом / отрицательной частицей (в зависимости от языка).

Ключевые слова: аттенуатив, компаратив, полисемия, сравнительная конструкция, финно-угорские языки

COMPARATIVE-ATTENUATIVE POLYSEMY IN FINNO-UGRIC LANGUAGES

Maria O. Cheremisinova, HSE University (Moscow); masha.cherem@gmail.com

The topic of this paper is comparative and attenuative polysemy in Finno-Ugric languages. In some Finno-Ugric languages comparative markers have attenuative meaning ('a little bit, slightly') as well. This paper focuses on three markers exhibiting such a polysemy: *-ges* in Beserman (Permic), *-šək* in Kazym Khanty (Ugric) and *-rak* in Hill Mari (Mari). These markers can be used in comparative constructions, have attenuative function and can be attached to the markers of negation. The markers can be used cross-categorically, therefore I also describe compatibility of the markers in different functions. In comparative constructions, all of the markers can be attached to adjectives and adverbs, while in Beserman it is also possible for the marker *-ges* to be combined with verbs, nouns and pronouns. However, in Kazym Khanty the marker *-šək* has attenuative meaning even in comparative constructions. In attenuative function, the markers in all three languages can attach to adjectives, adverbs, postpositions (or relational nouns), while in Beserman and Hill Mari verbs can be modified by the markers as well. Finally, comparative-attenuative markers in all of the languages can be attached to the markers of negation. In this function, the most common context allowing the use of the marker is the context meaning 'not yet': it appears in all of the described languages. In Kazym Khanty the marker *-šək* can also express the meaning of emphatic negation ('no way') when making a reference to future or present.

Keywords: attenuative, comparative construction, comparative marker, Finno-Ugric languages, polysemy

1. Введение¹

В некоторых финно-угорских языках показатели компаратива (или *компаративные показатели*) — сравнительной степени прилагательных и наречий — обладают еще одним значением, а именно аттенуативным. Аттенуативное значение — это значение невысокой степени: ‘немного, чуть-чуть’. Настоящая работа посвящена трем показателям, демонстрирующим такую полисемию: *-ges* в бесермянском, *-rak* в горномарийском и *-šək* в казымском хантыйском.

В литературе и, в частности, в грамматических описаниях компаративно-аттенуативной полисемии в финно-угорских языках уделялось мало внимания. Среди финно-угорских языков компаративно-аттенуативная полисемия зафиксирована в шести: удмуртском (пермские) [Перевошиков и др. 1962: 137—138; Efremov, Selmečzy 2018], коми-зырянском (пермские) [Coates 1982; Todesk 2015; Bartens 2000], бесермянском² (пермские) [Черемисинова 2019], горномарийском (марийские) [Саваткова 2002: 129, 134], луговом марийском (марийские) [Казанцева 2005: 157], хантыйском (угорские) [Каксин 2007: 43, 101]. Как правило, из указанных источников мы в лучшем случае узнаем лишь то, что такая полисемия есть (в сочетании с некоторыми прилагательными). Исключение составляют статьи Дж. Коутс [Coates 1982] и Д. Ефремова, Ш. Шелмеци [Efremov, Selmečzy 2018] о коми-зырянском показателе *-джык* и удмуртском показателе *ges* (*-гем*) соответственно, про которые известно, в частности, что они транскатегориальны: помимо прилагательных и наречий, эти показатели также могут присоединяться к существительным, местоимениям, послелогам и глаголам.

За пределами финно-угорских языков компаративно-аттенуативная полисемия встречается в тундровом ненецком [Буркова 2010: 261; Nikolaeva 2014: 134, 174; Berezovskaya 2019: 24], а также в некоторых тюркских языках (например, в чувашском [Павлов 1957: 94, 98] и мишарском татарском [Былинина 2017: 448]).

В работе предлагается описание для показателей компаратива в трех финно-угорских языках: бесермянском, горномарийском и казымском хантыйском. Выбор языков и диалектов обусловлен доступными нам данными. Исследование основано преимущественно на результатах полевой работы (мы также по возможности использовали корпусные данные и тексты).

Материалом для исследования послужили данные бесермянского языка, собранные в экспедициях в д. Шамардан (Юкаменский район, Удмуртия) в 2019 и 2020 году, данные корпусов (<http://beserman.ru>) и (<http://multimedia-corpus.beserman.ru/search>), а также примеры из «Тезауруса бесермянского наречия: имена и служебные части речи» [Усачева и др. 2017]. Данные казымского диалекта хантыйского языка были собраны в ходе полевой работы в с. Казым Ханты-Мансийского автономного округа в 2019 г.; часть примеров была получена онлайн в ходе интернет-созвонов с носителями языка в 2020 г.; помимо этого, мы также пользовались корпусом текстов на хантыйском языке³. Данные горномарийского языка были получены в экспедиции в с. Микряково⁴ Горномарийского района Республики Марий Эл в 2018 г. Все элицитированные примеры были собраны и проверены как минимум у трех носителей языка. Если не указано иное, пример элицитирован.

Работа устроена следующим образом. В разделе 2 мы опишем функционирование показателя *-ges* в бесермянском, в разделе 3 — показателя *-rak* в горномарийском, в разделе 4 — показателя *-šək* в казымском хантыйском. У каждого из этих показателей есть употребления как в компаративной конструкции, так и вне ее; с прилагательными, наречиями, послелогам, а также с глаголами; ограничения на употребления с прилагательными зависят от типа шкалы, ограничения на употребления с глаголами зависят от акциональности; в отрицательных предложениях у показателей появляются новые функции. Раздел 5 посвящен сравнению этих показателей (с привлечением там, где это возможно, коми-зырянского суффикса *-джык*). Раздел 6 представляет собой заключение.

¹ Автор выражает благодарность А. А. Козлову, М. Н. Усачевой и Т. А. Архангельскому за ценные комментарии на разных этапах работы. Мы также благодарны всем работавшим с нами носителям бесермянского, горномарийского и казымского хантыйского языков.

² Ранее считавшийся диалектом (или наречием) удмуртского языка, в 2022 году бесермянский получил статус самостоятельного языка России (см. http://jazykirf.iling-ran.ru/list_2022.shtml).

³ Корпус текстов разработан Е. В. Кашкиным и Н. А. Муравьевым в рамках программы фундаментальных исследований президиума РАН «Корпусная лингвистика» (2012—2014 г.).

⁴ Рассматриваемый нами говор принадлежит к правобережной группе говоров, емангашскому подговору (по классификации в [Саваткова 2002]). Как верно замечает рецензент, данные, представленные в этой работе, могут отличаться от данных об использовании показателя *-rak* в литературном языке.

2. Компаративно-аттенуативная полисемия в бесермянском языке

В удмуртском языке показатель компаратива *-ges*, помимо употребления в сравнительной конструкции (1), обладает также аттенуативным значением [Перевощиков и др. 1962: 138]⁵.

(1) УДМУРТСКИЙ

*километр иськем-лэсь вакчи-гес*⁶

километр верста-AVL короткий-COMP

‘Километр короче версты’ [Перевощиков и др. 1962: 138].

Помимо показателя *-ges*, в литературном удмуртском имеется также показатель *-гем*. Чем именно обусловлена дистрибуция показателей (если она есть), из описаний непонятно.

(2) ЮШ ЛЫМ ЯЗЬ ЛЫМ-ЛЭСЬ ЧЕСКЫТ-ГЕМ

окунь уха язь уха-AVL вкусный-COMP

‘Уха из окуня вкуснее уха из язя’ [Перевощиков и др. 1962: 137].

Как уже было упомянуто, в работах [Efremov, Selmečzy 2018, 2019] обращается внимание на широкую сочетаемость показателя *-ges* (*-gem*) в удмуртском языке (помимо сравнительной степени, авторы рассматривают превосходную степень и сравнивают *-ges* (*-gem*) с показателями в венгерском и чувашском языках соответственно): с существительными, прилагательными, наречиями, глаголами и послелогоми. Также авторы отмечают аттенуативное употребление показателя с прилагательными (*moderative degree markers* в терминологии авторов). Однако в работах не описаны ограничения на сочетаемость (в компаративной конструкции), также остается неизвестным, возможно ли употребление показателя в аттенуативной функции в сочетании с другими частями речи (помимо прилагательных), возможно ли употребление при отрицании (в [Efremov, Selmečzy 2019] упоминается употребление при отрицании, однако ни семантика, ни сочетаемость показателя в этом контексте не рассмотрены). В [Ефремов 2002] также была отмечена опциональность показателя *-ges* (*-gem*) в сравнительной конструкции.

В бесермянском языке представлен всего один показатель сравнения *-ges/-kes*⁷; ему и посвящен этот раздел. В разделе 2.1 будет рассмотрено употребление показателя *-ges* в компаративной конструкции, в разделе 2.2 описана его аттенуативная функция, раздел 2.3 посвящен употреблению показателя *-ges* при отрицании.

2.1. Употребление в компаративной конструкции

Показатель *-ges* факультативно сочетается с аспектом сравнения, выраженным прилагательным (3) или наречием⁸ (4). При этом различия в значении между предложениями с показателем *-ges* и без него не наблюдается.

(3) a. *vaša peťa-lâš žužât-kes*⁹

Вася Петя-GEN2 высокий-COMP

b. *vaša peťa-lâš žužât*

Вася Петя-GEN2 высокий

‘Вася выше Пети’.

⁵ Такое значение только упоминается, однако примеров авторы не приводят.

⁶ Примеры из литературы цитируются с сохранением системы источника.

⁷ Выбор варианта, как правило, зависит от глухости / звонкости предыдущего согласного: после звонких и после гласных — /ges/, после глухих — /kes/. Однако есть носители, которые предпочитают один из вариантов (*-ges* или *-kes*) вне зависимости от глухости / звонкости согласного перед показателем.

⁸ По крайней мере в некоторых финно-угорских языках зачастую сложно провести границу между прилагательными и наречиями. В данном случае мы опираемся на то, какой частью речи является вершина, от которой зависит рассматриваемое слово: прилагательные зависят от имен, наречия — от глаголов. Заметим, что по крайней мере для хантыйского языка проведенное таким образом разделение на прилагательные и наречия оказывается существенным для показателя *-šâk* (см. раздел 4.1).

⁹ Примеры из бесермянского языка приводятся в общепринятой финно-угорской транскрипции. Примеры из горномарийского и казымского хантыйского языков приводятся в соответствии с системой транскрипции, принятой в соответствующих экспедиционных проектах.

- (4) a. *vaša peía-lâś žog-ges biž-i-z*
 Вася Петя-GEN2 быстро-COMP бежать-PST-3(SG)
 b. *vaša peía-lâś žog biž-i-z*
 Вася Петя-GEN2 быстро бежать-PST-3(SG)
 ‘Вася бежал быстрее Пети’.

Помимо прилагательных и наречий показатели компаратива в бесермянском также могут присоединяться к финитным формам глаголов (5). Аналогичные примеры с выраженным наречием также возможны (6) и зачастую предпочитают примерам типа (5) (так как употребление показателя *-ges* в сочетании с глаголами опрошенные нами носители языка оценивают как «разговорное»¹⁰).

- (5) *so-len nâl-âz berd-e-ges maša śarâś*
 тот-GEN дочь-POSS.3SG плакать-PRS.3SG-COMP Маша о
 ‘Ее дочь плачет чаще, чем Маша’.
 (6) *so-len nâl-âz ċem(-ges) berd-e maša śarâś*
 тот-GEN дочь-POSS.3SG часто(-COMP) плакать-PRS.3SG Маша о
 ‘Ее дочь плачет чаще, чем Маша’.

