

**Вычисление степени морфологической близости  
восточных хантыйских диалектов (сургутского, ваховского и васюганского)  
на онлайн платформе ЛингвоДок**

**В. В. Воробьёва**

*Институт системного программирования им. В. П. Иванникова  
Российской академии наук, г. Москва, Российская Федерация;  
Национальный исследовательский Томский политехнический университет,  
г. Томск, Российская Федерация,  
vorobeva@tpu.ru*

**И. В. Новицкая**

*Национальный исследовательский  
Томский государственный университет  
г. Томск, Российская Федерация,  
irno2012@yandex.ru*

**АННОТАЦИЯ**

**Введение.** В статье затрагиваются обсуждаемые в научной литературе вопросы о принадлежности трёх диалектов хантыйского языка (ваховского, васюганского и сургутского) к одной или разным ветвям восточного кластера идиомов и о степени их морфологического размежевания относительно друг друга. Индивидуальные особенности каждого из анализируемых диалектов, выявленные до настоящего времени на разных уровнях языковой системы, не могут считаться полностью изученными, в связи с чем целесообразно применение современных цифровых инструментов для математической обработки корпусных данных, позволяющих выявить весь спектр различающихся показателей в их морфологических системах.

**Цель:** представление промежуточных результатов вычисления степени морфологической близости трёх восточных диалектов хантыйского языка на основе данных, исследовательской платформы ЛингвоДок.

**Материалы исследования:** корпуса ваховского, васюганского и сургутского диалектов хантыйского языка, размещённые на исследовательской платформе ЛингвоДок.

**Результаты и научная новизна.** Отправной точкой для анализа послужили наблюдения хантологов, которые ранее описали некоторые разноразличные особенности восточных диалектов. Использование опции ЛингвоДока позволило выявить список из 74 групп аффиксов-когнатов, а также 14 аффиксов, уникальных для одного из диалектов хантыйского языка. Из общего числа выявленных когнатов 77,02 % аффиксов имеют общие этимологические связи во всех трёх диалектах, что позволяет отнести эти идиомы к одной ветви восточных диалектов. Оставшиеся 21,62 % аффиксов этимологически связаны только в ваховском и васюганском диалектах, что обеспечивает их близкую к стопроцентной морфологическую близость. Из числа 14 уникальных аффиксов 12 аффиксов засвидетельствованы только в сургутском диалекте. Полученные данные свидетельствуют в пользу заключения о том, что сургутский диалект обнаруживает достаточную степень дистанцированности от ваховско-васюганского диалектного единства. Ранее вычисление морфологической близости восточных диалектов инструментальным путём в онлайн режиме не проводилось.

**Ключевые слова:** хантыйский язык, ваховский диалект, васюганский диалект, сургутский диалект, морфология, корпуса текстов, ЛингвоДок, анализ данных, документирование языков

**Благодарности:** Исследование выполнено при поддержке РФФ, проект № 20-18-00403 «Цифровое описание диалектов уральских языков на основании анализа больших данных».

**Для цитирования:** Воробьёва В. В., Новицкая И. В. Вычисление степени морфологической близости восточных хантыйских диалектов (сургутского, ваховского и васюганского) на онлайн платформе ЛингвоДок // Вестник угроведения. 2024. Т. 14. № 4 (59). С. 616–629.

**Calculation of the degree of morphological proximity  
of the Eastern Khanty dialects (Surgut, Vakh and Vasyugan)  
on the online platform LingvoDoc**

**V. V. Vorobyova**

*Ivannikov Institute for System Programming of the Russian Academy of Sciences,  
Moscow, Russian Federation,  
National Research Tomsk Polytechnic University, Tomsk, Russian Federation,  
vorobeva@tpu.ru*

I. V. Novitskaya

National Research Tomsk State University,  
Tomsk, Russian Federation,  
irno2012@yandex.ru

## ABSTRACT

**Introduction:** the article addresses some issues raised in the scientific literature, namely, about the affiliation Vakh, Vasyugan and Surgut Khanty to one or different branches of the Eastern cluster of idioms and the degree of their morphological variation in comparison with one another. The individual features of each dialect under analysis cannot be considered comprehensive, hence it is advisable to use modern digital tools to mathematically process large databases of text corpora with the purpose to identify an entire spectrum of differences in their morphological systems.

**Objective:** presentation of preliminary results of calculating the degree of morphological proximity of three Eastern dialects of the Khanty language based on data of the research platform LingvoDoc.

**Research materials:** the corpus of the Vakh, Vasyugan and Surgut dialects of the Khanty language, posted on the LingvoDoc research platform.

**Results and novelty of the research:** the starting point for the analysis was observations of specialists in the Khanty language who had previously described some of the different-level features of the Eastern dialects. Using the LingvoDoc option allowed us to identify a list of 74 groups of cognate affixes, as well as 14 affixes unique to one of the dialects of the Khanty language. Of the total number of identified cognates 77.02% of affixes have common etymological links in all three dialects, which makes it possible to attribute these idioms to one branch of the Eastern dialects. The remaining 21.62% of affixes are etymologically related only in the Vakh and Vasyugan dialects, which ensures their close morphological proximity. Of the 14 unique affixes, 12 affixes are attested only in the Surgut dialect. The data obtained support the conclusion that the Surgut dialect reveals a sufficient degree of distancing from the Vakh-Vasyugan dialect unity. Previously, the calculation of the morphological proximity of the Eastern dialects was not carried out instrumentally online.

**Key words:** Khanty language, Vakh dialect, Vasyugan dialect, Surgut dialect, morphology, text corpora, LingvoDoc, data analysis, language documentation

**Acknowledgments:** the study was funded by Russian Science Foundation, project № 20-18-00403 “Digital Description of Uralic Dialects Based on Big Data Analysis”.

**For citation:** Vorobyova V. V., Novitskaya I. V. Calculation of the degree of morphological proximity of the Eastern Khanty dialects (Surgut, Vakh and Vasyugan) on the online platform LingvoDoc // Vestnik ugrovedenia = Bulletin of Ugric Studies. 2024; 14 (4/59): 616–629.

## Введение

Согласно современной диалектной классификации [5], такие идиомы, как ваховский, васюганский и сургутский, относятся к восточному кластеру диалектов хантыйского языка и противопоставляются западным диалектам (приртышскому, среднеобскому, казымскому, шурышкарскому и приуральскому). Обозначения восточных идиомов связаны с названиями притоков реки Обь: Юган, Вах и Васюган. Юганский считается одним из говоров сургутского диалекта [5, 16]. Сургутский диалект в настоящем исследовании представлен именно этим говором (далее будет обозначаться как сургутский).

Статус ваховского и васюганского идиомов вызывает разногласия у хантологов. С одной точки зрения, эти идиомы образуют одно ваховско-васюганское наречие восточных ханты [5; 9, 8; 14, 1; 17]. Согласно другой, ваховский и васюганский относятся к разным восточно-хантыйским диалектам [8, 45; 13; 12, 172; 18; 19].

По мнению Н. И. Терёшкина, различие между ваховским и васюганским разновидностями языка устанавливается на фонологическом

уровне: первый диалект относится к *l*-типу, второй – к *j*-типу. Основное различие выявляется в использовании начальной фонемы в словоформах: ср. форму местоимения 3-го лица с [l] в ваховском *lōŷ* и форму с [j] в васюганском *jōŷ*. К числу других примеров указанной закономерности можно отнести следующие: *lol – jol* ‘зарубка (на лезвии), трещина’, *lop – jop* ‘удочка’, *lort – jort* ‘корень’ [7, 212]. В словарных материалах наблюдаются также и лексические расхождения, например, вах. *kārəntəŷəltä* ‘надевать обувь’, вас. *kärtä* ‘снимать шкуру с животного’, вас. *joləmnəŷ* ‘песня, исполняемая во время ритуала’, вах. *joŷətam* ‘после, затем’ [7] и другие. Тем не менее, указанная принадлежность только одному из диалектов не может исключать потенциального наличия его в другом из-за недостаточности сведений. Детальное описание каждого из идиомов по-отдельности не привело к подробному сопоставительному исследованию фонетических, лексических и грамматических явлений ваховского и васюганского диалектов с целью установления всего спектра параметров их расхождений.

Ваховскому и васюганскому диалектам противопоставляется кластер сургутских идиомов, которые относятся к  $\lambda$ -типу. Ещё один из диалектов, который принадлежал к восточной группе по морфологическим признакам – это салымский [7, 16]. Язык салымских ханты считается  $t$ -типом. вах., вас. *uləm* – сур. *uləm* – сал. *utəm* ‘сон, сон’ [7, 500], вах., вас. *weltä* – сур. *wälta* – сал. *wätta* ‘охотиться’ [7, 520]. Считается, что переход [ $\lambda$ ] в [t] начался в иртышско-кондинском диалекте около 200 лет назад и постепенно распространился на север и восток [2, 87; 16, 10]. Салымские ханты, среди которых живущих носителей сейчас не осталось, были первыми, кого затронул этот процесс, в то время как сургутские ханты этот процесс переживают в настоящее время.

