
РОССИЙСКАЯ АКАДЕМИЯ НАУК

ВОПРОСЫ ЯЗЫКОЗНАНИЯ



6

2024



НАУКА

— 1727 —

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Отделение историко-филологических наук

ВОПРОСЫ ЯЗЫКОЗНАНИЯ

Журнал основан в январе 1952 года

Выходит 6 раз в год

6

НОЯБРЬ — ДЕКАБРЬ

Москва

2024

Главный редактор:

В. А. Плунгян д. ф. н., проф., академик РАН, Институт русского языка им. В. В. Виноградова РАН; Московский государственный университет им. М. В. Ломоносова

Зам. главного редактора:

Н. Б. Вахтин д. ф. н., проф., чл.-корр. РАН, Европейский университет в Санкт-Петербурге; Институт лингвистических исследований РАН

В. И. Подлесская д. ф. н., проф., Институт языкознания РАН

Редколлегия:

В. М. Алпатов д. ф. н., проф., академик РАН, Институт языкознания РАН

И. М. Богуславский д. ф. н., проф., Институт проблем передачи информации им. А. А. Харкевича РАН; Мадридский политехнический университет, Испания

М. Д. Воейкова д. ф. н., Институт лингвистических исследований РАН

В. З. Демьянков д. ф. н., проф., Институт языкознания РАН

Д. О. Добровольский д. ф. н., проф., Институт русского языка им. В. В. Виноградова РАН; Институт языкознания РАН

А. Ф. Журавлёв д. ф. н., Институт славяноведения РАН; Московский государственный университет им. М. В. Ломоносова

П. В. Иосад Ph.D., Эдинбургский университет, Великобритания

Н. Н. Казанский д. ф. н., проф., академик РАН, Институт лингвистических исследований РАН

В. И. Киммельман Ph.D., Бергенский университет, Норвегия

Г. И. Кустова д. ф. н., проф., Институт русского языка им. В. В. Виноградова РАН

А. М. Молдован д. ф. н., академик РАН, Институт русского языка им. В. В. Виноградова РАН

М. С. Полинская Ph.D., проф., Мэрилендский университет, США

Е. В. Рахилина д. ф. н., проф., Национальный исследовательский университет «Высшая школа экономики»; Институт русского языка им. В. В. Виноградова РАН

Я. Г. Тестелец д. ф. н., проф., Российский государственный гуманитарный университет; Институт языкознания РАН

Л. А. Янда Ph.D., проф., Университет Тромсё — Арктический университет Норвегии, Норвегия

Зав. редакцией: Н. В. Ганнус

Зав. отделами: А. С. Кулева, А. Д. Подгорная

Отдел рецензий: М. И. Сатина

Статьи отбираются редколлегией журнала на основе анонимного независимого рецензирования.

Индексируется в: Российский индекс научного цитирования (РИНЦ); Elsevier BV Scopus; European Reference Index for the Humanities and Social Sciences (ERIH PLUS); Russian Science Citation Index (RSCI); Web of Science Core Collection's Emerging Sources Citation Index (ESCI).

Адрес редакции: 119019, Москва, ул. Волхонка, 18/2, Институт русского языка им. В. В. Виноградова РАН, редакция журнала «Вопросы языкознания»

Телефон: +7 495 637-25-16

E-mail: voprosy@mail.ru

Сайт: <https://vja.ruslang.ru>

© Российская академия наук, 2024

© Составление. Редколлегия журнала «Вопросы языкознания», 2024

RUSSIAN ACADEMY OF SCIENCES

Department of History and Philology

VOPROSY
JAZYKOZNANIJA
(TOPICS IN THE STUDY OF LANGUAGE)

Founded in January 1952
6 issues per year

6
NOVEMBER — DECEMBER

Moscow
2024

Editor-in-chief:

Vladimir A. PLUNGAN Vinogradov Russian Language Institute (RAS); Lomonosov Moscow State University, Moscow, Russia

Assistant editors:

Vera I. PODLESSKAYA Institute of Linguistics (RAS), Moscow, Russia
Nikolai B. VAKHTIN European University at St. Petersburg; Institute for Linguistic Studies (RAS), St. Petersburg, Russia

Editorial board:

Vladimir M. ALPATOV Institute of Linguistics (RAS), Moscow, Russia
Igor M. BOGUSLAVSKY Kharkevich Institute for Information Transmission Problems (RAS), Moscow, Russia; Universidad Politécnica de Madrid, Spain
Valery Z. DEMYANKOV Institute of Linguistics (RAS), Moscow, Russia
Dmitriy O. DOBROVOL'SKIY Vinogradov Russian Language Institute (RAS); Institute of Linguistics (RAS), Moscow, Russia
Pavel IOSAD University of Edinburgh / Oilthigh Dhùn Èideann, UK
Laura A. JANDA Universitetet i Tromsø: Norges arktiske universitet, Tromsø, Norway
Nikolai N. KAZANSKY Institute for Linguistic Studies (RAS), St. Petersburg, Russia
Vadim KIMMELMAN University of Bergen, Norway
Galina I. KUSTOVA Vinogradov Russian Language Institute (RAS), Moscow, Russia
Aleksandr M. MOLDOVAN Vinogradov Russian Language Institute (RAS), Moscow, Russia
Maria POLINSKY University of Maryland, College Park, USA
Ekaterina V. RAKHILINA HSE University; Vinogradov Russian Language Institute (RAS), Moscow, Russia
Yakov G. TESTELETS Russian State University for the Humanities; Institute of Linguistics (RAS), Moscow, Russia
Maria D. VOEIKOVA Institute for Linguistic Studies (RAS), St. Petersburg, Russia
Anatoly F. ZHURAVLEV Institute of Slavic Studies (RAS); Lomonosov Moscow State University, Moscow, Russia

Managing editor: Natalia V. GANNUS

Editorial staff: Anna S. KULEVA, Anastasia D. PODGORNAIA

Review editor: Maria I. SATINA

Articles are selected by the editorial board on the basis of anonymous double-blind independent peer review process.

Abstracting/Indexing: Elsevier BV Scopus; European Reference Index for the Humanities and Social Sciences (ERIH PLUS); Rossiiskii indeks nauchnogo tsitirovaniya (RINTs); Russian Science Citation Index (RSCI); Web of Science Core Collection's Emerging Sources Citation Index (ESCI).

Address: "Voprosy Jazykoznanija", editorial office, Vinogradov Russian Language Institute (RAS), Volkhonka street, 18/2, Moscow, 119019, Russia

Telephone: +7 495 637-25-16

E-mail: voprosy@mail.ru

Website: <https://vja.ruslang.ru>

Содержание

Е. А. Лютикова. Адъективные сказуемые в финитных и нефинитных клаузах	7
К. В. Филатов. О некоторых типологических характеристиках андийских глагольных систем	32
И. П. Новак. Карельский языковой ландшафт в диалектометрической парадигме . . .	58
В. Л. Васильев, И. И. Муллонен. Средневековые названия погостов Корельского уезда: на границе государств и языков	85
К. А. Студеникина. Взаимосвязь между приемлемостью и вероятностью высказываний: данные предикативного согласования с сочиненным подлежащим в русском языке	105
М. А. Шумилина. Референциальный выбор в устных и письменных рассказах: сопоставительное исследование на материале корпуса «Веселые истории из жизни»	133

Рецензии

Д. А. АЛЕКСЕЕВ. [Рец. на:] R. Meyer. <i>Iranian syntax in Classical Armenian: The Armenian perfect and other cases of pattern replication</i> . Oxford: Oxford University Press, 2023	160
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Contents

Ekaterina A. LYUTIKOVA. Adjectival predicates in finite and non-finite clauses	7
Konstantin V. FILATOV. On typologically relevant properties of Andic verbal grammatical systems	32
Irina P. NOVAK. The Karelian language landscape in the dialectometric paradigm	58
Valery L. VASILYEV, Irma I. MULLONEN. Medieval names of churchyards in Korelsky county: On the border of states and languages	85
Ksenia A. STUDENIKINA. Interaction between acceptability and probability: Evidence from predicate agreement with a coordinated subject in Russian	105
Marina A. SHUMILINA. Referential choice in spoken and written stories: A comparative study based on the corpus “Funny life stories”	133

Reviews

Danil A. ALEKSEEV. [Review of:] R. Meyer. <i>Iranian syntax in Classical Armenian: The Armenian perfect and other cases of pattern replication</i> . Oxford: Oxford University Press, 2023	160
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Адъективные сказуемые в финитных и нефинитных клаузах

© 2024

Екатерина Анатольевна Лютикова

Московский государственный университет имени М. В. Ломоносова, Москва, Россия;

lyutikova2008@gmail.com

Аннотация: В статье исследуется дистрибуция различных адъективных предикатов (кратких прилагательных, полных прилагательных в именительном и творительном падеже и пассивных причастий) в финитных и нефинитных клаузах в русском языке. Выявляются два фактора, предопределяющих приемлемость адъективных предикатов в определенной клаузе: (i) тип подлежащего клаузы (выраженная именная группа, А-след или PRO) и (ii) падеж подлежащего (номинатив или прочие падежи). Установлено, что адъективные предикаты отличаются друг от друга в отношении условий лицензирования: для использования краткого прилагательного необходимо, чтобы его подлежащее (выраженная именная группа, А-след или PRO) получило номинатив, а полные прилагательные в именительном падеже допустимы, только если их подлежащее — номинативная именная группа или А-след номинативной именной группы, но не номинативное PRO. Включение в общую картину причастных пассивов выявляет дополнительные расхождения в дистрибуции адъективных сказуемых: полная форма страдательного причастия не может образовывать пассив в финитной клаузе, но возможна в качестве пассивного предиката в нефинитных клаузах. В статье предлагается формальный анализ дистрибуции адъективных предикатов, который опирается на следующие допущения: (i) категориальный контраст краткой и полной форм; (ii) синтаксическое противопоставление контроля и подъема и (iii) способность PRO к получению падежа через управление со стороны комплементаризера или через унификацию падежного признака с контролером.

Ключевые слова: залог, именное сказуемое, контроль, краткие формы, падеж, пассив, подъем, полные формы, прилагательные, русский язык, синтаксис, финитность

Благодарности: Исследование выполнено в рамках проекта РНФ № 22-18-00037 «Параметрическая модель согласования в свете экспериментальных данных», реализуемого в МГУ имени М. В. Ломоносова.

Для цитирования: Лютикова Е. А. Адъективные сказуемые в финитных и нефинитных клаузах. *Вопросы языкознания*, 2024, 6: 7–31.

DOI: 10.31857/0373-658X.2024.6.7-31

Adjectival predicates in finite and non-finite clauses

Ekaterina A. Lyutikova

Lomonosov Moscow State University, Moscow, Russia; lyutikova2008@gmail.com

Abstract: The paper explores the distribution of various adjectival predicates (short form adjectives, long form adjectives in the nominative or instrumental case and passive participles) in Russian finite and non-finite clauses. Two factors turn out to determine the availability of adjectival predicates: (i) the type of the subject (overt DP vs. A-trace vs. PRO) and (ii) the case feature of the subject (nominative vs. other cases). Crucially, adjectival predicates differ as to the licensing factors: short form adjectives require that their subject (overt DP, A-trace or PRO) be nominative whereas nominative long form adjectives are only licit if their subject is a nominative overt DP or an A-trace of a nominative DP, but not a (nominative) PRO. Including adjectival passive into this picture gives rise to further discrepancies: the

long form of the passive participle cannot form a passive predicate of a finite clause but is licit as such a predicate in non-finite clauses. The paper provides a formal analysis of the distribution of adjectival predicates, which relies on the following assumptions: (i) the categorial contrast of short and long forms; (ii) the syntactically represented opposition of control and raising; and (iii) PRO's ability to receive the case feature via case assignment by complementizer or case transmission from the controller.

Keywords: adjective, case, control, finiteness, full forms, nominal predication, passive, raising, Russian, short forms, syntax, voice

Acknowledgements: This research is supported by Russian Science Foundation, RSF project No. 22-18-00037 "Parametric model of agreement in the light of experimental data" realized at Lomonosov Moscow State University.

For citation: Lyutikova E. A. Adjectival predicates in finite and non-finite clauses. *Voprosy Jazykoznanija*, 2024, 6: 7–31.

DOI: 10.31857/0373-658X.2024.6.7-31

1. Введение

В статье исследуются условия лицензирования различных классов адъективных сказуемых: кратких прилагательных, полных прилагательных в формах номинатива и инструменталиса, причастных пассивов — в финитных и нефинитных клаузах русского языка. Дистрибуция адъективных сказуемых разных классов в финитных клаузах подробно описывается в грамматиках [ГР 1952; РГ 1980; Кустова 2018] и специальных работах, посвященных прилагательным в предикативной позиции [Шведова 1952; Babby 1973; Nichols 1981; Никольс 1985; Никунласси 1993; Всеволодова 2013]. Основными вопросами в подобных исследованиях оказываются семантические и грамматические факторы, определяющие выбор между полной и краткой формой прилагательного, а также между падежными формами полного прилагательного в предикативной позиции. При этом значительно менее изученной оказывается дистрибуция адъективных сказуемых в разных типах нефинитных клауз; исключение составляют, пожалуй, лишь конструкции с ползунаменательными и функциональными глаголами, отдельно упоминаемые в [РГ 1980: §1956, §2346, §2357, §2365; Nichols 1981: 94–118; Всеволодова 2013]; см. также [Лютикова 2010; Гращенков 2022] об инфинитивных клаузах. При этом очевидно, что разные классы адъективных сказуемых по-разному взаимодействуют с финитностью клаузы. Настоящее исследование ставит перед собой задачу восполнения этого пробела в дескриптивной грамматике русского языка.

С теоретической точки зрения функционирование адъективных сказуемых в контекстах, различающихся с точки зрения финитности, важно как минимум в двух отношениях. Во-первых, как будет показано в статье, разные классы адъективных сказуемых имеют разную дистрибуцию в нефинитных клаузах. Соответственно, ограничения на функционирование в определенных контекстах могут послужить основанием для оценки предположений о категориальном статусе полных и кратких прилагательных и структуре проецируемых ими групп в составе именного сказуемого. Указанные вопросы активно обсуждаются в формальной русистике, и, по-видимому, консенсус в их решении пока не достигнут (см. подробный обзор современного состояния проблемы в [Гращенков 2022]).

Во-вторых, идентификация ограничений на разные классы адъективных сказуемых позволяет пролить свет на структуру нефинитных клауз. Если определенный класс адъективных сказуемых лицензируется в одном нефинитном контексте и не лицензируется в другом, то очевидно, что эти контексты имеют структурные различия: первый контекст содержит лицензирующий фактор, а второй — нет. Такого рода данные оказываются востребованы в дискуссии о структуре инфинитивных оборотов в русском языке и, шире, в моделировании конфигураций контроля и их отличий от конфигураций подъема.

Наконец, еще один важный аспект исследования адъективных сказуемых в русском языке связан с моделированием причастного пассива. Дистрибуция кратких и полных форм страдательного причастия во многом напоминает дистрибуцию соответствующих форм прилагательных [Babby 1973], однако имеются и характерные отличия, такие, например, как блокирование собственно пассивной интерпретации полной формы причастия в финитном сказуемом. Объяснение дистрибуции причастных пассивов в разных типах клауз — задача, которая может быть решена только при комплексном изучении адъективных сказуемых как единого структурного класса.

Дальнейшее изложение построено следующим образом. В разделе 2 будут кратко охарактеризованы классы адъективных сказуемых, доступных в финитных клаузах, и показана специфика причастных пассивов. Раздел 3 посвящен обзору нефинитных клауз русского языка и отбору релевантных для исследования контекстов. В разделе 4 обсуждаются ограничения на использование адъективных сказуемых разных классов в нефинитных контекстах и формулируются содержательные обобщения относительно лицензирующих их факторов. Раздел 5 посвящен дистрибуции причастных пассивов в нефинитных клаузах в сравнении с прочими адъективными сказуемыми. Раздел 6 содержит набросок формального анализа исследуемых классов адъективных сказуемых и предлагает модель, реализующую полученные в разделах 4 и 5 содержательные обобщения. В разделе 7 подводятся итоги исследования.

2. Адъективные сказуемые в финитных клаузах

В современном русском языке в предикативной позиции могут выступать следующие виды адъективов: краткое прилагательное (1a), полное прилагательное в формах именительного (1b) и творительного (1c) падежа, а также краткое страдательное причастие, образующее причастный пассив (1d)–(1e).

- (1) а. Пожалуй, я не была по-настоящему **талантлива** [НКРЯ].
б. Зарплата моя была **маленькая**, меньше, чем у технички, но мне хватало [НКРЯ].
с. Обстановка в городе была **тревожной** [НКРЯ].
д. Корреспонденция «Море пора спасать» была **опубликована** в «Правде» 22 января [НКРЯ].
е. Взгляд действующего лица в кадре падал на какую-либо вещь или деталь — эта вещь была **подаваема** крупным планом [НКРЯ].

Предикативные категории клаузы при адъективном сказуемом выражаются формами глагола-связки (в причастном пассиве — вспомогательного глагола) *быть*, представленного нулевой словоформой в настоящем времени индикатива. Адъективные компоненты сказуемого согласуются с подлежащим в числе и роде (в формах единственного числа).

Структурные ограничения на функционирование перечисленных форм в финитном контексте сводятся к следующим. Во-первых, форма творительного падежа полного прилагательного не используется при нулевой словоформе глагола-связки [РГ 1980: §2341, §2345]. Во-вторых, страдательные причастия в причастных пассивах в финитных клаузах используются только в кратких формах; полные формы как в именительном, так и в творительном падеже не имеют событийной интерпретации, что подтверждается их неспособностью присоединять типичные для пассивов адъюнкты временной локализации и агентивные дополнения (2a)–(2b).

- (2) а. *Корреспонденция «Море пора спасать» была **опубликованная** в «Правде» 22 января.
б. *Эта вещь была **подаваемой** режиссером крупным планом.

Появление полных форм страдательных причастий в предикативной позиции финитной клаузы обусловлено конверсией причастий в прилагательные; такие формы обозначают признак и не имеют событийной интерпретации, ср. (3a)–(3b).

- (3) а. Поверхность стойки была **полированной** (*в прошлом году) и отражала потолок [НКРЯ].
- б. Стоянка наша была **охраняемая** (*сотрудниками ЧОП), для сотрудников, но охранник или спал, или не видел преступника, а может быть, был с ним заодно [НКРЯ].

3. Нефинитные клаузы

В русском языке представлено несколько видов нефинитных клауз: инфинитивные¹, деепричастные, причастные, номинализованные и малые. В настоящем исследовании мы рассмотрим три вида нефинитных клауз: инфинитивные, деепричастные и малые клаузы, исключая из рассмотрения номинализованные и причастные клаузы. Это ограничение мотивировано следующими соображениями.

Номинализации в русском языке представлены как регулярными (на *-ниj(e)/-тиj(e)*), так и нерегулярными отглагольными дериватами (суффиксы *-Ø*, *-к(a)*, *-б(a)*, *-ом(a)*, *-циj(a)* и др.) [РГ 1980: §256–283; Пазельская 2009]. Для обоих типов характерна полисемия — неоднозначность между событийной и предметной или результативной интерпретацией, которая обычно анализируется как следствие того, что номинализации могут подвергаться как основы, так и расширенные проекции глагола [Alexiadou 2001]. Однако даже событийные номинализации в русском языке содержат небольшое, по сравнению с нефинитными клаузами других типов, количество клаузальной функциональной структуры [Schoorlemmer 1998; Alexiadou 2001; Rappaport 2001; Пазельская, Татевосов 2008]. Русская номинализирующая морфология принимает только глагольные составляющие, а глагол-связка (вспомогательный глагол в случае пассивов) появляется в структуре выше того уровня, который доступен номинализирующей вершине. Вследствие этого в русском языке клаузы с адъективным сказуемым (и причастным пассивом) не номинализуются (**бытие уверен / уверенная / уверенной / построен*). Таким образом, материал номинализаций нерелевантен для целей данной статьи.

Причастные клаузы в русском языке возглавляются действительными или полными страдательными причастиями и выступают как препозитивные или постпозитивные

¹ В русском языке возможны также независимые инфинитивные предложения (*России не быть под Антантой, Саду цвести*) [РГ 1980: §2560; Падучева 2017]. По меньшей мере часть из них допускает выражение категории времени и наклонения при помощи вспомогательного глагола, то есть является финитными: *Грузовикам было / будет / было бы не проехать* [Greenberg, Franks 1991]. О структуре таких предложений в формальной русистике идет непрекращающаяся дискуссия; среди дискуссионных моментов, в частности, вопрос о моноклаузальности или биклаузальности таких конструкций и, соответственно, о статусе инфинитивной клаузы как главной или зависимой, см. [Fleisher 2006; Tsedryk 2017; Моргунова 2021; Zimmerling 2024]. Возможный источник дательного падежа подлежащего в этих конструкциях также зависит от их структурной интерпретации: это может быть как (лексический) экспериенциальный датив, характерный для модальных предикатов и предикативов, так и (структурный) датив подлежащего, приписываемый ему со стороны финитной несогласуемой предикативной вершины Т или выбирающего такую вершину комплементаризера С. В связи с неясным статусом инфинитивной клаузы и ее подлежащего в независимых инфинитивных предложениях мы воздерживаемся от их обсуждения в этой статье, хотя поведение адъективных сказуемых в этих предложениях, по-видимому, совпадает с их поведением в целевых инфинитивных оборотах, имеющих собственное выраженное подлежащее.

определения в именной группе². Для причастных клауз в русском языке отсутствует консенсус относительно их внутренней структуры, в том числе и относительно наличия и формы подлежащего. С одной стороны, имеются свидетельства того, что причастные клаузы могут содержать собственное нулевое подлежащее [Тестелец 2001: 305–307]. Такие свидетельства включают способность подлежащих причастных клауз выступать связывателями анафоров (4a), контролировать PRO деепричастных оборотов (4b) и плавающие определители (4c). С другой стороны, возможность контроля плавающих определителей ограничена причастными оборотами в составе номинативных именных групп (подлежащих и именных сказуемых), ср. (4d). В этой связи диагностики внутренней структуры причастных клауз в русском языке неоднозначны.

- (4) а. ...Персикову прислали три посылки, [\emptyset_i содержащие в себе; зеркала]... [НКРЯ].
 б. По делу проходит сахалинский предприниматель, ... [\emptyset_i активно сотрудничающий со следствием, [PRO_i сдавая подельников]] [НКРЯ].
 в. Чем могли помочь ей эти юноши, ... [\emptyset_i едва сами_i спасшиеся от смерти]? [НКРЯ]
 д. *²Как она могла помочь этим юношам, [едва сами / самим спасшимся от смерти]?

Таким образом, мы исключаем из рассмотрения номинализованные и причастные клаузы и сосредоточиваемся на оставшихся типах нефинитных клауз: инфинитивных, деепричастных и малых.

Инфинитивные клаузы в русском языке используются в различных контекстах и, по-видимому, обладают разными структурными характеристиками. Можно выделить по меньшей мере следующие типы: актантные инфинитивные клаузы, выражающие сентенциальную валентность матричного предиката, сирконстантные инфинитивные клаузы, выражающие обстоятельственное значение цели, а также инфинитивные группы в составе аналитического будущего времени, фазовых и модальных конструкций. Из последних для текущего обсуждения представляют интерес эпистемические модальные конструкции, в которых инфинитивный оборот может содержать глагол-связку или вспомогательный глагол *быть*. Прочие модальные и фазовые конструкции, по-видимому, содержат глагольные проекции более низкого уровня и нерелевантны для текущего обсуждения; подробнее об их структуре см. [Лютикова 2022a].

Инфинитивные конструкции с модальными предикатами, выражающими «внешнюю» модальность, обнаруживают признаки реструктурирования (моноклаузальной структуры). Инфинитивный оборот в таком случае соответствует глагольной предикативной составляющей, включающей как минимум AspP; модальный предикат выступает как функциональная вершина высокого уровня. Важно подчеркнуть, что предикаты, выражающие внешнюю модальность, не проецируют собственных аргументов, и подлежащим в соответствующих конструкциях становится аргумент инфинитивного оборота, как это показано в (5a)–(5b). Таким образом, пустая категория при инфинитиве является следом от А-передвижения именной группы в позицию подлежащего, аналогичного тому, которое происходит в финитной клаузе без модального предиката (5c).

- (5) а. Я думаю, что мы можем задохнуться быстрее, чем помрем с голоду [НКРЯ].
 б. [_{TP} Мы_i можем [_{AspP} [_{VP} t_i задохнуться]]].
 в. [_{TP} Мы_i будем [_{VP} t_i задышаться]]].

² Отдельно следует упомянуть использование причастных клауз в конструкциях с предикатами подъема, см. [Лютикова 2022б].

Актантные инфинитивные клаузы полной структуры в русском языке используются исключительно в конфигурациях контроля (подъем всегда сопровождается реструктурированием, выражающимся в сокращении функциональной структуры инфинитивной и/или матричной клаузы³). Считается, что контроль — это семантико-синтаксическое отношение между именной группой-контролером и нулевой единицей (PRO) в позиции подлежащего нефинитной клаузы [Landau 2013]⁴. В литературе неоднократно отмечалось, что тип контроля коррелирует со структурными характеристиками инфинитивной клаузы [Козинский 1983; 1985; Babby 1998; Landau 2008; Lyutikova, Gerasimova 2023]. Конфигурации субъектного контроля (6a) противопоставляются конфигурациям объектного (6b) и произвольного (6c) контроля либо в отношении размера (VP/TP vs. CP), либо в отношении формальных признаков комплементаризера и его способности опосредовать согласование между контролером и контролируемой нулевой единицей — PRO инфинитивной клаузы.

- (6) а. Она_i ... явно хотела [PRO_i ей понравиться] [НКРЯ].
 б. Голос у меня был сильный, и однокурсники часто просили меня, [PRO_i что-нибудь спеть] [НКРЯ].
 в. Люди часто сбиваются с пути, важно [PRO_{ARB} вовремя найти правильный разворот] [НКРЯ].

Целевые инфинитивные клаузы можно разделить на два типа: простые целевые инфинитивные клаузы при глаголах движения и каузации движения (7a) и целевые инфинитивные клаузы с подчинительным союзом *чтобы* (7b). Клаузы первого типа обнаруживают свойства, сходные со свойствами актантных инфинитивных клауз субъектного и объектного контроля соответственно⁵. Клаузы второго типа, наряду с агентивно контролируемым PRO [Козинский 1983; Тестелец 2001: 326–328], могут содержать и выраженное подлежащее в дательном падеже.

- (7) а. Нас_i направили во Львов [PRO_i обеспечить избирательную кампанию] [НКРЯ].
 б. Как ты и хотела, я_i предпринял некоторые шаги, [чтобы (нам)/PRO_i перебраться все-таки в Москву] [НКРЯ].

Деепричастные клаузы также реализуют конфигурацию контроля: они содержат невыраженное подлежащее (PRO), которое нормативно контролируется подлежащим главной клаузы (8a) [Babby 1979; Ицкович 1982; Козинский 1983; Babby, Franks 1998; Тестелец 2001]. Однако известны и случаи субстандартного контроля PRO деепричастного оборота со стороны логического субъекта главной клаузы, не совпадающего с синтаксическим подлежащим (8b)–(8c) [Ицкович 1982; Йокояма 1983; Козинский 1983; Гловинская 1996; Онипенко, Биккулова 2016]. В современном русском языке такие употребления также встречаются, но признаются не отвечающими норме [РГ 1980: §2105].

- (8) а. [PRO_i Услышав мое предположение], дедушка_i засмеялся, но я все-таки попросил его быть настороже [НКРЯ].

³ Отсутствие подъема из инфинитивных клауз полной структуры в русском языке аргументируется в [Лютикова 2022a; 2022b]. Другие инфинитивные конструкции, предположительно использующие подъем дативного подлежащего, включают синтаксические идиомы с отрицательными и неопределенными местоимениями типа *Здесь неоткуда взяться деревьям*, *Здесь есть откуда взяться деревьям*, см. [Zimmerling 2024].

⁴ Об альтернативном подходе к анализу контроля через передвижение см. в разделе 7.

⁵ Подобное поведение простых целевых инфинитивных клауз при глаголах движения и каузации движения не уникально, ср., например, [Wurmbrand 2001] о немецком и [Lyutikova, Sideltsev 2021] о хеттском.

- b. [PRO_i Читая эти странные стихи], вам_i как бы слышится дух веков реформации [Ф. М. Достоевский; цит. по [Онипенко, Биккулова 2016]].
- c. [PRO_i Перечитывая «Наложницу»], меня_i всегда поражает легкость и верность ее слога [Е. А. Боратынский; цит. по [Онипенко, Биккулова 2016]].

Итак, инфинитивные клаузы в конфигурациях контроля и деепричастные клаузы содержат нулевое подлежащее — PRO. Важной характеристикой PRO является его падеж, который может быть выявлен на основании допущения, что так называемые плавающие определители — лексемы *сам, один, оба, все* — отражают падеж своего контролера, в том числе и нулевого [Comrie 1974; Babby 1998]. Эта диагностика позволяет определить падеж PRO в различных конфигурациях.

Актантные инфинитивные клаузы с субъектным контролем содержат PRO в именительном падеже (9а). При объектном и произвольном контроле плавающие определители указывают на дательный падеж PRO (9b)–(9c). Аналогично ведут себя простые целевые инфинитивные клаузы (9d)–(9e). Целевые инфинитивные клаузы с подчинительным союзом *чтобы* демонстрируют PRO в дательном падеже (9f), а при субъектном контроле — также и в именительном (9g).

- (9) a. Поэтому каждый артист_i хочет [PRO_i **сам** ставить], [PRO_i **сам** дергать за ниточки] [НКРЯ].
- b. Кутахов просил меня_i [PRO_i решать **самому** все вопросы по космосу] [НКРЯ].
- c. Снова практика привела меня к заключению, что в пьесах общественно-политического значения особенно важно [PRO_{ARB} **самому** зажить мыслями и чувствами роли], и тогда сама собой передастся тенденция пьесы [НКРЯ].
- d. Поэтому я_i пошла [PRO_i всё делать **сама** и набивать шишки] [НКРЯ].
- e. Серега попросил чаю, Пашута отправила его_i в кухню [PRO_i распоряжаться **самому**] [НКРЯ].
- f. Я_i очень осторожно переступал ногами, [чтобы PRO_i не споткнуться **самому**] [НКРЯ].
- g. Те_i достаточно велики, [чтобы PRO_i идти **сами**], но она несет их на руках [НКРЯ].

Итак, PRO в инфинитивных клаузах имеет признак именительного падежа в случае контроля со стороны номинативного подлежащего и дательного падежа во всех остальных случаях. Представляется разумным считать, вслед за [Babby 1998; Тестелец 2001; Landau 2008], что признак дательного падежа PRO в инфинитивном обороте получает по умолчанию от невыраженного или выраженного (*чтобы*) комплементаризера инфинитивного оборота. В случае же субъектного контроля PRO копирует признак именительного падежа своего контролера^{6,7}.

⁶ Более детальный анализ падежных характеристик PRO в зависимости от конфигурации контроля представлен в [Babby 1998; Тестелец 2001; Landau 2008; Лютикова 2010; Байков 2020]. В настоящей работе мы отвлекаемся от маргинальных случаев.

⁷ Анонимный рецензент статьи в этой связи ставит следующие вопросы: (i) в какой мере автору статьи нужен тезис о падежном PRO, не достаточно ли того, что сам факт наличия PRO в нефинитной клаузе с синтаксическим контролем блокирует выбор именительного падежа у именного компонента сказуемого; (ii) требуется ли для представленного в статье анализа тезис о том, что PRO получает в русском языке дательный падеж именно от комплементаризера, или это просто допущение, отражающее теоретические вкусы автора. Тезис о наличии падежа у PRO необходим ровно в той мере, в какой анализ призван не только объяснять запрет на лицензирование номинатива (с этим справлялись бы и теории беспадежного PRO и PRO, обладающего «нулевым падежом», ср. [Chomsky, Lasnik 1993]), но и предсказывать дательный (и реже — винительный) падеж у адъективных компонентов сказуемого. Источником датива рационально считать

Нулевое подлежащее деепричастных клауз при нормативном субъектном контроле имеет формальный признак именительного падежа (10а). В случае субстандартного контроля плавающие определители в деепричастном обороте затруднены (10б), что, предположительно, может свидетельствовать об отсутствии у PRO в таких случаях падежного признака.

- (10) а. Раньше, [PRO_i **сама** того не замечая], я_i жила в мире улыбок [НКРЯ].
 б. *³[PRO_i **Оба / обоим** читая эти странные стихи], друзьям_i как бы слышался дух веков реформации.

Обратимся теперь к конфигурациям с малыми клаузами — предикативными составляющими, лишенными глагола и неспособными самостоятельно выражать предикативные категории финитной клаузы. В работе [Лютикова 2022б] малые клаузы усматриваются в русских конструкциях с составным именным сказуемым, возглавляемых ползузнаменательными глаголами *становиться, делаться, считаться, оказаться, получиться, представляться* и т. п. (11а)–(11с) и структурно аналогичных им конструкциях с переходными глаголами мнения, восприятия, каузации, включающих присказуемые неглагольные составляющие (11d)–(11f) (см. также [Nichols 1981] о классах глаголов, сочетающихся с присказуемыми именами).

- (11) а. На фронте пацаны-подростки **становились** сыновьями полка [НКРЯ].
 б. Рана Швабрина **оказалась** не смертельна [А. С. Пушкин. Капитанская дочка (1836)].
 с. Памятник **считался** легендарным, установлен был более 50 лет назад на добровольные пожертвования комсомольцев [НКРЯ].
 д. Несмотря на выгоды, дворяне гнушались службою стрелецкою и **считали** оную пятном для своего рода [А. С. Пушкин. Заметка о русской истории (1833–1834)].
 е. Я старался вообразить себе капитана Миронова, моего будущего начальника, и **представлял** его строгим, сердитым стариком... [А. С. Пушкин. Капитанская дочка (1836)].
 ф. 28 июля они протестовали против избрания Станислава, **объявив** участников в оном врагами отечества [А. С. Пушкин. История Петра: Подготовительные тексты (1835–1836)].

Предлагаемый в [Лютикова 2022б] анализ конструкций с ползузнаменательными переходными и переходными глаголами как конструкций с подъемом из малых клауз подразумевает, что соответствующие глаголы обладают синтаксическими селективными признаками, позволяющими им принимать в качестве комплемента малые клаузы. Подлежащее малой клаузы, как и подлежащее глагольной области, нуждается в падежном лицензировании и претерпевает подъем в главную клаузу, в позицию подлежащего (12) или дополнения (13).

комплементаризер, поскольку именно его признаки «отвечают» за типирование инфинитивной клаузы как независимой, актантной или целевой. Альтернативный вариант, предлагаемый в [Sigurdsson 2008], — лицензирование падежного PRO нефинитной предикативной вершиной с определенными признаками. Поскольку между комплементаризером и предикативной вершиной T устанавливаются отношения селекции, выбор между двумя вариантами не принципиален. Безусловно, соображения такого рода релевантны только в том случае, если источниками структурного падежа являются функциональные вершины. В конфигурационной модели, допускающей дефолтные падежи для различных падежных областей, датив может рассматриваться как дефолтный падеж инфинитивной клаузы.

(12) Географический центр_i оказался [_{sc} *t_i* далеко за МКАД].

(13) Дворяне считали службу_i [_{sc} *t_i* пятном для своего рода].

Таким образом, пустая категория в позиции подлежащего малой клаузы имеет иную природу, чем подлежащее инфинитивных или деепричастных клауз: это не PRO, а след от передвижения именной группы (А-передвижения) в падежную позицию, аналогичный пустой категории в инфинитивных оборотах с эпистемическими модальными предикатами, представленных выше в этом разделе. В следующем разделе мы рассмотрим ограничения на использование адъективных сказуемых в нефинитных клаузах разных типов.

4. Адъективные сказуемые в нефинитных клаузах

В предыдущем разделе мы выделили два вида нефинитных клауз: клаузы, содержащие А-след (модальные конструкции с инфинитивом, конструкции с подъемом из малых клауз), и клаузы, содержащие PRO (актантные и сирконстантные инфинитивные клаузы, деепричастные клаузы). Сейчас мы исследуем допустимость разных классов адъективных сказуемых в нефинитных клаузах и выявим факторы, влияющие на их дистрибуцию.

4.1. Краткое прилагательное

Краткие прилагательные в позиции сказуемого допустимы в следующих типах конфигураций: модальные конструкции с инфинитивом (14а), конструкции с подъемом из малых клауз в позицию подлежащего (14b), актантные инфинитивные клаузы субъектного контроля (14с), деепричастные клаузы с каноническим контролером (14d). Кроме того, краткие прилагательные встречаются в позиции сказуемых целевых инфинитивных клауз с нулевым подлежащим, контролируемым номинативным подлежащим главной клаузы (14е)⁸.

- (14) а. Ученые из Карелии считают, что [отдых на Ладожском озере]_i может [*t_i* быть **опа-сен** для здоровья человека из-за пыли с каменных карьеров] [НКРЯ].
- б. Информация_i оказалась [*t_i* **доступна** и его знакомым] [НКРЯ].
- с. Угнетенный_i хочет [PRO_i быть **похож** на угнетателя], раб — на хозяина, слабейший противник — на сильнейшего [НКРЯ].
- д. Впрочем, Леха, сам, [PRO_i будучи **любопытен** и **неуклюж**], повалил стеллаж с хозяйственным товаром... [НКРЯ].
- е. ³Я_i ем, [чтобы PRO_i быть **счастлив**] [<https://www.meme-arsenal.com/create/meme/8912010>].

Краткие прилагательные неграмматичны в качестве сказуемых малых клауз при подъеме в позицию дополнения (15а), в актантных инфинитивах объектного контроля (15b), в целевом инфинитивном обороте с выраженным подлежащим или с нулевым подлежащим, контролируемым не-субъектом (15с), в деепричастных оборотах с субстандартным контролером (15d).

⁸ Подобные примеры отсутствуют в основном подкорпусе НКРЯ, но обнаруживаются в русскоязычном сегменте сети.

- (15) а. *Он считал эту информацию_i [*t_i доступна* и его знакомым].
 б. *Можно посоветовать угнетенному_i [*PRO_i не быть похож на угнетателя*].
 в. *Мне_i нужно немного, [чтобы *PRO_i быть счастлив*].
 д. **[PRO_i Будучи любопытен и неуклюж]*, Леха_i было неловко.

Таким образом, предварительное обобщение выглядит следующим образом: краткие прилагательные в качестве сказуемых допустимы в таких клаузах, подлежащее которых получает номинатив либо вследствие передвижения в позицию подлежащего финитной клаузы, либо благодаря контролю со стороны подлежащего финитной клаузы.

4.2. Полное прилагательное в именительном падеже

Полные прилагательные в именительном падеже в позиции сказуемого допустимы только в следующих типах конфигураций: модальные конструкции с инфинитивом (16а) и конструкции с подъемом из малых клауз в позицию подлежащего (16б).

- (16) а. А жизнь_i может [*t_i быть долгая, трудная*], когда человек ни о чем не думает, что относится к нему, а только о возможности служить другому человеку, другим людям [НКРЯ].
 б. Комната_i оказалась [*t_i небольшая, с высоким, смугло задымленным потолком*] [НКРЯ].

Все прочие типы конфигураций исключают использование полной формы прилагательного в именительном падеже в качестве сказуемого. Так, недопустимы соответствующие предложения с подъемом из малых клауз в позицию дополнения (17а), актантным инфинитивом субъектного и объектного контроля (17б), целевым инфинитивным оборотом (17с) и деепричастным оборотом (17д)⁹.

- (17) а. *Я представлял его_i [*t_i строгий и сердитый*].
 б. **[Шумеры и аккадцы]_i хотели / научили нас_j [PRO_{i,j} быть уверенные в завтрашнем дне]*.
 в. *Я_i отправилась в Москву, [чтобы *PRO_i быть готовая* подать документы].
 д. **[PRO_i Будучи любопытный и неуклюжий]*, Леха_i повалил стеллаж с хозяйственным товаром.

Обобщение о дистрибуции полных прилагательных в форме именительного падежа в позиции сказуемого может быть сформулировано так: эти формы возможны только в нефинитных клаузах, подлежащее которых поднимается в позицию номинативного подлежащего.

4.3. Полное прилагательное в творительном падеже

Полное прилагательное в форме творительного падежа в позиции сказуемого допустимо во всех нефинитных клаузах: в конфигурациях подъема (18а)–(18с) и конфигурациях

⁹ В основном подкорпусе НКРЯ обнаружен единственный пример такого рода, который, однако, представляется сомнительным:

(i) Полученные результаты показали, что приемник, будучи достаточно **универсальный**, вполне соответствует своему назначению... [НКРЯ].

(18) a. И легкая музыка_i может [_{t_i} быть **изумительной**], и серьезная_i может [_{t_i} быть **непотребной**] [НКРЯ].

b. Так что [встреча с уральскими специалистами]_i оказалась [_{t_i} весьма **своевременной**].

c. Для простоты будем пока считать их_i [_{t_i} абсолютно **равноправными**] [НКРЯ].

d. [Российское правительство]_i хочет [PRO_i быть **прозрачным**] [НКРЯ].

e. Прошу всех_i [PRO_i быть **готовыми** к обеду и не опаздывать в столовую в четырнадцать ноль-ноль] [НКРЯ].

f. Мы даже не знаем, были ли эти кружки_i с самого начала марксистскими или постепенно эволюционировали в этом направлении, [PRO_i изначально будучи скорее **народническими**] [НКРЯ].

g. Греки верили, что мы_i живем, [чтобы PRO_i быть **счастливыми**] [НКРЯ].

h. [?][PRO_i Будучи уже вполне **здоровым**], меня_i по-прежнему пугала мысль об осложнениях.

В таблице 1 представлены обобщения о дистрибуции адъективных сказуемых в нефинитных клаузах по сравнению с финитными. Мы видим, что все три класса адъективных сказуемых ведут себя по-разному: наиболее ограниченную дистрибуцию имеют полные прилагательные в именительном падеже, наиболее широкую — полные прилагательные в творительном падеже, а краткие прилагательные занимают промежуточную позицию.

Дистрибуция адъективных сказуемых в финитных и нефинитных клаузах: контексты

[illegible]

Легко видеть, что контексты для обоих классов адъективов с ограниченной дистрибуцией (для краткого прилагательного и для полного прилагательного в именительном падеже) определяются совокупностью значений двух параметров: типа пустой категории в позиции подлежащего нефинитной клаузы (след от А-передвижения vs. контролируемое PRO) и падежа пустой категории (номинатива при передвижении в позицию подлежащего или при номинативном контролере vs. косвенных падежей или отсутствия падежа). Соответственно, мы можем переформулировать ограничения на дистрибуцию адъективных сказуемых в виде признаковых ограничений на подлежащее клаузы (таблица 2).

Таблица 2

Дистрибуция адъективных сказуемых в финитных и нефинитных клаузах:
признаки подлежащего

Падеж	Подлежащее клаузы		
	выраженная ИГ	А-след	PRO
НОМ	каноническое финитное подлежащее	модальные конструкции с инфинитивом, подъем из малой клаузы в позицию подлежащего	инфинитив и деепричастие субъектного контроля
	функциональное подлежащее (напр., DAT при инфинитиве в независимом инфинитивном предложении или целевой клаузе)	подъем из малой клаузы в позицию дополнения	инфинитив объектного контроля, деепричастие с субстандартным контролером

прилагательное, п. ф.: INSTR

прилагательное, кр. ф.

прилагательное, п. ф.: НОМ

Итак, полное прилагательное в творительном падеже доступно в предикативной позиции любых нефинитных клауз, независимо от типа подлежащего и его падежных характеристик. Краткое прилагательное требует наличия подлежащего в именительном падеже, при этом его тип и, как следствие, способ получения падежного признака не имеет значения. Наконец, полное прилагательное в именительном падеже возможно только в таких клаузах, подлежащее которых получает номинатив вследствие А-передвижения.

5. Причастные пассивы в нефинитных клаузах

В этом разделе мы обсудим использование форм кратких и полных страдательных причастий в пассивах и сравним их дистрибуцию с дистрибуцией соответствующих форм прилагательного.

5.1. Краткое страдательное причастие

Выражение сказуемого краткой формой страдательного причастия допустимо в тех же контекстах, где может использоваться форма краткого прилагательного. Это модальные конструкции с инфинитивом (19a), конструкции с подъемом из малых клауз в позицию подлежащего (19b), актантные и целевые инфинитивные клаузы субъектного контроля (19c)–(19d), деепричастные клаузы с каноническим контролером (19e).

- (19) а. По требованию Банка настоящий Договор_i может [_i быть **расторгнут** судом при отсутствии операций по Счету в течение одного года] [НКРЯ].
- б. Вот и [тематика Года культуры в России]_i оказалась [_i быстро **сведена** к деньгам] [НКРЯ].
- с. Может быть, я_i хочу [PRO_i быть **убит**]... [НКРЯ].
- д. Речь_i рассчитана на то, [чтобы PRO_i быть **воспринимаема** не зрением, а слухом] [НКРЯ].
- е. [Кассета с лекцией о Вавилове]_i дошла до меня, [PRO_i будучи **передана** через многие руки] [НКРЯ].

Неноминативные подлежащие (выраженные, А-следы или PRO) при кратком страдательном причастии в позиции сказуемого недопустимы, ср. (20a)–(20d).

- (20) а. *Желаю вам_i [PRO_i быть **убит**].
- б. *Я предпринял некоторые шаги, [чтобы нам быть **прописаны** в Москве].
- с. *[PRO_i Будучи быстро **госпитализирован**], больного_i перестало тошнить.
- д. *Я считал его_i [_i **уволен** в прошлый понедельник].

5.2. Полное страдательное причастие

Дистрибуция форм полных страдательных причастий в предикативной позиции существенно отличается от дистрибуции полных прилагательных. Во-первых, полные страдательные причастия в именительном падеже не используются для образования причастного пассива не только в финитных клаузах, но и в нефинитных. Соответственно, в контекстах модальных конструкций с инфинитивами и подъема из малых клауз в позицию подлежащего возможны формы полных прилагательных (16a)–(16b), но не полных страдательных причастий в именительном падеже (21a)–(21b).

- (21) а. *По требованию банка договор, может [_i быть **расторгнутый** судом].
- б. *[Тематика Года культуры в России]_i оказалась [_i быстро **сведенная** к деньгам].

Во-вторых, несмотря на то, что использование полной формы страдательного причастия в творительном падеже в составе финитной клаузы не интерпретируется как пассив (ср. (2b) и (3a)), та же форма в нефинитных клаузах (22a)–(22g) допускает собственно пассивную интерпретацию, что поддерживается присоединением агентивных дополнений.

- (22) а. Человек_i может [_t_i быть **измороженным** судьбой], но для чего-то это все-таки дается... [Юрий Любимов. Мне чужд казенный пафос (2004) // «Театральная жизнь», 28.06.2004].
- б. «Роман о розе» придал всей этой культуре форму столь красочную, столь изощренную, столь богатую, что *pro*_i сделался [_t_i **читаемым** и **почитаемым** всеми (кто мог читать и почитать)] [НКРЯ].
- с. Он испугался и поспешно начал отступать, полагая себя_i [_t_i **отрезанным** от Казани] [НКРЯ].
- д. Глеб_i вывалился наружу, [*PRO*_i рискуя [*PRO*_i быть **сбитым** той машиной, что налетит сзади]] [НКРЯ].
- е. То есть не то чтобы он всплыл, он, собственно, и не погружался, он все время был у поверхности, просто теперь он позволил себе_i [*PRO*_i быть **увиденным**] [НКРЯ].
- ф. Впрочем, из автобуса *pro*_i не выходят, [чтобы *PRO*_i не быть **покусанными** гадюками] [НКРЯ].
- г. [*PRO*_i Будучи **допрошенной** по этому поводу], Якушева_i не отрицала факт изъятия ею денег из сумки Авраменко [НКРЯ].

Таким образом, в дистрибуции причастных пассивов наблюдаются определенные отличия от дистрибуции других адъективных сказуемых. Они обобщаются в таблице 3.

Таблица 3

Дистрибуция адъективных сказуемых и причастных пассивов в финитных и нефинитных клаузах

	финитная клауза	модальная конструкция с инфинитивом	малая клауза, подъем в позицию подлежащего	малая клауза, подъем в позицию дополнения	актантный инфинитив, субъектный контроль	актантный инфинитив, объектный контроль	целевой инфинитив, субъектный контроль	целевой инфинитив, не-субъектный контроль	депричастный оборот, стандартный контроль	депричастный оборот, субстандартный контроль
прилагательное, кр. ф.	+	+	+	–	+	–	+	–	+	–
причастный пассив, кр. ф.	+	+	+	–	+	–	+	–	+	–
прилагательное, п. ф.: NOM	+	+	+	–	–	–	–	–	–	–
причастный пассив, п. ф.: NOM	–	–	–	–	–	–	–	–	–	–
прилагательное, п. ф.: INSTR	+	+	+	+	+	+	+	+	+	+
причастный пассив, п. ф.: INSTR	–	+	+	+	+	+	+	+	+	+

Следовательно, при анализе дистрибуции форм причастного пассива можно опираться на единообразное поведение краткой формы прилагательных и страдательных причастий, но при этом необходимо дополнительно объяснить, почему полные формы причастия

невозможны в финитной клаузе с каноническим подлежащим и почему они возникают в нефинитных контекстах.

6. Обобщения и анализ

Итак, мы установили, что адъективные сказуемые имеют разные модели дистрибуции. Наиболее ограниченной в предикативной позиции является полная форма прилагательного в именительном падеже: она используется только в финитных клаузах с каноническим номинативным подлежащим и в клаузальных проекциях низкого уровня (в инфинитивных оборотах при модальных предикатах и в малых клаузах) при подъеме их подлежащего в позицию канонического номинативного подлежащего. Краткая форма прилагательного имеет более широкую дистрибуцию: кроме перечисленных выше случаев, она используется также в актантных и сирконстантных нефинитных клаузах, содержащих PRO, контролируемое каноническим номинативным подлежащим. Наконец, полная форма прилагательного в творительном падеже не ограничена в дистрибуции и встречается в предикативной позиции во всех рассмотренных контекстах.

Краткая форма страдательного причастия встречается в тех же контекстах, что и краткая форма прилагательного, и во всех случаях допускает собственно пассивную интерпретацию. Полная форма страдательного причастия в именительном падеже не используется в функции пассива ни в одном из типов клауз, а в творительном падеже — используется во всех нефинитных клаузах и в финитных клаузах с неканоническим подлежащим (дательное подлежащее в инфинитивных предложениях).

Объяснение дистрибуции адъективных сказуемых, предлагаемое в этой статье, является структурным и опирается на совокупность гипотез, независимо предложенных в формальной синтаксической литературе. Во-первых, это гипотеза о категориальном контрасте полной и краткой форм прилагательных и причастий [Babby 1973; Гращенков 2018; 2019; 2022], согласно которой группы краткой и полной форм представляют собой проекции разных синтаксических категорий и, соответственно, могут обладать разными наборами формальных признаков. Во-вторых, это гипотеза о противопоставлении контроля и подъема ([Rosenbaum 1965; Postal 1974; Chomsky 1981; Lasnik, Saito 1991]. [Лютикова 2022а; 2022б] для русского) как двух альтернативных друг другу способов организации референциальной зависимости между аргументами главной и зависимой клаузы в полипредикативной конструкции. В-третьих, это гипотеза о способности контролируемого PRO обладать формальным признаком падежа и означивать его разными способами в нефинитных клаузах разных типов [Babby 1998; Landau 2008; Lyutikova, Gerasimova 2023]. Данная гипотеза предполагает, что PRO может получать структурный дательный падеж от комплементаризеров либо копировать структурный номинатив (реже аккузатив) от своего контролера [Тестелец 2001: 295–305].

Ниже мы рассмотрим структурные характеристики, формальные признаки и лицензирование разных типов адъективных составляющих при опоре на эти гипотезы.

6.1. Краткая форма прилагательного

В данной работе принимается допущение, что краткое прилагательное — это лексическая категория, которая способна к проецированию собственного подлежащего [Geist 2010; Гращенков 2018; 2022]. Группа краткого прилагательного, AP, таким образом, содержит вершину A, ее лексические дополнения (при наличии) и подлежащее — именную группу или PRO, как это представлено в (23). Согласование краткого прилагательного

по роду и числу со своим подлежащим является локальным и происходит в конфигурации спецификатор-вершина.

- (23) Краткая форма прилагательного
[_{AP} DP/PRO [_{A'} A XP]]

Ограничения на дистрибуцию краткой формы прилагательного определяются тем, что AP нуждается в синтаксическом лицензировании. Лицензировать AP могут две конфигурации: либо непосредственное вложение AP в группу полной формы, *adjP*, возглавляемую функциональной вершиной *adj* (о ее структуре и функциях см. раздел 6.2), либо получение номинатива ее подлежащим. Последнее требование представляется легко согласуемым с историей кратких форм прилагательных, которые до утраты способности к склонению выступали в предикативной позиции в именительном падеже. Механизмы получения номинатива различны для подлежащих, выраженных именной группой и подлежащих, выраженных PRO.

Выраженное подлежащее должно получить номинатив как финитное подлежащее в результате взаимодействия с финитным T и вызванного этим взаимодействием A-передвижения, как это показано в (24).

- (24) а. Краткая форма в финитной клаузе

Маша была готова отвечать.

[_{TP} DP [_{NOMj}] T_{+FIN} [_{AuxP} Aux_{<быть>} [_{AP} t_i [_{A'} A XP]]]]

- б. Краткая форма в модальной конструкции с инфинитивом

Маша может быть готова отвечать.

[_{TP} DP [_{NOMj}] T_{+FIN} [_{ModP} Mod_{<мочь>} [_{AuxP} Aux_{<быть>} [_{AP} t_i [_{A'} A XP]]]]]

- с. Краткая форма при субъектном подъеме из малой клаузы

Маша оказалась готова отвечать.

[_{TP} DP [_{NOMj}] T_{+FIN} [_{vP} v_{intr} [_{VP} V_{<оказаться>} [_{AP} t_i [_{A'} A XP]]]]]

Нулевое подлежащее AP — PRO — не может быть подлежащим финитной клаузы, поскольку лицензируется нефинитным T [Sigurdsson 2002; 2008]. Следовательно, PRO не может получить номинатив тем способом, которым его получают финитные подлежащие — через согласование по ϕ -признакам с финитным T. Альтернативный способ получения структурного номинатива, доступный для контролируемого PRO, — согласование по падежу¹⁰ (или, скорее, унификация падежного признака) с контролером. Не вдаваясь в тонкости механизма такого согласования (возможные варианты включают, например, опосредование согласования падежным признаком комплементаризера [Landau 2008] или удаление оболочки CP зависимой клаузы как источника падежа [Lyutikova, Gerasimova 2023]), мы можем заключить, что PRO получает номинатив тогда и только тогда, когда его выраженный контролер — каноническое номинативное подлежащее. Лицензирование краткой формы прилагательного при субъектном контроле показано в (25).

- (25) Краткая форма при субъектном контроле

Маша хочет быть похожа на мать.

Будучи внешне похожа на мать, Маша унаследовала ум отца.

Маша занимается, чтобы быть здорова.

DP [_{NOMj}] ... [_{CP} C [_{TP} PRO [_{NOMj}] T_{-FIN} [_{AuxP} Aux [_{AP} t_i [_{A'} A XP]]]]]

¹⁰ Представление об унификации признаков падежа у двух различных именных составляющих как о согласовании сталкивается с определенными теоретическими осложнениями. В синтаксической литературе эксплуатируются различные механизмы операции согласования, например, разделение признаков (Feature sharing [Pesetsky, Torrego 2007]) или изменение направления согласования (Upward agree [Zeijlstra 2012]), чтобы избежать такого теоретического конструкта, как согласование по падежу.

6.2. Полная форма прилагательного

Полная форма прилагательного представляет собой проекцию атрибутивной вершины *adj*, которая деривирует полные прилагательные из АР (образуя качественные прилагательные) и других лексических категорий (образуя относительные прилагательные) [Гращенков 2018]. Соединение АР с вершиной *adj* лицензирует АР и приводит к изменениям аргументной структуры, признаковой матрицы и интерпретации.

Во-первых, *adjP*, в отличие от АР, не может иметь собственное подлежащее, поскольку неспособно приписать ему тета-роль (такое же различие между краткой и полной формой проводится в [Гращенков 2019]). Полагаем, что наследовать тематическое подлежащее от АР вершина *adj* также неспособна, а позиция ее спецификатора — это *A'*-позиция. Функция вершины *adj* аналогична функции комплементаризера в относительной клаузе¹¹: она притягивает подлежащее АР и создает условия для предикатной абстракции. Вследствие этого полные прилагательные интерпретируются как предикаты над индивидами и могут выступать в атрибутивной позиции¹².

Во-вторых, *adjP*, в отличие от АР, обладает формальным признаком падежа (ср. аналогичное признаковое различие, предлагаемое в [Babby 1973]). Признак падежа полного прилагательного может получить значение либо при конкорде — атрибутивном согласовании в именной группе, либо в предикативной конфигурации, которая обсуждается в разделе 6.3.

Структура полной формы прилагательного на основе АР представлена в (26).

- (26) Полная форма прилагательного
 $[_{adjP} Op_i adj_{[uCase]} [_{AP} t_i [_{A'} A XP]]]$

6.3. Падежные формы полного прилагательного

Будучи предикатом над индивидами, группа полного прилагательного нуждается в подлежащем, чтобы создать пропозицию. Возможность получить тематическое подлежащее создает предикативная вершина *Pred*, принимающая *adjP* в качестве комплемента и именную группу в качестве спецификатора [Bowers 1993; den Dikken 2006]. Подчеркнем, что основной функцией вершины *Pred* является именно лицензирование субъектно-предикатной структуры; эта функция независима от ее участия в приписывании творительного предикативного.

- (27) Полная форма прилагательного в составе *PredP*
 $[_{PredP} DP/PRO [_{Pred'} Pred [_{adjP} Op_i adj_{[uCase]} [_{AP} t_i [_{A'} A XP]]]]]$

¹¹ Ср. с идеей Л. Бэбби о деривации полных форм как редуцированных относительных предложений со сказуемым — краткой формой прилагательного [Babby 1973].

¹² Вопрос о том, почему абстракция происходит только по переменной, соответствующей подлежащему АР, достаточно сложен. Для объяснения этого факта можно считать, что вершина *adj* выступает как смешанный зонд, обладающий одновременно *φ*-признаками и *w/h*-признаками. Отметим, что это предположение открывает возможность объединения в естественном классе *adj* и финитного *T* как лицензоров АР. Отдельный вопрос состоит в том, почему в таких конфигурациях может выступать только нулевой оператор, но не выраженная *w/h*-проформа или группа. Эти вопросы мы вынуждены пока оставить без окончательных ответов; следует заметить, однако, что при анализе материала разноструктурных языков исследователи зачастую испытывают потребность в постулировании функциональных вершин, вызывающих предикатную абстракцию на уровне существенно ниже *TP/CP* (ср. [Harley 2020; Лютикова, Сидельцев 2020]).

В структуре (27) есть две составляющих, нуждающихся в падежном лицензировании. Это *adjP*, получившая неозначенный формальный признак падежа от вершины *adj*, создающей полную форму, и *DP* в позиции подлежащего *PredP*. Полагаем, что существует две стратегии падежного лицензирования этих составляющих: совместное лицензирование, когда падеж присваивается одновременно подлежащему и предикату *PredP*, и независимое лицензирование, когда подлежащее и предикат *PredP* получают падеж в результате двух разных операций.

При совместном падежном лицензировании структурный номинатив приписывается всей *PredP* и затем просачивается внутрь, достигая ее субъекта и предиката (ср. с идеей просачивания падежа в [Babby 1987]). В результате номинатив полной формы прилагательного в позиции сказуемого возможен только в том случае, когда подлежащее *PredP* также получает номинатив. Таким образом нам удастся проинтерпретировать различия в условиях лицензирования краткой формы и полной формы в именительном падеже. Первая требует, чтобы ее подлежащее получило номинатив, но получение номинатива подлежащим может произойти существенно позже, при унификации падежа PRO с падежом контролера из главной клаузы. Вторая же лицензируется только одновременно с номинативом подлежащего¹³. Деривация номинативной полной формы прилагательного в позиции сказуемого показана в (28) для различных лицензирующих ее конфигураций.

(28) а. Номинатив полной формы в финитной клаузе

Комната была небольшая.

$$[_{TP} DP_{[NOM]i} T_{+FIN} [_{AuxP} Aux_{<БЫТЬ>} [_{PredP} t_i [_{Pred'} Pred [_{adjP} Op_i adj_{[NOM]} [_{AP} t_i [_{A'} A XP]]]]]_{[NOM]j}]$$

б. Краткая форма в модальной конструкции с инфинитивом

Комната может быть небольшая.

$$[_{TP} DP_{[NOM]i} T_{+FIN} [_{ModP} Mod_{<МОЧЬ>} [_{AuxP} Aux_{<БЫТЬ>} [_{PredP} t_i [_{Pred'} Pred [_{adjP} Op_i adj_{[NOM]} [_{AP} t_i [_{A'} A XP]]]]]_{[NOM]j}]]$$

в. Краткая форма при субъектном подъеме из малой клаузы

Комната оказалась небольшая.

$$[_{TP} DP_{[NOM]i} T_{+FIN} [_{vP} v_{intg} [_{VP} V_{<ОКАЗАТЬСЯ>} [_{PredP} t_i [_{Pred'} Pred [_{adjP} Op_i adj_{[NOM]} [_{AP} t_i [_{A'} A XP]]]]]_{[NOM]j}]]$$

При независимом лицензировании группа полного прилагательного *adjP* в составе *PredP* получает творительный падеж (так называемый творительный предикативный) от функциональной вершины *Pred*, см. [Bailyn 2001; 2012; Pereltsvaig 2007; Madariaga 2007; Matushansky 2008]. Подлежащее *PredP* лицензируется независимо и может получать любой структурный падеж клаузы или соответствовать PRO в нефинитной клаузе. Поскольку лицензирование творительного падежа полной формы обусловлено присутствием вершины *Pred*, единственным требованием для такой конфигурации является возможность

¹³ Предложенный анализ деривации номинатива полной формы может быть распространен на конструкции с другими структурными падежами, доступными для *PredP* — аккузативом и дативом. Если полная форма прилагательного в составе *PredP* получает тот же падеж, что и подлежащее *PredP*, то в конструкциях с объектным подъемом из малых клауз мы ожидаем полной формы прилагательного в аккузативе, а в конструкциях с независимым инфинитивом — в дативе. И действительно, общий с подлежащим *PredP* падеж в подобных конструкциях изредка используется наряду с инструменталисом:

- (i) Там обычно я видел **его пьяного** (/пьяным): он честно ворочался в снегу, он никогда не сдавался и не замерзал, как многие другие пьяндыги [Владимир Маканин. Наше утро].
- (ii) Странное какое-то солнце... Красное, дымное... Да и **какому** (/каким) **ему** быть, ночному солнцу? [НКРЯ].

независимого лицензирования подлежащего. Поэтому полная форма прилагательного в творительном падеже может выступать в качестве сказуемого в клаузах любого типа¹⁴.

(29) Локальное лицензирование adjP

[_{PredP} DP/PRO [_{Pred'} Pred [_{adjP} Op_i adj]_[INSTR] [_{AP} t_i [_{A'} A XP]]]]]

Наконец, последний вопрос в деривации полной формы прилагательного в предикативной позиции, на котором мы остановимся, касается выбора между двумя стратегиями. Можно было бы предположить, что за каждую из них отвечает отдельная вершина Pred: приписывающая творительный предикативный отвечает за независимое лицензирование, а не приписывающая — за совместное. Однако более естественным кажется другое решение: совместная стратегия реализуется при инкорпорации вершины A+adj в вершину Pred. Подобную операцию — подъем вершины предиката (Raising the predicate head) предлагает М. ден Диккен [den Dikken 2006: 113–115]. Мотивацию для подъема вершины предиката ден Диккен усматривает в преодолении фазового ограничения: он считает Pred (в его терминологии — R, Relator) фазовой вершиной, так что условие непроницаемости фазы делает предикат недоступным для внешних зондов. Подъем вершины предиката позволяет обойти это ограничение.

Заметим, что совместная стратегия лицензирования требует означивания падежного признака у предиката PredP, значит, PredP при совместной стратегии является проницаемой для A-зондов. Следовательно, имеются теоретические аргументы в пользу подъема вершины предиката в русском языке. Если инкорпорация A+adj в вершину Pred имеет место, естественно предположить, что она меняет свойства вершины Pred и блокирует приписывание творительного предикативного¹⁵. Таким образом, падеж полной формы прилагательного (именительный или творительный) технически оказывается предопределен наличием или отсутствием подъема вершины предиката: при подъеме реализуется совместная стратегия (падеж подлежащего и предиката совпадает, (30a)), без подъема — независимая (предикат получает творительный предикативный, (30b)).

(30) а. Номинатив полной формы, анализ с подъемом вершины предиката

Комната была небольшая.

[_{TP} DP_[NOM] T_{+FIN} [_{AuxP} Aux_{<быть>} [_{PredP} t_i [_{Pred'} A_j+adj]_k+Pred_[NOM] [_{adjP} Op_i t_k [_{AP} t_i [_{A'} t_j XP]]]]]] [_{NOM}]]

б. Инструменталис полной формы

Комната была небольшой.

[_{TP} DP_[NOM] T_{+FIN} [_{AuxP} Aux_{<быть>} [_{PredP} t_i [_{Pred'} Pred [_{adjP} Op_i A_j+adj]_[INSTR] [_{AP} t_i [_{A'} t_j XP]]]]]]]

6.4. Причастный пассив

Перейдем к обсуждению причастного пассива. Начнем с обобщения о том, что для получения собственно пассивной интерпретации в финитной клаузе необходима форма краткого страдательного причастия.

¹⁴ Исключение составляют, как указано выше, финитные клаузы с нулевой словоформой глагола-связки. По-видимому, это ограничение не может быть объяснено без дополнительных допущений. Так, в [Bailyn 2012] и [Циммерлинг 2018] предполагается, что условием лицензирования творительного предикативного является наличие фонологически выраженной вершины сказуемого.

¹⁵ Допущение об изменении способности вершины к приписыванию падежа вследствие инкорпорации часто используется для моделирования падежного варьирования, в том числе и в формальной русистике. Например, в работе [Brown, Franks 1995] делается предположение, что

Вслед за [Пазельская, Татевосов 2008] мы предполагаем, что форма краткого страдательного причастия — это глагольная проекция уровня $vP/AspP$, включающая пассивный легкий глагол и вложенная под совокупность адъективирующих вершин, как это показано в (31).

(31) Краткое страдательное причастие

[_{PartP/AP} Part/A [_{AspP} Asp [_{vP} v_{PASS} [_{VP} DP/PRO V XP]]]]

Структура в (31) имеет много общего с краткой формой прилагательного, что отражает интуицию Л. Бэбби о категориальном единстве кратких прилагательных, кратких страдательных причастий и глаголов. В составе PartP имеется ровно одна именная группа, нуждающаяся в падежном лицензировании, — это пассивное подлежащее, внутренний аргумент глагольной основы. PartP может выступать компонентом вспомогательного глагола. Наконец, сама форма краткого страдательного причастия по своим морфологическим свойствам идентична краткому прилагательному. Категориальное сходство кратких прилагательных и кратких страдательных причастий обозначается в (31) смешанным ярлыком Part/A.