При присоединении к финитному глаголу показатель *-ges*, помимо частотности события (как в примере (5)), также может модифицировать качественную характеристику ситуации (7). В этом случае более вероятно эксплицитное выражение наречия, значение которого, собственно, и подвергается модификации.

- (7) *mesta-ja-z ul-âś ađami (đeś-kes) tod-e-ges tat-âś mesta-os-tâ*
 место-LOC-POSS.3SG жить-PTCP.ACT человек (хорошо-COMP) знать-PRS.3SG-COMP этот-EL место-PL-ACC.PL
 ‘Местный человек лучше (глубже) знает эти места’.

Среди финно-угорских языков возможностью маркировать глаголы в компаративной конструкции также обладает показатель *-džyk* в коми-зырянском языке [Coates 1982: 126; Todesk 2015].

- (8) КОМИ-ЗЫРЯНСКИЙ
mat-śys sijö bať dor-ś-ys myjlakö pol-is-džyk
 мать-EL.3SG тот отец сторона-EL-POSS.3SG почему-то бояться-PST.3SG-COMP
 ‘Почему-то он(а) боялся(-ась) мать больше, чем отца’ [Todesk 2015: 32]¹¹.

В обоих языках это употребление ограничено неопределёнными формами глаголов: состояниями (понятия *состояния, процесса, вхождения в состояние* мы понимаем так, как они определены у С. Г. Татевосова [Татевосов 2010]¹²) (9) и процессами (10), в то время как глаголы в предельной интерпретации с показателем *-ges* не сочетаются (11) (см. [Todesk 2015: 37] о коми-зырянском).

- (9) *tat-išen až-ik-e-ges*
 этот-EGR видеть-DETR-PRS.3SG-COMP
 ‘Отсюда лучше видно’.
 (10) *peía kalllen vera-śk-i-z-ges nâl murt-leś*
 Петя тихо говорить-DETR-PST-3(SG)-COMP девочка человек-GEN2
 ‘Петя говорил тише, чем девочка’.
 (11) *vasa žad-i-z (*-ges) peía śarâś*
 Вася уставать-PST-3(SG)(*-COMP) Петя о
 ‘Вася устал больше, чем Петя’.

¹⁰ Неясно, что означает эта оценка применительно к бесермянскому: бесермянское наречие не имеет письменности. Отметим, что, тем не менее, четкая граница в суждениях носителей есть: некоторые предложения возможны (хотя и называются разговорными), а некоторые категорически запрещаются.

¹¹ Глоссирование в примерах унифицировано.

¹² Для горномарийского и казымского хантыйского языков были использованы базы акциональности, собранные в экспедициях НИУ ВШЭ и МГУ под руководством С. Ю. Толдовой (в с. Микряково Горномарийского района Республики Марий Эл и в с. Казым Ханты-Мансийского автономного округа соответственно). Для бесермянского языка использовалась база акциональности, собранная автором работы в экспедиции в д. Шамардан Юкаменского района в Удмуртии под руководством М. Н. Усачевой и Т. А. Архангельского.

В число глаголов, которые могут сочетаться с показателем *-ges* в компаративной конструкции в непределельной интерпретации, входят *ažânâ* ‘видеть’, *berdânâ* ‘плакать’, *jaratânâ* ‘любить’, *kâškanâ* ‘бояться’, *mežmânâ* ‘скучать’, *tečânâ* ‘прыгать, танцевать’, *todânâ* ‘знать’, *veltânâ* ‘ходить’, *višânâ* ‘болеть’; в то же время глаголы *baštânâ* ‘брать, взять’, *pârânâ* ‘входить, войти’ с показателем *-ges* не сочетаются.

Исключения составляют глаголы с инкрементальным параметрическим свойством [Татевосов 2010: 118]¹³ в бесермянском¹⁴ (например, *amsêromânâ* ‘сузиться’, *pejmdčikânâ* ‘темнеть’, *perešmânâ* ‘стареть’), которые также сочетаются с показателем *-ges*, хотя и являются предельными (12). Вероятно, это может быть связано с тем, что глаголы с инкрементальным параметрическим свойством вводят шкалу свойства [Kennedy, Levin 2002], на которой показатель *-ges* может производить сравнение.

- (12) *šâd žog sija-z-ges žuk-leš*
суп быстро остывать-3(SG.PST)-COMP каша-GEN2
‘Суп остыл быстрее, чем каша’.

В отличие от глаголов с инкрементальным параметрическим свойством (12), глаголы с инкрементальной темой (13) и с инкрементальным путем¹⁵ (14) не сочетаются с показателем *-ges*.

- (13) **vaša (tros) šij-i-z-ges jablok jura šarâš*
Вася (много) есть-PST-3(SG)-COMP яблоко Юра о
‘Вася съел больше яблок, чем Юра’.

- (14) *vaša gurež-išen wašk-i-z kâđoke(-ges) / *wašk-i-z-ges kaťa-leš*
Вася гора-EGR спускаться-PST-3(SG) далеко(-COMP) / *спускаться-PST-3(SG)-COMP Катя-GEN2
‘Вася спустился с горы дальше, чем Катя’.

Имена существительные и местоимения, как правило, не сочетаются с показателем *-ges* (15).

- (15) **ataj-ges pi-z-e jarat-e čem anaj*
отец-COMP сын-POSS.3(SG)-ACC любить-PRS.3SG чем мать
‘Отец любит сына больше, чем мать’.

Однако показатель *-ges* может присоединяться к существительным и местоимениям в значении ‘в основном’ (вероятно, эта интерпретация также исходит из семантики сравнения: ‘больше остальных’ > ‘в основном’).

- (16) *kiň gurt-ân šij-on lešt-e? — anaj-ges*
кто дом-LOC есть-VN делать-PRS.3SG мать-COMP
‘Кто дома готовит? — В основном мама’.

- (17) *skal mi-ges kâsk-iško-m, anaj-len ataj-len dâr-zâ evâl*
корова мы.NOM-COMP доить-PRS-1PL мать-GEN отец-GEN время-POSS.3PL NEG.EXIST
‘В основном мы доим коров, у родителей нет времени’.

¹³ Согласно [Татевосов 2010: 113] (на основе работ М. Крифки [Krifka 1992, 1998]), «инкрементальность — это свойство (точнее, набор свойств), которыми обладает отношение между ситуацией и ее участником. Инкрементальность имеет место тогда, когда части события находятся в одно-однозначном соответствии с частями участника». Например, отношение в предикате *есть яблоко* инкрементальное, в то время как в предикате *нести яблоко* — нет, так как в первом случае развитие ситуации одно-однозначно соответствует частям яблока, которые исчезают (таким образом, когда яблоко будет съедено полностью, ситуация завершится), а во втором случае такой связи между участником и ситуацией нет. Различают как минимум три вида инкрементальных глаголов: глаголы с инкрементальной темой, инкрементальным путем и инкрементальным параметрическим свойством. Глаголы с инкрементальным параметрическим свойством — это такие глаголы, у которых за время события инкрементально изменяется параметрическое свойство аргумента (например, у глаголов *темнеть*, *толстеть*, *удлинять* и др. за событие изменяется темнота, толщина и длина соответственно) [Татевосов 2010: 118]. Многие из таких глаголов образованы от прилагательных, в литературе их обычно выделяют в особый класс градативов (*degree achievements*).

¹⁴ Заметим, что Т. Тодеск о глаголах этого класса ничего не говорит [Todesk 2015].

¹⁵ Глаголы с инкрементальной темой — глаголы, у которых в инкрементальном отношении с событием находится пациенс/тема [Татевосов 2010: 115] (например, *есть яблоко*, *читать книгу*, *строить дом*). Глаголы с инкрементальным путем — глаголы, у которых в инкрементальном отношении с событием находится траектория/путь аргумента (например, *возвращаться*) [там же: 117].

Среди финно-угорских языков такое употребление также было зафиксировано для показателя *-džyk* в коми-зырянском языке [Coates 1982].

(18) КОМИ-ЗЫРЯНСКИЙ

пестӧ ме-джык кыскал-а

дрова я.NOM-COMP приносить-PRS.1SG

‘Дрова приношу в основном я’ [Coates 1982: 124].

2.2. Аттenuативная функция

Вне компаративной конструкции показатель *-ges* обладает аттenuативным значением (‘немного’) в сочетании с некоторыми прилагательными, наречиями, реляционными именами, а также финитными и нефинитными формами глаголов.

Относительно прилагательных предварительно можно утверждать, что показатель *-ges* сочетается с прилагательными со шкалой, ограниченной минимумом (*lower-closed scale*) (19), а также с полностью открытой шкалой (*open scale*) (20). Термины и типы шкал мы понимаем вслед за определениями в [Kennedy, McNally 2005]. У прилагательных с открытой шкалой нет максимального или минимального значения. У шкал, ограниченных максимумом (*upper-closed scale*), есть максимальное значение; у шкал, ограниченных минимумом, — минимальное. Прилагательные со шкалой, ограниченной максимумом, сочетаются с модификатором *completely* ‘полностью, совершенно’, но не *slightly* ‘немного’. Обратное распределение модификаторов верно для прилагательных со шкалой, ограниченной минимумом. Прилагательные с открытой шкалой сочетаются с модификатором *very* ‘очень’, но не *completely* ‘полностью, совершенно’ и *slightly* ‘немного’ [Kennedy, McNally 2005]. В бесермянском языке для диагностики типа шкалы прилагательных использовался адвербиал *âžžât* ‘немного’¹⁶, а также сама семантика прилагательного (в этой работе мы вынуждены принять допущение, согласно которому тип шкалы прилагательного с заданной семантикой типологически универсален¹⁷).

В случае прилагательных с открытой шкалой для возможности аттenuативного прочтения необходимо указать на степень, относительно которой производится сравнение: «для чего», например, *ta uŝlâ* ‘для этой работы’ в (20), либо указание на такую степень должно быть задано прагматически (21). В противном случае прилагательные с открытой шкалой склонны интерпретироваться в компаративном значении (22). Данное ограничение касается только прилагательных с открытой шкалой и не распространяется на прилагательные со шкалой, ограниченной минимумом (19).

(19) *vaša gožt-i-z kârâž-ges linija*

Вася писать-PST-3(SG) кривой-COMP линия

‘Вася нарисовал кривоватую линию’.

(20) *ta pi murt ta uŝ-lâ lab-ges*

этот парень этот работа-DAT слабый-COMP

‘Этот парень слабоват для этой работы’.

(21) *ta pi murt lab-ges*

этот парень слабый-COMP

{Есть какая-то тяжелая работа, нужно тяжести поднимать, а мальчик не очень сильный, не потянет;} ‘Этот парень слабоват’.

(22) *ta pi murt lab-ges*

этот парень слабый-COMP

‘Этот парень слабее (другого)’.

Прилагательные со шкалой, ограниченной минимумом, которые также, по нашим данным, могут сочетаться с показателем *-ges*: *pož-ges* ‘грязноват(ый)’, *kot-kes* ‘сыроват(ый)’; с открытой шкалой: *piči-ges* ‘маловат’, *zeġ-kes* ‘большеват’, *lapeg-ges* ‘низковат(ый)’, *šekât-ges* ‘тяжеловат(ый)’, *žužât-kes* ‘высоковат(ый)’. Выборка невелика, но в дальнейшем мы попробуем расширить ее (не только для бесермянского, но для всех рассматриваемых нами языков).

¹⁶ Адвербиалы, выражающие высокую степень типа английского *completely* ‘полностью, совершенно’, нами найдены не были.

