

УДК 398.4
ББК 82.3(3)

Мифологические рассказы русского старожильческого населения Монголии в проекции цифровой фольклористики

София Андреевна Федотова

(Институт мировой литературы им. А. М. Горького РАН:
Российская Федерация, 121069, г. Москва, Поварская ул, д. 25а, стр. 1)

Аннотация. В статье исследуются закономерности в соотношении сюжетно-структурных особенностей текстов с пространственным распределением русского старожильческого населения Монголии. Источниками для исследования послужили транскрибированные мифологические рассказы (былички), собранные в 2018–2024 гг. в Улан-Баторе и Дархане. Работа является первым опытом изучения фольклорных текстов, записанных в ходе этнографических интервью с русскими старожилами Монголии, цифровыми методами. Выбор быличек в качестве объекта исследования обусловлен их относительной сохранностью по сравнению с другими жанрами фольклора в традиции русских Монголии. С опорой на отечественные и зарубежные исследования сформирован алгоритм ручной разметки особенностей мифологического рассказа как жанра фольклора. Поиск закономерностей среди размеченных категориальных признаков быличек осуществлялся с помощью метода кластерного анализа *k-modes*, статистического анализа посредством точного критерия Фишера и визуализации данных методом множественного анализа соответствий (MCA). Несмотря на то что выделенные алгоритмом кластеры текстов не соотнеслись с геолокацией происходящих в быличках событий, обнаружены статистически значимые зависимости между дистанцией рассказа («близко»/«далеко» от места записи) и такими признаками, как время действия и пол главного героя. Кроме того, результаты анализа подтвердили наблюдение об угасании фольклорной традиции в регионе.

Ключевые слова: русские старожилы Монголии, мифологические рассказы, цифровые методы в гуманитарных науках, категориальные признаки.

Исследование выполнено при поддержке РНФ. Проект № 23-18-00478 «Русские Монголии. Комплексное исследование культуры в иноэтническом окружении (фольклор, обрядовые традиции, языки)».

Дата поступления статьи: 16 сентября 2025 г.

Дата публикации: 25 декабря 2025 г.

Для цитирования: Федотова С. А. Мифологические рассказы русского старожильческого населения Монголии в проекции цифровой фольклористики // Традиционная культура. 2025. Т. 26. № 4. С. 87–98.

DOI: <https://doi.org/10.26158/TK.2025.26.4.007>

Фольклорная традиция русских старожилов Монголии изучена мало. Фактически она начала фиксироваться только с 2018 г. В. Л. Кляусом, и эта работа была продолжена в рамках проекта «Русские Монголии. Комплексное исследование культуры в иноэтническом окружении (фольклор, обрядовые традиции, язык)» в 2023–2025 гг., поддержанного под его руководством Российской научным фондом. Тем более для ее анализа никогда не применялись методы цифровой фольклористики. Настоящее исследование ставит целью выявить закономерности в распределении сюжетно-структурных особенностей фольклорных нарративов цифровыми методами и проверить гипотезу о соотнесении этих закономерностей с локализацией событий, отражающей историческое расселение русских старожилов Монголии, которое в 1920–1940-е гг. было в основном сконцентрировано в деревнях Селенгинского аймака, в 1950–1980-е — на станциях Улан-Баторской железной дороги¹.

Материалом для анализа послужили мифологические рассказы, записанные в 2018–2024 гг. от 16 русских старожилов Монголии, проживающих в городах Улан-Батор и Дархан. Выбор быличек в качестве объекта исследования связан с особенностями собранного в экспедициях в Монголию фольклорного материала. В целом среди этой группы, как отмечает В. Л. Кляус, наблюдается разрушение фольклорной традиции: в частности, некоторые жанры фольклора почти не фиксируются (песни, заговоры) или не фиксируются вовсе (сказки) [Кляус, Махова 2025, 24]. Мифологические рассказы — это единственный жанр, по которому удалось собрать количество материала, достаточное для изучения цифровыми методами. Но и он находится в упадке: в сравнении с родственными фольклорными традициями из Забайкалья отмечается уменьшение количества сюжетов, упрощение структуры быличек, унификация языка. Причины этого лежат в истории данной этнической группы (см.: [Сундуева 2024]).

¹ Автор благодарит Б. В. Орехова, О. В. Алиеву и других преподавателей магистерской программы «Цифровые методы в гуманитарных науках» НИУ ВШЭ за советы и помощь в работе над этим исследованием.

В советское время число русских старожилов в Монголии достигало нескользких тысяч. Согласно переписи населения страны 2020 г., их количество снизилось до 1600 человек [Биткеева, Исаков 2024, 872]. Подавляющее большинство живет в Улан-Баторе, столице Монголии; отдельные семьи до сих пор живут в Дархане и Дзунхаре. Но память о детстве и юношестве, проведенном в ныне исчезнувших с карты русских деревнях, передается через устные рассказы.

Столь интересный пример диаспоральной группы, десятки лет существующей в иноэтническом окружении и сохраняющей свои особенности, несомненно, привлекает внимание исследователей. О первых дореволюционных контактах Монголии и России и о первых поселениях пишут Н. Е. Единархова [Единархова 1999] и Е. М. Даревская [Даревская 1994]; «местнорусских» в социалистической Монголии, их политический статус, идентичность изучает А. В. Михалев [Михалев 2011]; о вопросах исторической памяти и травмы рассуждают А. Н. Биткеева и А. В. Исаков [Биткеева, Исаков 2024]; фольклорные сюжеты рассматривает В. Л. Кляус [Кляус 2024]; к социолингвистическим аспектам обращается Д. Б. Сундуева [Сундуева 2011]. Данная работа встраивается в междисциплинарное поле исследований этой этнической группы и является попыткой взглянуть на материалы, записанные от русских старожилов Монголии, через призму цифровых методов.