Как и краткое прилагательное, краткое страдательное причастие нуждается в синтаксическом лицензировании. Лицензорами могут выступать оболочка adjP при образовании полной формы причастия или номинативное подлежащее. Как и в случае с краткими прилагательными, допускается приписывание номинатива выраженному подлежащему PartP (32) или согласование PRO со своим номинативным контролером при субъектном контроле (33).

(32) а. Краткое пассивное причастие в финитной клаузе

Больной был госпитализирован.

[_{TP} DP [_{NOM}]i T_{+FIN} [_{AuxP} Aux_{<БЫТЬ>} [_{PartP/AP} Part/A [_{AspP} Asp [_{vP} v_{PASS} [_{VP} t_i V XP]]]]]]]]

б. Краткая форма в модальной конструкции с инфинитивом

Больной может быть госпитализирован.

[_{TP} DP [_{NOM}]i T_{+FIN} [_{ModP} Mod_{<МОЖЬ>} [_{AuxP} Aux_{<БЫТЬ>} [_{PartP/AP} Part/A [_{AspP} Asp [_{vP} v_{PASS} [_{VP} t_i V XP]]]]]]]]

в. Краткая форма при субъектном подъеме из малой клаузы

Больной оказался госпитализирован.

[_{TP} DP [_{NOM}]i T_{+FIN} [_{vP} v_{INTR} [_{VP} V_{<ОКАЗАТЬСЯ>} [_{PartP/AP} Part/A [_{AspP} Asp [_{vP} v_{PASS} [_{VP} t_i V XP]]]]]]]]

(33) Краткое пассивное причастие при субъектном контроле

Больной хочет быть госпитализирован.

Будучи госпитализирован, больной быстро пошел на поправку.

Больной явился в приемный покой, чтобы быть госпитализирован.

DP [_{NOM}]i ... [_{CP} C [_{TP} PRO [_{NOM}]i T_{-FIN} [_{AuxP} Aux [_{PartP/AP} Part/A [_{AspP} Asp [_{vP} v_{PASS} [_{VP} t_i V XP]]]]]]]]

При вложении PartP под адъективирующую вершину adj происходит предикатная абстракция по переменной, соответствующей пассивному подлежащему. При этом полная форма причастия сохраняет собственно пассивную интерпретацию, что видно по ее использованию в атрибутивной позиции, где она способна присоединять характерные для событийных пассивов зависимые (ср. *напечатанная в «Правде» на прошлой неделе статья; сцена, подаваемая режиссером крупным планом*). Предполагаем, что утрата собственно пассивной интерпретации и превращение полного причастия в результативное происходит при его использовании в PredP в качестве предиката в структуре (34).

приписывающая структурный аккузатив вершина Asp, инкорпорируясь в вершину Neg, меняет свои падежные свойства и выступает источником генитива отрицания.

(34) Полная форма страдательного причастия в составе PredP

[_{PredP} DP/PRO [_{Pred} Pred [_{adjP} Op_i adj_[uCase] [_{PartP/AP} Part/A [_{AspP} Asp [_{vP} v_{PASS} [_{VP} t_i V XP]]]]]]]]]

В финитной клаузе на основе такой структуры событийный пассив не может быть получен (**Статья была напечатанная/напечатанной в газете «Правда»*). Мы предполагаем, что на это есть и семантические, и синтаксические причины. С семантической точки зрения для пассивной интерпретации нужна событийная переменная, которая при вложении причастия в PredP оказывается связана. С синтаксической точки зрения подлежащим пассива должен выступать внутренний аргумент глагола, а в структуре (34) подлежащим оказывается аргумент вершины Pred. Таким образом, неспособность полных форм страдательного причастия выступать сказуемым в причастном пассиве получает структурное объяснение.

Тем не менее, мы обнаружили, что в нефинитных клаузах собственно пассивная интерпретация оказывается доступной и для полной формы страдательного причастия в творительном падеже. Мы полагаем, что использование этой формы мотивировано соображениями «крайнего средства» (last resort), как единственный способ лицензирования причастного пассива в нефинитных клаузах, не допускающих краткой формы (то есть обладающих неноминативным выраженным или нулевым подлежащим). Эта модель по аналогии распространяется на прочие нефинитные клаузы, в которых параллельно доступна и краткая форма страдательного причастия. Использование полной формы страдательного причастия в именительном падеже в нефинитных клаузах не могло бы увеличить число контекстов для причастного пассива, так как дистрибуция номинатива полной формы уже, чем краткой формы. Поэтому полная форма страдательного причастия в именительном падеже не используется в причастных пассивах ни в финитных, ни в нефинитных клаузах.

7. Выводы

В этой статье мы исследовали адъективные сказуемые в финитных и нефинитных клаузах русского языка. Выяснилось, что практически все структурные типы адъективных сказуемых: краткие прилагательные, полные прилагательные в именительном падеже, полные прилагательные в творительном падеже, краткие страдательные причастия, полные страдательные причастия в творительном падеже, — обладают собственными моделями дистрибуции. Совпадение структурных контекстов характерно только для кратких прилагательных и кратких страдательных причастий.

При описании дистрибуции адъективных сказуемых критически важными оказываются противопоставление канонического (номинативного) подлежащего и прочих видов именных групп (как неноминативных подлежащих, так и дополнений). Формальный признак падежа оказывается релевантен не только для выраженных именных групп, но и для пустых категорий — PRO и А-следов. Обнаруживаются свидетельства того, что унификация падежного признака PRO с падежным признаком контролера возможна не только в актантных, но и в сирконстантных нефинитных клаузах, причем наличие выраженного комплементаризера (*чтобы*) не создает препятствия для этого процесса.

Наибольший интерес представляет сравнение дистрибуции кратких прилагательных и полных прилагательных в именительном падеже. Исследование нефинитных конфигураций показывает, что эти типы адъективных сказуемых накладывают разные ограничения на подлежащее своей клаузы и позволяют противопоставить номинативный А-след (след от передвижения именной группы, получающей номинатив от финитной предикативной вершины) и номинативное PRO (нулевое подлежащее нефинитной клаузы, которое может копировать номинатив своего контролера при субъектном контроле). Тот факт, что для описания дистрибуции адъективных сказуемых необходимо различать А-след

и PRO, подтверждает теоретическую гипотезу о принципиальном структурном отличии конфигураций подъема и контроля. Этот вывод особенно значим ввиду всё возрастающей популярности теории обязательного контроля как А-передвижения именной группы в тематическую позицию (Movement theory of control [Hornstein 1999; 2001; 2003; Boeckx, Hornstein 2003; 2004; Hornstein, Polinsky 2010]). Очевидно, что даже если дистрибуция А-следов и PRO совпадает, их свойства оказываются разными.

Наконец, специфические ограничения на формы причастного пассива, доступные в финитных и нефинитных клаузах, свидетельствуют о нетривиальных синтаксических и семантических изменениях, происходящих при переходе от краткой формы к полной. Различия между ними, не всегда явные в случае прилагательных, наиболее ярко проявляются именно в причастиях, позволяя увидеть структурные и семантические корреляты «адъективности» и «атрибутивности» [Гращенков 2018].

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Байков 2020 — Байков Ф. В. Косвенный контроль в русском языке. *Rhema. Рема*, 2020, 1: 106–125. [Baykov F. V. Oblique control in Russian. *Rhema*, 2020, 1: 106–125.]
- Всеволодова 2013 — Всеволодова М. В. О грамматике полных и кратких форм прилагательных и причастий в русском языке. *Вопросы языкознания*, 2013, 6: 3–32. [Vsevolodova M. V. On the grammar of long and short adjectives and participles in Russian. *Voprosy Jazykoznanija*, 2013, 6: 3–32.]
- Гловинская 1996 — Гловинская М. Я. Активные процессы в грамматике. *Русский язык конца XX столетия (1985–1995)*. Земская Е. А. (ред.). М.: Языки русской культуры, 1996, 237–304. [Glovinskaya M. Ya. Active processes in grammar. *Russkii yazyk kontsa XX stoletiya (1985–1995)*. Zemskaya E. A. (ed.). Moscow: Yazyki russkoi kul'tury, 1996, 237–304.]
- ГР 1952 — Виноградов В. В., Истрина Е. С., Бархударов С. Г. (ред.). *Грамматика русского языка*. В 2 т. М.: Изд-во АН СССР, 1952. [Vinogradov V. V., Istrina E. S., Barkhudarov S. G. (eds.). *Grammatika russkogo yazyka* [Russian grammar]. In 2 vols. Moscow: Press of the Academy of Sciences of the USSR, 1952.]
- Гращенков 2018 — Гращенков П. В. *Грамматика прилагательного. Типология атрибутивности и адъективности*. М.: Языки славянской культуры, 2018. [Grashchenkov P. V. *Grammatika prilagatel'nogo. Tipologiya atributivnosti i ad'ektivnosti* [The grammar of the adjective. The typology of attributivity and adjectivity]. Moscow: Yazyki slavyanskoi kul'tury, 2018.]
- Гращенков 2019 — Гращенков П. В. Прилагательные в русской именной предикации. *Труды Института русского языка им. В. В. Виноградова*, 2019, 4(22): 51–66. [Grashchenkov P. V. Adjectives in Russian nominal predications. *Proc. of the V. V. Vinogradov Russian Language Institute*, 2019, 4(22): 51–66.]
- Гращенков 2022 — Гращенков П. В. О синтаксической селекции (группы) прилагательного. [Grashchenkov P. V. On syntactic selection of adjective (phrase).] *Acta Linguistica Petropolitana*, 2022, 18(3): 72–103.
- Ицкович 1982 — Ицкович В. А. *Очерки синтаксической нормы*. М.: Наука, 1982. [Itskovich V. A. *Ocherki sintaksicheskoi normy* [Essays on the syntactic norm]. Moscow: Nauka, 1982.]
- Йокояма 1983 — Йокояма О. В защиту запретных деепричастий. [Yokoyama O. In defence of prohibited gerunds.] *American Contributions to the Ninth International Congress of Slavists (Kiev, September 1983)*. Vol. 1: *Linguistics*. Flier M. (ed.). Columbus: Slavica, 1983, 373–381.
- Козинский 1983 — Козинский И. Ш. *О категории «подлежащее» в русском языке*. М.: ИРЯ АН СССР, 1983. [Kozinskii I. Sh. *O kategorii «podlezhashchee» v russkom yazyke* [On the category of Subject in Russian]. Moscow: Russian Language Institute of the Academy of Sciences of the USSR, 1983.]
- Козинский 1985 — Козинский И. Ш. Кореферентные связи инфинитивных оборотов в русском языке. *Типология конструкций с предикатными актантами*. Храковский В. С. (отв. ред.). Л.: Наука, 1985, 112–116. [Kozinskii I. Sh. Coreferential relations in Russian infinitival phrases. *Tipologiya konstruktii s predikatnymi aktantami*. Hrakovskij V. S. (ed.). Leningrad: Nauka, 1985, 112–116.]
- Кустова 2018 — Кустова Г. И. Прилагательное. *Материалы для проекта корпусного описания русской грамматики*. М., 2018. Рук. [Kustova G. I. Adjective. *Materialy dlya proekta korpusnogo opisaniya russkoi grammatiki*. Moscow, 2018. Ms.] <http://rusgram.ru>.

- Лютикова 2010 — Лютикова Е. А. К вопросу о категориальном статусе именных групп в русском языке. *Вестник Московского университета. Сер. 9: Филология*, 2010, 6: 36–76. [Lyutikova E. A. On the categorial status of noun phrases in Russian. *Lomonosov Philology Journal*, 2010, 6: 36–76.]
- Лютикова 2022а — Лютикова Е. А. Есть ли синтаксический подъем в русском языке? Часть 1: Инфинитивные клаузы. *Вестник Московского университета. Сер. 9: Филология*, 2022, 5: 27–45. [Lyutikova E. A. Does Russian attest syntactic raising? Part 1: Infinitival clauses. *Lomonosov Philology Journal*, 2022, 5: 27–45.]
- Лютикова 2022б — Лютикова Е. А. Есть ли синтаксический подъем в русском языке? Часть 2: Малые клаузы. *Вестник Московского университета. Сер. 9: Филология*, 2022, 6: 58–74. [Lyutikova E. A. Does Russian attest syntactic raising? Part 2: Small clauses. *Lomonosov Philology Journal*, 2022, 6: 58–74.]
- Лютикова, Сидельцев 2020 — Лютикова Е. А., Сидельцев А. В. Залоговые альтернации нефинитных форм глагола в хеттском языке. *Индоевропейское языкознание и классическая филология*, 2020, 24(1): 234–263. [Lyutikova E. A., Sideltsev A. V. Voice alternations in Hittite non-finite verbal forms. *Indo-European Linguistics and Classical Philology*, 2020, 24(1): 234–263.]
- Моргунова 2021 — Моргунова Е. В. *Синтаксис и семантика дативно-инфинитивных структур в русском языке*. Маг. дисс. М.: МГУ имени М. В. Ломоносова, 2021. [Morgunova E. V. *Sintaksis i semantika dativno-infinitivnykh struktur v russkom yazyke* [Syntax and semantics of dative-infinitive structures in Russian]. Master thesis. Moscow: Lomonosov Moscow State Univ., 2021.]. https://morgunova-katya.github.io/files/morgunova_MA_thesis.pdf.
- Никольс 1985 — Никольс Дж. Падежные варианты предикативных имен и их отражение в русской грамматике. *Новое в зарубежной лингвистике*. Вып. XV. *Современная зарубежная русистика*. Булыгина Т. В., Кибрик А. Е. (ред.). М.: Прогресс, 1985, 342–387. [Nichols J. Case forms of predicate nominals and their reflection in Russian grammar. *Novoe v zarubezhnoi lingvistike*. No. 15. *Sovremennaya zarubezhnaya rusistika*. Bulygina T. V., Kibrik A. E. (eds.). Moscow: Progress, 1985, 342–387.]
- Никунласси 1993 — Никунласси А. Именительный или творительный?: Синтаксические прилагательные при полнознаменательных глаголах в русском языке: Проблемы выбора падежа. Helsinki: Univ. of Helsinki, 1993. [Nikunlassi A. *Imenitel'nyi ili tvoritel'nyi?: Sintaksicheskie prilagatel'nye pri polnoznamenatel'nykh glagolakh v russkom yazyke: Problemy vybora padezha* [Nominative or Instrumental?: Syntactic adjectives construed with lexical verbs in Russian: Problems of case choice]. Helsinki: Univ. of Helsinki, 1993.]
- НКРЯ — Национальный корпус русского языка [Russian National Corpus]. <http://www.ruscorpora.ru>.
- Онипенко, Биккулова 2016 — Онипенко Н. К., Биккулова О. С. Проблема деепричастной нормы и категория субъекта. *Труды Института русского языка им. В. В. Виноградова*, 2016, 4(10): 220–231. [Onipenko N. K., Bikkulova O. S. The problem of norm in gerunds and the category of subject. *Proc. of the V. V. Vinogradov Russian Language Institute*, 2016, 4(10): 220–231.]
- Падучева 2017 — Падучева Е. В. Конструкция с независимым инфинитивом. *Материалы для проекта корпусного описания русской грамматики*. М., 2017. Рук. [Paducheva E. V. Independent infinitival construction. *Materialy dlya projekta korpusnogo opisaniya russkoi grammatiki*. Moscow, 2017. Ms.] <http://rusgram.ru>.
- Пазельская 2009 — Пазельская А. Г. Модели деривации отглагольных существительных: взгляд из корпуса. *Корпусные исследования по русской грамматике*. Киселева К. Л., Плунгян В. А., Рахилина Е. В., Татевосов С. Г. (ред.). М.: Пробел—2000, 2009, 65–91. [Pazel'skaya A. G. Derivation patterns of deverbal nouns: The view from the corpus. *Korpusnye issledovaniya po russkoi grammatike*. Kiseleva K. L., Plungian V. A., Rakhilina E. V., Tatevosov S. G. (eds.). M.: Probel—2000, 2009, 65–91.]
- Пазельская, Татевосов 2008 — Пазельская А. Г., Татевосов С. Г. Отглагольное имя и структура русского глагола. *Исследования по глагольной деривации*. Плунгян В. А., Татевосов С. Г. (отв. ред.). М.: Языки славянских культур, 2008, 348–379. [Pazel'skaya A. G., Tatevosov S. G. Deverbal noun and structure of Russian verb]. *Issledovaniya po glagol'noi derivatsii*. Plungian V. A., Tatevosov S. G. (eds.). Moscow: Yazyki slavyanskikh kul'tur, 2008, 348–379.]
- РГ 1980 — Шведова Н. Ю. (гл. ред.). *Русская грамматика*. В 2 т. Т. 2: *Синтаксис*. М.: Наука, 1980. [Shvedova N. Yu. (ed.). *Russkaya grammatika* [Russian grammar]. In 2 vols. Vol. 2: *Sintaksis* [Syntax]. Moscow: Nauka, 1980.]
- Тестелец 2001 — Тестелец Я. Г. *Введение в общий синтаксис*. М.: РГГУ, 2001. [Testeleys Ya. G. *Vvedenie v obshchii sintaksis* [Introduction to general syntax]. Moscow: Russian State Univ. for the Humanities, 2001.]
- Циммерлинг 2018 — Циммерлинг А. В. Два диалекта русской грамматики: корпусные данные и модель. *Компьютерная лингвистика и интеллектуальные технологии: по материалам*

- международной конференции «Диалог», 2018, 17(24): 818–830. [Zimmerling A. V. Two dialects of Russian grammar: Corpus data and the model. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue"*, 2018, 17(24): 818–830.]
- Шведова 1952 — Шведова Н. Ю. Полные и краткие формы имен прилагательных в составе сказуемого в современном русском языке. М.: Изд-во Моск. ун-та, 1952. [Shvedova N. Yu. *Polnye i kratkie formy imen prilagatel'nykh v sostave skazuemogo v sovremennom russkom yazyke* [Long and short forms of adjectives within the predicate in the modern Russian language]. Moscow: Moscow Univ. Press, 1952.]
- Alexiadou 2001 — Alexiadou A. *Functional structure in nominals: Nominalization and ergativity*. Amsterdam; Philadelphia: John Benjamins, 2001.
- Babby 1973 — Babby L. The deep structure of adjectives and participles in Russian. *Language*, 1973, 49(2): 349–360.
- Babby 1979 — Babby L. The syntax of gerunds in Russian. *Contributions to grammatical studies: Semantics and syntax*. van Coetsem F., Waugh L. (eds.). Leiden: Brill, 1979, 1–41.
- Babby 1987 — Babby L. Case, prequantifiers, and discontinuous agreement in Russian. *Natural Language and Linguistic Theory*, 1987, 5(1): 91–138.
- Babby 1998 — Babby L. Subject control as direct predication: Evidence from Russian. *Formal Approaches to Slavic Linguistics (FASL) 6: The Connecticut Meeting*. Bošković Ž., Franks S., Snyder W. (eds.). Ann Arbor: Michigan Slavic Publ., 1998, 17–37.
- Babby, Franks 1998 — Babby L., Franks S. The syntax of adverbial participles in Russian revisited. *The Slavic and East European Journal*, 1998, 42(3): 483–515.
- Bailyn 2001 — Bailyn J. The syntax of Slavic predicate case. *ZAS Papers in Linguistics* 22. Jäger G., Strigin A., Wilder Ch., Zhang N. (eds.). Berlin: ZAS, 2001, 1–23.
- Bailyn 2012 — Bailyn J. *The syntax of Russian*. Cambridge: Cambridge Univ. Press, 2012.
- Boeckx, Hornstein 2003 — Boeckx C., Hornstein N. Reply to “Control is not movement”. *Linguistic Inquiry*, 2003, 34: 269–280.
- Boeckx, Hornstein 2004 — Boeckx C., Hornstein N. Movement under control. *Linguistic Inquiry*, 2004, 35: 431–452.
- Bowers 1993 — Bowers J. The syntax of predication. *Linguistic Inquiry*, 1993, 24: 591–656.
- Brown, Franks 1995 — Brown S., Franks S. Asymmetries in the scope of Russian negation. *Journal of Slavic Linguistics*, 1995, 3(2): 239–287.
- Chomsky 1981 — Chomsky N. *Lectures on government and binding*. Dordrecht: Foris, 1981.
- Chomsky, Lasnik 1993 — Chomsky N., Lasnik H. The theory of principles and parameters. *An international handbook of contemporary research*. Vol. 1. Jacobs J., von Stechow A., Sternefeld W., Vennemann T. (eds.). Berlin; New York: De Gruyter Mouton, 1993, 506–569.
- Comrie 1974 — Comrie B. The second dative: A transformational approach. *Slavic transformational syntax*. Brecht R., Chvany C. (eds.). Ann Arbor: Univ. of Michigan, 1974, 123–150.
- den Dikken 2006 — den Dikken M. *Relators and linkers: The syntax of prediction, predicate inversion, and copulas*. Cambridge (MA); London: MIT Press, 2006.
- Fleisher 2006 — Fleisher N. *Russian dative subjects, case and control*. Berkeley: Univ. of California, 2006. Ms.
- Geist 2010 — Geist L. The argument structure of predicate adjectives in Russian. *Russian Linguistics*, 2010, 34(3): 239–260.
- Greenberg, Franks 1991 — Greenberg G., Franks S. A parametric approach to dative subjects and the second dative in Slavic. *The Slavic and East European Journal*, 1991, 35(1): 71–97.
- Harley 2020 — Harley H. Relative nominals and event nominals in Hiaki. *Nominalization: 50 years on from Chomsky's Remarks*. Alexiadou A., Borer H. (eds.). Oxford: Oxford Univ. Press, 2020, 203–230.
- Hornstein 1999 — Hornstein N. Movement and control. *Linguistic Inquiry*, 1999, 30: 69–96.
- Hornstein 2001 — Hornstein N. *Move! A minimalist theory of construal*. Oxford: Blackwell, 2001.
- Hornstein 2003 — Hornstein N. On control. *Minimalist syntax*. Hendrick R. (ed.). Oxford: Blackwell, 2003, 6–81.
- Hornstein, Polinsky 2010 — Hornstein N., Polinsky M. Control as movement. *Movement theory of control*. Hornstein N., Polinsky M. (eds.). Amsterdam; Philadelphia: John Benjamins, 2010, 1–42.
- Landau 2008 — Landau I. Two routes of control: Evidence from case transmission in Russian. *Natural Language and Linguistic Theory*, 2008, 26(4): 877–924.
- Landau 2013 — Landau I. *Control in generative grammar: A research companion*. Cambridge: Cambridge Univ. Press, 2013.

- Lasnik, Saito 1991 — Lasnik H., Saito M. On the subject of infinitives. *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*. Dobrin L., Nichols L., Rodriguez R. (eds.). Chicago: Chicago Linguistic Society, 1991, 324–343.
- Lyutikova, Gerasimova 2023 — Lyutikova E., Gerasimova A. Negative Concord and locality in Russian. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 2023, 68(1): 31–73.
- Lyutikova, Sideltsev 2021 — Lyutikova E., Sideltsev A. Voice neutrality in Hittite infinitives: A restructuring analysis. *Journal of Historical Syntax*, 2021, 5: Paper 15.
- Madariaga 2007 — Madariaga N. An economy approach to the triggering of the Russian instrumental predicative case. *Historical Linguistics 2005*. Salmons J., Dubenion-Smith S. (eds.). Amsterdam; Philadelphia: John Benjamins, 2007, 103–117.
- Matushansky 2008 — Matushansky O. A case study of predication. *Studies in formal Slavic linguistics. Contributions from Formal Description of Slavic Languages 6.5*. Marušić F., Žaucer R. (eds.). Frankfurt am Main: Peter Lang, 2008, 213–239.
- Nichols 1981 — Nichols J. *Predicate nominals: A partial surface syntax of Russian*. Berkeley; Los Angeles: Univ. of California Press, 1981.
- Pereltsvaig 2007 — Pereltsvaig A. *Copular sentences in Russian: A theory of intra-clausal relations*. Dordrecht: Springer, 2007.
- Pesetsky, Torrego 2007 — Pesetsky D., Torrego E. The syntax of valuation and the interpretability of features. *Phrasal and clausal Architecture: Syntactic derivation and interpretation*. Karimi S., Samiian V., Wilkins W. (eds.). Amsterdam; Philadelphia: John Benjamins, 2007, 262–294.
- Postal 1974 — Postal P. *On raising*. Cambridge (MA): MIT Press, 1974.
- Rappaport 2001 — Rappaport G. The geometry of the Polish nominal phrase: Problems, progress, and prospects. *Generative linguistics in Poland: Syntax and morphosyntax*. Bański P., Przepiórkowski A. (eds.). Warsaw: Polish Academy of Sciences, 2001, 173–189.
- Rosenbaum 1965 — Rosenbaum P. *The grammar of English predicate complement constructions*. Doctoral diss. Cambridge (MA): MIT, 1965.
- Schoorlemmer 1998 — Schoorlemmer M. Complex event nominals in Russian: Properties and readings. *Journal of Slavic Linguistics*, 1998, 6(2): 205–254.
- Sigurðsson 2002 — Sigurðsson H. To be an oblique subject: Russian vs. Icelandic. *Natural Language and Linguistic Theory*, 2002, 20(4): 691–724.
- Sigurðsson 2008 — Sigurðsson H. The case of PRO. *Natural Language and Linguistic Theory*, 2008, 26(2): 403–450.
- Tsedryk 2017 — Tsedryk E. Dative-infinitive constructions in Russian: Are they really biclausal? *Formal Approaches to Slavic Linguistics (FASL) 25: The Third Cornell Meeting*. Browne W., Despić M., Enzina N., Harmath-de Lemos S., Karlin R., Zec D. (eds.). Ann Arbor: Michigan Slavic Publ., 2017, 298–317.
- Wurmbrand 2001 — Wurmbrand S. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter, 2001.
- Zeijlstra 2012 — Zeijlstra H. There is only one way to agree. *The Linguistic Review*, 2012, 29(3): 491–539.
- Zimmerling 2024 — Zimmerling A. Microsyntax meets macrosyntax: Russian neg-words revisited. *Russian Linguistics*, 2024, 48(1): Paper 6.

О некоторых типологических характеристиках андийских глагольных систем

© 2024

Константин Вадимович Филатов

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; Институт лингвистических исследований РАН, Санкт-Петербург, Россия;
triolo@mail.ru

Аннотация: Несмотря на то, что андийские языки (< нахско-дагестанские) традиционно считаются агглютинативными, их глагольные системы практически не демонстрируют строго агглютинативных свойств. Наиболее важные характеристики этих систем — малая ортогональность грамматических категорий и связанные с ней свойства семантической неаддитивности или морфологической идиоматичности. В отношении морфотактической структуры андийские системы могут быть как формально-аддитивными, так и формально-неаддитивными, что связано с интенсивностью фузионных процессов в системе. Взаимосвязь структур семантических и формальных противопоставлений в таких системах является непрозрачной. Многие формативы в формально-аддитивных системах демонстрируют морфомные свойства, а их взаимная селективность проявляется в иерархической структуре парадигмы. Типологически такие особенности систем, выражающих преимущественно ТАМ-значения, не являются неожиданными. В статье предлагается ряд диахронических соображений, позволяющих объяснить наблюдаемые особенности систем. Неортогональность и семантическая неаддитивность может быть объяснена асинхронной грамматикализацией, которая предотвращает формирование ортогональных категорий с аддитивными средствами выражения и связана с разной стабильностью значений в системе и разной скоростью их обновления. Морфомные и иерархические структуры в парадигме возникают из-за сохранения в парадигме разных хронологических слоев морфологизованного материала и действия грамматического дрейфа, ведущего к образованию полисемичных и гетеросемичных структур значений показателей.

Ключевые слова: агглютинация, андийские языки, глагольные системы, грамматические категории, морфология, нахско-дагестанские языки, фузия, членимость

Благодарности: Мы благодарим анонимных рецензентов за ценные комментарии, которые позволили существенно углубить наш анализ. Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ.

Для цитирования: Филатов К. В. О некоторых типологических характеристиках андийских глагольных систем. *Вопросы языкознания*, 2024, 6: 32–57.

DOI: 10.31857/0373-658X.2024.6.32-57

On typologically relevant properties of Andic verbal grammatical systems

Konstantin V. Filatov

HSE University, Moscow, Russia; Institute for Linguistic Studies, Russian Academy of Sciences,
St. Petersburg, Russia; triolo@mail.ru

Abstract: Despite the fact that Andic languages (< East Caucasian) are widely recognised as agglutinating, their verbal systems barely demonstrate any agglutinating properties. The most important characteristics of these systems are the incomplete orthogonality of grammatical categories and the related

properties of semantic non-additivity or morphological idiomacity. In terms of morphotactic structure, different Andic systems can be either morphotactically complex or simplex. The difference is related to the intensity of fusional processes in the system. The relationship between the structures of semantic and formal oppositions in such systems is opaque. Many formatives in formally additive systems exhibit morphomic properties, and their decreased combinatorial potential is manifested in the hierarchical structure of the paradigm. From a typological perspective, such peculiarities are not unexpected for systems expressing predominantly TAM meanings. The paper proposes a number of diachronic considerations to explain the observed characteristics of the systems. Non-orthogonality and semantic non-additivity can be explained by asynchronous grammaticalization, which prevents the formation of orthogonal categories with additive means of expression and is related to different stability of meanings in the system and different rates of their diachronic renovation. Morphomic and hierarchical structures in the paradigm arise due to layering and grammatical drift leading to polysemic and heterosemic meaning structures of markers.

Keywords: agglutination, Andic, fusion, grammatical categories, morphology, Nakh-Daghestanian, segmentability, verbal systems

Acknowledgements: We are grateful to the anonymous reviewers for their valuable comments, which have enabled us to significantly enhance the depth of our analysis. Support from the Basic Research Program of the National Research University Higher School of Economics is gratefully acknowledged.

For citation: Filatov K. V. On typologically relevant properties of Andic verbal grammatical systems. *Voprosy Jazykoznanija*, 2024, 6: 32–57.

DOI: 10.31857/0373-658X.2024.6.32-57

1. Постановка проблемы

Проблема, которой посвящена эта статья, кратко сформулирована М. Баэрманом в работе [Baerman 2019: 124]: «Inflectional paradigms come in two possible flavours: either the forms line up with the featural system whose values they express, or they do not». Мы будем называть эту проблему **проблемой соотносённости категориальной и морфологической структур**. Более раннее обращение к ней можно найти в статье [Плунгян, Урманчиева 2004].

В. А. Плунгян и А. Ю. Урманчиева, развивая идею из работы [Welmers 1973: 343] о том, что глагольные системы языков нигер-конго лучше всего описывать в виде списка глагольных конструкций, а не сетки с пересечением категорий вида, времени и модальности, вводят противопоставление **линейных** и **многомерных** грамматических систем. Центральное понятие этой статьи, позволяющее определить два этих типа систем, — комбинаторный потенциал значений грамматических категорий. В многомерных системах грамматические формы способны комбинировать грамматические значения, в отличие от линейных — это системы с жестким, некомбинаторно заданным списком грамматических значений. Метафору, стоящую за этими терминами, можно интерпретировать, если обратиться к понятию **ортогональности** грамматических категорий, определенному в [Corbett 2012: 34]: ортогональными категориями называются такие, значения которых могут быть выражены в словоформе независимо (ср. также более ранние формулировки схожего принципа «невозможности двух однородных элементов в одной словоформе» в [Зализняк 1967/2002: 27], а также [Володин 1976: 85]). Таким образом, многомерными, по [Плунгян, Урманчиева 2004], являются системы, в которых можно выделить несколько ортогональных категорий.

Понятие комбинаторного потенциала ортогональных категорий позволяет переформулировать в терминах категориальной структуры понятия кумуляции и агглютинации: они возможны только в многомерных системах, но свой комбинаторный потенциал реализуют по-разному.

В **многомерно-кумулятивных** системах, таких как императив в латинском языке (табл. 1), показатели кумулятивно выражают комбинации из двух значений двух разных

категорий — числа и залога. Ортогональность категорий числа и залога обеспечивается в этом случае тем, что каждая морфема охарактеризована ровно двумя значениями. Поскольку все четыре комбинации значений выражаются различными морфемами (ACT + SG = -Ø / -e, ACT + PL = -(i)te, PASS + SG = -re, PASS + PL = -mini), можно считать, что категории числа и залога заданы независимо.

Таблица 1

Многомерно-кумулятивная система (императив в латинском языке, по [Harkness 1898: 98–99])

Залог	Число	
	SG	PL
ACT	-Ø / -e	-(i)te
PASS	-re	-mini

В узбекской глагольной системе (табл. 2) — **многомерно-агглютинативной** — каждая морфема выражает ровно одно значение категории, а комбинации значений выражаются при помощи комбинаций показателей, например, *кий-ин-тир-ма-яп-ман* [одеваться-REFL-CAUS-NEG-PROG-1SG] ‘я не одеваю в данный момент’ или *ур-иш-ма* [драться-RECP-NEG] ‘не дерись!’ [Ревзин, Юлдашева 1969: 47]. Как видно из приведенных примеров, в рамках каждого позиционного класса независимо выражаются значения ровно одной категории.

Таблица 2

Многомерно-агглютинативная система (узбекский язык, по [Ревзин, Юлдашева 1969])

Возвратность	Залог	Полярность	ТАМ	Лицо
-ин	-тир	-ма	-яп	-ман
-Ø	-иш	-Ø	-а / -й	-сан
	-л		-ди	-ти
			-ган	...
			...	

В отличие от многомерных систем, грамматические формы линейных систем не могут быть представлены при помощи композиции значений ортогональных категорий. В глагольной системе языка кпелле (< манде) (табл. 3, с. 35) каждая словоформа парадигмы представляет собой уникальный набор грамматических значений и морфологических средств их выражения.

При этом затруднительно сформулировать обобщения, при помощи которых дистрибуция грамматических показателей может быть соотнесена с однородными (в терминах [Зализняк 1967/2002: 23]) классами грамматических значений. Например, показатель *e* представлен в формах гортатива и прошедшего времени, а показатель *a* — в формах прогрессива, хабитуалиса, кондиционалиса и экспериенциального перфекта. В таких системах невозможно сгруппировать морфологические средства по семантическому признаку (аспекта, времени, модальности) и выделить на этом основании ортогональные категории. Можно лишь выделить одну «макрокатегорию» (называемую иногда ТАМ-категорией, см., например, [Виноградов 2013а]), в которую входят значения из разных семантических зон, таких как модальность, время, аспект.

Вместе с тем такая типология достаточно радикальна: во-первых, системы этих типов в чистом виде встречаются редко, а во-вторых, слишком ориентированы на элементарно-комбинаторные (англ. *Item-and-Arrangement* [Hockett 1954]) морфологические модели.

Таблица 3

Линейная глагольная система
(либерийский кпелле, глагол *kula* ‘выходить’, данные [Welmers 1973: 395]
цит. по [Плунгян, Урманчиева 2004])¹

Форма	Грамматический ярлык	Перевод
<i>kula</i>	Императив	‘выходи!’
<i>é kùla</i>	Гортатив	‘пусть выходит!’
<i>è kùla</i>	Прошедшее время	‘вышел’
<i>a kulâi</i>	Прогрессив	‘(в данный момент) выходит’
<i>a kùla</i>	Хабитуалис	‘(обычно) выходит’
<i>à kùla</i>	Кондиционалис	‘когда / если выйдет’
<i>à kula</i>	Экспериенциальный перфект	‘уже (не раз) выходил’
<i>aâ kula</i>	Результативный перфект	‘вышел (и находится снаружи)’
<i>èi kùla</i>	Дезидератив	‘хорошо бы вышел!’

Хотя латинская система часто используется в качестве хрестоматийного примера многомерно-кумулятивной системы, П. Х. Мэттьюз в монографии [Matthews 1972] показал, что она далеко не гомогенна в этом отношении: разные ее подпарадигмы демонстрируют разнообразные отклонения от кумулятивного принципа организации (см. [Ibid.: 56–96]) и во многом свидетельствуют против моделей словоизменения, основанных на представлении о морфемах как двусторонних единицах. Поэтому в современной морфологической теории скорее ставится вопрос о взаимной соотнесенности между структурой семантических противопоставлений и структурой формальных противопоставлений. Ср., например, формулировку М. Мэйдена [Maiden 2018: 34] (со ссылкой на [Matthews 1972: 67–77]) о латинской системе: «[T]he link between lexical and grammatical meanings and the internal structure of the word was frequently opaque, both paradigmatically and syntagmatically», — или следующее наблюдение Дж. Блевинса: «The structure of a language is defined by an association between a system of contrasts at the level of grammatical meaning and a system of contrasts at the level of form, but there need not be any correspondence between individual form contrasts and meaning contrasts» [Blevins 2016: 15].

В данной статье мы исследуем связь этих двух структур в глагольных системах некоторых андийских языков (< аваро-андо-цезские < нахско-дагестанские). Кавказоведческая традиция практически всегда характеризует эти языки как агглютинативные, рассматривая их самостоятельно либо в составе нахско-дагестанских языков (ср. [Услар 1889: 2; Бокарев, Климов 1967: 10; Назаров 1974: 19; Kibrik et al. (eds.) 1996: xiii; Алексеев, Атаев 1997: 39; van den Berg 2005: 162; Forker 2018: 3; 2020: 248; Creissels 2018: 167; Ферхеес 2019: 49; Ganenkov, Maisak 2020: 112] и др.). При этом нужно учитывать, что «агглютинация» — термин исторически многозначный (см. об этом, например, [Plungian 2001] или [Haspelmath 2009]). Не всегда ясно, имеется ли в виду агглютинация как разновидность изоморфизма между формальной и семантической организацией словоформы или же речь идет об агглютинации как отсутствии фузии.

Традиция описания, заложенная А. Е. Кибриком, также, в сущности, уходит от ответа на вопрос о характере связи между структурой словоформы и грамматическими значениями. Во введении к сборнику очерков глагольной морфологии [Кибрик, Кодзасов 1988: 9] ставится задача выявить состав грамматических категорий глагола в дагестанских языках, хотя в тексте самих очерков описываются лишь модели образования индивидуальных

¹ Ср. также более новое описание глагольной системы гвинейского кпелле [Коношенко 2017], не отличающееся в существенных для нас деталях от описания В. Велмерса.

глагольных форм, а парадигматические отношения между этими формами не обсуждаются. В схожем ключе выполнены, например, и более поздние грамматические описания [Kibrik et al. (eds.) 1996] и [Кибрик и др. 2001]. Показательна в этом отношении грамматика кванадинского диалекта багвалинского языка [Кибрик и др. 2001], где анализу морфологических структур глагольных словоформ [Там же: 67–105] и семантики глагольных грамматических значений [Там же: 232–366] принципиально посвящены разные разделы. Одна из немногих работ, прямо обсуждающих отношения между формальной структурой словоформы и структурой семантических противопоставлений («non-trivial form-meaning relations») в языках Кавказа, — статья [Arkadiev 2022: 19–26], однако материал нахско-дагестанских языков там представлен только образцами именной морфологии.

Характеристика нахско-дагестанских языков как агглютинативных встречается и в работах, специально посвященных типологии категориальных структур, что подразумевает понимание агглютинации как разновидности категориальной структуры. Так, в [Плунгян 2003; 2011; Плунгян, Урманчиева 2004] нахско-дагестанские глагольные системы описываются либо как многомерные без дальнейших уточнений, либо же как ранговые, то есть многомерно-агглютинативные. Стоит отметить, что эти работы не привлекали к детальному рассмотрению андийский материал, апеллируя скорее к системам типа лезгинских или даргинских.

Тем самым **цель** этой статьи — установить категориальные характеристики глагольных систем одной из групп нахско-дагестанских языков — андийских языков, вписать их в типологию грамматических систем и дать возможные диахронические объяснения архитектуры этих систем².

В разделе 2 будут кратко даны теоретические предпосылки. В разделе 3 мы проанализируем ряд глагольных систем андийских языков и покажем, что эти системы характеризуются структурой слабо ортогональных категорий и семантическая неаддитивность — их существенное свойство. Кроме того, мы обсудим и другие свойства этих систем: их различия в формальной аддитивности, иерархическую структуру парадигмы и морфомный характер некоторых формативов в ней. В разделе 4 будут приведены возможные диахронические объяснения наблюдаемых явлений.

2. Теоретические предпосылки

Прежде чем переходить к анализу андийских систем, определим круг теоретических предпосылок и понятий, из которых мы будем исходить. Наше изложение будет опираться на инференционно-реализационные морфологические модели [Stump 2001: 1]. В таких моделях план выражения морфологически неэлементарных словоформ состоит из **основы** и **экспонентов**. Основа, согласно [Ibid.: 33], представляет собой морфологическое выражение, к которому могут быть добавлены словоизменительные экспоненты. Экспонентами называются такие повторяющиеся в парадигме элементы словоформ, которые участвуют в реализации каких-либо словоизменительных признаков плана содержания. Частным случаем экспонентов являются **формативы**, то есть сегментные экспоненты (цепочки фонем). Формативы могут непосредственно выражать тот или иной словоизменительный признак, а могут быть **морфомными**, то есть такими, парадигматическое распределение

² Отметим, что глагольные системы даргинских и лезгинских языков отчасти рассматривались в работах [Сумбатова 2011] и [Майсак 2014]. Сравнительный анализ категориальных структур этих систем с андийскими выходит за рамки данного исследования, но предварительно можно заключить, что по крайней мере даргинские системы могут описываться как многомерные агглютинативные, ср. «[п]равила построения цепочек показателей могут быть построены в терминах грамматики порядков» [Сумбатова 2011: 134].

которых не связано ни с каким систематическим различием в значении или с выражением словоизменительного признака [Ibid.: 169]. В указанном выше смысле морфомным можно считать, например, распределение лексики по словоизменительным классам, см. детальную классификацию морфомных явлений в [Hesse 2023]. **Показателем** мы будем называть формально выделяемый сегмент независимо от того, морфемный он или морфомный.

Между экспонентами и словоизменительными признаками плана содержания устанавливается отношение выражения (англ. *expression*), которое часто формализуется в виде реализационных правил.

В [Blevins 2016: 267] выделяются морфосинтаксическое и морфотактическое «измерения» словоформы: **морфотактическое** относится к формальной структуре неэлементарных словоформ, а морфосинтаксическое, которое мы будем называть просто **грамматическим**, — к набору грамматических значений или словоизменительных признаков, которые соотнесены со словоформой.

Для ответа на вопрос о соотнесенности категориальной и морфотактической структур нам понадобятся три понятия: ортогональность категорий, формальная аддитивность и семантическая аддитивность. Под ортогональностью мы понимаем, как уже было указано выше, комбинаторный потенциал грамматических значений. При этом **ортогональность** может быть неполной — в случае, если не все комбинации двух постулируемых категорий возможны или не все из них имеют уникальное выражение (ср. понятие степени ортогональности (англ. *degree of orthogonality*), введенное в [Fedden, Corbett 2017]).

Под **формальной аддитивностью** мы будем иметь в виду возможность представить последовательность сегментов в виде нескольких формативов, противопоставляющих словоформы, при этом граница между этими формативами должна быть видна в поверхностной реализации словоформы. О **семантической аддитивности** формативов в словоформе мы будем говорить, если этим формативам можно поставить в соответствие некие элементы плана содержания, которые можно сформулировать в виде достаточно простых словоизменительных признаков («it can be captured with reference to a feature value» [Hesse 2023: 10]), или, иначе, если их дистрибуции не являются морфомными, а также если при этом грамматическое значение словоформы может быть получено из грамматических характеристик отдельных формативов («given (...) the characterization of the feature values in its feature specification, we can predict the meaning of the whole» [Round, Corbett 2017: 26]).

Позиционный класс мы понимаем вслед за [Stump 1991] как класс аффиксов, употребляющихся в определенной позиции в составе цепочки аффиксов, оформляющих словоформу.

3. Глагольные системы андийских языков

Андийские языки — группа в составе аваро-андо-цезской ветви нахско-дагестанских языков. Традиционно в составе этой группы выделяют восемь языков: андийский, ахвахский, багвалинский, ботлихский, годоберинский, каратинский, тиндинский, чамалинский. Зона исконного расселения носителей этих языков — северо-запад Горного Дагестана, а именно Ахвахский, Ботлихский, Цумадинский и отчасти Шамильский районы Республики Дагестан [Алексеев 1988: 3].

Несколько отвлекаясь от конкретноразговорных особенностей, ядерные, наиболее частотные формы глагольных систем андийских языков по данным [Kibrik et al. (eds.) 1996; Кибрик и др. 2001; Alexeyev, Verhees (to appear); Creissels 2018; Kaye, Maisak 2023; Магомедбекова 1967; 1971; Майсак 2016] можно охарактеризовать следующим образом. К формам плана прошедшего относятся нейтральный аорист (или претерит), противопоставленный эвиденциальному перфекту, который может иметь и другие значения (результативного состояния, текущей релевантности, см. подробнее в [Ферхеес 2019]). Среди форм с временной референцией к настоящему обычно противопоставлены хабитуалис

и немаркированный презенс, в некоторых идиомах им дополнительно противопоставлен прогрессив (например, в андийском языке [Кауе, Maisak 2023]), специализированно обозначающий актуальный процесс в момент речи. Зона будущего обслуживается формой футура или потенциалиса. Часто в системах также представлен проспектив. Формы со значением модальности, ориентированной на говорящего, включают императив и оптатив, иногда в системах представлены специализированные формы императива и оптатива множественного числа. В некоторых случаях мы будем привлекать к рассмотрению дополнительные формы, например, причастия и общие конвербы, морфология которых значительно различается в зависимости от системы [Алексеев 1988]. Поскольку в данной работе нас интересует в первую очередь морфологическая проблематика, мы ограничимся эмпирический материал только синтетическими формами.

Ярлыки для этих форм условны, поскольку не для всех идиомов доступны данные, которые бы позволили установить точные особенности их употребления. Выбор ярлыков во многом следует описательной традиции, заложенной работами [Kibrik et al. (eds.) 1996; Кибрик и др. 2001].

Теперь мы рассмотрим свойства нескольких андийских систем: на примере анчихского диалекта каратинского языка (с привлечением данных по кванадинскому диалекту багвалинского языка) мы исследуем полиформативную систему с преобладающей морфотактической неэлементарностью средств выражения грамматических значений (подраздел 3.1); а тиндинская и чамалинская системы послужат примером систем со значительной формальной неаддитивностью (подраздел 3.2).

3.1. Полиформативные системы

В качестве примера полиформативной системы рассмотрим глагольную систему анчихского диалекта каратинского языка, данные о формах и грамматических ярлыках приводятся по [Филатов 2020] и, в меньшей степени, по [Магомедбекова 1971] (табл. 4, с. 39). Оговоримся сразу, что нас будут интересовать только категориальные свойства суффиксального словоизменения, поскольку префиксально в глаголе выражается лишь согласовательный класс. Все словоформы в таблице ниже даны в форме среднего согласовательного класса *b-* 'N'.

В этой таблице даны основные глагольные формы анчихской системы на примере лексемы *b-ik* 'N-a: [N-быть-INF] 'быть'. Столбцы в таблице соответствуют количеству формативов, содержащихся в словоформах. Например, форма хабитуалиса *b-ik* 'ur-a [N-быть-IPFV-нав] вписана во второй столбец, поскольку содержит два форматива *-ur* и *-a*. Словоформы были сегментированы таким образом, чтобы выделить наименьшие формативы, способные противопоставлять различные формы.

Можно заметить, что формы, расположенные в столбцах правее, синтагматически более сложны, чем расположенные в столбцах левее (например, формы в третьем столбце содержат два форматива, а во втором — один). Причем для словоформ в таблице выполняется еще одно свойство: форма в столбце справа содержит цепочку формативов, «унаследованную» от формы в столбце слева, а также еще некоторые формативы (к этому свойству анчихской системы мы еще вернемся ниже в этом разделе). Можно сказать, что формы в столбце справа образованы от форм в столбце слева.

Пустая клетка означает, что автономной словоформы, которая бы выполняла эти требования, не существует. Форма имперфектива помечена звездочкой, поскольку она не является полностью автономной словоформой. Она выступает только в составе аналитических форм, таких как прогрессив *gah-ir gira* [делать-IPFV COP] 'делает (в настоящий момент)'. Глоссы, которые мы приписываем формативам (как будет видно из дальнейшего изложения), в значительной мере условные.

Таблица 4

**Глагольная система анчихского диалекта каратинского языка
(на примере глагола *bik'wa*: 'быть')**

Подсистема	Один форматив	Словоформа содержит	
		Два форматива	Три или четыре форматива
Подсистема перфектива	Аорист <i>b-ik'w-a</i> N-БЫТЬ-AOR 'БЫЛ'	Перфективное причастие <i>b-ik'w-a-b</i> N-БЫТЬ-PTCP-N 'БЫВШИЙ'	
		Перфективный конверб <i>b-ik'w-a-s</i> N-БЫТЬ-AOR-CVB 'БЫВ'	
		Отрицательный аорист <i>b-ik'w-a-č'e</i> N-БЫТЬ-AOR-NEG 'не был'	Отрицательное перфективное причастие <i>b-ik'w-a-č'-o-b</i> N-БЫТЬ-AOR-NEG-PTCP-N 'не бывший'
		Императив (мн. ч.) <i>b-ik'w-a-bi</i> N-БЫТЬ-AOR-IMP.PL 'будьте!'	Оптитив (мн. ч.) <i>b-ik'w-a-bi-λ'a</i> N-БЫТЬ-AOR-IMP.PL-OPT 'чтоб вы были!'
Подсистема имперфектива	Имперфектив <i>b-ik'-ur*</i> N-БЫТЬ-IPFV	Хабитуалис <i>b-ik'-ur-a</i> N-БЫТЬ-IPFV-HAB 'бывает'	Имперфективный конверб <i>b-ik'-ur-a-ra</i> N-БЫТЬ-IPFV-HAB-ICVB 'бывая'
			Имперфективное причастие <i>b-ik'-ur-o-b</i> N-БЫТЬ-IPFV-PTCP-N 'бывающий'
Подсистема отрицательного имперфектива		Отрицательный хабитуалис <i>b-ik'-u(b)a-č'e</i> N-БЫТЬ-NEG.IPFV(N)-NEG 'не бывает'	
		Отрицательный имперфективный конверб <i>b-ik'-u(b)a-ko</i> N-БЫТЬ-NEG.IPFV(N)-NEG.CVB 'не бывая'	
Подсистема ирреалиса		Прохибитив <i>b-ik'-ubi-s:e</i> N-БЫТЬ-IRR-PRON 'не будь!'	Оптитивный прохибитив <i>b-ik'-ubi-s:e-λ'a</i> N-БЫТЬ-IRR-PRON-OPT 'чтоб ты не был!'
		Отрицательный потенциалис <i>b-ik'-ubi-č'e</i> N-БЫТЬ-IRR-NEG 'не будет'	Отрицательное потенциальное причастие <i>b-ik'-ubi-č'-o-b</i> N-БЫТЬ-IRR-NEG-PTCP-N 'который не будет'

Подсистема	Словоформа содержит		
	Один форматив	Два форматива	Три или четыре форматива
Подсистема императива	Императив <i>b-ik' -u</i> N-БЫТЬ-IMP.INTR 'будь!'	Оптиматив <i>b-ik' -u-λ' a</i> N-БЫТЬ-IMP.INTR-OPT 'чтоб ты был!'	
Подсистема инфинитива	Инфинитив <i>b-ik' ^w-a:</i> N-БЫТЬ-INF 'быть'	Потенциалис <i>b-ik' ^w-a:-s: e</i> N-БЫТЬ-INF-POT 'будет'	Потенциальное причастие <i>b-ik' ^w-a:-λo-b</i> N-БЫТЬ-INF-POT.PTSP-N 'будущий'

Попытаемся ответить на интересующий нас вопрос о соотношении структуры формальных противопоставлений и структуры семантических противопоставлений в анчихской системе. Для этого требуется проанализировать формальную аддитивность формативов и то, насколько хорошо выделенные формативы соотносятся с какими-либо словоизменительными признаками (т. е. установить их семантическую аддитивность).

Начнем с вопроса о формальной аддитивности формативов. В целом можно признать, что большинство формативов в значительной степени аддитивны: границы между ними проводятся достаточно однозначно, а случаи фузии отсутствуют. Единственным отклонением от формальной аддитивности можно признать правило выпадения /e/ при сочетании формативов *-č' e* [-NEG] + *-o-b* [-PTSP-N] → *-č' -o-b* [-NEG-PTSP-N].

При этом установление семантической аддитивности вызывает гораздо большие трудности. Практически все грамматические значения словоформ в этой системе нельзя представить как композицию более простых семантических единиц, соотносимых с парадигматической дистрибуцией формативов. Например, форматив *-ubi* участвует в выражении значений прохибитива, отрицательного потенциалиса и отрицательного потенциального причастия. На этом основании можно предположить, что он выражает смысл 'отрицание'. Однако форма отрицательного потенциалиса *b-ik' -ubi-č' e* показывает, что это не так: второй форматив *-č' e*, исходя из его дистрибуции, тоже приходится признать соотносящимся со смыслом 'отрицание': он участвует в выражении таких грамматических значений, как отрицательный аорист (*b-ik' ^w-a-č' e*) и отрицательный имперфектив (*b-ik' -u(b) a-č' e*). Тем самым, можно сделать два вывода: (i) смыслу 'потенциалис' не удается поставить в соответствие никакого форматива в составе этой словоформы; (ii) парадигматическая дистрибуция показателя *-ubi* не описывается в семантических терминах, и его вклад в семантику грамматического значения формы отрицательного потенциалиса не может считаться аддитивным. Оба этих вывода свидетельствуют о семантической неаддитивности этого участка парадигмы.

Рассмотрим еще ряд примеров, чтобы убедиться в том, что семантическая неаддитивность — существенное свойство анчихской системы. Если сравнить формы инфинитива, потенциалиса и потенциального причастия, то можно заметить, что в них во всех есть форматив *-a:*, который в отсутствие других формативов образует форму инфинитива. При этом нельзя сказать, что форматив *-a:* — экспонент смысла 'инфинитив', поскольку значения форм потенциалиса и потенциального причастия никак нельзя получить при помощи аддитивной композиции смысла 'инфинитив' и еще некоторого добавления, вносимого, предположительно, формативами *-s: e* или *-λo-b*. Более того, в случае аддитивной организации парадигмы ожидался бы параллелизм в составе формативов между потенциалисом и потенциальным причастием, с одной стороны, и их отрицательными соответствиями — отрицательным потенциалисом и отрицательным потенциальным причастием — с другой. При каноническом агглютинативном устройстве системы они отличались бы только на показатель отрицания, тогда как в анчихской системе для выражения смыслов 'потенциалис' и 'отрицательный потенциалис' используются разные комбинации формативов — *-a:-s: e* и *-ubi-č' e* соответственно.

Еще более яркий пример семантической неаддитивности находим в образовании императива множественного числа (табл. 5). В случае семантической аддитивности этого участка парадигмы мы бы ожидали, что императив множественного числа содержал бы форматив императива единственного числа. Тем не менее табл. 5 показывает, что это не так — можно соотнести форму императива множественного числа с формами императива единственного числа и аориста, чтобы убедиться в том, что императив множественного числа образован при помощи морфонологически измененного формата аориста. Аорист образуется при помощи лексически распределенных формативов *-a / -e*, формирующих два словоизменительных класса: А-спряжение и Е-спряжение. Показатели императива (*-a / -i*) распределены в соответствии с переходностью глагола: переходные глаголы присоединяют форматив *-a*, а непереходные — форматив *-i*³. Данные в табл. 5 свидетельствуют о том, что распределение формата императива множественного числа (*-a / -i*) зависит от спряжения, а не переходности глагола.

Таблица 5

**Императив множественного числа и его связь с аористом
в анчихском диалекте каратинского языка**

	Императив ед. ч.	Императив мн. ч.	Аорист	Переходность	Спряжение
‘смотреть’	<i>haʔ-i</i>	<i>haʔ-a-bi</i>	<i>haʔ-a</i>	INTR	А
‘идти’	<i>beʔ-i</i>	<i>beʔ-i-bi</i>	<i>beʔ-e</i>	INTR	Е
‘жарить’	<i>bež-a</i>	<i>bež-a-bi</i>	<i>bež-a</i>	TR	А
‘брат’	<i>bah-a</i>	<i>bah-i-bi</i>	<i>bah-e</i>	TR	Е

Возвращаясь к табл. 4, в этом же смысле семантически неаддитивными приходится признать и сочетания формативов, образующих хабитуалис (*-ur-a* [-IPFV-НАВ]), имперфективный конверб (*-ur-a-ra* [-IPFV-НАВ-ICVB]), оптатив (*-u-ʔ'a* [-IMP.INTR-ОПТ]) и отрицательный оптатив (*-a-bi-ʔ'a* [-AOR-IMP.PL-ОПТ]).

Как и в примерах, рассмотренных выше, грамматические значения этих словоформ нельзя представить как композицию значений, приписываемых входящим в их состав формативам, по двум причинам. Во-первых, формативу невозможно приписать «значение», исходя из его парадигматической дистрибуции: таков случай «перфективного» формата *-a* или *-i / e*, который неожиданно находится в формах императива и оптатива множественного числа. Во-вторых, даже если условное «значение» приписать удастся (как в случае, например, с имперфективом — форматив *-ur* присоединяется только к имперфективным формам), то аддитивность все равно не соблюдена — это «значение» не позволяет композиционно получить грамматические значения словоформ, содержащих этот форматив. Все это указывает на то, что мы имеем дело с многочисленными случаями **морфологической идиоматичности**, иначе, морфологическими фразами [Мельчук 1997], или гештальт-выражениями (*gestalt-exponence*) в терминологии современных словесно-парадигматических моделей [Blevins et al. 2019: 276]: «An individual wordform is often analysable into parts that recur elsewhere in its inflectional paradigm or in the morphological system at large. But these parts may function solely to differentiate larger forms, so that the minimal parts that distinguish a pair of wordforms cannot be associated with the difference in grammatical meaning between these forms».

³ В табл. 4 выше представлен также показатель императива *-u*: *b-ik'-u* [N-БЫТЬ-IMP.INTR] ‘будь!’. Это морфонологический вариант показателя императива непереходных глаголов *-i* при лабиализованных основах. Ср. форму аориста *b-ik'-a* [N-БЫТЬ-АОР] ‘был’, где лабиализация основы видна в поверхностной форме.

В сущности, грамматические значения выражают не отдельные формативы, а идиоматические комплексы формально вычленимых формативов. Отдельные формативы в идиоматических комплексах являются **морфомными** — их парадигматическая дистрибуция не описывается в семантических терминах. Распространенный подход в реализационных моделях заключается в том, чтобы морфомные формативы относить к основе (см., например, [Spencer 2012; Blevins 2003]). Это соответствует традиционному дагестановедческому подходу, ср. термины «основа перфектива», «основа инфинитива» в [Кибрик и др. 2001]. В то же время в [Baerman, Corbett 2012] отмечается следующее: «[I]f there are multiple layers of formants we may struggle to determine whether we have a complex stem or a complex affix» [Ibid.: 65]. Морфомными формативами в анчихской системе являются, например, уже обсуждавшиеся выше *-a / -i* (условная глосса AOR), *-ubi* (условная глосса IRR), *-a:* (условная глосса INF). Вместе с тем дистрибуция показателя *-ur* (условная глосса IPFV), согласно некоторым взглядам на морфомные явления, может быть признана не морфомной, а морфемной: в [Herce 2023: 91] таким образом анализируется распределение перфективной и имперфективной основ в латинском языке: «[O]ne stem (*fāc-*) appears in imperfective tenses and another one (*fēc-*) in the perfective ones. This is, therefore, a natural/morphemic alternation». Если следовать данной логике, возникает дескриптивная асимметрия: в отличие от основ в латинском, анчихская «перфективная основа» *-a / -i* (условная глосса AOR) — морфомна (за счет присутствия ее форматива в формах императива мн. ч.), а «имперфективная» *-ur* (условная глосса IPFV) — морфемна (даже с учетом ее семантической неаддитивности, например, в составе формы *хабитуалиса*). На наш взгляд, эти проблемы связаны с общими трудностями однозначного определения границы между морфомными и морфемными явлениями (см. раздел 4 ниже).

Формативы не появляются в словоформе независимо друг от друга, то есть имеют **ограничения на взаимную селективность** — их комбинаторный потенциал невелик. Например, выбор *-ur* в качестве форматива первой позиции ограничивает множество формативов второй позиции до *-a* и *-o:* присоединение первого образует форму *хабитуалиса* или имперфективного конверба, второго — форму имперфективного причастия. Некоторые формативы, например форматив потенциального причастия *-λo* или форматив отрицательного имперфективного конверба *-ko*, изолированы в парадигме и не участвуют в образовании каких-либо других форм.

Какую же в таком случае категориальную структуру имеет система анчихских синтетических форм в целом? Рассмотренные примеры показывают, что далеко не все комбинации грамматических значений возможны в рамках синтетической парадигмы, что было бы условием для выделения двух (или нескольких) ортогональных категорий: напротив, большинство комбинаций фиксированы. Например, смыслы ‘перфектив’ + ‘прошедшее’ закреплены в системе лишь за аористом и отрицательным аористом, а к формам плана настоящего относятся лишь имперфективные формы — *хабитуалис* и отрицательный *хабитуалис*.

Наиболее подходящим кандидатом на формирование отдельной «прозрачной» категории, все значения которой выражаются аддитивно, в анчихской системе является отрицание, которое практически во всех словоформах имеет собственное выражение *-č'e / -č'*. Тем не менее ортогональность отрицания и ТАМ-значений является неполной. Не все отрицательные формы в анчихской системе дистрибутивно соответствуют своим утвердительным парам. В ряде случаев отрицательная форма соответствует нескольким утвердительным, то есть имеет более широкое распределение, чем каждая из них. Это можно заметить по отсутствию в табл. 4 отрицательного коррелята, к примеру, для формы перфективного конверба. Форма отрицательного аориста обслуживает как финитные (1), так и конвербиальные (2) употребления, так что специальная форма отрицательного перфективного конверба отсутствует.

- | | | | | |
|-----|---------------|----------------------|--------------------------|---------------------|
| (1) | <i>hark'a</i> | <i>q'ʷab-i-č'e</i> , | <i>hinč'-o-b</i> | <i>m-as-ir-a</i> |
| | глаз | моргать-AOR-NEG | быть_неправильным-PTCP-N | N-говорить-IPFV-NAВ |

hor-do-l

MED-OBL.HPL-ERG

‘Они врут, не моргнув глазом’. [Гаджимагомедов 2019: 83]

(2) *ho-guni-s: t:ɛj di-ja b-aq'-i-č'e*

MED-PLACE-ATTR вода я-DAT N-ПОДХОДИТЬ-AOR-NEG

‘Тамошняя вода мне не подошла’. [Там же: 92]

Аналогичным образом в системе нет отдельной морфологической формы отрицательного имперфективного причастия — соответствующие употребления имеет форма отрицательного перфективного причастия [Магомедбекова 1971: 153]. Эти примеры указывают на ограниченный комбинаторный потенциал отрицания: его сочетания с некоторыми грамматическими значениями (в данном случае имперфективного причастия или перфективного конверба) не запрещены, но и не выражаются специализированными средствами. Согласно [Corbett 2012: 156], такая ситуация является неканонической: «Canonical features and their values have dedicated forms (are ‘autonomous’)». Нельзя говорить и об аддитивности отрицания в этих формах — полное грамматическое значение, например, формы отрицательного аориста *b-ik'w-a-č'e* [N-быть-AOR-NEG] нельзя получить простой комбинацией смыслов ‘аорист’ и ‘отрицание’ — оно не будет включать информацию о том, что эта форма может употребляться и нефинитно.

Еще менее ортогональна, чем отрицание, возможная категория репрезентации, которая бы противопоставляла финитным формам формы конвербов и причастий, так как формы конвербов и причастий не образуются от неиндикативных форм⁴.

Таким образом, можно заключить, что устройство анчихской системы в значительной степени напоминает устройство линейных систем: ее точно так же невозможно разбить на несколько полностью ортогональных категорий, и даже отрицание и репрезентация в этой системе обладают неполной ортогональностью. При этом структура, задаваемая в парадигме противопоставляющими словоформы формативами, и структура семантических противопоставлений неизоморфны.

Стоит кратко коснуться еще одной особенности андийских систем, которая связана с ограничениями на взаимную селективность в идиоматичных комплексах формативов. В полиформативных системах наблюдается особый порядок организации парадигмы, который мы называем иерархическим. В таких системах наличие и наполнение каждой следующей позиции в словоформе определяется предыдущей позицией.

Иерархическую структуру парадигмы удобно рассмотреть на примере багвалинской системы (на рис. 1 мы приводим слегка модифицированную и сокращенную схему из [Кибрик и др. 2001: 69] на примере форм глагола *et-a*: [летать-INF ‘летать’]). Как видно, наличие в структуре словоформы форматива, выражающего значение императива, открывает следующую позицию, причем она может быть заполнена только формативом, участвующим в выражении опатива. Инфинитив в свою очередь допускает заполнение следующей позиции формативом будущего синтетического, деепричастия или причастия будущего времени. Позиционная структура, таким образом, не едина для всех словоформ, а определяется в значительной степени формативом первой послекорневой позиции, на основании чего все словоформы можно объединить в подсистемы (для багвалинской системы обозначены на рис. 1 обводками, а для анчихской — указаны в первом столбце табл. 4).

Из-за идиоматичности комплексов формативов иерархичность — чисто формальное свойство подобных систем: формы, имеющие общий семантический компонент,

⁴ Вообще говоря, тот факт, что в глагольных системах нередка ситуация запрещенных комбинаций грамем, известен и обсуждается, например, в монографии [Храковский, Мальчуков (ред.) 2020] и более ранних работах Санкт-Петербургской типологической школы, посвященных взаимодействию грамматических категорий. Тем не менее для наших целей пока достаточно ограничиться выводом о неполной ортогональности рассматриваемых категорий.

могут принадлежать к разным подсистемам. Так, будущее синтетическое в багвалинском относится к подсистеме инфинитива, а будущее отрицательное — к подсистеме имперфектива.