¹⁷ В дальнейшем данное утверждение должно проверяться на большей типологической выборке.

Однако показатель *-ges* не сочетается с прилагательными со шкалой, ограниченной максимумом (*upper-closed scale*), например, *šońer* ‘прямой’ (23), *dān* ‘чистый’. В этом случае возможна только интерпретация в компаративной функции (приближение к максимуму шкалы).

- (23) *vaša gožt-i-z šońer-ges lĩnija*
 Вася писать-PST-3(SG) прямой-COMP линия
 ‘Вася нарисовал более прямую линию (чем кто-то)’.
 *‘Вася нарисовал довольно-таки прямую линию’.

Естественным образом с прилагательными, не содержащими шкалу (*non-gradable adjectives*), показатель *-ges* также не сочетается, например, с *ulep* ‘живой’ (24).

- (24) *ta ađami ulep(*-kes)*
 этот человек живой(*-COMP)
 ‘Этот человек жив’.

Кроме этого, показатель *-ges* может выступать в аттенуативной функции при сочетании с наречиями:

- (25) *so kallēn-ges lāž-e no vičak valamon*
 тот медленно-COMP читать-PRS.3SG но все понимать.RES
 ‘Он читает немного медленно, но (зато) все понятно’.
- (26) *so ārod-ges až-e*
 тот плохо-COMP видеть-PRS.3SG
 ‘Он плоховато видит’.

Помимо наречий образа действия, показатель *-ges* также сочетается со следующими наречиями: *kāđokān* ‘далеко’ — *kāđokān-ges* ‘далековато’, *arten* ‘рядом’ — *arten-ges* ‘поближе’, *čēn* ‘часто’ — *čēn-ges* ‘частенько’, *šer* ‘редко’ — *šer-ges* ‘редковато’, *kal’* ‘сейчас’ — *kal’-ges* ‘только что’, *sre* ‘потом’ — *sre-ges* ‘попозже’. Предположительно, в наречиях с пространственной семантикой *-ges* оперирует на шкале пути (*kāđokān* ‘далеко’ — *kāđokān-ges* ‘далековато’, *arten* ‘рядом’ — *arten-ges* ‘поближе’), в наречиях, выражающих частотность, — модифицирует частотность (*čēn* ‘часто’ — *čēn-ges* ‘частенько’, *šer* ‘редко’ — *šer-ges* ‘редковато’), а в наречиях времени происходит сдвиг по временной шкале (*kal’* ‘сейчас’ — *kal’-ges* ‘только что’, *sre* ‘потом’ — *sre-ges* ‘попозже’).

Показатель *-ges* также продуктивно сочетается с рядом реляционных имен (реляционные имена по ряду свойств отличаются от существительных с одной стороны и от неизменяемых послелогов с другой стороны [Arkhangelskiy, Usacheva 2015]). В таких сочетаниях показатель *-ges*, по-видимому, также обладает аттенуативной семантикой: ‘немного’ в указанном направлении от заданного ориентира (28).

- (27) *mān-o so-os, až pal-a-zā kijon pot-i-z*
 идти-PRS.3PL тот-PL перед сторона-LOC-POSS.3PL волк выходить-PST-3(SG)
 ‘Едут они, а впереди волк вышел’ (К)¹⁸.

- (28) *so-os-len až pal-a-zā-ges mān-e so*
 тот-PL-GEN перед сторона-LOC-POSS.3PL-COMP идти-PRS.3SG тот
 ‘Она идет немного впереди них’ (К).

В некоторых случаях семантика показателя *-ges* в сочетании с реляционным именем менее предсказуема: так, из сочетания показателя *-ges* с реляционным именем *dor* ‘около’ получается значение ‘поближе (к чему-либо)’ (30).

- (29) *nāl-len velosiped-ez keńer dor-ān o-í sāl-e*
 девочка-GEN велосипед-POSS.3SG изгородь около-LOC тот-PROL стоять-PRS.3SG
 ‘Велосипед девочки стоит там возле изгороди’ (К).

- (30) *vu kiška-n dāđāk-jos dor-ān-ges keńer vāl-a-z*
 вода лить-VN голубь-PL около-LOC-COMP изгородь верх-LOC-POSS.3SG
 ‘Лейка поближе к голубям, на изгороди’ (К).

Согласно корпусным данным и результатам элицитации, *-ges* может сочетаться с такими реляционными именами, как *pal* ‘сторона’, *šor* ‘середина’, *sereg* ‘угол’, *šer / ber* ‘зад’, *dor* ‘около’, *vāl* ‘верх’, *wādes*

¹⁸ Помета (К) означает, что пример взят из корпуса (<http://beserman.ru> и <http://multimedia-corpus.beserman.ru/search>).

‘напротив’. Все реляционные имена, к которым может присоединяться показатель *-ges*, обладают пространственным значением. При этом показатель *-ges* не сочетается со следующими реляционными именами (хотя они также обладают пространственной семантикой): *jâl* ‘верхняя часть’, *visk-ân* ‘между’, *pâr* ‘сквозь, через’, *pumiť* ‘навстречу’. Описанным ограничениям только предстоит найти объяснение.

Показатель *-ges* также может сочетаться с глаголами в аттенуативной функции: с состояниями (например, *mez̄mânâ* ‘скучать’ в (31), *viťânâ* ‘ждать’, *viśânâ* ‘болеть’) и глаголами с инкрементальным параметрическим свойством (например, *kejomânâ* ‘толстеть’ в (32), *pejmâćikânâ* ‘темнеть’, *pereśmânâ* ‘стареть’, *sijanâ* ‘остывать’).

(31) *mon kal̄ mez̄m-iško-ges*
я.NOM сейчас скучать-PRS(1SG)-COMP
‘Я сейчас немножко скучаю’.

(32) *gužem-skân kaťa kejom-i-z-ges*
лето-за Катя толстеть-PST-3(SG)-COMP
‘За лето Катя немного потолстела’.

Однако *-ges* не сочетается с другими глаголами в интерпретации вхождения в состояние (например, *gožtânâ* ‘писать, написать’ в (33), *waśkânâ* ‘спускаться, спуститься’, *bejganâ* ‘успокоиться’, *śijânâ* ‘есть, съесть’) или в процессной интерпретации (например, *vijanâ* ‘капать’ в (34), *suredanâ* ‘рисовать’, *śuntânâ* ‘греть’).

(33) *peťa keňake gožt-i-z(*-ges)=no so ber-e vićak ćušâl-i-z*
Петя немного писать-PST-3(SG)(*-COMP)=ADD тот зад-ILL все зачеркивать-PST-3(SG)
‘Петя написал немного, а потом все зачеркнул’.

(34) *kran-âś vu vija(*-ges)*
кран-EL вода капать.PRS.3SG(*-COMP)
‘Вода немного капает из-под крана’.

Кроме того, показатель *-ges* также сочетается с некоторыми нефинитными формами в аттенуативной функции, однако ограничения на сочетаемость с нефинитными формами нами пока изучены не были.

(35) *mânat babam-e ćut-isa-ges veľt-e*
я.GEN дед-POSS.1SG хромать-CVB-COMP ходить-PRS.3SG
‘Мой дедушка ходит немного прихрамывая’.

(36) *ta ćuškon-ez keňake kwaśt-ono-ges=na*
этот полотенце-ACC немного сушить-DEB-COMP=еще
‘Это полотенце надо еще немного посушить’.

2.3. Употребление при отрицании

Контексты употребления показателя *-ges* в отрицательных предложениях шире, чем в положительных контекстах.

Во-первых, показатель *-ges* сочетается с отрицательным глаголом в контексте, содержащем клитику *=na* ‘все еще’¹⁹. В этом употреблении *-ges* может присоединяться к состояниям (например, *todânâ* ‘знать’ в (37), *oskânâ* ‘верить’, *potânâ* ‘хотеться (что-то сделать)’, *mez̄mânâ* ‘скучать’) и некоторым глаголам в интерпретации вхождения в состояние (например, *žadânâ* ‘устать’ в (38), *sijanâ* ‘остывать, остыть’, *amsâromânâ* ‘сузиться’, *waśkânâ* ‘спускаться, спуститься’, *bâdtânâ* ‘закончить’, *vuňâ* ‘прийти’, *târmânâ* ‘наполниться’, *kwaśmânâ* ‘сохнуть, высохнуть’). В обоих случаях *-ges* привносит компонент значения ‘немного’.

(37) *maď-leś kâl-jos-s-e ďeś ug-ges=na tod-išk-â*
песня-GEN2 слово-PL-POSS.3(SG)-ACC хорошо NEG.PRS-COMP=еще знать-PRS-SG(1)
‘Слова песни хорошо еще немного не знаю / плоховато еще знаю’.

(38) *mon marke e-j-ges=na žad-â kaď*
я.NOM что-то NEG.PST-1SG-COMP=еще устать-SG кажется
‘Я, кажется, еще не устал (досл. немного не устал)’.

¹⁹ Вопросом для дальнейшего изучения остается обязательность самой клитики *=na* в таких примерах: носители утверждают, что в большинстве случаев клитику опустить можно, но предложения с ней всегда «лучше» или «понятнее».

Однако некоторые глаголы вхождения в состояние, имеющие инкрементальную тему, сочетаются с показателем *-ges* только при употреблении сложной глагольной конструкции с глаголом *bâdtânâ* ‘заканчивать’ (39b, 40), в то время как сам глагол с показателем *-ges* при отрицании не сочетается (39a). Мы пока не в силах предложить объяснения этому явлению.

- (39) a. **vaša e-z-ges* *šij-â* *šâd-z-e*
 Вася NEG.PST-3-COMP есть-SG суп-POSS.3(SG)-ACC
 b. *vaša e-z-ges* *bâdt-â* *šâd-z-e* *šij-isa*
 Вася NEG.PST-3-COMP заканчивать-SG суп-POSS.3(SG)-ACC есть-CVB
 ‘Вася чуть-чуть не доел суп’.
- (40) *e-j-ges=na* *bâdt-â* *lâž-isa* *kniga-m-e*
 NEG.PST-1SG-COMP=еще заканчивать-SG читать-CVB книга-POSS.1(SG)-ACC
 ‘Я еще не дочитал эту книгу’.

Во-вторых, употребление *-ges* также возможно в контексте, содержащем клитику *=ni* ‘уже’²⁰. В этом случае показатель *-ges* переводится как ‘почти’. Можно предположить, что это не отдельное значение, а результат интерпретации семантики ‘немного’ в этом контексте: ‘почти’ обозначает такую ситуацию, в которой не достигнуто полностью некоторое состояние или событие, однако разница между текущим состоянием и результирующим небольшая (‘немного’ + ‘(совсем) не болит’ = ‘почти не болит’, ‘немного’ + ‘(совсем) не заходит’ = ‘почти не заходит’).

- (41) *gorlo-je* *tunne* *ug-ges=ni* *viš*
 горло-POSS.1SG сегодня NEG.PRS-COMP=уже болеть.3SG
 ‘Горло почти не болит сегодня’. (Комм.: «чуть-чуть побаливает еще».)
- (42) *so biž-em* *ber-a-z* *miññam ug-ges* *pâr-â=ni*
 тот выйти.замуж-NMLZ зад-LOC-POSS.3SG мы.ILL NEG.PRS-COMP входить-3SG=уже
 ‘После замужества она уже почти к нам не заходит (редко заходит)’.

3. Компаративно-аттенуативная полисемия в горномарийском языке

Показатель *-rak* был заимствован в горномарийский язык из чувашского [Майтинская 1979: 147], где он имеет вид *-pax / -pex*.