ЦИФРОВЫЕ МЕТОДЫ В ФОЛЬКЛОРИСТИКЕ

Изучение фольклора цифровыми методами — это новая, активно развивающаяся область гуманитарных исследований, которой уделяют внимание преимущественно в зарубежной науке. Все еще не устоялось однозначное название субдисциплины: зачастую в синонимичных контекстах употребляют сочетания *Digital Folkloristics* («цифровая фольклористика») и *Computational Folkloristics* («вычислительная фольклористика»). Однако наблюдается тенденция, что научные статьи в области «цифровой фольклористики»

в основном посвящены исследованиям сетевого фольклора, а «вычислительная фольклористика» — это преимущественно изучение традиционных фольклорных текстов цифровыми методами.

Основной задачей фольклористики, поддающейся решению цифровыми методами, является классификация текстов. Это становится особо актуальным в свете того, что кратно увеличивается объем доступного оцифрованного фольклорного материала, который из-за своего разнообразия уже не может быть систематизирован, например, по классическому указателю сюжетов Аарне — Томпсона — Утера. Следовательно, необходимо вдумчиво организовывать цифровое хранение текстов и метаданных, а также настраивать систему поиска, выдающую релевантные результаты по запросу пользователей.

Хранению коллекций в виде баз данных посвящено большое количество статей и выступлений на тематических конференциях: можно утверждать, что каждый институт сохранения культурного наследия решает похожие задачи по-своему из-за региональной специфики материала. Например, группа авторов, занимавшихся разработкой базы данных нидерландской сказочной прозы (*Dutch Folktale Database*), нашла способ автоматически присваивать сказкам такие метаданные, как язык (диалект), перечень именованных сущностей, ключевые слова и даже поджанры (сказка, шутка и т. п.), с помощью обучения классификаторов [Meder et al. 2016].

Отечественная наука тоже заинтересована в оцифровке фольклорных материалов и публикаций в виде баз данных. Фольклорные архивы из экспедиций с интеграцией геоинформационных систем, лингвистического поиска реализованы на факультете гуманитарных наук НИУ ВШЭ [Смирнова 2021] и в Лаборатории компьютерной филологии БашГУ [Галлямов, Орехов 2016]. На сайте ЦТСФ РГГУ в электронном виде представлены базы данных различных жанров фольклора не только русской фольклорной традиции, но и других, там же хранятся каталоги и указатели, сборники статей по релевантным темам.

Оцифровка и систематизация фольклорных коллекций дает возможность проводить анализ корпусов методами обработки естественного языка (NLP). В статье исследователей из Германии [Lindemann et al. 2019] предлагается алгоритм обучения классификатора выделять место действия в сказках (один из элементов, влияющий на отнесение сюжета к тому или иному номеру в ATU).

Фольклористы применяют цифровые методы и в исследованиях других жанров. Например, исследование песен можно проводить с использованием тематического моделирования: алгоритмы различают темы в текстах и позволяют визуализировать семантическую структуру корпуса. В кейсе с изучением словенских фольклорных песен [Strle, Marolt 2016] на основе тематического моделирования проведена иерархическая кластеризация, доказавшая, что тематическая структура, выделенная LDA, соответствует жанровой структуре, принятой в фольклористике.

Методы поиска кластеров в данных «без учителя» тоже применяются в вычислительной фольклористике. Результативны как филогенетические методы, разбивающие на группы матрицы с закодированными признаками с целью найти закономерности распространения мотивов [Ross, Atkinson 2016] или найти различия в двух схожих сюжетах, бытующих в разных частях света [Tehrani 2013], так и методы машинного обучения (упомянутая выше иерархическая кластеризация [Strle, Marolt 2016], кластеризация эмбеддингов посредством k-средних [Ning, Chen 2023]). Их преимущество заключается в том, что количество документов для анализа может быть сравнительно небольшим, в то время как почти все вышеперечисленные методы (классификация, тематическое моделирование и т. д.) требуют крупные корпуса фольклорных текстов (в среднем около двух тысяч документов и более) для получения релевантного результата.

Обобщая вышеизложенное, можно предположить, что именно метод кластеризации, который заключается в автоматическом разбиении данных на группы, и другие статистические методы позволяют выявить скрытые закономерности в распределении текстов.

ПОДГОТОВКА ДАННЫХ

Данными для исследования выступают 115 транскрибированных быличек русских старожилов Монголии [ФК ИМЛИ]. Тексты сопровождаются следующими метаданными: полное имя респондента, год записи интервью, место записи и место событий, а также сведения о респондентах: место рождения, местоположение школы, год рождения, пол. В результате сформирована полная база сведений о факторах, предположительно соотносимых с закономерностями распределения текстов, выделенными статистическими алгоритмами.

Целесообразно обратиться к методу ручной разметки категориальных признаков, отражающих особенности быличек как жанра, чтобы найти вероятные закономерности в том, как место событий и его удаленность от рассказчика могут влиять на структурную целостность текста. В результате был определен 21 признак, важный для построения сюжета, оформления повествования, убеждения слушателя в правдивости былички.

Первые семь признаков непосредственно являются элементами сюжетного повествования, складывающегося из взаимодействия главного героя, второстепенного героя и сверхъестественной силы.

1. Количество главных героев² (далее — ГГ): один/два/больше.

2. Пол ГГ: мужской/женский/мужской и женский.

3. Этническая принадлежность ГГ: русский из Монголии/россиянин/китаец/монгол/несколько.

4. Количество второстепенных героев³: ноль/один/два/больше.

5–6. Пол и этническая принадлежность второстепенных героев: аналогично пунктам 3, 4.