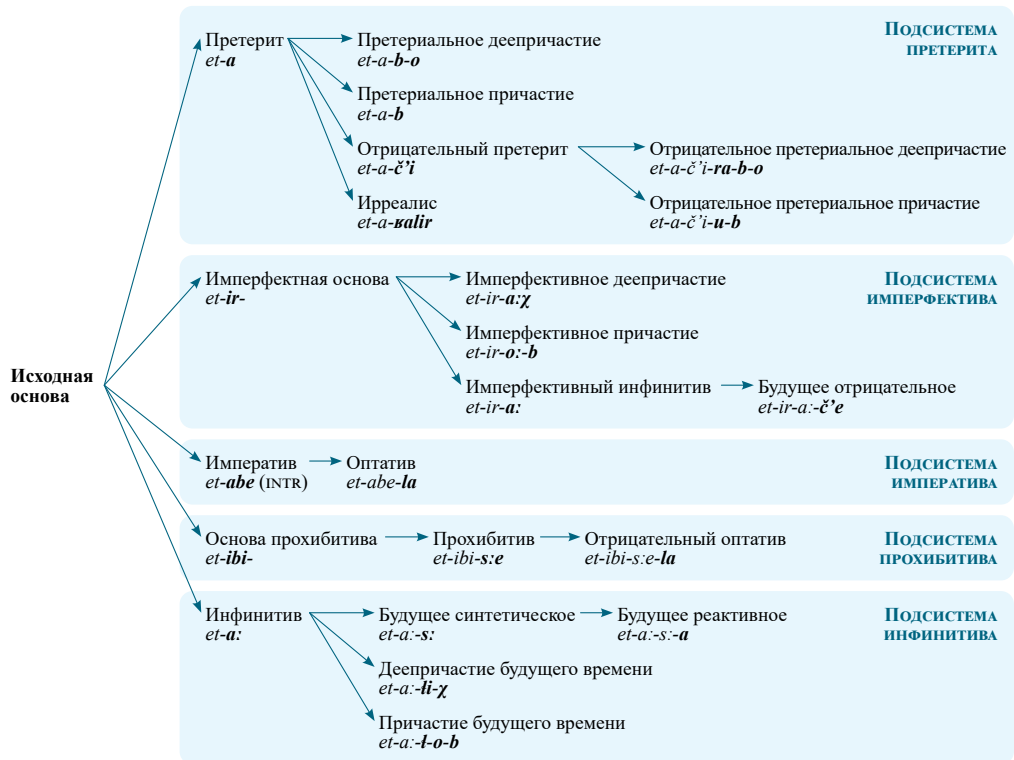


Рис. 1. Глагольная система багвалинского языка по [Кибрик и др. 2001: 69], формы глагола *et-a:* [летать-INF] 'летать'

Отметим, что похожим свойством обладает словоизменительная система латинского глагола, в которой представлены три морфомных основы, из которых, например, т. н. третья основа, образованная от перфектного причастия (которое зачастую имеет пассивное значение), служит для образования причастия будущего времени, которое, в свою очередь, значение пассива никогда не выражает. Ср. замечание в [Aronoff 1994: 31–33] о несовпадении принадлежности к подсистеме и выражаемой формой семантики: «If we derive the future participle from the perfect participle, then what do we do with the passive meaning and syntax of the perfect participle?»⁵

3.2. Формально-неаддитивные системы

Рассмотрим теперь фрагменты систем двух других андийских языков — тиндинского и чамалинского, которые обладают свойством формальной неаддитивности.

Начнем с рассмотрения тиндинской системы (табл. 6; данные даются по описанию собственно-тиндинского диалекта [Authier (to appear)]). К синтетическим глагольным формам

⁵ Мы благодарим одного из анонимных рецензентов за указание на эту аналогию.

индикатива тиндинского языка, согласно описанию Ж. Отье, можно отнести семь форм: инфинитив, хабитуалис, отрицательный хабитуалис, аорист, отрицательный аорист, перфект и дебитив. Из этих форм две — перфект и дебитив — имеют пары дублетных форм и могут выражаться как синтетически, так и аналитически. По [Authier (to appear): 18–19], эвентуалис — грамматическая форма, совмещающая значения хабитуалиса и будущего времени с низкой степенью уверенности («less than certain future»). Дебитив — форма, обозначающая необходимость или будущее время с высокой степенью уверенности («obligation or certainty in the future»). Формы эвентуалиса и аориста могут также функционировать и как имперфективный и перфективный конвербы соответственно.

Тиндинская система, на наш взгляд, наиболее близка по своим свойствам к прототипическим линейным системам типа системы кпелле (табл. 3). Если не учитывать дублетные аналитические формы, все формативы в словоформах, выражающие грамматические значения, формально нечленимы. Поскольку комбинаторный потенциал грамматических значений в системе практически отсутствует, выделить несколько ортогональных категорий здесь затруднительно. Единственной категорией, которая могла бы быть введена, является отрицание. Ее значения в двух случаях (отрицательный эвентуалис и отрицательный аорист) комбинируются с ТАМ-значениями, однако в этом случае, по всей видимости, можно говорить лишь о маргинальной ортогональности — отрицание остальных грамматических значений выражается неморфологически [Authier (to appear): 19–20]. Ясно также, что и аддитивность в такой системе наблюдается только между лексическим значением, выражаемым корнем, и грамматическим значением целиком, выражаемым нечленимой флексией.

Таблица 6

**Синтетические индикативные формы глагола *ihilja* ‘делать’ в тиндинском языке
(по [Authier (to appear)])**

Грамматический ярлык	Форма
Инфинитив	<i>ih-ilja</i> делать-INF ‘делать’
Эвентуалис (Имперфективный конверб)	<i>ih-a:</i> делать-NAV ‘(обычно) делает’
Отрицательный эвентуалис	<i>ih-ih</i> делать-NAV.NEG ‘(обычно) не делает’
Аорист (Перфективный конверб)	<i>ih-o</i> делать-AOR ‘сделал’
Отрицательный аорист	<i>ih-e:</i> делать-AOR.NEG ‘не сделал’
Эвиденциальный перфект	<i>ih-o: / ih-o ija</i> делать-PF / делать-AOR COP ‘сделал (говорят)’
Дебитив	<i>ih-ile / ih-ilja ija</i> делать-DEV / делать-INF COP ‘(обязательно) сделает’

На наш взгляд, отсутствие формальной аддитивности грамматических формативов в ряде случаев может объясняться действием фузионных процессов. Как отмечается в [Магомедова 2012] и [Authier (to appear)], синтетические формы перфекта (*ih-o:*) и дебитива (*ih-ile:*) не что иное, как стяженные аналитические формы, то есть *ih-o:* [делать-РФ] ‘сделал (незасв.)’ < *ih-o ija* [делать-АОР СОР] ‘сделал (незасв.)’, *ih-ile:* [делать-ДЕВ] ‘сделает’ < *ih-ijja ija* [делать-ІНФ СОР] ‘сделает’. Это подтверждают также и данные акнадинско-ангидинского диалекта тиндинского языка (приводятся по нашим полевым данным), где соответствующие когнатные формы сохранили полиформативную структуру: *ih-o-ja* [делать-АОР-СОР] ‘сделал (незасв.)’, *ih-ila-ja* [делать-ІНФ-СОР] ‘сделает’.

Глагольную систему, претерпевшую фузионные преобразования, можно наблюдать и в чамалинском языке (табл. 7).

Таблица 7

Фрагмент глагольной системы чамалинского языка (по [Бокарев 1949])

Глагол	‘молоть’	‘жать’	‘шить’
Перфективный конверб	<i>aq-i</i>	<i>ak^w-a</i>	<i>bak[’]:^w-ĩ</i>
Перфект	<i>aqi:da</i>	<i>ak^we:da</i>	<i>bak[’]:ũjda</i>
Футур	<i>aqida</i>	<i>akuda</i>	<i>bak[’]:unna</i>
Хабитуалис	<i>aqide:da</i>	<i>akude:da</i>	<i>bak[’]:unne:da</i>

Как показывают данные А. А. Бокарева [1949], формы перфекта и хабитуалиса диахронически представляли собой сочетание перфективного конверба и футура⁶ со связкой. В процессе грамматикализации эти формы подверглись фузионной компрессии, в результате которой морфемный шов стал скрыт. Так, для форм хабитуалиса можно предположить следующее развитие: *aq-ida + ida > aqide:da* [молоть:НАВ] ‘мелет’, *akuda + ida > akude:da* [жать:НАВ] ‘жнет’, *bak[’]:unna + ida > bak[’]:unne:da* [шить:НАВ] ‘шьет’. При этом формы перфекта еще более показательны, поскольку в этих формах в разных словоизменительных классах взаимодействие бывшей связки с основой произошло по-разному, ср.: *aqi + ida > aqi:da* [молоть:РФ] ‘смолол’, *ak^w-a + ida > ak^we:da* [жать:РФ] ‘сжал’, *bak[’]:^w-ĩ + ida > bak[’]:ũjda* [шить:РФ] ‘сшил’. В результате полиформативность перфекта и хабитуалиса здесь устанавливается только на уровне глубинного (морфонологического) представления.

4. Типологические и диахронические наблюдения над андийскими системами

Теперь попытаемся сделать некоторые диахронические наблюдения над рассмотренными выше особенностями андийских систем и провести к ним возможные типологические параллели.

Формальная аддитивность как шкала. Мы установили, что андийские глагольные системы можно предварительно разбить на три подтипа: полиформативный (анчихская система), фузионный (чамалинская система) и моноформативный (тиндинская система). Полиформативный тип характеризуется значительной формальной аддитивностью формативов при малой семантической аддитивности; фузионный тип также не обладает семантической аддитивностью, а его формальная аддитивность частична: в силу действия

⁶ В данном случае мы отвлекаемся от семантического сдвига, который претерпели формы футура («старого хабитуалиса») и хабитуалиса («старого презенса»), известного также как презентный дрейф, об этом см. [Bybee et al. 1994; Майсак 2008] среди проч.

фузионных морфонологических правил аддитивность может быть установлена лишь для глубинных форм формативов, а в поверхностном выражении проведение границ между ними затруднено; моноформативный тип не обладает формальной аддитивностью, а его семантическая аддитивность тем самым тривиальна.

Нетрудно заметить, что три этих типа систем располагаются вдоль континуума аддитивно-фузионных свойств словоформы [Плунгян 2016: 70–71]. Как кажется, этот факт позволяет выдвинуть гипотезу о том, что возникновение этих типов может объясняться разной интенсивностью процессов морфологизации и фузионной компрессии. Идиоматические комплексы формативов образуются в условиях морфологизации, более интенсивной, чем процессы фузионной компрессии, то есть тогда, когда неаддитивность значения уже достигнута⁷, а форму все еще можно представить через комплекс парадигматически противопоставленных формативов.

Неизоморфность структур и асинхронность грамматикализации. Как показал анализ, ни одна из рассмотренных систем не обладает полной изоморфностью между структурами формальных и семантических противопоставлений. На наш взгляд, причины этого несоответствия кроются, с одной стороны, в составе грамматических значений этих систем, а с другой — в особенностях процессов грамматикализации, протекающих в таких системах.

Если не считать отрицания, которое в целом лучше других категорий подвергается ортогональному анализу, то словоформы рассмотренных андийских систем организованы вокруг ТАМ-значений. В этом смысле андийские данные подтверждают наблюдение Б. Эрсе о малой типологической засвидетельствованности ортогонально организованных категорий, состоящих из ТАМ-значений: «Whereas some features, like person and number, or case and number, are generally well-behaved regarding their (logical) orthogonality, others like TAM are more troublesome. Thus, it is often difficult to find a perfect orthogonality of tense and aspect, aspect and mood, or tense and mood» [Herce 2023: 123]⁸.

Данные показывают, что в этих системах протекают достаточно разнородные процессы грамматикализации: каждая форма проходит свой путь грамматикализации и новейший слой морфологии в этих формах грамматикализуется из источника независимо от других форм. Это видно, например, по отсутствию общего деепричастного показателя у форм претеритального (-b-o) и имперфективного (-a:χ) деепричастий в багвалинской системе (рис. 1). Об этом же говорят разные источники грамматикализации будущего синтетического (-s:) и нефинитных форм с временной референцией к будущему, деепричастия и причастия (-li/-l) — это результат морфологизации разного лексического материала в ходе разных циклов грамматикализации.

Такая материальная «неоднородность», по-видимому, может объясняться разной диахронической стабильностью грамматических значений в системе. Дж. Николс указывает: «Diversity arises when some element is relatively unstable and therefore prone to replacement in various ways» [Nichols 2003: 283], — а значит, стабильные и нестабильные граммемы в системе можно устанавливать сравнительно-историческим методом. Например, по данным сравнительной грамматики М. Е. Алексеева, аористные формы в андийских языках

⁷ Строго говоря, неаддитивность значения может быть достигнута и без морфологизации: данные в [Ферхеес 2019] указывают на то, что, хотя обсуждавшийся выше эвиденциальный перфект в ряде андийских языков (багвалинском, каратинском, северно-ахвахском) является аналитическим, его значение тем не менее нельзя композиционально представить через значение копулы и лексического глагола.

⁸ Внимательный читатель мог заметить, что в рассмотренном нами примере многомерно-агглютинативной системы в узбекском языке (табл. 2) все ТАМ-значения также организованы в **единую** категорию в рамках одного позиционного класса аффиксов. Более того, пример линейной системы кпелле также содержит исключительно ТАМ-значения. Все это подтверждает наблюдение Б. Эрсе.

достаточно стабильны и могут быть реконструированы на праандийском уровне: «Единственной неописательной формой, наблюдаемой более или менее последовательно во всех языках, является форма аориста» [Алексеев 1988: 107]. В то же время формы будущего времени затруднительно реконструировать на праандийском уровне, что означает, что в разных языках произошли разные, более поздние и, возможно, независимые замены этих форм: «В значительной степени проблематичной представляется реконструкция общеандийского показателя буд[ущего] времени» [Там же: 110]. Неодинаковая стабильность грамматических значений в системе ведет к тому, что они обновляются асинхронно, а асинхронность может приводить к тому, что грамматикализированным оказывается материал из разных конструкций. Связанная с этим проблема — сохранение более архаичных форм в неассертивных контекстах [Bybee et al. 1994; Гусев 2006]. К примеру, в анчиской системе результат действия этого принципа можно предполагать в уже обсуждавшейся в подразделе 3.1 асимметрии между различием аориста и перфективного конверба в утвердительных формах и совмещением этих функций в одной отрицательной форме. М. Е. Алексеев [1988: 112] предполагает «исконное совпадение финитного и нефинитного оформления общеандийских глаголов», а также то, что «нынешняя система деепричастий сложилась уже в каждом из андийских языков независимо» [Там же: 111]. Как кажется, это можно интерпретировать в пользу того, что более архаичное состояние системы — аористо-конвербиальная полисемия — сохранилось в контексте отрицания, способствующем сохранению архаичных структур, в то время как более инновативные утвердительные формы обновились независимо.

Кроме того, образованию ортогональных категорий препятствуют более новые слои морфологизации. Они могут нарушать уже сложившиеся позиционные структуры. Это можно увидеть на примере утвердительных и отрицательных форм футура и хабитуалиса в сравнении с аористом и прогрессивом в гагатлинском диалекте андийского языка (табл. 8).

Таблица 8

Формы глагола *roxudi* ‘чесать’ в гагатлинском диалекте андийского языка
(по [Kaye, Maisak 2023])

Грамматический ярлык	Утвердительные формы	Отрицательные формы
Хабитуалис	<i>roxu-do</i>	<i>roxu-do-s:u</i>
Футур	<i>roxu-dʰa</i> (< <i>*-do-ja</i>)	<i>roxu-dos:ʰa</i> (< <i>*-do-s:u-ja</i>)
Аорист	<i>roxo</i>	<i>roxo-s:u</i>
Прогрессив	<i>roxo-rado</i>	<i>roxo-rado-s:u</i>

Если хабитуалис образует отрицательную форму вполне регулярно и конкатенативно — при помощи показателя *-s:u*, характерного и для большинства других глагольных форм, то в футуре обнаруживаются фузионные показатели *-dʰa* и *-dos:ʰa*, которые в [Kaye, Maisak 2023] возводятся к сочетаниям **-do-ja* и **-do-s:u-ja*. Во-первых, фузионное взаимодействие между **-do* и **-ja* или **-s:u* и **-ja* нарушает формальную аддитивность формативов. Во-вторых, в этом случае континуант форматива отрицания *-s:* находится в несвойственной ему линейной позиции, поскольку обычная позиция показателя отрицания финитных форм в андийском языке — крайняя правая (ср. отрицательные формы аориста и прогрессива в табл. 8). Тем самым позиционный класс показателя отрицания оказывается нарушен.

С нашей точки зрения, как разнородность морфологизируемого материала, так и нарушения позиционной структуры словоформ при морфологизации можно рассматривать как свидетельства асинхронной морфологизации, предотвращающей формирование единой позиционной структуры для всех словоформ.

Диахрония морфомных и иерархических структур. Иерархическая структура парадигмы в андийских полиформативных системах также отражает известное свойство грамматикализации, заключающееся в сосуществовании внутри системы разных хронологических слоев (многослойность, англ. *layering* в терминах [Hopper 1991]): показатели, находящиеся ближе к концу словоформы, представляют более новые слои, имеющие менее размытые значения и линейно расположенные дальше от корня. Соответственно, показатели, расположенные ближе к корню, — это более старые слои морфологизованного материала, имеющие более размытые значения и меньшую морфологическую автономность.

Попытаемся теперь соотнести свойства формативов, которые мы в подразделе 3.1 выше идентифицировали как морфомные с тем, что известно о типологии и диахронии морфом. По определению, например [Blevins 2003: 737], морфомные формативы «участвуют в построении разнородных форм» («underlie a heterogeneous class of forms») или «не вносят непосредственного семантического вклада в грамматические формы» («make no constant morphosyntactic contribution to the forms that they underlie»). Встречаются и более «радикальные» определения, прямо отрицающие какую-либо семантическую морфомных формативов, как, например, в работе [Spencer 2012: 88]: «основы, представляющие собой чистые формы и не реализующие каких-либо словоизменительных признаков» («stems that are pure forms and which are not the realization of any feature or property set»). Иными словами, граница между морфемами и морфемами проводится в зависимости от того, насколько множество клеток парадигмы, в которых представлен форматив, признается «семантически гомогенным» или «естественным» [Herce 2020]. В действительности же, дистрибуция того или иного форматива может быть в разной степени семантически мотивированной (англ. *coherent* в терминологии [Smith 2013]) — это так называемый **градиент** между морфемной и морфомной дистрибуциями [Smith 2013; Herce 2020]. Например, дистрибуция форматива *-a* [-AOR] и его алломорфов в подсистеме перфектива анчихской системы (табл. 4) с трудом может быть названа семантически мотивированной: его наличие в формах императива и оптатива множественного числа делает класс форм системы перфектива неестественным, не описываемым при помощи словоизменительных признаков. С другой стороны, распределение анчихского форматива имперфектива *-ir* демонстрирует значительную семантическую мотивацию: он встречается только в формах с имперфективной семантикой: имперфективе, хабицуалисе, имперфективном конвербе и причастии. Хотя даже при таком анализе мы сталкиваемся с трудностями: смысл ‘имперфектив’ сам по себе не моносемичен и представляет собой аспектуальный кластер из дуратива и хабицуалиса [Плунгян 2011: 402], в то время как форма имперфективного конверба в анчихском диалекте каратинского языка выражает не собственно имперфектив, а контемпоралис, т. е. строгую одновременность двух ситуаций, по классификации [Муравьев 2017]. Насколько эти значения формируют «естественный класс», остается неясным.

В работе [Smith 2013: 250–251] такие морфомы, возникающие в ТАМ-системах, выделяются в отдельный класс морфом (ТАМ-морфомы) на том основании, что они являются одним из наиболее семантически мотивированных классов морфом. Подобные парадигматические дистрибуции формативов с остаточной семантической мотивированностью обычно образуются в ходе развития полисемии показателя [Heine et al. 1991: 225–229; Майсак 2002: 20], т. е. при **грамматическом дрейфе** — семантическом развитии показателей уже внутри грамматической системы, а не в ходе превращения из лексической единицы в грамматическую [Урманчиева 2015]. При этом разрастание семантической сети значений показателя часто ведет к тому, что эти значения становится невозможно свести к одному семантическому инварианту (такая ситуация называется гетеросемией в [Lichtenberk 1991: 476–477]) или, по [Herce 2019: 113], происходит «утрата семантического единства множества прежде семантически близких глагольных форм» («loss of semantic unity on a set of formerly semantically related tenses»).

В некотором смысле андийские идиоматические комплексы словоизменительных показателей можно рассматривать как недистантный аналог полиаффиксов, выделяемых

в языках банту [Плунгян 2016: 95–97] (или конфи́ксов в более традиционной терминологии [Аксенова 1997]). Вместе с тем эти системы имеют и существенные отличия, поскольку в языках банту есть как префиксальные, так и суффиксальные позиционные классы, а андийские языки все ТАМ-значения выражают суффиксально. Это обстоятельство, очевидно, влияет на диахроническую судьбу данных систем, поскольку сценарий фузионной компрессии полиаффиксов естественным образом блокирован присутствием корня между префиксальной и суффиксальной частями полиаффикса.

Более отдаленный (но не менее интересный с точки зрения грамматической типологии) аналог андийских иерархических эффектов обнаруживается в глагольных системах ряда языков майя (юкатекских и чоль) [Виноградов 2013а; 2013б]. Эти системы можно считать линейными, поскольку в них неортогонально противопоставлены не только аспектуальные, но также модальные и временные значения [Виноградов 2013б: 124]. Отличительной особенностью этих систем является так называемый двухуровневый механизм выражения грамматических значений: системы некоторых языков требуют одновременного присутствия двух видо-временных показателей в глагольной форме — суффиксального, выражающего более общие и размытые значения (например, перфектив, имперфектив, ирреалис), и препозитивного, выражающего более частные значения (дуратив, перфект, проспектив, авертив и пр.). Как отмечает И. А. Виноградов, между уровнями наблюдается нетривиальное синтагматическое (иерархическое) взаимодействие: «[Г]раммемы одного уровня могут сочетаться не с любыми граммемами другого уровня, а только со строго определенными» [Виноградов 2013б: 59]. Так, в аспектуальной системе языка чоль (рис. 2) граммема перфектива препозитивного уровня сочетается только с граммемой перфектива суффиксального уровня, а суффиксальный имперфектив может принимать в качестве грамем препозитивного уровня имперфектив, дуратив и авертив.

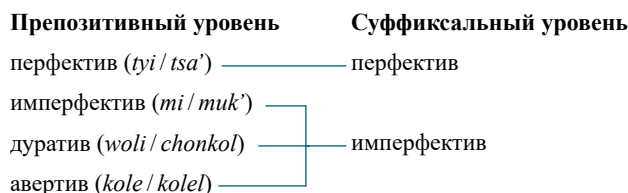


Рис. 2. Аспектуальная система языка чоль (по [Виноградов 2013б: 58])

Для грамем препозитивного уровня в [Виноградов 2013б: 63] вводится особое понятие интерпретирующей категории: «интерпретирующую категорию можно определить как грамматическую категорию, обязательно употребляющуюся совместно с какой-либо другой грамматической категорией, значения которой она интерпретирует (уточняет или модифицирует)». Однако из следующего пассажа видно, что интерпретирующая категория выделяема только на формальных основаниях — в силу того, что показатели интерпретирующей категории формируют хорошо выделяемый позиционный класс. Однако семантические свойства этих показателей не позволяют выделить две ортогональные категории: «Поскольку в каждом конкретном языке граммема интерпретирующей категории сочетается со строго определенной граммемой базовой категории и только с ней, рассмотреть семантику грамем этих двух категорий по отдельности не представляется возможным. Поэтому ⟨...⟩, говоря о семантике отдельной грамемы, мы имеем в виду семантику сочетания интерпретирующей и базовой грамем» [Там же: 125]. Тем самым грамматические формы в такой системе не обладают семантической аддитивностью, поскольку не удастся рассмотреть семантику грамем по отдельности. Интерпретирующая и базовая категории не ортогональны, их комбинаторный потенциал неполон из-за ограничений взаимной селективности между базовой граммемой и интерпретирующей — по сути дела, интерпретирующая категория подчинена базовой.

Отчасти похожие явления можно наблюдать и в андийских системах как с точки зрения семантики показателей, так и с точки зрения их селективности. Как мы уже отмечали выше, нередко показатель с более широкой дистрибуцией или общим значением находится ближе к корню, чем показатель с более узкой дистрибуцией или частным значением (например, уже обсуждавшиеся имперфектив и хабитуалис). Как и в майяских системах, комбинаторный потенциал формативов ограничен: выбор ближайшего к корню форматива определяет выбор формативов следующих позиций. Тем не менее между этими типами систем есть и важное различие. В структуре майяской формы все грамматическое значение с обязательностью распределено между двумя позициями — препозитивной и суффиксальной [Виноградов 2013б: 61], так что единая для всех форм позиционная структура позволяет говорить о двух позиционных классах, которые можно описывать в виде двух категорий, пусть и не ортогональных. В андийских языках нет единой позиционной структуры для всех словоформ, а грамматическое значение может выражаться наборами из разного количества формативов. По этим причинам для андийских полиформативных систем речь о введении интерпретирующей категории все же не идет. Большая формальная «стройность» майяских систем по сравнению с андийскими, при некотором сходстве их семантической организации, на наш взгляд, может объясняться меньшей морфологизованностью показателей препозитивного уровня (в значительной мере это аналитические вспомогательные частицы [Там же: 97]) и дистантным расположением показателей разных уровней (слева и справа от корня), способствующим формированию двух позиционных классов и препятствующим возможной фузии.

Добавление о дивергенции. Из-за того, что большое количество грамматических элементов в полиформативных системах типа анчихской и багвалинской могут выступать и как автономные словоформы, эти автономные основы демонстрируют процесс расхождения, или дивергенции (англ. *divergence*, по П. Хопперу), который в [Hopper 1991: 22] описывается применительно к лексике так: «When a lexical form undergoes grammaticization to a clitic or affix, the original lexical form may remain as an autonomous element and undergo the same changes as ordinary lexical items». Здесь показателен багвалинский пример из [Кибрик и др. 2001: 78, сноска 1], описывающий диахронию деепричастия и причастия будущего времени (см. также соответствующие формы на рис. 1). Футуральные формы деепричастия *et-a:-li-χ* [летать-INF-FUT-CVB] ‘будучи тем, который полетит’ и причастия *et-a:-l-o-b* [летать-INF-FUT-PTCP-N] ‘который полетит’ образуются от основы инфинитива, которая способна выступать как автономная словоформа *et-a:* [летать-INF] ‘летать’. Источником грамматикализации показателей деепричастия и причастия *-li-χ* и *-l-o-b*, по [Кибрик и др. 2001], являются, соответственно, деепричастная и причастная формы глагола **le-* ‘идти’, от которого в современном языке остались только формы причастия *l-o-b* [идти-PTCP-N] ‘идуший’ и гортатива *le* [идти.НОРТ] ‘пойдем’. Этот пример показывает, что возможна и «обратная дивергенция», когда при грамматикализации конструкции лексический источник показателя (глагол ‘идти’) исчезает, а форма-реципиент (инфинитив) сохраняется в виде полноценной словоформы.

5. Заключение

Итак, мы проанализировали несколько глагольных систем андийских языков и показали, что они весьма далеки от того, что принято считать агглютинативным прототипом: противопоставленные значения в системе слабо ортогональны, чем напоминают системы, организованные по линейному принципу. При этом можно выделить как системы с преобладающей формальной аддитивностью формативов (например, анчихская и багвалинская системы), так и с существенной формальной неаддитивностью (таковы тиндинская и чамалинская системы).

Важным свойством андийских полиформативных систем является семантическая неаддитивность: значительному числу формативов в них не удастся приписать таких значений, из которых можно было бы композиционно получить общее грамматическое значение словоформы — такие формативы являются морфомными. Ослабленная ортогональность грамматических категорий в системе ведет к образованию иерархической структуры парадигмы. Иерархическое (древесное) устройство системы проявляется в том, что более близкие к корню показатели накладывают селективные ограничения на более периферийные показатели. Судя по всему, такая организация глагольной системы не уникальна в языках мира — неортогональность ТАМ-значений отмечается типологически [Herce 2023].

По всей видимости, большинство наблюдаемых особенностей андийских систем могут быть объяснены диахронически. Разница между формально-аддитивными и формально-неаддитивными системами, как можно предполагать, заключается в том, что в формально-неаддитивных системах протекают фузионные процессы, более интенсивные, чем процессы морфологизации нового материала. Неизоморфность категориальной и морфологической структур может объясняться асинхронной грамматикализацией разных элементов в системе, которую, в свою очередь, можно связать с неодинаковой стабильностью грамматических значений в системе. Иерархическое устройство парадигмы отражает сосуществование хронологически разных слоев грамматикализованного материала (*layering*, по [Hopper 1991]). Морфомные показатели образуются из семантически однозначных в результате разрастания семантической сети под действием грамматического дрейфа, ведущего к полисемии и гетеросемии таких показателей.

Исходя из такой перспективы, приходится признать недостаточной гипотезу [Blevins et al. 2019: 275] о возникновении агглютинации: «An agglutinating pattern could arise in a language whose morphotactic organization preserved the structure determined by successive waves of morphologization». Данные андийских полиформативных систем вполне соответствуют описанной ситуации: в ходе нескольких последовательных волн морфологизации были сформированы морфотактически неэлементарные средства выражения грамматических значений. Однако разрастание семантических сетей, ведущее к гетеросемии и морфомности показателей в ходе грамматического дрейфа ТАМ-форм в системе, не позволяет сохранять семантическую аддитивность между старыми и новыми слоями морфологизованного материала.

Таким образом, типологические особенности андийских систем в основном могут быть объяснены своеобразием представленных в них грамматических значений и связанных с этими значениями диахронических процессов. По всей видимости, интуиция Дж. Байби о влиянии семантики на формирование тех или иных языковых структур может быть продуктивно применена и к области парадигматики: «[T]he typological differences among languages in this view derive from the differences in substance that give rise to differences in structure» [Bybee 1995: 244]. Такой взгляд, во-первых, позволяет исследовать способность разных классов грамматических значений формировать разные парадигматические структуры и, во-вторых, позволяет установить характер диахронического взаимодействия между категориальной и морфологической структурами. При этом нельзя сказать, что только семантический фактор играет роль в формировании парадигматических структур. Ясно, что и структурные параметры, такие как склонность / несклонность языковой системы к фузии или предпочтение словоизменительной префиксации / суффиксации, также могут значительно влиять на свойства грамматической системы.

СПИСОК СОКРАЩЕНИЙ

1 — первое лицо
АСТ — активный залог
АОР — аорист

ATTR — атрибутивизатор
CAUS — каузатив
COP — копула

CVB — конверб
 DAT — датив
 DEB — дебитив
 ERG — эргатив
 FUT — будущее время
 HAB — хабитуалис
 HORT — гортатив
 HPL — личный множественный
 согласовательный класс
 ICVB — имперфективный конверб
 IMP — императив
 INF — инфинитив
 INTR — непереходность
 IPFV — имперфектив
 IRR — ирреалис

MED — средний демонстратив
 N — средний согласовательный класс
 NEG — отрицание
 OBL — косвенная основа
 OPT — опатив
 PASS — пассивный залог
 PF — перфект
 PL — множественное число
 PLACE — место
 POT — потенциалис
 PROG — прогрессив
 PRON — прохибитив
 PTCP — причастие
 REFL — рефлексив
 SG — единственное число

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Аксенова 1997 — Аксенова И. С. *Категории вида, времени и наклонения в языках банту*. М.: Наука, 1997. [Aksenova I. S. *Kategorii vida, vremeni i nakloneniya v yazykakh bantu* [The categories of aspect, tense, and mood in Bantu languages]. Moscow: Nauka, 1997.]
- Алексеев 1988 — Алексеев М. Е. *Сравнительно-историческая морфология аваро-андийских языков*. М.: Наука, 1988. [Alekseev M. E. *Sravnitel'no-istoricheskaya morfologiya avaro-andiiskikh yazykov* [Comparative and historical morphology of Avar-Andic languages]. Moscow: Nauka, 1988.]
- Алексеев, Атаев 1997 — Алексеев М. Е., Атаев Б. М. *Аварский язык*. М.: Academia, 1997. [Alekseev M. E., Ataev B. M. *Avarskii yazyk* [The Avar language]. Moscow: Academia, 1997.]
- Бокарев 1949 — Бокарев А. А. *Очерк грамматики чамалинского языка*. М.; Л.: Изд-во АН СССР, 1949. [Bokarev A. A. *Ocherk grammatiki chamalinskogo yazyka* [A sketch of Chamalal grammar]. Moscow; Leningrad: Press of the Academy of Sciences of the USSR, 1949.]
- Бокарев, Климов 1967 — Бокарев Е. А., Климов Г. А. *Иберийско-кавказские языки. Введение. Языки народов СССР. Иберийско-кавказские языки*. Бокарев Е. А., Ломтатидзе К. В. (отв. ред.). М.: Наука, 1967, 7–14. [Bokarev E. A., Klimov G. A. *Iberian-Caucasian languages. Introduction. Yazyki narodov SSSR. Iberiisko-kavkazskie yazyki*. Bokarev E. A., Lomtadidze K. V. (eds.). Moscow: Nauka, 1967, 7–14.]
- Виноградов 2013а — Виноградов И. А. «Интерпретирующие» грамматические категории: к определению понятия. *Вопросы языкознания*, 2013, 5: 28–45. [Vinogradov I. A. “Interpretive” grammatical categories: Defining the concept. *Voprosy Jazykoznanija*, 2013, 5: 28–45.]
- Виноградов 2013б — Виноградов И. А. *Семантические грамматические категории глагола в языках майя: сравнительный и типологический анализ*. Дис. ... канд. филол. наук. М.: МГУ имени М. В. Ломоносова, 2013. [Vinogradov I. A. *Semanticheskie grammaticheskie kategorii glagola v yazykakh maia: sravnitel'nyi i tipologicheskii analiz* [Semantic grammatical categories of the verb in Mayan languages: A comparative and typological analysis]. Candidate diss. Moscow: Lomonosov Moscow State Univ., 2013.]
- Володин 1976 — Володин А. П. *Ительменский язык*. Л.: Наука, 1976. [Volodin A. P. *Itel'menskii yazyk* [The Itelmen language]. Leningrad: Nauka, 1976.]
- Гаджимагомедов 2019 — Гаджимагомедов М. З. *Анчихский язык. Словарь*. Махачкала, 2019. Рук. [Gadzhimagomedov M. Z. *Anchikhskii yazyk. Slovar'* [The Anchiq language. Dictionary]. Makhachkala, 2019. Ms.]
- Гусев 2006 — Гусев В. Ю. О сохранении архаичных форм в неассертивных контекстах: материал самодийских языков. *Проблемы типологии и общей лингвистики: международная конференция, посвященная 100-летию со дня рождения проф. А. А. Холодовича. Материалы. Санкт-Петербург, 4–6 сентября 2006 г.* Храковский В. С., Дмитренко С. Ю., Заика Н. М. (ред.). СПб.: Нестор-История, 2006, 41–45. [Gusev V. Yu. On the retention of archaic forms in non-assertive contexts: Data of Samoyedic languages. *Problemy tipologii i obshchei lingvistiki. International conf. on the 100th anniversary of the birthday of Prof. A. A. Xolodovich. Proceedings. St. Petersburg, September 4–6, 2006*. Xrakovskij V. S., Dmitrenko S. Yu., Zaika N. M. (eds.). St. Petersburg: Nestor-Istoriya, 2006, 41–45.]

- Зализняк 1967/2002 — Зализняк А. А. Русское именное словоизменение. «Русское именное словоизменение» с приложением избранных работ по современному русскому языку и общему языкознанию. Зализняк А. А. М.: Языки славянской культуры, 2002, 1–370. 1-е изд. 1967 г. [Zaliznyak A. A. Russian nominal declension. «Russkoe imennoe slovoizmenenie» s prilozeniem izbrannykh rabot po sovremennomu russkomu yazyku i obshchemu yazykoznaniiu. Zaliznyak A. A. Moscow: Yazyki slavyanskikh kul'tur, 2002, 1–370. 1st publ. 1967.]
- Кибрик и др. 2001 — Кибрик А. Е. (ред.-сост.), Казенин К. И., Лютикова Е. А., Татевосов С. Г. (ред.). *Багвалинский язык. Грамматика. Тексты. Словари.* М.: Наследие, 2001. [Kibrik A. E. (ed.-comp.), Kazenin K. I., Lyutikova E. A., Tatevosov S. G. (eds.). *Bagvalinskii yazyk. Grammatika. Teksty. Slovarei* [The Bagvalal language. Grammar. Texts. Dictionaries]. Moscow: Nasledie, 2001.]
- Кибрик, Кодзасов 1988 — Кибрик А. Е., Кодзасов С. В. *Сопоставительное изучение дагестанских языков. Глагол.* М.: Изд-во Московского ун-та, 1988. [Kibrik A. E., Kodzasov S. V. (eds.). *Sopostavitel'noe izuchenie dagestanskikh yazykov. Glagol* [Comparative dictionary of Daghestanian languages. Verb]. Moscow: Moscow Univ. Press, 1988.]
- Коношенко 2017 — Коношенко М. Б. Кпелле язык. *Языки мира: Языки манде.* Выдрин В. Ф., Мазурова Ю. В., Кибрик А. Е., Маркус Е. Б. (ред.). СПб.: Нестор-История, 2017, 284–342. [Konoshenko M. B. The Kpelle language. *Yazyki mira: Yazyki mande.* Vydrin V. F., Mazurova Yu. V., Kibrik A. E., Markus E. B. (eds.). St. Petersburg: Nestor-Istoriya, 2017, 284–342.]
- Магомедбекова 1967 — Магомедбекова З. М. *Ахвахский язык: грамматический анализ, тексты, словарь.* Тбилиси: Мецниереба, 1967. [Magodbekova Z. M. *Akhvakhskii yazyk: grammaticheskii analiz, teksty, slovar'* [The Akhvakh language: Grammar analysis, texts, dictionary]. Tbilisi: Mecniereba, 1967.]
- Магомедбекова 1971 — Магомедбекова З. М. *Каратинский язык: грамматический анализ, тексты, словарь.* Тбилиси: Мецниереба, 1971. [Magodbekova Z. M. *Karatinskii yazyk: grammaticheskii analiz, teksty, slovar'* [The Karata language: Grammar analysis, texts, dictionary]. Tbilisi: Mecniereba, 1971.]
- Магомедова 2012 — Магомедова П. Т. *Тиндинский язык.* Махачкала: ИЯЛИ ДНЦ РАН, 2012. [Magomedova P. T. *Tindinskii yazyk* [The Tindi language]. Makhachkala: Tsadasa Institute of Language, Literature and Art, 2003.]
- Майсак 2002 — Майсак Т. А. *Типология грамматикализации конструкций с глаголами движения и глаголами позиции.* Дис. ... канд. филол. наук. М.: МГУ имени М. В. Ломоносова, 2002. [Maisak T. A. *Tipologiya grammatikalizatsii konstruksii s glagolami dvizhenia i glagolami pozitsii* [Typology of grammaticalization paths with verbs of motion and position]. Candidate diss. Moscow: Lomonosov Moscow State Univ., 2002.]
- Майсак 2008 — Майсак Т. А. Семантика и происхождение глагольных форм настоящего и будущего времени в удинском языке. *Удинский сборник: грамматика, лексика, история языка.* Алексеев М. Е., Майсак Т. А. (отв. ред.). М.: Academia, 2008, 162–222. [Maisak T. A. Semantics and origin of present and future verb forms in Udi. *Udinskii sbornik: grammatika, leksika, istoriya yazyka.* Alekseev M. E., Maisak T. A. (eds.). Moscow: Academia, 2008, 162–222.]
- Майсак 2014 — Майсак Т. А. Глагольные системы лезгинских языков: структура, семантика и пути развития перфективных форм. *Вестник Института языка, литературы и искусства им. Г. Цадасы*, 2014, 6: 26–42. [Maisak T. A. Verbal systems of Lezgian languages: Structure, semantics, and ways of the evolution of perfective forms. *Bulletin of the G. Tsadasa Institute of Language, Literature and Art*, 2014, 6: 26–42.]
- Майсак 2016 — Майсак Т. А. *Глагольная система рикванинского говора андийского языка: формообразование.* М., 2016. Рук. [Maisak T. A. *Glagol'naya sistema rikvaninskogo govora andiiskogo yazyka: formoobrazovanie* [Verbal system of the Rikvani variety of Andi: Forms]. Moscow, 2016. Ms.]
- Мельчук 1997 — Мельчук И. А. *Курс общей морфологии.* Т. I: Введение. Слово. Перцов Н. В. (ред.). Перцова Н. Н., Саввина Е. Н. (пер.). М.; Вена: Языки русской культуры, 1997. [Mel'chuk I. A. *Kurs obshchei morfologii* [Course in general linguistics]. Vol. 1: *Vvedenie. Slovo* [Introduction. Word]. Pertsov N. V. (ed.), Pertsova N. N., Savvina E. N. (trans.). Moscow; Vienna: Yazyki slavyanskoi kul'tury, 1997.]
- Муравьев 2017 — Муравьев Н. А. *Таксис и таксисные значения в языках мира: теория и типология.* Дис. ... канд. филол. наук. М.: МГУ имени М. В. Ломоносова, 2017. [Muraviev N. A. *Taksis i taksisnye znacheniya v yazykakh mira: teoriya i tipologiya* [Taxis and taxis meanings in the world's languages: Theory and typology]. Candidate diss. Moscow: Lomonosov Moscow State Univ., 2017.]
- Назаров 1974 — Назаров В. П. *Разыскания в области исторической морфологии восточнокавказских языков.* Махачкала: Дагучпедгиз, 1974. [Nazarov V. P. *Razyskaniya v oblasti istoricheskoi*

- morfologii vostochnokavkazskikh yazykov* [Studies on historical morphology of East Caucasian languages]. Makhachkala: Daguchpedgiz, 1974.]
- Плунгян 2003 — Плунгян В. А. Африканские глагольные системы: заметки к типологии. *Основы африканского языкознания: Глагол*. Виноградов В. А., Топорова И. Н. (ред.). М.: Восточная литература, 2003, 5–40. [Plungian V. A. African verb systems: Remarks on a typology. *Osnovy afrikanskogo yazykoznanija: Glagol*. Vinogradov V. A., Toporova I. N. (eds.). Moscow: Vostochnaya literatura, 2003, 5–40.]
- Плунгян 2011 — Плунгян В. А. *Введение в грамматическую семантику: грамматические значения и грамматические системы языков мира*. М.: РГГУ, 2011. [Plungian V. A. *Vvedenie v grammaticheskuyu semantiku: grammaticheskie znacheniya i grammaticheskie sistemy yazykov mira* [Introduction to grammatical semantics: Grammatical meanings and grammatical systems of the world's languages]. Moscow: Russian State Univ. for the Humanities, 2011.]
- Плунгян 2016 — Плунгян В. А. *Общая морфология: Введение в проблематику*. М.: ЛЕНАНД, 2016. [Plungian V. A. *Obshchaya morfologiya: Vvedenie v problematiku* [General linguistics: Introduction to the problematics]. Moscow: LENAND, 2016.]
- Плунгян, Урманчиева 2004 — Плунгян В. А., Урманчиева А. Ю. «Линейные» глагольные системы: к постановке проблемы. *Грамматические категории и единицы: Сб. научных статей к 75-летию А. Б. Копелиовича*. Пименова М. В. (ред.). Владимир: ВГПУ, 2004, 121–138. [Plungian V. A., Urmanchieva A. Yu. “Linear” verb systems: Towards the problem statement. *Grammaticheskie kategorii i edinitsey. Coll. of scientific papers on the 75th anniversary of A. B. Kopielovich*. Pimenova M. V. (ed.). Vladimir: Vladimir State Pedagogical Univ., 2004, 121–138.]
- Ревзин, Юлдашева 1969 — Ревзин И. И., Юлдашева Г. Д. Грамматика порядков и ее использования. *Вопросы языкознания*, 1969, 1: 42–56. [Revzin I. I., Yuldasheva G. D. The grammar of orders and its use. *Voprosy Jazykoznanija*, 1969, 1: 42–56.]
- Сумбатова 2011 — Сумбатова Н. Р. Показатели предикативных категорий в даргинском языке. *Московский лингвистический журнал*, 2011, 11: 151–173. [Sumbatova N. R. Markers of predicative categories in Dargwa. *Moscow Journal of Linguistics*, 2011, 11: 151–173.]
- Урманчиева 2015 — Урманчиева А. Ю. Как грамматическая система управляет семантической эволюцией показателей (на примере эвиденциальной системы тазовского селькупского). *Вопросы языкознания*, 2015, 6: 54–77. [Urmanchieva A. Yu. How much impact can grammatical system have on the semantic evolution of grams (a case study on the Tas Selkup system of evidential markers). *Voprosy Jazykoznanija*, 2015, 6: 54–77.]
- Услар 1889 — Услар П. К. *Аварский язык*. Тифлиς, 1889. [Uslar P. K. *Avarskii yazyk* [The Avar language]. Tiflis, 1889.]
- Ферхеев 2019 — Ферхеев Я. Х. Эвиденциальность как часть видо-временной системы глагола в нахско-дагестанских языках. Дис. ... канд. филол. наук. М.: НИУ ВШЭ, 2019. [Verhees J. H. *Evidential 'nost' kak chast' vido-vremennoi sistemy glagola v nakhsko-dagestanskikh yazykakh* [Evidentiality as part of tense-aspect in East Caucasian languages]. Moscow: HSE Univ., 2019.]
- Филатов 2020 — Филатов К. В. Система индикатива анчихского диалекта каратинского языка. ВКР бакалавра. М.: НИУ ВШЭ, 2020. [Filatov K. V. *Sistema indikativa anchikhskogo dialekta karatinskogo yazyka* [Indicative system of Anchiq Karata]. BA thesis. Moscow: HSE Univ., 2020.]
- Храковский, Мальчуков (ред.) 2020 — Храковский В. С., Мальчуков А. Л. (ред.). *Иерархия и взаимодействие грамматических категорий глагола*. СПб.: ИЛИ РАН, 2020. [Xrakovskij V. S., Malchukov A. L. (eds.). *Ierarkhiya i vzaimodeistvie grammaticheskikh kategorii glagola* [Hierarchy and interaction of verbal categories]. St. Petersburg: Institute for Linguistic Studies of the Russian Academy of Sciences, 2020.]
- Alexeyev, Verhees (to appear) — Alexeyev M., Verhees S. Botlikh. *Handbook of Caucasian languages*. Vol. 2: *Nakh-Daghestanian*. Koryakov Yu. B., Maisak T. A. (eds.). Berlin; New York: De Gruyter Mouton. To appear.
- Arkadiyev 2022 — Arkadiyev P. Morphology of the Caucasian languages: A typological overview. *Jezikoslovní zapiski*, 2022, 28(1): 7–38.
- Aronoff 1994 — Aronoff M. *Morphology by itself: Stems and inflectional classes*. Cambridge (MA): MIT Press, 1994.
- Authier (to appear) — Authier G. Tindi. *Handbook of Caucasian languages*. Vol. 2: *Nakh-Daghestanian*. Koryakov Yu. B., Maisak T. A. (eds.). Berlin; New York: De Gruyter Mouton. To appear.
- Baerman 2019 — Baerman M. Feature duality. *Morphological perspectives: Papers in honour of Greenville G. Corbett*. Baerman M., Bond O., Hippisley A. (eds.). Edinburgh: Edinburgh Univ. Press, 2019, 124–137.

- Baerman, Corbett 2012 — Baerman M., Corbett G. G. Stem alternations and multiple exponence. *Word Structure*, 2012, 5(1): 52–68.
- Blevins 2003 — Blevins J. P. Stems and paradigms. *Language*, 2003, 79: 737–767.
- Blevins 2016 — Blevins J. P. *Word and paradigm morphology*. Oxford: Oxford Univ. Press, 2016.
- Blevins et al. 2019 — Blevins J. P., Ackerman F., Malouf R. Word and paradigm morphology. *Oxford handbook of morphological theory*. Audring J., Masini F. (eds.). Oxford: Oxford Univ. Press, 2019, 265–284.
- Bybee 1995 — Bybee J. Diachronic and typological properties of morphology and their implications for representation. *Morphological aspects of language processing*. Feldman L. (ed.). Hillsdale (NJ): Lawrence Erlbaum Associates, 1995, 225–246.
- Bybee et al. 1994 — Bybee J. L., Perkins R. D., Pagliuca W. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: Univ. of Chicago Press, 1994.
- Corbett 2012 — Corbett G. G. *Features*. New York: Cambridge Univ. Press, 2012.
- Creissels 2018 — Creissels D. Perfective tenses and epistemic modality in Northern Akhvakh. *The semantics of verbal categories in Nakh-Daghestanian languages*. Forker D., Maisak T. A. (eds.). Leiden; Boston: Brill, 2018, 166–187.
- Fedden, Corbett 2017 — Fedden S., Corbett G. G. Gender and classifiers in concurrent systems: Refining the typology of nominal classification. *Glossa: a journal of general linguistics*, 2017, 2(1): 1–47.
- Forker 2018 — Forker D. Introduction. *The semantics of verbal categories in Nakh-Daghestanian languages*. Forker D., Maisak T. A. (eds.). Leiden; Boston: Brill, 2018, 1–25.
- Forker 2020 — Forker D. Avar. *The Oxford handbook of languages of the Caucasus*. Polinsky M. (ed.). Oxford; New York: Oxford Univ. Press, 2020, 243–279.
- Ganenkova, Maisak 2020 — Ganenkova D., Maisak T. Nakh-Dagestanian languages. *The Oxford handbook of languages of the Caucasus*. Polinsky M. (ed.). Oxford; New York: Oxford Univ. Press, 2020, 87–146.
- Harkness 1898 — Harkness A. *A complete Latin grammar*. New York; Cincinnati; Chicago: American Book Company, 1898.
- Haspelmath 2009 — Haspelmath M. An empirical test of the Agglutination Hypothesis. *Universals of language today*. Scalise S., Magni E., Bisetto A. (eds.). Dordrecht: Springer Netherlands, 2009, 13–29.
- Heine et al. 1991 — Heine B., Claudi U., Hünnemeyer F. *Grammaticalization: A conceptual framework*. Chicago; London: Univ. of Chicago Press, 1991.
- Herce 2019 — Herce B. Morpheme interactions. *Morphology*, 2019, 29(1): 109–132.
- Herce 2020 — Herce B. On morphemes and morphemes: Exploring the distinction. *Word Structure*, 2020, 13(1): 45–68.
- Herce 2023 — Herce B. *The typological diversity of morphemes: A cross-linguistic study of unnatural morphology*. Oxford: Oxford Univ. Press, 2023.
- Hockett 1954 — Hockett C. F. Two models of grammatical description. *Word*, 1954, 10(2–3): 210–234.
- Hopper 1991 — Hopper P. J. On some principles of grammaticalization. *Approaches to grammaticalization*. Traugott E. C., Heine B. (eds.). Amsterdam; Philadelphia: John Benjamins, 1991, 17–36.
- Kaye, Maisak 2023 — Kaye S., Maisak T. The two verbal stems in Andi. *Talk at Imperfective Modalities in the Caucasus and Beyond. Collège de France, Paris, September 12, 2023*.
- Kibrik et al. (eds.) 1996 — Kibrik A., Tatevosov S., Eulenberg A. (eds.). *Godoberi*. Munich; Newcastle: Lincom Europa, 1996.
- Lichtenberk 1991 — Lichtenberk F. Semantic change and heterosemy in grammaticalization. *Language*, 1991, 67(3): 475–509.
- Maiden 2018 — Maiden M. *The Romance verb: Morphomic structure and diachrony*. Oxford: Oxford Univ. Press, 2018.
- Matthews 1972 — Matthews P. H. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge Univ. Press, 1972.
- Nichols 2003 — Nichols J. Diversity and Stability in Language. *The Handbook of historical linguistics*. Joseph B. D., Janda R. D. (eds.). Malden (MA): Blackwell Publ., 2003.
- Plungian 2001 — Plungian V. A. Agglutination and flection. *Language typology and language universals: An international handbook*. Vol. 1. Haspelmath M., König E., Oesterreicher W., Raible W. (eds.). Berlin; New York: Walter de Gruyter, 2001, 669–678.
- Round, Corbett 2017 — Round E. R., Corbett G. G. The theory of feature systems: One feature versus two for Kayardild tense-aspect-mood. *Morphology*, 2017, 27(1): 21–75.
- Smith 2013 — Smith J. C. The morpheme as a gradient phenomenon: Evidence from Romance. *The boundaries of pure morphology: Diachronic and synchronic perspectives*. Cruschina S., Maiden M., Smith J. Ch. (eds.). Oxford: Oxford Univ. Press, 2013, 247–261.

- Spencer 2012 — Spencer A. Identifying stems. *Word Structure*, 2012, 5(1): 88–108.
- Stump 1991 — Stump G. T. On the theoretical status of position class restrictions on inflectional affixes. *Yearbook of Morphology 1991*. Booij G., van Marle J. (eds.). Amsterdam: Kluwer Academic Publ., 1991, 211–241.
- Stump 2001 — Stump G. T. *Inflectional morphology. A theory of paradigm structure*. Cambridge: Cambridge Univ. Press, 2001.
- van den Berg 2005 — van den Berg H. The East Caucasian language family. *Lingua*, 2005, 115(1–2): 147–190.
- Welmers 1973 — Welmers Wm. E. *African language structures*. Berkeley: Univ. of California Press, 1973.

Получено / received 11.03.2024

Принято / accepted 24.09.2024

Карельский языковой ландшафт в диалектометрической парадигме

© 2024

Ирина Петровна Новак

Институт языка, литературы и истории Карельского научного центра РАН,
Петрозаводск, Россия; novak@krc.karelia.ru

Аннотация: Изучение карельской диалектной речи ведется силами отечественных и зарубежных фенистов на протяжении полутора столетий, но такие проблемы карельской диалектологии, как территориальная неясность распространения диалектных единиц, выбор основного принципа диалектного членения, определение языкового статуса отдельных идиомов и др., до настоящего момента оставались нерешенными. В статье обобщены результаты исследования карельской диалектной речи, проведенного на основе архивных рабочих материалов 1930–1970-х гг. «Диалектологического атласа карельского языка» (1997) с применением диалектометрического метода кластерного анализа: перечислены основные сопоставительные явления фонологического, морфонологического, морфологического и лексического языковых уровней, описаны ареалы и дана типология их членов, представлена актуальная трехуровневая диалектная классификация, а также периодизация истории карельского языка. На разработанной диалектной карте намечено три наречия: собственно карельское, ливвиковское и людиковское. Фундаментом для их формирования послужили изоглоссы морфологической и морфонологической систем, восходящие к раннему периоду развития языка. В основе собственно карельского наречия лежат прибалтийско-финские праязыковые и общекарельские инновационные черты, ливвиковского — прибалтийско-финские архаизмы, системно представленный древневепсский субстрат, древнекарельские инновации и собственные диалектные особенности, людиковского — общекарельские инновационные черты и ярко выраженный древневепсский субстрат. В составе собственно карельского наречия было вычленено три диалекта: северный, южный (говоры Центральной России) и характеризующийся промежуточным между наречиями положением переходный. В составе людиковского — два: исконно людиковский и ощутивший сильное позднее влияние со стороны контактных ливвиковских говоров пряжинский (людиковские говоры Пряжинского района Карелии). Говоры ливвиковского наречия продемонстрировали относительное единство. Сложение диалектов основано на поздних древнекарельских диалектных различиях и результатах интенсивных контактов в пограничной между наречиями зоне, тогда как выделение двадцати групп говоров — на поздних фонетических и лексических инновациях и влиянии соседних языков.

Ключевые слова: диалектология, карельский язык, лингвистическая география, лингвистический атлас, прибалтийско-финские языки

Благодарности: Исследование выполнено в рамках государственного задания КарНЦ РАН.

Для цитирования: Новак И. П. Карельский языковой ландшафт в диалектометрической парадигме. *Вопросы языкознания*, 2024, 6: 58–84.

DOI: 10.31857/0373-658X.2024.6.58-84

The Karelian language landscape in the dialectometric paradigm

Irina P. Novak

Institute of Linguistics, Literature and History, Karelian Research Centre
of the Russian Academy of Sciences, Petrozavodsk, Russia; novak@krc.karelia.ru

Abstract: Karelian dialectal speech has been studied by linguists of Finnic languages in Russia and elsewhere for a century and a half, but such issues in Karelian dialectology as the fuzzy distribution of dialectal

units over the territory, selection of the fundamental principle for dialectal division, determination of the language status of specific varieties, etc., have so far remained unresolved. This article summarizes the results of a study of Karelian dialectal speech based on archival materials from the 1930s–1970s in the “Dialectological Atlas of Karelian” (1997) performed using the dialectometric method of cluster analysis: the main contrastive phenomena of phonological, morphonological, morphological and lexical linguistic levels are listed, the distribution ranges are described and a typology of their members is provided, the current three-level dialect classification and the periodization of the history of the Karelian language are presented. Three supra-dialects are outlined on the dialect map: Karelian Proper, Livvi, and Ludic. They are distinguished on the basis of isoglosses of the morphological and morphonological systems inherited from the early period in the development of the language. The Karelian Proper supra-dialect rests upon Balto-Finnic Protolanguage and pan-Karelian novel traits; Livvi shows Balto-Finnic archaisms, features coming from the Old Veps substrate, which has affected different language levels, Old Karelian innovations, and its own dialectal features; Ludic incorporates Old Karelian innovations and a pronounced Old Veps substrate. The Karelian Proper supra-dialect was subdivided into three dialects: northern, southern (variants spoken in Central Russia), and a transitional dialect occupying an intermediate position between supra-dialects. Ludic is differentiated into two dialects: the original Ludic and the Pryazha dialect (Ludic variants spoken in the Pryazha District of Karelia), which was heavily influenced in later periods by neighboring Livvi sub-dialects. Variants of the Livvi supra-dialect proved to be relatively uniform. The making of the dialects was shaped by late Old Karelian dialectal differences and intensive contacts in the border areas of the supra-dialects, whereas the twenty sub-dialectal groups became differentiated through relatively recent phonetic and lexical innovations and contact influence of the neighboring languages.

Keywords: dialectology, Finnic, Karelian, linguistic atlas, linguistic geography

Acknowledgements: The study was carried out under the state order of the Karelian Research Centre of the Russian Academy of Sciences.

For citation: Novak I. P. The Karelian language landscape in the dialectometric paradigm. *Voprosy Jazykoznanija*, 2024, 6: 58–84.

DOI: 10.31857/0373-658X.2024.6.58-84

Введение

Фиксация и изучение карельской диалектной речи ведется российскими, финляндскими и эстонскими феннистами непрерывно на протяжении полутора столетий. Результаты этой работы обобщены в виде «Диалектологического атласа карельского языка» [ДАКЯ], диалектных словарей, дескриптивных описаний отдельных групп говоров. Однако последние затронули далеко не все территориальные варианты карельского языка. Междиалектные исследования проведены лишь по отдельным явлениям и категориям карельской диалектной речи. Имеющиеся лакуны до настоящего момента не позволяли подойти вплотную к решению актуальных проблем карельской диалектологии, касающихся состава и истории сложения диалектной системы карельского языка, границ и статуса образующих ее элементов и требующих переработки имеющейся диалектной карты.

Согласно традиционной в российском финно-угроведении классификации, карельский язык подразделяется на три наречия: собственно карельское, ливвиковское и людиковское, обнаруживающие существенные различия на всех языковых уровнях [Бубрих 1950: 86; Zaikov 2017; Новак 2019]. В работах финляндских феннистов людиковские говоры относятся к самостоятельному прибалтийско-финскому языку [Virtaranta 1972: 26; Pahomov 2017: 285]. В последнее десятилетие проблема определения языкового статуса распространилась также на язык карелов-ливвиков и тверских карелов [Сюрюн и др. 2021: 13–22, 42–43].

Современный карельский языковой ландшафт — результат длительного процесса, включившего в себя внутренние языковые изменения, продолжавшиеся на протяжении нескольких столетий влияния со стороны близкородственных вепсского и финского языков и неродственного русского языка, а также вызванные интенсивными разнонаправленными миграциями карелов межнаречные и междиалектные контакты.

Сложение народности корела в северо-западном Приладожье относится к началу II тысячелетия. Войны XVI–XVIII вв. повлекли за собой массовые переселения карелов с исторической родины на территории Северной и Средней Карелии, а также в Центральную Россию, где оформились диалекты собственно карельского наречия, а также на Олонецкий перешеек, населенный носителями древнеливвиковских и древнелюдиковских говоров [История Карелии 2001: 103–134].

В основу разработанной в 1980–1990 гг. системы диалектного членения карельских наречий (рис. 1) положен предложенный Э. Лескинен административный принцип [KKN II: 5–6], заключающийся в использовании волостных границ Карелии начала XX в. в качестве диалектных. Далеко не между всеми диалектами из этой классификации удается обнаружить какие-либо существенные отличия. Кроме того, довольно часто наблюдается несовпадение границ отдельных из них с изоглоссами диалектно-дифференцирующих явлений [Новак 2022].

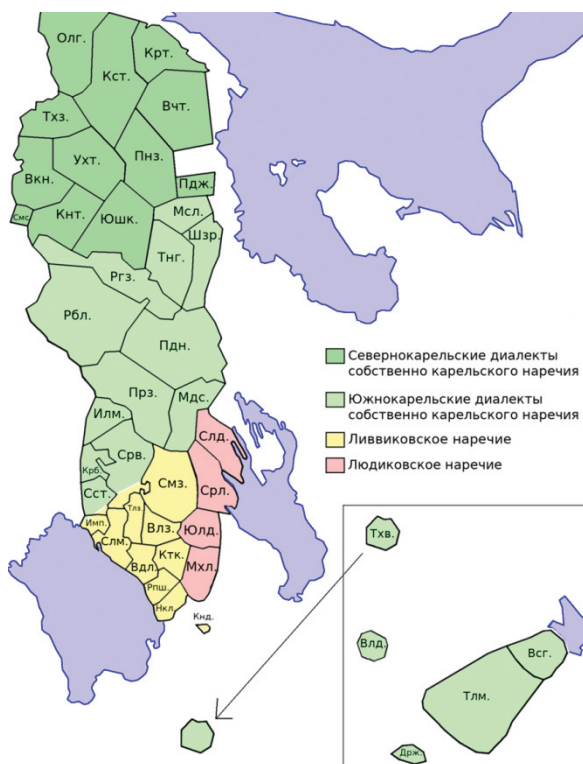


Рис. 1. Традиционная диалектная классификация карельского языка¹

¹ Диалекты: вdl. — видлицкий, вkn. — вокнаволоцкий, влд. — валдайский, влз. — ведлозерский, всг. — весьегонский, вчт. — вычетабольский, држ. — держанский, илм. — иломантсинский, имп. — импилахтинский, кнд. — кондушский, кнт. — контоковский, крб. — корбисельгский, крт. — керетский, кст. — кестеньгский, ктк. — коткозерский, мдс. — мяндусельгский, мсл. — маслозерский, мхл. — михайловский, нкл. — некульский, олг. — олангский, пдж. — подужемский, пдн. — паданский, пнз. — панозерский, прз. — поросозерский, рбл. — ребольский, ргз. — ругозерский, рпш. — рыпушкальский, слд. — севернелюдиковский, слм. — салминский, смз. — сямозерский, смс. — суомуссалмский, срв. — суоярвский, срл. — среднелюдиковский, сст. — суйстамский, тлз. — тулмозерский, тлм. — толмачевский, тнг. — тунгудский, тхв. — тихвинский, тхз. — тихтозерский, ухт. — ухтинский, шзр. — шуезерский, юлд. — южнолудиковский, юшк. — юшкозерский.

Ситуацию существенно усугубляет процесс стремительной утраты карельским языком диалектного разнообразия, что связано как с сокращением числа носителей, так и с угасанием целых групп говоров. Согласно итогам Всероссийской переписи населения 2020 г. число носителей карельского языка составило менее 14 тыс. человек [Итоги]. «Уснувшими» к настоящему моменту являются валдайские, кондушские, зубцовские, кандалакшские карельские говоры. Серьезные опасения вызывает и ситуация в остальных регионах компактного проживания карелов.

Настоящее исследование предпринято с целью комплексного анализа фонологической, морфонологической, морфологической и лексической систем локальных вариантов карельского языка методами диалектометрии, дополненного лингвогеографической и сравнительно-исторической интерпретацией результатов. Итогом этого анализа является построение многоуровневой диалектной классификации.

В качестве теоретической основы выступили работы по карельской диалектологии [Turunen 1946; 1950; Беляков 1958; Itkonen 1971; Virtaranta 1972; 1985; ДАКЯ; Leskinen 1998; Zaikov 2017] и др., труды Московской школы лингвистической географии [Вопросы теории; Пшеничнова 2008; Бромлей 2010] и др., материалы по методике диалектометрии американских и финляндских лингвистов [Wiik 2004a; 2004b; Wieling, Nerbonne 2015; Honkola et al. 2019] и др.

Основными источниками данных для выявления членов междиалектных соответствий и определения ареалов их распространения явились хранящиеся в Научном архиве КарНЦ РАН рабочие материалы 1930–1970-х гг. «Диалектологического атласа карельского языка» (1997) [Программы]. Вопросы для атласа были заполнены более чем в 200 населенных пунктах Карелии и Тверской области. В каждом из них представлены ответы на около 2 тыс. вопросов по фонетике, фонологии, морфологии и лексике. В атлас была отобрана лишь десятая часть из них. Дополнительный сбор данных был произведен по отсутствующим в атласе тихвинскому, валдайскому и кондушскому говорам.

В целях реконструкции истории становления карельской диалектной системы были привлечены представленные в диалектном подкорпусе Открытого корпуса вепсского и карельского языков тексты образцов карельской речи, собранные начиная с конца XIX в. [ВепКар], а также тексты восьми памятников карельской письменности [Мещерский 1961; Муллонен, Панченко 2013; Мызников 2010; Перевод Евангелия от Марка; Перевод Евангелия от Матфея; Перевод некоторых молитв; Савельева (ред.) 2023; Сравнительные словари]. Кроме того, были использованы полевые данные, полученные автором в период с 2014 по 2023 г. в ходе экспедиционных работ в 13 традиционных ареалах проживания карелов Республики Карелия, Тверской, Ленинградской и Мурманской областей.

Исследование основано на применении диалектометрического метода кластерного анализа, заключающегося в разбиении множества говоров (населенных пунктов, в которых были записаны диалектные данные) по заданному набору диалектных различий на группы (кластеры), компоненты которых обнаруживают между собой минимальное число отличий. Для применения иерархических (полной связи, одиночной связи, центроидный, Соллина) и итеративного (к-средних) методов кластеризации² [Новак, Крижановская 2022] была разработана и наполнена рабочими материалами атласа база данных [Murreh]. Соответствующий модуль базы представляет результаты кластеризации в виде кластерных карт, содержащих наглядную информацию о границах ареалов распространения членов междиалектных соответствий.

В работе использована методика многоступенчатой кластеризации. На первом этапе на основе совокупности вопросов и ответов, содержащих анализируемую фонему или морфему, были разработаны карты отдельных явлений и категорий, представляющие собой

² Результаты кластеризаций, полученных разными методами, не обнаруживают существенных отличий между собой, но наилучшим образом (при сравнении визуализированных данных вручную) проявили себя методы полной связи и к-средних [Новак, Крижановская 2022].

традиционные лингвогеографические карты. В данном случае применение кластерного анализа позволило максимально объективно провести изоглоссы в переходных зонах. Далее на основе данных полученных карт были обчислены сбалансированные сводные кластерные карты по вокализму, консонантизму, фонологии, именному словоизменению, глагольному словоизменению и лексике. Анализ диалектных различий этих уровней играет решающую роль в процессе выделения единиц разных ярусов иерархической структуры языка. На заключительном этапе на основе данных сводных карт была произведена итоговая кластеризация, демонстрирующая усредненные границы диалектных ареалов.