- (43) ЧУВАШСКИЙ
йытпă кашкăр-тан пĕчĕк-рех
 собака волк-ABL маленький-COMP
 ‘Собака меньше (поменьше) волка’ [Павлов 1957: 97].

В грамматике А. А. Саватковой указывается, что *-rak* образует сравнительную степень прилагательных (44), а также с помощью него может выражаться значение ‘неполноты качества’ (тоже у прилагательных) (45).²¹ Кроме того, значение ‘неполноты качества’ отмечается у показателя *-rak* в [Казанцева 2005] для лугового марийского языка. Как мы покажем в последующих разделах, этим употребление показателя *-rak* не ограничивается.

- (44) a. *nărăštä* ‘молодой’ — *nărăštä-räk* ‘более молодой, моложе’
 b. *törsär* ‘неровный’ — *törsär-räk* ‘более неровный’ [Саваткова 2002: 129]

²⁰ Аналогично клитике *=na* ‘все еще’, *=ni* можно опустить, но предложения с ней обычно оцениваются как более предпочтительные.

²¹ Кроме того, в докладе Ю. В. Синицыной также рассматривался показатель *-rak* в горномарийском языке [Синицына 2019] (материалом исследования являются данные по говорам с. Кузнецово и окрестных деревень). В работе описаны факторы, влияющие на необязательность показателя *-rak*, аттенуативное значение показателя в сочетании с прилагательными и отрицательными формами глаголов, числительными, именными локативными формами. Результаты автора относительно аттенуативной функции в сочетании с прилагательными, послелогоми и отрицательными формами глаголов сходны с нашими, однако остальные употребления рассмотрены в меньшей подробности, сочетание *-rak* с глаголами в положительной форме не рассмотрено вовсе. Также все-таки имеются некоторые различия между говорами: например, согласно нашим данным, показатель *-rak* не сочетается с числительными.

- (45) a. *tärvä-vlä-žä kăžgä-räk-vlä*
 губа-PL-POSS.3SG толстый-COMP-PL
 ‘губы толстоватые’
 b. *văckăž-räk kăp*
 тонкий-COMP тело
 ‘худоватая фигура’ [Саваткова 2002: 134]

В разделе 3.1 мы рассмотрим употребление показателя *-rak* в компаративной конструкции, в разделе 3.2 — его аттенуативную функцию, раздел 3.3 посвящен употреблению показателя *-rak* в отрицательных предложениях.

3.1. Употребление в компаративной конструкции

В горномарийском языке показатель *-rak*, так же как и показатель *-ges* в бесермянском, факультативен в сравнительной конструкции в сочетании с аспектом сравнения, выраженным прилагательным (46) или наречием (47).

- (46) *vas'a maša găc nelă(-răk)*
 Вася Маша из тяжелый(-COMP)
 ‘Вася тяжелее Маши’.
- (47) *maša mirovoj rekord găc=ăt căn'(-răk) kărgăž-ăn*
 Маша мировой рекорд из=ADD быстро(-COMP) бежать-PRF[3SG]
 ‘Маша пробежала быстрее мирового рекорда’.

Однако, в отличие от бесермянского, в горномарийском языке сочетание глагола с показателем *-rak* в компаративной конструкции невозможно.

- (48) *vas'a sâlăklan-a a masa eče bol'să-răk sâlăklan-a(*-rak)*
 Вася грустить-NPST.3[SG] а Маша еще больше-COMP грустить-NPST.3[SG](*-COMP)
 ‘Вася грустит, а Маша еще сильнее грустит’.

3.2. Аттенуативная функция

Показатель *-rak* также обладает аттенуативным значением: в сочетании с некоторыми прилагательными, наречиями, послелогоми, а также с глаголами.

Так же, как и в бесермянском, в горномарийском языке этот показатель сочетается только с такими прилагательными, которые имеют шкалу, ограниченную минимумом (например, *l'ävărän* ‘грязный’ в (49), *sâlăk* ‘грустный’), или открытую шкалу²² (например, *izi* ‘маленький’ в (50), а также *songă* ‘старый’, *glupă* ‘глупый’, *kužă* ‘длинный’, *kogo* ‘большой’). Заметим, что, опять же, в случае прилагательных с открытой шкалой носители предпочитают предложения с выраженным аргументом-функцией («для чего»).

- (49) *ploška-vlä l'ävărän-răk-vlä eče ikănă măš*
 тарелка-PL грязный-COMP-PL еще раз мыть[IMP.2SG]
 ‘Тарелки грязноваты, помой еще раз’.
- (50) *vas'a texen' kn'igă-m lăd-aš izi-răk âl-eš*
 Вася немного книга-ACC читать-INF маленький-COMP быть-NPST.3SG
 ‘Вася маловат для такой книжки’.

При этом с прилагательными со шкалой, ограниченной максимумом, *-rak* не сочетается (например, *ire* ‘чистый’ в (51), *kukšă* ‘сухой’).

- (51) *ploška-vlä ire(*-răk)-vlă*
 тарелка-PL чистый(*-COMP)-PL
 ‘Тарелки довольно-таки чистые’.

²² Как и в случае с бесермянским, для определения типа шкалы прилагательных использовался адвербиал со значением *iziš* ‘немного’, подходящего аналога для адвербиала *completely* найдено не было, вследствие чего также приходилось опираться на семантику самих прилагательных.

Помимо прилагательных, показатель *-rak* также сочетается как минимум с качественными наречиями в аттенуативной функции.

(52) *vas'a olen-räk sir-ä*
 Вася медленно-COMP писать-NPST.3[SG]
 'Вася медленновато (немного медленно) пишет'.

(53) *mäm-nä-n äškal-vlä-nä čädä(-räk)*
 мы-POSS.1PL-GEN корова-PL-POSS.1PL мало(-COMP)
 'У нас маловато коров'.

Кроме того, показатель *-rak* сочетается с некоторыми пространственными послелогоми (например, *anzäl* 'перед' в (54), *väl* 'верх', *saga* 'рядом')²³.

(54) *maša kat'a anzäl-nä-rak šalg-a*
 Маша Катя перед-IN-COMP стоять-NPST.3[SG]
 'Маша стоит немного перед Катей'.

Показатель *-rak* также может сочетаться с некоторыми неопределёнными глаголами в аттенуативном значении (например, *amalaš* 'спать' в (55) и *karštaš* 'болеть'), а также с предельными глаголами с инкрементальным параметрическим свойством или инкрементальной темой (например, *ükšäš* 'остывать, остыть' (56), *kačkaš* 'есть, съесть'). Однако другие глаголы, имеющие предельную интерпретацию (такие, как *pačäš* 'открывать, открыть' в (57)), с *-rak* не сочетаются.

(55) *tädä eče amal-al-žä-rak vara kän'ält-e-nä*
 он еще спать-ATT-IMP.3SG-COMP потом будить-NPST-1PL
 'Пусть он поспит еще немножко, потом разбудим'.

(56) *lem ükš-äläl-än-räk ugäc äräktä*
 суп остывать-ATT-PRF[3SG]-COMP снова греть[IMP.2SG]
 'Суп немного остыл, подогрей снова'.

(57) *okn'a-m iziš pač-äldal(*-rak)*
 окно-ACC немного открыть-ATT[IMP.2SG](*-COMP)
 'Приоткрой окно!'

Интересно также, что, по-видимому, для употребления показателя *-rak* в этой функции глаголу необходимо иметь показатель аттенуатива (глагольный аттенуатив имеет значение неполноты действия, подробнее см. [Макарчук 2019]). Объяснение этому ограничению еще предстоит найти.

(58) **vuj-em (iziš) karšt-a-rak*
 голова-POSS.1SG (немного) болеть-NPST.3[SG]-COMP
 'Голова болит немного'.

(59) *vuj-em karšt-alal-eš-räk*
 голова-POSS.1SG болеть-ATT-NPST.3[SG]-COMP
 'У меня голова немного побаливает'.

3.3. Употребление при отрицании

Показатель *-rak* также сочетается с глаголами в отрицательных предложениях. В отличие от показателя *-ges* в бесермянском языке, *-rak* располагается на смысловом глаголе, а не на отрицательном.

(60) *lem a-k šol-rak*
 суп NEG.NPST-3 кипеть[SG]-COMP
 'Суп еще не кипит'.

²³ Также нами был зафиксирован один пример (в текстах из корпуса), в котором показатель *-rak* сочетается с существительным в предикативной позиции. Данная тема нуждается в дальнейшем изучении.

(a) *iziš xul'igan-rak*
 немного хулиган-COMP
 'чуть-чуть хулиган'

Контекстом, в котором возможно употребление показателя *-rak* при отрицании, является контекст со словом *eče* ‘все еще’ (напомним, что аналогичный контекст был зафиксирован для бесермянского). В этой функции показатель *-rak* сочетается с вхождениями в состояние и означает, что времени до достижения предела осталось немного. Зачастую этот показатель переводится как «еще не» или «пока не», при том что само слово *eče* ‘все еще’ можно опустить (61).²⁴

- (61) *maša (eče) tol-de-räk*
 Маша (еще) приходит-NEG.PRF[3SG]-COMP
 ‘Маша еще не пришла’. (Комм.: «но скоро должна».)

4. Компаративно-аттенуативная полисемия в казымском диалекте хантыйского языка

Показатель *-šək* заимствован в хантыйский язык из коми-зырянского языка [Fuchs 1949: 169—170; Raun 1971: 120].

- (62) КОМИ-ЗЫРЯНСКИЙ
tani ol-yš-yd tatčössa-jas-tö töd-ö-džyk
 здесь жить-PTCP.ACT-POSS.2SG местный-PL-ACC знать-PRS.3SG-COMP
 ‘Человек, живущий здесь, знает местных лучше’ [Todesk 2015: 32].

В грамматике казымского диалекта хантыйского языка А. Д. Каксина говорится, что показатель *-šək*²⁵ может факультативно образовывать сравнительную форму прилагательных и наречий (63) [Каксин 2007: 101]; кроме того, зафиксировано употребление этого показателя в «относительно-уменьшительном» значении (64) [там же: 43].

- (63) *apše-n näj ewält-an-a kārās-šək*
 младший.брат-POSS.2SG ты от-POSS.2SG-DAT высокий-ATT
 ‘Твой младший брат выше тебя’ [там же: 101].

- (64) *niwi-šək* ‘беловатый’, *waš-šək* ‘узковатый’, *aj-šək* ‘маловатый’ [там же: 43]

Наши данные несколько отличаются от приведенных выше примеров, к тому же в последующих разделах мы покажем и некоторые другие функции показателя *-šək*.

В разделе 4.1 будет рассмотрено употребление показателя *-šək* в компаративной конструкции, в разделе 4.2 описана его аттенуативная функция, раздел 4.3 посвящен употреблению показателя *-šək* при отрицании.

4.1. Употребление в компаративной конструкции

В отличие от показателей *-ges* в бесермянском и *-rak* в горномарийском, показатель *-šək* в казымском хантыйском в сравнительной конструкции не является показателем компаратива. Так, согласно нашим данным, *-šək* в сочетании с прилагательными обладает аттенуативным значением²⁶ (65), в то время как компаративная конструкция образуется без этого показателя (66) (заметим, что примеры типа (66) также были возможны в бесермянском и горномарийском языках из-за факультативности показателей, однако наличие показателей не приносило аттенуативной семантики, как в (65)).