7. Природа сверхъестественной силы: хозяева места/люди/животные/покойники/современные персонажи.

В исследовательских работах о быличках в качестве основополагающих элементов рассказа обычно выделяют «объект» (соответствующий понятию «сверхъестественной силы» в моей разметке) и «субъект» (например: [Шумов 1993, 8]). В статье используются иные понятия для обозначения схожих явлений, чтобы разделить главных и второстепенных героев по признаку активности их участия в событиях. Хотя второстепенные герои по моему определению пассивны и не влияют на сюжет, само их присутствие делает описание более насыщенным, что должно усиливать веру в «достоверность» былички.

Следующая группа признаков — обобщенные сведения о сюжетах-мотивах мифологических рассказов.

8. Собственно сюжет-мотив: 19 групп, в их числе: чудится в доме, колдовка оборачивается в животное, умершие приходят к живым и т. д. Группировка отчасти основана на указателях сюжетов-мотивов быличек, однако не совпадает с ней из-за того, что имеющаяся классификация не подходит к видоизмененной традиции русских старожилов Монголии.

9. Фольклорная традиция: русская/монгольская/китайская. Этот пункт отражает разнообразие фольклорного репертуара русских старожилов: в их традиции «смешиваются» русские сюжеты, принесенные предками респондентов из Забайкалья, и более близкие и актуальные для них сюжеты из легендарной прозы соседей-монголов или китайцев.

² «Главным героем» считается активный участник события либо человек, обладающий сверхъестественными способностями (следовательно, он становится сверхъестественной силой, однако в подобных рассказах о людях заметно смещается фокус с переживаний героев от встречи с неизведанным на интерес к заонарским или колдовским манипуляциям).

³ «Второстепенным героем» считается пассивный участник события.

10. Место действия: неясно/дорога/ многоквартирный дом / частный дом и пространство вокруг / общественные здания (школа, больница) / общественное пространство (улица, кладбище) / природа / иной мир / несколько — обозначает пространственный компонент сюжета, не привязанный непосредственно к локализации (а именно частные дома встречаются и в городе (Улан-Баторе), и в деревнях (все остальные места событий)).

Отметим особенности строения текста, сигнализирующие о сохранности традиции:

11. От чьего лица ведется повествование с указанием источника рассказа: от первого — про себя/от первого — про третье лицо/от третьего лица — близкий родственник/от третьего — все остальные. Этот пункт демонстрирует организацию повествования с точки зрения того, сколько ступеней нарратии обнаружено в тексте [Черванёва 2017, 12], а также отражает отмеченное Э. В. Померанцевой разделение мифологических рассказов на меморат и фабулат (на этом основании она проводила границу между быличкой и бывальщиной [Померанцева 1975, 6], однако в этом исследовании вслед за К. Э. Шумовым и другими фольклористами различие игнорируется).

12. Есть ли в тексте прямая речь или диалог: нет/есть прямая речь/есть диалог. Наличие подобных форм в организации нарратива демонстрирует вовлеченность рассказчика.

Следующий комплекс признаков объединяет в себе виды утверждения достоверности рассказа — одного из ключевых элементов композиции былички [Зиновьев 2019, 284].

13. Ссылка на конкретный источник информации: нет/есть. Под «конкретным» подразумевается упоминание ФИО источника, степени родства или знакомства либо повествование от первого лица.

14. Есть указание на место действия: нет/есть. Виды локализации следующие: название города или конкретного района, расстояние от населенного пункта («...а

там осталось 15 километров до этого, до Ховда...» (Зап. от Марины Геннадьевны Таракановской, 1963 г. р., Улан-Батор, Монголия. Соб. В. Л. Кляус. 2018 г.) [ФК ИМЛИ]), описание конкретного месторасположения дома («*А соседним домом, там у бабушки маленки дом был, а соседний домом он...*» (Зап. от Тамары Пугачёвой, 1951 г. р., Улан-Батор, Монголия. Соб. В. Л. Кляус, Д. Б. Сундуева. 2024 г.) [ФК ИМЛИ]), употребление наречия «здесь» как маркера дома, упоминание общеизвестных социальных объектов (больница, русский садик).

15. Тип хронологической отсылки: нет/ возраст одного из героев/событие жизненного цикла (например, похороны)/год или десятилетие/событие календарного цикла (например, праздник).

16. Время действия: день/ночь/неясно/утро/вечер. Это важно не только с точки зрения обозначения достоверности, но и с точки зрения следования традиционным устоям жанра, для которого характерны преимущественно «ночные» события [Померанцева 1975, 23].

17. Прямое утверждение о достоверности: нет/есть.

18. Реалистичная детализация рассказа: нет/есть. Под этим пунктом понимается упоминание подробностей, которые создают не влияющий на сюжет антураж, в частности, описание пейзажа или обстановки в доме, ощущений персонажей, подробностей смерти и т. д. Обычно выражается эпитетами или сравнительными оборотами.

Следующие пункты разметки тоже связаны с компонентами строения былички.

19. Упоминание поверья либо описание ритуала или обычая: нет/есть.

20. Тональность концовки: негативная/нейтральная/позитивная. Определяется согласно тому, нанесла ли сверхъестественная сила вред герою (негативная) или, наоборот, помогла (позитивная), или история повествует о простом взаимодействии, или отношение нарратора к финалу неясное (нейтральная).

Финальный пункт взят из метаданных о респондентах:

21. Возраст респондента (десятилетие): 1930/1940/1950/1960/1980. В интервью участвовали представители разных поколений русских старожилов, и, предположительно, их возраст становится одним из факторов влияния на перечень сюжетов, которые они знают, и на то, как они их рассказывают и где локализуют.