Отличительной чертой ваховско-васюганской подгруппы является сингармонизм, который в некоторой степени также присутствует в сургутском и салымских идиомах [6, 17; 10, 11, 24; 15, 16, 19]. Ваховский и васюганский имеют значительные расхождения с сургутским диалектом в области морфологии. При работе с двумя восточными кластерами сразу становится очевидным количественное варьирование временных форм от одной до четырёх, о чём упомянуто в работах Н. И. Терёшкина и М. Чепреги [6; 9]. Однако для выявления всех морфологических расхождений между восточными подгруппами хантыйских диалектов требуется последовательный анализ морфологических форм на базе большого массива данных.

Современные цифровые инструменты исследовательской платформы ЛингвоДок позволяют обработать имеющиеся корпуса текстовых и словарных данных на восточных диалектах хантыйского, размещённых на ресурсе, и выявить весь спектр морфологических показателей, указывающих на векторы индивидуального развития каждого из идиомов.

### Материалы и методы

Исследование выполнено на корпусах ваховского, васюганского и сургутского диалек-

тов восточных ханты, которые находятся в открытом доступе на исследовательской платформе ЛингвоДок ([https://lingvodoc.ispras.ru/language\\_databases](https://lingvodoc.ispras.ru/language_databases)).

Корпус ваховского диалекта хантыйского языка<sup>1</sup> объёмом около 12 тыс. словоформ включает в себя сказки, сказания, грамматические анкеты по заданным тематикам, прозаические тексты и тексты других жанров. Этот корпус отражает современное состояние языка, т. к. собран по материалам экспедиции в 2017–2019 гг. в п. Корлики ХМАО-Югры В. В. Воробьёвой и С. В. Ковылиным.

Корпус васюганского диалекта хантыйского языка<sup>2</sup> в объёме около 1,5 тыс. словоформ включает в себя памятник «Материалы для изучения языка васюганских инородцев (остяков) Нарымского края» св. Макария (Невского) (1887) и тексты, собранные в Каргасокском районе Томской области и переведённые на русский язык Л. И. Калининой в 1966 г. Тексты глоссированы А. Ю. Фильченко в 2009 г. и обработаны для системы ЛингвоДок В. В. Воробьёвой.

Материалы для сургутского диалекта хантыйского языка<sup>3</sup> объёмом около 1 тыс. словоформ на платформе ЛингвоДок были заимствованы из проекта «Обско-угорская база данных: анализ текстовых корпусов и словари наименее описанных обско-угорских диалектов» Е. К. Скрибник<sup>4</sup>. Вошедшие в корпус тексты сказок были собраны Х. Паасоненом в 1901 г., позже отредактированы, переведены на немецкий язык и опубликованы Э. Вертеш в 2001 г. Тексты конвертированы в формат.odt для размещения на ЛингвоДоке М. П. Безеновой.

Анализ был проведён с применением традиционных методов (интерпретация, обобщение и др.) и методов корпусного анализа (диверсифицированный поиск, составление контекстов и конкордансов, подсчёт частотности употребления и составление статистики, анализ разметки текста, конвертация и др.)

<sup>1</sup> Корпус текстов ваховского диалекта хантыйского языка. Автор В. В. Воробьёва. URL: <https://lingvodoc.ispras.ru/dictionary/7349/279/perspective/7349/283/view/> (дата обращения: 11.09.2024). Корпус тестов ваховского диалекта хантыйского языка. Автор С. В. Ковылин. URL: <https://lingvodoc.ispras.ru/dictionary/3021/71/perspective/3021/75/view/> (дата обращения: 05.09.2024).

<sup>2</sup> Корпус текстов васюганского диалекта хантыйского языка. Памятник св. Макария (Невского). Автор С. В. Ковылин. URL: <https://lingvodoc.ispras.ru/dictionary/3021/71/perspective/3021/75/edit/> (дата обращения: 01.09.2024); Корпус тестов васюганского диалекта хантыйского языка. Тексты Л. И. Калининой. Автор обработки материала В. В. Воробьёва. URL: <https://lingvodoc.ispras.ru/dictionary/2488/8764/perspective/2488/8764/view/> (дата обращения: 12.09.2024).

<sup>3</sup> Корпус юганского говора сургутского диалекта хантыйского языка Автор обработки материала М. П. Безенова. URL: <https://lingvodoc.ispras.ru/dictionary/743/10734/perspective/743/10737/edit/> (дата обращения: 10.10.2024).

<sup>4</sup> Корпус текстов юганского диалекта хантыйского языка Мюнхенского университета. Автор Е. К. Скрибник. URL: [https://www.babel.gwi.uni-muenchen.de/index.php?abfrage=PA\\_corpus&subnavi=corpus\\_pub/](https://www.babel.gwi.uni-muenchen.de/index.php?abfrage=PA_corpus&subnavi=corpus_pub/) (дата обращения: 11.09.2024).

## Результаты Исследовательские инструменты платформы ЛингвоДок

Вычисление степени близости восточных диалектов было проведено в виртуальной лаборатории ЛингвоДок ([www.lingvodoc.ispras.ru](http://www.lingvodoc.ispras.ru)).

ЛингвоДок – это сервис, ориентированный на документацию языков с помощью создания словарей любой структуры, установления этимологических связей между ними и осуществления расчётов по языковым данным [3; 4, 62]. После регистрации любой пользователь платформы имеет возможность создавать свои текстовые корпуса и словари, а также анализировать материалы в открытом доступе, разработанные другими авторами.

В настоящее время на платформе доступны около двух тысяч словарей и корпусов, созданных на основании аудиословарей, собранных в полевых условиях в формате *.wav*, и архивных записей по уральским и алтайским языкам [1, 252].

При разработке корпусов исследовательская платформа ЛингвоДок поддерживает разметку текстов в формате ELAN. Также система обеспечивает хранение, просмотр и обработку корпусов, размеченных в ELAN и подгруженных в формате *.odt* на ЛингвоДок. Просмотр текстов обеспечивается специальной веб-программой, которая показывает текст в многослойном формате. Наглядное представление корпуса можно увидеть на рис. 1:

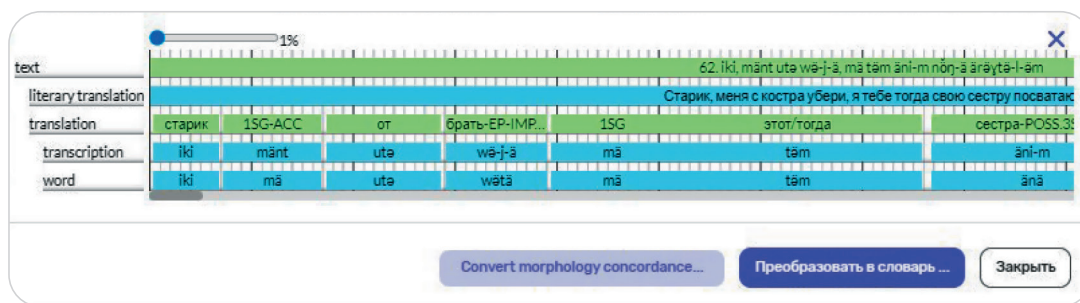


Рис. 1. Просмотр текста в формате *.odt* на платформе ЛингвоДок

Словари также представлены как настраиваемые многослойные структуры, при использовании которых пользователь может иметь полный или ограниченный доступ и, в зависимости от этого, самостоятельно определять тип требуемой информации в столбцах и количество последних. Например, можно выбрать столбцы звук, разметка, когнаты,

праформа и др. На ЛингвоДоке есть также встроенная программа просмотра корпусов в формате ELAN, которая позволяет конвертировать корпуса текстов в конкордансы с лексическим или парадигматическим входом. Каждый морфологический вход, например, разделён на 4 колонки, что проиллюстрировано на рис. 2:

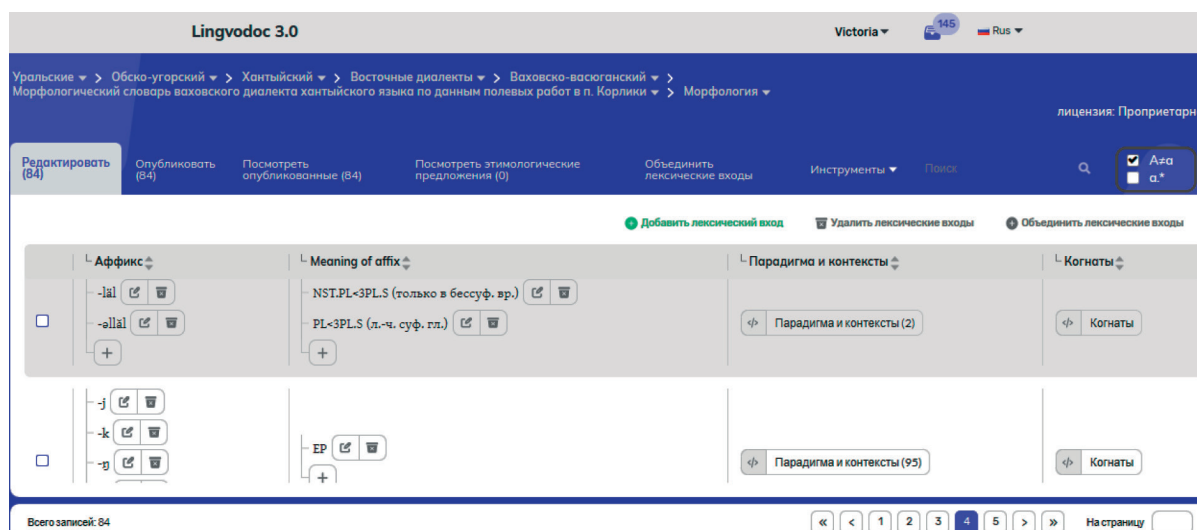


Рис. 2. Морфологический словарь на платформе ЛингвоДок



Первая колонка «Аффикс» представляет сам суффикс, который можно вручную дополнить алломорфами. Во второй колонке «Значение суффикса» указываются все его значения, которые можно вручную корректировать: добавлять или изменять значение

суффикса. В третьей колонке «Парадигма и контексты» подгружаются все контексты употребления суффикса, а также указывается их количество. Корректировка в третьей колонке не допускается. Для наглядности см. рис. 3.

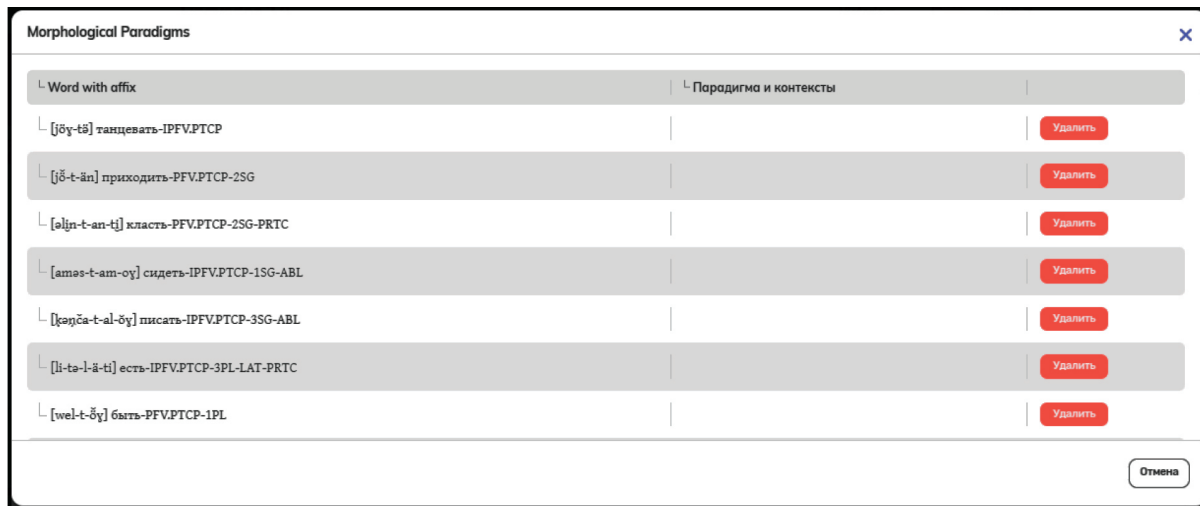


Рис. 3. Морфологические парадигмы

После конвертации текста в морфологический словарь требуется ручная обработка словаря, которая заключается в снятии омонимии. Словарь можно постоянно обновлять, дополняя данными из другого корпуса или вновь загруженного на платформу, после каждого обновления требуется пройти про-

цедуру снятия омонимии аффиксов. Четвёртая колонка «Когнаты» даёт возможность связать каждый аффикс полученного словаря этимологически с другими диалектами и языками ручным способом, выбрав родственный аффикс в нужном диалекте или языке (см. рис. 4):

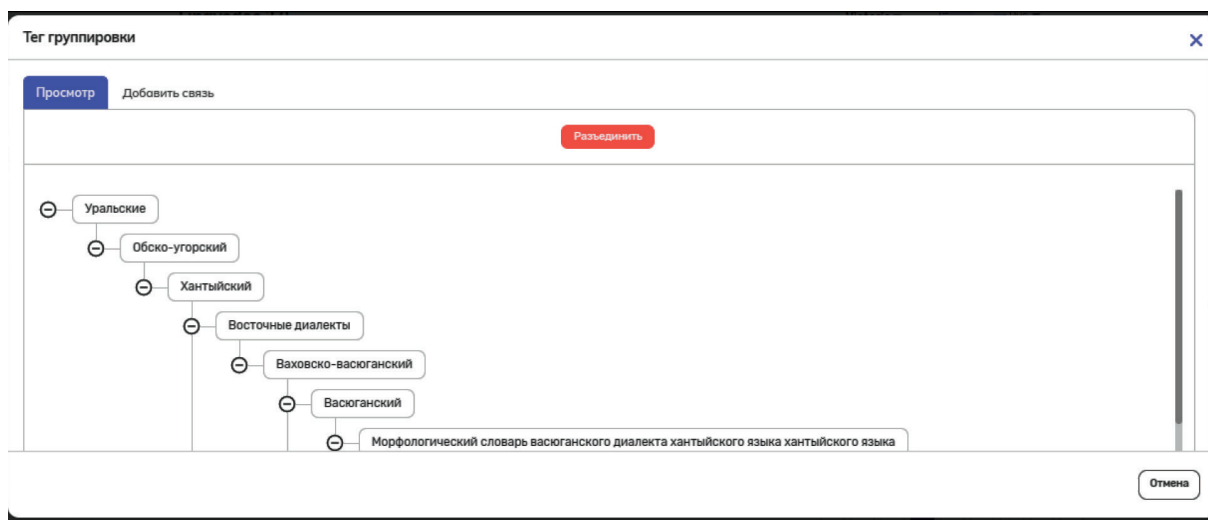


Рис. 4. Ручной способ связи этимологий

В каждом словаре есть меню «Инструменты», в котором в настоящее время авторам доступны 15 опций для анализа, из них несколько программ доступны сторонним пользователям [4, 65]. Один из инструментов «Поиск результатов в разных языках или диалектах» позволяет

связать этимологии машинным способом. Однако собственный опыт работы на платформе показывает, что машина связывает корректно только около 10 % этимологий, остальные этимологии приходится связывать вручную через поиск когнатов в разных диалектах или языках.

В будущем эта проблема может быть устранена в большой мере за счёт создания универсальной системы обозначения глосс и детализации разметки, а также замены или удаления некоторых символов, нечитаемых машиной. В любом случае, машине всегда нужна будет помощь лингвиста, поскольку ошибки при разметке текстов, связанные с человеческим фактором, вряд ли можно будет полностью исключить.

Программные инструменты ЛингвоДока позволяют выполнить фонетический, морфологический и этимологические анализы с большой скоростью и производительностью при условии выполнения всех ранее описанных этапов. Ранее подобные анализы выполнялись лингвистами вручную и занимали колоссальное количество времени по сравнению с машинным математическим подсчётом языковых данных. Стоит отметить, что в системе есть также диалоговое окно поиска, которое позволяет в любом словаре найти нужную морфему точно и просмотреть все её парадигмы и когнаты.

### Результаты вычисления и их интерпретация

Расчёт и анализ близости восточных диалектов был выполнен на платформе ЛингвоДок с применением инструмента «Степень морфологической близости разных диалектов». Программа высчитывает отдалённость разных диалектов на основе математического анализа и представляет результаты на онлайн платформе моментально в виде широкой базы связанных или несвязанных аффиксов, которая требует описания и обобщения промежуточных или окончательных результатов. Результаты также можно скачать в формате EXCEL. Следует отметить, что вычисление близости восточных диалектов хантыйского языка с помощью онлайн инструментов было проведено впервые.

Машине была поставлена задача обчислить морфологическую близость восточных диалектов. В выборку вошли ваховский, васюганский и юганский говор сургутского диалекта. После обчёта данных машина выдаёт общие сведения о задействованных источниках, что показано на рис. 5.