- (65) *lenaj-en kat'aj-en kińsa kārās-šək*
 Лена-POSS.2SG Катя-POSS.2SG чем высокий-ATT
 ‘Лена чуть повыше Кати’.

²⁴ Кроме того, *-rak* сочетается с нефинитными формами глаголов, также добавляя компонент значения ‘еще не’, ‘пока не’: в (а) приведен пример с отрицательным причастием. Однако более подробно сочетания *-rak* с нефинитными формами пока нами изучены не были.

(а) *lem (eče) šol-dät̄(-rak)-ät̄ maša už-eš*
 суп (еще) кипеть-PTCP.NEG(-COMP)-ACC Маша видеть-NPST.3SG
 ‘Маша видит, что суп еще не кипит.’

²⁵ В данном случае мы также цитируем, сохраняя систему записи источника.

²⁶ Аналогично ведет себя в сравнительной конструкции компаративный показатель *-rka* в тундровом ненецком [Berezovskaya 2019: 24].

- (66) *lenaj-en kat'aj-en kińsa kārás*
Лена-POSS.2SG Катя-POSS.2SG чем высокий
'Лена выше Кати'.

Дополнительным подтверждением наличия этого компонента значения у *-šək* является его несовместимость с наречиями, обозначающими высокую степень, например, *wera* 'очень' (67).

- (67) *lenaj-en kat'aj-en kińsa wera kārás(*-šək)*
Лена-POSS.2SG Катя-POSS.2SG чем очень высокий(*-АТТ)
'Лена намного выше Кати'.

Употребление показателя *-šək* в компаративной конструкции в сочетании с наречиями отличается от его употребления с прилагательными. Так, все носители предпочитают вариант с показателем *-šək*, а некоторые носители запрещают наречия без показателя *-šək* в компаративной конструкции: *pāsta-šək* 'быстрее' (68), *siri-šək* 'раньше' (69), то же верно как минимум для наречия *jelən-šək* 'дальше'.

- (68) *vas'aj-en pet'aj-en kińsa pāsta-šək / ??pāsta χəχəl-əs*
Вася-POSS.2SG Петя-POSS.2SG чем быстро-АТТ / ??быстро бежать-PST[3SG]
'Вася бежал быстрее Пети'.

- (69) *vas'aj-en gal'aj-en kińsa siri-šək / ??siri juχt-əs*
Вася-POSS.2SG Галя-POSS.2SG чем раньше-АТТ / ??раньше прийти-PST[3SG]
'Вася пришел раньше Гали'.

Кроме того, в сравнительной конструкции в сочетании с наречием *-šək* аттенуативным значением не обладает (68). В частности, в отличие от прилагательных, наречия, модифицированные показателем *-šək*, сочетаются с наречием *wera* 'очень' (70).

- (70) *vas'aj-en wera pāsta(-šək) pet'aj-en kińsa χəχəl-əs*
Вася-POSS.2SG очень быстро(-АТТ) Петя-POSS.2SG чем бежать-PST[3SG]
'Вася бежал намного быстрее Пети'.

Так же, как и в горномарийском языке, в казымском хантыйском сочетании глагола с показателем *-šək* в компаративной конструкции невозможно (таким образом, эта функция засвидетельствована у компаративного показателя только в бесермянском и коми-зырянском языках).

- (71) a. **mašaj-en (χɯw) ul-əl-šək vas'aj-en ewəlɬ*
Маша-POSS.2SG (долго) спать-NPST[3SG]-АТТ Вася-POSS.2SG от
b. *mašaj-en χɯw-šək ul-əl vas'aj-en ewəlɬ*
Маша-POSS.2SG долго-АТТ спать-NPST[3SG] Вася-POSS.2SG от
'Маша спит больше, чем Вася'.

4.2. Аттенуативная функция

Вне компаративной конструкции показатель *-šək* также обладает аттенуативным значением в сочетании с некоторыми прилагательными, наречиями и послелогоми.

Как и рассмотренные показатели в бесермянском и горномарийском языках, показатель *-šək* в казымском хантыйском обладает аттенуативным значением в сочетании с прилагательными, вводящими шкалу, ограниченную минимумом (72), или открытую шкалу²⁷ (73), и не сочетается с прилагательными со шкалой, ограниченной максимумом²⁸ (74).

- (72) *vas'aj-en tāmxātəl šɯmal'-šək*
Вася-POSS.2SG сегодня грустный-АТТ
'Вася сегодня грустноват'.

²⁷ Как и в случае с бесермянским и горномарийским языками, для определения типа шкалы прилагательных использовался адвербиал *jāša* 'немного', подходящего аналога для адвербиала *completely* найдено не было, вследствие чего также приходилось опираться на семантику самих прилагательных.

²⁸ Как отмечает рецензент, в других вариантах хантыйского языка показатель *-šək* может сочетаться с прилагательными со шкалой, ограниченной максимумом.

- (73) *tām laraś-en lawərt-šək*
 этот ящик-POSS.2SG тяжелый-АТТ
 ‘Эта коробка тяжеловата’.
- (74) *tām wenš mənχ-ti suχ-en sorəm / #sorəm-šək*
 этот лицо вытирать-NFIN.NPST ткань-POSS.2SG сухой / #сухой-АТТ
 ‘Это полотенце сухое / #посуше (чем другое)’.

В нашей выборке есть следующие прилагательные, сочетающиеся с показателем *-šək*: со шкалой, ограниченной минимумом: *χələŋ-šək* ‘грязноват(ый)’, *nivəŋ-šək* ‘влажноват(ый)’, *jijki-šək* ‘мокрват(ый)’, *sakkar-əŋ-šək* ‘сладковат(ый)’; с открытой шкалой: *ləŋsi-šək* ‘тепловат(ый)’, *tinəŋ-šək* ‘дороговат(ый)’, *ləl-šək* ‘низковат(ый)’, *kārəs-šək* ‘высоковат(ый)’.

Помимо прилагательных, некоторые наречия также сочетаются с показателем *-šək* в аттенуативной функции.

- (75) *lɔw ajlta-šək kera lɔŋət-λ təp isa jəm-a uš-a wər-λ*
 он медленно-АТТ хоть читать-NPST[3SG] только всегда хороший-DAT ум-DAT делать-NPST[3SG]
 ‘Он читает немного медленно, но зато все понимает’.
- (76) *vas’aj-en χɔw-šək kiŋškaj-əl lɔŋət-əs*
 Вася-POSS.2SG долго-АТТ книжка-POSS.3SG читать-PST[3SG]
 ‘Вася долговато читал книжку’.

Показатель *-šək* также может сочетаться с некоторыми наречиями времени. Наречие *mulχa* ‘раньше’ в сочетании с показателем *-šək* образует наречие со значением ‘недавно’ (78). Кроме того, *-šək* сочетается со следующими наречиями времени: *isa* ‘всегда’ — *isa-šək* ‘частенько’, *χɔwəŋ* ‘давно’ — *χɔwəŋ-šək* ‘давненько’, *in tām* ‘сейчас’ — *in tām-šək* ‘недавно’. Семантика временных наречий с показателем *-šək* сходна с аналогичными наречиями с показателем *-ges* в бесермянском языке (см. раздел 2.2).

- (77) *mulχa śiw jəŋχ-s-əm*
 раньше DEM.ILL ездить-PST-1SG
 ‘Раньше я туда ездил.’
- (78) *mulχa-šək śiw jəŋχ-s-əm*
 раньше-АТТ DEM.ILL ездить-PST-1SG
 ‘Я недавно туда ездил.’

Однако не все темпоральные адвербиалы сочетаются с показателем *-šək*.

- (79) *mulχatəl(*-šək) jert wə-s*
 вчера(*-АТТ) дождь быть-PST[3SG]
 ‘Вчера шел дождь’.
- (80) *vas’aj-en mɔŋ χuś-ew-a murt śos(*-šək) lɔŋt-ijəl-λ*
 Вася-POSS.2SG мы к-POSS.1PL-DAT редкий час(*-АТТ) заходить-FREQ-NPST[3SG]
 ‘Вася редко к нам заходит’.

Некоторые пространственные послелогои также сочетаются с показателем *-šək*: например, *jəlpi* ‘перед’ (81), *jɔri* ‘за’ (82), *lipi* ‘внутри’, *ilpi* ‘под’. В этой функции показатель *-šək* функционирует аналогично показателям в бесермянском и горномарийском.

- (81) *mašaj-en kat’aj-en jəlpij-ən-šək ləl’-λ*
 Маша-POSS.2SG Катя-POSS.2SG перед-LOC-АТТ стоять-NPST[3SG]
 ‘Маша стоит немного впереди Кати’.
- (82) *pet’aj-en mɔŋ jɔre-w-n-šək šəš-əs*
 Петя-POSS.2SG мы за-POSS.1PL-LOC-АТТ шагать-PST[3SG]
 ‘Петя шел немного позади нас’.

Однако *-šək* сочетается не со всеми пространственными послелогоми, в частности невозможно его употребление с *ewəl* ‘из’ в примере (83), *χonəŋ* ‘рядом’ в (84), *əχti* ‘на’, *tuχti* ‘мимо’. Предварительно можно предположить, что *-šək* возможен только с такими послелогоми, которые позволяют указывать на расстояние между объектами или на продвижение объекта по некоторой шкале (как получается в паре *lipi-ja* внутри-DAT ‘внутри’ — *lipi-ja-šək* внутри-DAT-АТТ ‘поглубже внутри’).

- (83) *λuw ban'a ewəλt(*-šək) kim pit-əs*
он баня из(*-АТТ) на.улицу стать-PST[3SG]
'Он из бани вышел'.
- (84) *kāti-n lipət χonəŋ-ən(*-šək) oməs-λ*
кот-POSS.2SG цветок рядом-LOC(*-АТТ) сидеть-NPST[3SG]
'Кот сидит рядом с цветком'.

Значительное отличие употребления показателя *-šək* от употребления показателей *-ges* в бесермянском и *-rak* в горномарийском заключается в том, что *-šək* не сочетается с финитными формами глагола без отрицания.

- (85) *ma jāša-šək wuli ewəλt pāl-λ-əm(*-šək)*
я немного-АТТ олень от бояться-NPST-1SG(*-АТТ)
'Я немного боюсь оленей'.

С нефинитными формами показатель *-šək* также не сочетается: в (86) причастие не может быть маркировано аттенуативным показателем.

- (86) *ew-en jāša χəχλə-m-əl(*-šək) jurij-ən juχi mān-əs*
девочка-POSS.2SG немного бегать-NFIN.PST-3SG(*-COMP) после.того.как-LOC домой идти-PST[3SG]
'После того как девочка немножко побегала, она пошла домой'.

4.3. Употребление при отрицании

В казымском хантыйском употребление показателя *-šək* также возможно в отрицательных предложениях. Показатель *-šək* сочетается с отрицательным элементом *ān(t)* (*ān(t)-šək*). Как и в бесермянском и горномарийском языках, это употребление довольно продуктивно и разнообразно по значениям.

В казымском хантыйском, как и в бесермянском, показатель *-šək* сочетается с показателем отрицания при глаголах, обозначающих состояния (*kāλti* 'видеть' (88), *wəti* 'знать', *ńəmti* 'помнить', *λāŋχati* 'хотеть').

- (87) *ma sem-ŋəl-am ān kāλ-λ-əŋən*
я глаз-DU-POSS.1DU NEG видеть-NPST-3DU
'Я не вижу (досл. мои глаза не видят)'.
- (88) *ma sem-ŋəl-am ān-šək kāλ-λ-əŋən*
я глаз-DU-POSS.1DU NEG-АТТ видеть-NPST-3DU
'Я плоховато вижу (досл. мои глаза не совсем видят)'.