Сперва обратимся к методу кластерного анализа как к инструменту поиска скрытых закономерностей. Оценка соответствия результатов кластеризации ожиданиям производится посредством сравнения предсказанного моделью распределения текстов с ожидаемыми метками при наличии размеченной выборки. Метрика гомогенности показывает, содержит ли кластер элементы только одного класса. Метрика полноты отвечает на вопрос, все ли элементы одного класса попали в один кластер. Метрика v -меры вычисляется как гармоническое среднее между предыдущими двумя метриками, поэтому является сбалансированным показателем «качества» кластеризации. Индекс Adjusted Rand оценивает совпадение «истинных» и предсказанных меток с поправкой на случайность (особенно актуально для несбалансированных выборок).

В соответствии с исследовательской гипотезой «истинными» метками должны выступить места событий, однако так как геолокация текстов распределена крайне неравномерно (например, почти половина рассказов (45) повествует о событиях в Улан-Баторе, при этом всего в метаданных указано 20 уникальных названий мест), классы укрупнены путем присвоения новых лейблов по принципу географической близости:

«Группа 1»: Улан-Батор, Амгалан (48 текстов).

«Группа 2»: Дзунхара, Дархан, Сухэ-Батор, Корнаковка, Алтан-Булак, Шамара, Оджук (32 текста).

«Группа 3»: Керулен, Чойболсан (7 текстов).

«Группа 4»: Ховд, Монгольский Алтай (2 текста).

«Группа 5»: Булган (2 текста).

«Группа 6»: Братск, Улан-Удэ (5 текстов).

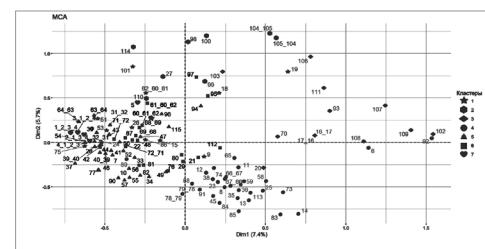
«Группа 7»: Китай (1 текст).

N/A: неизвестно-близко, неизвестно-далеко, русская деревня (18 текстов) — не выделяются в отдельную группу, потому что потенциально кластеризация могла бы указать на локализацию этих текстов относительно любой из семи групп.

КЛАСТЕРИЗАЦИЯ

Одним из широко используемых алгоритмов для кластеризации категориальных данных является k-modes. Принцип работы заключается в том, что алгоритм итеративно определяет наиболее часто встречающиеся значения признаков «центрами» кластеров (модами) и разбивает объекты на кластеры, основываясь на сходстве их признаков с модами. Кластеризация осуществлялась на языке программирования R в среде RStudio с применением имплементации алгоритма из пакета klaR.

Итак, в соответствии с задачей сравнить кластеры с пространственными метками документы разделены на семь кластеров. Для удобства восприятия результаты визуализированы посредством MCA (Multiple Correspondence Analysis — способ представления связей между категориальными переменными в пространстве сниженной размерности).



Визуализация кластеризации (k-modes) посредством MCA

The visualization of clustering (k-modes) using MCA

Разделение на семь кластеров кажется избыточным: три выделенных кластера по левую сторону оси координат накладываются друг на друга. Тем не менее, основываясь на признаках, выделенных алгоритмом в качестве модов, можно сделать интересные наблюдения:

1. Самый плотный кластер 4 объединил в себе 29 документов (чуть менее тре-

ти корпуса), по которым можно судить об угасании фольклорной традиции в регионе: «усредненный портрет» текстов — это пересказ истории в третьем лице, из всех разновидностей средств убеждения в достоверности присутствует только ссылка на источник информации, концовка нейтральная, сюжет строится на взаимодействии главного героя (второстепенные отсутствуют) и человека, обладающего сверхъестественными способностями. Алгоритм верно выделил в качестве моды характеристики «слабых» текстов, находящихся на грани перехода в область бытового рассказа.

2. Выделяются группы текстов, преимущественно относящихся к монгольской традиции (кластеры 2 и 3). Для них характерно наличие диалогов или прямой речи, негативная концовка (сюжет нередко построен на нарушении запрета и последующем наказании). Интересно, что в кластер 2 попали былички, повествующие преимущественно о женщинах, а в кластер 3 — о мужчинах.

Проверка на сходство между кластеризациями по категориальным признакам и лейблами места событий в быличках с использованием метрик показала, что немногочисленные совпадения мест событий и групп текстов практически случайны:

Таблица 1. Значения метрик соответствия кластеризации категориальных признаков (k -modes, R) ожидаемому распределению меток

Метрика	Значение
Гомогенность	0.1964052
Полнота	0.1601505
V-мера	0.1764346
Индекс Adjusted Rand	0.05436365

Для исключения вероятности негативного влияния некоторых признаков на сходство кластеризации с местами событий методом автоматизированного перебора всех комбинаций из 10–20 признаков были подобраны наилучшие значения метрик. Оптимальнее всего показала себя комбинация из 15 признаков (исключены: пол и этничность второстепенного героя,

наличие прямой речи, тип хронологической отсылки, реалистичная детализация рассказа, тональность концовки). По статистическим меркам улучшение показателей совсем незначительное.

Таблица 2. Значения метрик соответствия кластеризации выборочных категориальных признаков (k -modes, R) ожидаемому распределению меток

Метрика	Значение
Гомогенность	0.304879
Полнота	0.2709981
V-мера	0.2869419
Индекс Adjusted Rand	0.1712878

Однако выросшее значение индекса Adjusted Rand сигнализирует о том, что все-таки удалось найти слабую связь между локализацией рассказа и кластеризацией.