**Рис. 5.** Обобщение данных, вошедших в обсчёт

В нашем случае было проанализировано 3 словаря, 74 группы когнатов и 223 транскрипции. В общей сложности в словарях было проанализировано 239 лексических входов: 71 в сургутском, по 84 в ваховском и васюганском. Программой предусмотрено также предоставить отчёт о не вошедших в обсчёт данных и причинах их исключения. В приведённом примере одна группа когнатов была исключена из-за отсутствия лексических записей в двух выбранных словарях, что отразилось на результатах анализа в виде небольшой погрешности.

Группы когнатов, выделенные программой, распределяются по подгруппам: 1) 54 когната, этимологически связанных во всех трёх словарях, 2) 16 когнатов, связанных только в двух словарях. Также программе удалось выявить

14 аффиксов, которые представлены только в одном из трёх диалектов.

Программный аппарат платформы ЛингвоДок показывает морфологическую близость между диалектами по данным обработанных словарей схематически, в виде этимологического дерева расстояний между восточными диалектами (см. рис. 6). На схеме красная точка представляет «Морфологический словарь юганского говора сургутского диалекта», сиреневая точка – «Морфологический словарь ваховского диалекта» и зелёная точка «Морфологический словарь васюганского диалекта». Схема демонстрирует морфологическую близость ваховского и васюганского диалектов и отдалённость от них, в некоторой степени, сургутского диалекта.

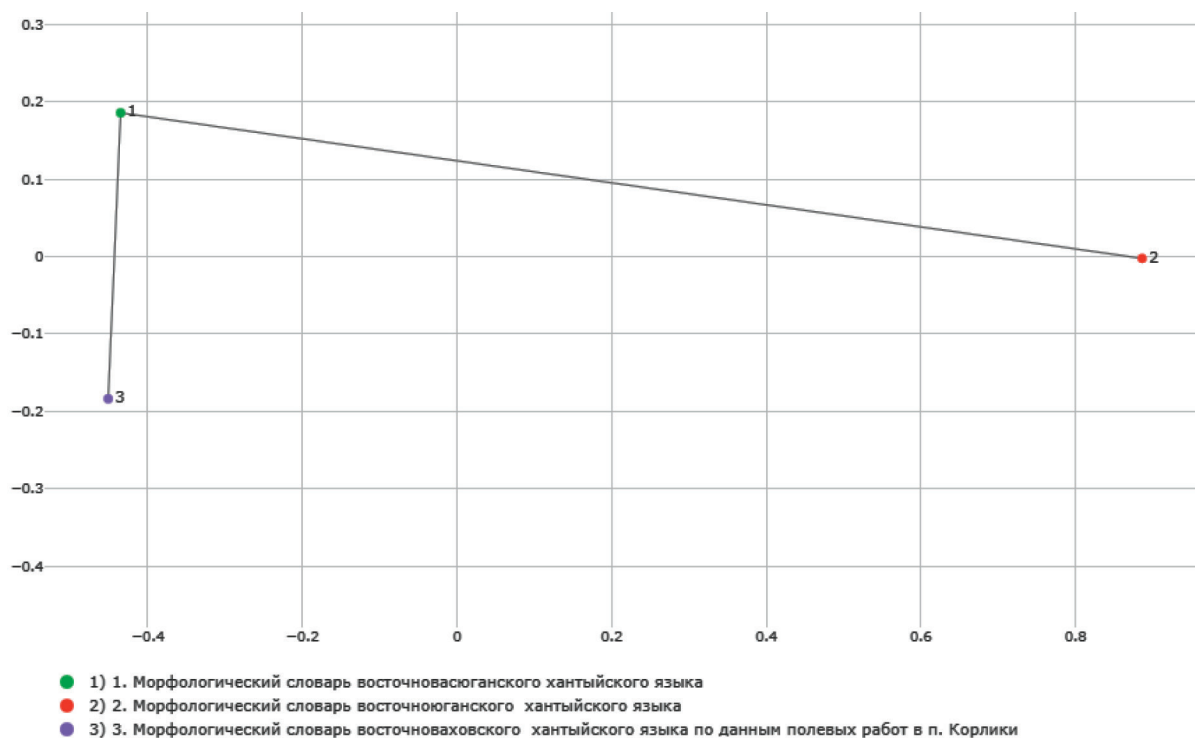


Рис. 6. Этимологическое дерево расстояний

Результат математического анализа, проделанного машиной, представлен также в процентном соотношении. На рис. 7 показано, что морфологическая близость ваховского и васюганского диалекта составляет 100 %, в то время как близость морфологических систем васюганского диалекта и сургутского диалекта составляет 85 %, а ваховского и сургутского – 87 %.

ганского диалекта составляет 100 %, в то время как близость морфологических систем васюганского диалекта и сургутского диалекта составляет 85 %, а ваховского и сургутского – 87 %.

	1. Морфологический словарь васюганского диалекта хантыйского языка хантыйского языка	2. Морфологический словарь юганского говора сургутского диалекта хантыйского языка	3. Морфологический словарь ваховского диалекта хантыйского языка по данным полевых работ в п. Корлики
1	n/a	0.99 (87%)	-0.00 (100%)
2	1.12 (85%)	n/a	0.99 (87%)
3	-0.00 (100%)	0.88 (88%)	n/a

Рис. 7. Морфологический обcчёт близости восточных диалектов в процентах

По результатам анализа словарных и корпусных данных трёх восточных диалектов хантыйского языка, была сформирована общая сводная база аффиксов. Результаты обcчёта показывают, что большая часть аффиксов (54) обнаруживает общие этимологические связи. Если аффикс из данной группы имеет общую этимологию только с одним из выбранных для обcчёта словарей, то колонка в словаре, в котором этимология отсутствует, остаётся пустой (см. рис 8а). Например, васюганский предикативный аффикс *-iki-* не обнаруживает соответствия в сургутском диалекте, но имеет параллельные форму в ваховском диалекте (см. рис. 8а). Схожая ситуация и с глагольным аффиксом немаркированного времени *-iyän-/iyän-* для 3SG субъектного спряжения, который не представлен в сургутском.

Общее количество проанализированных групп аффиксов, этимологически связанных в словарях, включает 74 подгруппы. Из них 54 аффикса этимологически связаны во всех трёх диалектах, 16 аффиксов связаны только в двух – ваховском и васюганском. Группа аффиксов, уникальных для какого-то одного диалекта, включает 14 морфем.

В общей выборке из 54 показателей, этимологически связанных во всех трёх диалектах, именные и глагольные аффиксы образуют количественно равные подгруппы, лишь небольшое число аффиксов включает в себя другие показатели, например, показатель формы числительного (ORD). Большая часть этимологически связанных глагольных аффиксов являются показателями лично-числовых форм глагола в не прошедшем или прошедшем времени (NPST, PST),

1. Морфологический словарь важоганского диалекта хантыйского языка хантыйского языка	2. Морфологический словарь юганского говора сургутского диалекта хантыйского языка	3. Морфологический словарь ваховского диалекта хантыйского языка по данным полевых работ в п. Корлики
-ali ( DIM )	-li ( DIM.MEL )	-alji; -āli; -ali; -āli ( DIM ) -li ( ATTR )
-at ( PASS.3PL )	-et ( PASS.3PL )	-át; -āt; -āt; -āt; -āt ( PASS (только в 3PL) )
-i; -j ( PASS.3SG )	-i ( PASS.3SG; PST.PASS.3SG )	-j; -i ( PASS.3SG; NST.PASS (только в бессуф. вр. PASS) )
-iki ( PRD )		-ikā; -iki; -jki; -ki; -ki; -q; -qā; -qi; -iki; -āqi; -āki; -ēki; -ēki; -ēqā; -ēqi; -ēkj; -ēki; -ēki ( PRD )
-il ( MULT (асп. суф.) )	-it ( INTR (асп. суф.) )	-il; -j; -ilē ( MULT (асп. суф.) )
-im ( SG<1SG.S ) -ēm; -m; -im; -ām; -am; -ēm; -jm ( NST.1SG.S (только в бессуф. вр.); POSS.1SG; 1SG.S; 1SG )	-m; -ēm; -em; -ēm ( POSS.1SG; 1SG; 1SG.S; PST.1SG.S ) -im ( S<1SG.S )	-im; -im ( NST.SG<1SG.S (только бессуф. вр.); SG<1SG.S ) -im; -ēm; -im; -am; -m; -ām; -ām; -ēm; -jm ( NST.1SG.S (только в бессуф. вр.); POSS.1SG; 1SG; 1SG.S )
-iyān; -agan ( NST.3SG.S )		-iyān; agān; -iyān; -iyān; -āyān; -āyān ( NST.3SG.S )
-jmān; -imān ( SG<1DU.S ) -имен; -ымынь; ман; -ыминь ( 1DU.S; POSS.1DU )	-ēmān; -min ( POSS.1DU; 1DU.S; 1DU )	-ēmān; -ēmān; -amān; -mān; -mān ( POSS.1DU; 1DU; 1DU.S; NST.1DU.S (только в бессуф. вр.) ) -jmān; -imān ( 1SG<DU.S )
-jw ( SG<1PL.S ) -ōw ( SG<1PL.S; 1PL.S; POSS.1PL; 1PL; NST.1PL.S (только в бессуф. вр.); 1PL )	-oβ ( 1PL.S; )	-uγ; -ōγ; -ōγ; -oγ; -ōγ; -ōγ; -ōγ; -ūγ; -ūγ ( NST.1PL.S (только в бессуф. вр.); 1PL; 1PL; SG<1PL.S; POSS.1PL; 1PL.S )