В качестве дополнительных в процессе исследования были привлечены диалектометрический метод анализа когнатов лингвистической платформы ЛингвоДок [LingvoDoc], использованный для определения основных фонологических сопоставительных явлений, сравнительно-исторический метод с приемами внешней и внутренней реконструкции, филологический метод, а в ходе полевой работы также методы непосредственного наблюдения за речью носителей диалектов, прямого опроса, интервьюирования с записью на цифровые носители.

В настоящей статье представлен комплекс сопоставительных явлений, послуживших основой для разработки диалектной карты карельского языка, основанной исключительно на структурных критериях.

1. Фонологические диалектные различия

Для карельской фонологической системы характерна доминирующая роль гласных над согласными. Привлечение диалектометрических алгоритмов позволило выявить несколько комплексов диалектных различий в **вокалической системе** карельского языка.

1.1. Наиболее богатой на членов сопоставительных явлений оказалась система расширяющихся дифтонгов. В карельских говорах представлено по шесть рефлексов прибалтийско-финских долгих гласных **aa*, **ää* ударного слога, например, *maa / tua / moa / tuo / too / mja* ‘земля’, *pää / piä / peä / pie / pee / pöä* ‘голова’, см. Приложение (табл., строка 1). Пару рефлексов *aa ~ ää* следует относить к архаизмам, широко распространенные пары *oa ~ eä*, *ua ~ iä* — к общекарельской инновации, тогда как остальные члены представляются поздними образованиями.

Рефлексы заударных **aʏa*, **äʏä*, **ada*, **ädä* в собственно карельских и ливвиковских говорах в целом совпадают с ситуацией ударного слога, например, *elää / eleä / elyä / eliä / elee / elöä / elie / elöä* ‘жить’ (строка 2). На месте праязыковых **uʏa*, **uʏä*, **ida*, **idä*, **oʏa*, **oda*, **eʏa*, **eʏä*, **eda*, **edä*, **iʏa*, **iʏä*, **ida*, **idä* (строки 3, 4) выступают пракарельские *ua*, *yä*, *ia*, *iä* (*koivua* парт. ед. ч. от *koivu* ‘береза’, *pimiä* ‘темный’), общекарельские инновационные *uo*, *yö*, *ie* (*koivuo*, *pimie*), древнеливвиковский *ei* (*pimei*), а также в собственно карельских держанских говорах поздние *uu*, *yy*, *i* (*koivuu*, *kivi*). Людиновское наречие отличается сохранением интервокальных согласных *d*, *g* в заударной позиции (*elädä*, *koivud*, *pimed*), что возводится к древневепсскому наследию.

1.2. Рефлексы сужающихся дифтонгов представлены более единообразно. Однако для южных собственно карельских говоров Карелии и северных людиновских говоров характерен относимый к вепсскому субстратному явлению переход второго компонента дифтонгов на *y* в заднерядное соответствие *u* (*keuh(ä/e)* < *keyhä* ‘бедный’), для тихвинских собственно карельских — поздняя монофтонгизация (*lookko* < *loukko* ‘дырка’), для ряда собственно карельских южных и основной части ливвиковских говоров — позднее развитие полугласного *w* / согласного *v* в позиции перед *v* (*uv/wvet* ‘новые’ < *uuvet*) (строка 5).

Сужающиеся дифтонги на *i* ударного слога утратили второй компонент перед *sk* / *šk*, *st* / *št* (строка 6) в ливвиковском (кроме северных говоров) и людиновском наречиях (*las* / *šk(u)* ‘ленивый’, ср. с.к.: *laiska*). Выпадение второго компонента дифтонга затронуло также

формы имен на *-ise (строка 7) в собственно карельских и ряде ливвиковских говоров (*kuldaneli* ‘золотой’, *kuldaz/žet* ‘золотые’, ср. люд.: *kuldaine*, *kuldaizet* / *kuldaizet*). Михайловским людиковским говорам свойственно позднее явление монофтонгизации (*kuldaane*, *kuldaazet*).

Сохранение праязыковых ауслутных дифтонгов *oi, *öi (строка 8) свойственно ливвиковским и северным людиковским говорам (*reboi* ‘лиса’). В собственно карельском наречии второй компонент дифтонга не представлен (*rebo*), в южно- и среднелюдиковских говорах обнаружены трифтонги (*rebuoi* > *rebuo*), а в михайловских — монофтонги (*rebuu*) позднего происхождения.

1.3. Важнейшие диалектные различия обнаруживает **конечная огласовка начальных форм номинатива единственного числа одноосновных имен на -a, -ä** (строка 9). В двусложных именах с долгим³ первым слогом гласные сохраняются в собственно карельском наречии (*leibä* ‘хлеб’), в современных ливвиковских говорах продолжение нашел древнеливвиковский переход *a > u, *ä > y (*leiby*; в кондушских говорах — в o, ö (*leibö*)), в среднелюдиковских и южнолюдиковских говорах — в e (*leibe*), в остальных говорах людиковского наречия и держанском собственно карельском говоре конечные гласные бесследно исчезли (*leib*). Выпадение конечных гласных *a, *ä также характерно для ряда многосложных имен ливвиковских и людиковских говоров (*ikkun* ‘окно’, ср. с.к.: *ikkuna*).

1.4. Анализ диалектных данных выявил наличие более и менее обильных нерегулярных случаев **отступления от закона гармонии гласных** (строка 10) в севернелюдиковских, среднелюдиковских и михайловских говорах, что объединяет их с вепским языком (*ižand(e)* ‘хозяин’, *kezrata* ‘прясть’, ср. с.к., ливв.: *ižändä/y*, *kežrätä*).

1.5. В позиции ударного слога в положении после губно-губных (*p, b, m*) и губно-зубного (*v*) согласных и перед ними или перед лабиализованными гласными (*u, y*) **лабиализация гласного i** (строка 11), имеющая явные древневепские корни, прослеживается в говорах ливвиковского и людиковского наречий и примыкающих к ним южных собственно карельских говорах (*pystäy* / *pyštäb* ‘колет’, ср. с.к.: *pistäy*). Гласный *e* в заударной позиции перед согласными *m, v* (строка 12), напротив, подвержен регрессивной лабиализации в основной массе собственно карельских говоров Карелии и центральных толмачевских говорах (*vägöväť* ‘сильные’, *lugomah* III инфинитив, иллатив от ‘читать’, ср. ливв., люд.: *vägevät*, *lugemah/i*), что дает основание относить явление к древнекарельской инновации.

Карельский диалектный консонантизм отличается от вокализма отсутствием многочисленных диалектных различий.

1.6. Для северных собственно карельских диалектов характерно отсутствие **оппозиции согласных по глухости / звонкости** (*koti* ‘дом’, *vesi* ‘вода’, ср. южнокарельские диалекты с.к., ливв., люд.: *kodi*, *vezi*), что объясняется результатом позднего влияния со стороны соседних говоров финского языка (строка 13).

1.7. Один из важнейших диалектно-дифференцирующих признаков карельского консонантизма — **дистрибуция переднеязычных щелевых согласных s / š, z / ž** (строки 14–29). Она зависит от нескольких факторов, в одних диалектах выступающих независимо друг от друга, в других — в совокупности и при этом в различных возможных комбинациях: позиция в слове, фонетическое окружение, качество гласных слова. Согласно результатам сводной кластеризации по данному фонологическому явлению, современные карельские говоры подразделяются на две крупные зоны: собственно карельскую, характеризующуюся использованием свистящих согласных в позиции после гласного *i* (*šeižou* ‘стоит’, *laiska* ‘ленивый’) и широким употреблением шипящих согласных в остальных позициях (*šeinä* ‘стена’, *uuži* ‘новый’, *lapši* ‘ребенок’), и ливвиковско-людиковскую, демонстрирующую противоположную ситуацию (*seižov*, *lašku/e*, *seiny/(e)*, *uuzi*, *laps(i)*). Формирование двух зон с полностью противоположной картиной дистрибуции переднеязычных щелевых

³ Долгий слог — слог, содержащий долгий гласный, дифтонг или закрытый слог.

указывает на их древнекарельский инновационный характер. На территории Средней Карелии после массовых переселений карелов в результате перекрестного влияния этих двух зон произошло смешение случаев употребления свистящих и шипящих согласных, что привело к появлению здесь широкой полосы переходных говоров.

1.8. К отличительным чертам, объединяющим южнокарельские говоры собственно карельского наречия Карелии, северо-западные и кондушские говоры ливвиковского наречия, а также все людиковские говоры, относится **переход $j > d'$** (*d'yvā* 'зерно', *kard'u* 'стадо', ср. севернокарельские говоры с.к.: *jyvā*, *karja*). Данный переход следует возводить к довольно молодым явлениям карельской диалектной речи (примерно конец XVII — начало XVIII в.) с чисто ареальной природой (строки 30, 31).

1.9. Степень **палатализации** переднеязычных согласных (строки 32–44) в карельских говорах значительно варьируется и зависит как от их артикуляционных свойств и положения в слове, так и от непосредственного фонетического окружения. В ливвиковском наречии явление представлено минимально (*lehmy* 'корова', *tedri* 'тетерев', *nenä* 'нос', *n'ivel* 'сустав'). Людиковское наречие обнаруживает здесь общие черты с вепскими говорами (*l'ehm*, *t'edri* / *tedri*, *nenä*, *n'ivel* / *n'ivel*). Наиболее широкое распространение палатализация получила в южнокарельских говорах собственно карельского наречия (*l'ehmä*, *t'edr'i*, *n'en'ä*, *n'ivel*), что возводится еще к древнекарельскому периоду развития языка. В его старокарельский период двуязычие карелов способствовало расширению явления на востоке собственно карельской территории и в Центральной России, тогда как влияние финского языка на северо-западе, наоборот, сдерживало его развитие.

1.10. Ряду ливвиковских и людиковских лексем (наречиям, числительным и частицам) свойственна **утрата конечного согласного n** (*kaheksa* 'восемь', *enne* / *ende* 'прежде', ср. с.к.: *kahekšan*, *enne(i)n*), что относится к довольно древнему фонетически обусловленному явлению (строка 45).

1.11. В **системе удвоенных согласных** (строка 46) удалось очертить ареал относительно позднего сокращения постсонорных смычно-взрывных геминат говорами собственно карельского наречия Центральной России, юго-западной половиной ливвиковских, михайловскими и отдельными средне- и севернолюдиковскими говорами, например, *värtin(ä)* 'веретено', *čirk(u)* / *čirkoi* 'воробей', ср. говоры с.к. Карелии: *värttinä*, *čirkku*.

1.12. Для михайловского людиковского говора характерны такие поздние инновационные черты, вызванные влиянием со стороны соседних говоров вепского языка, как нерегулярный **переход $l > u$** в словах заднерядного вокализма⁴ (*peud* / *pseudo* 'поле', ср. с.к., ливв.: *peldo*) (строка 47) и **использование геминированной аффрикаты cc** (*mecc* 'лес', ср. с.к., ливв.: *meččä/y*) (строка 48).

1.13. Южнокарельским говорам собственно карельского наречия, северо-восточным и кондушским ливвиковским, а также людиковским говорам свойственно позднее инновационное явление **метатезы сочетания согласных nh** (строка 49), например, *tahn* / *tahnut* / *tahnno* 'двор', ср. севернокарельские говоры с.к.: *tanhut* / *tanhuo*.

1.14. Собственно карельские пада́нские и восточные ругозерские говоры обнаруживают позднее явление **выпадения интервокального согласного v** (*hyä* 'хороший', ср. остальные говоры: *huvä*) (строка 50).

Результаты **сводной кластеризации по вокализму** разбили говоры карельского языка на три группы, соответствующие наречиям с традиционной диалектной карты. Формирование основной части диалектно-дифференцирующих явлений вокалической системы восходит к древнекарельскому периоду функционирования языка. На более дробном уровне кластеризации (рис. 2) из красного собственно карельского кластера обособились фиолетовая группа пада́нских и синяя группа поросозерских и мяндусельгских говоров,

⁴ В словах заднерядного вокализма обязательно наличие гласных заднего ряда *o*, *a*, *u* при их возможном соседстве с переднерядными *e*, *i*.

из оранжевого ливвиковского — коричневая группа северных сямозерских, темно-зеленая группа кондушских и зеленая группа видлицких и южных тулмозерских говоров, людиковский кластер разбился на три группы — малиновую группу севернолудиковских и северных среднелудиковских, голубую группу южнолудиковских и южных среднелудиковских и бирюзовую группу михайловских говоров.

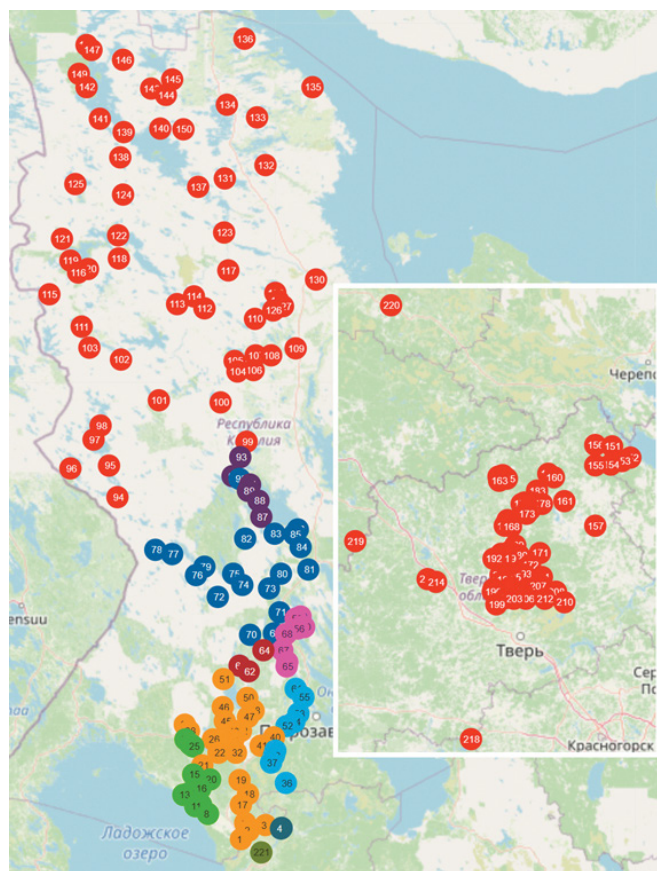


Рис. 2. Сводная кластеризация по вокализму

Консонантная диалектная система карельского языка определенно представляется более молодой, чем вокалическая. Большая часть выявленных диалектных различий — поздние инновации ареального характера. К древнекарельским восходят лишь переход свистящих щелевых в шипящие, отдельные случаи палатализации и выпадение конечного *л*. На **сводной кластерной карте по консонантизму** сформировалось три крупные группы говоров: собственно карельский с севернолудиковскими и северными среднелудиковскими говорами, ливвиковско-лудиковский и михайловский (коричневый). На более дробных уровнях кластеризации (рис. 3) из первого красного кластера выделилась малиновая группа паданских говоров и синяя группа южных говоров Карелии с северными лудиковскими говорами, из голубого ливвиковско-лудиковского — зеленая группа северных тулмозерских и западных ведлозерских говоров и оранжевая группа видлицких и южных тулмозерских говоров. Таким образом, за счет большого числа поздних инноваций сводные изоглоссы по консонантизму разошлись с аналогичными по вокализму и, соответственно, с границами наречий.

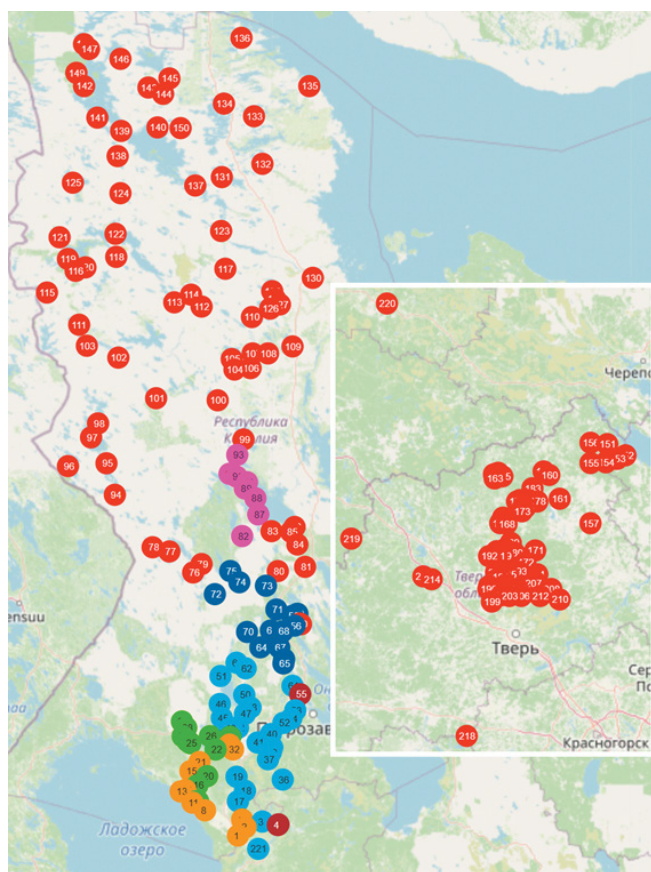


Рис. 3. Сводная кластеризация по консонантизму

2. Морфонологические диалектные различия

Характерной особенностью людиковской морфонологии, объясняемой результатом влияния древневепсского строя, является практически полное отсутствие в ней качественного вида альтернации согласных, заключающейся в выпадении или замене другими согласными в определенных словоизменительных формах смычно-взрывных *k*, *g*, *t*, *d*, *p*, *b*. Собственно карельская и ливвиковская системы чередования ступеней согласных обнаруживают здесь ряд диалектных различий.

2.1. Слабоступенными альтернантами одиночных *g*, *d* в зависимости от непосредственного фонетического окружения выступают нуль звука (0) или носящие вставочный характер альтернаты *j* (после *i*, *e*) и *v* (при наличии рядом лабиализованного гласного).

По позиции после кратких гласных *i*, *e* (строки 51, 52) выделилась группа собственно карельских северных говоров и говоров Центральной России, например, *hiješšä* ‘в поту’ (< *hige-*), *vejet* ‘воды’ (< *vede-*), ср. южнокарельские говоры с.к. Карелии, ливв.: *hieššä* / *hies*, *viet* / *veet* / *veit*. Очевидно, изначально вставочные согласные в слабой степени

в древнекарельском языке были представлены лишь в позиции после долгих ударных слогов, но постепенно они стали проникать и в другие интервокальные позиции, что привело к формированию членов рассмотренного междialeктного соответствия.

Паданские собственно карельские говоры, в свою очередь, выделяются поздними инновационными явлениями. К ним относятся использование в слабой ступени нуля звука даже при наличии в непосредственной близости лабиализованного гласного (*hauat* ‘ямы’ (< *hauda*), *maot* ‘черви’ (< *mado*), ср. остальные говоры с.к. и ливв.: *hauv/vvat*, *mavot*) (строка 53), а также отсутствие чередования *b : v* (*hebot* ‘лошади’ (< *hebo*), ср. остальные говоры с.к. и ливв.: *hevot*) (строка 54).

2.2. Слабоступенными соответствиями сочетаний *lg, rg* (строка 55) в северной половине говоров собственно карельского наречия выступают одиночные согласные *l, r* (*jalat* ‘ноги’ (< *jalga*), *arat* ‘робкие’ (< *arga*)), тогда как в остальных собственно карельских и ливвиковских говорах представлены геминированные *ll, rr* (*jallat / d'allat, arrat*). **Сочетания согласных *ld, rd, nd, mb*** (строка 56) в слабой ступени ведут себя везде одинаково, но в ливвиковских говорах важную роль сохранило положение содержащего чередующийся согласный слога в слове (*pellot* ‘поля’ (< *peldo*), но ливв.: *izändät* ‘хозяева’ (< *izändä-*), ср. с.к.: *izännät*) — чередование возможно только в позиции после ударного слога⁵.

2.3. Наличие диалектных различий демонстрирует также альтернация смычно-взрывных, выступающих в позиции после глухих согласных. В ливвиковском наречии, как и в людиковском, чередование здесь не представлено (*pitkät* ‘длинные’ (< *pitkä-*), *lehtet* ‘листья’ (< *lehte-*), *mustat* ‘черные’ (< *musta-*), *kosket* ‘пороги’ (< *koske-*)). В собственно карельских говорах смычно-взрывные согласные сочетаний *tk, ht* выпадают (*pität, lehet*) (строка 57), дентальный согласный из сочетания *st / št* (за исключением ряда позиций в говорах Центральной России) ассимилируется с переднеязычным щелевым (*muššat / mussat*) (строка 58), велярный согласный сочетания *sk / šk* в южнокарельских говорах подвержен ассимиляции (*kosset / koššet*), а в севернокарельских — выпадению (*košet / kožet / kosei*) (строка 59). Широкое распространение этих видов чередования во всех собственно карельских говорах позволяет возводить их к древнекарельскому периоду функционирования языка.

На верхнем уровне **сводной кластерной карты по качественной альтернации согласных** образовалось три группы говоров, изоглоссы которых совпали с границами наречий. Практически лишенная внутренних диалектных различий ливвиковская альтернация (рис. 4, с. 68, голубой кластер) выделилась за счет своего архаичного характера. Людиковская морфонология (малиновый кластер) продемонстрировала наличие сильного древневеппского субстрата. Определенная диалектная вариативность древнекарельской морфонологии послужила причиной разбиения собственно карельского кластера на более дробном уровне кластеризации на четыре группы: оранжевую группу севернокарельских говоров, коричневую группу говоров Средней Карелии и две группы (красную и зеленую) говоров Центральной России.

3. Морфологические диалектные различия

Членами морфологических соответствий карельского языка выступают как отдельные словоизменятельные аффиксы (алломорфы и их фонемный состав), так и целые словоизменятельные парадигмы.

⁵ В карельском языке фиксированное ударение, оно всегда падает на первый слог слова.

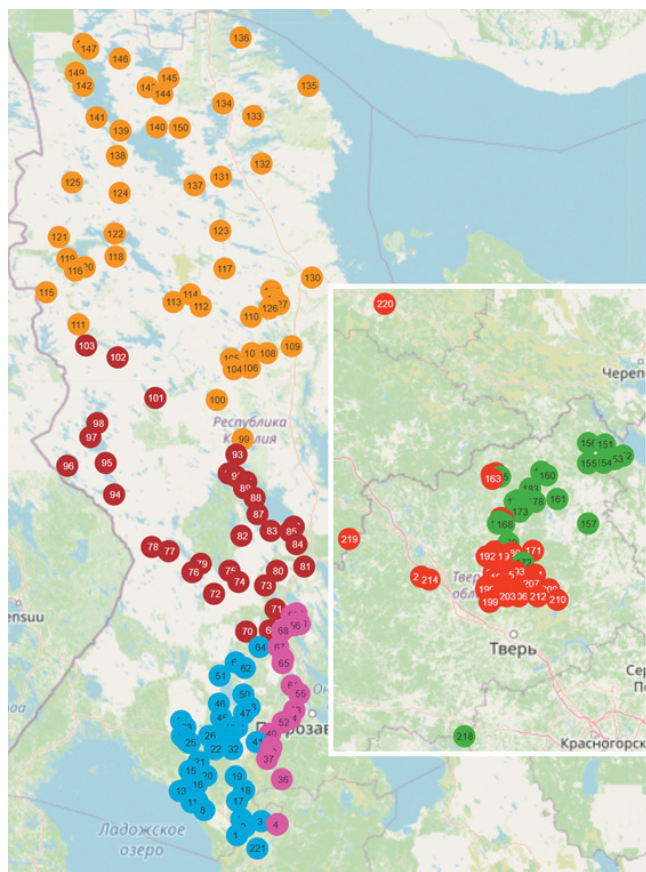


Рис. 4. Сводная кластеризация по морфологии

3.1. Количество падежей в именной словоизменительной системе в диалектах карельского языка варьируется, что объясняется результатом слияния нескольких падежных аффиксов в один или разной степенью продуктивности отдельных из них. Диалектные различия обнаруживает структура показателей следующих **падежей**:

- генитива множественного числа: говоры с.к. Карелии, ливв.: *lapši-en* ‘ребенок-PL.GEN’, говоры с.к. Центральной России: *lapšii-n* / *lapšiloi-n*, люд.: *lapši-den* (строка 60);
- партитива единственного числа: с.к., ливв.: *kivi-e* / *kive-e* / *kivi-ä* / *kivi-i* ‘камень-PART’, люд.: *kive-d* / *kive-t*; с.к.: *korgielia-da* ‘высокий-PART’, ливв.: *korgielia-du*, люд.: *korgeda-d* / *korgeda-t* (строка 61);
- партитива множественного числа: говоры с.к. Карелии: *šano-ja* ‘слово-PL.PART’, говоры с.к. Центральной России: *šanoi-da*, ливв.: *sanoi* / *sanoi-du*, люд.: *sanoi-d* / *sanoi-t* (строка 62);
- эссива: с.к.: *paimeni-na* ‘пастух-PL.ESS’, ливв.: *paimoloi-nni*, люд.: *paimoloi-n* (строка 63);
- транслатива: с.к.: *opastaja-kši* ‘учитель-TRANS’, ливв.: *opettaja-kše*, люд.: *opendaja-ks* (строка 64);
- инессива и адессива: с.к.: *kylä-ššä* ‘деревня-INESS’, *veiče-llä* ‘нож-AD’, ливв., люд.: *kylä-s*, *veiče-l* (строка 65);

- элатива и аблатива: с.к.: *mečä-štä* ‘лес-EL’, *kaivo-lda* ‘колодец-AVL’, ливв., люд.: *mečä-s / mečä-späi / mečä-spiäi*, *kaivo-l / kaivo-lpäi / kaivo-lpiäi* (строка 66);
- аллатива: с.к.: *uko-lla* ‘старик-ALL’, ливв.: *uko-l*, люд.: *uko-le* (строка 67);
- иллатива: с.к., ливв., южнолюдиковские говоры: *kyly-h* ‘баня-ILL’, паданские и тихвинские говоры с.к.: *kyly-y*, северно- и среднелюдиловские говоры: *kyly-i*, михайловские говоры люд.: *kyly-he / kyly-hy*; михайловские говоры люд.: *venehe-ze* ‘лодка-ILL’, остальные говоры: *venehe-h / venehe-e / venehe-i* (строка 68);
- абессива: с.к.: *denga-tta / raha-tta* ‘деньги-AVV’, ливв.: *d'enga-ttah*, южные говоры люд.: *den'ga-ta*, северные говоры люд.: *denga-ttai* (строка 69);
- комитатива: говоры с.к. Карелии: *akko-neh* ‘старуха-СОМ’, говоры с.к. Центральной России, ливв., люд.: *tuatto-nke(n'nal/lr)* ‘отец-СОМ’ (строка 70). В говорах собственно карельского наречия Карелии значение комитатива передается также послеложной конструкцией: *tuaton kera* ‘с отцом’.

Падеж послеложного образования аппроксиматив с формантами *-lluo / -llyö* представлен в ливвиковских и южнокарельских говорах собственно карельского наречия Карелии и южнолюдиковских говорах, с формантом *-ppo, -ppu* — в людиковских говорах, тогда как в остальных собственно карельских говорах для выражения данного значения используется послеложная конструкция (строка 71). Терминатив на *-ssah / -ssäh* характерен только для ливвиковских и соседних с ними южнолюдиковских говоров (строка 72).

3.2. В системе глагольного словоизменения наличие диалектных различий демонстрируют **лично-числовые окончания**:

- 3 лица единственного числа: с.к., восточные говоры ливв., люд.: *šano-u / šano-v* ‘сказать-PRES.3SG’, западные говоры ливв.: *šano-h / sano-o*, михайловские говоры люд.: *sanu-b* (строка 73);
- 1 и 2 лица множественного числа: с.к.: *löyvä-ttä* ‘находить-PRES.2PL’, *kando-ma* ‘нести-PRES.1PL’, ливв.: *löyvä-ttö, kannoi-mmo*, люд.: *löydä-tte, kandoi-mme*, михайловские говоры люд.: *löydä-t, kandoi-m* (строки 74, 75);
- 3 лица множественного числа: паданские и тихвинские говоры с.к.: *pu-rra-a* ‘кусать-PRES.3PL’, северно- и среднелюдиловские говоры: *pur-da-u / pur-da-v*, михайловские говоры люд.: *pur-da-ze*, остальные говоры: *pur-ra-h / pur-da-h* (строка 76).

Кроме того, отличия обнаруживает образование форм 3 лица множественного числа от одноосновных глаголов с основой на *-e*. В ливвиковских, собственно карельских говорах Центральной России и в северо-восточной части южнокарельских говоров Карелии присоединение показателя осуществляется напрямую к форме I инфинитива (*lugie-ta-h* ‘читать-PRES.3PL’ при форме I инфинитива *lugie*), тогда как в остальных говорах — к основе слова (с.к.: *leve-ta-h*, люд.: *luge-da-h*, ср. *luve-n / luge-n* ‘читать-PRES.1SG’) (строка 77).

3.3. Наличие диалектного различия показало **образование форм имперфекта** от двусловных стяженных глаголов (строка 78). Формант *-(i)zi- / -ži-* в них представлен только в собственно карельских говорах Карелии и людиковских говорах (*ava-(i)zi-t / ava-ži-t* ‘открыть-IMPERF-2SG’, остальные говоры: *ava-i-t*). Для собственно карельских говоров Центральной России, юго-восточных ливвиковских и михайловских людиковских говоров, в отличие от всех остальных, характерно также употребление геминированного показателя имперфекта *-ii-* (*ot-ii-n* ‘брать-IMPERF-1SG’, ср. остальные говоры: *ot-i-n*) (строка 79).

3.4. Из косвенных наклонений наибольшее число диалектных различий обнаружили показатели 2 лица множественного числа **императива**: с.к.: *anda-kkoa / anda-kkua / anda-kkaa / anda-kkuo / anda-kkja* ‘давать-IMPER.2PL’, люд.: *anda-gat / an-kat / anda-gatte*, ливв.: *ando-at / anda-at / andu-at / ando-ot / andu-ot / andu-ade / ando-ade / anda-atto* (строка 80). Кроме того, южнокарельские говоры собственно карельского наречия и отдельные ливвиковские

говоры демонстрируют возможность употребления слабой ступени согласных компонентов лексической основы слова в позиции перед показателем императива (строка 81).

Формы 3 лица императива в говорах собственно карельского наречия образуются с помощью аффикса *-kkah / -kkäh (elä-kkäh / elä-kkää* ‘жить-IMPER.3’), в западных ливвиковских говорах — *-hez(e) (elettä-hez)*, тогда как в остальных говорах синтетические формы не сохранились (строка 82).

Отличия демонстрируют также анлаутная огласовка отрицательных форм императива: с.к., северные людиковские говоры: *elgyä / elgiä / elgeä / elgat* ‘NEG.IMPER.2PL’, ливв., южнолюдиковские говоры: *älgeä*, михайловские говоры люд.: *algat* (строка 83).

3.5. Кластеризация позволила выделить ареалы употребления в говорах карельского языка полного (люд.: *anda-iži(i)-n / anda-izi-n* ‘давать-COND-1SG’, с.к., люд.: *anda-is / anda-iž / anda-iš* ‘давать-COND.3SG’) и сокращенного (с.к., ливв.: *anda-zi(i)-n / anda-ži-n*; ливв.: *anda-s / anda-š / anda-ž*) показателей **кондиционала** (строки 84, 85). Кроме того, ливвиковское и людиковское наречия дифференцируют возводимое к вепсскому языку наличие особых синтетических форм имперфекта кондиционала, не свойственных собственно карельским говорам (строка 86).

3.6. Диалектные различия демонстрирует показатель **потенциала**, выступающий в говорах карельского языка с геминированным или негеминированным согласным: говоры с.к. Центральной России: *anda-nno-u* ‘давать-ROT-3SG’, говоры с.к. Карелии, ливв., люд.: *anda-no-u* (строка 87).

3.7. Из **именных глагольных форм** диалектно-дифференцирующим потенциалом обладают показатели:

- I инфинитива одноосновных глаголов с основой на дифтонг: с.к., ливв.: *sua-ha / soa-ha* ‘достать-INF’, люд.: *su/oa-da*, некульские и рыпушкальские говоры ливв.: *suo-ja* (строка 88);
- II инфинитива: с.к.: *oštu-o-šša / oštu-o-s š’a* ‘покупать-INF2-INESS’, ливв.: *osta-je-s / osta-e-s*, люд.: *osta-de-s/z* (строка 89);
- I причастия актива: с.к.: *lugi-ja* ‘читать-IPFV.PTCP.ACT’, ливв., люд.: *lugi-i* (строка 90);
- II причастия актива: тунгудские, паданские, ребольские, поросозерские говоры с.к. и видлицкие говоры ливв.: *et tul-lut* ‘NEG.2SG приходить-PFV.PTCP.ACT’, остальные говоры с.к.: *et tul-lun*, остальные говоры ливв.: *et tul-luh*, люд.: *et tul-nu* (строка 91);
- партитивной формы II причастия пассива: говоры с.к. Карелии и южные говоры ливв.: *ošš-ttuio / ošte-ttuio* ‘купить-PFV.PTCP.PASS.PART’, говоры с.к. Центральной России, северные говоры ливв., люд.: *osta-huu / osta-hui / osta-huo / osta-hut*, северно- и среднелюдиковские говоры: *osta-ttuhut* (строка 92).

3.8. Характерной чертой, выделяющей ливвиковские и людиковские говоры на фоне собственно карельских, является наличие в них **возвратного спряжения**, обладающего особым набором лично-числовых окончаний (строка 93), что вновь объясняется влиянием со стороны древневепсского языка.

Сводные кластерные карты по именному и глагольному словоизменению на верхних уровнях продемонстрировали наличие четких пучков изоглосс, совпадающих с границами между тремя наречиями. Формирование основной части выведенных изоморфов восходит, очевидно, к раннему периоду функционирования языка.

На более подробном уровне **сводной кластеризации по именному словоизменению** (рис. 5) из оранжевого собственно карельского кластера отделились красная группа собственно карельских говоров Центральной России с юго-восточными южнокарельскими говорами Карелии, бирюзовая группа южных мяндусельских говоров, а также малиновая группа держанских говоров, из зеленого ливвиковского — фиолетовая группа кондушских говоров, из коричневого людиковского — голубые южнолюдиковские говоры.

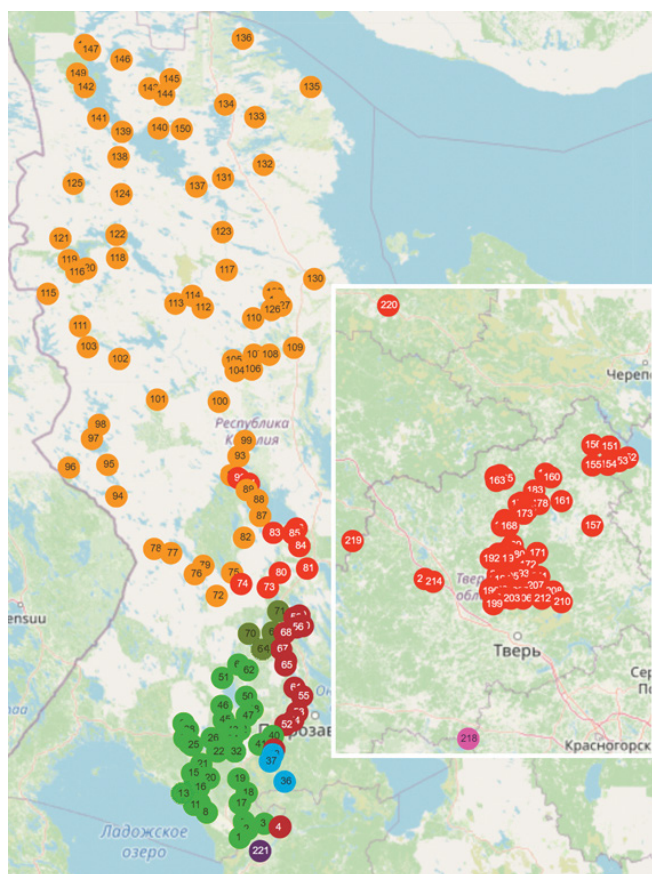


Рис. 5. Сводная кластеризация по именному словоизменению

На соответствующей карте **по глагольному словоизменению** (рис. 6, с. 72) из оранжевой собственно карельской группы обособились красный кластер собственно карельских говоров Центральной России и синий кластер собственно карельских восточных поросозерских и западных мяндусельских говоров с северными ливвиковскими говорами, людиковская группа разбилась на малиновый кластер севернолюдиковских и северных среднелюдиковских говоров, голубой кластер южнолюдиковских и южных среднелюдиковских говоров и коричневый кластер михайловских говоров, тогда как зеленая ливвиковская группа сохранила свое единство.

4. Лексические диалектные различия

В целях разработки диалектной карты карельского языка, основанной на данных максимального числа языковых уровней, к представленным в статье сводным кластеризациям была добавлена также карта по лексике, разработанная на основе материалов соответствующего раздела диалектной базы Murreh. На верхнем уровне этой карты говоры карельского

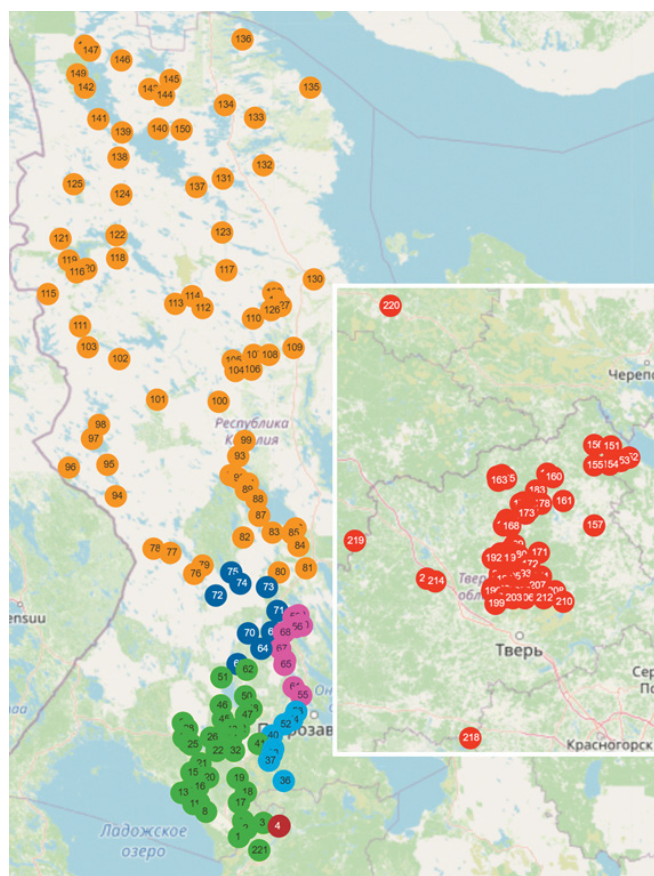


Рис. 6. Сводная кластеризация по глагольному словоизменению

языка разбились на два кластера: собственно карельский (например, *ovi* ‘дверь’, *tuah* ‘навоз’, *tujeh* ‘ряпушка’, *valehella* ‘лгать’, *kuotella* ‘пробовать’) и ливвиковско-людиковский (например, *uksi* ‘дверь’, *höštö* ‘навоз’, *riäpöi* ‘ряпушка’, *kielastua* / *tuaniittada* ‘лгать’, *oppiä* ‘пробовать’).

При более дробной кластеризации (рис. 7, с. 73) собственно карельский кластер разделился на три группы: обнаруживающие довольно большое число общих черт с финским языком северные говоры собственно карельского наречия (оранжевый кластер, например, *hiät* ‘свадьба’, *laulu* ‘песня’), ощутившие заметное влияние со стороны ливвиковского и людиковского наречий южные собственно карельские говоры Карелии (голубой кластер, не обнаруживает каких-либо ярких собственных черт) и демонстрирующие наличие высокого процента поздних заимствований из русского языка собственно карельские говоры Центральной России (красный кластер, например, *broni* ‘ворона’, *juablokka* ‘картофель’). Ливвиковско-людиковская группа разбилась на два кластера: зеленый ливвиковско-южнолюдиковский, выделившийся за счет наличия характерного исключительно для него пласта лексики (например, *vuitti* ‘доля’, *sälgy* ‘жеребенок’, *leuhkia* ‘хвастаться’, *suittoa* ‘копить’), и малиновый людиковский (без южнолюдиковских говоров), обнаруживающий наибольшую концентрацию общих с вепским языком лексем (например, *ägez* ‘борона’, *verduda* ‘сердиться’).

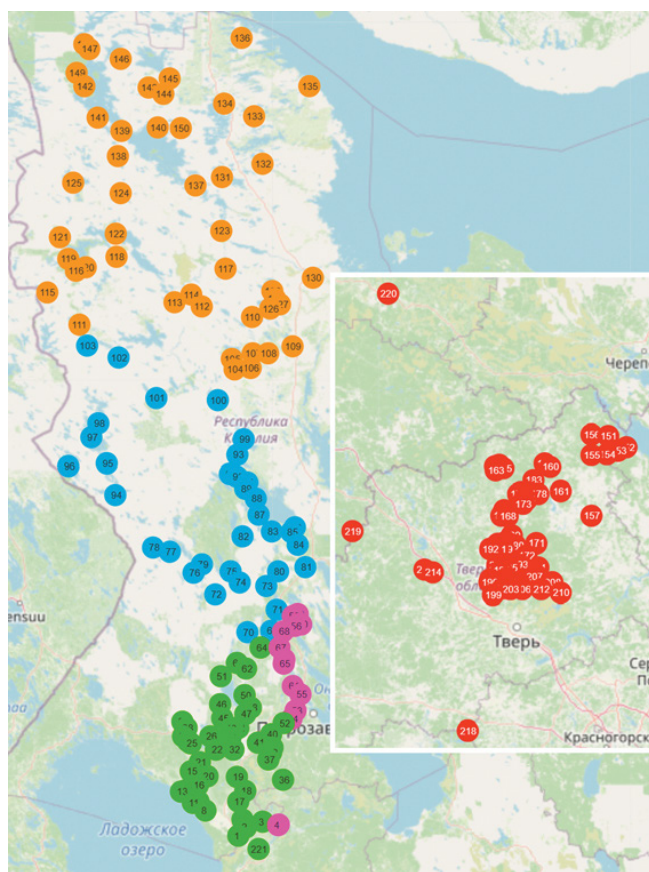


Рис. 7. Сводная кластеризация по лексике

Заключение

Карельский языковой ландшафт — сложная многоуровневая система, сформировавшаяся в результате одновременных и разнонаправленных миграций карелов, сопровождавшихся контактами с носителями близкородственных и неродственных языков. В ходе настоящего исследования кроме большого числа диалектных различий карельские говоры продемонстрировали общность многих черт всех языковых уровней. В совокупности с этим, убедительно проявивший себя на картах дискретно-континуальный характер карельской диалектной речи позволяет сделать вывод об отсутствии сравнительно-исторических оснований для отделения людиковских, ливвиковских и тем более тверских карельских говоров от карельского языка.

На основе данных шести сводных карт была проведена **итоговая кластеризация** (рис. 8), на верхнем уровне которой оформилось три группы, соответствующие наречиям карельского языка из традиционной классификации: собственно карельскому, ливвиковскому и людиковскому. Оформление наречий базируется на совокупности изоглосс обладающих устойчивым характером морфологической и морфонологической систем, восходящих к раннему периоду развития языка (задолго до периода массовых переселений карелов).

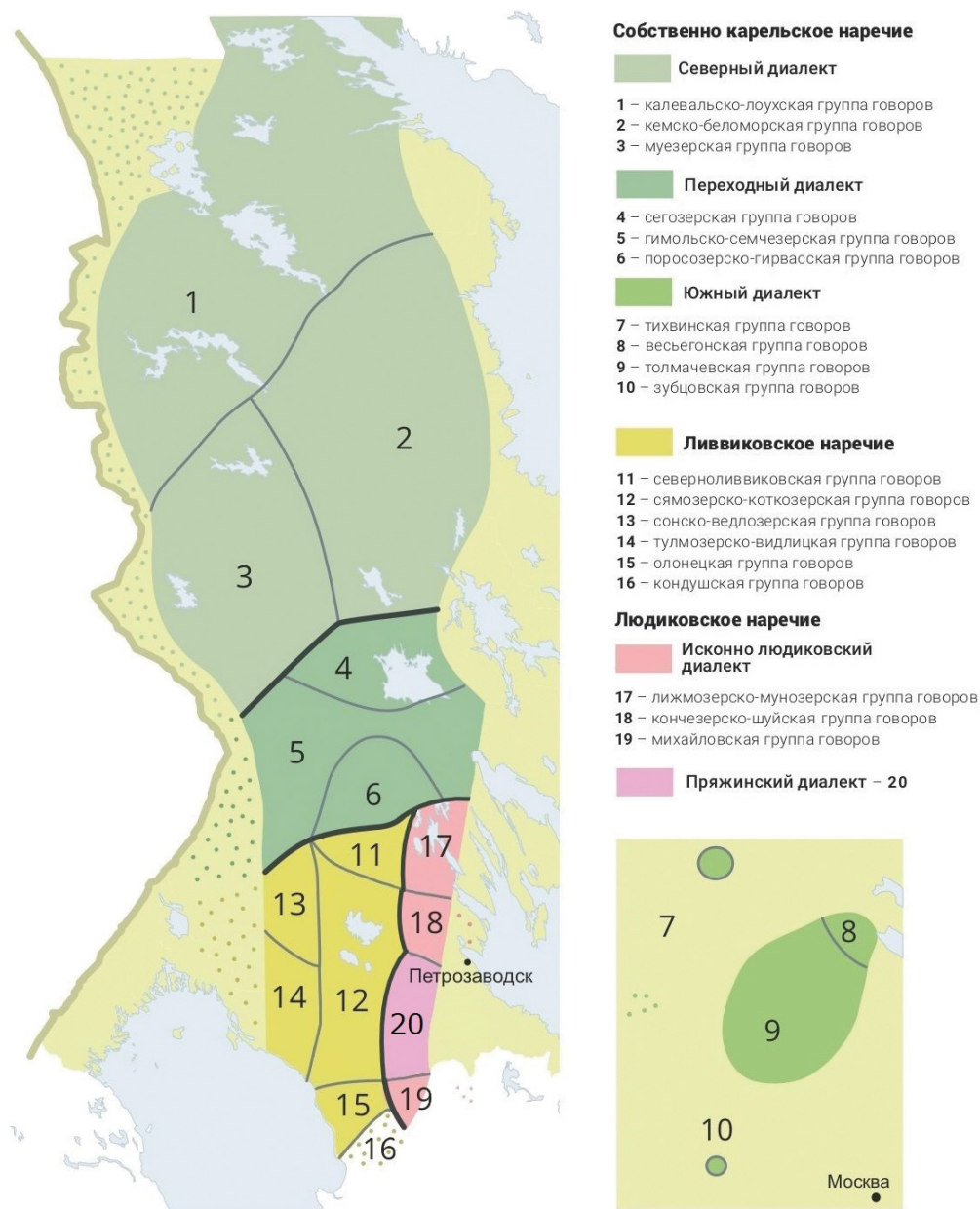


Рис. 8. Диалектная карта карельского языка

По пучкам изоглосс морфонологической и фонологической (в большей степени вокалической) систем, отражающих поздние древнекарельские диалектные различия, а также результаты интенсивных контактов в пограничных между наречиями зонах, в собственно карельском наречии оформилось три диалекта: северный (говоры Лоухского, Калевальского, Кемского, Беломорского, Муезерского районов и Костомукшского ГО Карелии), южный (говоры Тверской, Ленинградской и Новгородской областей) и характеризующийся промежуточным положением между наречиями переходный (говоры северной части

Людиковское наречие разделилось на два диалекта: исконно людиковский (людиковские говоры Кондопожского и Олонецкого районов Карелии) и демонстрирующий наличие сильного влияния со стороны контактных ливвиковских говоров пряжинский (людиковские говоры Пряжинского района Карелии). Внимание на себя обращает довольно однородное в языковом плане ливвиковское наречие.

По изоглоссам фонологической (особенно консонантной) и лексической систем, изобилующих поздними инновациями (особенно в сегозерских, зубцовских, тихвинских, кондушских и михайловских говорах), следам как позднего влияния соседних финского (в северных и западных собственно карельских говорах Карелии), вепсского (в самых южных людиковских и ливвиковских говорах) и русского (в собственно карельских северо-восточных говорах и говорах Центральной России) языков, так и продолжавшегося межнаречного взаимодействия, диалекты карельского языка разбились на 20 групп говоров (рис. 9).

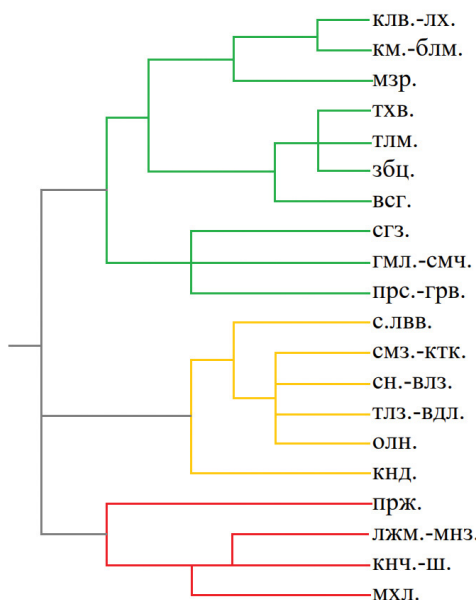


Рис. 9. Дендрограмма групп говоров карельского языка

Намеченные на итоговой карте границы диалектных ареалов, основанные на комплексе изоглосс всех языковых уровней, существуют не сами по себе. Повторяя географические, этнические, исторические, топографические границы, они четко укладываются в единую систему, подтверждая тем самым достоверность полученных результатов.

Работа в экспедициях 2014–2024 гг. показала, что разработанная трехуровневая диалектная классификация с учетом более не заселенных носителями территорий, отмеченных на итоговой карте точечной штриховкой, актуальна для современного карельского языка.

Проведенное исследование не просто помогло существенно расширить состав карельских диалектно-дифференцирующих явлений и уточнить границы ареалов, обнаруживающих их членов, но в совокупности с материалами памятников письменности и данными близкородственных языков позволило систематизировать эти диалектные различия во времени, благодаря чему была разработана **периодизация карельского языка**, представленная четырьмя периодами:

- 1) пракарельским (начало II тысячелетия — XIV в.), на который приходится формирование ижорского и древнекарельского языков, а также восточных диалектов финского языка;
- 2) древнекарельским (XIV — конец XVI в.), характеризующимся консолидацией наречий древнекарельского языка, а также началом сложения их диалектных различий. В северо-западном Приладожье оформилось древнее собственно карельское наречие, в котором наряду с прибалтийско-финским праязыковым компонентом реконструируется значительное число инноваций. Вдоль побережья Онежского озера на основе ярко выраженного древневепского субстрата в результате увеличивавшегося влияния со стороны языка древней корелы выделилось древнелюдиковское наречие. Древнеливвиковское наречие Олонецкого перешейка отразило в своем составе прибалтийско-финские архаичные черты, системно представленный древневепский субстрат, древнекарельские инновации и собственные диалектные особенности;
- 3) старокарельским (начало XVII — конец XIX в.), в ходе которого развитие диалектно-дифференцирующих явлений привело к вычленению в составе наречий диалектов и групп говоров;
- 4) новокарельским (с начала XX в. до настоящего времени), характеризующимся относительной стабильностью диалектной системы, объясняемой постепенным переходом на русский язык, остановившим процессы внутриязыкового взаимодействия между носителями диалектов.

Полученные в ходе исследования результаты имеют важное практическое значение в процессе становления нормированных вариантов карельского языка, число которых достигло к настоящему моменту шести (ливвиковский, михайловский людиковский, пряжинский людиковский, северный собственно карельский, тверской собственно карельский и основанный на говорах Приграничной Карелии южный собственно карельский). Они позволяют создать комплекс рекомендаций по редактированию их правил и норм, а также определить реально необходимое количество таких вариантов.

Приложение

Таблица (Распределение членов сопоставительных явлений карельской диалектной речи)

Приложение содержит данные о следующих группах говоров карельского языка: всг. — весьегонская, гмл.-смч. — гимольско-семчезерская, збц. — зубцовская, клв.-лх. — калевальско-лоухская, км.-блм. — кемско-беломорская, кнд. — кондушская, кнч.-ш. — кончезерско-шуйская, лжм.-мнз. — лижмозерско-мунозерская, мзр. — муезерская, мхл. — михайловская, олн. — олонецкая, прж. — пряжинская, прс.-грв. — поросозерско-гирвасская, с.лвв. — северноливвиковская, сгз. — сегозерская, смз.-ктк. — сямозерско-коткозерская, сн.-влз. — сонско-ведлозерская, тлз.-вдл. — тулмозерско-видлицкая, тлм. — толмачевская, тхв. — тихвинская. Разными цветами в строках залиты члены междиалектных соответствий.

В таблице использованы следующие сокращения: акт. — актив, безуд. — безударный слог, г.г. — гармония гласных, гл. — глухой, ед. — единственное число, з.р. — заднерядный вокализм, зауд. — заударный слог, зв. — звонкий, имп. — императив, имперф. — имперфект, инф. — инфинитив, к.с. — конец слова, конд. — кондиционал, л. — лицо, мн. — множественное число, н.с. — начало слова, одноосн. — одноосновный, п.р. — переднерядный вокализм, пасс. — пассив, пот. — потенциал, прич. — причастие, с.с. — середина слова, стяз. гл. — стяженный глагол, уд. — ударный слог, А — а, ä, С — любой согласный, D — d, t, O — o, ö, U — u, y, V — любой гласный.

Наречие Диалект	Собственно карельское						Ливвиковское				Людиковское								
	северный		южный		переходный		Ливвиковское				Людиковское								
Группа говоров	клев.- лх.	км.- блм.	тхв.	всг.	тлм.	збц.	стз.	гмл.- смч.	прс.- гров.	слвв.	смз.- ктк.	сн.- влз.	тлз.- вдл.	оли.	кнд.	прж.	лжм.- миз.	кнч.- ш.	мхл.
	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.	исконно люд.
1 *AA ул.	ua, ia	ua, ia	ua, ie / oa, öä	ia, ää	ua, ia	ua, ia	oa, cä / oo, ee	oa, cä / oa, cä	oa, cä	ua, ia	oa, cä / oa, ia	oa, ee	aa, ää	uo, ie / ua, ia	ia, ää	ua, ia	ua, ia	uo, ie / oa, cä	ua, ia
2 *ACA безул.	ua, ya	ua, ya	ua, ie / oa, öä	ia, ää	ua, ia	u, i	oo, ee / öö	ua, ia / oa, cä	oa, cä	ua, ia	oa, cä / oa, ia	oo, ee	aa, ää	uo, ie / ua, ia	ia, ää	ada, ää	ada, ää	ada, ää	ada, ää
3 *U/OCA безул.	uo, yä	uo, yä	ua, yä / uo, uö	uo, yä	uo, yä	uo, yä	ua, yä	uo, yä	uo, yä	uo, yä	uo, yä	uo, yä	ua, yä	uo, yä	ia, yä	Vda, Vda	Vda, Vda	Vda, Vda	Vda, Vda
4 *e/ICA безул.	ie	ie	ie	ie	ie	i	ia, ää	ie	ie	ie	ie, ei	ie, ei	ia, ää, ei	ie, ei	ie, ei	id	id	id	id
5 *VU	VU, VUv	VU, VUv	VV/VU, Vv	VU, VUv	VU, VUv	VU, Vv	VU, VUv	VU, VUv / Vv	Vu, VUv	VU, VUv	VU, Vv	VU, Vv	VU, Vv	VU, Vv	VU, Vv	VU, VUv	VU, VUv	VU, VUv	VU, VUv
6 *Vist/k ул.	Vi	Vi	Vi	Vi	Vi	Vi	Vi	Vi	Vi	Vi	Vi	V	V	V	V	V	V	V	V
7 *Vise заул.	V(2), V(3)	V(2), V(3)	V	V	V	V	V	V	V	Vi	Vi	V	Vi / V	Vi / V	Vi / V	Vi	Vi	Vi	Vv
8 *Oi к.с.	O	O	O	O	O	O	O	O	O	Oi	Oi	Oi	Oi	Oi	Oi	uoi, yoi	uoi, yoi	uoi, yoi	uoi, yoi
9 *A к.с.	a, ä	a, ä	a, ä	a, ä	a, ä	0	a, ä	a, ä	a, ä	u, y, 0	u, y, 0	u, y, 0	u, y, 0	u, y, 0	u, y, 0	e, 0	e, 0	e, 0	0
10 гл.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+/-
11 i ул.	i	i	i	i	i	i	y	i / y	y	y	y	y	y	y	y	y	y	y	y
12 e безул.	o	o	e / o	e	e / o	0	o	e / o	e	e	e	e	e	e	e	e	e	e	e
13 зв./гл.	гл.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.	зв.
14 *s н.с. з.р.	š	š	š	š	š	š	š	s	s	s	s	s	s	s	s	s	s	s	s
15 *si н.с.	š	š	š	š	š	š	š	š / š	š	s	s	s	s	s	s	š	š	š	s
16 *s н.с. п.р.	š	š	š	š	š	š	š	š / š	š	s	s	s	s	s	s	š	š	š	s
17 *is	s	z / ž	z	z	z	z	z	z / ž	z	z / ž	z / ž	z / ž	z / ž	z	z	z	z	z	z
18 *si з.р.	š	z / ž	ž	ž	ž	ž	ž	ž	ž	z	z	z	z	z	z	ž	ž	ž	z

Воканализм

Консонантизм

Наречие Диалект	Собственно карельское							Ливвиковское				Людиковское							
	северный		южный			переходный													
	клев.- лх.	км.- блм.	тхв.	всг.	тлм.	збл.	стз.	гмл.- смч.	прс.- гров.	с.лвв.	смз.- ктк.	сн.- влз.	тлз.- вдл.	оли.	кид.	прж.	лжм.- миз.	кнч.- ш.	млх.
Группа говоров																			
19 *si n.p.	s	z	z	z	z	z	z	z	z	z	z	z	z	z	z	z	z	z	z
20 *VšV n.p.	š	ž/z	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž
21 *VšV з.р.	š	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž	ž
22 *iks	iks	iks	iks	iks	iks	iks	iks	iks/iks	iks	iks	iks	iks	iks	iks	iks	iks	iks	iks	iks
23 *k/psi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi	Csi
24 *k/psV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV	CšV
25 *rnsV	CšV	CzV	CzV	CzV	CzV	CzV	CzV	CzV/CzV	CzV/CzV	CzV	CzV	CzV	CzV	CzV	CzV	CzV	CzV/CzV	CzV	CzV
26 *stki	sCi	sCi	sCi	sCi/sCi	sCi	sCi	sCi	sCi/sCi	sCi/sCi	sCi	sCi	sCi	sCi	sCi	sCi	sCi	sCi/sCi	sCi	sCi
27 *isk	isk	isk	isk	isk	isk	isk	isk	isk/isk	isk/isk	isk	isk	isk	isk	isk	isk	isk	isk	isk	isk
28 *ist, *is k.c.	ist	ist	ist	ist	ist	ist	ist	ist/ist	ist/ist	ist	ist	ist	ist	ist	ist	ist	ist	ist	ist
29 *Vš k.c.	š	š	š	š	š	š	š	š/s	š/s	š	š	š	š	š	š	š	š	š	š
30 *j n.c.	j	j	d'	j	j	j	j	d'	d'	j/d'	j/d'	j/d'	j	d'	d'	d'	d'	d'	d'
31 *hŋ	j	j	j	j	j	j	j	d'	d'	j/d'	j	j/d'	j	j	d'	d'	d'	d'	d'
32 *li, *mi	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i	li, n'i
33 *le/y/ä/ö n.c.	l	l	l'	l'	l'	l'	l'	l'	l'	l'	l'	l	l	l	l'	l'	l'	l'	l
34 *ne/y/ä/ö n.c.	n	n'	n'	n'	n'	n'	n'	nV	n'	n'	n'/n	n'/n	n'/n	n'	n'	n'	n'	n'	n
35 *ti n.c.	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i	t'i
36 *te/y/ä/ö	t	t	t'	t'	t'	t'	t'	tV	t'	t	t	t	t	t	t	t	t'/t	t'/t	t
37 *ni n.c.	ni	ni	ni	ni	ni/r'i	ni	ni	r'i	r'i	ni/r'i	ni	ni	ni	ni	ni	r'i	r'i	ni/r'i	ni
38 *ne/y/ä/ö	r	r	r	r	r'/r	r'/r	r'/r	r	r	r	r	r	r	r	r	r	r'/rV	r'/r	r
39 *lne/y/ä/ö c. c.	l, n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n	l', n
40 *nen c. c.	n'e	n'e	n'e	n'e	n'e/n'e	n'e	n'e	ne	ne	ne	ne	ne	ne	ne	ne	ne	n'e	ne	ne

Константин

Наречие Диалект	Собственно карельское						Ливвиковское						Людиковское						
	северный		южный		переходный								исконно люд.						
Группа говоров	клев.- лх.	км.- блм.	тхв.	всг.	тлм.	збл.	стз.	гмл.- смч.	прс.- грв.	с.лвв.	смз.- ктк.	си.- влз.	тлз.- вдл.	оли.	кнд.	прж.	лжм.- миз.	кнч.- ш.	мхл.
	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.
41 *и с. з. р.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.
42 *и с. с. п. р.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.	тх. дх.
43 *и п. с.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
44 *и к. с.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
45 *и к. с.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
46 *и п. м. р. м. с. с.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
47 *и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
48 *и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
49 *и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
50 *и в. в.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
51 *и к. т.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
52 *и к. т.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
53 *и к. т. о, *и к. т. у.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
54 *и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
55 *и к.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
56 *и п. м. р. м. с. с.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
57 *и к. *и т.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
58 *и т.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и
59 *и к.	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и	и

Консонантизм

Алфавитная соразличия

Наречие Диалект	Собственно карельское							Ливвиковское				Людиковское									
	северный		южный			переходный		Ливвиковское				Людиковское									
Группа говоров	кзв.- лх.	км.- блм.	мзр.	тхв.	всг.	тлм.	збл.	стз.	гмл.- смч.	прс.- гров.	с.лвв.	смз.- ктк.	си.- влз.	тлз.- вдл.	олн.	кнд.	прж.	лжм.- миз.	кнч.- ш.	мхл.	
	-(j)en	-(j)en	-(j)en	-(e)n	-n	-n	-n	-(e)n	-(j)en, -n	-(e)n	-(e)n	-(e)n	-(e)n	-(e)n	-(e)n	-(e)n	-(e)n	-den	-den	-den	
60 гегитив мн.	-V, -DA	-V, -DA	-V, -DA	-V, -DA	-V, -DA	-V, -DA	-V, -D	-V, -DA	-V, -DA	-V, -DA	-V, -DU	-V, -DU	-V, -DU	-V, -DU	-V, -DO	-V, -DO	-D, -De	-D, -De	-D, -De	-D, -De	
61 партияив ед.	-e, -jA, -DA	-e, -jA, -DA	-e, -jA, -DA	-e, -jA, -DA	-e, -jA, -DA	-e, -jA, -DA	-e, -DA, -DA	-e, -jA, -DA	-e, -jA, -DA	-e, -DA, -DA	-e, -DA, -DU	-e, -DA, -DU	-e, -DA, -DU	-e, -DA, -DU	-e, -DU, -DU	-e, -DU, -DU	-D	-D	-D	-D	
62 партияив мн.	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	-nA	
63 эсив	-ksi	-ksi	-ksi	-ksi	-ksi	-ksi	-ksi	-ksi	-ksi	-ksi	-kse	-kse	-kse	-kse	-kse	-kse	-kse	-kse	-ks	-ks	-ks
64 транслатив	-ssA, -jA	-ssA, -jA	-ssA, -jA	-ssA, -jA	-ssA, -jA	-ssA, -jA	-ss, -jA	-ssA, -jA	-ssA, -jA	-ssA, -jA	-ss, -jA	-ss, -jA	-ss, -jA	-ss, -jA	-ss, -jA	-ss, -jA	-ss, -jA	-ss, -jA	-ss, -jA	-ss, -jA	
65 инессив, адресив	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-st, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	-stA, -jA	
66 зыгитив, аблатив	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-jA	
67 азалатив	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	-h, -hV	
68 иллатив	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	
69 абессив	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	
70 комитатив	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	-neA	
71 аппроксиматив	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
72 терминатив	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
73 3 л. ед.	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	-U	
74 1, 2 л. мн. презенса	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmA, -tA	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmme, -tte	-mmme, -tte	-mmme, -tte	-mmme, -tte	
75 1, 2 л. мн. имперф.	-mA, -jA	-mA, -jA	-mA, -jA	-mA, -jA	-mA, -jA	-mA, -jA	-mA, -jA	-mA, -jA	-mA, -jA	-mA, -jA	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmO, -tO	-mmme, -tte	-mmme, -tte	-mmme, -tte	-mmme, -tte	

Именные словоизменения

Лозовые словоизменения

Наречие Диалект	Собственно карельское						Ливвиковское				Людиковское		
	северный		южный		переходный		с.л.в.	с.и.-в.л.	т.л.з.-в.д.л.	о.л.н.	к.н.д.	п.р.ж.	л.ж.м.-к.н.ч.-м.л.ш.
Группа говоров	к.л.в.-лх.	к.м.-б.л.м.	т.х.в.	в.сг.	т.л.м.	з.б.п.	с.г.з.	г.м.л.-с.м.ч.	п.р.-г.р.в.				
76 3 л. мл.	-h	-V/-h	-h	-V/-h	-h	-h	-V	-V/-h	-h	-h	-h	-h	-U, -v/-h/-j
77 3 л. мл. одноосн. глаголов на -е	-etAh	-ietAh	-etAh/-ietAh	-ietAh	-ietAh	-ietAh	-etAh	-etAh	-etAh	-ietAh	-ietAh	-etAh	-etAh
78 имперф. стиж. пл.	-si-	-(i)zi-	-i-	-i-	-i-	-i-	-(i)zi-	-(i)zi-	-i-	-i-	-i-	-(i)zi-	-(i)zi-
79 имперф.	-i-	-i-	-i/-ji-	-ji-	-i/-ji-	-ji-	-i-	-i-	-i-	-i-	-ji-	-i-	-i-
80 2 л. мл. импер.	-(k)kVV	-(k)kVV	-VV(i)V	-(k)kVV	-(k)kVV	-(k)kVV	-(k)kVV	-(k)kVV	-VV(i)V	-VV(i)V	-VV(i)V	-gat	-gat
81 ступень + kVV	сильн.	сильн.	сильн.	сильн.	сильн.	сильн.	слаб.	слаб.	сильн.	слаб.	сильн.	сильн.	сильн.
82 3 л. имп.	-kkAh	-	-	-kkAh	-	-kkAh	-kkAh	-kkAh	-hVs, -hVze	-	-	-	-
83 отрицательная форма имп.	elgVV	elgVV	elgVV	elgVV	elgVV	elgVV	elgVV	elgVV	älgVV	älgVV	älgVV	elgat	algat/elgat
84 конц.	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-	-zi-
85 3 л. ед. конц.	-is	-is	-is	-is	-is	-is	-is	-is	-s	-s	-s	-is	-s
86 имперф. конц.	аналит.	аналит.	аналит.	аналит.	аналит.	аналит.	аналит.	-nuz(i)-	-nuz(i)-	-nuz(i)-	-nuz(i)-	-nuz(i)-	-nuz(i)-
87 пот.	-nOU, -neh	-nOU, -neh	-nOU, -meh	-nOU, -meh	-nOU, -meh	-nOU, -meh	-nOU, -neh	-nOU, -neh	-nOU, -meh	-nOU, -meh	-nOU, -meh	-nOU, -meh	-nOU, -meh
88 I инф.	-hA	-hA	-hA	-hA	-hA	-hA	-hA	-hA	-hA	-hA	-dA	-dA	-dA
89 II инф.	-VV.	-VV.	-VV.	-VV.	-VV.	-VV.	-VV.	-VV.	-je-	-je-	-je-	-de	-de
90 I прич. акт.	-jA/-Va	-jA	-jA	-jA	-jA	-jA	-jA	-jA	-i	-i	-i	-i	-i
91 II прич. акт.	-nUn	-nUn	-nUn	-nUn	-nUn	-nUn	-nUn/-nUt	-nUn/-nUt	-nUh/-nUt	-nUh	-nU	-nU	-nU
92 II прич. пассив.	-DUO	-DUO	-DUO	-hUO	-hUO	-hUO	-DUO	-DUO/-hUO	-DUO/-hUO	-DUO	-hUt	-hUt	-DUt
93 возвратное страдание	-	-mÄcci	-	-	-	-	-	-	-(m)mos	-(m)mos	-(m)mos	-	- / -(m)mos