ДИСТАНЦИЯ

Основываясь на разметке сюжетно-структурных особенностей мифологических рассказов, представляется возможным проверить, существует ли связь между элементами быличек и дистанцией между местом события и местом записи рассказа. Подобная зависимость уже отмечена, например, С. С. Шляховой [Шляхова 2003]: она утверждает, что дистанция влияет на фоносемантическую структуру былички («чем дальше во времени и пространстве описываемое событие, тем меньше <фоносемантических> аномалий содержит текст»). Мое же предположение о связи строения былички и дистанции от рассказчика основано на неоднократно упоминавшейся главной характеристике жанра — установке на достоверность, которая подталкивает носителя традиции наполнять быличку подробностями. Возможно ли, что близость к месту события упрощает респонденту задачу насыщенного описания произошедшего? Или, наоборот, отдаленность заставляет «припомнить» больше деталей, чтобы убедить слушателя в правдивости истории? Не проявляются ли также отличия в сюжетном наполнении «далеких» рас-

сказов, учитывая, что под категорией «далеко» зачастую подразумеваются деревни, в которых выросли сами респонденты или их старшие родственники?

Чтобы проверить гипотезу, былички обозначены лейблом «близко», если место события и место записи совпадают, и «далеко» — если различаются. Распределение относительно равномерное: 52 былички повествуют о «близких событиях», 63 — об «отдаленных».

Категориальные данные можно исследовать на предмет наличия или отсутствия статистической значимости распределения переменных с применением критерия хи-квадрат или точного критерия Фишера. Критерий хи-квадрат предназначен для таблиц с большими частотами (> 5), из-за чего не применяется к таблицам с сильной диспропорцией между значениями в ячейках, как, например, в моих данных.

Точный критерий Фишера — это тест, проверяющий зависимость между категориальными переменными в таблицах сопряженности размеров $2 \times n$ (стандартно таблица должна иметь размер 2×2 , но реализация алгоритма на языке программирования R позволяет оценивать

таблицы с большим числом столбцов). Он подходит для работы с разреженными и неравномерными выборками (значения частот могут быть менее 5).

Гипотезы, проверяемые точным тестом Фишера, формулируются следующим образом:

H_0 : Две переменные независимы.

H_1 : Две переменные не являются независимыми.

Нулевая гипотеза опровергается, когда p-value (вероятность случайно получить наблюдаемый результат) меньше, чем 0,05, следовательно, в подобных случаях тест демонстрирует наличие устойчивой статистической значимости различий в распределениях.

Таблицы сопряженности (т. е. таблицы распределения двух категориальных переменных) построены для всех признаков (в столбцах) в соотношении с переменной «дистанция» (в строках). В большинстве случаев тест Фишера подтвердил нулевую гипотезу — опровергнуть независимость переменных не удалось. Однако было выявлено несколько интересных статистически значимых зависимостей.

1. Время действия × дистанция.

Таблица 3. Сопряженность между временем действия и дистанцией

Дистанция	Время действия				
	Неясно	День	Ночь	Утро	Вечер
Близко	31	5	8	7	1
Далеко	35	6	21	0	1

Значение p составило 0,006848, при удалении столбца «неясно» оно еще сильнее уменьшилось — до 0,002434 — это показывает высокую значимость результата. Наибольший вклад в низкое p-value вносит переменная «утро»: в исследуемом корпусе повествующие о событиях в это время суток тексты локализуются исключительно как «близкие». А былички о «ночных» событиях в два раза чаще происходят на отдалении. «Утренние» тексты, за исключением одного, рассказывают о личном опыте респондента от первого лица либо о ближайших родственниках, но с высоким процентным содержанием диалога в тексте. «Ночные»

тексты представляют собой либо варианты «традиционных» сюжетов (несколько записаны в единственном варианте — про колдовку, оборачивающуюся в животное, про ночь ведьм; остальные — про гадания, про лечение и снятие порчи), либо мистические рассказы о столкновениях с духами места на дорогах. Наличие связи между временем действия и дистанцией — наглядная иллюстрация концепции хронотопа, сформулированной М. М. Бахтиным, о неразрывной взаимосвязи времени и пространства в художественных произведениях [Бахтин 1975, 398].

2. Пол главного героя × дистанция.

Таблица 4. Сопряженность между полом главного героя и дистанцией

Дистанция	Пол		
	мужской	женский	мужской и женский
Близко	15	36	1
Далеко	35	21	7

P-value равняется 0,0003049, что тоже указывает на высокую достоверность связи. Наблюдается почти ровно противоположное распределение героев-мужчин и героев-женщин по «близким» и «далеким» текстам.

Среди сюжетов о мужчинах в «далеких» текстах превалируют рассказы о слазах, знахарях, знающих, колдунах (в подобных сюжетах главными героями чаще всего оказываются именно люди, обладающие сверхъестественными способностями), а в «близких» текстах их крайне мало, следовательно, можно предположить, что мужская бытовая магия почти исчезла вместе с русскими деревнями старожилов, о которых остались лишь далекие воспоминания. Помимо этого, группа сюжетов о том, как чудилось на дорогах, попала в перечень «далеких быличек о мужчинах» в полном составе: очевидно, это демонстрирует гендерный перекос в профессии шофера, которые сталкивались с духами места и покойниками далеко от дома.

На контрасте сюжеты о слазах, знахарках и колдовках встречаются и в «близких», и в «далеких» быличках о женщинах в одинаковом процентном соотношении (55–57 %), а в абсолютных значениях их стало даже больше в «близкой» среде, преимущественно совпадающей с городской. Это свидетельствует о том, что именно женщины дольше, чем мужчины, сохраняют традиционную культуру при изменении уклада жизни, следовательно, такие сферы, как народная медицина, постепенно становятся чисто женским занятием.