Рис. 8а. Этимологически связанные аффиксы восточных диалектов

-kāl; -yāl ( DU<3SG.S; NST.DU<3SG.S (только в бессуф. вр.); POSS.DU<3SG )	-qet; -yēt ( POSS.DU<3SG )	DU<3SG.S; POSS.DU<3SG; NST.DU<3SG.S (только в бессуф. вр.) )
-i-in; -i-jn ( POSS.PL<3DU; POSS.PL<2PL; POSS.2DU.PL )	-t-ēn; -t-in ( POSS.2DU.PL; POSS.PL<2PL; POSS.PL<3DU )	-ijn; -lin ( POSS.PL<2PL; POSS.2DU.PL; POSS.PL<3DU )
-i-āmān ( POSS.PL<1DU )	-t-amān ( POSS.PL<1DU )	-i-āmān ( POSS.PL<1DU )
-i-ōw ( POSS.PL<1PL )	-t-uβ ( POSS.PL<1PL )	-i-ōγ; -i-ōγ ( POSS.PL<1PL )
-i; -l ( NPST (кроме 3SG, 2SG, 3PL) )	-ētt; -ēt; -ēt; -t; -tt; -t ( NPST; NPST.3SG )	-i; -j ( NPST )
-la ( PL<SG.S (чис. суф. IMP) )	-t-e ( PL<SG.S (чис. суф. IMP) )	-lā; -la ( PL<SG.S (чис. суф. IMP) )
-tə ( CAUS )	tə ( CAUS )	-tə ( CAUS )
-m ( SMF )	-m ( SMF )	-m ( SMF )
-m; -m; -im; -m ( PFV.PTCP )	-ēm; -m ( PFV.PTCP )	-mā; -mā; -ēm; -ēm; -m ( PFV.PTCP )
-mtā ( MONT )	-mtā ( MONT )	-lā; -mtā ( MONT )
-mat ( ORD )	-mēt ( ORD )	-māt ( ORD )
-na; -na ( COM )	-niet; -net ( COM )	-na; -nāt; -nā ( COM )
-ntē ( MULT.INTR (асп. суф.) )	-ntē ( MULT.INTR (асп. суф.) )	-ntē; -ēnt ( MULT.INTR (асп. суф.) )
-sā; -sā; -s ( COLL )	-sē; ( COLL )	-sā; -sā; -s ( COLL )
-t ( TR (асп. суф.) )	-t ( TR (асп. суф.) )	-tē; -t; -tē ( TR (асп. суф.) )
-ti ( PR.INTN )		-ti ( PR.INTN )
-tē; -tē ( PL.S<SG.O (ч. суф. IMP) )		-tē; -tē ( PL.S<SG.O (ч. суф. IMP) )
-tāl ( POSS.3PL )		-tāl; -ētāl; -tāl; -tāl; -tā ( POSS.3PL )
-u ( PASS (кроме 3SG, 3PL) )	-o; -o; j ( PASS (кроме 3SG, 3PL); PASS.2SG )	u ( PASS (кроме 3SG, PL) )
-wāytā ( MONT.DIM (асп. суф.) )		-wāytā ( MONT.DIM (асп. суф.) )
-ŋ ( PROPR )	-ŋ ( PROPR )	-ēŋ ( PROPR )
-ēntā ( NEG )	-ēntā ( NEG )	-ēntim; -ēntā; -ēnt ( NEG; NEG.EX )
-ētē; -ē; -ēt; -ētā; -tē ( SG<3SG.S; NST.SG<3SG.S (только в бессуф. вр.) )		-ēt; -tē; -ē; -ētē; -ētā ( NST.SG<3SG.S (только в бессуф. вр.); SG<3SG.S )

Рис. 8б. Этимологически связанные аффиксы восточных диалектов



-ətən; -ətən; -ətən; -tən ( POSS.3DU; NST.2DU.S (только в бессуф. вр.); POSS.3PL; SG<3DU.S; SG<DU.S (ч. суф. IMP); POSS.2DU; DU.S (ч. суф. IMP); 2DU.S )	jtən; -jttən; -tən; -əttən; -itən ( PL<DU.S (ч. суф. IMP); SG<3DU; SG<3DU.S; SG<DU.S (ч. суф. IMP); DU<DU.S (ч. суф. IMP); 3SG<2DU.S; 3SG<2PL.S; 2DU.S; 2DU )	-ətən; -tən; -ətən ( DU.S (ч. суф. IMP); SG<3DU.S; S<DU.S (ч. суф. IMP); POSS.2DU; NST.2DU.S (только в бессуф. вр.); POSS.2PL; POSS.3DU; 2DU.S )
-əḳətə ( INCH (асп. суф. )		-əḳə; -ḳət; -ḳətə; -əḳə; -əḳətə; -yətə; -əkə; -əkətə; -kə; -kətə ( INCH (асп. суф. )
-əyən ( 3DU.S )	-əyən; -əyyən; -yən; -iyən ( 3DU.S; PST.3DU.S )	-əyən; -kən; -əyən; -əyən; -əyən; -yən; -yən; -yən ( NST.3DU.S (только в бессуф. вр.); 3DU.S; NST.3DU.S )
-yī-āmən ( POSS.DU<1DU )	-yət-amən ( POSS.DU<1DU )	-yəl-āmən ( POSS.DU<1DU )
-yī-ōw ( POSS.DU<1PL )	-yət-uβ ( POSS.DU<1PL ) -yət-əβ ( POSS.DU<1PL )	-yī-ōy ( POSS.DU<1PL )
-yī-ən ( POSS.DU<2SG )		-yī-ən; -əyī-ən ( POSS.DU<2SG )
-yāi; -yal; -kāl; -kāl ( PST3; PST3.3SG.S )		-yāi; -yal; -kāl; -ḳal ( PST3.3SG.S )
-yāl-ām ( POSS.DU<1SG )	-yət-am ( POSS.DU<1SG )	-əkāl-ām; -yī-ām; -əyī-ām; -yī-ām; -əyī-ām; -yāl-ām; -kāl-ām ( POSS.DU<1SG )
-a ( OBL (пад. суф.); K. )	-et ( INSC )	-tə; -ā; -tā; -ā; -ə; -ə ( OBL (пад. суф.); K. )
-a ( REFL )		-a; -ā ( REFL )
-a(t) ( ACC (мест. пад. суф. )	-t ( ACC (мест. пад. суф. )	-tā; -t; -ət ( ACC (мест. пад. суф. )
-ā; -я; -a ( IMP.SG.S (только в SG суб. спр., т.к. ч. суф. нулевой )	-e ( IMP.SG.S (только в SG.S )	-ā; -ā; -a; -ā ( IMP.SG.S (только в SG суб. спр., т.к. ч. суф. нулевой )
-w ( NPST (только в 2SG суб. спр. )		-w ( NPST (только в 2SG суб. спр. )
-wəly; -ууль; -вуль; -выл; -wəl; -вул ( NPST.3SG.S )		-wəl; -wəl ( NPST.3SG.S; NPST (только в 3PL суб. сп. )