Литературное словоизменение

СПИСОК СОКРАЩЕНИЙ

1, 2, 3 — 1, 2, 3 лицо	IMPERF — имперфект
ливв. — ливвиковское наречие	INESS — инессив
люд. — людиковское наречие	INF — I инфинитив
с.к. — собственно карельское наречие	INF2 — II инфинитив
ABB — абессив	IPFV — имперфектив
ABL — аблатив	NEG — отрицательная форма
ACT — актив	PART — партитив
AD — адессив	PASS — пассив
ALL — аллатив	PFV — перфектив
COM — комитатив	PL — множественное число
COND — кондиционал	POT — потенциал
EL — элатив	PRES — презенс
ESS — эссив	PTCP — причастие
GEN — генитив	SG — единственное число
ILL — иллатив	TRANS — транслатив
IMPER — императив	

СПИСОК ИСТОЧНИКОВ

- ВепКар — *Открытый корпус вепсского и карельского языков (ВепКар)*. 2009–2022. Электронный ресурс: <http://dictorpus.krc.karelia.ru/ru>.
- Мещерский 1961 — Мещерский Н. А. Русско-карельские словарные записи XVI — начала XVIII в. *Прибалтийско-финское языкознание*, 1961, 23: 16–32.
- Муллонен, Панченко 2013 — Муллонен И. И., Панченко О. В. *Первый карельско-русский словарь и его автор афонский архимандрит Феофан*. Петрозаводск: ПетрГУ, 2013.
- Мызников 2010 — Мызников С. А. Карельско-вепские заговоры Олонецкого сборника. *Русские заговоры из рукописных источников XVII — первой половины XIX века*. Топорков А. Д. (сост.). М.: Индрик, 2010, 286–310.
- Перевод Евангелия от Марка — Перевод Евангелия от Марка на тверское наречие карельского языка 1820 г. *Тверские переводные памятники карельской письменности начала XIX века*. Громова Л. Г., Новак И. П. (ред.). Петрозаводск: КарНЦ РАН, 2020, 185–224.
- Перевод Евангелия от Матфея — Перевод Евангелия от Матфея на тверское наречие карельского языка 1820 г. *Тверские переводные памятники карельской письменности начала XIX века*. Громова Л. Г., Новак И. П. (ред.). Петрозаводск: КарНЦ РАН, 2020, 31–132.
- Перевод некоторых молитв — *Перевод некоторых молитв и сокращенного катехизиса на олонецком языке*. СПб.: Синодальная тип., 1804.
- Программы — *Программы по собиранию материала для диалектологического атласа карельского языка*. Научный архив КарНЦ РАН. Ф. 1. Оп. 38. № 16–170, 261–263, 267–268, 276; Оп. 5. № 79–94, 105–141, 161–166; Оп. 43. № 43–74, 116–132, 138, 140, 179–181, 195, 219–225, 228, 229, 274–275. Петрозаводск, 1937–1950.
- Савельева (ред.) 2023 — *Памятники лексикографии в «Цветнике» Прохора Коломятина 1668 г. Исследования и тексты*. Гл. 5: Карельско-русский словарь. Савельева Н. В. (ред.). М.; СПб.: Альянс-Архео, 2023, 397–417.
- Сравнительные словари — *Сравнительные словари всех языков и наречий, собранные десницею Всевысочайшей особы. Отделение первое, содержащее в себе европейские и азиатские языки*. Ч. 1–2. СПб.: Тип. у Шнора, 1787–1789.
- KKN II — *Karjalan kielen näytteitä II*. Leskinen E. (ed.). Helsinki: Suomalaisen Kirjallisuuden Seura, 1934.
- LingvoDoc — *LingvoDoc: лингвистическая платформа*. М.: ИЯз РАН; ИСП РАН, 2012–2024. Электронный ресурс: <https://lingvodoc.ispras.ru/>.
- Murreh — *Murreh: диалектная база карельского языка*. Петрозаводск: КарНЦ РАН, 2022. Электронный ресурс: <http://murreh.krc.karelia.ru>.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Беляков 1958 — Беляков А. А. Языковые явления, определяющие границы диалектов и говоров карельского языка в КарАССР. *Прибалтийско-финское языкознание*, 1958, 12: 49–62. [Belyakov A. A. Linguistic phenomena that define the boundaries of dialects and smaller varieties of the Karelian language in the Karelian ASSR. *Pribaltiisko-finskoe yazykoznanie*, 1958, 12: 49–62.]
- Бромлей 2010 — Бромлей С. В. *Проблемы диалектологии, лингвогеографии и истории русского языка*. М.: Азбуковник, 2010. [Bromlei S. V. *Problemy dialektologii, lingvogeografii i istorii russkogo yazyka* [Issues in dialectology, linguistic geography, and history of the Russian language]. Moscow: Azbukovnik, 2010.]
- Бубрих 1950 — Бубрих Д. В. Не достаточно ли емских теорий? *Известия Карело-Финской научно-исследовательской базы Академии наук СССР*, 1950, 1: 80–92. [Bubrikh D. V. Aren't yem's theories enough? *Izvestiya Karelo-Finskoi nauchno-issledovatel'skoi bazy Akademii nauk SSSR*, 1950, 1: 80–92.]
- Вопросы теории — Аванесов Р. И. (ред.). *Вопросы теории лингвистической географии*. М.: АН СССР, 1962. [Avanesov R. I. (ed.). *Voprosy teorii lingvisticheskoi geografii* [Issues in linguistic geography]. Moscow: Academy of Sciences of the USSR, 1962.]
- ДАКЯ — Бубрих Д. В., Беляков А. А., Пунжина А. В. *Диалектологический атлас карельского языка*. Helsinki: Suomalais-Ugrilainen Seura, 1997. [Bubrikh D. V., Belyakov A. A., Punzhina A. V. *Dialektologicheskii atlas karel'skogo yazyka* [Dialectological atlas of the Karelian language]. Helsinki: Suomalais-Ugrilainen Seura, 1997.]
- История Карелии 2001 — Кораблев Н. А. (ред.). *История Карелии с древнейших времен до наших дней*. Петрозаводск: Периодика, 2001. [Korablev N. A. (ed.). *Istoriya Karelii s drevneishikh vremen do nashikh dnei* [History of Karelia from ancient times to the present day]. Petrozavodsk: Periodika, 2001.]
- Итоги — *Итоги Всероссийской переписи населения 2020 г.* Т. 5: *Национальный состав и владение языками*. https://rosstat.gov.ru/vpn/2020/Tom5_Natsionalnyj_sostav_i_vladenie_yazykami.
- Новак 2019 — Новак И. П. Карельский язык и его диалекты. *Народы Карелии: историко-этнографические очерки*. Винокурова И. Ю. (отв. ред.). Петрозаводск: Периодика, 2019, 56–65. [Novak I. P. Karel'skii yazyk i ego dialekty [Karelian language and its dialects]. *Narody Karelii: istoriko-etnograficheskie ocherki*. Vinogradova I. Yu. (ed.). Petrozavodsk: Periodika, 2019, 56–65.]
- Новак 2022 — Новак И. П. Проблемы диалектной классификации карельского языка. *Ежегодник финно-угорских исследований*, 2022, 2: 204–213. [Novak I. P. Problems of dialectal classification of the Karelian language. *Yearbook of Finno-Ugric Studies*, 2022, 2: 204–213.]
- Новак, Крижановская 2022 — Новак И. П., Крижановская Н. Б. Система восходящих дифтонгов в говорах карельского языка Карелии: в поисках алгоритма кластеризации. *Вестник угроведения*, 2022, 3: 486–496. [Novak I. P., Krizhanovskaya N. B. The system of ascending diphthongs in dialects of the Karelian language: Comparison of clustering methods. *Bulletin of Ugric Studies*, 2022, 3: 486–496.]
- Пшеничникова 2008 — Пшеничникова Н. Н. *Лингвистическая география*. М.: Азбуковник, 2008. [Pshenichnova N. N. *Lingvisticheskaya geografiya* [Linguistic geography]. Moscow: Azbukovnik, 2008.]
- Сюрюн и др. 2021 — Сюрюн А. А., Давидюк Т. И., Евстигнеева А. П. *Уточнение списочного состава языков России*. М.: ИЯз РАН, 2021. [Syuryun A. A., Davidyuk T. I., Evstigneeva A. P. *Utochnenie spisochnogo sostava yazykov Rossii* [A refined list of languages of Russia]. Moscow: Institute of Linguistics of the Russian Academy of Sciences, 2021.]
- Honkola et al. 2019 — Honkola T., Santaharju J., Syrjänen K., Pajusalu K. Clustering lexical variation of Finnic languages based on Atlas Linguarum Fennicarum. *Linguistica Uralica*, 2019, 3: 161–184.
- Itkonen 1971 — Itkonen T. Aunuksen äänneopin erikoispiirteet ja aunukselaismurteiden synty [Phonetic features of the Olonets dialect of the Karelian language and the composition of Olonets dialects]. *Virittäjä*, 1971, 75: 153–185.
- Leskinen 1998 — Leskinen H. Karjala ja karjalaiset kielentutkimuksen näkökulmasta [Karelia and the Karelians from the point of view of linguistics]. *Karjala: historia, kansa, kulttuuri*. Nevalainen P., Sihvo H. (eds.). Helsinki: Suomalaisen Kirjallisuuden Seura, 1998, 352–382.
- Pahomov 2017 — Pahomov M. *Lyyläiläiskysymys: Kansa vai heimo, kieli vai murre?* [The Ludic question: Nation or tribe, language or dialect?] Helsinki: Helsingin yliopisto, 2017.
- Turunen 1946 — Turunen A. *Lyyläiläismurteiden äännehistoria I. Konsonantit* [Historical phonetics of Ludic dialects I. Consonants]. Helsinki: Suomalais-Ugrilainen Seura, 1946.

- Turunen 1950 — Turunen A. *Lytydiläismurteiden äännehistoria II. Vokaalit* [Historical phonetics of Lude dialects I. Vowels]. Helsinki: Suomalais-Ugrilainen Seura, 1950.
- Virtaranta 1972 — Virtaranta P. Die Dialekte des Karelischen. *Советское финно-угроведение*, 1972, 8: 7–27. [Virtaranta P. Die Dialekte des Karelischen. *Sovetskoe finno-ugrovedenie*, 1972, 8: 7–27.]
- Virtaranta 1985 — Virtaranta P. Kriterien zur Klassifizierung der Dialekte des Karelischen. *Dialectologia Uralica*, 1985, 20: 117–137.
- Wieling, Nerbonne 2015 — Wieling M., Nerbonne J. Advances in dialectometry. *Annual Review of Linguistics*, 2015, 1: 243–264.
- Wiik 2004a — Wiik K. *Karjalan kielen murteet. Kvantitatiivinen tutkimus* [Dialects of the Karelian language. A quantitative study]. Turku: Kalevi Wiik, 2004.
- Wiik 2004b — Wiik K. *Suomen murteet. Kvantitatiivinen tutkimus* [Dialects of the Finnish language. A quantitative study]. Helsinki: Suomalaisen Kirjallisuuden Seura, 2004.
- Zaikov 2017 — Zaikov P. M. *Karjalan kielen murteet* [Dialects of the Karelian language]. Petrozavodsk: Petrozavodsk State Univ., 2017.

Получено / received 06.07.2024

Принято / accepted 24.09.2024

Средневековые названия погостов Корельского уезда: на границе государств и языков

© 2024

Валерий Леонидович Васильев

Институт русского языка им. В. В. Виноградова РАН, Москва, Россия; vihnn@mail.ru

Ирма Ивановна МуллоненИнститут языка, литературы и истории Карельского научного центра РАН, Петрозаводск,
Россия; Уральский федеральный университет, Екатеринбург, Россия;
irma.mullonen@hotmail.com

Аннотация: В статье всесторонне анализируются происхождение и история названий средневековых новгородских погостов, прилегавших к северным берегам Ладожского озера. Имена погостов рассматриваются в географической последовательности — с севера на юг: *Ильинский Иломанский, Никольский Сердовольский, Воскресенский Соломянский, Богородицкий Кирьяжский, Воскресенский Городенский, Михайловский Сакульский* и *Васильевский Ровдужский*. Авторы сосредоточены преимущественно на анализе вторых прилагательных в приведенных парах. В основном это образования от карельских названий тех селений, которые новгородцы сделали центрами своих погостов и церковных приходов. Название *Городенского* погоста отличается от остальных исконным, древнерусским происхождением (< *городь*, *городькъ* ‘укрепленное место’), оно относилось только к погосту-округу. В средневековой документации (русской, шведской, финской) топонимы имеют множество вариантных форм, что связано со спецификой языковой адаптации и письменной передачи. Большинство первых письменных упоминаний о погостах относятся к 1500 г., кроме Кирьяжского и Сердовольского погостов, свидетельства о которых известны с XIV–XV вв. Некоторые наименования погостов (*Кирьяжский, Ровдужский, Сердовольский*) в фонетическом развитии ушли весьма далеко от реконструируемых карельских топонимов-прототипов: **Kurgijogi, *Raudu, *Sordavala*. Древнерусская адаптация этих карельских топонимов, помимо регулярных фонетических закономерностей, включала вторичное сближение с созвучными лексемами и антропонимами. Данный тезис можно применить также к наименованию *Соломянского* погоста, которое возникло путем замещения карельского **Salmi* равнозначным древненовгородским *Соломя* (= *соломя* ‘пролив’). Исходные карельские топонимы трактуются по-разному. В одних воспроизводится обозначение природного объекта, при котором стояло карельское селение, ставшее погостским центром. Таковы названия погостов *Кирьяжский* (< **Kurgijogi* ‘журавлиная река’), *Соломянский* (< *Salmi* ‘пролив’) и *Иломанский* (*Ilomantsi* < прасаам. **ēlēmāñce* ‘самый верхний’). Другие — *Сакульский* (*Sakkula*), *Ровдужский* (**Raudu*), *Сердовольский* (**Sordavala*) — содержат в основе антропоним, который должен квалифицироваться как патроним, личное имя или прозвище первопоселенца.

Ключевые слова: антропонимы, карельский язык, Новгород, севернорусские диалекты, старорусский язык, погосты, прибалтийско-финские языки, скандинавские языки, топонимия, финский язык, шведский язык

Благодарности: Исследование выполнено при поддержке гранта РНФ № 23-18-00439 «Ономастикон и лингвокультурная история Европейской России».

Для цитирования: Васильев В. Л., Муллонен И. И. Средневековые названия погостов Корельского уезда: на границе государств и языков. *Вопросы языкознания*, 2024, 6: 85–104.

DOI: 10.31857/0373-658X.2024.6.85-104

Medieval names of churchyards in Korelsky county: On the border of states and languages

Valery L. Vasilyev

Vinogradov Russian Language Institute, Russian Academy of Sciences, Moscow, Russia;
vihnn@mail.ru

Irma I. Mullonen

Institute of Linguistics, Literature and History, Karelian Research Centre
of the Russian Academy of Sciences, Petrozavodsk, Russia;
Ural Federal University, Ekaterinburg, Russia; irma.mullonen@hotmail.com

Abstract: The article comprehensively analyzes the origin and history of the names of medieval Novgorod churchyards (pogosts) adjacent to the northern shores of Lake Ladoga. The churchyard names are considered in geographical sequence — from north to south: *Il'inskij Ilomanskij, Nikol'skij Serdovol'skij, Voskresenskij Solomjanskij, Bogorodickij Kir'jažskij, Voskresenskij Gorodenskij, Mixajlovskij Sakul'skij* and *Vasil'evskij Rovdužskij*. The authors focus primarily on the analysis of the second adjectives in the given pairs. These are mainly derivatives of the Karelian names of those villages that the Novgorodians made the centers of their churchyards and church parishes. The name of the *Gorodenskij* pogost differs from the others in its vernacular, ancient Russian origin (< *gorodъ, gorodъkъ* 'fortified place'), it referred only to the territorial district. In medieval documentation (Russian, Swedish, Finnish), place names have many variant forms due to the peculiarities of language adaptation and writing systems. Most of the first written mentions of churchyards date back to 1500, except for the *Kir'jažskij* and *Serdovol'skij* pogosts, evidence of which goes back to the 14th–15th centuries. Some churchyard names (*Kir'jažskij, Rovdužskij, Serdovol'skij*) in their phonetic development went very far from the reconstructed Karelian prototype place names: **Kurgijogi*, **Raudu*, **Sordavala*. The Old Russian adaptation of these Karelian toponyms, in addition to regular phonetic correspondences, included a secondary convergence with assonant lexemes and anthroponyms. This thesis can also be applied to the name *Solomjanskij* pogost, which arose by replacing the Karelian **Salmi* with the equivalent Old Novgorod *Solomja* (= *solomja* 'strait'). The originally Karelian place names are interpreted differently. Some reproduce the designation of a natural site near which a Karelian village that became a territorial center was located. These are the churchyard names: *Kir'jažskij* (< **Kurgijogi* 'crane river'), *Solomjanskij* (< *Salmi* 'strait') and *Ilomanskij* churchyards (*Ilomantsi* < Proto-Saami **elēmānce* 'topmost'). Others — *Sakul'skij* (*Sakkula*), *Rovdužskij* (**Raudu*), *Serdovol'skij* (**Sordavala*) — contain an anthroponym which should qualify as a patronymic, personal name, or nickname of the first settler.

Keywords: anthroponymy, Finnic, Finnish, Karelian, Middle Russian, North Russian, Novgorod, pogosts, Scandinavian, Swedish, toponymy

Acknowledgements: The research was supported by the Russian Science Foundation (grant project No. 23-18-00439 "Onomasticon and Linguocultural History of European Russia").

For citation: Vasilyev V. L., Mullonen I. I. Medieval names of churchyards in Korelsky county: On the border of states and languages. *Voprosy Jazykoznanija*, 2024, 6: 85–104.

DOI: 10.31857/0373-658X.2024.6.85-104

Северо-Западное Приладожье, где на рубеже I–II тыс. н. э. происходило формирование карельского этноса, было на протяжении многих веков территорией шведско-русского противостояния и неоднократно переходило из рук в руки. На разных этапах истории оно входило в состав земель Великого Новгорода, Московского княжества, Шведского королевства, Финляндской республики, Российского государства. Нередко такой переход сопровождался частичной или полной сменой населения, как это было, например, в начале XVII в., когда территория отошла к Швеции и сюда начало стекаться финское население, или после Второй мировой войны, когда финское и карельское население покинуло территорию полностью.

Соответственно, местная карельская топонимия находила в разные годы отражение в русских, шведских и финских источниках, каждый из которых адаптировал иноязычное для него карельское название к своему языку и топонимической системе. Это нередко затухивало исходный облик названий. Сопоставление разноязычных источников позволяет при учете особенностей адаптационных процессов того или иного языка восстановить в ряде случаев первоначальные исконные карельские топонимы и предложить их этимологии, а также проследить модели интеграции карельских топонимов в русскую систему именований.

Первая относительно массовая фиксация топонимии (преимущественно ойконимии) на территории Северо-Западного Приладожья состоялась в самом конце XV в. благодаря повсеместному описанию новыми московскими властями земельных владений Водской пятины. Данная топонимия представлена в новгородской писцовой книге 1500 г. [ПОКВП]. В эпоху первых московских описаний территория Северо-Западного Приладожья административно составляла Корельский уезд с центром в крепости Корела (см. рис.). Уезду предшествовала более обширная Корельская земля (Корела), которая с XII в. поддерживала с Новгородом отношения союзнической конфедерации, а затем вассалитета. По условиям Ореховецкого мирного договора 1323 г. часть Корельской земли досталась Швеции, а другая, приладожская часть напрямую вошла в подчинение Новгороду. Из приладожской части, подразделенной на Переднюю Корелу и Заднюю Корелу (соответственно — ближнюю и дальнюю относительно Новгорода), после присоединения Новгородской республики к Москве окончательно выделился Корельский уезд, включавший, согласно ПОКВП, семь погостов-округов: Воскресенский Городенский, Михайловский Сакульский, Васильевский Ровдужский, Богородицкий Кирьяжский, Никольский Сердобольский, Воскресенский Соломянский и Ильинский Иломанский.

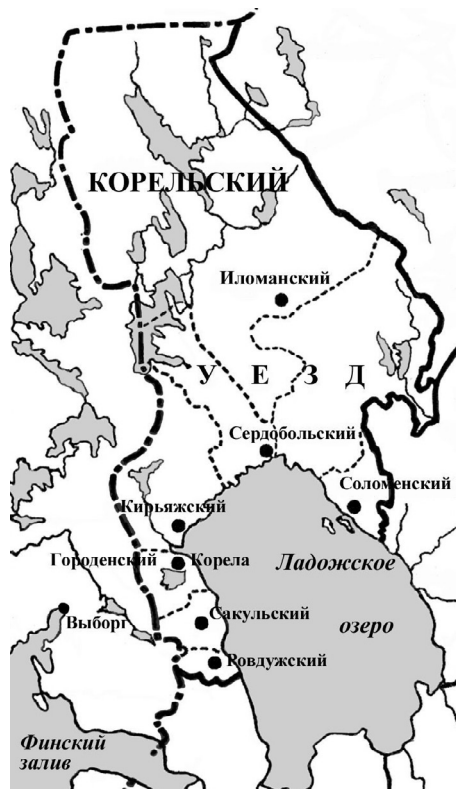


Рис. Карта Корельского уезда XVI в.

Название уезда отсылает к этнониму — племенному названию *корела* (др.-русс. *корѣла*) — карел. *karjala, kariela*. Этноним, будучи трехсложным, явно выбивается из ряда двухсложных именовании прибалтийско-финских народов. Происхождение его дискуссионно. Среди многих гипотез, собранных в [Grünthal 1997: 73–96], наиболее активно обсуждается условно германская теория, в соответствии с которой в основе лежит германское заимствование *karja*, но не в современном значении ‘стадо’, а в утраченном — ‘группа людей, вооруженный отряд’, которое присуще источнику заимствования в германских языках: прагерм. **χarja* (гот. **harjis*, др.-норв. **herr*, др.-швед. **haer*). Параллельно обсуждается возможность балтских корней, ср. др.-прус. *karja* ‘воинский отряд, подразделение’ [SSA, 1: 313–314]. При этом истоки конечного *-la* рассматриваются как результат абстрагирования или переразложения, произошедшего на базе прилагательного *karjalainen* (с суффиксом *-lainen*), букв. ‘принадлежащий к народу *karja*’ → *karjala* (см., например, [Nissilä 1962b: 365]), что, однако, не укладывается в ряд многочисленных образований подобного рода (*suomalainen* ‘финн’, *hämäläinen* ‘принадлежащий к племени емь’, *savolainen* ‘житель провинции Саво (Savo)’ и др.), ни одно из которых не пошло по пути подобного переразложения. Более вероятно предположение о том, что концовка *-la* в этнониме — это прибалтийско-финский суффикс *-la* с локативной семантикой, т. е. первоначально *Karjala* — место расселения той самой «группы» или того самого «отряда» *karja*, распространившееся со временем на более обширную территорию и население, заселявшее ее. При этом предполагается, что этот первоначальный центр располагался в устье р. Вуоксы, служившей на рубеже I–II тыс. н. э. оживленным торговым путем из Финского залива в Ладожское озеро, т. е. фактически являлся предшественником города-крепости Корела, см., например, [Kirkinen 1981: 12; Grünthal 1997: 73–83]. Версия отойконимного происхождения этнонима при всех сопутствующих вопросах представляется перспективной для дальнейшего исследования.

Далее с привлечением русских и шведских документов обратимся к рассмотрению наименований всех семи погостов Корельского уезда, взятых в географической последовательности — с севера на юг. В области пятинного деления Великого Новгорода для разных пятин традиционно использовались разные модели именования погостов. Модель номинации, типовая для Водской пятины, к которой относился Корельский уезд, представляет собой структуру, состоящую из производных адъективов-приложений на *-ский/-ской* (Васильевский Ровдужский погост, Михайловский Сакульский погост и др.). На первое место ставилось прилагательное, отсылавшее к наименованию приходской погостской церкви, на второе место — прилагательное с топографическим признаком указания на исконное собственное имя селения (иногда — имя реки, водоема), при котором был обустроен погостский центр. Церковный признак, считавшийся приоритетным и регулярным при описании погостов-округов, часто не указывался при описании погостов-мест, где допускались разные варианты номинации. Такая схема описаний устойчиво воспроизводится в письменной документации на протяжении нескольких столетий вплоть до упразднения системы погостов в конце XVIII в. [Васильев, Вихрова 2014: 73–74]. В настоящей статье мы сосредоточимся главным образом на ранней письменной истории и языковой эволюции наименований погостских центров Корельского уезда, включая этимологии стоящих за ними прибалтийско-финских топонимов.

Самым северным погостом бывшего Корельского уезда являлся **Ильинский Иломанский**. На заре своей истории этот погост прошел через те же этапы русско-шведского противостояния, что и более южные погосты. Но, в отличие от последних, после Второй мировой войны центральное селение погоста осталось в составе Финляндии и сейчас административно принадлежит финской провинции Северная Карелия.

Новгородская писцовая книга 1500 г., описывая селения, принадлежавшие погосту *Ильинской Иломанской* с церковью Ильи Пророка, сообщает также о д. *Иломанец у Ильи святого у погоста* [ПОКВП: 170, 174]. Надо полагать, производное наименование погоста образовалось при помощи форманта *-ской* от топонима *Иломанец*. В шведской,

финской и русской документации более позднего времени обнаруживаются записи *Ilmans Sochn* [ПККУ 1590: 270], *Ilomantsi Pogost* [ПККУ 1631: 506], с середины XVII в. — *Ilomans Pågosth*, *Ilomantsi*, *Иломанской погост* [ПККЛ 1637: 264, 438, 631]. По сообщению К. А. Неволлина [1853: 130], на картах 1-й половины XIX в. центром погоста указан кирхшпил *Иломаитс*, иначе *Иломанц* на одном из небольших озер в Куопиоской губернии Великого княжества Финляндского.

Топоним *Иломантси* (совр. фин. *Ilomantsi*) оставался загадкой вплоть до недавнего времени. Народная этимология связывала его с наименованием погостской церкви Ильи Пророка, на такую возможность осторожно указывал и зачинатель изучения карельской топонимии В. Ниссиля [Nissilä 1975: 200]. В начале 2000-х гг. убедительная расшифровка топонима *Ilomantsi* была предложена финским исследователем А. Айкио. Он возвел его к прасаамскому прилагательному **elēmāñce* (совр. саам. *alimuš*) ‘самый верхний’, в котором праязыковая основа **elē* ‘верхний’ выступает в связке с показателем суперлатива **-māñce* [Aikio 2003: 101–102]. Будучи усвоенным в карельское бытование, оно приобрело в нем закономерный вид *Ilomančči* [Иломанччи], соответственно, фин. *Иломантси*. Первоначально название относилось к озеру (совр. *Ilomantsinjärvi*), при котором возникло селение и которое действительно является верхним в цепочке озер, стекающих в систему р. Вуоксы.

Этимология А. Айкио поддерживается не только фактором топографии, но и тем обстоятельством, что топоним не является уникальным, такая мотивационная модель повторяется неоднократно на Северо-Западе России. Своеобразными «тезками» *Иломантси* выступают имена рек *Ильмаж*, *Ильмеж*, *Ильмеш*, *Ильмаза* на Русском Севере, озер *El'muz*, *Elmüsjärvi*, *Elmüsjärvi* и рек *Илеменз* и *Ilmetjoki* (< **Ilmesjoki*) в Карелии. Аналогичная модель бытует и в прибалтийско-финской гидронимии. Она представлена озерными именованиями *Ylimäini* ~ *Ylimäisjärvi* в Карелии, *Ylimmäinen* в Финляндии: *yli* ‘верхний’ + показатель суперлатива *-mäinen* (*-mäini*), который в формообразовательной парадигме приобретает вид *-mäise*. Видимо, в этом же ряду стоит и название оз. *Ильмень* < **Ylimäine* [Муллонен 2018: 12–17].

Никольский Сердоловский погост с центром в селении Сердолово, совр. гор. Сортавала, располагался на северо-восточном берегу Ладожского озера. В русских источниках впервые его упоминает писцовая книга 1500 г.: *погост Никольской Сердоловской*, там же отмечена *волость Спаская Валамского монастыря въ Сердолови* [ПОКВП: 143, 166]. В XVII в. наряду с *Сердоловом* появляется вариант *Сердоболь* [Левашов 1988: 75], подписанный единственным на картах 1-й половины XIX в. [Неволин 1853: 130]. В шведских источниках название имело вид *Sordawala* 1554 г., *Sordawala* 1581 г., *Sordabolschi eller Micholski Sochn* 1589 г., *Sordowala* 1648 г., *Sordawala* 1767 г. [Lehikoinen 2018: 36]. Первое упоминание — в шведском документе о нарушении русско-шведского пограничного соглашения — относится к 1468 г. В нем топоним представлен в трех разных написаниях: *Sartwal*, *Sortewala* и *Sortawala* [Hausen 1924: 308–309]. Они воспроизводят освоенную уже финнами форму, близкую или практически совпадающую с современной. Это и понятно: шведы, владевшие в тот период этой территорией, просто использовали в своей части договора финское написание карельского топонима (новгородский оригинал договора, к сожалению, не сохранился). Финнизированный вариант устоялся и продолжает употребляться в совр. *Сортавала*, как правило, со смещением на третий слог ударением.

Принципиальный вопрос, возникающий в связи с поисками этимологических истоков названия погоста, — как соотносятся между собой совр. *Сортавала* (фин. *Sortavala*) и устаревшее *Сердолово/Сердоболь*? Считается, что они имеют единые истоки, но мнения о том, какой из вариантов первичен, разнятся: от примарности *Sortavala* до осторожного признания первичности русской формы *Сердоболь* [Левашов 1988: 75]. Между тем акценты должны быть смещены: какое из названий — финское или русское — адекватнее отражает карельский оригинал? Мысль о первичности русск. *Сердоболь* не вписывается в общую схему номинации приладожских погостов Водской пятины, производные наименования которых (за исключением *Городенского*, см. ниже) отсылают к прибалтийско-финскому

названию места или местности. Более того, на основе финской орфографической записи хорошо реконструируется карельский исходный ойконим, который в соответствии с древнекарельской фонетикой имел вид **Sordavala* и соответствовал типологически средневековым карельским ойконимам Приладожья.

Вообще происхождение топонима *Sortavala* < карел. **Sordavala* неоднократно привлекало к себе внимание, при этом спектр интерпретаций широк. Среди них встречаются сугубо народноэтимологические сближения основы *Sort-* с русск. *черп*¹. В научном дискурсе² обсуждалось возведение топонима к финскому причастию *sortawa* ‘рассекающий’ с топографическим обоснованием: город стоит на берегу залива, глубоко вдающегося в берег Ладожского озера, рассекающего его [Грот 1899: 198]. Значительно более обоснован отантропонимный подход, тем более что формант *-la* является маркером ойконимов, который к тому же всегда в традиционной прибалтийско-финской топонимии функционировал в связке с антропонимом. В случае с *Сортавалой* это древнее прибалтийско-финское имя или прозвище *Sortava*, выраженное формой причастия от глагола *sortaa* ‘ломать, крушить, валить (например, лес)’, но ему присуща и более абстрактная семантика ‘угнетать, тиранить, подавлять, притеснять’. Подобные причастные именованья были обычны для древнекарельского именослова, например *Tietävä* ‘знающий’, *Kuuleva* ‘слышащий’, *Horjuva* ‘болтающийся, шатающийся’, *Ottava* ‘берущий’, *Vihava* ‘злобствующий’ и др. [Nissilä 1975: 123]. В качестве мотива выдвигалась версия о том, что имя имело семантику ‘рубящий подсеку’ [Ibid.], хотя с точки зрения типологии прозвищных именованья, как правило, негативных по содержанию, не менее, а может, даже более обоснованно исходить из значения ‘тиран, угнетатель’. При этом реальное бытование антропонима в прошлом подтверждается фамилиями *Sordava* и *Sordavainen* (*per sortaua*, *Joan Sordauainn*, *Oloff Sordauainen*, *Kauppi Sordauann*), присутствие которых засвидетельствовано документами XVI в. в общине *Uusikirkko* (совр. Поляны Выборгского района Ленинградской области), в которую, кстати, входила до Второй мировой войны и д. *Sortavala*, полная «тезка» приладожской Сортавалы [Lehikoinen 2018: 37].

Косвенной поддержкой отантропонимной этимологии следует считать и ту ойконимическую среду, неотъемлемой частью которой является название *Сортавала*. Книга Водской пятины 1500 г. [ПОКВП] упоминает в ближайших окрестностях этого города десятки названий деревень, оформленных концевкой *-la*, которая присоединяется к основе, выраженной именем, как правило, некалендарным, или прозвищем (возможно, родовым). Большинство из этих ойконимов, кстати, дожили, несмотря на все перипетии истории, до сегодняшнего дня: дд. *Гелюля*, *Гимбола*, *Кармола*, *Пуйкала*, *Рекала*, *Тулола*, *Марила* из ПОКВП на современной карте Сортавалы обозначены как *Хелюля*, *Хюмпеля*, *Кармала*, *Пуйккола*, *Риеккала*, *Тулола*, *Марила*. Название гор. *Сортавала* самым естественным образом встраивается в этот ряд.

Таким образом, топоним *Сортавала* (< карел. **Sordavala*) и структурно, и семантически входит в число традиционных карельских ойконимов.

Какова в этом контексте позиция названия *Сердоволь*? Оно не может быть объяснено из лексических ресурсов русского языка и в целом — славянских языков. Вместе с тем *Сортавала* непосредственно не определяет появление *Сердоволь*, т. е. русский вариант

¹ В соответствии с примитивно-этимологическим осмыслением, *Сортавала* — это «место, где установилась власть (фин. *valta* ‘власть’) черта, изгнанного монахами с Валаама при освящении острова». В иной версии этого толкования топоним выводили из заклинания монахов «*Черт, вали!»*, произносившегося при изгнании.

² Вне научного обсуждения находится абсолютно надуманная интерпретация названия *Сортавала* как сложного слова, состоящего из двух элементов (фин. *sortua* ‘падать, рушиться’ и имени языческого бога *Ваала*) и связанного якобы с уничтожением места языческого культа [Михайлова 2016]. Она противоречит номинативным принципам и типологии прибалтийско-финской ойконимии.

не модифицирован прямо из финского. Как *Сортавала* (фин. *Sortavala*) — это финнизированная форма, так и *Сердоволь* — русифицированная форма исконно карельского топонима **Sordavala*, появившегося на ранней территории проживания карельского народа и первоначально не знакомого ни собственно финнам, ни новгородцам.

Средневековые новгородцы вряд ли познакомились с карел. *Sordavala* до учреждения в этом карельском селении Никольского Сердовольского погоста. Это произошло не ранее XIV в., который отмечен созданием в Северном Приладожье сети церковных приходов и началом широкого распространения православия среди местного карельского населения [Очерки 2001: 280]. Предпочтительнее вести речь о 2-й половине XIV в., опираясь на дату основания ближайшего к Сердовольскому погосту монастыря на острове Валаам, появление которого относят обычно к концу XIV — началу XV в. [Там же: 276, 280]. Предполагаемая эпоха знакомства новгородцев с карел. *Sordavala* вполне отвечает особенностям русской адаптации этого названия. По крайней мере в XIV в. уже деактуализовались древнерусские (конкретнее — древненовгородские) закономерности, под действие которых мог попасть карельский топоним. В частности, сегмент *Sorda-* уже не мог преобразоваться в *Сорода-*, поскольку осталось в прошлом действие древнерусского полногласия, не действовал в указанную эпоху и переход приб.-фин. *v* > др.-новг. *б*, хотя приб.-фин. *a*, вероятно, еще давало др.-русс. *о*. Разумеется, не могло быть никакого, даже случайного, смешения гласного *о* в **Sordavala* с др.-русс. гласными *ѣ* или *ь*, поскольку в указанную эпоху редуцированные уже исчезли.

Глубокую фонетическую переработку карел. **Sordavala* в русск. *Сердоболь* нельзя объяснить ничем иным, кроме устно-разговорного переосмысления «темного» иноязычного названия по облику понятного созвучного русского слова. Вообще говоря, подобного рода народная этимологизация в истории топонимии (как, впрочем, и в истории лексики) — частое, но не всегда отслеживаемое явление. Отдельные факты хорошо известны, например, древнерусское имя шведского гор. *Стокгольм*, усвоенное как *Стекольна* по связи со *стекло*, *стекольный*, или гор. *Выборг* (< швед. *Viborg*), которое порой изменялось в др.-русс. *Выборъ*, ст.-русс. *Выбор* по народноэтимологической связи с *выборъ* [Ан. РЭС, 9: 124].

Еще М. Фасмер писал о том, что название *Сердоболь* появилось благодаря сближению фин. *Sortavala* с русск. *сердобольный* [Фасмер, III: 605]. Эту важную мысль Фасмера следует уточнить и подкрепить некоторыми дополнительными соображениями. Как выше уже говорилось, нужно исходить не из финнизированной формы, а из карел. **Sordavala*, давшего старое русск. *Сердоволь* (а не *Сердоболь*). При переосмыслении карельский топоним воспринимался русскими как некое сложное слово, в котором сегмент *Sorda-* был переименован по облику русск. *сердо-* (: *сердце*) некоторых сложений. В самом деле, компонент *сердо-* уже оказывался результатом вторичных, народноэтимологических преобразований. Отличной иллюстрацией аналогичного процесса выступает русск. *сердолик*, др.-русс. *сердоликъ* (*середоликъ*) — обозначение полудрагоценного камня, который в древнерусских церковнославянских текстах известен в вариантах *сардониксъ*, *сардоникши*, *сардионъ*, *сардий*, см. [СлРЯ XI–XVII, 23: 64–65; 24: 78] (< греч. *σαρδόνιξ*, *σάρδιον* ‘сардский камень’)³. Таким образом, изменение сегмента *Sorda-* > *Сердо-* в карельском топониме осуществилось по образцу уже состоявшегося древнерусского незакономерного изменения *сядо-* > *сердо-*, пережитого обозначением полудрагоценного камня⁴. Данное изменение древнерусской лексемы произошло не позднее 1-й половины XIV в. и поначалу затронуло только первую часть лексемы; ср. производное *сердониный* ‘сделанный

³ В данном случае переосмысление произошло под влиянием «красного, а точнее, мясного цвета камня, символически связанного с сердцем» [Кучко 2023: 258].

⁴ В этой связи обращает на себя внимание любопытный термин *сердоболит* ‘горная порода [какая?]. Петерб., 1865’ [СРНГ, 37: 188], сближающий название гор. *Сердоболь* с обозначением некоего минерала (быть может, как раз с *сердолик*?).

из сердолика⁵ в грамоте около 1339 г. [СлРЯ XI–XVII, 24: 78] (< **сердоникъ* < *сардониксъ*). Целиком измененные *сердолик*, производное *сердоличный*, в которых, помимо первой, была модифицирована и вторая часть — по слову *лик* (: *лицо*), закрепились в языке не сразу: судя по сохранившимся записям, это случилось не ранее XV–XVI вв. Сначала произошла частичная (*сердоничный*), а затем полная (*сердоличный*) ремотивация лексемы, причем модифицированный сегмент *сердо-* повлек за собой переосмысление второго сегмента. Сходный процесс последовательной ремотивации в русской речи, надо полагать, пережил и карельский топоним **Sordavala*. Вначале была переосмыслена первая часть *Сердо-* (< *Sorda-*), что спровоцировало дальнейшее переосмысление сегмента *-vala* в *-воль*, явно через сближение с корнем *вол*² — многих русских слов, таких как *воля*, *вольный* и т. п. Квалифицировать появление *-воль* (< *-vala*) закономерностями фонетики (например, в силу перехода приб.-фин. *a* > др.-русс. *o* и редуцирования гласной концовки при ударном начальном слоге четырехсложного имени) крайне сомнительно хотя бы потому, что данные процессы в известной мере факультативны, а самое главное — не объясняют смягчение *l* > *л*³. Переосмысленное *Сердоволь* воспринималось как сложение двух известных русских основ⁵, однако не выражало законченного смысла всего композита. Поэтому на следующем этапе, уже к концу XVI в. (судя по швед. *Sordabolschi eller Micholski Sochn* 1589 г., см. выше), произошла окончательная ремотивация топонима — в *Сердоболь*, по облику др.-русс. *сердоболь* ‘родственник, близкий по крови человек; соплеменник’, ‘близкий человек’ [СлРЯ XI–XVII, 24: 77], русск. *сердобольный*. Впрочем, вариант *Сердоболь* не смог вытеснить полностью раннюю форму *Сердоволь*, оба варианта продолжали конкурировать вплоть до XX в.

Воскресенский Соломянский погост располагался на северо-восточном берегу и прибрежных островах Ладожского озера, в устье реки Тулемайоки. В писцовой книге Водской пятины 1500 г. содержится характерная для эпохи перехода новгородских земель под власть Москвы запись: «Въ Заднѣй же Корелѣ Воскресенскомъ же погостѣ въ Соломянскомъ Великого Князя волость Соломяна, что была за владыкою» [ПОКВП: 180]. Земли, управлявшиеся новгородским архиепископом, отошли великому князю Ивану Третьему. С середины XVI в. сохранилась запись о погосте *Воскресенской Соломанской* [ПКВП 1568: 171]; в 1637 г., когда земли погоста, как и всего Корельского уезда, уже отошли Швеции, было составлено русскоязычное описание его северной части и в нем упомянут *Соломенской погост* [ПККЛ 1637: 720]. К. А. Неволин [1853: 125] на месте главного селения средневекового Соломянского погоста указывает отмеченный на картах 1-й половины XIX в. кирхшпиль *Салмис* на одной из губ северо-восточной оконечности Ладожского озера. В свою очередь, в шведских документах XVI–XVII вв. использовались записи *Salmis Sochn* [ПККУ 1590: 282], *Salmis pogost* [ПККУ 1631: 542]; *Salmis Pâgosth* или *Sallmis* 1637 г. [ПККЛ 1637: 341, 446], которые позволяют возводить топоним к карельскому термину *salmi* ‘пролив’. Название в полной мере оправдано топографически: выход из ладожского залива, на берегу которого стоит населенный пункт, в открытое озеро осуществлялся через узкий пролив между материком и островом Лункулансаари.

⁵ Первая из основ — сегмент *серд(о)-* — встречается как народноэтимологический модификат не только рассмотренных выше соотносимых лексических и топонимических фактов, но и лексической основы *серед-* (: *середина*). Ср. русск. диал. *сердовач*, *сердович* ‘человек средних лет’, *сердохрѣстье* ‘среда на четвертой неделе Великого поста’, перм. *сердолицы́й* с невыясненным значением и др. [СРНГ, 37: 188–189], новг. *серди́нник* ‘вид пирога с начинкой из крупы и селедки’ [НОС: 1077]. Вряд ли изменение *серед-* > *серд-* во всех случаях такого рода надо списывать на обыкновенную редуцию с утратой гласного, скорее речь должна идти о сближении с *сердце*, например, новг. *серди́нник* допускает мотивировку не столько по признаку срединного расположения начинки, сколько по форме пирога, напоминающей сердце. Данные факты отражают повышенный потенциал незакономерного появления сегмента *сердо-* в русском языке, хотя и не имеют прямого отношения к истории топонима *Сердоволь* / *Сердоболь*.

Русское наименование погоста в самой ранней его записи 1500 г. — *Соломянской* — восходит к топониму *Соломя*. Он фактически является др.-русск. термином *соломя* ‘пролив, протока’ [Срезневский, III: 461; СлРЯ XI–XVII, 26: 135], закрепленным в проприальной функции за конкретным локусом — селением возле одного из проливов на ладожском побережье. Селение, именовавшееся по-русски *Соломя*, а по-карельски *Salmi* (карельский ойконим сохраняется до сих пор — пос. *Салми* Питкярантского р-на), новгородцы сделали центром своего погостского округа, поставив там приходскую Воскресенскую церковь. Название *Соломя* (как и стоящий за ним термин *соломя*) изменялось по падежам с наращением *-ен/-ян-*, которое оставалось перед суффиксальным формантом *-ской* также в производных колеблющихся формах *Соломенской* / *Соломянской* / *Соломанской*⁶. Данные погостские наименования отметили обустройство погоста в *Соломяни* / *Соломени*. От косвеннопадежной основы *Соломян-* ойконима было образовано (с помощью суффикса-флексии *-а*) название р. *Соломяна*, известное из писцовой книги 1500 г. (совр. р. Тулемайоки), а по реке, в свою очередь, стали именовать волость *Соломяна*, упомянутую этим же источником. Хотя ойконим *Соломя* и карел. *Salmi* имеют единый этимологический источник (< карел. *salmi* ‘пролив’), облик древненовгородского ойконима нельзя считать результатом прямой фонетической адаптации *Salmi*. Дело в том, что заимствование восточными славянами приб.-фин. *salmi* ‘пролив’ случилось весьма рано, скорее еще до начала II тыс. н. э. (см., например, [Филин 1972: 533]). На древность заимствованного термина *соломя* намекают, во-первых, преимущественно дописменная хронология языковых процессов его славянизации, таких как полногласная рефлексия основы⁷, оформление по древнему консонантному склонению на **-men*, во-вторых, сравнительно раннее проникновение *соломя* в письменность (конец XIV в.), в-третьих, остаточные ареальные проявления *соломя* на рано славянизированной новгородско-псковской территории, включая юго-западные пределы⁸. Похоже, славяне освоили термин *соломя* ‘пролив’ еще в Ильмень-Волховском бассейне. Когда новгородцы добрались до северо-восточного побережья Ладоги и познакомились с карел. *Salmi*, они, понимая апеллятивный смысл данного ойконима-термина, заместили его созвучным и равнозначным др.-русск. *Соломя*. Таким образом, предпочтительнее вести речь не о фонетической адаптации, а о топонимической субституции карел. *Salmi*.

Богородицкий Кирьяжский погост впервые появляется в письменных источниках с XIV в. Комиссионный список Новгородской I-й летописи под 1396 г. упоминает этот погост под формой *Кюрьеский* вместе с Кюлолашским погостом: «Того же лѣта пришедше Нѣмци в Корѣльскую землю, и повоеваша 2 погоста: Кюрьеский и Кюлолашский» [НПЛ: 387]. Наименование погоста варьируется среди списков самой этой летописи и по иным, более поздним новгородским летописям XV в.: кроме *Кюрьеский* Комиссионного и Академического списков НПЛ, имеется еще форма *Кюрьеский* Толстовского списка [Там же], в Новгородской IV-й летописи — погост *Кюрьевский*, в Софийской I-й летописи — *Кирьевский*, в Летописи Авраамки — *Кирьеский* [Материалы 1941: 89]. Новгородская грамота

⁶ Нарращение *-ен-* является первичным. Начиная с XV в. в говорах, особенно севернорусских, наблюдается конкуренция *-ен-* с новым *-ян-*, которое возникло путем контаминации флексивных элементов (*солом-я* + *солом-ен-и* = *солом-ян-и*). Поэтому в письменности XV–XVI вв. написания типа *времяни*, *имяни*, *племяни* и т. п. не редкость [Хабургаев 1990: 76–77]. Это объясняет вариантность погостских наименований *Соломенской* / *Соломянской*. Что же касается варианта *Соломанской*, то он возник как результат спорадического отверждения согласного в *Соломянской*.

⁷ У восточных славян, впрочем, этот процесс завершился достаточно поздно: по мнению С. Б. Бернштейна [2005: 218–219], не ранее X в.

⁸ Таковы новг. (1963 г.) и пск. *сѡлѡ́мя* ‘узкий пролив, протока’, *сѡлѡ́менный* ‘проточный’ 1902–1904 гг. [СРНГ, 39: 293, 296], оз. *Соломенное* / *Соломино* близ Торопца, в самом деле представляющее собой расширение русла р. Торопы с узким проливом, р. *Соломенка* с руч. *Соломский* в районе Старой Руссы и др.

на бересте № 248, стратиграфически датированная 80-ми — 90-ми годами XIV в., представляет собой письмо, открывающееся словами: «Бѣють челомъ корила погоскаа, Кюлоласкаа и Кюриескаа господину Новугороду». Речь идет о письме карел, живущих на территории Кюлолакшского и Кирьяжского погостов, которые обращаются в Новгород за защитой от шведских грабежей [Зализняк 2004: 623–624]. Из этого документа следует, что интересующий нас погост именуется *Кюриеский*. Писцовая книга Водской пятины рубежа 1500 г. отмечает погостские наименования *Богородицкой Кирьяжской*, *Кирьяшской*, а также *Курьяжской* («В Курьяжскомъ же погостѣ великаго князя волости»), кроме того, упомянута р. *Кирьеша*: «на рѣцѣ на Кирьешѣ» [ПОКВП: 120, 122, 124, 127, 128]. Финское название этого новгородского погоста — *Kurkijoki* — впервые обнаруживаем в шведской переписной книге Корельского уезда 1590 г., где оно передано в виде *Kurchiochi* [ПККУ 1590: 265], причем книга отмечает финский топоним совместно с русским: *Kiriaschoi* или *Kurchiochi Sochn* (т. е. *Кирья(ж)ской* или *Куркиёки*), не оставляя сомнений в их тождестве. Далее форма *Kurchiochi* закрепляется в документах писцового делопроизводства и на картах шведского периода. В XVII в. селение получило примерно на 100 лет статус города со шведским названием *Kronoborg* ‘коронный город’. Современное название — пос. *Куркиёки* Лахденпохского р-на — является транслитерацией финского варианта *Kurkijoki*.

«Колеблющиеся» наименования Кирьяжского погоста в новгородской письменности XIV–XVI вв. (*Кюриеский* / *Кюрьеский* / *Кюрьский* / *Кюрьевский* / *Кирьевский* / *Кирьеский* / *Кирьяжской* / *Кирьяшской* / *Курьяжской*) и фиксации *Kurchiochi* / *Kurkijoki* (совр. *Куркиёки*) шведских и финских письменных документов подразумевают общий исторический карельский топоним с разнообразием его фонетической адаптации и письменной передачи в разноразличных средах. Опора на шведско-финские данные сразу же позволяет восстановить исходный карельский топоним в виде **Kurgijogi* со значением ‘журавлиная река’. Карельская форма до сих пор передается в финском облике с оглушением звонкого карел. *g* (*Kurkijoki*), что отражают и шведские средневековые документы. Более сложен вопрос появления многочисленных русских вариантов, которые тоже возводимы к карел. **Kurgijogi*.

Вообще говоря, возникновение русских топонимных вариантов с гласной меной первого слога (*Кур-*, *Кир-*, *Кюр-*) можно трактовать двояко — не только на древнерусской, но отчасти и на карельской языковой почве. В самом деле, закономерная для карельских говоров Приладожья утрата гласного *i* второго слога [Новак и др. 2019: 45] могла привести к исчезновению *г* и далее к полной редукции второго слога, что разгружало консонантный кластер на стыке двух основ сложения: **Kurgijogi* > **Kurgjogi* > **Kurjogi*. Это влекло за собой палатализацию *r* > *r'*, а в дальнейшем передвижку гласного *u* > *y* в первом слоге, т. е. **Kur'jogi* > **Kyr'jogi*. Однако, учитывая, что фин. *Kurkijoki* отсылает к карел. **Kurgijogi*, не затронутому вторичными фонетическими изменениями, логичнее исходить из того, что карел. **Kurgijogi* изменили не сами карелы, а средневековые новгородцы, после того как познакомились с названием карельского селения и учредили при нем погостский центр. Его существенную фонетическую русификацию оптимально связать с эпохой падения редуцированных. Карел. *i* второго слога воспринималось тогда еще как слабый *ь*, подлежащий утрате, а упрощение труднопроизносимых групп согласных, образовавшихся после падения редуцированных, относится к числу регулярных процессов русского языка XIII–XIV вв. Следовательно, на первом этапе карел. **Kurgijogi* обратилось в др.-русс. **Курьѣга* или **Курьѣга*⁹, от которого закономерно, с морфонологическим чередованием *г/ж*, образовалось наименование погоста *Курьяжской*, зафиксированное писцовой книгой 1500 г. На следующей стадии русификации возникло специфическое варьирование топоосновы *Кур-/Кюр-/Кир-*, которое демонстрируют новгородские летописи и писцовые книги. Такое

⁹ Ударение русифицированной формы сохранялось на первом слоге, что обеспечило передачу карел. *u* в виде др.-русс. *y*. Детерминант *jogi* ‘река’ при русской адаптации легко модифицировался в *-ѣга* или *-ѣга*. Севернорусская гидронимия прибалтийско-финского происхождения содержит массу примеров этих двух типов преобразования.

колебание основы невозможно считать результатом закономерного звукового развития на русской языковой почве, однако оно вполне объяснимо тривиальной «подгонкой» усвоенного новгородцами иноязычного топонима под облик хорошо известных древне- и старорусских календарных имен на *Кюр-/Кур-/Кир-*. Новгородское берестяное письмо знает «колеблющиеся» личные имена *Кюръ* и *Куръ*, *Кюрила* и *Курила*, *Кюрикъ* (см. [Зализняк 2004: 752–754, гр. 373, 138, 729, 220, Пск. 6])¹⁰, дополняемые по другим, обычно более поздним письменным источникам, вариантами *Кир*, *Кирил(а)*, *Кирик/Курик*, а также и самостоятельным именем *Кириак/Курьяк* и т. п., см., например, [Суперанская: 209].

Способом формальной «подгонки» под варианты формы *Кюръ/Киръ/Куръ* известного календарного имени (< греч. κύριος ‘господин, владыка; хозяин’) новгородцы явно пытались истолковать обсуждаемое иноязычное название. Наименование погоста *Курьяжской*, хотя и зафиксированное поздно — под 1500 г., но фактически отсылающее к первоначальному **Курьяга* (< карел. **Kurgijogi*), новгородцы сближали с антропонимом в варианте *Куръ*. С иным вариантом этого же антропонима — *Кюръ* — сблизились погостские наименования *Кюрьеский/Кюриеский/Кюрьеский/Кюрьский*, донесенные новгородской документацией XIV–XV вв. Антропонимная форма *Киръ*, в свою очередь, обусловила появление наименований *Кирьеский/Кирьеский* новгородских летописей XV в. и *Кирьяжской/Кирьяжской* писцовой книги 1500 г. и последующей документации. Показательно, что погостские наименования на *Кюр-* появляются только в ранних записях конца XIV — 1-й половины XV в., тогда как в более поздних находим формы на *Кир-* (< *Кюр-*) и на *Кур-*. Эти факты не случайны, они согласуются с хронологией форм в системе антропонимов: древнерусские формы на *Кюр-* (*Кюръ*, *Кюрикъ*, *Кюриль* и т. п.) в старорусское время замещались на *Кур*, *Кир*, *Кирик*, *Кирил*, *Кирил*.

Наряду с колебанием гласных первого слога, письменные фиксации показывают фонетическое варьирование второй основы карельского топонима, восходящей к детерминанту *-jogi*. Написания *Кюрьеский/Кюриеский/Кирьеский* отражают упрощение мягкого шипящего *ж'* перед суффиксом *-ск-* (< *-ьск-*), которое имеет ассимилятивную природу. Оно могло реализоваться только после утраты редуцированного в суффиксе, причем диалектная древненовгородская возможность смешения мягких *ж'* и *з'* способствовала ассимиляции *ж'ск > з'ск > ск*. На месте упрощенного *ж'* иногда наблюдается вставка *в* (*Кюрьеский/Кирьеский*), благодаря чему возникает псевдосуффикс *-ев-*. Полученные формы имитируют структуру русских адективных дериватов от личных имен с финалями *-ей*, *-ий* (типа *Кюрией*, *Кирей*) и, следовательно, наиболее отчетливо сообщают о русской антропонимной ремотивации иноязычного топонима. Фиксации наименований без согласного *ж* принадлежат эпохе XIV в., отчасти XV в., когда этот звук еще мог оставаться неотвердевшим. Сравнительно поздние фиксации 1500 г. и XVI в. демонстрируют уже окончательно отвердевший *ж*, не исчезающий перед *-ск-*: *Кирьяжской/Кирьяжской/Курьяжской*.

Центральным погостом-округом Корельского уезда был **Воскресенский Городенский**, а его главным селением — город-крепость *Корела* на западном берегу Ладожского озера, давший название уезду. Из Софийской I-й летописи известно, что в 1295 г. шведский отряд проник к устью Вуоксы и «поставиша Свѣйские Нѣмци городъ въ Корѣлѣ» [ПСРЛ, V: 202], который в хронике Эрика называется Кексгольмом (*Kekes-holm*) [Рыдзевская 1978: 112]. В том же году крепость была уничтожена новгородцами. Далее на месте обветшалых укреплений в 1310 г. новгородцы «срубиша городъ на порозѣ новъ, ветхый

¹⁰ Производные от древненовгородских личных имен *Кюръ*, *Кюрикъ* встречаются в писцовых книгах XV–XVI вв., которые содержат сведения о владельцах земель с патронимом *Кюров* (Кюров Емельян и Кюров Митрофан в Шелонской пятине, Кюров Михаил в Бежецкой пятине) и о д. *Кюриково* [НПК, V: 388; VI: 400; IV: 276]. В средневековой письменности встречались еще варианты имена *Чюръ*, *Чюриль* (откуда пошли производные фамилии *Чуров*, *Чурилов*, *Чурилин*), происходившие из форм *Кюръ*, *Кюриль* благодаря замещению мягкого *к'* мягким *ч* (уподобление процессу первой палатализации).

сметавше» [НПЛ: 92–93, 333], который в последующих сообщениях XIV в. Новгородской I-й летописи устойчиво именуется *Корельский городок* / *Корельский город* или сокращенно *Корельский* [НПЛ: 92–94, 96, 335, 338, 346, 367, 379, 477]. В писцовой книге Водской пятины 1500 г. Корельский уезд открывается описанием селений *Городенского* погоста, его центр именуется *Корельской городок*, или *городок Корела*: «у городка у Корель» [ПОКВП: 8, 16, 30].

Производное *Городенский* — единственное из наименований приладожских погостов-округов, которое не отсылает к прибалтийско-финскому топониму. Оно образовалось от терминов *городок*, *город* (< др.-русск. *городькъ*, *городь*), постоянно сопровождавших название погостского центра, судя по средневековым новгородским источникам. Наименование *Городенский* было закреплено еще за несколькими погостами Новгородской земли, центры которых располагались в *городах*, т. е. в *городных* (укрепленных) местах. На территории Водской пятины, помимо обсуждаемого *Воскресенского Городенского* погоста, значились *Дмитриевский Городенский*, *Успенский Городенский* и *Спасский Городенский* погосты, в Деревской пятине был свой погост *Городенский*. Топонимия на *Город* (*Городно*, *Городня*, *Городец*, *Городок* и т. п.), указывавшая на ранее укрепленные селения или просто на огороженные участки, наблюдается в сотнях мест в области бывшего новгородского пятинного деления¹¹.

Шведское название городского центра Городенского погоста, в свою очередь, запечатлено в виде *Kexholm* 1468 и 1482 гг., *Käkisholm*, *Käkssholm* 1473 г., *Kexholm* 1542 г., *Käkisalmj* 1556 г., *Käkisholm* 1603 г., в дальнейшем — *Kexholm* [Lehikoinen 2018: 23–24]. Оно перекликается с карельским и финским *Käkisalmi* ‘кукушкин (кукушений) пролив’, хотя и не соответствует ему полностью. Видимо, за *Kexholm* (швед. *-holm* ‘островок’) мог скрываться утраченный карельский прототип **Käkisoari* (карел. *-soari* ‘остров’), отражающий расположение крепости на острове в начале пролива. При этом совсем необязательно видеть за *käki* ‘кукушка’ апеллятив, это может быть и антропоним, на что указывают некоторые шведские топонимические формы, в которых *-s-* в первом элементе топонима (*Käkis-*, *Käks-*, да и собственно *Kex-*) может интерпретироваться как шведское окончание генитива с функцией посессивности [FSB]. Родовой патроним (фамилия) *Käki* и производные от него *Käkinen*, *Käkönen* неоднократно встречаются в документах XVI–XVII вв. на Карельском перешейке: *Peder Käki*, *Iwan Käköinen*, *Jöns Käköinen* и др. [Sukunimet: 249].

Михайловский Сакульский погост с центром в селении Сакула (фин. *Sakkola*), ныне пос. Громово Приозерского района Ленинградской области, этимологически одно из самых прозрачных наименований погостов Корельского уезда. Его письменная история берет начало с переписей XVI в.: *погост Михайловской Сакульской над Сванским озером*, *Сакула над Сванским озером* под 1500 г. [ПОКВП: 33, 53], *дер. Сакула у озера у Сванского* [ПКВП 1568: 123]. В шведском списке *Sackula Sochn* 1589 г., *Sackula by* 1592 и 1613 гг., *Sackola* 1748 г., *Sackola* 1767 г. [Kepsu 2018: 405–406]. В краеведческих изданиях принято считать, что смена первоначального *и* на *о* в шведских источниках была вызвана ошибочным написанием. В действительности за ней стоят правила шведского правописания XVII в., требующие именно такого написания. В дальнейшем же облик *Sakkola* закрепился и в произношении, несмотря на протесты местных жителей и рекомендации Географического общества Финляндии [Wallin 1896: 43].

Этимологию названия предложил в свое время В. Ниссиля, связавший его с карельским антропонимом *Sakku*, как народным вариантом одного из двух православных имен — *Исак* и *Захарий* [Nissilä 1975: 97, 121]. Интерпретация нашла поддержку и у современных исследователей [Kepsu 2018: 406; Lehikoinen 2018: 33]. Однако не стоит упускать из внимания

¹¹ В частности, многочисленность топонимов *Городок* позволила обосновать существование на территории Новгородской, Псковской и Тверской областей живого диалектного термина *городок* со значением ‘возвышенное место, урочище, с которым связывают пребывание древних людей’ [Васильев, Вихрова 2020], не замечавшегося прежде.

и возможность нехристианских истоков имени, частых в древней ойконимии Корельского уезда. В этом случае речь может идти о прозвище, в котором воплотился апеллятив *sak(k)a* ‘черт (особ. в ругательствах); шут, посмешище; проказник’ [SKES, 4: 951] в сочетании с диминутивным формантом *-u*. При этом безусловная антропонимная природа ойконима подтверждается присутствием в нем суффикса *-la*, который в традиционной карельской топонимии был маркером ойконима (первоначально именования дома, в дальнейшем — (однодворного) селения) и всегда присоединялся к антропониму [Nissilä 1962a].

Сакульский погост дает убедительные примеры того, что здешняя топонимия складывалась значительно раньше рубежа XV–XVI вв., когда она оказывается задокументированной. Именно здесь, на территории погоста, на древнем волоке между озером Сванским, иначе Свань, входившим в систему р. Вуоксы, и Ладожским побережьем располагался небольшой средневековый торговый городок *Волочек Сванской*, упоминаемый книгой 1500 г. [ПОКВП: 7] как купленный Великим князем у Валаамского монастыря. Его название отсылает к оз. *Сванскому* (совр. оз. Суходольское), по-фински *Suvanto* — из фин., карел. *suvanto* ‘плёс’ [СКНГТ: 208]: озеро представляло собой фактически участок, разлившееся в течении р. Вуоксы. Пара *Suvanto* — *Сванское* позволяет реконструировать в основе русского варианта утраченный редуцированный гласный (**Съван-*) и констатировать в результате, что топоним был усвоен в русское бытование еще в древнерусское время, до падения редуцированных. Изучение берестяных грамот позволило установить, что неконечные редуцированные массово исчезали в первые десятилетия XIII в. и практически сошли на нет к концу этого столетия [Зализняк 2004: 59–60, диаграмма IX]. Следовательно, новгородцы должны были познакомиться с озером и его карельским названием *Suvanto* не позднее XIII в., а скорее еще раньше — в XI или XII в., когда редуцированные были вполне регулярны. Для позднерусской эпохи восстанавливалась бы основа **Суван-* (а не **Съван-* > *Сван-*), сохраняющая прибалтийско-финский гласный. Волок и селение при нем играли важную роль для сокращения пути и доставки товаров из внутренних областей Корельского уезда в Новгород и обратно, так что раннее возникновение новгородского варианта вполне объяснимо. Надо полагать, что при интеграции карельского топонима **Suvando* (> фин. *Suvanto*) произошло упрощение образовавшейся при добавлении суффикса *-sk* группы согласных, приведшее к утрате *d*. В дальнейшем же из русского прилагательного уже вторичным образом абстрагировался вариант *Свань*, занесенный в ПОКВП.

Васильевский Ровдужский погост с центром в селении Рауту (фин. *Rautu*), на картах 1-й половины XIX в. — кирхшпиль Раутус к югу от оз. Сувандо [Неволин 1853: 133], ныне пос. Сосново Приозерского района Ленинградской области, — самый южный из погостов Корельского уезда. За финским названием *Rautu* стоит его карельский оригинал **Raudu*. Наиболее ранние упоминания обнаруживаются в писцовой книге Водской пятины 1500 г.: *погост Васильевской Ровдужской*, иногда в форме *Ровдожской* [ПОКВП: 22, 30, 76 и сл.] и книгах середины XVI в.: *в Ролдужском погосте* [ПКВП 1539, 37–39], *в Корельском присуде погост Васильевской Ровдужской* [ПКВП 1568: 148]. Шведские записи XVI–XVII вв. приведены в [Kersu 2018: 380] и отражают исходную карельскую или финскую огласовку первого слога (*Rautha*, *Rauto*), а также закономерную неустойчивость смычно-взрывного *d* второго слога, который в ситуации закрытого слога в соседстве с лабиализованным гласным преобразовывался в *v* [Новак и др. 2019: 125–127]: *Raw*, *Rauu*, *Rawu*. Концовка *-s* (*Rautus*, *Rautis*) — окончание генитива, которое в древнешведском употреблялось с существительными мужского и среднего рода и широко использовалось в названиях хуторов и дворов. В шведских вариантах финских топонимов оно стало восприниматься как топонимический показатель [FSB].