ЗАКЛЮЧЕНИЕ

Результаты анализа не позволяют подтвердить изначальную гипотезу о том, что закономерности в распределении текстов соотносятся напрямую с их простран-

ственным распределением. Документы формировали кластеры по наборам структурных характеристик, однако не выявлено связи между набором сюжетов, а также их структурным оформлением в процессе рассказа и локализацией происходящего в быличке события.

Однако и опровергнуть гипотезу окончательно нельзя. Разметка отличительных признаков быличек хоть и не привела к искомому выявлению закономерностей в распределении текстов по кластерам и пространственным метаданным, но стала основой для подтверждения неоспоримой связи пространства и сюжетов, фиксируемой не только эмпирически, но и статистически. Более того, кластерный анализ признаков с нового ракурса подтвердил наблюдение об угасании фольклорной традиции русских старожилов Монголии: беднеет круг «традиционных» для русского фольклора сюжетов, теряется разнобразие и в целом структура быличек.

Неоднозначность достигнутого результата открывает перспективы для дальнейших исследований фольклорного материала. Расширение списка категориальных признаков, вероятно, позволит углубить понимание связи между структурой повествования и локализацией текстов. Можно также выявлять закономерности методом кластеризации текстов на основе векторного представления лексической информации (TF-IDF, word2vec, BERT и т. д.) и/или применять к документам другие методы анализа (тематическое моделирование, классификация и т. д.). Для продолжения работы в любом направлении необходимо увеличить объем корпуса быличек, что, надеюсь, будет достигнуто в ближайшие годы благодаря новым полевым исследованиям среди русских старожилов Монголии.

Источники и материалы

ФК ИМЛИ — Фольклорная коллекция Института мировой литературы им. А. М. Горького РАН. Фонд В. Л. Кляуса, Экспедиции 2018, 2019 гг. Фонд гранта РНФ № 23-18-00478. Экспедиции 2023–2024 гг.

Исследования

Бахтин 1975 — *Бахтин М. М. Формы времени и хронотопа в романе* // Бахтин М. М. Вопросы литературы и эстетики: Исследования разных лет. М.: Худож. лит., 1975.

Биткеева, Исаков 2024 — *Биткеева А. Н., Исаков А. В. Дискурсы конструирования исторической памяти о местнорусских в Монголии* // *Oriental Studies*. 2024. Т. 17. № 4. С. 870–882.

Галлямов, Орехов 2016 — *Галлямов А. А., Орехов Б. В. Об электронном научном издании «Фольклорный архив Башкирского государственного университета»* // Вестник РГГУ. Сер. Литературоведение. Языкоизнание. Культурология. 2016. № 12 (21). С. 140–149.

Даревская 1994 — *Даревская Е. М. Сибирь и Монголия: очерки русско-монгольских связей в конце XIX — начале XX веков*. Иркутск: Изд-во Иркут. ун-та, 1994.

Единархова 1999 — *Единархова Н. Е. Из истории пребывания русских в Монголии (до 1917 г.)* // Диаспоры. 1999. № 2–3. С. 81–99.

Зиновьев 2019 — *Зиновьев В. П. Русский фольклор Восточной Сибири*. Т. 1 / Под ред. Р. П. Матвеевой. Иркутск, 2019.

Кляус 2024 — *Кляус В. Л. Кукла с золотыми глазами: Об одном уникальном забайкальском сказочно-мифологическом сюжете в фольклоре местнорусских Монголии* // Языки и фольклор коренных народов Сибири. 2024. № 2 (Вып. 50). С. 131–144.

Кляус, Махова 2025 — *Кляус В. Л., Махова М. П. Песенный фольклор русских старожилов Монголии* // Живая старина. 2024. № 3. С. 24–29.

Михалев 2011 — *Михалев А. В. Местнорусское население Монголии. Статус «Другого» в условиях политических трансформаций XX века* // *Tartaria Magna*. 2011. № 1. С. 54–64.

Померанцева 1975 — *Померанцева Э. В. Мифологические персонажи в русском фольклоре*. М.: Наука, 1975.

Смирнова 2021 — *Смирнова А. А. Электронный фольклорный архив НИУ ВШЭ* // Живая старина. 2021. № 2 (110). С. 41–42.

Сундуева 2011 — *Сундуева Д. Б. Речевые портреты соотечественников в*

Монголии

// Вестник ЧитГУ. 2011 № 4 (71). С. 63–69.

Сундуева 2024 — *Сундуева Д. Б. Русские старожилы Монголии: страницы истории* // *Русские старожилы Монголии в воспоминаниях, очерках и фотографиях* / Сост. Д. Б. Сундуева, В. Л. Кляус. М.: ИМЛИ РАН, 2024. С. 15–18.

Черванёва 2017 — *Черванёва В. А. «Рассуждающий» повествователь в мифологическом нарративе* // Вестник РГГУ. Сер. Литературоведение. Языкоизнание. Культурология. 2017. № 12 (33). С. 9–16.

Шляхова 2003 — *Шляхова С. С. Фоносемантическая структура текста былички* // Фольклорный текст-2002: Матер. науч.-практ. семинара (г. Чайковский, 15 окт. 2002 г.; г. Пермь, 10 янв. 2003 г.) / [Науч. ред. И. А. Подюков]. Пермь: Перм. гос. пед. ун-т, 2003. С. 34–43.

Шумов 1993 — *Шумов К. Э. Современное состояние былички в Северном Прикамье. Опыт машинного анализа сюжетного состава локальной традиции: Автореф. дис. ... канд. филол. наук: 10.01.09. Екатеринбург, 1993.*

Lindemann et al. 2019 — *Lindemann M., Grünwald S., Declerck T. Annotation and classification of locations in folktales* // Proceedings of the Second Workshop on Corpus-Based Research in the Humanities, Vienna, Austria, Gerastree Proceedings, GTP. 2019. Vol. 1.