Рис. 8в. Этимологически связанные аффиксы восточных диалектов

-əly ( PL<3SG.S )		-āt; -at; -atə; -ə; -ət; -at; -ət; -ā; -tə; -t ( PL )
-lāny; -lāly; -lāl ( POSS.PL<3SG )	-ətt; -tt ( PL<3SG.S )	-lə; -li; -l; -āl; -āl; -əl; -āl; -āl; -jī; -əl; -lā; -l; -l ( 3SG; POSS.3SG )
-əl; -lā; -l; -al; -əl ( 3SG; POSS.3SG )	-et; -t; -ət ( POSS.3SG; 3SG )	-lə; -əl-lə; -l-lə; -l-lə; -l-lə; -əl-lə ( PL<3SG.S; POSS.PL<3SG; NST.PL<3SG.S (только в бессуф. вр. )
-t ( PL )	-t; -ət ( PL )	
-ətax; -тых; -тыг; -тах; -итах ( 2PL.S )		-təy; -təy; -təy ( PLS (ч. суф. пов. накл.); 2PL.S )
-i; -ā; -a; -e; -я ( LAT; . )	-e; -e; -je ( DLAT )	-a; -ā; -ā; -ā ( LAT )
-i; -й; -ы ( IMP (со всеми ч. суб. и об. спр. кроме SG.S )	-j; -j; -i ( IMP )	-i; -j ( IMP (со всеми ч. суб. и об. спр. кроме SG.S); IMP.SG<SG.S (только в SG об. спр., т.к. ч. суф. нулевой )
-il; -ely; -l; -ily; -ily ( 3PL; SG<3PL.S )	-tt ( SG<3PL.S; PST.S<3PL.S; 3PL )	-il; -il; -jī; -l; -il; -il ( 3PL; SG<3PL.S; NST.SG<3PL.S (только в бессуф. вр.); . )
-kan ( DU )	-yən; -qən; -ən ( DU )	-əyən; -kən; -kkən; -yən; -kən; -əyən; -kən ( DU )
-kas; -gac ( PST2; PST2.3SG )		-yās; -yas; -kās; -ḳas; -kas ( PST2.3SG.S; PST2 )
-lāl; -l-ya; -l-ān ( PL<3PL.S; POSS.PL<3PL )	-t-ət ( POSS.PL<3PL )	-lāl; -əllāl ( NST.PL<3PL.S (только в бессуф. вр.); PL<3PL.S ) -lāl; -l-al; -l-əl; -l-āl ( POSS.PL<3PL )
-l-āny ( PL<2PL.S )	-ət-ən ( PL<2PL.S ) -jt-ən; -it-in; -t-ən ( PL<PL.S (ч. суф. IMP); PL<2PL.S; SG<DU.S (ч. суф. IMP); PL<3DU.S; PL<2DU.S; SG<PL.S (ч. суф. IMP )	-il-ən; -əl-in; -əl-jī; -jī-ən ( PL<PL.S (ч. суф. IMP); PL<2DU.S; PL<DU.S (ч. суф. IMP); PL<2PL.S; PL<3DU.S )
-l-īm; -l-ām; -l-ya ( PL<1SG.S; POSS.PL<1SG )	-ət-əm ( PL<1SG.S )	-l-ām; -l-ām ( POSS.PL<1SG; PL<1SG.S )
-l-ya; -l-ān ( POSS.PL<2SG )		-l-ān; -l-ān ( POSS.PL<2SG )

Рис. 8г. Этимологически связанные аффиксы восточных диалектов

аспектуальных форм глагола (MULT, MONT, SMF, TR, CAUS), форм причастий (IPFV.PTCP, PFV.PTCP), форм пассива (PASS) и императива (IMP), аффикс деепричастия (CNV). Перечисленные виды аффиксов частично приведены на рис. 8б.

В подгруппе именных аффиксов общую этимологию обнаруживают посессивные показатели (POSS), падежные (LAT, LOC, OBL, ACC), числовые аффиксы (PL, DU), и формообразующие аффиксы (PROPR, COLL). Частично указанные аффиксы приведены на рис. 8в.

В связи с тем, что проанализированные текстовые данные и, соответственно, сформированные морфологические словари опубликованы либо на латинице, либо на кириллице, то в сводной таблице этимологически связанных аффиксов данные приводятся в смешанном формате (см. рис. 8г.), средствами обеих письменных систем, что несколько затрудняет точную расшифровку и интерпретацию материала. Анализ фонетических корреляций аллофонов одного и того же аффикса может в таких

случаях требовать дальнейшего уточнения. Например, показатель лативного падежа на рис. 8 г-д. в васюганском словаре приводится в виде -и, -а, -ä, -е, -я, а в ваховском как -а, -ä, и др.

### Обсуждение и заключение

Проведённый экспериментальным путём анализ морфологических словарей ваховского, васюганского и сургутского диалектов показывает, что 16 аффиксов в васюганском и ваховском диалектах, связанных этимологически, не засвидетельствованы в сургутском диалекте. В данную выборку аффиксов попали именные падежные показатели (ABL, ALL), маркеры рядов местоимений (PR.INTN, REFL, INDEF), посессивные суффиксы (POSS.3PL, POSS.DU<2SG), глагольные темпоральные маркеры (NST.3SG.S, PST2, PST3, NPST.3SG.S, NPST (только в 2SG суб. спр.), NPST (только в 3PL суб. спр.), аспектуальные суффиксы (MONT.DIM, INCH); числовые суффиксы императива (PL.S<SG.O, PL.S) и повторительный суффикс числительного (REP) (табл. 1).

-л-им; -л-ам; -л-ям ( PL<1SG.S; POSS.PL<1SG )	-et-em ( PL<1SG.S )	-lām; -lām ( POSS.PL<1SG; PL<1SG.S )
-л-ян; -л-ан ( POSS.PL<2SG )		-lān; -lān ( POSS.PL<2SG )
-ин; -ын; -in ( SG<2SG.S; 2DU; 2PL; 3DU )	-in ( 3DU; POSS.3DU )	-n; -ān; -ēn; -an; -ān; -ēn ( 2SG.S; 2SG; NST.2SG.S (только в бессоф. вр.); POSS.2SG )
-ан; -анн; -ань; -н ( POSS.2SG; 2SG.S )	-e; -en ( PST.2SG.S; 2SG; 2SG.S; POSS.2SG )	-in; -in ( ; 2PL; 3DU; 2DU; SG<2SG.S )
-лихь ( ABS )		-lōy; -lōy; -løy; -løy; -lūy ( ABS )
-ман; -мин ( CVB )	-min ( CNV )	-min; -mīn ( CVB )
-н ( COND )	-en ( COND )	-n ( COND (усл. накл.) )
-нь; -ән; -ан; -ань; -нә; -на ( LOC )	-nә ( LOC )	-nә; -nә; -ān; -ān; -ēn; -ēn; -n; -na; -nā; -nā; -ni; -nә ( LOC )
-ох; -ог; -охь ( ABL )		-oy; -ūy; -uy; -ōy; -ēy; -ōy; -ūy; -ōy; -ōy ( ABL )
-нә; -на; -рә ( REP )		-rә; -pa; -pā ( REP )
-на ( ALL )		-pa; -pā; -pā; -āpa; -āp; -āpā; -āpā; -apa; -pā ( ALL )
-сь; -сы; -с; -са ( PST1; PST1.3SG )	-es; -es; -s ( PST1; PST1.3SG )	-ēs; -ēs; -s ( PST1; PST1.3SG.S )
-т; -нт ( IPFV.PTCP )		-tā; -ta ( INF )
-тә; -та ( INF )	-t; -tә; -ettә ( IPFV.PTCP )	-t; -tā; -tә; -tә; -et ( ; IPFV.PTCP; IPFV.PTCP )
-ты ( PTCL )	-t; -ta ( PTCL )	-ti; -ti; -t ( PTCL )
-х; -г; -к ( EP )	-j; -y ( EP )	-n; -j; -k; -y ( EP )
-ыг; -г; -хь ( K.; TRL )	-yә; -ә; -kә ( TRL )	-y; -ēy; -kkә; -ay; -kә ( TRL )
-әт; -әт; -t ( NST.3PLS (только в бессоф.вр.); NST.3PLS (только в бессоф.вр.) )	-t; -әt ( 3PLS; NST.3PLS (только в бессоф.вр.) )	-әt; -әt; -t ( 3PLS; NST.3PLS (только в бессоф.вр.) )
i ( ATTR )	i ( ATTR )	-j ( ATTR )
ḱem ( INDEF )		-ḱem ( INDEF )
yәl ( MULT.DIM )	-yәt ( MULT.DIM (асп. суф.) )	-yәl ( MULT.DIM (асп. суф.) )

Рис. 8д. Этимологически связанные аффиксы восточных диалектов

**Аффиксы ваховского и васюганского диалектов,  
которые не имеют этимологий в сургутском диалекте**

	Морфологический словарь васюганского диалекта	Морфологический словарь юганского говора сургутского диалекта	Морфологический словарь ваховского диалекта
1.	-iḡän; -аган (NST.3SG.S)	–	-iḡän; -äḡän; -iḡən; -aḡən (NST.3SG.S)
2.	-ti (PR.INTN)	–	-ti (PR.INTN)
3.	-tə; -tə̄ (PL.S<SG.O (ч. суф. IMP))	–	-tə; -tə̄ (PL.S<SG.O (ч. суф. IMP))
4.	-təl (POSS.3PL)	–	-təl; -ätäl; -täl; -tla (POSS.3PL)
5.	-wäytä (MONT.DIM (асп. суф.))	–	-wäytä (MONT.DIM (асп. суф.))
6.	-əḡətə (INCH (асп. суф.))	–	-əḡə; -ḡətə; -kätä; -əḡä; -äkätä; -äkä; -əḡətə (INCH (асп. суф.))
7.	-ḡl-än (POSS.DU<2SG)	–	-ḡl-än; -əḡl-än (POSS.DU<2SG)
8.	-ḡäl; -ḡal; -käl'; -käl (PST3; PST3.3SG.S)	–	-ḡäl; -ḡal; -käl; -ḡal PST3; PST3.3SG.S)
9.	-a; -ä (REFL)	–	-a; -ä (REFL)
10.	-w (NPST (только в 2SG суб. спр.))	–	-w (NPST (только в 2SG суб. спр.))
11.	-wäl; -wəl (NPST.3SG.S; NPST (только в 3PL суб. спр.))	–	-wäl; -wəl (NPST.3SG.S; NPST (только в 3PL суб. спр.))
12.	-təḡ; -təḡ̄ (PL.S (ч. суф. пов. накл.); 2PL.S)	–	-təḡ; -təḡ̄ (PL.S (ч. суф. пов. накл.); 2PL.S)
13.	-ḡäs; -ḡas; -käs; -ḡas; -kas (PST2.3SG.S; PST2)	–	-ḡäs; -ḡas; -käs; -ḡas; -kas (PST2.3SG.S; PST2)
14.	-oḡ (ABL)	–	-oḡ; -öḡ; -äḡ; -öḡ; (ABL)
15.	-pa (REP)	–	-pä; -pa (REP)
16.	-pa (ALL)	–	-pa; -pä; -äpä; -apa (ALL)
17.	-ḡəm (INDEF)	–	-ḡəm (INDEF)