Фиксируемое документацией XVI в. наименование погоста *Ровдужской* / *Ровдожской* / *Ролдужской* восходит к названию карельского селения **Raudu*, в котором новгородцы устроили погостский центр. Поздние письменные фиксации топонимов (не ранее 1500 г.), конечно, не свидетельствуют о позднем учреждении погоста, напротив, расположение его, самое близкое к Новгороду среди остальных погостов Корельского уезда, подразумевает

древность освоения новгородцами этой местности. В указанных производных формах на *-ской* отражено несколько моментов ранней адаптации карел. *Raudu*. Здесь наблюдается рефлекс приб.-фин. *a* в виде *o* как свидетельство ранней интеграции топонима в официальное русское бытование. Приб.-фин. *u* после гласного перед согласным передано билабиальным *w*, который сравнительно долго сохранялся в новгородско-псковской зоне [Зализняк 2004: 55] и выступал заместителем гласного *u* в указанной слабой позиции, будучи согласным, очень близким к гласному по интенсивности колебания голосовых связок¹². Форма *Ролдужской* отражает спорадическое смещение *w* (> *в*) с боковым *л* в слабой предконсонантной позиции, известное субстратной севернорусской топонимии и заимствованной лексики; ср., например, варианты *Гавдозеро/Галдозеро* при вепс. *haud* ‘яма’, *Евчозеро/Елчозеро* при приб.-фин. *jouſen* ‘лебедь’ или сев.-русс. *рѡвдуга* ‘выделанная кожа оленя, лося, кабарги без шерсти, оленья, лосяная замша’ и т. п. наряду с *рѡлдуга* ‘шкура дикой козы, выделанная без шерсти’ [СРНГ, 35: 111–112, 170]¹³, ст.-русс. *ровдога* и *ролдуга* ‘ровдуга, выделанная оленья, лосяная или иная подобная шкура’, 1586 г. [Срезневский, III: 127].

Затруднительной остается только трактовка согласного *ж*, устойчиво закрепившегося при акте производства вариантов *Ровдужской/Ролдужской/Ровдожской* из карел. **Raudu* (> *Ровду-*). Этот согласный напоминает специфическую интерфиксальную вставку между гласным *u* и суффиксом *-ск-*, но скорее всего, мы имеем хорошо известный в русской лексике и топонимии суффикс *-уш-*, подвергшийся озвончению. Суффикс мог присоединиться к основе карел. **Raudu*, что при русификации топонима давало бы **Ровдуш-*, кроме того, он мог быть абстрагирован из элемента *-s-* прибалтийско-финских окончаний. Так, средневековые названия *Келтушского* погоста Водской пятины (совр. *Колтуши*) и д. *Кондуши* в Соломянском погосте, объясняемые из фин. *kelttu*, *keltto* ‘неплодородная, песчаная, сухая местность’ и карел. *kondu* ‘крестьянский двор; однодворная деревня’, вероятно, были получены путем присоединения русского суффикса *-уш* к прибалтийско-финским основам. Иную возможность появления *-ш-* иллюстрирует наименование находившегося рядом с Ровдужским *Куйвошского* погоста с главным селением *Куйвоша*. По-фински оно именуется *Kuivosii*, что выглядит как форма родительного падежа множественного числа патронима *Kuivonet* (ср. фамилию *Kuivonen*), указывавшая на ‘(селение) рода Куйвонен’. Значит, основа *Куйвош-* происходит здесь из *Kuivos(ii)*, а предпосылкой изменения *s* > *ш* могло стать недостаточное различие, вплоть до полного смещения, мягких свистящих и шипящих во многих говорах древненовгородского диалекта. Обе указанные возможности русской адаптации карел. **Raudu* ведут к появлению основы **Ровдуш-*, которая в дальнейшем могла превратиться в *Ровдуж-* (: *Ровдужской*). Новгородских фактов озвончения *ш* > *ж* в финалях топооснов обнаруживается немало¹⁴. Более того, нельзя исключать и вторичного нивелирующего влияния лексической основы *ровдуж-*, известной средневековым новгородцам; ср. вышеупомянутое сев.-русс. *рѡвдуга*, ст.-русс. *ровдога*,

¹² Аналогичное фонетическое изменение знакомо даже исконно русским лексемам, например русск. *завтра*, *завтрак* возникли из др.-русс. *заутра*, *заутрокъ* [Фасмер, II: 73] через промежуточное *заутр-*. Новгородская берестяная грамота № 417 содержит написание *завѣтрѣ*, где отразилось [zaw't'r-], развившееся из *заутр-* [Зализняк 2004: 81, 545]. После широкого закрепления губно-зубного *в* (< *w*) в русском языке гласный *у* в безударной поствокальной позиции обычно сохранялся неизменным.

¹³ Ср. еще известную фамилию *Ролдугин*, восходящую к этому слову.

¹⁴ Приведем лишь некоторые, из новгородской гидронимии: руч. *Моглажа* рядом с селением *Моглошь* в Среднем Поволховье (с суф. *-ош-*), р. *Любша* (с суф. *-ъш-*) с вариантами *Любжа*, *Люжба*, залив *Бронижа* в Верхневолжском вхр. на месте р. *Бранишня* (с суф. *-ш-*), р. *Апажа* с вариантом *Апаша* в бассейне Мсты, новгородская р. *Мологжа* при ярославской р. *Молокша*, р. *Сорогожа* бассейна Мологи, некогда называвшаяся *Сорогошина* (там же была волость *Сорогошино*).

ролдога — обозначение выделанной шкуры, но особенно его производные *ровдужий*, *ровдужина*, *ровдужный* [СРНГ, 35: 112]. К сожалению, письменные источники поздно упоминают древний Ровдужский погост (не ранее 1500 г.) и не свидетельствуют о более ранней основе **Ровдуш-*, как и о топониме **Ровдуж(а)*, от которого вполне ожидаемо производится наименование погоста *Ровдужской/Ролдужской/Ровдожской*.

Для исходного карельского топонима **Raudu* существуют по крайней мере две интерпретации, обе в определенной степени сомнительные. В свое время В. Ниссиля, крупнейший исследователь топонимии Карельского перешейка, предлагал, правда, под вопросом, сопоставлять его с фин. диал. *rautu* ‘речная форель’, которое реализовалось в названии оз. *Rautjärvi*, расположенного на территории погоста, а через него — в именовании селения [Nissilä 1975: 63]. Его оппонент Сауло Кепсу резонно обратил внимание на то, что термин бытует в говорах Приботнии и не известен на Карельском перешейке, и предложил, в свою очередь, реконструировать в названии озера карел. *rauda* ‘железо’, признавая при этом, что водоем расположен на самой окраине погоста, в отдалении от селения [Kerpsu 2018: 380].

Для этимологии полезно вписать **Raudu* (= фин. *Rautu*) в ряд других ойконимов Карельского перешейка с концовкой *-u*. Они немногочисленны, но составляют, тем не менее, однородную группу: *Tervu / Терву*, *Kirvu / Кирву*, *Harlu / Харлу*, *Paussu* (в составе пос. Харлу) и др. Она явно соотносится с многочисленными карельскими антропонимами на *-u*: *Hiltu* — Филат, *Jemmu* — Емельян, *Jeru* — Ерофей, *Mašu* — Мария, *Miku* — Михаил, *Ogru* — Агриппина, *Os'u* — Осип, *Riku* — Григорий, сюда же упомянутый в связи с историей названия Сакульского погоста карельский антропоним *Sakku*, соотносящийся с русск. *Захар* или *Исак*. В статье [Кузьмин 2016: 70], приводящей этот список, концовка *-u* вполне обоснованно считается суффиксом с диминутивной семантикой, ставшей своеобразным маркером личного имени. На апеллятивном уровне суффикс малопродуктивен и функции его несколько размыты, хотя в некоторых примерах (*härky* ‘бычок’ при *härkä* ‘бык’ и др.) и просматривается его диминутивное назначение [Hakulinen 2000: 184–185]. Видимо, перед нами пример универсальной закономерности, в соответствии с которой форманты, востребованные в онимах, теряют свою продуктивность в качестве апеллятивных суффиксов, и наоборот [Теория 1986: 167].

В этом контексте ойконимы следует рассматривать как восходящие к соответствующим антропонимам — именам, патронимам, прозвищам первооселенцев. Такой способ образования ойконимов, т. е. переход антропонима в ойконим при нулевой аффиксации, хорошо известен именованиям карельских селений, появившихся в более позднее время. Полагаем, что в карел. **Raudu* скрывается, соответственно, некалендарное имя или патроним **Raudu*, в основе которого карел. *rauda* ‘железо’. Показательно в связи с этим, что фамилия *Rauta* и суффиксальные *Rautanen*, *Rautava*, *Rautio*, *Rautala* широко бытуют в именослове Финляндии, будучи при этом зафиксированными и в ранних документах с территории Карельского перешейка [Sukunimet: 592–493]. Мотив именования не до конца понятен. Не исключено, что по крайней мере некоторые из патронимических образований могут указывать на профессию кузнеца (ср. фин. *rautio*, карел. *rautivo* ‘кузнец’).

Подведем некоторые **выводы**. Средневековые русские (древне- и старорусские), шведские и финские письменные материалы фиксируют названия семи новгородских приладожских погостов Карельского уезда в соответствии с фонетическими, лексическими и словообразовательными закономерностями своих языков. В разноязычной документации наблюдаются значительные колебания облика топонимов, что связано не только с собственно языковой спецификой их адаптации, но и с особенностями передачи топонимов в письменности — от более или менее правильного следования сложившейся орфографической норме вплоть до описок и ошибок, которыми нередко грешат и современные документы. Самые отличительные и многочисленные топонимные варианты запечатлены в средневековой русской (новгородской) письменности.

Наименования погостов Корельского уезда на *-ский/-ской* в систематизированной полноте дает новгородская писцовая книга 1500 г. [ПОКВП], она же содержит их первые письменные фиксации, за исключением Кирьяжского и Сердовольского погостов, древненовгородские и шведские свидетельства о которых идут с конца XIV в. и с середины XV в. Формы на *-ский/-ской* преимущественно были образованы от исконно карельских названий населенных пунктов, которые новгородцы учредили центрами погостских округов и церковных приходов. Название *Иломанского* погоста, самого северного в Корельском уезде, имеет предположительно древние саамские корни (прасаам. **elēmāñce* ‘самый верхний’), однако ко времени писцовых описаний оно уже было основательно переработано карельским языком и именно через карельское посредство попало в русские источники. Наименование *Городенского* погоста отличается от всех остальных не только исконным, древнерусским происхождением (< *городъ*, *городькъ* ‘укрепленное место’), но и тем, что оно относилось только к погосту-округу. Укрепленный центр Городенского погоста носит разные имена в русской и шведской документации: новг. *Корельской город(ок)* и швед. *Kekes-holm* и т. п.

Некоторые из погостских наименований, прежде всего те из них, которые сильно «колеблются» на письме (*Кирьяжский*, *Ровдужский*, *Сердовольский*), в своем фонетическом развитии ушли весьма далеко от реконструируемых карельских топонимов-прототипов (**Kurgijogi*, **Raudu*, **Sordavala*). В существенной степени это объясняется тем, что древненовгородская языковая адаптация данных карельских топонимов, помимо регулярных фонетических моментов, включала вторичное, народноэтимологическое сближение с созвучными лексемами и антропонимами. Это же справедливо и по отношению к названию погоста *Огреба* рядом с Корельским уездом (см. ниже). Вообще говоря, следует констатировать, что подобного рода ремотивация играет заметную (хотя обычно недооцениваемую) роль при русской адаптации иноязычных топонимов. К регулярности фонетических переходов нередко примешиваются нерегулярные внешние лексические воздействия. Идея о лексическом влиянии при адаптации топонимии применима также к трактовке наименования *Соломянского* погоста, которое, как мы полагаем, возникло путем замещения исходного карельского термина-топонима **Salmi* (= *salmi* ‘пролив’) древнерусским термином *соломя* в этом же значении, который был ранее усвоен славянами Новгородско-Псковского региона от прибалтийско-финского населения.

Исходные карельские топонимы имеют двоякое происхождение. В одних воспроизводится обозначение природного объекта, при котором стоит центр погоста, а первоначально — то карельское селение, которое стало таким центром. Таковы названия *Кирьяжского* (**Kurgijogi*), *Соломянского* (*Salmi*) и *Иломанского* (*Ilomantsi*, *Ilomančči*) погостов. Другие — *Сакульский* (*Sakkula*), *Ровдужский* (**Raudu*), *Сердовольский* (**Sordavala*) — содержат в основе антропоним, который должен квалифицироваться как патроним, личное имя или прозвище первопоселенца.

Явно антропонимные истоки имели названия еще двух древних карельских погостов — *Яскы* (карел. *Jeäski*, фин. *Jääski*) и *Огреба* (карел. *Äkräpeä*, фин. *Äyräpää*)¹⁵, отошедших по условиям Ореховецкого мирного договора к Швеции в 1323 г. Предполагается, что первое содержит в основе соответствующее карельское православное имя, восходящее к русск. *Ефим* или, возможно, *Иосиф* [Kerpu 2018: 166]. Однако с учетом древнего возраста погоста, продуктивности в ранней ойконимии Приладожья

¹⁵ Показательно, что фонетически в паре карел. *Äkräpeä* — русск. *Огреба* воплотилась та же модель ранней фонетической адаптации, которая известна в паре вепс. *Ä(ä)nižjärvi* — русск. *Онежское озеро*, т. е. в позиции начала слова приб.-фин. *ä* > др.-русс. *о*. Впрочем, применительно к паре *Äkräpeä* — *Огреба* нельзя не обратить внимания на дополнительное лексическое сближение с русск. *зреб-* (: *огребати*). В обоих случаях речь идет о древнерусских заимствованиях раннего периода.

нехристианского именослова, да и просто в силу фонетического облика наименования логичнее реконструировать в нем карельский прозвищный антропоним **Jeäski* — из *jeäski* ‘упрямый, упертый, с плохим нутром человек’ [KKS, 1: 495, 498]. Второе принято связывать с языческим карельским божеством по имени *Akris* (см. эту и другие сомнительные интерпретации в [Lehikoinen 2018: 47]), хотя с точки зрения ойконимической типологии значительно надежнее исходить из карельского антропонима прозвищного характера **Äkräpeä* из *äkräpeä* ‘упрямый, строптивый, сердитый’, букв. ‘упрямая голова’. Показательно, что оба эти погоста, наряду с *Сакульским* и *Ровдужским*, тоже сохраняющими в своих именах память о карельских основателях селений, находились на Вуоксе — главной средневековой карельской реке, в то время как вторичные, восходящие к обозначениям природных объектов карельские названия оказались закреплены за новгородскими погостами, тяготеющими к удаленной северо-восточной окраине уезда, а в прошлом Задней Корелы. В их ряду и название третьего утраченного карельского погоста — *Севилакша* или *Саволак* — из карел. *Savilakši*, букв. ‘глинистый залив’, на территории современной Юго-Восточной Финляндии. Такая ареальная дистрибуция, по-видимому, связана с особенностями карельского расселения и поэтапным освоением территории. В прибалтийско-финской ойконимии отмечено, что антропонимные топонимы преобладают в местах наиболее раннего оседлого земледельческого освоения [Nissilä 1962a: 92]. Вуокса была тем центром, откуда происходила миграция и освоение северных окраин Приладожья карелами. Конечно, нельзя абсолютизировать такую дистрибуцию, однако отметить ее как тенденцию стоит. Заманчиво встроить в этот же ряд ранние карельские центры на Вуоксе с антропонимными именованиями и *Корелу* или *Карельский городок*.

СПИСОК ИСТОЧНИКОВ

- Ан. РЭС, 9 — Аникин А. Е. *Русский этимологический словарь*. Вып. 9. М.: ИРЯ РАН; ИФЛ СО РАН, 2015.
- Материалы 1941 — *Материалы по истории Карелии XII–XVI вв.* Гейман В. Г. (ред.). Петрозаводск: Гос. изд-во КФ ССР, 1941.
- Неволин 1853 — Неволин К. А. *О пятинах и погостах Новгородских в XVI в. С приложением карты*. СПб.: Тип. Императорской академии наук, 1853.
- НОС — *Новгородский областной словарь*. Левичкин А. Н., Мызников С. А. (изд. подгот.). СПб.: Наука, 2010.
- НПК — *Новгородские писцовые книги, изданные Императорской археографической комиссией*. Т. I–VI. СПб.: Сенатская тип., 1859–1910.
- НПЛ — *Новгородская первая летопись старшего и младшего изводов*. Насонов А. Н. (ред.). М.; Л.: Изд-во АН СССР, 1950.
- ПКВП 1539 — Писцовая книга Водской пятины 1539 г. *История Карелии XVI–XVII вв. в документах*. Коваленко Г. М., Чернякова И. А., Салохеймо В. (сост.). Петрозаводск; Йоенсуу: б. и., 1987, 19–52.
- ПКВП 1568 — Писцовая книга Водской пятины 1568 г. *История Карелии XVI–XVII вв. в документах*. Коваленко Г. М., Чернякова И. А., Салохеймо В. (сост.). Петрозаводск; Йоенсуу: б. и., 1987, 52–278.
- ПККУ 1590 — Переписная книга Корельского уезда 1590 г. *История Карелии XVI–XVII вв. в документах*. Коваленко Г. М., Чернякова И. А., Салохеймо В. (сост.). Петрозаводск; Йоенсуу: б. и., 1987, 265–283.
- ПККУ 1631 — Переписная книга Корельского уезда 1590 г. *История Карелии XVI–XVII вв. в документах*. Коваленко Г. М., Чернякова И. А., Салохеймо В. (сост.). Петрозаводск; Йоенсуу: б. и., 1987, 388–568.
- ПККЛ 1637 — Поземельная книга Кексгольмского лена 1637 г. *История Карелии XVI–XVII вв. в документах. II*. Катаяла К., Хирвонен С. (ред.). Йоенсуу: б. и., 1991.

- ПОКВП — Переписная окладная книга по Новгороду Вотьской пятины Корельского городка с уездом 7008 года. Сообщ. д. ч. кн. М. А. Оболенским. *Временник Императорского Московского общества истории и древностей российских*. Кн. 12. Беляев И. Д. (ред.). М.: Университетская тип., 1852.
- ПСРЛ, V — *Полное собрание русских летописей*. Т. 5. V. VI. Псковская и Софийская летописи. СПб.: Тип. Эдуарда Праца, 1851.
- СКНГТ — Кузьмин Д. В. *Словарь карельской народной географической терминологии*. Петрозаводск: Периодика, 2020.
- СЛРЯ XI–XVII — *Словарь русского языка XI–XVII вв.* Вып. 1–. Бархударов С. Г. (гл. ред.). М.: Наука, 1975–.
- Срезневский — Срезневский И. И. *Материалы для словаря древнерусского языка*. Т. I–III. СПб.: Отд. рус. яз. и словесности Императорской академии наук, 1893–1912.
- СРНГ — *Словарь русских народных говоров*. Вып. 1–. Филин Ф. П., Сороколетов Ф. П., Мызников С. А. (гл. ред.). М.; Л.; СПб.: Наука, 1965–.
- Суперанская — Суперанская А. В. *Словарь русских личных имен*. М.: АСТ, 2010.
- Фасмер — Фасмер М. *Этимологический словарь русского языка*. Т. I–IV. Трубачев О. Н. (пер. с нем. и доп.). М.: Прогресс, 1986–1987.
- FSB — *Finlandssvenska bebyggelsenamn*. Электронный архив шведских ойконимов Финляндии на сайте Шведского литературного общества Финляндии: <https://bebyggelsenamn.sls.fi/namnelement/49/s/> (дата обращения 25.04.2024).
- Hausen 1924 — *Finlands Medeltidsurkunder. Samlade och i tryck utgifna af Finlands Statsarkiv genom Reinh. Hausen. IV. 1451–1480*. Helsingfors: Kejsarliga Senatens Tryckeri, 1924.
- KKS — *Karjalan kielen sanakirja*. 1–6. Virtaranta P., Koponen R. (toim.). Helsinki: Suomalais-Ugrilainen Seura, 1968–2005.
- SKES — *Suomen kielen etymologinen sanakirja*. 1–7. Toivonen Y. H. (toim.). Helsinki: Suomalais-Ugrilainen Seura, 1955–1981.
- SSA — *Suomen sanojen alkuperä. Etymologinen sanakirja*. 1–3. Itkonen E., Kulonen U.-M. (toim.). Helsinki: Kotimaisten kielten tutkimuskeskus; Suomalaisen Kirjallisuuden Seura, 1992–2000.
- Sukunimet — Mikkonen P., Paikkala S. *Sukunimet*. Keuruu: Otava, 1992.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Бернштейн 2005 — Бернштейн С. Б. *Сравнительная грамматика славянских языков: учебник*. 2-е изд. М.: Изд-во Московского ун-та; Наука, 2005. [Bernstein S. B. *Sravnitel'naya grammatika slavyanskikh yazykov: uchebnik* [Comparative grammar of Slavic languages: Textbook]. 2nd edn. Moscow: Moscow Univ. Press; Nauka, 2005.]
- Васильев, Вихрова 2014 — Васильев В. Л., Вихрова Н. Н. Типология наименований новгородских погостов (по данным писцовых книг конца XV–XVI в.). *Вопросы языкознания*, 2014, 3: 67–81. [Vasilyev V. L., Vixrova N. N. The typology of Medieval Novgorod pogost names (based on the data of land inventories of late 15–16th cc.). *Voprosy Jazykoznanija*, 2014, 3: 67–81.]
- Васильев, Вихрова 2020 — Васильев В. Л., Вихрова Н. Н. Топонимы и термины *Городок* в Новгородско-Псковско-Тверских землях. *Ономастика Поволжья*. Т. 2. Неганова Г. Д. (отв. ред.). Кострома: Костромской гос. ун-т, 2020, 21–30. [Vasilyev V. L., Vixrova N. N. Toponyms and terms *Gorodok* in the Novgorod-Pskov-Tver lands. *Onomastika Povolzh'ya*. Vol. 2. Neganova G. D. (ed.). Kostroma: Kostroma State Univ., 2020, 21–30.]
- Грот 1899 — Грот Я. К. *Филологические разыскания. II*. СПб.: Тип. Министерства путей сообщения, 1899. [Grot Ya. K. *Filologicheskie razyskaniya. II*. [Grot Ya. K. Philological studies. II]. St. Petersburg: Press of the Ministry of Railways, 1899.]
- Зализняк 2004 — Зализняк А. А. *Древненовгородский диалект*. 2-е изд. М.: Языки славянской культуры, 2004. [Zaliznyak A. A. *Drevnenovgorodskii dialekt* [Old Novgorod dialect]. 2nd edn. Moscow: Yazyki slavyanskoi kul'tury, 2004.]
- Кузьмин 2016 — Кузьмин Д. В. Христианские имена карелов. *Вопросы ономастики*. 2016, 13(2): 56–86. [Kuzmin D. V. Christian Names of Karelians. *Voprosy Onomastiki*, 2016, 13(2): 56–86.]
- Кучко 2023 — Кучко В. С. Многоликий сердолик: история названия камня в русской языковой традиции. *Известия УрФУ. Сер. 2. Гуманитарные науки*, 2023, 25(4): 251–263. [Kuchko V. S. Many-faced

- carnelian: The history of the names of the stone in the Russian language tradition. *Izvestiya. Ural Federal Univ. Journal. Series 2. Humanities and Arts*, 2023, 25(4): 251–263.]
- Левашов 1988 — Левашов Е. А. Сортавала (материалы к этимологии топонима). *Прибалтийско-финское языкознание. Вопросы лексикологии и грамматики*. Керт Г. М., Зайцева Н. Г. (отв. ред.). Петрозаводск: Карельский филиал АН СССР, 1988, 66–79. [Levashov E. A. Sortavala (materials on the etymology of the toponym). *Pribaltiisko-finskoe yazykoznanie. Voprosy leksikologii i grammatiki*. Kert G. M., Zaitseva N. G. (eds.). Petrozavodsk: Karelian branch of the Academy of Sciences of the USSR, 1988, 66–79.]
- Михайлова 2016 — Михайлова Л. В. Сортавала — Сердоболь. *Русская речь*, 2016, 3: 102–107. [Mikhailova L. V. Sortavala — Serdobol'. *Russkaya Rech'*, 2016, 3: 102–107.]
- Муллонен 2018 — Муллонен И. И. Фонетические варианты древней топоосновы *Ylä- 'верхний' и их генезис в гидронимии Карелии. *Вопросы ономастики*, 2018, 15(2): 7–27. [Mullonen I. I. Phonetic variants of the ancient toponymic stem *Ylä- 'upper' and their genesis in the hydronymy of Karelia. *Voprosy Onomastiki*, 2018, 15(2): 7–27.]
- Новак и др. 2019 — Новак И., Пенттонен М., Руусканен А., Сиилин Л. *Карельский язык в грамматиках. Сравнительное исследование фонетической и морфологической систем*. Петрозаводск: КарНЦ РАН, 2019. [Novak I., Penttonen M., Ruuskanen A., Siilin L. *Karel'skii yazyk v grammatikakh. Sravnitel'noe issledovanie foneticheskoi i morfologicheskoi sistem* [Karelian in grammars. Comparative study of phonetic and morphological systems]. Petrozavodsk: Karelian Research Centre of the Russian Academy of Sciences, 2019.]
- Очерки 2001 — Герд А. С., Лебедев Г. С. (ред.). *Очерки исторической географии: Северо-Запад России: Славяне и финны*. СПб.: Изд-во Санкт-Петербургского ун-та, 2001. [Gerd A. S., Lebedev G. S. (eds.). *Ocherki istoricheskoi geografii: Severo-Zapad Rossii: Slavyane i finny*. [Essays on historical geography: North-West Russia: Slavs and Finns]. St. Petersburg: St. Petersburg Univ. Press, 2001.]
- Рыдзевская 1978 — Рыдзевская Е. А. *Древняя Русь и Скандинавия в IX–XIV вв. (материалы и исследования)*. М.: Наука, 1978. [Rydzevskaia E. A. *Drevnyaya Rus' i Skandinaviya v IX–XIV vv. (materialy i issledovaniya)* [Ancient Rus' and Scandinavia in the IX–XIV centuries (materials and research)]. Moscow: Nauka, 1978.]
- Теория 1986 — Непокупный А. П. (отв. ред.). *Теория и методика ономастических исследований*. М.: Наука, 1986. [Nepokupnyi A. P. (ed.). *Teoriya i metodika onomasticheskikh issledovanii* [Theory and methodology of onomastic research]. Moscow: Nauka, 1986.]
- Филин 1972 — Филин Ф. П. *Происхождение русского, украинского и белорусского языков. Историко-диалектологический очерк*. Л.: Наука, 1972. [Filin F. P. *Proiskhozhdenie russkogo, ukrainskogo i belorusskogo yazykov. Istoriko-dialektologicheskii ocherk* [The origin of the Russian, Ukrainian and Belarusian languages. A historical and dialectological essay]. St. Petersburg: Nauka, 1972.]
- Хабургаев 1990 — Хабургаев Г. А. *Очерки исторической морфологии русского языка. Имена*. М.: Изд-во Московского ун-та, 1990. [Khaburgaeu G. A. *Ocherki istoricheskoi morfologii russkogo yazyka. Imena* [Essays on Russian historical morphology. Nominal morphology]. Moscow: Moscow Univ. Press, 1990.]
- Aikio 2003 — Aikio A. Suomen saamelaisperäisistä paikannimistä [Finnish place names of Sámi origin]. *Virittäjä*, 2003, 107(1): 99–106.
- Grünthal 1997 — Grünthal R. *Livvistä liiviin. Itämerensuomalaiset etnonymit* [From Livvi to Liivi. Baltic Finnish ethnonyms]. Helsinki: Suomalaisen Kirjallisuuden Seura, 1997.
- Hakulinen 2000 — Hakulinen L. *Suomen kielen rakenne ja kehitys* [The structure and development of the Finnish language]. 5th edn. Helsinki: Helsingin yliopiston suomen kielen laitos, 2000.
- Kepsu 2018 — Kepsu S. *Kannaksen kylät* [The villages of the Karelian Isthmus]. Helsinki: Kotimaisten kielten keskus, 2018.
- Kirkinen 1981 — Kirkinen H. Portti itään ja länteen [Gate to the east and west]. *Karjala*, 1981, 1: 11–23.
- Lehikoinen 2018 — Lehikoinen L. *Antreasta Äyräpähän. Luovutetun Karjalan pitäjien nimet* [From Antrea to Äyräpää. The names of the keepers of the Surrendered Karelia]. Helsinki: Kotimaisten kielten keskus, 2018.
- Nissilä 1962a — Nissilä V. *Suomalaista nimistöntutkimusta* [Study of Finnish place names]. Helsinki: Suomalaisen Kirjallisuuden Seura, 1962.
- Nissilä 1962b — Nissilä V. *Karjalan nimestä* [On the name of Karelia]. *Virittäjä*, 1962, 66(4): 345–363.

- Nissilä 1975 — Nissilä V. *Suomen Karjalan nimistö* [Toponymy of Finnish Karelia] Joensuu: Karjalaisen kulttuurin edistämissäätiö, 1975.
- Wallin 1896 — Wallin V. *Suomalaisten kuntain nimet*. Suomen Maantieteellisen Seuran kokoomain tietojen yhteensovitus [Names of Finnish municipalities. Revision of data compiled by the Finnish Geographical Society]. *Fennia*, 1896, 14 (3): 1–54.

Получено / received 23.05.2024

Принято / accepted 24.09.2024

Взаимосвязь между приемлемостью и вероятностью высказываний: данные предикативного согласования с сочиненным подлежащим в русском языке

© 2024

Ксения Андреевна Студеникина

Московский государственный университет имени М. В. Ломоносова, Москва, Россия;
xeanst@gmail.com

Аннотация: Данное исследование направлено установить, способны ли языковые модели, обученные на неразмеченных текстовых данных, моделировать согласовательную вариативность. Мы провели сравнение суждений о приемлемости, вынесенных носителями языка, и вероятностных метрик, предсказанных двунаправленной языковой моделью ruBERT без тонкой настройки. В качестве лингвистического феномена рассматривалось предикативное согласование с сочиненным подлежащим в русском языке. Мы подробно проанализировали, какие синтаксические, морфологические и семантические факторы оказывают влияние на приемлемость и вероятность высказываний. Использование результатов синтаксических экспериментов позволило выявить роль каждого фактора по отдельности и их взаимодействия. Помимо стандартной метрики для оценки вероятности последовательности, использовались ее вариации, учитывающие либо длину предложения, либо вероятность каждого токена, либо оба этих параметра. Мы предполагали, что модель будет предсказывать наибольшую вероятность той стратегии, которая является наиболее приемлемой с точки зрения носителей языка. Однако гипотеза не подтвердилась: наличие вариативного согласования с сочиненным подлежащим снижает корреляцию между приемлемостью и вероятностью. Линейное расположение элементов предложения — позиция подлежащего и сказуемого, порядок конъюнктов — оказалось единственным фактором, который в равной степени оказывает влияние на приемлемость и вероятность высказывания. Совпадение рода конъюнктов повышает приемлемость предикативного согласования по единственному числу, однако не меняет его вероятность. Одушевленность конъюнктов и симметричность предиката не влияют ни на приемлемость, ни на вероятность высказывания. Наше исследование показывает, что ruBERT не может быть использован для моделирования предикативного согласования с сочиненным подлежащим. Приемлемость высказывания опирается на более тонкие языковые конформансы, которые не значимы при автоматической оценке его вероятности.

Ключевые слова: грамматическая приемлемость, русский язык, синтаксис, согласование, сочинение, эксперимент, языковая модель

Благодарности: Автор выражает благодарность Екатерине Анатольевне Лютиковой, Анастасии Алексеевне Герасимовой и другим участникам Московской группы экспериментального синтаксиса за плодотворное сотрудничество при проведении экспериментальных исследований по изучению вариативного согласования в русском языке. Именно их результаты послужили материалом для данной работы. Исследование выполнено при финансовой поддержке Некоммерческого Фонда развития науки и образования «Интеллект».

Для цитирования: Студеникина К. А. Взаимосвязь между приемлемостью и вероятностью высказываний: данные предикативного согласования с сочиненным подлежащим в русском языке. *Вопросы языкознания*, 2024, 6: 105–132.

DOI: 10.31857/0373-658X.2024.6.105-132

Interaction between acceptability and probability: Evidence from predicate agreement with a coordinated subject in Russian

Ksenia A. Studenikina

Lomonosov Moscow State University, Moscow, Russia; xeanst@gmail.com

Abstract: The study aims to establish whether language models trained on unlabeled text data can parametrize agreement variation. We compared the acceptability judgments made by native speakers and the probability metrics predicted by the language model ruBERT without fine-tuning. As a specific linguistic phenomenon, we considered predicate agreement with a coordinated subject in Russian. We analyzed in detail which syntactic, morphological, and semantic factors influenced sentence acceptability and probability. The experimental data enables us to reveal the role of each factor and their interaction. Besides the standard logarithmic probability, we considered sentence length and unigram probability. We assumed that the model would assign the highest probability to the most acceptable agreement strategy. However, our hypothesis was not confirmed: the correlation between probability and acceptability is lower for sentences with agreement variation than for sentences without variation. The linear position — the subject-predicate order and the conjuncts' order — turned out to be the only factor which equally influences the acceptability and probability of a sentence. If the gender features of conjuncts match, the acceptability of singular agreement increases while the probability does not change. The animacy of conjuncts and the predicate symmetry influence neither acceptability nor probability. Our research demonstrates that ruBERT cannot be used to parametrize predicate agreement with a coordinated subject. The acceptability of a sentence is based on subtle linguistic contrasts which are not significant for the computer evaluation of its probability.

Keywords: acceptability, agreement, coordination, experiment, language model, Russian, syntax

Acknowledgements: The author expresses gratitude to Ekaterina A. Lyutikova, Anastasia A. Gerasimova and other participants of the Moscow Experimental Syntax Group for productive cooperation in conducting experimental studies on agreement variation in Russian. The current paper is based on the results of these studies. This work was supported by Non-Profit Foundation for the Advancement of Science and Education "INTELLECT".

For citation: Studenikina K. A. Interaction between acceptability and probability: Evidence from predicate agreement with a coordinated subject in Russian. *Voprosy Jazykoznanija*, 2024, 6: 105–132.

DOI: 10.31857/0373-658X.2024.6.105-132

1. Введение

Умение оценивать грамматическую правильность предложения на родном языке — ключевое свойство языковой способности человека [Chomsky 2014]. Под грамматичностью подразумевается интуитивная оценка, основанная на врожденной грамматике. Помимо грамматической правильности, на общую приемлемость предложения влияют факторы употребления: частотность слов, длина предложения, семантическая сочетаемость слов, объем рабочей памяти говорящего. Суждения о приемлемости предложений называют «оценками приемлемости». Они показывают, в какой степени высказывание является естественным по мнению носителей языка. В лингвистике оценки приемлемости используются для проверки грамматических теорий. Они выступают как один из основных типов данных для моделирования языковой способности людей и анализа восприятия говорящими определенного языкового выражения. Для извлечения оценок приемлемости проводится синтаксический эксперимент.

Суждения о приемлемости представляют интерес как для бинарно-категориальной, так и для вероятностной теорий грамматичности [Sprouse, Schütze 2014; Sprouse 2015].

Они являются градиентными по своей природе и поэтому не могут быть непосредственно учтены с помощью бинарных ограничений формальной грамматики. Важный вопрос состоит в том, можно ли свести приемлемость высказывания к его вероятности, которая широко исследуется в области компьютерной обработки языка.

Вероятность некоторой цепочки слов можно вычислить с помощью языковой модели. Чтобы она могла корректно рассчитывать вероятности, модель предварительно обучают на неразмеченном корпусе текстов. Алгоритм анализирует, какие комбинации слов и в каком порядке чаще всего встречаются вместе. Чем больше и разнообразнее набор текстов, на которых модель обучается, тем более сложные зависимости она улавливает и воспроизводит на новых данных. Первые языковые модели были построены на основе частотности *n*-грамм. С распространением машинного обучения возросла популярность языковых моделей на основе рекуррентных нейронных сетей и на основе архитектуры Трансформер.

В данной работе мы представим сравнение оценок приемлемости, выставленных носителями языка, и вероятностных метрик, предсказанных нейронной языковой моделью BERT. Материалом для исследования послужили данные синтаксических экспериментов по изучению предикативного согласования с сочиненным подлежащим. Анализ этой конфигурации позволяет выявить синтаксические, морфологические и семантические факторы, которые определяют выбор стратегии согласования при наличии множественного контролера. Сравнение суждений о приемлемости со значениями вероятности показывает, насколько обнаруженные корреляции релевантны для нейронных сетей. Мы намерены ответить на вопрос, способны ли языковые модели, обученные на неразмеченных данных без эксплицитной грамматической информации, усваивать тонкие лингвистические различия, которые важны при наличии согласовательной вариативности.

Статья имеет следующую структуру. В разделе 2 мы рассмотрим ряд предшествующих исследований, которые изучали связь между приемлемостью и вероятностью предложений. Раздел 3 посвящен проблеме вариативного согласования и особенности сбора данных для его изучения. В разделе 4 сосредоточимся на одном из контекстов согласования с несколькими контролерами — предикативном согласовании с сочиненным подлежащим. Раздел 5 содержит технические детали нашего эксперимента, а именно аргументацию выбора языковой модели и набор используемых вероятностных метрик. В разделе 6 мы описываем проведенный эксперимент по сравнению оценок приемлемости носителей и значений вероятности нейронной сети. Раздел 7 подводит итоги исследования.

2. Соотношение понятий «приемлемость» и «вероятность» высказывания

Вопрос о том, каким образом связаны приемлемость высказывания и вероятность его появления в речи, является фундаментальным. Еще Н. Хомский [Chomsky 1956] указывал, что установление соответствия между этими понятиями затруднительно. Так, предложения (1a) и (1b) одинаково невозможно встретить в английской речи, что делает их маловероятными. Однако при отсутствии семантической осмысленности обоих высказываний пример (1a) обладает грамматической правильностью, в отличие от примера (1b), что делает его более приемлемым. Утверждение об одинаково низкой вероятности примеров (1a) и (1b) опровергается в работе Ф. Перейра [Pereira 2000], в которой предлагается использовать биграммную языковую модель на основе классов для подсчета вероятности отсутствующих в обучающем корпусе биграмм. В результате вероятность высказывания (1a) оказывается в несколько раз выше, чем вероятность предложения (1b), что показано в выражении (2).

- (1) a. *Colorless green ideas sleep furiously.*
 b. *Furiously sleep ideas green colorless.*
 ‘Бесцветные зеленые идеи яростно спят’.

$$(2) \frac{P(\textit{Colorless green ideas sleep furiously.})}{P(\textit{Furiously sleep ideas green colorless.})} \approx 2 \times 10^5 \quad [\textit{Pereira 2000: 1245}]$$

Первое более подробное исследование взаимосвязи между приемлемостью и вероятностью представлено в статье Дж. Лау, А. Кларка и Ш. Лаппина [Lau et al. 2017]. На национальном корпусе английского языка были обучены *n*-граммная модель, скрытая марковская модель и рекуррентная нейронная сеть. При помощи краудсорсинга были собраны суждения о приемлемости 2500 предложений: из них 2000 предложений получены циклическим переводом 500 предложений из Британского национального корпуса на китайский, японский, норвежский, испанский и обратно. Суждения о приемлемости предложений собирались по шкалам трех типов: 1–2, 1–4, 1–100. Значения вероятности, порожденные моделью, сравнивались со средними оценками приемлемости носителей с помощью коэффициента корреляции Пирсона. Оказалось, что наблюдается значительная корреляция между предсказанными вероятностями и человеческими суждениями о приемлемости.

Тем не менее использование предложений, полученных с помощью циклического машинного перевода, не совсем корректно, поскольку они не отражают тонкие противопоставления, которые традиционно рассматриваются в экспериментальном синтаксисе. Данный недостаток был исправлен в работе Дж. Спрауза и коллег [Sprouse et al. 2018]. Рассматривалось три датасета, каждый из которых представляет интерес с точки зрения синтаксической теории. Первый набор данных состоял из минимальных пар из журнала «Linguistic Inquiry», в которых изменению подвергался только один грамматический параметр. Во второй датасет были добавлены предложения из учебника по теоретическому синтаксису Д. Аджера [Adger 2003]. Третья выборка содержала 120 предложений, которые получились путем перестановки слов в предложении (1a). Для экспериментов были выбраны *n*-граммная модель и рекуррентная нейронная сеть. Спрауз и коллеги выяснили, что данные языковые модели не способны предсказать вариативность, которая предсказывается категориальными оценками в учебнике Аджера, и не различают 23 % явлений, представленных в журнале «Linguistic Inquiry». Следовательно, при обращении к более сложным синтаксическим феноменам взаимосвязь между приемлемостью и вероятностью становится менее прозрачной.

Важным недостатком рассмотренных нами работ является тот факт, что они не приводят детальное сравнение оценок приемлемости и вероятностных метрик на материале конкретных лингвистических феноменов. Отсутствует информация о том, какие именно факторы учитываются носителями языка при вынесении суждения о естественности предложения и насколько они релевантны для языковых моделей.

Развернутая синтаксическая оценка была осуществлена в работе А. Варштадта и коллег [Warstadt et al. 2020] на материале корпуса лингвистических минимальных пар BLiMP. Он состоит из 67 отдельных наборов данных, каждый из которых содержит 1000 минимальных пар по приемлемости (3). Исследователи проводили эксперименты с *n*-граммной моделью, рекуррентной нейронной сетью и трансформерной моделью GPT-2. Наилучший результат продемонстрировала последняя модель. Было обнаружено, что наиболее точно идентифицируются морфологические контрасты, связанные с согласованием, но более тонкие семантические и синтаксические явления, такие как элементы отрицательной полярности, эллипсис и вопросительные выносы, вызывают трудности.

- (3) a. *The cats annoy Tim.* (приемлемо)
 b. **The cats annoys Tim.* (неприемлемо)
 ‘Кошки раздражают Тима’. [Warstadt et al. 2020: 377]

Таким образом, для простых случаев вроде предложений с предикативным согласованием подлежащего и сказуемого языковые модели довольно успешно предсказывают выбор наиболее приемлемого высказывания, тогда как при наличии дистантных зависимостей точность моделирования падает. Важно отметить, что предложения с предикативным согласованием и вопросительными выносами отличаются не только языковым феноменом, но и сложностью синтаксической структуры.

Интересующий нас вопрос состоит в том, насколько успешным будет моделирование предикативного согласования при наличии нестандартных или множественных контролеров, что делает возможным варьирование согласовательных вариантов. В следующем разделе мы детально рассмотрим проблематику согласования и ее релевантность для анализа взаимосвязи между приемлемостью и вероятностью высказываний.

3. Изучение вариативности при согласовании: проблемы и методы

Согласование (операция Agree) является одной из базовых операций в синтаксисе наряду с операцией соединения элементов (операция Merge) для построения предложения [Chomsky 1995]. В стандартном случае форма мишени согласования однозначно предсказывается на основании признаков контролера согласования, таких как лицо, число, род. Помимо контекстов, в которых выбор согласовательных признаков определяется однозначно, выделяются конструкции, где возможно вариативное согласование. В этом случае наблюдается неоднозначное вычисление согласовательных признаков мишени при фиксированных значениях признаков контролера. Варьирование возникает в ситуациях согласования с неканоническим контролером, таким как управляющий квантификатор (4), и при наличии нескольких потенциальных контролеров, как в сочинительных конструкциях (5).

(4) *Двое из нас придут / придет / *придем в гости.*

(5) *Петя и я идут / идем / *идут / *идет домой.*

Большинство исследований согласования в рамках русской научной традиции рассматривают контексты варьирования как нестандартный вариант, противопоставленный общему случаю. Л. Л. Иомдин [1990] называет их «осложненным согласованием», И. А. Мельчук [1993] — согласованием «по смыслу», А. Е. Кибрик [1999] — отклонением от ядерной модели. Однако простое перечисление случаев варьирования и осложняющих факторов недостаточно для разработки формальной модели согласования, основанной на единой системе признаков.

Эксплицитное моделирование вариативного согласования предполагает повышенные требования к эмпирическим данным. Выбор между согласовательными вариантами обусловлен множеством параметров, таких как лексические характеристики контролера и мишени, их линейный порядок, эффект синкретизма форм. Систематический сбор количественных данных о влиянии конкретных факторов осуществлялся в ряде предшествующих исследований, среди которых частотно-стилистический словарь по материалам газетного корпуса Л. К. Граудиной, В. А. Ицкович и Л. П. Катлинской [2001], монография по русскому синтаксису В. З. Санникова [2008], корпусные исследования конкретных контекстов варьирования [Пекелис 2013а; Кувшинская 2013; Добрушина, Сидорова 2019]. Однако данные источники содержат только описание приемлемых контекстов и не включают перечень неприемлемых случаев согласования. Чтобы разграничить допустимые и недопустимые стратегии согласования, необходимо подтвердить неприемлемость определенного контекста.

Моделирование согласовательной вариативности наиболее эффективно с опорой на данные экспериментальных исследований. В качестве стимулов выступают сконструированные

предложения, включающие набор всех исследуемых факторов, что позволяет учесть и статистически оценить их влияние. Кроме того, экспериментальные данные содержат сведения о допустимости всех типов контекстов, включая неприемлемые. Оценка некоторой конфигурации согласования как недопустимой подтверждает, что данный фактор ограничивает возможную вариативность. Методы экспериментального синтаксиса широко применяются для изучения вариативного согласования на материале различных языков, таких как словенский [Marušič et al. 2015], сербохорватский [Arsenijević, Mitić 2016], французский [An, Abeillé 2019], немецкий [Himmelreich, Hartmann 2023].

Проблематика вариативного согласования в русском языке активно разрабатывается Московской группой экспериментального синтаксиса¹. Ее участники занимаются изучением и формальным моделированием грамматических ограничений на основании экспериментальных данных. Исследованиям научной группы посвящен специальный выпуск журнала «Rhema. Рема» за 2023 год (№ 2). Статья Е. А. Лютиковой и А. А. Герасимовой [2023] рассматривает методологические проблемы, связанные с исследованием согласовательной вариативности. Результатом исследований Д. Д. Беловой и Т. И. Давидюк [2023], Д. Д. Врубель [2023], Л. И. Паско [2023] и К. А. Студеникиной [2023а] являются качественные эмпирические описания конкретных контекстов варьирования, сопровождаемые подробным статистическим анализом.

Результаты исследований Московской группы экспериментального синтаксиса могут выступить в качестве надежной эмпирической базы для анализа корреляции между приемлемостью и вероятностью. В качестве методики для всех экспериментов использовалась оценка приемлемости по шкале Ликерта от 1 до 7 [Likert 1932]. Данный метод является наиболее мощным в сравнении с другими методами извлечения суждений, как показали Дж. Спрауз, К. Шютце и Д. Алмеида [Sprouse et al. 2013] с помощью симуляции выборов. Экспериментальные исследования основываются на факторном подходе к изучению языковых явлений. Контролируя уровни факторов, мы можем оценивать влияние каждого из них по отдельности, а также их взаимодействия. Попарные сравнения позволяют понять, обнаруживаются ли значимые различия в приемлемости экспериментальных условий. Аналогичная процедура статистического анализа может быть применена и для вероятностных метрик. Если контрасты, наблюдаемые на материале оценок приемлемости, также выявляются для значений вероятности, можно говорить, что языковая модель способна успешно предсказывать выбор согласовательной стратегии. В следующем разделе будет подробно описан набор факторов, релевантных для моделирования вариативного согласования при сочинении.

4. Предикативное согласование с сочиненным подлежащим

Для автоматического моделирования варьирования было выбрано предикативное согласование с сочиненным подлежащим. Данный феномен позволяет проанализировать влияние разнообразных синтаксических, морфологических и семантических факторов. Исследователями Московской группы экспериментального синтаксиса накоплен значительный объем эмпирических данных о приемлемости, которые могут быть использованы для сравнения с вероятностными метриками языковых моделей. В подразделе 4.1 мы рассмотрим основные теоретические исследования данного явления на материале русского языка. Подраздел 4.2 посвящен экспериментальным исследованиям, результаты которых были использованы для анализа корреляции между приемлемостью и вероятностью высказываний.

¹ Сайт Московской группы экспериментального синтаксиса: <https://expsynt.com/>.

4.1. Теоретические исследования согласования при сочинении

В том случае, если подлежащее выражено сочиненной конструкцией, наблюдается несколько потенциальных контролеров предикативного согласования. Следовательно, возможны различные стратегии вычисления признаков для сказуемого — мишени согласования. Если подлежащее выражено сочиненными существительными единственного числа, для сказуемого наблюдается варьирование при вычислении признака числа. Если один из конъюнктов является местоимением и признаки лица конъюнктов различаются, глагол демонстрирует вариативность признака лица. В обоих случаях признаки глагола либо просто копируются от одного из конъюнктов, либо вычисляются в соответствии с правилами разрешения (англ. *resolution rules*), то есть некоторым образом выбираются исходя из того, какими значениями обладают конъюнкты [Corbett 1983].

Для конструкций, где подлежащее является сочиненной именной группой, возможны две стратегии согласования: полное или частичное. При полном согласовании признак числа вычисляется как сумма признаков обоих конъюнктов и сказуемое стоит во множественном числе (6). При частичном согласовании глагол получает признак единственного числа путем копирования признака только одного из конъюнктов (7). Рассмотрим факторы, определяющие выбор стратегии согласования в русском языке, на материале исследований В. З. Санникова [2008] и О. Е. Пекелис [2013б].

- (6) *Звук и свет **находятся** на единой шкале электромагнитных колебаний.* [«Волга», 2016]²
- (7) *Это был последний фильм, где звук и видеоряд **записывался** одновременно и вживую!* [Форум: 17 мгновений весны (2005–2010)]

В первую очередь, отмечается роль линейного порядка подлежащего и сказуемого. Преппозиция сказуемого (порядок VS) способствует его согласованию с ближайшим конъюнктом по единственному числу (8), а постпозиция (порядок SV) — согласованию со всей сочиненной группой (9).

[Пекелис 2013б]

- (8) *И вот уже **исчезли** / **исчез** и актер, и женщина.*
- (9) *И женщина, и актер **исчезли** / ***исчез**.*

Вероятность частичного согласования также повышается при совпадении признака рода сочиненных существительных. В этом случае согласовательные формы предиката совпадают для обоих конъюнктов, и любой из них способен выступать контролером согласования. Частичное согласование оказывается приемлемо, если оба конъюнкта имеют признак женского рода (10), однако недопустимо при сочинении существительного женского и мужского рода (11).

[Санников 2008: 158]

- (10) *Белена и крапива **росли** / **росла** прямо под окнами.*
- (11) *Белена и бурьян **росли** / ***росла** прямо под окнами.*

На выбор стратегии согласования влияет семантика глагола и конъюнктов. Для предикатов оказывается важной симметричность — семантическое свойство, которое предполагает аргументную структуру особого типа. Симметричным называется предикат, который обязательно имеет более одного участника с одинаковой семантической ролью. Так,

² Примеры (6)–(7) взяты из Национального корпуса русского языка (<https://ruscorpora.ru/>).

частичное согласование оказывается недопустимо, если сказуемое выражено симметричным предикатом, соотносящим субъекты друг с другом (12). Существительные выстраиваются в иерархию по способности индуцировать частичное согласование в зависимости от одушевленности / неодушевленности и конкретности / абстрактности: имена собственные (13) > одушевленные нарицательные > неодушевленные конкретные (14) > неодушевленные абстрактные (15). Чем выше по иерархии семантический тип существительного, тем выше приемлемость единственного числа глагола.

[Санников 2008: 157–159]

(12) Там целуются / *целуется Коля и его невеста.

(13) Отсюда мне видны / *виден Коля и Маша.

(14) Отсюда мне виден / видны дом и опушка леса.

(15) Во всем был виден / *были видны точный расчет и удивительная целеустремленность.

Для конструкций, где конъюнкты различаются по признаку лица, также возможно несколько стратегий предикативного согласования. Согласно нормативным грамматикам [РГ 1980; Розенталь и др. 1999], глагол получает признак лица конъюнкта, расположенного выше на личной иерархии $1 > 2 > 3$ [Zwicky 1977]. Если одним из конъюнктов является личное местоимение 1 лица, то согласование будет осуществляться по 1 лицу (16)–(17). Если в состав сочиненного ряда входит личное местоимение 2 лица, а второй конъюнкт имеет признак 3 лица, то «выигрывает» 2 лицо (18). Нарушение иерархии лиц возможно при порядке слов VS. В этом случае предикат копирует признак лица ближайшего конъюнкта (19). Кроме того, при сочинении местоимений 1 и 2 лица возможно согласование по 3 лицу. Данная стратегия называется «дефолтной»: ни один из конъюнктов не обладает признаком 3 лица, поэтому можно считать, что оно применяется по умолчанию.

[РГ 1980: 244]

(16) Ни я, ни ты не едем.

(17) Я и он придём.

(18) Ты и твои родные не едете.

(19) Остаются учителя старших классов и ты. [Пекелис 2013б]

Рассмотренные работы позволяют выделить релевантные факторы, определяющие выбор согласовательного варианта, однако они описывают отдельные сюжеты без учета полной картины параметров согласования. Чтобы осуществить статистическую оценку значимости факторов и их взаимодействия, были проведены экспериментальные исследования.

4.2. Экспериментальные исследования согласования при сочинении

Исследователи Московской группы экспериментального синтаксиса проанализировали влияние различных факторов на выбор стратегии согласования, среди которых тип конъюнктов (два местоимения / два существительных / одно местоимение и одно существительное), линейная позиция подлежащего и сказуемого (SV / VS), соотношение рода конъюнктов (оба мужского рода / оба женского рода / один мужского рода и один женского), симметричность предиката (симметричный / несимметричный), одушевленность конъюнктов (одушевленные / неодушевленные).

В данной статье мы сравним, в какой степени влияние перечисленных факторов на приемлемость высказывания совпадает с влиянием на его вероятность. Будут использованы

данные экспериментов Д. Д. Беловой, Д. Д. Врубель, Т. И. Давидюк, Л. И. Паско и К. А. Студеникиной, представленные на учебной конференции «Экспериментальные исследования языка» в 2022–2023 гг.

В табл. 1 кратко представлена характеристика указанных работ: рассматриваемый тип конъюнктов и независимые переменные. Эксперименты Д. Д. Врубель [2022], Л. И. Паско [2022] и К. А. Студеникиной [2023б] посвящены предикативному согласованию с именными конъюнктами по числу. Д. Д. Белова [2022] и Т. И. Давидюк [2022] изучали предикативное согласование по лицу и числу с конъюнктами, содержащими местоимение. Во всех исследованиях применялась методика извлечения суждений при помощи шкалы Ликерта от 1 до 7. Количество участников каждого эксперимента составило от 75 до 126 человек.

Таблица 1

Описание используемых экспериментальных исследований

Тип конъюкта	Согласование	Исследуемый фактор	Источник
существительные	числовое	соотношение рода конъюнктов	[Врубель 2022]
		симметричность предиката, позиция конъюнктов и предиката	[Паско 2022]
		одушевленность конъюнктов, позиция конъюнктов и предиката	[Студеникина 2023б]
местоимения <i>я, ты</i>	лично-числовое	порядок конъюнктов, позиция конъюнктов и предиката	[Белова 2022]
местоимение <i>я</i> , существительное		порядок конъюнктов	[Давидюк 2022]

Помимо стимульных предложений, каждый эксперимент содержал филлеры. Их соотношение к экспериментальным предложениям составляло 1:1. Филлеры не демонстрировали контекстов варьирования и были разделены на два типа: грамматичные представляли собой полностью корректные предложения, неграмматичные содержали ошибку в согласовании (20) или в предложном управлении (21). При анализе оценок стимульных предложений можно опираться на оценки для филлеров в качестве границы (не)грамматичности. Вероятностные метрики для филлеров также могут выступать в качестве бейзлайна: они показывают, насколько успешно модель оценивает невариативное согласование.

[Белова, Давидюк 2023: 73]

(20) *Дима и Витя убираем на кухне.*

(21) *Кирилл и Федя ушли по киоск.*

Влияние рода конъюнктов на выбор стратегии согласования рассмотрела Д. Д. Врубель. Для эксперимента были выбраны конструкции с повторяющимся союзом и порядком слов VS, способствующие возникновению частичного согласования. Пример экспериментального блока представлен в табл. 2 (с. 114). Приемлемость частичного согласования оказывается значимо выше для стимулов с совпадающим признаком рода у обоих конъюнктов.

Значимость симметричности предиката и линейного расположения подлежащего и сказуемого проанализировала Л. И. Паско. Примеры стимульных предложений содержатся в табл. 3 (с. 114). Симметричность предиката не оказала значимого влияния на выбор стратегии согласования. Частичное согласование при порядке SV было оценено ниже, чем при порядке VS, но значимо выше, чем неграмматичные филлеры.

Таблица 2

Пример экспериментального блока для исследования Д. Д. Врубель [2022]

Род конъюнктов	Число предиката	Стимул
м. р. + м. р.	ед. ч.	В длинном проигрыше в песне звучал и орган, и барабан.
	мн. ч.	В длинном проигрыше в песне звучали и орган, и барабан.
ж. р. + ж. р.	ед. ч.	В длинном проигрыше в песне звучала и электрогитара, и скрипка.
	мн. ч.	В длинном проигрыше в песне звучали и электрогитара, и скрипка.
ж. р. + м. р.	ед. ч.	В длинном проигрыше в песне звучала и электрогитара, и орган.
	мн. ч.	В длинном проигрыше в песне звучали и электрогитара, и орган.
м. р. + ж. р.	ед. ч.	В длинном проигрыше в песне звучал и орган, и электрогитара.
	мн. ч.	В длинном проигрыше в песне звучали и орган, и электрогитара.

Таблица 3

Пример экспериментального блока для исследования Л. И. Паско [2022]

Тип предиката	Линейная позиция субъекта и предиката	Число предиката	Стимул
симметричный	SV	ед. ч.	На старой фотографии сливается лицо и фон.
		мн. ч.	На старой фотографии сливаются лицо и фон.
	VS	ед. ч.	Лицо и фон сливается на старой фотографии.
		мн. ч.	Лицо и фон сливаются на старой фотографии.
несимметричный	SV	ед. ч.	На старой фотографии стирается лицо и фон.
		мн. ч.	На старой фотографии стираются лицо и фон.
	VS	ед. ч.	Лицо и фон стирается на старой фотографии.
		мн. ч.	Лицо и фон стираются на старой фотографии.

Роль одушевленности конъюнктов и линейной позиции сказуемого исследовала К. А. Студеникина. Табл. 4 (с. 115) включает примеры стимулов. Одушевленность не повлияла на приемлемость стратегии предикативного согласования. Роль линейной позиции подтвердилась: порядок VS более приемлем при согласовании по единственному числу, а порядок SV — по множественному.

Распределение стратегий предикативного согласования с подлежащим, выраженным сочиненными местоимениями *я* и *ты*, исследовала Д. Д. Белова. Было проведено два эксперимента с одинаковыми факторами, различающиеся порядком субъекта и предиката. Примеры экспериментальных блоков представлены в табл. 5 (с. 115). Нормативное согласование по 1 лицу множественному числу получает наиболее высокие оценки в обоих экспериментах. Дефолтное согласование по 3 лицу множественному числу оказывается приемлемым вне зависимости от порядка конъюнктов при SV и VS. Согласование с ближайшим конъюнктом является допустимым только при инверсии субъекта и предиката. После глагола в форме 1 лица единственного числа сочиненная конструкция *я и ты* получает значимо более высокие оценки, чем *ты и я*. Эта же закономерность обнаруживается с формой 2 лица единственного числа.

Таблица 4

Пример экспериментального блока для исследования К. А. Студеникиной [2023б]

Одушевленность конъюнктов	Линейная позиция субъекта и предиката	Число предиката	Стимул
одушевленные	SV	ед. ч.	Воспитатель и ребёнок присел на скамейку в саду.
		мн. ч.	Воспитатель и ребёнок присели на скамейку в саду.
	VS	ед. ч.	На скамейку в саду присел воспитатель и ребенок.
		мн. ч.	На скамейку в саду присели воспитатель и ребенок.
неодушевленные	SV	ед. ч.	Кабачок и арбуз созрел на огороде у бабушки.
		мн. ч.	Кабачок и арбуз созрели на огороде у бабушки.
	VS	ед. ч.	На огороде у бабушки созрел кабачок и арбуз.
		мн. ч.	На огороде у бабушки созрели кабачок и арбуз.

Таблица 5

Пример экспериментального блока для исследования Д. Д. Беловой [2022]

Линейная позиция субъекта и предиката	Порядок конъюнктов	Форма предиката	Стимул
SV (эксперимент 1)	<i>я и ты</i>	1 л., мн. ч.	Я и ты строим крепость из снега.
		1 л., ед. ч.	Я и ты строю крепость из снега.
		2 л., ед. ч.	Я и ты строишь крепость из снега.
		3 л., мн. ч.	Я и ты строят крепость из снега.
	<i>ты и я</i>	1 л., мн. ч.	Ты и я строим крепость из снега.
		1 л., ед. ч.	Ты и я строю крепость из снега.
		2 л., ед. ч.	Ты и я строишь крепость из снега.
		3 л., мн. ч.	Ты и я строят крепость из снега.
VS (эксперимент 2)	<i>я и ты</i>	1 л., мн. ч.	Крепость из снега строим я и ты.
		1 л., ед. ч.	Крепость из снега строю я и ты.
		2 л., ед. ч.	Крепость из снега строишь я и ты.
		3 л., мн. ч.	Крепость из снега строят я и ты.
	<i>ты и я</i>	1 л., мн. ч.	Крепость из снега строим ты и я.
		1 л., ед. ч.	Крепость из снега строю ты и я.
		2 л., ед. ч.	Крепость из снега строишь ты и я.
		3 л., мн. ч.	Крепость из снега строят ты и я.

Изучением согласования по лицу и числу с сочиненным подлежащим, выраженным местоимением *я* и именем собственным, занималась Т. И. Давидюк. Для всех стимульных предложений был зафиксирован порядок слов VS. Табл. 6 содержит пример экспериментального блока. Наиболее высоко оценивается согласование по 1 лицу множественному числу. Также приемлемым оказывается согласование с ближайшим конъюнктом, однако для порядка *я и X* оно оценивается значимо выше, чем для порядка *X и я*. Согласование

по 3 лицу множественному числу является значимо более приемлемым, чем согласование со вторым конъюнктом, оценки для которого ниже, чем для неграмматичных филлеров.

Таблица 6

Пример экспериментального блока для исследования Т. И. Давидюк [2022]

Порядок конъюнктов	Форма предиката	Стимул
<i>я и X</i>	1 л., мн. ч.	В районном центре живём я и Вася.
	1 л., ед. ч.	В районном центре живу я и Вася.
	3 л., ед. ч.	В районном центре живёт я и Вася.
	3 л., мн. ч.	В районном центре живут я и Вася.
<i>X и я</i>	1 л., мн. ч.	В районном центре живём Вася и я.
	1 л., ед. ч.	В районном центре живу Вася и я.
	3 л., ед. ч.	В районном центре живёт Вася и я.
	3 л., мн. ч.	В районном центре живут Вася и я.

Для автоматического моделирования вариативного согласования результаты экспериментов были объединены в единый датасет. Каждому предложению была присвоена средняя оценка приемлемости на основе ответов всех респондентов. Датасет включает следующую информацию: автора эксперимента, текст предложения, среднюю оценку приемлемости, тип (стимул / филлер), для стимулов — тип субъекта (два существительных / *я и ты* / *ты и я* / *я и X* / *X и я*), тип согласования (ед. ч. / мн. ч. / 1 л. мн. ч. / 1 л. ед. ч. / 2 л. ед. ч. / 3 л. ед. ч. / 3 л. мн. ч.), линейная позиция (SV / VS), симметричность предиката (симметричный / несимметричный), одушевленность конъюнктов (одушевленные / неодушевленные), род конъюнктов (совпадает / не совпадает), для филлеров — тип (грамматичный / неграмматичный). Фрагмент датасета проиллюстрирован в табл. 7, весь набор данных доступен по ссылке³.

Таблица 7

Фрагмент набора данных для моделирования вариативного согласования

author	sentence	score	type	subject	agree- ment	order	sym- metry	ani- macy	gender	filler- type
Pasko	Земля и небо размывается на далеком горизонте.	5	stimul	noun_ noun	sg	SV	nonsym	NA	nonmatch	NA
Pasko	Земля и небо размываются на далеком горизонте.	5.64	stimul	noun_ noun	pl	SV	nonsym	NA	nonmatch	NA
Pasko	Земля и небо сходится на далеком горизонте.	3.2	stimul	noun_ noun	sg	SV	sym	NA	nonmatch	NA
Pasko	Земля и небо сходятся на далеком горизонте.	6.2	stimul	noun_ noun	pl	SV	sym	NA	nonmatch	NA

³ https://github.com/Xeanst/acceptability_probablitiy_coordination

author	sentence	score	type	subject	agreement	order	symmetry	animacy	gender	filler-type
Pasko	Маша ценит доброту и честность в других людях.	6.60	filler	NA	NA	NA	NA	NA	NA	good
Pasko	Петя пишет сочинение и диктант на этом уроком.	2.79	filler	NA	NA	NA	NA	NA	NA	bad

Всего в датасет вошло 1660 предложений, среди которых 1404 стимула и 256 филлеров. В следующем разделе мы подробно рассмотрим процедуру выбора модели и подсчета метрик для оценки вероятности примеров из полученного набора данных.

5. Выбор языковой модели и метрик для оценки вероятности высказываний

В настоящее время наилучшие результаты компьютерной обработки текста достигаются с помощью больших языковых моделей на основе архитектуры Трансформер [Vaswani et al. 2017]. Благодаря механизму внимания подобные модели учитывают широкий контекст высказывания и улавливают дистантные синтаксические зависимости. Ожидается, что корреляция между оценками приемлемости и вероятностными метриками для языковых моделей на основе Трансформера должна быть значительно выше, чем для моделей менее сложной архитектуры.