Так же, как и показатели в бесермянском и горномарийском языках, показатель *-šək* сочетается с глаголами в интерпретации вхождения в состояние (например, *kawərti* 'кипеть, вскипеть' в (89), *juχətti* 'приходить, прийти', *jeišəptəti* 'заканчивать, закончить', *tiwti* 'получиться', *tārəmti* 'хватать, хватить') и глаголами с инкрементальным параметрическим свойством (*pātləti* 'темнеть' в (90), *enəmti* 'расти') в контексте со словом *χилна* 'еще'²⁹.

- (89) *lant jīŋk χилна ānt-šək kawr-əs*
мука вода еще NEG-АТТ кипеть-PST[3SG]
'Суп еще немного не вскипел'. (Комм.: «еще немного, пять минут, подожди».)
- (90) *χилна ān-šək pātlə-s*
еще NEG-АТТ темнеть-PST[3SG]
'Еще не стемнело (но скоро стемнеет)'.

С вхождениями в состояние, являющимися глаголами созидания, показатель *-šək* не сочетается даже при отрицании (*jontti* 'шить, сшить' в (91), 'строить, построить' в (92)). Однако в таких случаях глагол можно заменить на *jeišəptəti* 'закончить', и тогда сочетание *ān-šək* возможно. Заметим, что аналогичный способ превратить неграмматичное предложение в грамматичное путем использования глагола с семантикой 'заканчивать' также был нами зафиксирован в бесермянском языке (см. раздел 2.3).

²⁹ Как и в случае с бесермянским и горномарийским языками, само слово *χилна* 'еще' можно опустить, однако носители отдадут предпочтение предложениям с эксплицитно выраженной частицей.

(91) *jernas-em ǎn-šək *jont-s-em / jetšaptə-s-em*
 платье-POSS.1SG NEG-ATT *шить-РСТ-1SG>SG / закончить-РСТ-1SG>SG
 ‘Платье чуть-чуть не закончила, немного не дошила’.

(92) *joχ-λ-an χot-əλ ǎn-šək jetšapt-əs-el*
 люди-PL-POSS.2SG дом-POSS.3SG NEG-ATT закончить-РСТ-3PL>SG
 ‘Мужики еще немного не достроили дом’.

В казымском хантыйском при референции к плану будущего или настоящего времени возникает значение категорического отрицания, которое на русский язык носителями, как правило, переводится как ‘ни за что’. Это значение было зафиксировано нами только в казымском хантыйском, в остальных рассматриваемых нами языках оно невозможно. Можно предположить, что это значение возникает в результате того, что аттенуатив находится в сфере действия отрицания: NEG[АТТ[VP]]. Таким образом, отрицается небольшая степень — получаем некоторую высокую степень AUG[VP], которая при референции к плану будущего дает значение категорического отрицания. Почему в данном случае аттенуатив находится в сфере действия отрицания, а в примерах выше — наоборот (88)—(92), еще только предстоит установить.

(93) *λmw ǎn-šək kašas-λ*
 он NEG-ATT соглашаться-NPST[3SG]
 ‘Он ни за что не согласится / ни за что не соглашается’.

(94) *vas’aj-en učit’el’-a ǎn-šək pit-λ*
 Вася-POSS.2SG учитель-DAT NEG-ATT стать-NPST[3SG]
 ‘Вася ни за что не станет учителем’. (Комм.: «сам категорически не хочет».)

При референции к плану прошедшего времени такое значение невозможно.

(95) **λmw ǎn-šək kašas-əs*
 он NEG-ATT соглашаться-РСТ[3SG]
 Ожид.: ‘Он ни за что не соглашался / наотрез не согласился’.

Для возможности употребления *ǎn-šək* также важно, чтобы у референта было (не)желание, то есть он должен быть агентивным (96). В частности, с неодоушевленными участниками *ǎn-šək* в значении «категорического отрицания» не сочетается (97)³⁰.

(96) *ta məš-a χəp / *ǎn-šək ji-λ-əm*
 я болезнь-DAT NEG.EMPH / NEG-ATT стать-NPST-1SG
 {Мама говорит надеть шапку, а то заболеешь, а я ей отвечаю:} ‘Я точно не заболею!’

(97) *pās-an-en ǎn / *ǎn-šək šnkal-əλ*
 стол-POSS.2SG NEG / *NEG-ATT сломаться-NPST[3SG]
 ‘Стол не сломается. / Ожид.: *Стол точно (ни за что) не сломается’.

5. Обсуждение результатов

В данном разделе мы подводим итоги и сравниваем показатели с компаративно-аттенуативной полисемией в трех рассматриваемых нами языках: *-ges* в бесермянском, *-šək* в казымском хантыйском и *-rak* в горномарийском.

Все три показателя способны выполнять функции, обсуждавшиеся в этой работе: они употребляются в компаративной конструкции, имеют аттенуативную функцию, а также могут сочетаться с отрицательным глаголом / отрицательной частицей (в зависимости от языка).

В таблице 1 представлены основные функции показателей *-ges*, *-rak*, *-šək* в сравнении с языками, в которых, по данным литературы, имеет место компаративно-аттенуативная полисемия (безотносительно частеречной принадлежности слова, к которому присоединяется показатель). Как было упомянуто ранее, такой тип полисемии зафиксирован в еще одном финно-угорском языке, а именно в коми-зырянском: показатель *-džuk* может употребляться в компаративной конструкции, иметь аттенуативную функцию (как минимум в сочетании с прилагательными [Bartens 2000: 135]), а также сочетается с отрицательным глаголом. Помимо финно-угорских языков, компаративно-аттенуативная полисемия встреча-

³⁰ Однако это употребление также нуждается в дальнейшем изучении в поисках возможных ограничений.

ется как минимум еще в одном уральском языке (самодийской ветви): показатель *-rka* в тундровом ненецком употребляется в компаративной конструкции и обладает аттенуативной функцией, однако нам неизвестно, возможно ли употребление этого показателя в сочетании с отрицательным глаголом. Наконец, этот тип полисемии встречается в некоторых тюркских языках, однако о возможности употребления при отрицании опять же нам ничего неизвестно.

Таблица 1. Основные функции показателей *-ges*, *-rak* и *-šək* в типологической перспективе

	Компаративная конструкция	Аттенуативная функция	Употребление при отрицании
Бесермянский <i>-ges</i>	+	+	+
Горномарийский <i>-rak</i>	+	+	+
Казымский хантыйский <i>-šək</i>	+	+	+
Коми-зырянский <i>-džyk</i> [Coates 1982: 124—127; Todesk 2015; Bartens 2000: 135]	+	+	+
Ненецкий <i>-rka</i> [Nikolaeva 2014: 134, 174; Буркова 2010: 261] (самодийские < уральские)	+	+	?
Чувашский <i>-rak</i> [Павлов 1957: 94], мишарский татарский <i>-rak</i> [Былина 2017: 448] и др. (тюркские) ³¹	+	+	?

Далее в этом разделе мы сравним показатели в трех языках, описанию которых были посвящены разделы 2—4, а именно в бесермянском, казымском хантыйском и горномарийском (отмечая сходства с другими языками в тех случаях, когда нам про них что-либо известно).

5.1. Обобщения: компаративные конструкции

В компаративной конструкции в сочетании с прилагательными показатель *-šək* в казымском хантыйском обладает аттенуативным значением. Согласно П. Березовской [Berezovskaya 2019: 24], так же ведет себя в сравнительной конструкции показатель *-rka* в тундровом ненецком. Вероятно, эти показатели довольно далеки от прототипических показателей компаратива и скорее ближе к показателям аттенуатива.

Во всех остальных рассмотренных нами языках показатели в компаративных конструкциях сочетаются с аспектом сравнения, выраженным прилагательным, только факультативно (при этом данные показатели не имеют аттенуативного компонента значения, так как сочетаются с наречиями высокой степени, например, ‘очень’).

Помимо прилагательных и наречий, показатели компаратива в бесермянском и коми-зырянском языках также могут присоединяться к финитным формам глаголов. В обоих языках это употребление ограничено непредельными формами глаголов (за исключением предельных глаголов с инкрементальным параметрическим свойством в бесермянском, которые также сочетаются с показателем *-ges*). Также показатель *-ges* в бесермянском языке может присоединяться к существительным и местоимениям в значении ‘в основном’. Это употребление также было зафиксировано для показателя *-džyk* в коми-зырянском языке.

Особенности употребления показателей *-ges*, *-rak* и *-šək* и их сочетаемость с различными частями речи в компаративной конструкции приведены в таблице 2.

Таблица 2. Употребление показателей *-ges*, *-rak* и *-šək* в компаративной конструкции

	Бесермянский <i>-ges</i>	Горномарийский <i>-rak</i>	Казымский хантыйский <i>-šək</i>
Прилагательные	(+) ³²	(+)	АТТ
Наречия	(+)	(+)	+
Глаголы	(+)	–	–

³¹ Показатель *-rak* встречается во многих тюркских языках [Левитская 1976: 147]. По-видимому, диахроническое аттенуативное употребление у этого показателя в тюркских языках также присутствует [Тенишев 2002: 311, 434, 525].

³² Скобки (+) в этой таблице обозначают факультативность показателя.

5.2. Обобщения: аттенуативная функция

Во всех рассматриваемых нами языках показатели обладают аттенуативным значением, в котором они могут сочетаться с прилагательными, наречиями и реляционными именами / послелогоми. В бесермянском и горномарийском языках показатели могут сочетаться еще и с глаголами. В таблице 3 приведена сочетаемость показателей *-ges*, *-rak* и *-šək* в аттенуативной функции.

Таблица 3. Употребление показателей *-ges*, *-rak* и *-šək* в аттенуативной функции

	Бесермянский <i>-ges</i>	Горномарийский <i>-rak</i>	Казымский хантыйский <i>-šək</i>
Прилагательные	+	+	+
Наречия	+	+	+
Реляционные имена / послелогои	+	+	+
Глаголы	+	+	–

Несмотря на то, что выборка прилагательных, находящаяся в нашем распоряжении, пока невелика, мы можем с осторожностью заметить, что показатели *-ges*, *-rak* и *-šək* скорее сочетаются с прилагательными с открытой шкалой, а также со шкалой, ограниченной минимумом. Такой же сочетаемостью обладает аффикс *-ovat* в русском языке [Kagan, Alexeyenko 2011].

Во всех рассматриваемых нами языках показатели сочетаются с некоторыми наречиями. Кроме того, показатели сочетаются с пространственными послелогоми.

Употребление показателя *-ges* в сочетании с глаголами в бесермянском и показателя *-rak* в горномарийском схоже: оба показателя сочетаются с состояниями и глаголами с инкрементальным параметрическим свойством, но не сочетаются с другими глаголами, обозначающими вхождение в состояние³³.

5.3. Обобщения: употребление при отрицании

Рассматриваемые нами показатели в контексте глагольного отрицания имеют больше значений, чем в положительных контекстах.

Общим контекстом является контекст со значением ‘еще не’: он встречается во всех трех языках. Глаголы в таком контексте чаще всего являются вхождениями в состояние. В казымском хантыйском и бесермянском *-šək* и *-ges* также сочетаются с показателем отрицания при глаголах, обозначающих состояния или являющихся глаголами с инкрементальным параметрическим свойством.

Кроме того, в бесермянском есть контекст с клиткой *=ni* ‘уже’, в котором при добавлении показателя *-ges* получается значение, переводимое на русский язык как ‘почти’. В казымском хантыйском при референции к плану будущего или настоящего возникает значение категорического отрицания ‘ни за что’.