Отдельную группу аффиксов образуют морфемы, которые являются характерной особенностью только одного из анализируемых диалектов и не обнаруживают этимологических

связей с другими, всего 14 аффиксов, из их числа 12 в сургутском, по одному – в васюганском и ваховском. Автономные морфемы восточных диалектов представлены на рис. 9.

1. Морфологический словарь восточновасюганского хантыйского языка	2. Морфологический словарь восточноюганского хантыйского языка	3. Морфологический словарь восточноваховского хантыйского языка по данным полевых работ в п. Корлики
-əḡiŋ ( MONT.INTR (асп. суф.))		
	-nəm ( ALL )	
	-tamin ( SG<1DU.S )	
	-ətəβ ( SG<1PL.S )	
	-e ( IMP.SG<SG.S )	
	-təḡ ( PTCP.NEG )	
	-tiŋki ( DIM )	
	-əpti ( EXPL (Honti 1985: 65) )	
	-i; -ji ( ABL )	
	-ətəḡ; -təḡ ( SG<3SG.S; PST.SG<3SG.S )	
	-ə ( PR.INTN )	
	-ḡət-a ( POSS.DU<2SG )	
		-ətəl; -əl; -təl; -ätäl ( DISTR (под. суф.))

Рис. 9. Аффиксы, которые встречаются только в одном из восточных диалектов



Самое большое количество автономных аффиксов было выявлено в сургутском диалекте – 12 морфем, что подтверждает его морфологическую отдалённость от восточных ваховского и васюганского. В данную группу аффиксов, не связанных этимологически с аффиксами других восточных диалектов, входят именные падежные показатели *-nvm* (ALL), *-i*; *-ji* (ABL) и *apti* (EXPL), посессивные суффиксы *-yət-a* (POSS.DU<2SG), *-it* (POSS.3PL), лично-числовые суффиксы глаголов объектного спряжения *-tamin* (SG<1DU.S), *-atəβ* (SG<1PL.S), *-atəy*; *-təy* (S<3SG.S), числовой суффикс императива *-e* (IMP.SG<SG.S), суффикс причастия отрицания *-ləy* (PTCP.NEG), который омонимичен суффиксу абессивного падежа во всех восточных диалектах, а также суффикс образования разряда лично-усилительных местоимений *-ə* (PR.INTN).

В ваховском и васюганском диалектах было выявлено по одному автономному аффиксу, которые по данным автоматического анализа по трём словарям, не связаны этимологически с другими восточными диалектами. В ваховском диалекте это суффикс разделительного падежа *-əl*; *-ätäl*; *-tal*; *-ətal* (DISTR), а в васюганском – аспектуальный суффикс *-əyiŋ* (MONT.INTR).

На основе полученных данных можно сделать только предварительные выводы, поскольку корпуса по трём языковым системам имеют неравный по количеству объём языкового материала, размещённого на платформе ЛингвоДок. Ваховский корпус имеет самый весомый объём, васюганский и сургутский значительно уступают. Восточнохантыйские тексты на платформе ЛингвоДок не относятся к одному времени, поэтому сравнивать их сложно, например, на реке Юган, в начале XX в. ещё употреблялось прошедшее время с суффиксом *-s*, которого уже нет нигде в сегодняшнем сургутском диалекте. При уравнивании объёмов и дополнении современными языковыми данными васюганского и сургутского корпусов подсчёт данных будет более точным и достовер-

ным и позволит установить степень близости восточных идиомов хантыйского языка с меньшей погрешностью. Погрешность в подсчёте данных также может быть вызвана расхождением в разметке словоформ у авторов разных корпусов: детализация разметки (некоторые аффиксы не выделяются), разное значение глосс, ошибки при разметке из-за человеческого фактора, нечитаемость машиной некоторых символов и др. Все эти факторы позволяют нам предоставить только промежуточный результат вычисления степени близости восточных идиомов, который в будущем будет пересмотрен с получением новых данных.

На основании полученных данных подсчёта морфологических показателей трёх диалектов можно констатировать незначительное расхождение в области морфологии между ваховским и васюганским диалектами и степень близости их морфологических систем в районе 99 %. Указанные промежуточные результаты позволяют поддержать точку зрения о том, что данные языковые системы формируют единое диалектное пространство и представляют собой один вах-васюганский диалект с двумя разными говорами: ваховский и васюганский. Окончательные выводы в отношении степени морфологической близости ваховского и васюганского хантыйского будет возможно сделать лишь при расширении текстового корпусов на васюганском хантыйском языке, который в настоящее время ограничен лишь материалами текста XIX в.

Сургутский диалект, который в настоящем исследовании представлен юганским говором, имеет отличия в области морфологии в 14–15% случаях, поэтому может быть достоверно дистанцирован от ваховско-васюганского диалектного единства.

В связи с несбалансированностью объёма доступных для цифрового анализа корпусов, полученные данные носят промежуточный результат вычисления степени близости восточных идиомов, который в будущем будет пересмотрен с получением новых данных.

#### Список источников и литературы

1. Норманская Ю. В. Возможности «Лингводока» для выявления степени близости языков и диалектов и описания графико-фонетических особенностей // Информационные технологии в гуманитарных исследованиях: материалы Междунар. науч.-практ. конф. Красноярск, 25–28 сентября 2023 г. / отв. ред. М. А. Лаптева. Красноярск: Сиб. федер. ун-т, 2023. С. 251–260. URL: <https://bik.sfu-kras.ru/ft/LIB2/ELIB/b71/free/i-419654163.pdf> (дата обращения: 10.09.2024).
2. Норманская Ю. В. Графико-фонетические различия в хантыйских диалектах по данным источников XVIII в. // Вестник угроведения. 2022. Т. 12. № 1. С. 84–93.



3. Норманская Ю. В. Первые тюркские кириллические книги на платформе ЛингвоДок // Родные языки и культуры в современном изменяющемся мире. 2022. № 1. URL: <https://rodnoyyazyk.ru/index.php/journal/article/view/9> (дата обращения: 10.10.2024 г.).
4. Программная система LingvoDoc и возможности, которые она предлагает для документирования и анализа обско-угорских языков / А. И. Аветисян, И. Б. Белобородов, О. Д. Борисенко, Ю. В. Норманская // Доклады Российской академии наук. Математика, информатика, процессы управления. 2022. Т. 504. № 1. С. 60–82.
5. Соловар В. Н., Нахрacheва Г. Л., Шиянова А. А. Диалекты хантыйского языка. Ханты-Мансийск: Ижевск: Принт-2, 2016. 349 с.
6. Терёшкин Н. И. Очерки диалектов хантыйского языка (ваховский диалект). Л.: Наука, 1961. 204 с.
7. Терёшкин Н. И. Словарь восточнохантыйских диалектов. Л.: Наука, 1981. 544 с.
8. Хайду П. Уральские языки и народы. М.: Прогресс, 1985, 425 с.
9. Чепреги М. Сургутский диалект хантыйского языка. Ханты-Мансийск: Печатный мир г. Ханты-Мансийск, 2017. 275 с.
10. Collinder B. Comparative Grammar of the Uralic Languages. Stockholm: Almqvist & Wiksell, 1960. 425 p.
11. Honti L. Chrestomathia Ostiastica. Budapest: Tankonyvkiado, 1986. 432 p.
12. Honti L. Die ostjakische Sprache // The Uralic Languages. Description, History and Foreign Influences. Handbuch der Orientalistik. / Ed. D. Sinor. Leiden; New York; København; Köln: Brill, 1988. Pp. 172–196.
13. Karjalainen K. F. Grammatikalische Aufzeichnungen aus ostjakischen Mundarten. Bearbeitet und herausgegeben von Edith Vertes. MSFOu, 128. Helsinki: Finno-Ugrian Society, 1964. 341 p.
14. Sosa S. Functions of morphosyntactic alternations, and information flow in Surgut Khanty Discourse. Doctor of Philology Thesis, Helsinki, 2017. 254 p.
15. Steinitz W. Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache. Berlin: Akademie – Verlag, 1966–1993. Vol. 1–15. 1504 p.
16. Steinitz W. Ostjakologische Arbeiten. Beiträge zur Sprachwissenschaft und Ethnographie, Für die Herausgabe wissenschaftlich bearbeitet von Ewald Lang, Gert Sauer und Renate Steinitz. Budapest: Akademiai Kiado, 1980. Band IV. 498 p.
17. Vértés E. Einwandfrei Aufhebbare Widersprüche zweier Mitteilungen über die Vach-Mundart // Linguistica Uralica. 1965. Vol. 1 (3). Pp. 213–217.
18. Vorobeva V., I. Novitskaya. Inflectional morphology of nouns in Eastern Khanty (Vakh, Vasyugan, Surgut, Salym) // Ural-Altai Studies. 2020. № 38. Pp. 33–70.
19. Vorobeva V. V., Novitskaya I. V. Functional and semantic analysis of the case markers in Vakh Khanty (a case study of the latest field data on LingvoDoc) // Труды института системного программирования РАН. 2024. Т. 36. Вып. 2. Pp. 199–210. DOI: 10.15514/ISPRAS-2023-35(2)-15.