Для современных больших языковых моделей важным становится разделение на однонаправленные (генеративные) и двунаправленные (автоэнкодерные). Исходно языковые модели, начиная с n -граммных, были однонаправленными. Поэтому вероятность P предложения s определялась как произведение вероятности каждого слова w_i с учетом левого контекста $w_{<i}$ (22). При обучении на размеченных текстовых данных для каждого слова предсказывается вероятность встретить его при условии, что известно предыдущее слово: w_2 при условии w_1 , w_3 при условии w_1 и w_2 и т. д. К однонаправленным или генеративным моделям на основе архитектуры Трансформер относятся модели семейства GPT [Radford et al. 2018]. Двунаправленные языковые модели наподобие BERT [Devlin et al. 2019] позволяют учитывать одновременно левый контекст $w_{<i}$ и правый контекст $w_{>i}$ при подсчете вероятности каждого слова w_i (23). В основе обучения лежит алгоритм маскированного языкового моделирования: случайно выбранные слова заменяются на спецсимвол [MASK], модель должна предсказать наиболее вероятный токен на месте маски.

$$(22) \vec{P}(s) = \prod_{i=0}^{|s|} P(w_i | w_{<i})$$

$$(23) \vec{P}(s) = \prod_{i=0}^{|s|} P(w_i | w_{<i}, w_{>i})$$

Корреляция между оценками приемлемости и вероятностными метриками для языковых моделей на основе Трансформера исследовалась в работе Дж. Лау и коллег [Lau et al. 2020]. Одни и те же высказывания рассматривались в трех конфигурациях: (i) изолированно, без контекста; (ii) с подходящим контекстом, в роли которого выступали три предшествующих предложения из исходного текста; (iii) с неуместным контекстом, в качестве которого использовались три случайных предложения. Приемлемые

предложения были взяты из Википедии, неприемлемые были получены с помощью циклического машинного перевода. Для предсказания вероятности последовательности использовались однонаправленная языковая модель GPT-2⁴ и двунаправленная модель BERT⁵. Результаты показали, что корреляция между оценками приемлемости и вероятностными метриками оказывается больше для двунаправленной языковой модели, чем для однонаправленной. Наличие релевантного контекста повышает приемлемость и вероятность предложений.

Сравнение приемлемости и вероятности проводилось также на материале русского языка. В исследовании В. Михайлова и коллег [Mikhailov et al. 2022] на материале корпуса лингвистической приемлемости для русского языка RuCoLA реализуется бинарная классификация предложений по приемлемости. Для применения двунаправленных моделей осуществляется тонкая настройка. Без дообучения применяется только однонаправленная модель ruGPT-3⁶. С помощью кросс-валидации исследователи находят пороговое значение для предсказанных ей вероятностных метрик. Бинарная классификация по приемлемости с помощью ruGPT-3 оказывается практически случайной: доля правильных ответов составила 55.79 %.

Вопрос о том, как связаны оценки приемлемости по шкале Ликерта и значения вероятности генеративной модели ruGPT-3, также был рассмотрен на материале русских именных групп с сочиненными модификаторами [Studenikina 2023]. Рассматриваемая языковая модель предсказывает корректные вероятностные метрики для предложений, где выбор согласовательной стратегии определяется однозначно. Для контекстов свободного варьирования вероятностные метрики также совпадают с оценками приемлемости. Однако предсказанные моделью вероятности не позволяют выявить контрасты в приемлемости, которые демонстрируют оценки носителей при наличии одной предпочтительной стратегии согласования.

Рассмотренные исследования демонстрируют, что вероятностные метрики однонаправленных моделей не вполне коррелируют с оценками приемлемости носителей. Поскольку они учитывают только левый контекст предложения, значительная часть информации о приемлемости теряется. В частности, генеративные модели не позволяют рассчитать, насколько вероятным будет появление определенной формы глагола при сочиненном подлежащем, если сказуемое расположено в препозиции. Поэтому для нашей задачи корректнее будет использовать вероятностные метрики, предсказанные двунаправленной языковой моделью. В нашем исследовании была использована модель ruBERT⁷ в базовой конфигурации, учитывающая регистр. Для удобства вычислений вместо исходной вероятности используется логарифмическая вероятность слов:

$$(24) LP = \sum_{i=0}^{|s|} \log P(w_i | w_{<i}, w_{>i})$$

Важно отметить, что вероятность предложения $P(s)$, подсчитанная с помощью языковой модели, не всегда напрямую связана с его приемлемостью. Влияние оказывают также факторы употребления. Поскольку с увеличением длины предложения его вероятность уменьшается, применяется нормализация по длине $|s|$ с возможным учетом коэффициента масштабирования α . Для учета частотности слов используется униграммная вероятность предложения $P_u(s)$. Она рассчитывалась по русскоязычному корпусу Araneum Russicum⁸ (120 миллионов вхождений).

⁴ <https://github.com/openai/gpt-2>

⁵ <https://github.com/google-research/bert>

⁶ https://huggingface.co/ai-forever/rugpt3medium_based_on_gpt2

⁷ <https://huggingface.co/DeepPavlov/rubert-base-cased>

⁸ <http://aranea.juls.savba.sk/guest/>

Помимо стандартной метрики для оценки вероятности последовательности (LP), были опробованы ее вариации, учитывающие либо длину последовательности ($MeanLP$, $PenLP$), либо вероятность каждого токена ($NormLP$), либо оба этих параметра ($SLOR$). Формулы для расчета метрик представлены в табл. 8.

Таблица 8

**Метрики для подсчета вероятности предложения,
предложенные Дж. Лау и коллегами [Lau et al. 2020]**

Метрика	Формула	Расшифровка
LP	$\log P(s)$	логарифмическая вероятность предложения
MeanLP	$\frac{\log P(s)}{ s }$	логарифмическая вероятность предложения, нормализованная по длине
PenLP	$\frac{\log P(s)}{\left(\frac{5 + s }{5 + 1}\right)^\alpha}$	логарифмическая вероятность предложения, нормализованная по длине с учетом коэффициента масштабирования
NormLP	$-\frac{\log P(s)}{\log P_u(s)}$	логарифмическая вероятность предложения, нормализованная с учетом униграммной вероятности
SLOR	$\frac{\log P(s) - \log P_u(s)}{ s }$	логарифмическая вероятность предложения, нормализованная по длине с учетом униграммной вероятности

В следующем разделе мы продемонстрируем, насколько каждая из рассмотренных метрик коррелирует с оценками приемлемости. Также будет проанализировано, способны ли данные метрики улавливать контрасты в приемлемости стратегий согласования, которые оказываются релевантными для носителей языка.

6. Анализ корреляции между вероятностными метриками ruBERT и оценками приемлемости

Для анализа величины линейной связи между оценками приемлемости носителей и вероятностными метриками модели ruBERT мы использовали коэффициент корреляции Пирсона. Собранный набор данных содержит как предложения, где наблюдается вариативное согласование (стимулы), так и примеры, где выбор стратегии согласования однозначен (филлеры). Чтобы понять, насколько нейронные сети способны моделировать разные типы согласования, необходимо оценить коэффициент корреляции для каждой из этих групп. В табл. 9 представлены полученные значения. Вероятностные метрики гораздо лучше предсказывают невариативное согласование, значение коэффициента корреляции при наличии вариативности снижается. Для вариативных контекстов наибольшая корреляция наблюдается для метрики SLOR, учитывающей длину и частоту слов. Однозначное согласование лучше моделируется метрикой LP, которая представляет собой логарифмическую вероятность без учета дополнительных параметров.

Таблица 9

Значение коэффициента корреляции Пирсона для каждой из вероятностных метрик

	LP	MeanLP	PenLP	NormLP	SLOR
Наличие вариативности	0.27	0.31	0.30	0.39	0.45
Отсутствие вариативности	0.63	0.49	0.58	0.52	0.51

В ходе экспериментальных исследований предикативного согласования с сочиненным подлежащим были изучены различные факторы: тип конъюнктов (два существительных / два местоимения / местоимение и существительное), порядок слов (SV / VS), соотношение рода конъюнктов (совпадает / различается), тип предиката (симметричный / несимметричный), тип конъюнктов (одушевленные / неодушевленные), порядок конъюнктов (*я и ты, ты и я, я и X, X и я*). Статистический анализ оценок приемлемости, которые выставлены носителями языка, позволяет оценить значимость каждого из них. Мы рассмотрим, насколько данные факторы оказались важными для языковой модели *tuBERT*, осуществив аналогичный статистический анализ для вероятностных метрик. Для выявления значимости использовался многофакторный дисперсионный анализ ANOVA. В подразделе 6.1. мы рассмотрим фактор порядка слов, который был задействован во всех экспериментальных исследованиях. В подразделе 6.2 будут проанализированы факторы, которые влияют на выбор стратегии согласования с именными конъюнктами: соотношение рода, одушевленность существительных и симметричность глагола. Подраздел 6.3 описывает стратегии согласования для местоименных конъюнктов. В подразделе 6.4 представлен анализ предикативного согласования с подлежащим, содержащим местоимение и существительное. Подраздел 6.5 сравнивает результаты для разных вероятностных метрик.

6.1. Влияние порядка слов на выбор стратегии согласования

На рис. 1 (с. 121) представлены средние значения для оценок приемлемости и вероятностных метрик в зависимости от линейного порядка подлежащего и сказуемого. По оценкам можно заметить, что множественное число глагола более приемлемо при любом порядке слов, однако приемлемость единственного числа повышается при обратном порядке подлежащего и сказуемого. Дисперсионный анализ выявил значимое влияние на приемлемость предложений фактора числа ($p < .0001$), порядка слов ($p < .01$) и их взаимодействия ($p < .0001$). Фактор числа также оказывает значимое влияние на вероятность высказываний ($p < .0001$ для всех метрик), однако значимость фактора порядка слов не наблюдается для большинства метрик ($p < .05$ для MeanLP, $p > .05$ для LP, PenLP, NormLP и SLOR). Взаимодействие числа и порядка слов также значимо влияет на вероятность предложений ($p < .0001$ для LP, PenLP, NormLP и SLOR, $p < .001$ для MeanLP).

Разница в приемлемости и вероятности единственного и множественного числа при порядке SV оказывается значимой ($p < .0001$ для оценок носителей и всех вероятностных метрик). При порядке VS различие в приемлемости и вероятности единственного и множественного числа также значимо ($p < .0001$ для оценок, $p < .001$ для SLOR; $p < .01$ для LP, PenLP и NormLP; $p < .05$ для MeanLP). Контексты с единственным числом глагола значимо отличаются по приемлемости и вероятности в зависимости от порядка слов ($p < .0001$ для оценок, LP и MeanLP; $p < .001$ для NormLP; $p < .01$ для PenLP и SLOR). В большинстве случаев значимое различие не наблюдается между прямым и обратным порядком слов при множественном числе глагола ($p < .001$ для LP; $p > .05$ для оценок, MeanLP, PenLP, NormLP и SLOR).

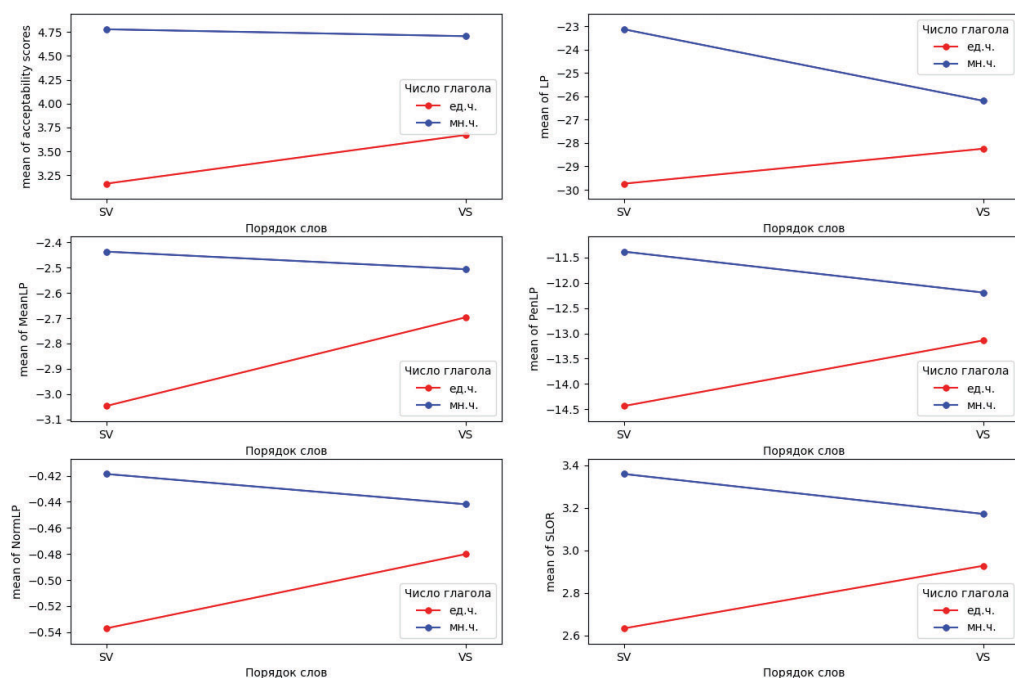


Рис. 1. Средние значения оценок приемлемости и вероятностных метрик в зависимости от порядка слов

Контрасты в приемлемости, наблюдаемые при различном порядке слов, оказываются релевантны и для значений вероятности. Большинство метрик показывают схожий результат, отличие демонстрирует только метрика LP, выявившая избыточный контраст при одном из попарных сравнений. Различия в линейной позиции слов оказываются значимыми при выборе стратегии согласования как для носителей, так и для нейросетевой модели.

6.2. Анализ морфологических и семантических факторов для именных конъюнктов

Рис. 2 демонстрирует средние значения оценок приемлемости и вероятностных метрик с учетом соотношения рода именных конъюнктов. Влияние рода на оценки приемлемости аналогично влиянию линейного порядка: множественное число глагола одинаково приемлемо при любом соотношении признаков рода, приемлемость единственного числа повышается при совпадении рода конъюнктов. Дисперсионный анализ показывает, что приемлемость предложений зависит от числа сказуемого ($p < .0001$), рода подлежащих ($p < .0001$) и взаимодействия этих факторов ($p < .05$). Для вероятностных метрик подтверждается только значимость фактора числа ($p < .01$ для NormLP и SLOR; $p < .05$ для LP, MeanLP и PenLP). На вероятность не оказывают значимого влияния фактор рода конъюнктов ($p > .05$ для всех метрик) и взаимодействие факторов ($p > .05$ для всех метрик).

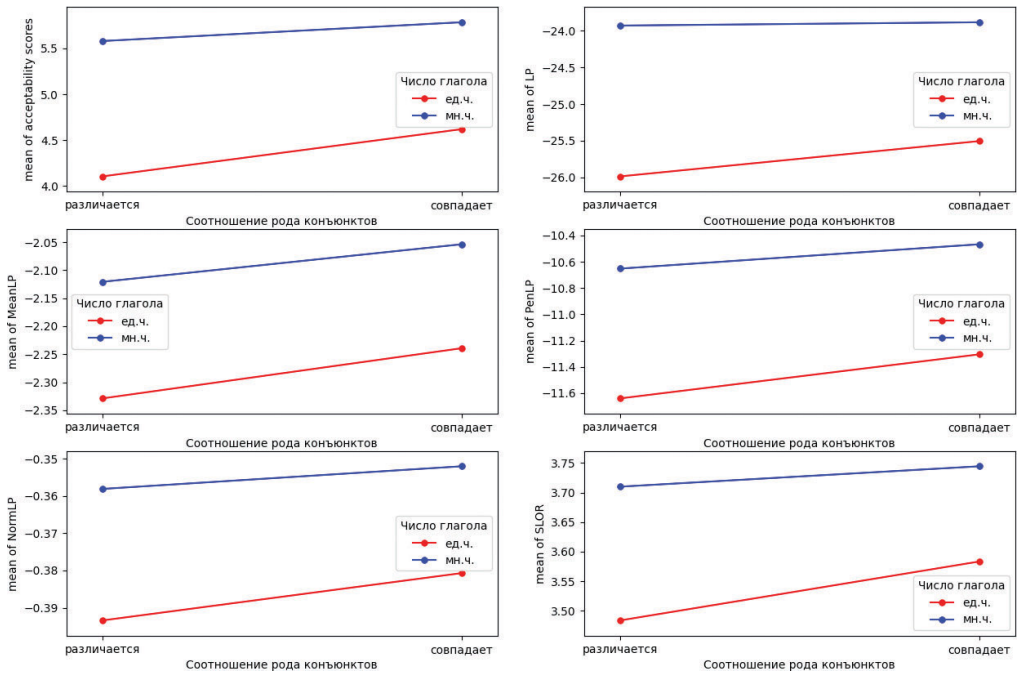


Рис. 2. Средние значения оценок приемлемости и вероятностных метрик в зависимости от соотношения рода конъюнктов

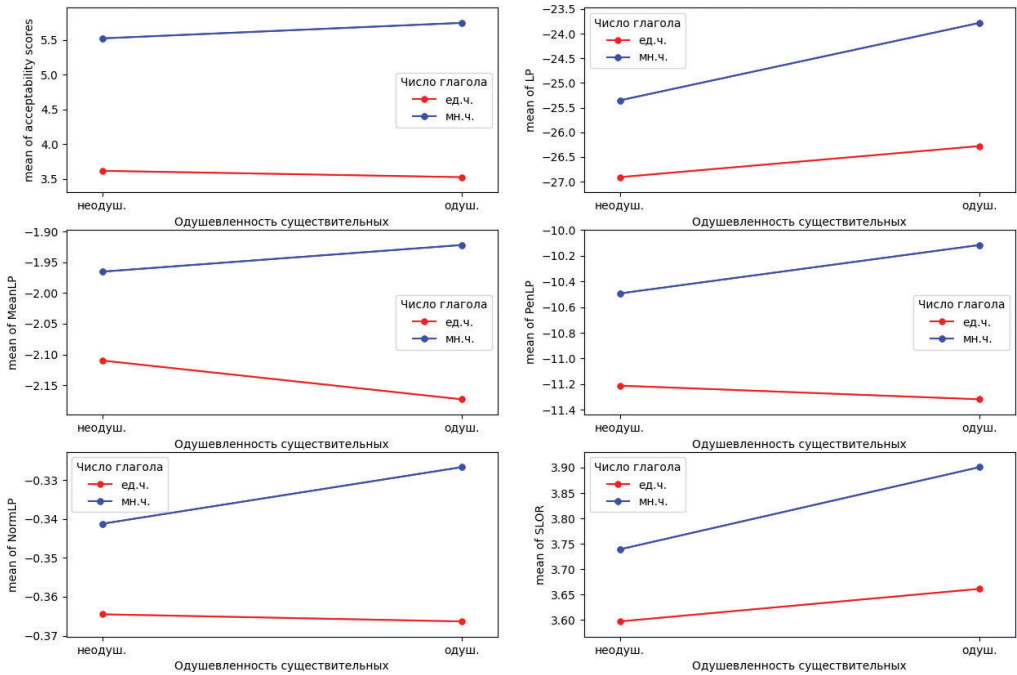


Рис. 3. Средние значения оценок приемлемости и вероятностных метрик в зависимости от одушевленности конъюнктов

Попарные сравнения методом Тьюки выявили значимую разницу в приемлемости единственного и множественного числа при совпадении рода конъюнктов ($p < .0001$) и при различии рода ($p < .0001$). Оценки приемлемости значимо меняются при совпадении и при различии рода, если сказуемое стоит в единственном числе ($p < .0001$), однако не меняются при множественном числе сказуемого ($p > .05$). Что касается контрастов для вероятностных метрик, статистический анализ не выявил значимых различий ни для одной пары условий ($p > .05$ для всех метрик). В отличие от приемлемости, вероятность единственного числа предиката не зависит от соотношения рода конъюнктов. Кроме того, вероятностные метрики не отражают большую приемлемость множественного числа при фиксированном соотношении рода конъюнктов.

На рис. 3 (с. 122) изображены средние значения оценок приемлемости и вероятностных метрик при различном семантическом типе именного подлежащего. Дисперсионный анализ суждений носителей и вероятностных метрик выявил значимость фактора числа ($p < .0001$ для оценок; $p < .05$ для всех метрик). Фактор одушевленности и взаимодействие факторов оказываются незначимыми ($p > .05$ для оценок и всех метрик).

Попарные сравнения выявили значимые различия в приемлемости между единственным и множественным числом глагола при одушевленных ($p < .0001$) и при неодушевленных существительных ($p < .0001$). Однако предложения с одушевленными и неодушевленными конъюнктами значимо не отличаются в приемлемости как в единственном ($p > .05$), так и во множественном числе ($p > .05$). В отношении вероятностных метрик попарные сравнения не выявили значимых контрастов ни для одного из условий ($p > .05$ для всех метрик). Следовательно, одушевленность конъюнктов не влияет ни на приемлемость, ни на вероятность при выборе числа предиката. Большая приемлемость множественного числа глагола при фиксированном семантическом типе существительных не отражается с помощью вероятностных метрик.

Рис. 4 показывает средние значения оценок приемлемости и вероятностных метрик при различном семантическом типе сказуемого. Дисперсионный анализ продемонстрировал

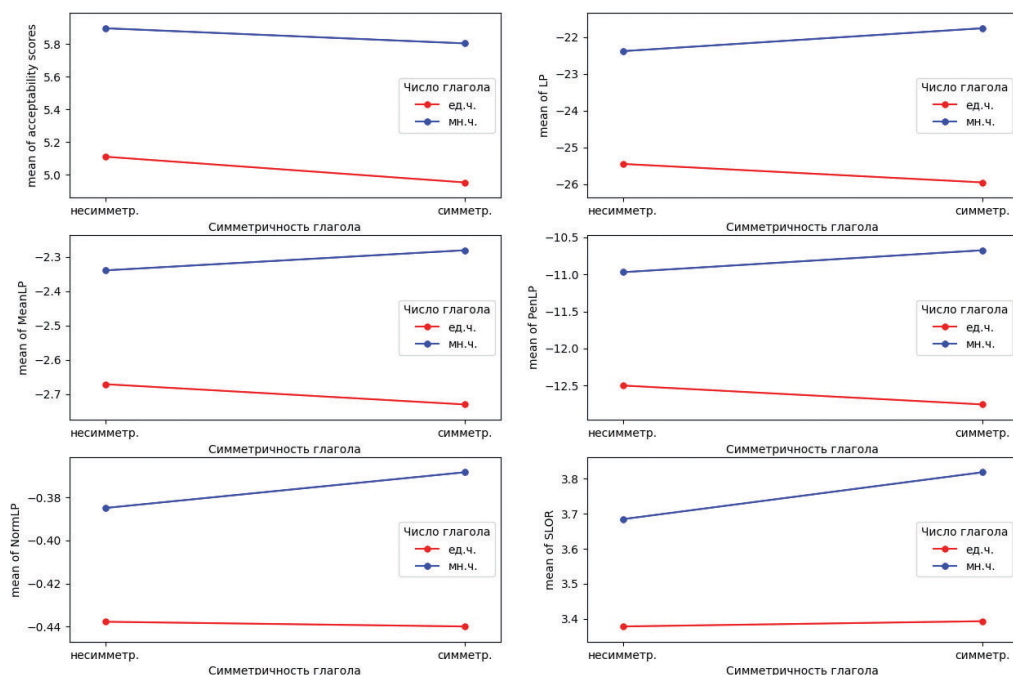


Рис. 4. Средние значения оценок приемлемости и вероятностных метрик в зависимости от симметричности предиката

значимое влияние фактора числа на приемлемость и вероятность высказываний ($p < .0001$ для оценок; $p < .001$ для LP, PenLP, NormLP и SLOR; $p < .01$ для MeanLP). Однако приемлемость и вероятность высказывания не зависят ни от симметричности предиката, ни от взаимодействия этих факторов ($p > .05$ для оценок и всех метрик).

Попарные сравнения показывают, что предложения с единственным и множественным числом глагола получают значимо разные оценки приемлемости при наличии как симметричного ($p < .0001$), так и несимметричного предиката ($p < .0001$). Предложения, которые отличаются только семантическим типом глагола, не различаются по приемлемости ни в единственном ($p > .05$), ни во множественном числе ($p > .05$). Вероятность предложений с симметричными предикатами значимо отличается при единственном и множественном числе глагола ($p < .05$ для всех метрик). Остальные пары условий значимо не различаются по вероятности ($p > .05$ для всех метрик). Таким образом, симметричность предиката не оказывает влияния ни на приемлемость, ни на вероятность той или иной стратегии согласования. Для симметричных предикатов выявляется большая вероятность и большая приемлемость множественного числа. Аналогичный контраст в приемлемости для несимметричных предикатов не подкрепляется при сравнении значений вероятности.

6.3. Стратегии согласования для местоименных конъюнктов

Рассмотрим различные стратегии согласования для местоименных конъюнктов. На рис. 5 изображены средние значения оценок приемлемости и вероятностных метрик при прямом порядке подлежащего и сказуемого. Дисперсионный анализ оценок выявил значимое влияние фактора стратегии согласования ($p < .0001$), однако порядок местоимений

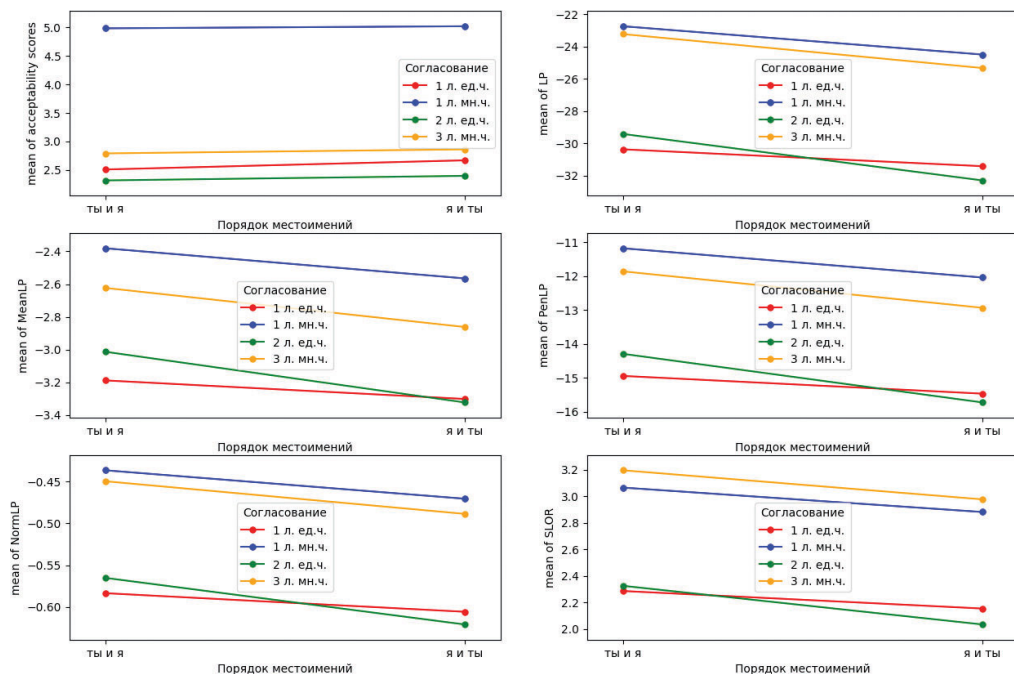


Рис. 5. Средние значения оценок приемлемости и вероятностных метрик для местоимений при порядке слов SV

и взаимодействие факторов оказываются не значимы ($p < .05$). В то же время на вероятность значимо влияют оба фактора: и форма глагола ($p < .0001$ для всех метрик), и тип подлежащего ($p < .01$ для MeanLP, PenLP, NormLP и SLOR; $p < .05$ для LP), тогда как взаимодействие факторов не значимо ($p < .05$ для всех метрик).

Попарные сравнения оценок приемлемости не выявили значимых различий между предложениями, которые отличаются только порядком конъюнктов при фиксированном типе согласования ($p > .05$ для 1 л. мн. ч., 1 и 2 лиц ед. ч., 3 л. мн. ч.). Вне зависимости от порядка местоимений, приемлемость предложений, где глагол принимает форму 1 л. ед. ч., значимо отличается от примеров с остальными типами согласования ($p < .0001$ при попарных сравнениях с формой 1 л. ед. ч., 2 л. ед. ч. и 3 л. мн. ч.). Также значимое различие в приемлемости наблюдается между согласованием по 2 л. ед. ч. и 3 л. мн. ч. для обоих порядков конъюнктов ($p < .01$). Приемлемость остальных стратегий отличается незначимо при любом типе подлежащего ($p > .05$ для форм 1 л. ед. ч. в сравнении с формами 2 л. ед. ч. и 3 л. мн. ч.).

В отношении вероятностных метрик значимые различия между примерами с различным порядком конъюнктов также не наблюдаются ($p > .05$ для всех метрик для каждой из глагольных форм). Для обоих типов сочиненных подлежащих формы 1 л. мн. ч. значимо более вероятны, чем формы 1 л. ед. ч. (*ты и я*: $p < .0001$ для всех метрик; *я и ты*: $p < .0001$ для NormLP, $p < .001$ для LP, MeanLP, PenLP и SLOR) и чем формы 2 л. ед. ч. (*ты и я*: $p < .0001$ для NormLP и SLOR, $p < .001$ для LP и PenLP, $p < .01$ для MeanLP; *я и ты*: $p < .0001$ для всех метрик). Отсутствует значимое различие в вероятности согласования по 1 л. и по 2 л. в ед. ч. (*ты и я*: $p > .05$ для всех метрик; *я и ты*: $p > .05$ для всех метрик), по 1 л. и 3 л. во мн. ч. (*ты и я*: $p > .05$ для всех метрик; *я и ты*: $p > .05$ для всех метрик). Вероятность предложений с формами 3 л. мн. ч. почти всегда значимо выше, чем с формами 1 л. ед. ч. (*ты и я*: $p < .0001$ для NormLP и SLOR, $p < .001$ для LP, $p < .01$ для PenLP, $p < .05$ для MeanLP; *я и ты*: $p < .0001$ для SLOR, $p < .001$ для NormLP, $p < .01$ для LP, $p < .05$ для PenLP, $p > .05$ для MeanLP) и 2 л. ед. ч. (*ты и я*: $p < .0001$ для SLOR, $p < .001$ для NormLP, $p < .01$ для LP, $p < .05$ для PenLP, $p > .05$ для MeanLP; *я и ты*: $p < .0001$ для NormLP и SLOR, $p < .001$ для LP, $p < .01$ для PenLP, $p > .05$ для MeanLP).

При порядке слов SV как приемлемость, так и вероятность той или иной стратегии согласования не зависят от расположения местоименных конъюнктов. Наиболее приемлемым и наиболее вероятным (кроме метрики SLOR) оказывается согласование по 1 лицу множественному числу, описываемое в нормативных грамматиках. Предложения с дефолтным согласованием по 3 лицу множественному числу являются значимо менее приемлемыми, однако различие в вероятности оказывается незначимым. Приемлемость дефолтного согласования незначимо отличается от приемлемости согласования по 1 лицу единственному числу, однако различие в вероятности является значимым (за исключением метрики MeanLP при подлежащем *я и ты*). Ни оценки приемлемости, ни вероятностные метрики не демонстрируют значимых различий между формами 1 и 2 лица единственного числа. Значимым оказывается различие в вероятности и приемлемости при сравнении нормативного согласования (1 лицо множественное число) с согласованием по единственному числу (1 или 2 лица). Дефолтное согласование (3 лицо множественное число) является значимо более вероятным и приемлемым, чем согласование по 2 лицу единственному числу (кроме метрики MeanLP).

Рис. 6 демонстрирует средние значения оценок приемлемости и вероятностных метрик при обратном порядке подлежащего и сказуемого. Значимое влияние на приемлемость оказывает выбор стратегии согласования ($p < .0001$), но не порядок местоимений ($p > .05$). Взаимодействие этих факторов также является значимым ($p < .0001$). Дисперсионный анализ выявил значимое влияние формы предиката только для некоторых вероятностных метрик ($p < .001$ для SLOR, $p < .05$ для NormLP, $p > .05$ для LP, MeanLP и PenLP). Значимого влияния на вероятность высказываний не оказывают порядок конъюнктов, а также взаимодействие факторов ($p > .05$ для всех метрик).

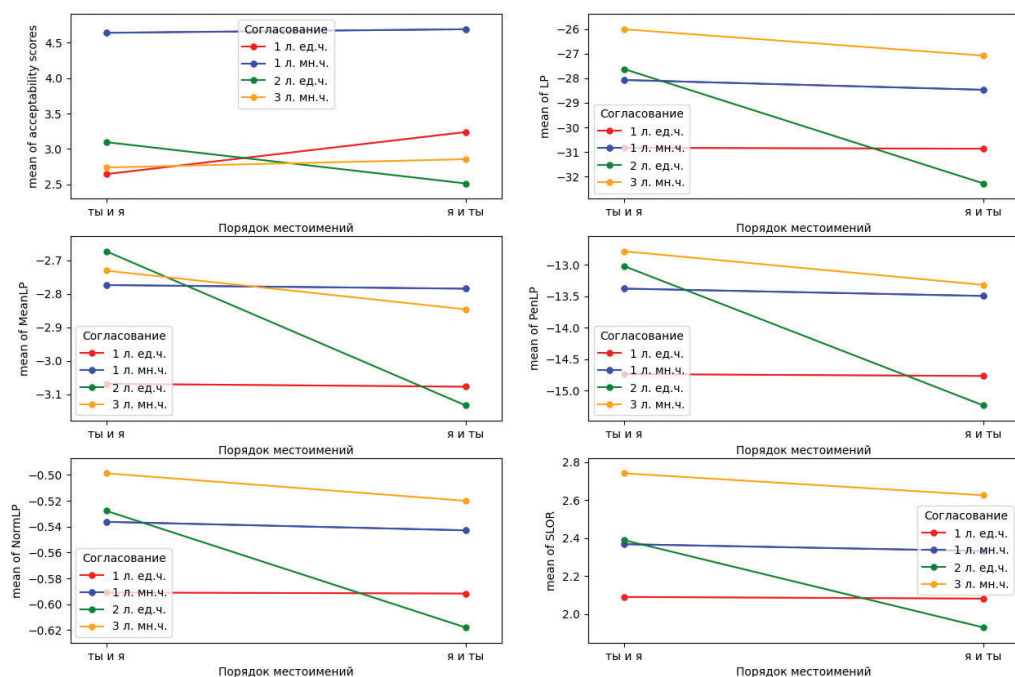


Рис. 6. Средние значения оценок приемлемости и вероятностных метрик для местоимений при порядке слов VS

Попарные сравнения демонстрируют, что приемлемость не зависит от порядка местоимений при нормативном согласовании по 1 л. мн. ч. ($p > .05$) и дефолтном согласовании по 3 л. мн. ч. ($p > .05$). Нормативное согласование все еще значимо более приемлемо, чем все остальные стратегии, при любом порядке конъюнктов ($p < .0001$ при сравнении 1 л. мн. ч. с 1 и 2 лицом ед. ч., 3 л. мн. ч.). В то же время согласование с ближайшим конъюнктом оказывается гораздо более допустимо при порядке VS. Предложения с предикатом в 1 л. и 2 л. ед. ч. значимо различаются по приемлемости (*ты и я*: $p < .01$, *я и ты*: $p < .0001$). Согласование по 2 л. ед. ч. значимо более приемлемо для подлежащего *ты и я* ($p < .0001$), а согласование по 1 л. ед. ч. — для подлежащего *я и ты* ($p < .0001$). При порядке конъюнктов *ты и я* дефолтное согласование значимо менее приемлемо, чем согласование по 2 л. ед. ч. ($p < .05$), но не отличается от согласования по 1 л. ($p > .05$). Для предложений с подлежащим *я и ты* выявляется значимая разница в приемлемости 3 л. мн. ч. по сравнению с ед. ч. 1 л. ($p < .05$) и 2 л. ($p < .05$). В отличие от оценок приемлемости, единственная пара условий, для которой различие в вероятности оказалось значимым, — 2 л. ед. ч. и 3 л. мн. ч. для подлежащего *я и ты* ($p < .05$ для SLOR; $p > .05$ LP, MeanLP, PenLP и NormLP). При попарном сравнении остальных условий вероятностные метрики не демонстрируют значимых контрастов ($p > .05$ для всех метрик).

В двух экспериментах с местоименными конъюнктами были использованы стимульные предложения с одинаковым лексическим содержанием, которые различались только порядком подлежащего и сказуемого (SV / VS). Изменение порядка при сохранении лексического материала оказывает сильное влияние на то, как соотносятся оценки приемлемости и вероятностные метрики. При порядке SV ряд контрастов, выявленных для оценок приемлемости, также демонстрируются при попарном сравнении вероятностных метрик. Однако при порядке VS почти ни для одного из условий, которые различаются по приемлемости, не наблюдается значимого различия по вероятности.

6.4. Согласование с подлежащим, содержащим местоимение и существительное

Рис. 7 демонстрирует средние значения оценок приемлемости и вероятностных метрик для подлежащего, выраженного местоимением 1 лица единственного числа (*я*) и именем собственным (*Х*). Все предложения имели одинаковый порядок слов (VS), но различались порядком конъюнктов и формой глагола. Дисперсионный анализ оценок приемлемости выявил значимое влияние стратегии согласования ($p < .0001$), типа подлежащего ($p < .01$), а также взаимодействия этих факторов ($p < .0001$). В отношении вероятностных метрик значимой оказывается только форма предиката ($p < .001$ MeanLP и PenLP; $p < .01$ LP, NormLP и SLOR), вероятность не зависит от порядка конъюнктов и взаимодействия факторов ($p > .05$ для всех метрик).

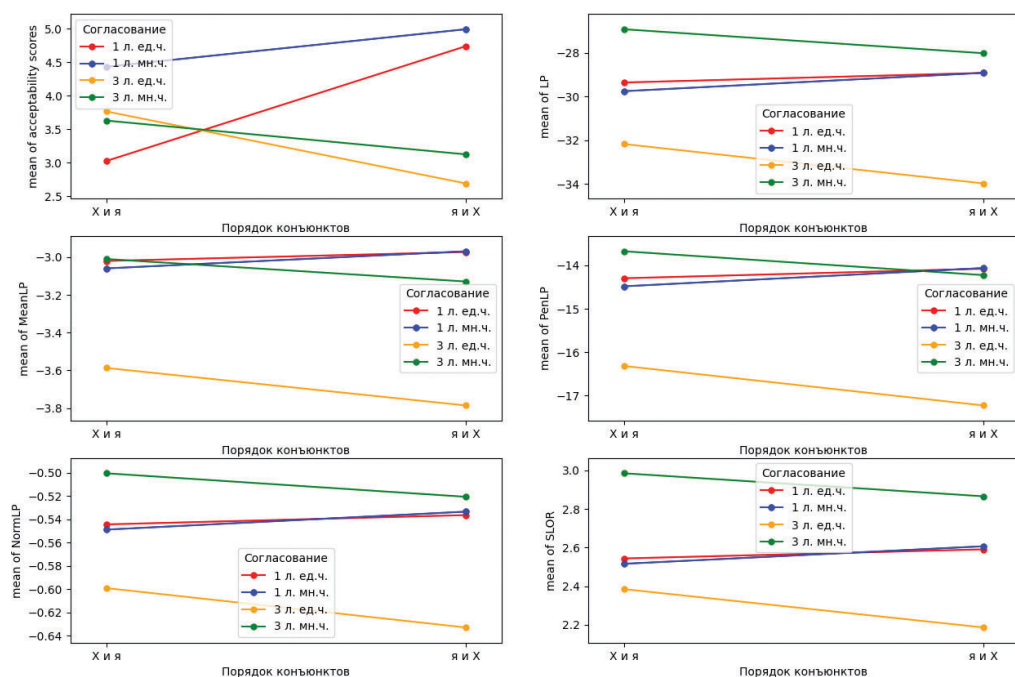


Рис. 7. Средние значения оценок приемлемости и вероятностных метрик для подлежащего *я* и *Х* при порядке слов VS

Попарные сравнения оценок приемлемости показывают, что предложения с различным типом подлежащего значимо отличаются друг от друга ($p < .0001$ для 1 л. и 3 л. ед. ч., $p < .001$ для 1 л. мн. ч., $p < .01$ для 3 л. мн. ч.). При любом порядке конъюнктов нормативная стратегия согласования (1 л. мн. ч.) значимо отличается от согласования по 3 л. ед. ч. (*Х и я*, *я и Х*: $p < .0001$) и 3 л. мн. ч. (*Х и я*, *я и Х*: $p < .0001$). Для подлежащего *я и Х* согласование по 1 л. мн. ч. и 1 л. ед. ч. одинаково приемлемы ($p > .05$), для подлежащего *Х и я* приемлемость нормативной стратегии значимо выше ($p < .0001$). Приемлемость согласования по 1 л. ед. ч. при подлежащем *Х и я* значимо ниже, чем по 3 л. ед. ч. ($p < .0001$) и 3 л. мн. ч. ($p < .0001$). Для порядка конъюнктов *я и Х* согласование по 1 л. ед. ч., наоборот, значимо выше, чем по 3 л. ед. ч. ($p < .0001$) и 3 л. мн. ч. ($p < .0001$). Предложения с предикатом 3 лица при подлежащем *Х и я* занимают среднюю позицию по приемлемости вне зависимости от числа ($p > .05$). При подлежащем *я и Х* приемлемость 3 л. ед. ч. значимо ниже, чем

3 л. мн. ч. ($p < .05$). Парные сравнения вероятностных метрик выявили отличия только для двух пар условий. Предложения с формой 3 л. ед. ч. при подлежащем *я и X* оказываются значимо менее вероятны, чем с формой 1 л. ед. ч. ($p < .05$ для MeanLP; $p > .05$ для LP, PenLP, NormLP и SLOR) и 1 л. мн. ч. ($p < .05$ для MeanLP; $p > .05$ для LP, PenLP, NormLP и SLOR). Различия в вероятности между остальными условиями являются незначимыми ($p > .05$ для всех метрик).

Вероятностные метрики выявляют низкую приемлемость согласования по 3 лицу единственному числу, однако наибольшая приемлемость нормативного согласования (по 1 лицу множественному числу) не подтверждается высокой вероятностью. Кроме того, предложения с различным порядком конъюнктов отличаются по приемлемости, но не по вероятности.

6.5. Сравнение результатов для разных вероятностных метрик

Результат сравнения оценок приемлемости и значений вероятности представлен в табл. 10. Для всех 64 пар условий было посчитано количество истинных и ложных ответов. Истинно положительными исходами считались те, где контраст выявляется как для приемлемости, так и для вероятности. Истинно отрицательный исход наблюдается, если контраст отсутствует и для приемлемости, и для вероятности. Если условия не различаются по приемлемости, но им присваивается разная вероятность, то такие исходы считаются ложно положительными. Пары, для которых значимое различие есть для оценок приемлемости, но не для вероятностной метрики, являются самыми частотными и названы ложно отрицательными. Чем больше истинных исходов и чем меньше ложных исходов демонстрирует вероятностная метрика, тем лучше она коррелирует с приемлемостью.

Таблица 10

Количество контрастов между парами условий для вероятностных метрик в сравнении с контрастами в приемлемости

Тип исхода	Значимое различие		LP	MeanLP	PenLP	NormLP	SLOR
	приемлемость	вероятность					
истинно положительный	+	+	10	10	10	10	11
истинно отрицательный	–	–	16	18	17	17	17
ложно положительный	–	+	3	1	2	2	2
ложно отрицательный	+	–	35	35	35	35	34

Наибольшее количество истинно положительных ответов и наименьшее количество ложно отрицательных ответов получено для метрики SLOR. Больше всего истинно отрицательных исходов и меньше всего ложно положительных исходов продемонстрировала метрика MeanLP. Следовательно, для более точного отражения контрастов в приемлемости с помощью вероятностных метрик необходимо учитывать количество слов и их частотность (SLOR). Однако нормализация только по длине предложения без учета частоты встречаемости слов также показывает хороший результат (MeanLP).

7. Заключение

В данной статье мы проанализировали, как связаны между собой понятия «приемлемость» и «вероятность» высказываний на материале вариативного согласования в русском языке. Мы подробно рассмотрели, какие синтаксические, морфологические и семантические факторы оказывают влияние на приемлемость высказываний с точки зрения носителей языка, а какие — на вероятность, предсказываемую языковыми моделями. В качестве языкового феномена было выбрано предикативное согласование с сочиненным подлежащим. Использование данных синтаксических экспериментов позволило проанализировать влияние каждого фактора по отдельности, а также их взаимодействия.

Мы оценивали способность двунаправленной модели ruBERT-base моделировать вариативное согласование. Было использовано пять различных вероятностных метрик: логарифмическая вероятность предложения без нормализации (LP), с нормализацией по длине (MeanLP), с нормализацией по длине с учетом коэффициента масштабирования (PenLP), с учетом униграммной вероятности (NormLP), с нормализацией по длине и учетом униграммной вероятности (SLOR).

Ожидалось, что модель будет предсказывать наибольшую вероятность стратегии, которая является наиболее приемлемой с точки зрения носителей языка. Однако предположение не подтвердилось. Для примеров без вариативного согласования вероятностные метрики более точно отражают различия в приемлемости. При наличии вариативного согласования с сочиненным подлежащим корреляция между приемлемостью и вероятностью падает.

Вероятностные метрики наиболее точно отражают контрасты в приемлемости при изменении линейного порядка подлежащего и сказуемого. Согласование по единственному числу оказывается более вероятно и приемлемо при порядке VS, тогда как приемлемость и вероятность множественного числа не зависят от линейной позиции предиката.

Для именных конъюнктов значимые различия в вероятности не выявляются, если приемлемость зависит от морфологических характеристик подлежащего. Согласование по единственному числу более приемлемо при совпадении рода конъюнктов, однако на вероятность высказываний этот фактор не влияет. Семантический тип конъюнктов и предиката не оказывает влияния ни на приемлемость, ни на вероятность высказываний.

Для местоименных конъюнктов при порядке SV наиболее приемлемым и вероятным оказывается нормативное согласование (1 л. мн. ч.). Дефолтной стратегии (3 л. мн. ч.) приписывается слишком высокая вероятность при среднем уровне приемлемости. При порядке VS повышается приемлемость согласования с ближайшим конъюнктом (2 л. ед. ч. для *ты* и *я*, 1 л. ед. ч. для *я* и *ты*), однако эта тенденция не выявляется при анализе значений вероятности.

Если подлежащее выражено местоимением *я* и именем собственным, при порядке VS наиболее приемлемо согласование по 1 л. мн. ч. Также растет приемлемость согласования с ближайшим конъюнктом по 1 л. ед. ч. при подлежащем *я* и *Х*. Предсказанные языковой моделью значения вероятности не отражают эти закономерности.

При наличии вариативного согласования приемлемость и вероятность высказывания в значительной степени зависят от линейного расположения элементов предложения: позиции подлежащего и сказуемого, порядка конъюнктов. Морфологические факторы оказывают влияние на выбор наиболее приемлемой стратегии согласования, однако не являются значимыми при сравнении вероятности высказываний. Семантические особенности подлежащего и сказуемого не вызывают различий в приемлемости и вероятности. Наилучший результат показали метрики MeanLP и SLOR.

Можно сделать вывод, что при обучении на большом количестве размеченных текстовых данных двунаправленная языковая модель ruBERT не усваивает ряд важных аспектов грамматики и не может быть использована для моделирования вариативного согласования

сказуемого с сочиненным подлежащим без тонкой настройки. Понятие «вероятность» лишь опосредованно связано с понятием «приемлемость»: суждения о приемлемости отражают более тонкие контрасты при выборе стратегии согласования, которые не выявляются с помощью предсказываемых моделью значений вероятности.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Белова 2022 — Белова Д. Д. Предикативное согласование с подлежащим, выраженным местоименными конъюнктами. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Belova D. D. Predicative agreement with the subject expressed by pronominal conjuncts. *Talk at the student conference “Experimental language studies”, Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Белова, Давидюк 2023 — Белова Д. Д., Давидюк Т. И. Согласование с сочиненным подлежащим, содержащим личное местоимение: экспериментальное исследование на материале русского языка. *Rhema. Рема*, 2023, 2: 53–88. [Belova D. D., Davidyuk T. I. Agreement with coordinated subjects containing a personal pronoun: Experimental data from Russian. *Rhema*, 2023, 2: 53–88.]
- Врубель 2022 — Врубель Д. Д. Синкретичное согласование по роду в конструкциях с повторяющимся союзом *и*. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Vrubel D. D. Syncretic gender agreement in constructions with a repeated conjunction *i*. *Talk at the student conference “Experimental language studies”, Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Врубель 2023 — Врубель Д. Д. Эффект синкретизма при предикативном согласовании с сочинительными конструкциями с повторяющимся союзом *и*. *Rhema. Рема*, 2023, 2: 104–118. [Vrubel D. D. The role of syncretism in the predicate agreement with coordinate constructions with the Russian correlative conjunction *i...i*. *Rhema*, 2023, 2: 104–118.]
- Граудина и др. 2001 — Граудина Л. К., Ицкович В. А., Катлинская Л. П. *Грамматическая правильность русской речи: Опыт частотно-стилистического словаря вариантов*. М.: Наука, 2001. [Graudina L. K., Itskovich V. A., Katlinskaya L. P. *Grammatical correctness of speech: Towards a frequency and stylistic dictionary of variants*. Moscow: Nauka, 2001.]
- Давидюк 2022 — Давидюк Т. И. Влияние аргументной структуры предиката на лично-числовое согласование с сочиненным подлежащим. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Davidyuk T. I. The influence of argument structure on predicative agreement in person and number with compound subjects. *Talk at the student conference “Experimental language studies”, Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Добрушина, Сидорова 2019 — Добрушина Е. Р., Сидорова М. И. Число предиката в конструкциях типа «те/все, кто пришел/пришли» и падеж вершины-корпусное исследование. *Вестник Православного Свято-Тихоновского гуманитарного университета. Сер. 3: Филология*, 2019, 59: 22–35. [Dobrushina E. R., Sidorova M. I. Number of the predicate in phrases *te/vse, kto prishel/prishli* ‘those/all who came (sg./pl.)’ and the case of the head: A corpus-based study. *St. Tikhon’s Univ. Review. Series III: Philology*, 2019, 59: 22–35.]
- Иомдин 1990 — Иомдин Л. Л. *Автоматическая обработка текста на естественном языке: модель согласования*. М.: Наука, 1990. [Iomdin L. L. *Avtomaticheskaya obrabotka teksta na estestvennom yazyke: model’ soglasovaniya* [Automatic natural language processing: Agreement model]. Moscow: Nauka, 1990.]
- Кибрик (ред.) 1999 — Кибрик А. Е. (ред.-сост.). *Элементы цахурского языка в типологическом освещении*. М.: Наследие, 1999. [Kibrik A. E. (ed.-comp.) *Elementy tsakhurskogo yazyka v tipologicheskoy osveshchenii* [Elements of Tsakhur in a typological perspective]. Moscow: Nasledie, 1999.]
- Кувшинская 2013 — Кувшинская Ю. М. Согласование сказуемого с подлежащим, выраженным именной группой с количественным значением (по данным НКРЯ за 2000–2010 гг.). *Русский язык в научном освещении*, 2013, 26: 112–151. [Kuvshinskaya Yu. M. Predicate agreement with the quantifier phrase in Russian (according to the data of the National Russian Corpus for the period of 2000–2010). *Russkij jazyk v nauchnom osveshchenii*, 2013, 26: 112–151.]

- Лютикова, Герасимова 2023 — Лютикова Е. А., Герасимова А. А. Исследование вариативного согласования в русском языке: проблемы и методы. *Rhema. Рема*, 2023, 2: 9–27. [Lyutikova E. A., Gerasimova A. A. Studying agreement variation in Russian: Problems and methodology. *Rhema*, 2023, 2: 9–27.]
- Мельчук 1993 — Мельчук И. А. Согласование, управление, конгруэнтность. *Вопросы языкознания*, 1993, 5: 16–58. [Mel'čuk I. A. Agreement, regimen, congruence. *Voprosy Jazykoznanija*, 1993, 5: 16–58.]
- Паско 2022 — Паско Л. И. Предикативное согласование с сочиненным подлежащим и симметричность предиката. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Pasko L. I. Predicative agreement with compound subject and predicate symmetry. *Talk at the student conference "Experimental language studies", Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Паско 2023 — Паско Л. И. Против АТВ-анализа частичного согласования в русском языке: экспериментальное исследование. *Rhema. Рема*, 2023, 2: 89–103. [Pasko L. I. Against AТВ-analysis of partial agreement in Russian: An experimental study. *Rhema*, 2023, 2: 89–103.]
- Пекелис 2013а — Пекелис О. Е. «Частичное согласование» в конструкции с повторяющимся союзом: корпусное исследование основных закономерностей. *Вопросы языкознания*, 2013, 4: 55–86. [Pekelis O. E. Partial agreement with subjects linked by a correlative conjunction: A corpus-based study of main regularities. *Voprosy Jazykoznanija*, 2013, 4: 55–86.]
- Пекелис 2013б — Пекелис О. Е. Сочинение. *Материалы для проекта корпусного описания русской грамматики*. М., 2013. Рук. [Pekelis O. E. Coordination. *Materials for the project of corpus description of Russian grammar*. Moscow, 2013. Ms.] <http://rusgram.ru/Сочинение>.
- РГ 1980 — Шведова Н. Ю. (ред.). *Грамматика русского языка*. Т. 2. М.: Наука, 1980. [Shvedova N. Yu. (ed.). *Grammatika russkogo yazyka* [Russian grammar]. Vol. 2. Moscow: Nauka, 1980.]
- Розенталь и др. 1999 — Розенталь Д. Э., Джанджакова Е. В., Кабанова Н. П. *Справочник по правописанию, произношению, литературному редактированию*. М.: ЧеРо, 1999. [Rozenal D. E., Dzhandzakova E. V., Kabanova N. P. *Spravochnik po pravopisaniiyu, proiznosheniiyu, literaturnomu redaktirovaniyu* [Handbook of spelling, pronunciation, literary editing]. Moscow: CheRo, 1999.]
- Санников 2008 — Санников В. З. *Русский синтаксис в семантико-прагматическом пространстве*. М.: Языки славянских культур, 2008. [Sannikov V. Z. *Russkii sintaksis v semantiko-pragmaticheskom prostranstve* [Russian syntax in a semantic-pragmatic aspect]. Moscow: Yazyki slavyanskikh kul'tur, 2008.]
- Студеникина 2023а — Студеникина К. А. Об идентичности морфологических признаков при эллипсисе в русском языке: данные именных групп с сочиненными прилагательными. *Rhema. Рема*, 2023, 2: 28–52. [Studenikina K. A. Towards the feature identity for ellipsis in Russian: Evidence from noun phrases with coordinated adjectives. *Rhema*, 2023, 2: 28–52.]
- Студеникина 2023б — Студеникина К. А. Влияние одушевленности конъюнктов и линейной позиции сказуемого на выбор стратегии предикативного согласования. *Устный доклад на учебной конференции «Экспериментальные исследования языка», МГУ имени М. В. Ломоносова, Москва, 16 июня 2022 г.* [Studenikina K. A. The role of animacy of conjuncts and predicate linear position in the choice of predicate agreement strategy. *Talk at the student conference "Experimental language studies", Lomonosov Moscow State Univ., Moscow, June 16, 2022.*]
- Adger 2003 — Adger D. *Core syntax: A minimalist approach*. Oxford: Oxford Univ. Press, 2003.
- Arsenijević, Mitić 2016 — Arsenijević B., Mitić I. Effect of animacy and agentivity on the processing of agreement in Serbo-Croatian. *Studies in languages and mind. Selected papers from the 3rd Workshop in Psycholinguistics, Neurolinguistic and Clinical Linguistic Research*. Halupka-Rešetar S., Martínez-Ferreiro S. (eds.). Novi Sad: Univ. of Novi Sad, 2016, 41–77.
- An, Abeillé 2019 — An A., Abeillé A. Number agreement in French binomials. *Empirical Issues in Syntax and Semantics*, 2019, 12: 31–60.
- Chomsky 1956 — Chomsky N. Three models for the description of language. *IRE Transactions on Information Theory*, 1956, 2(3): 113–124.
- Chomsky 1995 — Chomsky N. *The Minimalist Program*. Cambridge (MA): MIT Press, 1995.
- Chomsky 2014 — Chomsky N. *Aspects of the theory of syntax*. Cambridge (MA): MIT Press, 2014.
- Corbett 1983 — Corbett G. Resolution rules: Agreement in person, number, and gender. *Order, concord and constituency*. Gazdar G., Klein E., Pullam G. K. (eds.). Berlin, Boston: De Gruyter Mouton, 1983, 175–205.
- Devlin et al. 2019 — Devlin J., Chang M. W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. *Proc. of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019, 1: 4181–4186.

- Himmelreich, Hartmann 2023 — Himmelreich A., Hartmann K. Agreement with disjoined subjects in German. *Glossa: a journal of general linguistics*, 2023, 8(1): 1–44.
- Lau et al. 2020 — Lau J. H., Armendariz C., Lappin S., Purver M., Shu C. How furiously can colorless green ideas sleep? Sentence acceptability in context. *Transactions of the Association for Computational Linguistics*, 2020, 8: 296–310.
- Lau et al. 2017 — Lau J. H., Clark A., Lappin S. Grammaticality, acceptability, and probability: A probabilistic view of linguistic knowledge. *Cognitive Science*, 2017, 41(5): 1202–1241.
- Likert 1932 — Likert R. A technique for the measurement of attitudes. *Archives of Psychology*, 1932, 22(140): 5–55.
- Marušič et al. 2015 — Marušič F., Nevins A. I., Badecker W. The grammars of conjunction agreement in Slovenian. *Syntax*, 2015, 18(1): 39–77.
- Mikhailov et al. 2022 — Mikhailov V., Shamardina T., Ryabinin M., Pestova A., Smurov I., Artemova E. RuCoLA: Russian Corpus of Linguistic Acceptability. *Proc. of the 2022 Conference on Empirical Methods in Natural Language Processing*. Goldberg Y., Kozareva Z., Zhang Y. (eds.). Association for Computational Linguistics, 2022, 5207–5227.
- Pereira 2000 — Pereira F. Formal grammar and information theory: together again? *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 2000, 358(1769): 1239–1253.
- Radford et al. 2018 — Radford A., Narasimhan K., Salimans T., Sutskever I. *Improving language understanding by generative pre-training*. Technical Report, OpenAI, 2018.
- Sprouse 2015 — Sprouse J. Three open questions in experimental syntax. *Linguistics Vanguard*, 2015, 1(1): 89–100.
- Sprouse et al. 2013 — Sprouse J., Schütze C. T., Almeida D. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Lingua*, 2013, 134: 219–248.
- Sprouse et al. 2018 — Sprouse J., Yankama B., Indurkha S., Fong S., Berwick R. C. Colorless green ideas do sleep furiously: gradient acceptability and the nature of the grammar. *The Linguistic Review*, 2018, 35(3): 575–599.
- Sprouse, Schütze 2014 — Schütze, C., Sprouse, J. Judgment data. *Research methods in linguistics*. Sharma D., Podesva R. (eds.). Cambridge: Cambridge Univ. Press, 2014, 27–50.
- Studenikina 2023 — Studenikina K. A. Parametrizing number variation in Russian noun phrases with experimental studies and language modeling. *Research Result. Theoretical and Applied Linguistics*, 2023, 9(1): 192–205.
- Vaswani et al. 2017 — Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017, 30: 5998–6008.
- Warstadt et al. 2020 — Warstadt A., Parrish A., Liu H., Mohananey A., Peng W., Wang S. F., Bowman S. R. BLiMP: The benchmark of linguistic minimal pairs for English. *Transactions of the Association for Computational Linguistics*, 2020, 8: 377–392.
- Zwicky 1977 — Zwicky A. Hierarchies of person. *Papers from the 13th Regional Meeting, Chicago Linguistics Society*. Beach W. A., Fox S. E., Philosoph Sh. (eds.). Chicago: Chicago Linguistics Society, 1977, 714–733.

Референциальный выбор в устных и письменных рассказах: сопоставительное исследование на материале корпуса «Веселые истории из жизни»

© 2024

Марина Андреевна Шумилина

Московский государственный университет имени М. В. Ломоносова, Москва, Россия;
Институт языкознания РАН, Москва, Россия; mari.an.shum@iling-ran.ru

Аннотация: Исследование посвящено сравнению референциального выбора в устном и письменном дискурсе на русском языке. Референциальный выбор рассматривается в виде тройного противопоставления полных именных групп, местоимений и нулевых именных групп. Работа выполнена на материале дискурсов, каждый из которых изложен рассказчиком дважды — в устной и письменной формах. В рассказах все именные группы были идентифицированы и охарактеризованы по 29 параметрам. На собранных выборках были обучены модели мультиномиальной логистической регрессии и деревья решений, на основе деревьев решений были построены диаграммы значимости факторов. Интерпретация моделей и диаграмм показывает, что некоторые факторы референциального выбора по-разному проявляют себя в устных и письменных дискурсах, например грамматическая роль, семантическая гиперроль и неполная кореферентность между анафором и антецедентом. Также модели демонстрируют, что наборы факторов, значимых для двух выборок, неодинаковы: в частности, одушевленность референта и семантическая гиперроль анафора присутствуют в дереве решений только для письменного дискурса.

Ключевые слова: дискурс, референция, русский язык, устный дискурс

Благодарности: Автор выражает искреннюю благодарность А. А. Кибрику и двум анонимным рецензентам за полезные замечания по структуре и стилю текста, а также Г. Доброву, А. Большой и К. Студеникиной за консультации по предварительной обработке данных, подбору и применению методов машинного обучения.

Для цитирования: Шумилина М. А. Референциальный выбор в устных и письменных рассказах: сопоставительное исследование на материале корпуса «Веселые истории из жизни». *Вопросы языкознания*, 2024, 6: 133–159.

DOI: 10.31857/0373-658X.2024.6.133-159

Referential choice in spoken and written stories: A comparative study based on the corpus “Funny life stories”

Marina A. Shumilina

Lomonosov Moscow State University, Moscow, Russia;
Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia; mari.an.shum@iling-ran.ru

Abstract: The study is concerned with referential choice in spoken and written discourse in the Russian language. I consider referential choice to consist in a threefold opposition between full noun phrases, pronouns, and zero noun phrases. The study is based on discourses each of which was presented by its narrator twice, namely in the spoken and written forms. In each story, all the noun phrases were identified and described according to 29 parameters. I trained logistic regression models and decision trees on the collected samples and analyzed factor importance diagrams built on the basis of the decision trees. The interpretation of the models and diagrams shows that some factors have different impact

on referential choice in spoken and written discourses, for instance, grammatical role, semantic hyperrole and sloppy identity between the anaphor and the antecedent. Besides, the models also demonstrate that the sets of significant factors for the two samples are not identical: in particular, the referent's animacy and the anaphor's semantic hyperrole are present solely in the decision tree for written discourse.

Keywords: discourse, oral discourse, reference, Russian

Acknowledgements: I am very grateful to A. A. Kibrik and the two anonymous reviewers for their extremely helpful suggestions and remarks on this work, as well as to G. Dobrov, A. Bolshina, and K. Studenikina for their recommendations on preliminary data processing, and also choice and application of machine-learning algorithms.

For citation: Shumilina M. A. Referential choice in spoken and written stories: A comparative study based on the corpus "Funny life stories". *Voprosy Jazykoznanija*, 2024, 6: 133–159.

DOI: 10.31857/0373-658X.2024.6.133-159

Введение

Референциальный выбор — это выбор, совершаемый говорящим в процессе порождения дискурса: для каждого референта, который говорящий хочет упомянуть, он должен выбрать план выражения. Например, один и тот же объект можно назвать *розовая стеклянная ваза*, *ваза*, *она* или *это*, а в некоторых случаях упоминание может быть фонологически не выражено, как в дееспричастном обороте во фразе *ваза стояла на подоконнике и сияла, пропуская сквозь себя лучи света*.

Явление референциального выбора имеет достаточно длительную историю изучения как в синтаксисе, так и в дискурсивном анализе. В качестве примеров широко известных подходов к его моделированию можно назвать теорию центрирования [Grosz et al. 1986], теорию доступности [Ariel 1990] и теорию когнитивных статусов [Gundel et al. 1993].

В 1996 году А. А. Кибрик в статье [А. А. Кибрик 1996] предложил моделировать референциальный выбор в когнитивном ключе. Это означает, что в процессе референциального выбора решающей силой признаются когнитивные процессы внимания и активации в рабочей памяти говорящего [А. А. Кибрик 2011: 377], а также их грамматические и дискурсивные соответствия, например одушевленность и линейное расстояние в абзацах. Помимо этого, в рамках данного подхода каждый фактор получает количественную оценку того, насколько существенно его влияние на результат референциального выбора.

В когнитивной количественной модели референциальный выбор представлен как классический многофакторный процесс [А. А. Кибрик 1997: 94], моделирование которого состоит из нескольких этапов:

- 1) Сначала необходимо сформировать набор значимых факторов и подобрать числовые веса для значений каждого фактора.
- 2) Далее для каждого конкретного случая референциального выбора суммируются веса его факторных значений; полученное число отражает степень активации референта в данной точке дискурса и именуется коэффициентом активации.
- 3) После этого следует проинтерпретировать коэффициент с помощью специальной шкалы: при малых значениях выбор производится в пользу полной именной группы, при больших — в пользу местоимения или нулевой именной группы, промежуточные значения предписывают использовать полную ИГ либо местоимение [Там же: 101].
- 4) Заключительный функциональный элемент модели — набор фильтров. Фильтр — это условие, которое принуждает говорящего использовать полную ИГ вне зависимости от степени активации референта (подробнее см. п. 3.2).

Настоящая работа посвящена референциальному выбору в рассказах на русском языке. Исследование выполнено на основе 11 рассказов из корпуса «Веселые истории из жизни»,

далее ВИЖ¹ (подробнее см. п. 2.1). Материал рассматривается в рамках когнитивного количественного подхода.

Цель исследования — выявить различия в процессах референциального выбора (далее РВ) в рассказах, которые относятся к разным модусам. Под модусом в данной работе понимается канал передачи информации [Fox 1995; А. А. Кибрик 2011: 11–12]. Предлагается говорить о двух основных модусах — устном и письменном. Сравнение устных и письменных дискурсов представляется интересным направлением в свете того, что в современном языкознании еще сильна идея первостепенной важности письменного модуса: в частности, закономерности референциального выбора в письменном дискурсе могут быть неосторожно экстраполированы на устный дискурс. Как показано в [Fox 1995: 136–152], тип дискурса влияет на закономерности анафоры: автор исследует спонтанные устные диалоги и короткие тексты публицистического стиля, тем самым анализируя выборки из различных модусов.

В настоящей работе выдвинуты четыре гипотезы:

- 1) Существует **качественное** отличие в наборе факторов, значимых для РВ в устном и письменном модусах дискурса.
- 2) Для некоторых факторов, значимых в обоих модусах, существует **количественное** отличие между степенью влияния на РВ в каждом из модусов.
- 3) Для обоих модусов верно, что риторическое расстояние оказывает большее влияние на РВ, чем линейное расстояние².
- 4) Помимо факторов, которые уже учитывались в других машинных моделях РВ (например, в [Strube, Wolters 2002] и [van Vliet 2012]), в данной работе делается попытка ввести несколько теоретически значимых факторов, которые в этих моделях представлены не были. Такие факторы предлагается называть **экспериментальными**, большинство из них помогает более детально учитывать синтаксический контекст упоминания. Предполагается, что все они продемонстрируют значимость на материале исследования. Более подробно об этих факторах рассказано в п. 3.1.

Гипотезы 1, 2 и 3 мотивированы сопоставительным характером исследования. В основе гипотезы 4 заложено предположение, что для моделирования референциального выбора может быть полезно принимать во внимание свойства локального контекста упоминания.

Все ключевые материалы работы — размеченные дискурсы, их риторические деревья, база данных в табличной форме — доступны в облачном хранилище по ссылке: <https://clck.ru/3EUWN7>.