6. Заключение

В этой работе мы описали семантику и функционирование показателей, обладающих компаративно-аттенуативной полисемией, в трех финно-угорских языках: *-ges* в бесермянском, *-rak* в горномарийском и *-šək* в казымском диалекте хантыйского языка.

Эти показатели употребляются в компаративной конструкции, обладают аттенуативной функцией и сочетаются с отрицанием. В этой работе мы также описали сочетаемость показателей в различных функциях. Показатели имеют много схожих функций и употреблений, однако по многим особенностям и отличаются друг от друга.

Описание показателей *-ges*, *-rak* и *-šək*, предложенное в работе, неполно. Нами были выявлены определенные тенденции, не зафиксированные ранее в литературе, которые и представлены в этой статье. Таким образом, настоящая работа может стать основой для дальнейших исследований в этой области.

³³ Исключение составляют процессы: в то время как в бесермянском процессы не сочетаются с показателем *-ges* в аттенуативной функции, некоторые процессы в горномарийском языке сочетаются с показателем *-rak*. На данный момент мы не до конца понимаем, от чего зависит это распределение.

Сокращения

Глоссы

1, 2, 3 — 1, 2, 3 лицо	IMP — императив
ABL — аблатив	IN — инессив
ACC — аккузатив	INF — инфинитив
ADD — аддитивная частица	LOC — локатив
ATT — аттенуатив	NEG — отрицание
AUG — аугментатив	NFIN — нефинитная форма
COMP — компаратив	NMLZ — номинализация
CVB — конверб	NOM — номинатив
DAT — датив	NPST — непрошедшее время
DEV — дебитив	PL — множественное число
DEM — демонстратив	POSS — посессив
DETR — детранзитив	PRF — перфект
DU — двойственное число	PROL — пролатив
EGR — эгрессив	PRS — настоящее время
EL — элатив	PST — прошедшее время
EMPH — эмфатическая частица	PTCP.ACT — действительное причастие
EXIST — экзистенциальный глагол	PTCP.NEG — отрицательное пассивное причастие
FREQ — фреквентатив	RES — результатив
GEN — генитив	SG — единственное число
GEN2 — второй генитив	VN — отглагольное имя
ILL — иллатив	

Литература

- Буркова 2010 — С. И. Буркова. Краткий очерк грамматики тундрового диалекта ненецкого языка // Диалектологический словарь ненецкого языка / Ред. Н. Б. Кошкарева. Екатеринбург, 2010. С. 179—349.
- Былинина 2017 — Е. Г. Былинина. Сравнительные конструкции // Элементы татарского языка в типологическом освещении. Мишарский диалект / Ред. С. Г. Татевосов, А. Г. Пазельская, Д. Ш. Сулейманов. М., 2017. С. 448—466.
- Ефремов 2002 — Д. А. Ефремов. Грамматические свойства имени прилагательного в удмуртском языке. Дисс. ... канд. филол. наук. Ижевск, 2002.
- Казанцева 2005 — И. В. Казанцева. Функционально-семантическое поле компаративности в марийском литературном языке. Дисс. ... канд. филол. наук. Йошкар-Ола, 2005.
- Каксин 2007 — А. Д. Каксин. Казымский диалект хантыйского языка. Ханты-Мансийск, 2007.
- Левитская 1976 — Л. С. Левитская. Историческая морфология чувашского языка. М., 1976.
- Майтинская 1979 — К. Е. Майтинская. Историко-сопоставительная морфология финно-угорских языков. М., 1979.
- Макарчук 2019 — И. В. Макарчук. Глагольный аттенуатив в горномарийском языке // Acta Linguistica Petropolitana. Труды ИЛИ РАН. 2019, 15 (2). С. 130—162.
- Павлов 1957 — Материалы по грамматике современного чувашского языка. Часть первая. Морфология / Ред. И. П. Павлов. Чебоксары, 1957.
- Первошиков и др. 1962 — Грамматика современного удмуртского языка / Ред. П. Н. Первошиков, В. М. Вахрушев, В. И. Алатырев, А. А. Поздеева, И. В. Тараканов. Ижевск, 1962.
- Саваткова 2002 — А. А. Саваткова. Горное наречие марийского языка. Сомбатхей, 2002.
- Синицына 2019 — Ю. В. Синицына. Морфосинтаксис и семантика горномарийского компаративного суффикса *-rak* // XVI Конференция по типологии и грамматике для молодых исследователей (Санкт-Петербург, 21—23 ноября 2019 г.).
- Татевосов 2010 — С. Г. Татевосов. Акциональность в лексике и грамматике. Дисс. ... д. филол. наук. М., 2010.
- Тенишев 2002 — Сравнительно-историческая грамматика тюркских языков. Региональные реконструкции / Отв. ред. Э. Р. Тенишев. М., 2002.
- Усачева и др. 2017 — Тезаурус бесермянского наречия: Имена и служебные части речи (говор деревни Шамардан) / Ред. М. Н. Усачева, Т. А. Архангельский, О. Л. Бирюк, В. А. Иванов, Р. И. Идрисов. М., 2017.
- Черемисинова 2019 — М. О. Черемисинова. Компаративно-аттенуативная полисемия в бесермянском удмуртском // Типология морфосинтаксических параметров. 2019, 2 (2). С. 128—150.

- Arkhangelskiy, Usacheva 2015 — T. Arkhangelskiy, M. Usacheva. Syntactic and Morphosyntactic Properties of Postpositional Phrases in Beserman Udmurt as Part-of-Speech Criteria // SKY Journal of Linguistics. 2015. Vol. 28. P. 103—137.
- Bartens 2000 — R. Bartens. Permiläisten kielten rakenne ja kehitys [The structure and development of Permic languages]. Helsinki, 2000.
- Berezovskaya 2019 — P. Berezovskaya. Comparison constructions in Nenets: about a language that has degrees and also uses them // Урал-алтайские исследования. 2019, 4 (35). С. 18—35.
- Coates 1982 — J. G. Coates. The -джык Comparative Suffix in Contemporary Komi Usage // Transactions of the Philological Society. 1982. Vol. 80, 1. P. 119—129.
- Efremov, Selmezy 2018 — D. Efremov, S. Selmezy. The comparative and the superlative in Udmurt and Hungarian // Ежегодник финно-угорских исследований. 2018, 12 (4). С. 6—14.
- Efremov, Selmezy 2019 — D. Efremov, S. Selmezy. On the question of adjective quality degrees in the Udmurt and Chuvash languages // Ежегодник финно-угорских исследований. 2019, 13 (4). С. 591—598.
- Fuchs 1949 — D. R. Fuchs. Der Komparativ und Superlativ in den finnischen-ugrischen Sprachen // Finnisch-ugrischen Forschungen. 1949. Vol. 30. P. 147—230.
- Kagan, Alexeyenko 2011 — O. Kagan, S. Alexeyenko. Degree modification in Russian morphology: The case of the suffix *-ovat* // Proceedings of Sinn und Bedeutung 15 / Eds. Reich, Ingo et al. Germany, 2011. P. 321—336.
- Kennedy, Levin 2002 — C. Kennedy, B. Levin. Telicity corresponds to degree of change. Unpublished Ms. Northwestern University, Stanford University, 2002.
- Kennedy, McNally 2005 — C. Kennedy, L. McNally. Scale structure and the semantic typology of gradable predicates // Language. 2005. Vol. 81. P. 1—37.
- Krifka 1992 — M. Krifka. Thematic relations as links between nominal reference and temporal constitution // Lexical matters / Eds. I. Sag, A. Szabolsci. Stanford, 1992.
- Krifka 1998 — M. Krifka. The origins of telicity // Events and Grammar / Ed. S. Rothstein. Dordrecht, 1998.
- Nikolaeva 2014 — I. Nikolaeva. A Grammar of Tundra Nenets. Berlin, 2014.
- Raun 1971 — A. Raun. Essays in Finno-Ugric and Finnic Linguistics. Bloomington, 1971.
- Todesk 2015 — T. Todesk. The verbal augmentative and the inherent properties of verbs in Komi // Linguistica Uralica. 2015. Vol. 51, 1. P. 28—43.

References

- Arkhangelskiy, Usacheva 2015 — T. Arkhangelskiy, M. Usacheva. Syntactic and Morphosyntactic Properties of Postpositional Phrases in Beserman Udmurt as Part-of-Speech Criteria. In: SKY Journal of Linguistics. 2015. Vol. 28. P. 103—137.
- Bartens 2000 — R. Bartens. Permiläisten kielten rakenne ja kehitys. Helsinki, 2000. {R. Bartens. The structure and development of Permic languages. Helsinki, 2000.}
- Berezovskaya 2019 — P. Berezovskaya. Comparison constructions in Nenets: about a language that has degrees and also uses them. In: Ural-Altai Studies. 2019, 4 (35). P. 18—35.
- Burkova 2010 — S. I. Burkova. Kratkij ocherk grammatiki tundrovogo dialekta neneckogo yazyka. In: Dialektologicheskii slovar' nenetskogo yazyka / Red. N. B. Koshkareva. Ekaterinburg, 2010. S. 179—349. {S. I. Burkova. A concise grammatical sketch of Tundra Nenets grammar. In: Dialect dictionary of Nenets / Ed. N. B. Koshkareva. Ekaterinburg, 2010. P. 179—349.}
- Bylinina 2017 — E. G. Bylinina. Sravnitel'nye konstrukcii. In: Elementy tatarskogo yazyka v tipologicheskom osveshchenii. Misharskij dialekt / Red. S. G. Tatevosov, A. G. Pazel'skaya, D. S. Suleymanov. M., 2017. S. 448—466. {E. G. Bylinina. Comparative constructions. In: Elements of Tatar language in typological perspective. Mishar dialect / Eds. S. G. Tatevosov, A. G. Pazel'skaya, D. S. Suleymanov. Moscow, 2017. P. 448—466.}
- Cheremisina 2019 — M. O. Cheremisina. Komparativno-attenuativnaja polisemija v besermyanskom udmurtskom. In: Tipologiya morfosintaksicheskikh parametrov. 2019, 2 (2). S. 128—150. {M. O. Cheremisina. Comparative and attenuative polysemy in Beserman Udmurt. In: Typology of Morphosyntactic Parameters. 2019, 2 (2). P. 128—150.}
- Coates 1982 — J. G. Coates. The -джык Comparative Suffix in Contemporary Komi Usage. In: Transactions of the Philological Society. 1982. Vol. 80, 1. P. 119—129.
- Efremov 2002 — D. A. Efremov. Grammaticheskie svojstva imeni prilagatel'nogo v udmurtskom yazyke. Diss. ... kand. filol. nauk. Izhevsk, 2002. {D. A. Efremov. Grammatical properties of adjectives in Udmurt. PhD thesis. Izhevsk, 2002.}
- Efremov, Selmezy 2018 — D. Efremov, S. Selmezy. The comparative and the superlative in Udmurt and Hungarian. In: Yearbook of Finno-Ugric Studies. 2018, 12 (4). P. 6—14.