Meder et al. 2016 — *Meder T., Karsdorp F., Nguyen D., Theune M., Trieschnigg D., Muiser I. E. C. Automatic enrichment and classification of folktales in the Dutch Folktale Database* // Journal of American Folklore. 2016. Т. 129. № 511. P. 78–96.

Ning, Chen 2023 — *Ning H., Chen Z. Fusion of the word2vec word embedding model and cluster analysis for the communication of music intangible cultural heritage* // Scientific Reports. 2023. Vol. 13. № 1.

Ross, Atkinson 2016 — *Ross R. M., Atkinson Q. D. Folktale transmission in the Arctic provides evidence for high bandwidth social learning among hunter-gatherer groups* // Evolution and Human Behavior. 2016. Vol. 37. № 1. P. 47–53.

Strle, Marolt 2016 — *Strle G., Marolt M. Language Technologies in Humanities: Computational Semantic Analysis in Folkloristics* // Conference on Language Technologies & Digital Humanities, Ljubljana. 2016. P. 227–229.

Tehrani 2013 — *Tehrani J. J. The Phylogeny of Little Red Riding Hood* // PloS one. 2013. Vol. 8. № 11.

© С. А. Федотова, 2025

СВЕДЕНИЯ ОБ АВТОРЕ

Федотова С. А. <https://orcid.org/0009-0000-0264-8829>

Младший научный сотрудник отдела фольклора Института мировой литературы им. А. М. Горького РАН: Российская Федерация, 121069, г. Москва, Поварская ул., д. 25а, стр. 1; тел.: +7 (495) 690-50-30; e-mail: fedotovy4.11@gmail.com

Mythological Tales of Russian Old-Time Settlers of Mongolia in the Projection of Computational Folkloristics

Sofia A. Fedotova

(M. Gorky Institute of World Literature of the Russian Academy of Sciences:

25a, Povarskaya str., Moscow, 121069, Russian Federation)

Summary. This study undertook an examination of the correlation between the distribution of texts (and their plot and structural features) with the spatial distribution of Russian old-time settlers in Mongolia. The data consists of transcribed mythological tales (*bylichki*) collected in 2018–2024 in Ulaanbaatar and Darkhan. For the first time the study of folklore texts recorded during ethnographic interviews with Russian settlers of Mongolia was conducted using methods of digital humanities. The choice of mythological tales as the object of study is due to their relatively good preservation in the tradition of the Russians of Mongolia compared to other genres of folklore. With reliance on Russian and international studies in the field of computational folkloristics, an algorithm was developed for the manual coding of the features of mythological tales as a genre of folklore. The search for patterns among the categorized features was carried out using the k-modes clustering method, statistical analysis (Fisher's exact text) and data visualization via multiple correspondence analysis (MCA). Although the text clusters identified by the algorithm did not correlate with the geolocation of the events described in the stories, statistically significant dependencies were found between the distance of the story's action from the place of recording ("near" or "far") and such characteristics as the time of action and the gender of the main character. In addition, the results of the analysis support the observation that the folklore tradition of the region is fading away.

Key words: Russian old-time settlers of Mongolia, mythological tales, digital humanities, categorial features.

Acknowledgements. The study was carried out within the framework of RFS project No. 23-18-00478, "Russian Mongolia. A Comprehensive Study of Culture in a Non-Ethnic Environment (Folklore, Ritual Traditions, Language)".

Received: September 16, 2025.

Date of publication: December 25, 2025.

For citation: Fedotova S. A. Mythological Tales of Russian Old-Time Settlers of Mongolia in the Projection of Computational Folkloristics. *Traditional Culture*. 2025. Vol. 26. No. 4. Pp. 87–98. In Russian.

DOI: <https://doi.org/10.26158/TK.2025.26.4.007>

References

- Bakhtin M. M. (1975) Formy vremeni i khronotopa v romane [Forms of Time and Chronotope in the Novel]. In: Bakhtin M. M. Voprosy literatury i estetiki [Questions of Literature and Aesthetics]. Moscow: Khudozh. lit. In Russian.
- Bitkeeva A. N., Isakov A. V. (2024) Diskursy konstruirovaniyu istoricheskoi pamyati o mestno-russkikh v Mongoli [Russian Diaspora in Mongolia: Discourses of Historical Memory Construction]. *Oriental Studies*. 2024. Vol. 17. No. 4. Pp. 870–882. In Russian.
- Chervaneva V. A. (2017) «Rassuzhdayushchii» povestvovatel' v mifologicheskem narrative [The “Reasoning” Narrator in a Mythological Narrative]. *Vestnik RGGU. Seriya Literaturovedenie. Yazykoznanie. Kul'turologiya* [RSUH/RGGU Bulletin: “Literary Theory. Linguistics. Cultural Studies” series]. 2017. No. 12 (33). Pp. 9–16. In Russian.
- Darevskaya E. M. (1994) Sibir'i Mongoliya: Ocherki russko-mongol'skikh svyazei v kontse XIX — nachale XX vekov [Siberia and Mongolia: Essays on Russian-Mongolian Relations in the Late Nineteenth and Early Twentieth Centuries]. Irkutsk: Izd-vo Irkut. un-ta. In Russian.
- Edinarkhova N. E. (1999) Iz istorii prebyvaniya russkikh v Mongoli (do 1917 g.) [From the History of Russian Presence in Mongolia (Before 1917)]. *Diaspora* [Diasporas]. 1999. No. 2–3. Pp. 81–99. In Russian.
- Gallyamov A. A., Orekhov B. V. (2016) Ob elektronnom nauchnom izdanii «Fol'klornyi arkhiv Bashkirskogo gosudarstvennogo universiteta» [On the Electronic Edition of the Folklore Archive of Bashkir State University]. *Vestnik RGGU. Seriya Literaturovedenie. Yazykoznanie. Kul'turologiya* [RSUH/RGGU Bulletin: “Literary Theory. Linguistics. Cultural Studies” series]. 2016. No. 1 (34). Pp. 10–17. In Russian.