## References

1. Normanskaya Yu. V. *Vozmozhnosti “LingvoDoka” dlja vyjavlenija stepeni blizosti jazykov i dialektov i opisanija grafiko-foneticheskikh osobennostej* [Possibilities of “LingvoDoc” for identifying the degree of closeness of languages and dialects and describing graphic and phonetic features]. *Informacionnye tekhnologii v gumanitarnyh issledovaniyah: materialy Mezhdunar. nauch.-prakt. konf. Krasnoyarsk, 25–28 sentyabrya 2023 g.* [Information technologies in humanitarian research: proceedings of the International scientific and practical conference (Krasnoyarsk, September 25–28, 2023)]. Ed. M. A. Lapteva. Krasnoyarsk: Sibirskij federalnyj universitet Publ., 2023. Pp. 251–260. Available at: <https://bik.sfu-kras.ru/ft/LIB2/ELIB/b71/free/i-419654163.pdf> (accessed September, 2024). DOI: 10.5281/zenodo.7385533. (In Russian)
2. Normanskaya Yu. V. *Grafiko-foneticheskie razlichiya v hantyskikh dialektah po dannym istochnikov XVIII v.* [Graphic and phonetic differences in the Khanty dialects according to sources of the XVIII century]. *Vestnik ugrovedenia* [Bulletin of Ugric Studies], 2022, no. 12 (1), pp. 84–93. (In Russian)
3. Normanskaya Yu. V. *Pervye tjurkskie kirillicheskie knigi na platforme LingvoDok* [First Turkic cyrillic books on the LingvoDoc platform]. *Rodnye jazyki i kul'tury v sovremennom izmenyayushchemsya mire* [Native languages and cultures in the modern changing world], 2022, no. 1. Available at: <https://rodnoyyazyk.ru/index.php/journal/article/view/9> (accessed October 10, 2024). (In Russian)
4. Avetisyan A. I., Beloborodov I. B., Borisenko O. D., Normanskaya Yu. V. *Programmaja sistema LingvoDOC i vozmozhnosti, kotorye ona predoagaet dlja dokumentirovanija i analiza obsko-ugorskich jazykov* [The LingvoDOC software system and the possibilities it offers for documenting and analyzing the Ob-Ugric languages]. *Doklady Rossijskoj akademii nauk. Matematika, informatika, processy upravleniya* [Reports of the Russian Academy of Sciences. Mathematics, informatics, control processes], 2022, no. 504 (1), pp. 60–82. (In Russian)
5. Solovar V. N., Nakhacheva G. L., Shiyanova A. A. *Dialekty hantyskogo jazyka*. [Dialects of the Khanty language]. Khanty-Mansiysk; Izhevsk: Print-2 Publ., 2016. 349 p. (In Russian)
6. Tereshkin N. I. *Oчерки dialektov hantyskogo jazyka (vahovskij dialekt)* [Essays on dialects of the Khanty language (Vakh dialect)]. Leningrad: Nauka Publ., 1961. 204 p. (In Russian)
7. Tereshkin N. I. *Slovar' vostochnohantyskikh dialektov* [Dictionary of the Eastern Khanty dialects]. Leningrad: Nauka Publ., 1981. 544 p. (In Russian)

8. Hajdú P. *Ural'skie jazyki i narody* [The Uralic languages and peoples]. Moscow: Progress Publ., 1985. 425 p. (In Russian)
9. Chepregi M. *Surgutskij dialekt hantyjskogo jazyka* [Surgut dialect of the Khanty language]. Khanty-Mansiysk: Pechatnyj mir g. Khanty-Mansijsk Publ., 2017. 275 p. (In Russian)
10. Collinder B. *Comparative Grammar of the Uralic Languages*. Stockholm: Almqvist & Wiksell, 1960. 425 p. (In English)
11. Honti L. *Chrestomathia Ostiacica*. Budapest: Tankonyvkiado, 1986. 432 p. (In Hungarian)
12. Honti L. Die ostjakische Sprache. *The Uralic Languages. Description, History and Foreign Influences. Handbuch der Orientalistik*. Ed. D. Sinor. Leiden; New York; København; Köln: Brill, 1988. Pp. 172–196. (In German)
13. Karjalainen K. F. *Grammatikalische Aufzeichnungen aus Ostjakischen Mundarten*. Bearbeitet und herausgegeben von Edith Vertes. Helsinki: Suomalais-Ugrilaisen Seura, 1964. 234 p. (In German)
14. Sosa S. *Functions of morphosyntactic alternations, and information flow in Surgut Khanty Discourse*. Helsinki, 2017. 254 p. (In English)
15. Steinitz W. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin: Akademie – Verlag, 1966–1993. Vol. 1–15. (In Khanty, German)
16. Steinitz W. *Ostjakologische Arbeiten. Beiträge zur Sprachwissenschaft und Ethnographie*. Für die Herausgabe wissenschaftlich bearbeitet von Ewald Lang, Gert Sauer und Renate Steinitz. Budapest: Akademiai Kiado, 1980. Band IV. 498 p. (In German)
17. Vértés E. Einwandfrei Aufhebbare Widersprüche zweier Mitteilungen über die Vach-Mundart. *Linguistica Uralica*, 1965, no. 1 (3), pp. 213–217. (In German)
18. Vorobeva V., Novitskaya I. Inflectional morphology of nouns in Eastern Khanty (Vakh, Vasyugan, Surgut, Salym). *Ural-Altai Studies*, 2020, no. 38, pp. 33–70. (In English)
19. Vorobeva V. V., Novitskaya I. V. Functional and semantic analysis of the case markers in Vakh Khanty (a case study of the latest field data on LingvoDoc). *Trudy instituta sistemnogo programirovaniya RAN* [Proceedings of the Institute for System Programming of the RAS], 2024, no. 36 (2), pp. 199–210. DOI: 10.15514/ISPRAS-2023-35(2)-15. (In English)

#### ИНФОРМАЦИЯ ОБ АВТОРАХ

**Воробьева Виктория Владимировна**, научный сотрудник, Институт системного программирования им. В. П. Иванникова РАН (109004, Российская Федерация, г. Москва, ул. Александра Солженицына, д. 25); доцент, Национальный исследовательский Томский политехнический университет; кандидат филологических наук.

vorobeva@tpu.ru

ORCID ID: 0000-0002-8729-0375

**Новицкая Ирина Владимировна**, профессор, Национальный исследовательский Томский государственный университет (634050, Российская Федерация, г. Томск, пр. Ленина, д. 36), доктор филологических наук.

irno2012@yandex.ru

ORCID ID: 0000-0003-1559-8810

#### ABOUT THE AUTHORS

**Vorobyova Viktoria Vladimirovna**, Researcher, Ivannikov Institute for System Programming of the Russian Academy of Sciences (109004, Russian Federation, Moscow, Aleksander Solzhenitsyn Str., 25); Associate Professor, National Research Tomsk State University (634050, Russian Federation, Tomsk, Lenin Avenue, 30), Candidate of Philological Sciences.

vorobeva@tpu.ru

ORCID ID: 0000-0002-8729-0375

**Novitskaya Irina Vladimirovna**, Professor, National Research Tomsk State University (634050, Russian Federation, Tomsk, Lenin Avenue, 30), Doctor of Philological Sciences.

irno2012@yandex.ru

ORCID ID: 0000-0003-1559-8810