- Efremov, Selmeczy 2019 — D. Efremov, S. Selmeczy. On the question of adjective quality degrees in the Udmurt and Chuvash languages. In: Yearbook of Finno-Ugric Studies. 2019, 13 (4). P. 591—598.
- Fuchs 1949 — D. R. Fuchs. Der Komparativ und Superlativ in den finnischen-ugrischen Sprachen. In: Finnisch-ugrischen Forschungen. 1949. Vol. 30. P. 147—230.
- Kagan, Alexeyenko 2011 — O. Kagan, S. Alexeyenko. Degree modification in Russian morphology: The case of the suffix -ovat. In: Proceedings of Sinn und Bedeutung 15 / Eds. Reich, Ingo et al. Germany, 2011. P. 321—336.
- Kaksin 2007 — A. D. Kaksin. Kazym'skij dialekt khantyjskogo yazyka. Hanty-Mansijsk, 2007. {A. D. Kaksin. Kazym dialect of Khanty language. Hanty-Mansijsk, 2007.}
- Kazantseva 2005 — I. V. Kazantseva. Funkcional'no-semanticheskoe pole komparativnosti v marijskom literaturnom yazyke. Diss. ... kand. filol. nauk. Yoshkar-Ola, 2005. {I. V. Kazantseva. Functional-semantic map of comparativity in Standard Mari. PhD thesis. Yoshkar-Ola, 2005.}
- Kennedy, Levin 2002 — C. Kennedy, B. Levin. Telicity corresponds to degree of change. Unpublished Ms. Northwestern University, Stanford University, 2002.
- Kennedy, McNally 2005 — C. Kennedy, L. McNally. Scale structure and the semantic typology of gradable predicates. In: Language. 2005. Vol. 81. P. 1—37.
- Krifka 1992 — M. Krifka. Thematic relations as links between nominal reference and temporal constitution. In: Lexical matters / Eds. I. Sag, A. Szabolsci. Stanford, 1992.
- Krifka 1998 — M. Krifka. The origins of telicity. In: Events and Grammar / Ed. S. Rothstein. Dordrecht, 1998.
- Levitskaya 1976 — L. S. Levitskaya. Istoricheskaya morfologiya chuvashskogo yazyka. M., 1976. {L. S. Levitskaya. Historical morphology of Chuvash language. Moscow, 1976.}
- Makarchuk 2019 — I. V. Makarchuk. Glagol'nyj attenuativ v gornomarijskom yazyke. In: Acta Linguistica Petropolitana. Trudy ILI RAN. 2019, 15 (2). S. 130—162. {I. V. Makarchuk. Verbal attenuative in Hill Mari. In: Acta Linguistica Petropolitana. 2019, 15 (2). P. 130—162.}
- Maytinskaya 1979 — K. E. Maytinskaya. Istoriko-sopostavitel'naya morfologiya finno-ugorskih yazykov. M., 1979. {K. E. Maytinskaya. Historical comparative morphology of Finno-Ugric languages. Moscow, 1979.}
- Nikolaeva 2014 — I. Nikolaeva. A Grammar of Tundra Nenets. Berlin, 2014.
- Pavlov 1957 — Materialy po grammatike sovremennogo chuvashskogo yazyka. Chast' pervaya. Morfologiya / Red. I. P. Pavlov. Cheboksary, 1957. {Materials on the grammar of modern Chuvash. Part 1. Morphology / Ed. I. P. Pavlov. Cheboksary, 1957.}
- Perevoshchikov et al. 1962 — Grammatika sovremennogo udmurtskogo yazyka / Red. P. N. Perevoshchikov, V. M. Vakh-rushev, V. I. Alatyrev, A. A. Pozdeeva, I. V. Tarakanov. Izhevsk, 1962. {A grammar of modern Udmurt / Eds. P. N. Perevoshchikov, V. M. Vakh-rushev, V. I. Alatyrev, A. A. Pozdeeva, I. V. Tarakanov. Izhevsk, 1962.}
- Raun 1971 — A. Raun. Essays in Finno-Ugric and Finnic Linguistics. Bloomington, 1971.
- Savatkova 2002 — A. A. Savatkova. Gornoe narechie marijskogo yazyka. Sombatkhei, 2002. {A. A. Savatkova. Hill dialect of the Mari language. Szombathely, 2002.}
- Sinitsyna 2019 — J. V. Sinitsyna. Morfosintaksis i semantika gornomarijskogo komparativnogo suffiksa -rak. In: XVI Konferenciya po tipologii i grammatike dlya molodyh issledovatelej (Sankt-Peterburg, 21—23 noyabrya 2019 g.). {J. V. Sinitsyna. Morphosyntax and semantics of Hill Mari comparative marker -rak. In: 16th Conference on typology and grammar for young scholars (St. Petersburg, 21—23 November 2019).}
- Tatevosov 2010 — S. G. Tatevosov. Aktsional'nost' v leksike i grammatike. Diss. ... doct. filol. nauk. M., 2010. {S. G. Tatevosov. Actionality in lexicon and grammar. Doctoral thesis. Moscow, 2010.}
- Tenishev 2002 — Sravnitel'no-istoricheskaya grammatika tyurkskih yazykov. Regional'nye rekonstrukcii / Red. E. R. Tenishev. M., 2002. {Comparative grammar of Turkic languages. Regional reconstructions / Ed. E. R. Tenishev. Moscow, 2002.}
- Todesk 2015 — T. Todesk. The verbal augmentative and the inherent properties of verbs in Komi // Linguistica Uralica. 2015. Vol. 51, 1. P. 28—43.
- Usacheva et al. 2017 — Tezarus besermyanskogo narechiya: Imena i sluzhebnye chasti rechi (govor derevni Shamardan) / Red. M. N. Usacheva, T. A. Arkhangelskiy, O. L. Biryuk, V. A. Ivanov, R. I. Idrisov. M., 2017. {Beserman thesaurus: nouns, adjectives, numerals and functional parts of speech (Shamardan idiom) / Eds. M. N. Usacheva, T. A. Arkhangelskiy, O. L. Biryuk, V. A. Ivanov, R. I. Idrisov. Moscow, 2017.}

Требования к оформлению статей / Style sheet

Резюме

Просьба приложить к статье два резюме: 1) написанное на языке статьи, 2) написанное на одном из языков журнала (напоминаем, что это русский, английский, немецкий), но не на языке статьи.

Оформление

Допустимые форматы файла: .rtf или .doc.

Просим Вас также обязательно присылать текст в формате .pdf или распечатку.

Должны использоваться шрифты, поддерживающие Unicode.

В начале статьи указываются фамилия автора и место его работы.

Заголовки выделяются **полужирным шрифтом**.

Языковые примеры выделяются *курсивом*; смысловые выделения отмечаются **полужирным шрифтом** или **разрядкой**.

Примеры, которые занимают отдельную строку или абзац, просьба нумеровать; номера ставятся в начале в круглых скобках.

Длинные цитаты даются как отдельный абзац и отделяются в начале и в конце одной пустой строкой.

Цитаты заключаются в «кавычки».

Переводы и значения выделяются ‘одинарными (марровскими) кавычками’.

Просьба не расставлять переносы.

Язык и орфография

В статьях на английском языке просьба последовательно придерживаться британского или американского варианта и соответствующей орфографии; в статьях на немецком языке — старых либо новых орфографических правил.

Если статья написана не на родном языке автора, мы рекомендуем дать текст на вычитку носителю языка.

Библиографические ссылки

В тексте статьи по образцу:

в квадратных [] скобках Фамилия год: страница (страницы) (например, [Aijmer 1996: 22—25]).

В конце статьи дается список использованной литературы по приводимым ниже образцам.

Если статья написана не на русском языке, то в библиографии к ней работы на русском и других языках с кириллическим шрифтом можно дать отдельным блоком, без латинской транскрипции. В любом случае ссылка в тексте статьи должна соответствовать написанию в списке литературы.

Summaries

Please provide two summaries: 1) in the language of your paper, 2) in one of the languages of the journal (Russian, English, German), but not in the language your paper is written in.

Formatting

The file is to be submitted in one of the following formats: .rtf or .doc.

Please also add either a .pdf version or a hard copy of the text.

Please use Unicode fonts.

Please indicate the author's name and affiliation on the title page.

Headings have to be **boldfaced**.

Use *italics* for language data; **boldface** or `spacing` for emphasis.

Please number language examples which are not in the body of the text; enclose each number in parentheses.

Please format long quotations as separate paragraphs; add one blank line before and after them.

Please use «double quotes» for quotations.

Please use 'single quotes' for meanings.

Please do not hyphenate the document.

Language and spelling

In English articles please use consistently either British or American norm and the respective spelling; in German articles please follow either the old or the new orthography rules.

If the paper is not written in the author's native language, we highly recommend to have it checked by a native speaker.

References

In the body of the text please use the following format:

[Name year: page(s)] (e.g. [Aijmer 1996: 22—25]).

In papers not in Russian works in Cyrillic can be listed separately, without Latin transcription. In any case, their spelling must be identical with the references cited in the text.

At the end of the paper add a list of references as exemplified below.

Монография, словарь с именами авторов на титульном листе / Monographic work or dictionary with author's (authors') name(s) on the title page

Грунина 1991 — Э. А. Грунина. Историческая грамматика турецкого языка. М., 1991.

Janhunen 1977 — J. Janhunen. Samojedischer Wortschatz: Gemeinsamojedische Etymologien. Helsinki, 1977.

Коллективная монография, словарь с именем ответственного редактора на титульном листе / Collection of papers or dictionary with editor's (editors') name(s) on the title page

Этническая история 1982 — Этническая история народов Севера / Отв. ред. И. С. Гурвич. М., 1982.

Lessing 1960 — Mongolian-English Dictionary / Ed. F. Lessing. Berkeley; Los Angeles, 1960.

Статья в коллективной монографии, сборнике / Paper in an (edited) book

Дёрфер 1986 — Г. Дёрфер. О языке гуннов // Зарубежная тюркология. Вып. 1. Древние тюркские языки и литературы. М., 1986. С. 71—134.

Дмитриев 1958 — Н. К. Дмитриев. О тюркских элементах русского словаря // Лексикографический сборник. Вып. 3. М., 1958. С. 3—47.

Kiefer, Gyuris 2006 — F. Kiefer, B. Gyuris. Szemantika // A magyar nyelv kézikönyve / Ed. F. Kiefer. Budapest, 2006.

Nikolaeva 2005 — I. Nikolaeva. Agreement and linguistic construal // Uralic languages today / Ed. J. Fernandez-Vest. Paris, 2005.

Статья в журнале / Journal paper

Иллич-Свитыч 1963 — В. М. Иллич-Свитыч. Алтайские дентальные: *t, d, δ* // ВЯ. 1963, 6. С. 37—56.

Clark 1980 — L. V. Clark. Turkic Loanwords in Mongol. I: The Treatment of non-initial *s, z, š, č* // Central Asiatic Journal. 1980. Vol. 24, 1—2. P. 23—45.

Doerfer 1969 — G. Doerfer. Ein altosmanisches Lautgesetz im Kurdischen // Wiener Zeitschrift für die Kunde des Morgenlandes. 1969. Bd. 62. S. 250—263.

Цифровые и интернет-источники / Digital and Internet sources

Helimski 2007 — E. Helimski. База данных энецкого языка // www.helimski.com, 2007.

Редактор
С. И. Переверзева

Адрес редакции
125009, г. Москва, Б. Кисловский пер., 1, корп. 1,
Отдел урало-алтайских языков

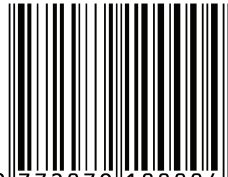
Телефон
+7 (495) 691-63-06

Editor
Svetlana Pereverzeva

Editorial office
125009, Moscow, B. Kislovskiy sidestr., 1, 1,
Department of Uralo-Altaic languages

Phone
+7 (495) 691-63-06

ISSN 2079-1003



9 772079 100004 >