- istics. Cultural Studies" Series]. 2016. No. 12 (21). Pp. 140–149. In Russian.
- Klyaus V. L.** (2024) Kukla s zolotymi glazami: Ob odnom unikal'nom zabaikal'skom skazochno-mifologicheskem syuzhete v fol'klore mestnorusskikh Mongolii [A Doll with Golden Eyes: About a Unique Fairytale-Mythological Plot in the Folklore of Russians in Mongolia]. *Yazyki i fol'klor korennykh narodov Sibiri* [Languages and Folklore of Indigenous Peoples of Siberia]. 2024. No. 2 (50). Pp. 131–144. In Russian.
- Klyaus V. L., Makhova M. P.** (2025) Pesennyyi fol'klor russkikh starozhilov Mongolii [Folklore Songs of Russian Old-Time Settlers of Mongolia]. *Zhivaya starina* [Living Antiquity]. 2024. No. 3. Pp. 24–29. In Russian.
- Lindemann M., Grünwald S., Declerck T.** (2019) Annotation and Classification of Locations in Folktales. Proceedings of the Second Workshop on Corpus-Based Research in the Humanities. Vienna, Austria, Gerastree Proceedings. GTP. Vol. 1. In English.
- Meder T., Karsdorp F., Nguyen D., Theune M., Trieschnigg D., Muiser I. E. C.** (2016) Automatic Enrichment and Classification of Folktales in the Dutch Folktale Database. *Journal of American Folklore*. 2016. Vol. 129. No. 511. Pp. 78–96. In English.
- Mikhailov A. V.** (2011) Mestnorusskoe naselenie Mongolii. Status «Drugogo» v usloviyakh politicheskikh transformatsii XX veka [The Local Russian Population of Mongolia: The Status of 'Outsider' in Conditions of Political Transformations of the Twentieth Century]. *Tartaria Magna*. 2011. No. 1. Pp. 54–64. In Russian.
- Ning H., Chen Z.** (2023) Fusion of the Word2vec Word Embedding Model and Cluster Analysis for the Communication of Music Intangible Cultural Heritage. *Scientific Reports*. Vol. 13. No. 1. In English.
- Pomerantseva E. V.** (1975) Mifologicheskie personazhi v russkom fol'klore [Mythological Characters in Russian Folklore]. Moscow: Nauka. In Russian.
- Ross R. M., Atkinson Q. D.** (2016) Folktales Transmission in the Arctic Provides Evidence for High Bandwidth Social Learning among Hunter-gatherer Groups. *Evolution and Human Behavior*. 2016. Vol. 37. No. 1. Pp. 47–53. In English.
- Shlyakhova S. S.** (2003). Fonosemanticskaya struktura teksta bylichki [Phonosemantic Structure of the Text of the Bylichka]. In: *Fol'klornyi tekst-2002* [Folklore Text-2002]. Ed. by I. A. Podyukov. Perm': Perm' gos. ped. un-t. Pp. 34–43. In Russian.
- Shumov K. E.** (1993) Sovremennoe sostoyanie bylichki v Severnom Prikame. Opyt mashinnogo analiza syuzhetnogo sostava lokal'noi traditsii [The Current State of the Bylichka in the Northern Kama Region. An Experiment in Computational Analysis of the Plot Composition of a Local Tradition]. Abstract of a PhD dissertation. Ekaterinburg. In Russian.
- Smirnova A. A.** (2021). Elektronnyi fol'klornyy arkhhiv NIU VSHEH [Electronic Folklore Archive of HSE University]. *Zhivaya starina* [Living Antiquity]. 2021. No. 2. Pp. 41–42. In Russian.
- Strle G., Marolt M.** (2016) Language Technologies in Humanities: Computational Semantic Analysis in Folkloristics. Conference on Language Technologies & Digital Humanities. Ljubljana. Pp. 227–229. In English.
- Sundueva D. B.** (2011) Rechevye portrety sootechestvennikov v Mongolii [Verbal Portraits of Mongolian Russians]. *Vestnik ChitGU* [Bulletin of Chita State University]. 2011. No. 4 (71). Pp. 63–69. In Russian.
- Sundueva D. B.** (2024) Russkie starozhily Mongolii: stranitsy istorii [Russian Old-Time Settlers of Mongolia: Pages of History]. In: *Russkie starozhily Mongolii v vospominaniyah, ocherkakh i fotografiyah* [Russian Old-Time Settlers of Mongolia in Memoirs, Essays and Photographs]. Comp. by D. B. Sundueva, V. L. Klyaus. Moscow: IMLI RAN. Pp. 15–18. In Russian.
- Tehrani J. J.** (2013) The Phylogeny of Little Red Riding Hood. *PloS one*. 2013. Vol. 8. No. 11. In English.
- Zinov'yev V. P.** (2019) Russkii fol'klor Vostochnoi Sibiri [Russian Folklore of Eastern Siberia]. Ed. by R. P. Matveeva. Vol. 1. Irkutsk. In Russian.

© S. A. Fedotova, 2025

ABOUT THE AUTHOR

Sofia A. Fedotova <https://orcid.org/0009-0000-0264-8829>

E-mail: fedotovy4.11@gmail.com

Tel.: +7 (495) 690-50-30

25a, Povarskaya str., Moscow, 121069, Russian Federation

Junior Researcher, Department of Folklore, A. M. Gorky Institute of World Literature, Russian Academy of Sciences



This is an open access article distributed under the Creative Commons Attribution 4.0 International (CC BY 4.0)