Статья организована следующим образом. В разделе 1 введены и объяснены ключевые понятия работы. В разделе 2 описаны сбор и разметка данных для моделей. Раздел 3 посвящен построению и интерпретации моделей, в его рамках отдельное внимание уделяется факторам (п. 3.1), фильтрам (п. 3.2) и данным (п. 3.3), на основе которых обучаются модели; устройству и анализу моделей посвящен п. 3.4. В разделе 4 приводится обсуждение полученных результатов. Заключительный раздел содержит краткие выводы и перспективы данного направления исследований.

¹ Сокращение принято вслед за [Буденная 2018]. Корпус доступен по ссылке: <http://spokencorpora.ru/showcorpus.py?dir=02funny>.

² Существует большое разнообразие способов измерить линейное расстояние между рассматриваемым упоминанием и его ближайшим линейным antecedentом. Так, в исследовании [Loukachevitch et al. 2011: 522] предложены измерения в словах, предложениях, абзацах; в статье [Strube, Wolters 2002] вводится специальная единица **major clause unit**. Кроме того, в [Carlson et al. 2003] был предложен способ выделения ЭДЕ (**элементарной дискурсивной единицы**) для письменных текстов. В настоящей работе линейное расстояние измерялось в ЭДЕ по концепции, изложенной в [А. А. Кибрик и др. 2009а]; см. подробнее в п. 2.2.

1. Ключевые понятия и явления

1.1. Основные термины

Прежде всего, следует определить понятия **референта** и **референтности именной группы**. В настоящей работе приняты определения, предложенные в книге [А. А. Кибрик 2011]. Референт — это понятие в сознании говорящего и адресата, соответствующее некоторому объекту во внеязыковом мире, реальном или сконструированном. Именная группа считается референтной, если имеет один из субстантивных референциальных статусов, конкретный либо неконкретный (подробнее о референциальных статусах см. [Ibid. 32]).

В рамках данной работы рассматривается именная референция, поэтому считается, что референт не может иметь предикатный или автонимный статус. Вслед за [А. А. Кибрик 1997] референты, упомянутые более чем однократно, называются **значимыми**.

Чтобы ввести еще два термина, обратимся к рабочей классификации ИГ, которая отражает некоторые свойства референта (рис. 1). Она разработана специально для данного исследования.

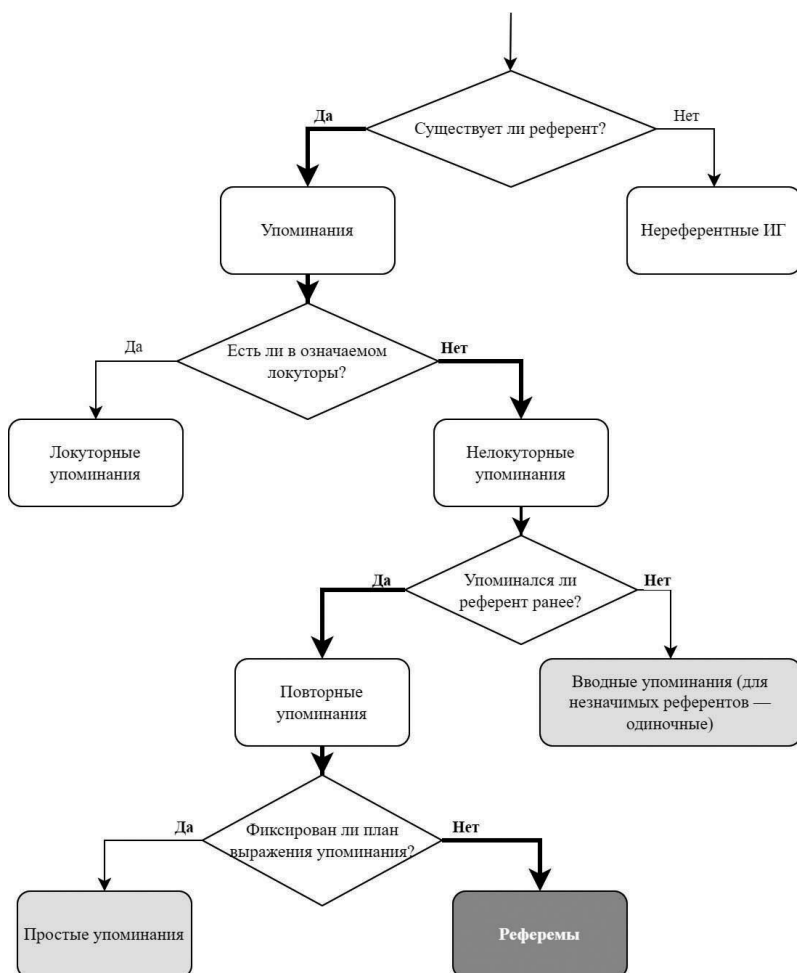


Рис. 1. Классификация именных групп

Все референтные ИГ предлагается называть **упоминаниями**. Если референт включает кого-либо из коммуникантов, то упоминание называется **локуторным** (*locutor reference* в [А. А. Кибрик 2011: 43]). **Нелокуторные** упоминания подразделяются на **вводные** (с ними также группируются **одиночные**) и **повторные**. Если говорящий должен самостоятельно выбрать план выражения для повторного упоминания, то оно является **реферемой**.

Реферема — одно из центральных понятий для данной работы. В качестве примера рассмотрим фрагмент устного рассказа из ВИЖ³, сегментированный на элементарные дискурсивные единицы (далее ЭДЕ):

- (1) FS_04-f_sp // № 4sp⁴:
 27. *моя подружка —*
 28. *Кристина,*
 29. *смотрит,*
 30. *а там у него цветы всякие в горшках,*
 31. *ну и мы спросили 0,*
 32. *мол это е = I он сам сажает 0.*

В ЭДЕ 32 при помощи референциального нуля повторно упоминаются цветы. В данном примере можно постулировать нуль, так как глагол «сажать» является переходным и предполагает наличие прямого объекта, при этом объект в предложении не выражен. Рассказчица имела возможность упомянуть их с помощью другого **референциального средства** (далее РС). В (2) показано несколько сконструированных альтернативных вариантов этой реферемы:

- (2) 32(а). *мол это е = I он сам их сажает,*
 32(б). *мол это е = I он сам сажает их,*
 32(в). *мол это е = I он сам эти цветы сажает,*
 32(г). *мол это е = I он сам сажает эти цветы.*

Таким образом, **реферема** — это нелокуторное повторное упоминание, для которого говорящий должен был выбрать план выражения.

Если упоминание обладает фиксированным планом выражения, то оно называется **простым**⁵. Другими словами, говорящий вынужден использовать то или иное РС в силу недискурсивных факторов (которые часто относятся к сфере грамматики). В качестве примера нулевого простого упоминания можно назвать подлежащее при инфинитивном обороте:

- (3) FS_17-f_sp // № 9sp:
 7. *Как потом оказалось,*
 8. *в день свадьбы —*
 9. *нашей с ним,*
 10. — *его мама хотела 0 взять <его | с собой> паспорт в ЗАГС.*

³ Здесь и далее каждый фрагмент из ВИЖ начинается с идентификатора рассказа, откуда он был взят. Слева от двух косых черт указан идентификатор в корпусе, справа — идентификатор в выборке в рамках исследования. Полный список дискурсов корпуса с их идентификаторами для настоящей работы приведен в табл. 1.

Каждая новая строка в примере — самостоятельная ЭДЕ. Нумерация ЭДЕ сохраняется от начала целого рассказа. Сохраняется исходная дискурсивная разметка. Кроме того, во всех примерах (не только из ВИЖ) ИГ, представляющие интерес, отмечены жирным шрифтом, а нулевые ИГ обозначаются как 0.

⁴ Сохранена транскрипция, представленная в корпусе.

⁵ Отсутствие стандартной процедуры референциального выбора сближает простые упоминания с ограничениями — классом явлений, который рассмотрен в статье [Залманов, Кибрик 2023].

Вводное упоминание и кореферентные ему реферемы вместе составляют **референциальную цепочку**, далее **РЦ** (термин «referential chain» широко используется в исследованиях референции, в частности в [Loukachevitch et al. 2011]).

1.2. Реферемы и простые упоминания

В работе уделено особое внимание проведению границы между реферемами и простыми упоминаниями — это было необходимо, чтобы данные в выборках были однородными, что важно для качества машинных моделей. Отдельную сложность представляла работа с нулевыми повторными упоминаниями. С опорой на классификацию, изложенную в [А. А. Kibrik 1996: 262], для работы был составлен следующий список нулевых реферем:

- Подлежащие:
 - нулевое подлежащее независимой клаузы;
 - нулевое подлежащее придаточного финитного предложения, например *Он позволит вам, когда **0** закончит работу.*
- Прочие:
 - нулевое прямое дополнение (аккузативное или генитивное при отрицании);
 - нулевой дативный принципал (подробнее о системе семантических ролей см. п. 3.1) при категории состояния и модальных предикатах;
 - нулевая вершина с адъективными модификаторами (включая числительные).

Одно из ключевых отличий простых упоминаний от реферем — изначальная безальтернативность референциального средства. Типы нулевых ИГ, не попавшие в список выше, были недостаточно представлены в проанализированных данных, поэтому не было возможности проверить их на наличие вариативности РС (то есть определить, возможно ли отнести их к реферемам). По этой причине в данной работе они отнесены к простым упоминаниям.

Кроме того, некоторые референтные упоминания полностью исключаются из рассмотрения. Во-первых, личные местоимения локуторов, которые затруднительно считать анафорическими потому, что РС упоминаний говорящего и адресата регулируются не их antecedентами, а степенью влияния на дискурс каждого из коммуникантов [А. А. Kibrik 2011: 43]. Разумно предположить, что вариативность локуторных упоминаний регулируется иными механизмами, чем вариативность прототипических реферем. Во-вторых, это референция непредметного типа. Состояния, события, а также локативная и темпоральная референция обладают особыми наборами РС⁶:

- (4) Локация: {На улице Менжинского / Там / В том месте} *никогда не было театра.*
- (5) Время: {В 2020-м году / Тогда / В то время} *занятия в школах проходили дистанционно.*
- (6) Событие: {Сборка мебели / Она / Это} *не займет много времени.*

Похожая специфика отмечается у количественных групп: они обладают расширенным набором РС, что затрудняет их анализ. Разнообразие способов упомянуть такой референт показано в примерах (7)–(10):

- (7) *Я училась в этой школе **пять лет**. Они прошли незаметно.*
- (8) ***Десять лет** назад я училась в этой школе. **Столько лет** прошло с тех пор!*

⁶ Примеры, при которых не указан источник, являются сконструированными.

- (9) Для исследования было нужно **тридцать участников**, но **столько их** было найти негде, поэтому пришлось опрашивать знакомых по несколько раз.
- (10) За нарушение штраф **пять тысяч рублей**. Но у нас не было с собой **столько 0**.

В рассмотренных дискурсах количественные группы встречаются достаточно редко. Имена существительные в составе предложений, маркирующих начало и конец рассказа, не расцениваются как упоминания, ср.: ...у меня тут на работе такая странная **история** произошла (FS_13-f_sp // № 6sp); В этой **истории** я была главная звезда (FS_16-f_sp // № 9sp); Вот такая вот **история** (FS_02-f_sp // № 2sp).

Исключаются любые местоимения, кроме личных, указательных, притяжательных и возвратных. В прямой речи не рассматриваются обращения. Как уже было сказано в п. 1.1, исключены из рассмотрения ИГ, референты которых обладают предикатным и автономным денотативными статусами.

Кроме того, несколько классов упоминаний рассматриваются исключительно как антецеденты реферем:

- возвратные местоимения;
- ИГ, при которых применяется один из фильтров (см. п. 3.2);
- вынесенные топики;
- нули при сочинительном сокращении;
- локуторные упоминания референтов в цитации, которые вне цитации не являются локуторами, например:

- (11) FS_5-m_sp // № 11sp:
 22. пятого числа **папа** доложил,
 23. и шестого **0** пошел **0** бегать по всяким авиакассам,
 24. **0** объяснять,
 25. что «Вот,
 26. тыры-пыры,
 27. **мне** надо **0** уехать.»,
 28. ну **ему** говорят

В процитированном фрагменте диалога (ЭДЕ 25–28) папа рассказчика выступает как говорящий, и в разборе рассказов подобные случаи учитываются как простые упоминания. Такое решение принято в силу эмпирической закономерности: если считать линейным антецедентом реферемы в ЭДЕ 28 упоминание в ЭДЕ 24 (то есть ближайшее упоминание вне прямой речи), то линейное расстояние между анафором и антецедентом не вполне объясняет фактический выбор РС (по данным устной выборки, при линейном расстоянии 4 полные ИГ более частотны, чем местоимения).

2. Сбор материала

2.1. Корпусные дискурсы

Корпус «Веселые истории из жизни» уже применялся для исследования РВ, в частности в диссертации [Буденная 2018]. Его особенность заключается в том, что каждый информант рассказывал одну и ту же историю дважды: сначала производилась запись рассказа в устном формате, затем рассказчик излагал тот же сюжет письменно. Устные представления рассказов имеют разметку в трех степенях подробности, для данной работы взят

минимальный вариант. Проанализировано 11 сюжетов, каждый рассмотрен в обоих моду-сах; извлечено и рассмотрено 211 реферем из письменных текстов и 280 из устных. Как представляется, количество реферем, собранное из такого объема текста, составляет ми-нимум для статистических подсчетов.

В настоящей работе не ставилась цель получить выборку, сбалансированную по полу и/или возрасту рассказчиков, поэтому тексты рассматривались сплошным образом, од-нако некоторые были исключены за счет большого присутствия пространственной, вре-менной и/или количественной референции (FS_07-f и FS_09-f); большого присутствия событийной референции (FS_10-m и FS_11-m); малого числа реферем при достаточно большом количестве ЭДЕ (FS_08-f и FS_12-f); большого числа хезитаций и других рече-вых сбоев (FS_14-f).

Все рассмотренные рассказы получили новые идентификаторы. В табл. 1 приводится соответствие этих идентификаторам исходным.

Таблица 1

Полный список проанализированных дискурсов

Новый ID	Исходный ID	Новый ID	Исходный ID
1	FS_01-f	6	FS_13-f
2	FS_02-f	7	FS_15-m
3	FS_03-m	8	FS_16-f
4	FS_04-f	9	FS_17-f
5	FS_06-f	10	FS_18-f
		11	FS_05-m

Среди 11 рассказчиков восемь женщин и трое мужчин, возраст рассказчиков — от 18 до 22 лет (средний возраст 20,4 года).

В дальнейшем для каждого текстового примера из корпуса приводится только **новый** идентификатор дискурса-источника.

2.2. Линейная сегментация дискурса

Обсудим основные термины анализа дискурсивной структуры, принятые в работе. Для локальной сегментации используется понятие ЭДЕ — элементарной дискурсивной еди-ницы: «EDUs are quanta, or moments, of discourse time: discourse progresses in steps equalling EDUs» [A. A. Kibrik 2011: 377]. Преимущество такой сегментации заключается в ее ком-плексном характере [А. А. Кибрик и др. 2009а: 57]:

- с точки зрения координации речепроизводства и дыхания ЭДЕ произносится за один выдох;
- с когнитивной точки зрения ЭДЕ содержит один «фокус сознания» [Chafe 1994];
- с позиции семантики ЭДЕ описывает одно событие или одно состояние;
- с позиции синтаксиса ЭДЕ часто является клаузой;
- с точки зрения просодии ЭДЕ имеет все признаки самостоятельной «произноситель-ной конфигурации» (единый интонационный контур, характерная динамика темпа и громкости, паузация).

Понятие ЭДЕ применимо не только к устному дискурсу, но и к письменному; один из способов делить письменные предложения на ЭДЕ представлен в [Carlson et al. 2003: 87–89]: как правило, ЭДЕ совпадает с клаузой, однако есть три особых случая.

- 1) Некоторые клаузы не составляют отдельных ЭДЕ — это сентенциальные подлежащие, дополнения и комплементы клауз.
- 2) Некоторые неклаузальные элементы размечаются как ЭДЕ — это предложные группы, которые начинаются с сильного «дискурсивного маркера», см. [Ibid.: 98–99], например *because, in spite of, according to*.
- 3) Авторы вводят понятие вложенной ЭДЕ: таким образом размечаются относительные придаточные предложения, именные модификаторы в постпозиции (*nominal post-modifiers*), а также любая другая клауза, которая располагается внутри полноценной ЭДЕ. Если в полноценную ЭДЕ вложена клауза, то фрагмент ЭДЕ до этой клаузы и фрагмент после нее рассматриваются как отдельные ЭДЕ, единство которых на уровне риторической структуры восстанавливается с помощью специального отношения.

Важно отметить, что ЭДЕ в [Carlson et al. 2003] и ЭДЕ в [А. А. Кибрик и др. 2009a] соответствуют разным явлениям, поскольку эти понятия изначально разработаны для разных модусов дискурса.

2.3. Риторическая структура

Для описания смыслового устройства рассказов применяется теория риторической структуры [Mann, Thompson 1987]. В этой теории при помощи фиксированного набора смысловых отношений (например, «Обстоятельство», «Уступка», «Последовательность») устанавливаются связи между единицами дискурса:

- самой малой единицей в оригинальной теории является клауза (в настоящей работе — ЭДЕ);
- предложения соединяются смысловыми отношениями и объединяются в группы;
- группы можно соединять с единицами и с другими группами;
- самой большой группой является целый дискурс.

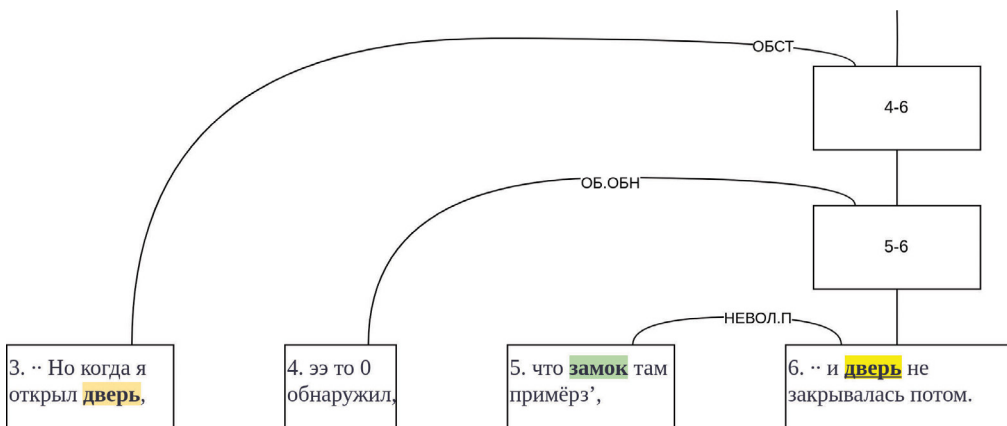


Рис. 2. Пример риторической структуры (часть рассказа № 7sp)⁷

⁷ Обозначения отношений: НЕВОЛ.П — неволевая причина; ОБСТ — обстоятельство; ОБ.ОБН — обстоятельство обнаружения. На рис. 2 и 3 для наглядности из дерева убрана дополнительная разметка, однако она присутствует в дереве данного рассказа в облачном хранилище.

Результат анализа представляет собой дерево, в котором учтены все ЭДЕ данного дискурса. Фрагмент такого дерева показан на рис. 2 (с. 141). Все диаграммы были построены вручную в draw.io — приложении для создания схем, графиков, деревьев решений и других логических и структурных визуализаций.

Различается два типа смысловых отношений.

1. Симметричное, или многоядерное, отношение соединяет два и более элементов структуры. Соединенные ЭДЕ являются одинаково значимыми в дискурсе. В дереве элементы соединены дугами между собой и косыми чертами с общим узлом группы. Примеры отношений: конъюнкция, последовательность. Также симметричное отношение показано на рис. 3.
2. Асимметричное отношение соединяет два элемента. Один из них является главным и называется ядром, другой второстепенным и называется сателлитом. Ядро соединено дугой с сателлитом и прямой чертой с общим обозначением группы. Примеры отношений: фон, обстоятельство, уступка. Такие отношения фигурируют на рис. 2 и 3.

Инвентарь риторических отношений может варьироваться в зависимости от типа рассматриваемых дискурсов. В настоящей работе задействована существенная часть перечня, который использовался в [Литвиненко и др. 2009], с добавлением трех новых отношений.

1. **Вставка:** симметрично, используется для зависимой клаузы, вложенной в главную, исключительно в письменных текстах; это отношение соединяет в один узел все три компонента, которые получаются в результате вложения клаузы⁸. Пример применения этого отношения изображен на рис. 3.

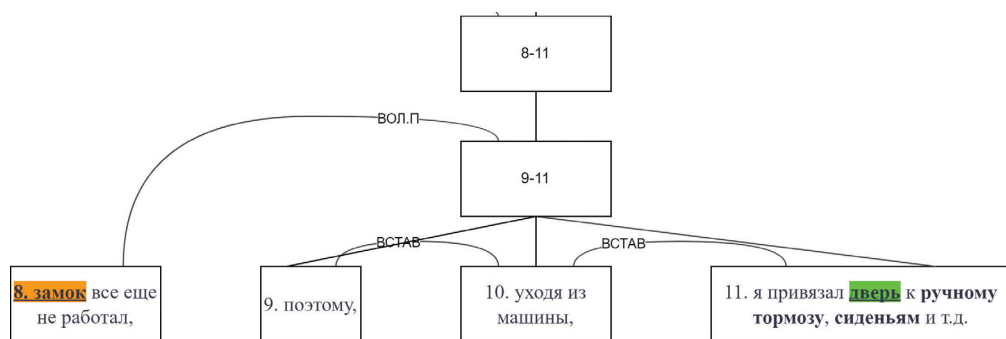


Рис. 3. Использование риторического отношения «Вставка» (обозначено как «ВСТАВ», соединяет ЭДЕ 9, 10 и 11) во фрагменте рассказа № 7wr

2. **Поиск:** симметрично, соединяет маркер поиска с полнозначными элементами дискурса до маркера и после него. Используется, когда говорящий останавливается, чтобы вспомнить нужное слово, и в процессе вспоминания порождает регуляторные ЭДЕ⁹,

⁸ Другой способ учитывать вложенные клаузы — специальное отношение SAME-UNIT [Carlson et al. 2003], но при таком подходе две несоседние единицы — части главной клаузы — оказываются соединенными «в обход» вложенной. С другой стороны, возможно присоединение вложенной клаузы к одной из частей главной клаузы: это помогает сохранить смысловые взаимоотношения между двумя клаузами, однако не всегда бывает легко выбрать часть главной клаузы, к которой будет естественно присоединить зависимую. Предложенное в настоящей статье решение можно назвать упрощенным в том отношении, что оно не отражает смысловых связей между клаузами, однако его преимущество состоит в том, что оно находится в рамках классической концепции риторического отношения.

⁹ Термин «регуляторные ЭДЕ» используется в понимании, изложенном в [А. А. Кибрик и др. 2009б].

например как *его*¹⁰. Данное отношение имеет компонент разрывности («фокус сознания» уже оформлен, возникает лишь проблема с подбором конкретного слова), поэтому представляется возможным считать его симметричным. Пример использования такого отношения показан на рис. 4¹¹.

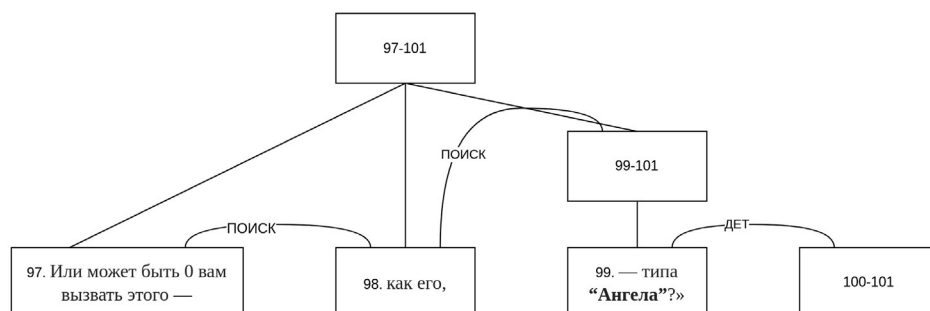


Рис. 4. Фрагмент рассказа № 2sp с применением отношения «Поиск» (соединяет ЭДЕ 97, 98 и группу 99–101)

3. Обращение: асимметрично, в сателлите вводит обращение в рамках прямой речи, в качестве ядра присоединяет всю остальную прямую речь.

2.4. Разметка дискурсов для исследования

2.4.1. Общая и сегментная разметка

Каждый дискурс в выборке был скопирован в отдельный docx-файл, размечен и снабжен анкетой, которая отражает состав его референтов и распределение РС. Это позволило лучше контролировать последовательность и качество сбора данных. Пример анкеты показан в табл. 2 (с. 144):

Дополнительная разметка референциальных элементов состоит из следующих действий. Письменные дискурсы разбиты на ЭДЕ вручную в соответствии с методикой, предложенной в [Carlson et al. 2003]. Устные рассказы дополнительно размечены по эпизодам (с опорой на деление их письменных аналогов на абзацы). В каждом тексте отмечены все нулевые упоминания, также жирным шрифтом размечены все нелокуторные упоминания и разными цветами — все РЦ. Внутри РЦ реферемы отмечаются более ярким цветом и дополнительно подчеркиваются, а простые упоминания окрашиваются более бледным оттенком того же цвета.

¹⁰ Такие ЭДЕ также называются плейсхолдерами (placeholders) [Podlesskaya 2010].

¹¹ Как справедливо отметил рецензент, существует другой вариант риторического анализа таких сегментов — маркер поиска можно считать сателлитом при одном из полнозначных элементов. Аналогичный подход предлагается, например, в [Литвиненко и др. 2009] для фальстартов. Однако в случае с плейсхолдерами, как и при отношении «Вставка», полнозначные элементы присутствуют по обе стороны от маркера речевого сбоя, а значит, возникает вопрос, к какому из них будет более правильно присоединять маркер в качестве сателлита. Симметричное трехместное отношение «Поиск», предлагаемое в данной статье, снимает эту проблему.

Анкета рассказа № 8wr

Таблица 2

Общее число референтов:	6
Из них значимых:	4
На них приходится упоминаний:	23
— в том числе вводных упоминаний:	4
— в том числе простых упоминаний:	5
— в том числе реферем:	14
Перечисление значимых референтов (в порядке появления):	мальчик, сухари, актер, билетерши
Среди реферем значимых референтов:	
Число полных ИГ ¹² :	5
— из них дескрипций с местоимением <i>этот</i> :	1
— из них прочих распространенных дескрипций:	0
— из них имен собственных:	0
Число местоимений:	9
Число нулей:	0
Комментарий (опционально)	—

Затем для каждого рассказа в каждом его варианте было вручную построено дерево риторической структуры. Часто в риторических деревьях горизонтально выравниваются промежуточные узлы, ср. пример (7) из [Mann, Thompson 1987: 25]:

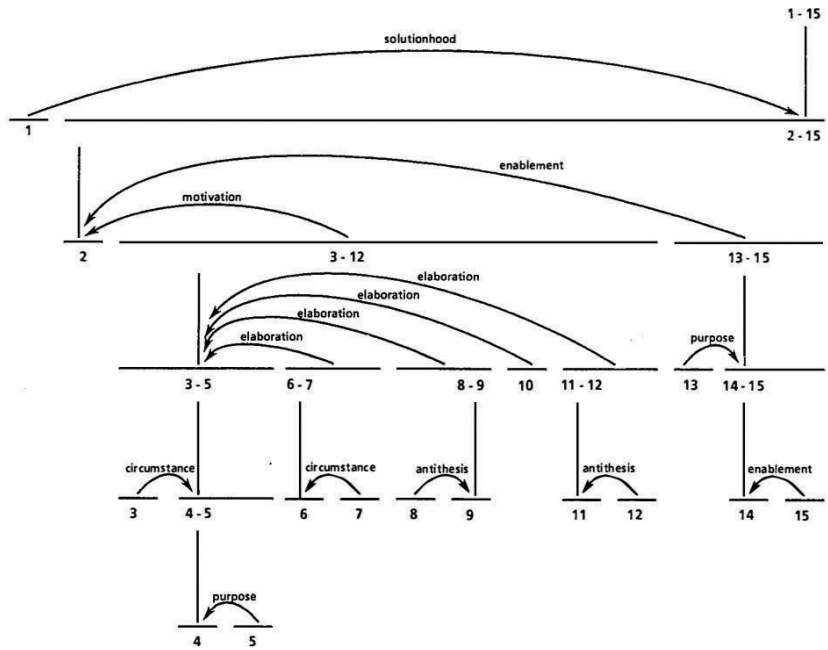


Рис. 5. Классическое дерево риторической структуры

¹² Разновидности полных ИГ подсчитывались для общей картины, итоговые показатели не были использованы для задач данной работы.

В настоящей работе, напротив, выравниванию подлежал самый нижний уровень структуры (см. рис. 2–4), уровень ЭДЕ — так было сделано для удобства чтения текста по диаграмме. Вся разметка упоминаний и реферем была перенесена из docx-файлов в риторические деревья для удобства подсчета расстояний.

2.4.2. Факторная разметка

Разметка реферем и их антецедентов по факторам выполнялась не в файлах с текстами, а в специальных xlsx-книгах. Данные распределяются в две книги.

1. Каждая реферема в дискурсе получает уникальный идентификатор — трехчисловой код: первое число отражает номер дискурса в выборке, второе — номер референта по порядку появления, третье — номер ЭДЕ в референциальной цепочке данного референта. Перечисление дискурсов, референтов и реферем, присвоение ID реферемам и контрольные подсчеты были произведены в **таблицах ключей** (по одной для каждого модуса). Таблицы объединены в **книгу ключей**.

2. Факторная разметка каждой реферемы производилась в таблице, где первый столбец содержит ID всех реферем, далее следуют столбцы для 29 факторов (о них см. в следующем разделе), а также столбец, где отмечается наличие действия фильтров на данную реферему. Такая таблица называется **таблицей данных**, для каждого модуса предусмотрена своя таблица с таким содержанием.

Вместе две таблицы составляют **базу данных**.

3. Моделирование

3.1. Факторы

В качестве источников факторов, влияющих на референциальный выбор, привлекались исследования [Чейф 1982; А. А. Kibrik 1996; Strube, Wolters 2002; Arnold 2008; Loukachevitch et al. 2011; van Vliet 2012; Same, van Deemter 2020] и др. Всего отобрано 29 факторов, в том числе пять экспериментальных, значимость которых ранее не проверялась методами машинного обучения. Факторы сгруппированы в факторные блоки по аналогии с [Loukachevitch et al. 2011]:

Факторы референта:

1. Конкретность референта (см. п. 1.1),
2. Одушевленность,
3. Число,
4. Род,
5. Объединенное свойство род-число,
6. Протагонизм 1,
7. Протагонизм 2.

Факторы реферемы:

8. Грамматическая роль,
9. Подлежащность,
10. Семантическая гиперроль,
11. Порядковый номер в РЦ.

Факторы антецедентов:**Линейный антецедент:**

12. РС,
13. Грамматическая роль,
14. Подлежащность,
15. Семантическая гиперроль,
16. Неполная кореферентность с анафором (sloppy identity) [Dahl 1973],

Риторический антецедент:

17. РС,
18. Грамматическая роль,
19. Подлежащность,
20. Семантическая гиперроль,
21. Неполная кореферентность с анафором [Ibid.].

Дистанции:

22. Линейное расстояние в ЭДЕ,
23. Линейное расстояние в реферемах — число реферем между анафором и антецедентом (все РЦ),
24. Расстояние в эпизодах для устного дискурса, в абзацах — для письменного,
25. Риторическое расстояние,

Факторы клаузы:

26. Наличие параллелизма,
27. Наличие контрастивности,

Референциальная информативность предиката:

28. Отражает ли предикат род референта,
29. Отражает ли предикат число референта.

Остановимся подробнее на некоторых факторах.

Протагонизм-1 и протагонизм-2 (факторы 6 и 7). Алгоритмы подсчета этих мер протагонизма были представлены в [Loukachevitch et al. 2011: 523]: в первом случае протагонизм — это отношение длины данной РЦ к длине самой большой РЦ в дискурсе, во втором случае — отношение длины данной РЦ к сумме длин всех РЦ в дискурсе.

Синтаксическая роль анафора, линейного и риторического антецедентов (факторы 8, 13 и 18). Каждый из этих факторов является категориальным и допускает три значения: «подлежащее», «прямое дополнение» и «прочее». Каждый фактор имеет бинарный аналог (9, 14 и 19 соответственно), в котором подлежащее противопоставлено прочим ролям.

Семантические гиперроли анафора, линейного и риторического антецедентов (факторы 10, 15 и 20). В качестве рабочего набора гиперролей приняты принципал, пациентив и датив по определениям в статье [А. Е. Kibrik 1997], а также инструмент по определению в статье [Fillmore 1968: 46].

Линейное расстояние в реферемах (фактор 23) отражает количество реферем других РЦ между анафором и антецедентом. Так, в примере (1), повторенном в (12), расстояние между упоминанием цветов в ЭДЕ 32 и ЭДЕ 30 равно 2, так как между ними присутствуют два упоминания того человека, который эти цветы посадил (нуль в 31 и «он» в 32):

- (12) FS_04-f_sp // № 4sp:
27. *моя подруга* —
28. *Кристина*,

29. *смотрит,*
 30. *а там у него цветы всякие в горшках,*
 31. *ну и мы спросили 0,*
 32. *мол это е=1 он сам сажает 0,*

Этот фактор устроен аналогично фактору «Расстояние в маркабулах» (*distance in markables*) [Loukachevitch et al. 2011: 522], однако понятия реферемы и маркабулы не тождественны¹³.

Риторическое расстояние (фактор 25). Подсчет был произведен в соответствии с системой, предложенной в [А. А. Kibrik, Krasavina 2005]:

- 1) за каждый горизонтальный переход по дуге между ядром и сателлитом (либо наоборот) начисляется единица;
- 2) однако за переход между клаузами, тесно связанными синтаксически (*syntactically tightly knit clauses* [Ibid.: 568]), начисляется 0,5 балла; столько же начисляется за переход от основного нарратива к цитированию (отношение «Содержание»);
- 3) за переход по симметричному отношению между сестринскими узлами начисляется 1 балл;
- 4) за подъем или спуск по симметричному отношению (переход от родительского узла к дочернему или наоборот) добавляется 0,5 балла.

Перейдем к рассмотрению экспериментальных факторов. Как уже было сказано во введении, все эти факторы представлены в теоретических исследованиях РВ, но все они, за исключением параллелизма, не были ранее проверены при помощи машинных моделей.

Конкретность референта (фактор 1) является обобщением над возможными денотативными статусами. Это бинарный фактор: конкретные статусы (определенный и неопределенный) противопоставлены неконкретным (экзистенциальному, универсальному, атрибутивному и родовому).

Параллелизм (фактор 26) как значимый фактор упоминается, в частности, в статье [Arnold 2008]. В [Strube, Wolters 2002] этот фактор присутствует, однако понимается в узком смысле — предлагается считать реферему и ее антецедент параллельными в том случае, если они имеют одинаковые синтаксические роли [Ibid.: 20]. В [Fahnestock 2003: 131] представлен более широкий взгляд на это явление: параллелизм рассматривается на шести уровнях — 1) длительности, 2) интонации, 3) фонетики (аллитерация и ассонанс), 4) грамматики (повтор последовательности синтаксических элементов), 5) семантики (элементы с одинаковой синтаксической ролью относятся к одной «семантической категории»¹⁴),

¹³ Обсудим соотношение понятий «маркабула» и «реферема». В [А. А. Kibrik 2011: 472] маркабула определяется как элемент, который следует аннотировать в корпусе, после определения приводятся все классы ИГ и разряды местоимений, которые составляют множество маркабул: «The markables to be annotated include definite, demonstrative, and possessive full NPs (headed by a common noun), proper nouns, personal pronouns, and demonstratives. In addition, other NPs are annotated if they serve as antecedents of anaphors (in particular, indefinite full NPs)». Реферема — это ИГ с определенным семантическим свойством (референт не локутор), повторное упоминание некоторого референта, план выражения которого не предопределен (главным образом грамматически).

Все классы ИГ и разряды местоимений, включенные во множество маркабул, могут быть реферемами, однако из факторного блока [Ibid.] можно заключить, что в состав маркабул входят вводные упоминания. Кроме того, не каждая реферема попадает во множество маркабул — прежде всего это касается нулевых ИГ. Таким образом, множество, которое задается определением маркабулы, пересекается со множеством, которое задается определением реферемы, но ни одно из них не является подмножеством другого.

¹⁴ Исходя из примеров в данной статье, можно заключить, что автор называет семантическими категориями не семантические роли, а принадлежность слов к одному онтологическому

6) повторения. В настоящей работе клаузы считаются параллельными, если демонстрируют грамматический или семантический признак из данного перечня.

Контрастивность (фактор 27) рассматривается в статье [Чейф 1982]: исходя из эмпирических данных, контрастивные референты чаще бывают выражены полными ИГ.

Референциальная информативность предиката с точки зрения рода и числа (факторы 28 и 29). Понятие референциальной информативности глагола было введено в диссертации [Ефимова 2006] на материале японского языка. Этот фактор сообщает о наличии или отсутствии бенефактивной конструкции, которая состоит из полнозначного глагола и следующего за ним вспомогательного. З. В. Ефимова отмечает, что в японском языке бенефактивный референт часто бывает выражен нулевой ИГ. В таких случаях вспомогательный глагол может выступать единственным показателем присутствия этого референта, поэтому может быть полезно учитывать наличие такой конструкции. С опорой на идею влияния данного фактора в настоящей работе сделана попытка учесть согласование предиката с подлежащим по роду и числу:

(13) № 4wr:

24. **Он** ответил,

25. что **0** купил

26. и **0** посадил их сам,

(14) (Пример, сконструированный на основе (13)):

24. **Он** ответил,

25. что **0** покупает

26. и **0** сажает их сам,

В (13) оба нулевых упоминания являются актантами предиката в прошедшем времени: в каждом случае предикат содержит информацию о роде и о числе референта. В (14) прошедшее время предикатов в ЭДЕ 25 и 26 заменяется на настоящее: такие предикаты несут информацию о числе референта, но не передают сведений о его роде.

3.2. Фильтры

В когнитивной количественной модели **фильтр** — это проверка, которая проводится после подсчета коэффициента активации референта; фильтр имеет условия действия, и при наличии этих условий он переопределяет выбор РС. В [А. А. Kibrik 1996: 280–282] описано два фильтра: референциальный конфликт и смена границ действительности (*world boundary filter*). Оба они предписывают рефереме план выражения в виде полной ИГ вне зависимости от величины коэффициента. Первый фильтр срабатывает, когда использование редуцированного РС может спровоцировать референциальную неоднозначность. Второй фильтр действует, когда в рамках цитирования некоторого дискурса в виде прямой речи (как правило, диалога или размышлений) говорящий впервые упоминает референт, но при этом в рамках всего дискурса данное упоминание не является вводным.

классу. Семантическое сходство в статье понимается достаточно узко, например, оно постулируется в паре предложений:

(i) *During February, ice storms occur frequently, creating havoc with the traffic.*

(ii) *Around August, tornadoes happen occasionally, causing danger on the roadways.*

Из объяснений автора можно сделать вывод, что семантический параллелизм в этих предложениях представлен в виде трех пар: названия месяцев, стихийные бедствия, проблемы на дорогах.

В настоящей работе к ним добавляется третий фильтр — недопустимость нулевой предположной группы¹⁵. Необходимость его введения обоснована эмпирически: существует несколько классов предложений (особенно распространенных в устной речи), в которых референция происходит при помощи предположной группы и в случае замены этой предположной группы на нуль предложение становится неясным или неграмматичным, как, например, в ЭДЕ 39:

(15) № 11sp:

36. *Дождь* *он* *шестого* *числа*,

37. *0* *пришел*,

38. *сажают* *на* *самолет*.

39. *Папа* *со своим билетом*¹⁶.

3.3. Данные

Из корпусного материала были сформированы две выборки — устная и письменная, — куда вошли все реферемы из проанализированных текстов. Подсчеты референтов и РС приведены в табл. 3.

Таблица 3

Статистика референтов и реферем по двум выборкам

	Письменная выборка	Устная выборка
Всего референтов	119	154 ¹⁷
Значимых референтов	51	58
Упоминаний значимых референтов	337	483
Реферем, в том числе:	211	280
— полных ИГ	109 (52 %)	121 (43 %)
— местоимений	75 (35 %)	115 (41 %)
— нулевых ИГ	27 (13 %)	44 (16 %)

Перед построением моделей в среде R¹⁸ был выполнен поиск отсутствующих значений — в обеих таблицах все значения заполнены. Кроме того, проверка методом nearZeroVar показала, что конкретность референта, контрастивность и параллелизм являются незначимыми на материале выборок обоих модусов. Таким образом, три из пяти экспериментальных фактора были исключены из дальнейшего рассмотрения.

При анализе выборок было выявлено, что значение «средний род» факторов «род» и «род-число», а также значение «инструмент» семантических факторов встречаются

¹⁵ Возможно, этот фильтр было бы полезно рассматривать как ограничение референциального выбора, см. [Залманов, Кибрик 2023].

¹⁶ В настоящей статье предлоги не включаются в упоминания, однако такое решение можно назвать вынужденным упрощением. Изучение взаимосвязей между предлогами и референциальными выражениями является темой для дальнейшего исследования.

¹⁷ Разница между выборками в общем количестве референтов составляет 35, в количестве значимых референтов — 7. Таким образом, в устных текстах незначимых референтов на 24 больше, чем в письменных, что примечательно.

¹⁸ Доступна по ссылке <https://cran.r-project.org>.

крайне редко — модели не смогли бы корректно обучиться на единицах с этими значениями. Реферемы с данными признаками были исключены из выборки.

3.4. Построение и интерпретация моделей

Для моделирования РВ были применены два метода классического машинного обучения — мультиномиальная логистическая регрессия и дерево решений. Оба метода не предъявляют жестких требований к количеству материала и могут работать с малыми выборками. Кроме того, на основе деревьев решений получены диаграммы значимости факторов (Variable Importance in Projection, сокр. VIP) — в них ранжированы по степени важности все факторы, в том числе те, которые не вошли в финальные деревья.

Регрессия построена в статистическом пакете jamovi¹⁹, деревья решений и диаграммы — в среде R.

3.4.1. Логистическая регрессия

Данный метод машинного обучения применяется во многих исследованиях РВ, например в [Loukachevitch et al. 2011; van Vliet 2012; Rosa, Arnold 2017; Same, van Deemter 2020]. В настоящей работе использована его мультиномиальная разновидность, предназначенная для случаев, когда зависимая переменная имеет более двух значений. Ниже приведено краткое описание устройства этой модели; оно может быть полезно для понимания того, как были получены веса факторных значений, а также для объяснения принципов их интерпретации.

Прежде чем описать устройство мультиномиальной модели, рассмотрим ее бинарный прототип. Бинарная логистическая регрессия позволяет определить для каждого целевого объекта, принадлежит ли он к некоторому целевому классу или нет. В ее основе лежит линейная функция вида $y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \beta_0$, где n — это количество факторов, x_i — это некоторое факторное значение, каждый β_i из множества $\beta_1 \dots \beta_n$ — это вес значения x_i в модели, или численное выражение его вклада в общий результат, а β_0 — свободный коэффициент. Веса также принято называть бета-коэффициентами. Входными значениями для логистической модели являются характеристики некоторого объекта; в процессе ее работы рассчитывается вероятность, что данный объект относится к целевому классу. Взвешенная сумма факторных значений может быть любым действительным числом, однако необходимо, чтобы ее результат был отображен в пределах отрезка $[0; 1]$ — для этого она проходит через функцию сигмоиды. Полученная вероятность сравнивается с некоторым порогом, и если она превышает его, то исследуемый объект относится к целевому классу. В состав логистической модели входит много компонентов, однако для задач исследования первостепенное значение имеют веса $\beta_1 \dots \beta_n$; именно они подвергаются интерпретации.

В качестве примера бинарной логистической регрессии рассмотрим модели, представленные в гл. 6 книги [van Vliet 2012]. Автор моделирует референциальный выбор между местоимением и именем собственным. Так как изначально логистическая модель предсказывает принадлежность/непринадлежность объекта к целевому классу, выбор между двумя РС адаптируется в виде вариантов «имя собственное» и «не имя собственное», причем второй вариант полностью совпадает с классом «местоимение», поскольку другие РС автор в своей модели не учитывает. В настоящей работе класс, не являющийся целевым, предлагается называть базовым; в представленных ниже логистических моделях в качестве базового класса выбрано местоимение, как и в моделях в [Ibid.]. Однако в настоящей

¹⁹ Доступен по ссылке <https://www.jamovi.org>.

работе у зависимой переменной — референциального средства — не два возможных значения, а три. Это значит, что у нее два не-базовых значения: полная ИГ и нулевая ИГ. Следовательно, в модели отдельно вычисляется вероятность полной ИГ относительно местоимения и вероятность нулевой ИГ относительно местоимения. Таким образом, при интерпретации моделей речь будет идти не о трех значениях совместно, а о двух парах, поскольку веса для каждой пары рассчитываются отдельно. Универсальных правил выбора базового значения не существует; выбор в пользу местоимения сделан для удобства интерпретации.

Логистическая регрессия может принимать факторы нескольких типов: категориальные (например, семантическую гиперроль), количественные дискретные (например, расстояние в ЭДЕ или в абзацах) и количественные непрерывные (к этому типу в моделях относятся только протогонизм). Их веса устроены следующим образом:

- у категориального фактора одно из значений выбирается базовым, его вес фиксирован и составляет 0. Остальные значения получают свои веса в процессе обучения модели. Если не-базовых значений больше одного, то каждое получает свой собственный вес;
- количественный фактор имеет один вес, который умножается на его величину; базового значения для нулевой величины не предусмотрено.

В [Molnar 2022] приведены рекомендации, как следует интерпретировать веса логистической регрессии:

- для категориальных факторов:
 - базовое значение не вносит вклад в результат работы модели;
 - не-базовое значение, имеющее вес n , повышает вероятность не-базового значения зависимой переменной в $\exp(n)$ раз;
- для количественных факторов:
 - нулевое значение величины не вносит вклад в результат работы модели;
 - если фактор имеет вес n и величину t , то вероятность не-базового значения зависимой переменной возрастает в $t \cdot \exp(n)$ раз;
- свободный коэффициент β_0 не подвергается интерпретации.

Если вес n факторного значения отрицательный, то $\exp(n)$ меньше единицы, и тогда, наоборот, повышается вероятность базового значения зависимой переменной.

Программа `jamovi`, в которой были получены модели регрессии на материале выборов, рассчитывает не только сами веса, но также и их p -уровни значимости — в первом приближении, степени их недоверности: это значит, что чем меньше p -уровень, тем достовернее величина. Анализ были подвергнуты только веса с p -уровнем значимости не более 0,1. Важно отметить, что уровень значимости выше 0,1 говорит не об отсутствии влияния фактора на результат, а о том, что на материале выборки невозможно точно вычислить его вес.

Прежде чем перейти к интерпретации весов, рассмотрим процесс выбора базовых значений для категориальных факторов. Как и в случае с зависимой переменной, универсальной процедуры нет, поэтому везде, где это возможно, был сделан обоснованный лингвистически выбор. Например, для фактора одушевленности базовым значением выбрана неодушевленность: таким образом в модель закладывается идея о повышении активации в случае, если референт одушевленный [А. А. Kibrik 2011: 413]. Несколько сложнее было выбрать базовое значение для числа, однако есть основания предполагать, что множественное число усложняет РВ и определенным образом влияет на него ([Ibid.: 554–555]; наблюдения за дистрибуцией РС в зависимости от числа референта также представлены в статье [Zalmanov, A. A. Kibrik 2021]), поэтому в качестве базы было выбрано единственное число.

Достаточно сложно было выбрать базовое значение для грамматической роли: известно, что «подлежащее» повышает активацию, но каково влияние значений «прямое дополнение» и «прочее»? Представляется важным, чтобы под не-базовым факторным значением объединялись однородные опции, которые единообразно влияют на РВ; если категория «прочее» окажется не-базовой, то ее интерпретация будет затруднена в силу разнообразия значений, которые в нее входят. Так, у фактора грамматической роли в состав «прочего» включены дативные дополнения и несогласованные определения; представляется более разумным принять его как базовое. Таким же образом была выбрана категория «прочее» как базовое значение для всех семантических факторов.

Для факторов рода и рода-числа выбрать базовое значение обоснованным образом оказалось невозможно, поэтому они не вошли в регрессии.

Перейдем к анализу весов логистических моделей. В табл. 4 и 5 представлено сравнение вероятностей РС в парах «полная ИГ — местоимение» и «нуль — местоимение» соответственно. Третий столбец таблицы — «Значение» — заполнен только для категориальных факторов, так как у них каждое значение имеет свой собственный вес; модуль коэффициента соответствует силе влияния данного факторного значения на РВ. Также для каждого категориального фактора в скобках приведено его базовое значение, которое не вносит вклад в работу модели. В столбцах «РС» указывается средство, вероятность которого повышается за счет данного фактора или факторного значения. В таблицах размещены только коэффициенты, уровень значимости которых оказался приемлемым в обоих модусах.

Таблица 4

Значимые коэффициенты регрессий в паре «полная ИГ — местоимение»

Группа	Фактор	Значение	Письм. модус		Уст. модус	
			Коэфф.	РС	Коэфф.	РС
Референт	протагонизм-1		−4.64	мест.	−2.28	мест.
	протагонизм-2		15.53	полная ИГ	7.84	полная ИГ
Реферема	грамматическая роль (прочее)	подлеж.	−4.8	мест.	1.08	полная ИГ
	подлежащность (нет)	да	−4.8	мест.	1.08	полная ИГ
Лин. ант.	неполная кореферентность (нет)	да	−18.7	мест.	28.81	полная ИГ
Рит. ант.	неполная кореферентность (нет)	да	35	полная ИГ	−27.59	мест.
Дист.	риторическое расстояние		0.83	полная ИГ	0.74	полная ИГ
Клауза	отражает ли предикат число референта (нет)	да	−3.38	мест.	−1.31	мест.

Таблица 5

Значимые коэффициенты регрессий в паре «нуль — местоимение»

Группа	Фактор	Значение	Письм. модус		Уст. модус	
			Коэфф.	РС	Коэфф.	РС
Референт	одушевленность (неодуш.)	одуш.	−2.59	мест.	−3.27	мест.
Реферема	грамматическая роль (прочее)	подлеж.	16.3	нуль	−10.35	мест.
	подлежащность (нет)	да	16.3	нуль	−10.35	мест.
	семантическая гиперроль (прочее)	принципал	−15.75	мест.	23.47	нуль
Лин. ант.	неполная кореферентность (нет)	да	−4.99	мест.	−16.8	мест.
Рит. ант.	неполная кореферентность (нет)	да	−3.9	мест.	−17.54	мест.
Дист.	линейное расстояние в абзацах/эпизодах		−13.6	мест.	−27.15	мест.

В паре «полная ИГ — местоимение» в обоих модусах вероятность полной ИГ повышают протагонизм-2 и риторическое расстояние, вероятность местоимения повышают протагонизм-1 и референциальная информативность предиката. Разнонаправленное влияние демонстрируют четыре факторных значения, в том числе неполная кореферентность с линейным и риторическим антецедентами. Это может быть связано с малой представленностью обоих факторов в выборках. Интерес также представляет различие во влиянии двух вариантов протагонизма — это явление нуждается в дальнейшем исследовании.

В паре «нуль — местоимение» одушевленность, неполная кореферентность и линейное расстояние в абзацах/эпизодах повышают вероятность местоимения, три других факторных значения действуют на модусы различным образом. Примечательно, что гиперроль принципал у реферемы повышает вероятность местоимения в письменном дискурсе и нуля — в устном.

Таким образом, при анализе логистических моделей обнаруживаются некоторые расхождения по силе и направлению влияния факторов на два модуса. Проведенное сопоставление позволяет говорить об истинности второй гипотезы: один и тот же значимый фактор может влиять на два модуса с разной силой.

3.4.2. Дерево решений и диаграмма VIP

Данный метод выявляет наборы факторных значений, наиболее характерные для каждого РС, разбивая всю выборку на подмножества — узлы. Начальный узел, в состав которого входит вся выборка, называется корневым, конечные узлы, которые не подвергаются дальнейшему делению, — листьями.

Деревья построены в среде R при помощи пакетов `rpart` и `rpart.plot`. Они показаны на рис. 6 и 7 (с. 154).

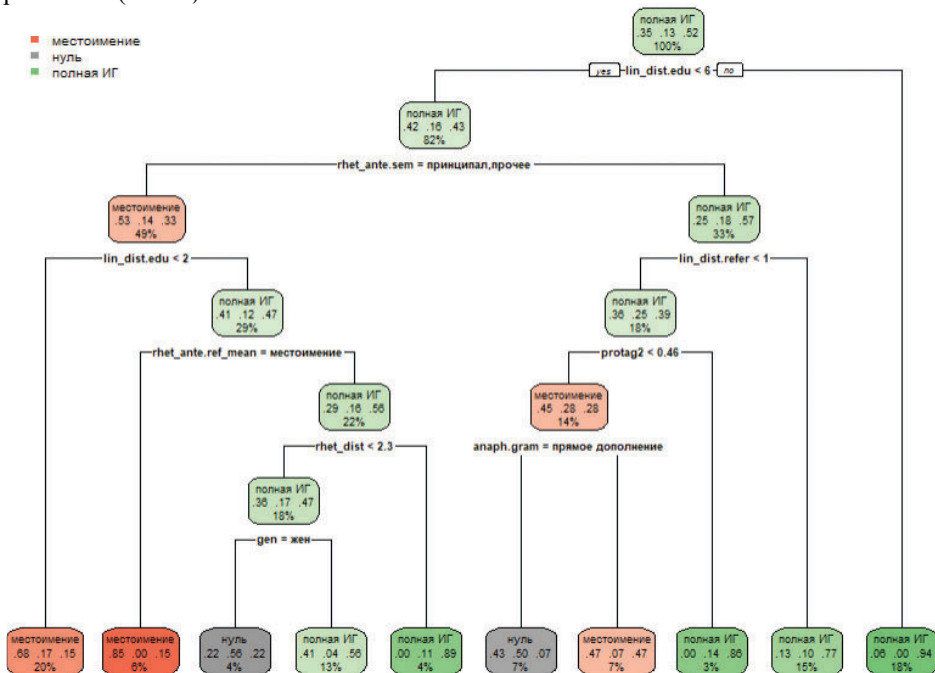
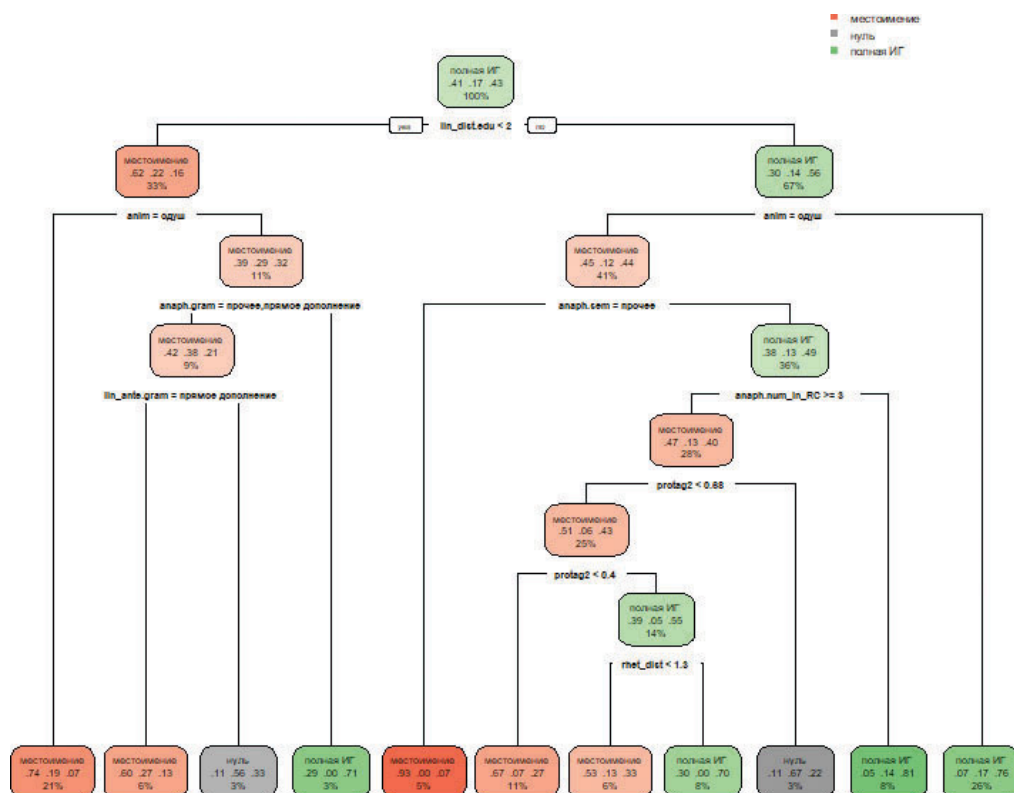


Рис. 6. Дерево решений для письменного модуса²⁰

²⁰ Сокращения в алфавитном порядке: **anaph.gram** — грамматическая роль анафора, **gen** — род референта, **lin_dist.edu** — линейное расстояние в ЭДЕ, **lin_dist.refer** — линейное расстояние

Рис. 7. Дерево решений для устного модуля²¹

В состав дерева входят:

— Узлы. Каждый узел дерева — это подмножество выборки с определенными свойствами. Для каждого узла имеется четыре показателя:

- 1) название, которое присваивается по самому частотному РС в данном подмножестве;
- 2) цвет, который означает РС, преобладающие в данном подмножестве: подмножествам, в которых большинство рефером выражено полными ИГ, соответствуют зеленые листья, местоимениями — оранжевые, нулевыми ИГ — серые. Чем насыщеннее цвет, тем сильнее перевес в сторону самого частотного РС;
- 3) распределение рефером в данном подмножестве между местоимениями, нулями и полными ИГ (процентные доли указаны именно в таком порядке);
- 4) процент, который данное подмножество составляет от всей выборки. Например, первый слева нижний узел на рис. 6 содержит 20 % всей выборки. 68 % его рефером являются местоимениями, 17 % — нулями, 15 % — полными ИГ. Он озаглавлен как «местоимение».

в реферемах, **protag2** — протагонизм-2, **rhet_ante.sem** — семантическая гиперроль риторического антецедента, **rhet_ante.ref_mean** — РС риторического антецедента, **rhet_dist** — риторическое расстояние.

²¹ Сокращения, которых нет на рис. 6, в алфавитном порядке: **anaph.num_in_RC** — порядковый номер анафора в РЦ, **anaph.sem** — семантическая гиперроль анафора, **anim** — одушевленность, **lin_ante.грам** — грамматическая роль линейного антецедента.

- Условия деления (или ветвления). Каждое дерево содержит оптимальное количество и состав условий, который был выведен алгоритмом. Каждое деление бинарно.
- «Ветки», или переходы. Для корневого узла отмечено, что переход влево осуществляется при соблюдении условия деления, переход вправо — при несоблюдении. Переходы при остальных делениях устроены так же.

Первое ветвление обоих деревьев происходит по линейному расстоянию, что говорит о его высокой значимости. Также в обоих деревьях присутствуют деления по риторическому расстоянию, протагонизму-2, грамматической роли риторического antecedента. Дерево для устной выборки содержит больше семантических факторов (одушевленность и семантическая роль реферемы), факторный набор для письменной выборки более разнообразен.

На рис. 8 и 9 (с. 156) показаны VIP-диаграммы, построенные на базе деревьев решений.

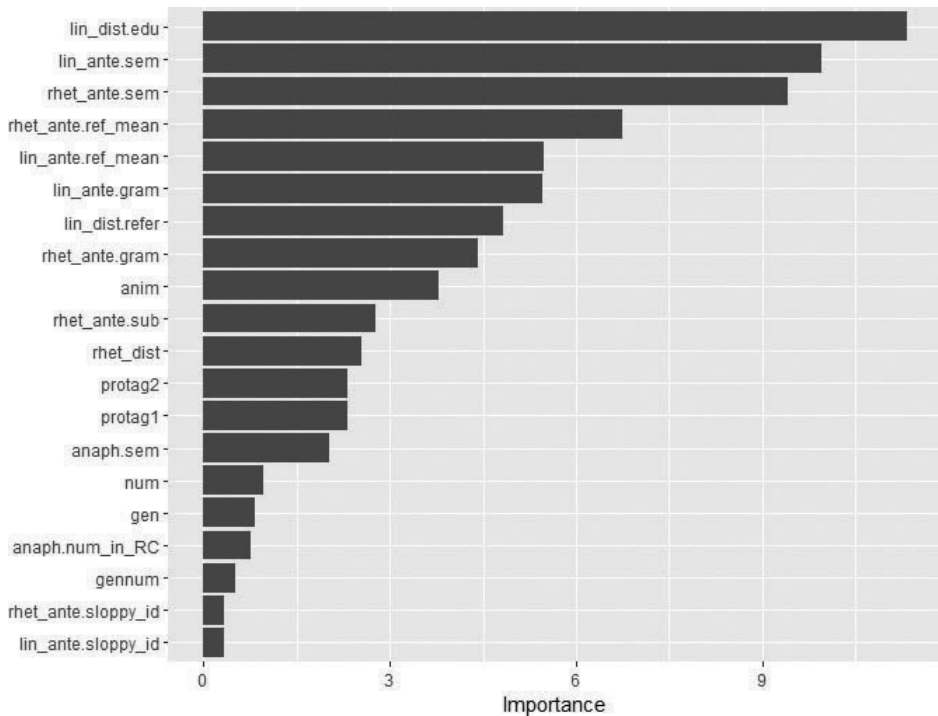


Рис. 8. VIP для письменного модуса

Для обеих моделей риторическое расстояние (обозначено как `rhet_dist`) является значимым. Однако на графике для письменной выборки оно занимает 7 место и уступает некоторым свойствам antecedентов; на графике для устной выборки оно занимает 8 место и уступает достаточно разнородному набору факторов, в том числе всем трем семантическим свойствам (реферемы и обоих antecedентов). Как представляется, имеет смысл обращать больше внимания на позицию фактора в рейтинге, чем на его величину VIP, так как эта величина, по-видимому, является безразмерной, что затрудняет ее интерпретацию. Интересно отметить, что оба экспериментальных фактора, оставшихся в выборках после фильтрации — референциальная информативность предиката по роду и по числу — фигурируют в диаграмме устного модуса и демонстрируют незначительное влияние, в то время как в письменном модусе референциальная информативность предиката по роду (`pred:shows_gen`) отсутствует как значимый фактор РВ.

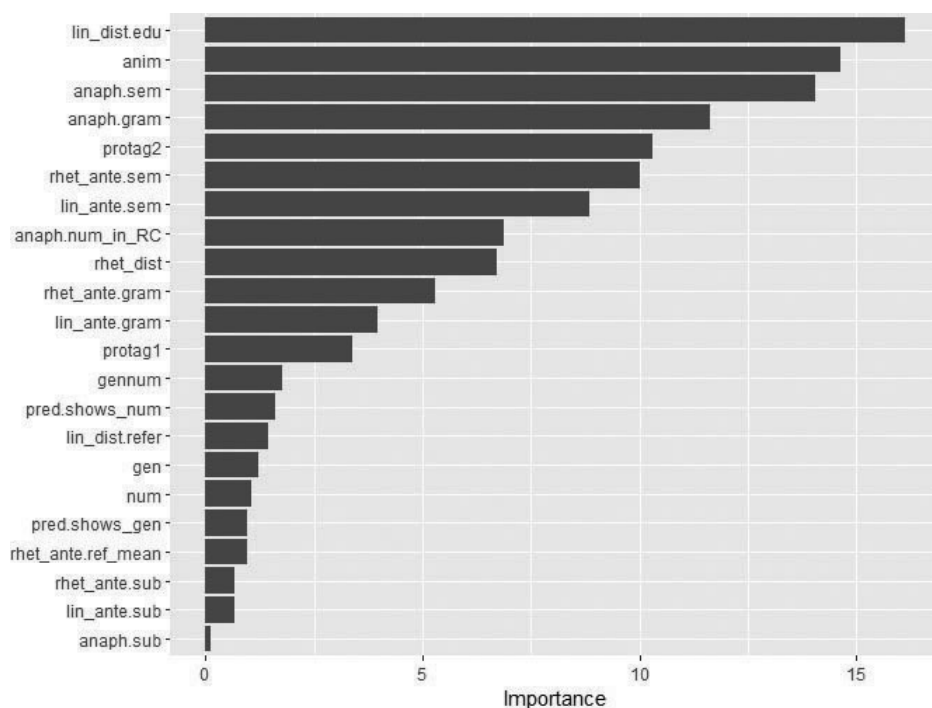


Рис. 9. VIP для устного модуса

4. Обсуждение результатов

Проверка гипотез дала следующие результаты:

Гипотеза 1: наборы факторов, значимых для РВ в двух модусах, различаются (показано с помощью деревьев решений).

Гипотеза 2: для некоторых факторов, значимых в обоих модусах, существует количественное отличие в степени влияния на каждый из модусов (наблюдается в моделях регрессии и VIP).

Гипотеза 4 подтверждена частично: модели позволяют говорить о значимости только одного экспериментального фактора — референциальной информативности предиката с точки зрения числа.

Что касается гипотезы 3 об относительной значимости линейного и риторического расстояния, примененные методы не позволяют делать однозначных выводов. Анализ бета-коэффициентов логистической регрессии не позволил оценить влияние линейного расстояния. Деревья решений демонстрируют, что для обеих выборок первым по приоритетности из всех факторов является линейное расстояние в ЭДЕ, в то время как риторическое расстояние уступает по степени влияния многим факторам. Вероятно, такой противоречивый результат был получен из-за невыясненного до конца характера связи этих факторов с остальными. Так, в когнитивной количественной модели для письменных нарративов риторическое расстояние учитывается дважды: как самостоятельный фактор и в составе фактора «одушевленность». Возможно, похожим образом эти факторы-дистанции ведут себя и в устном модусе, однако раздельного применения регрессии и деревьев решений оказалось недостаточно, чтобы выявить такое влияние.

Заключение

Целью настоящей работы было изучение сходств и различий референциального выбора в устных и письменных нарративах на русском языке. Для исследования были привлечены рассказы из специального корпуса, где каждый сюжет содержится в двух изложениях — устном и письменном. Было взято 11 сюжетов, то есть 22 рассказа; все рассказы были размечены согласно принципам теории риторической структуры. В рассказах были найдены реферемы: собрано около 280 единиц из устных текстов и около 210 из письменных. Каждая реферема получила характеристику по 29 факторам; результирующие базы данных были переданы машинным алгоритмам, отдельно для устной и письменной выборок были построены модели логистической регрессии и деревья решений. Интерпретация моделей показала, что наборы значимых факторов в устных и письменных нарративах различаются; кроме того, факторы, имеющие большую значимость в одном модусе, могут оказывать малый эффект на РВ во втором модусе. Ряд заключений на основе построенных моделей нуждается в рассмотрении на большем массиве данных.

Референциальный выбор продолжает оставаться актуальной темой не только в прикладном, но и в когнитивном ключе. Более глубокое понимание различий между механизмами РВ в устных и письменных дискурсах можно назвать промежуточным шагом к сопоставительному исследованию когнитивных процессов при их порождении. В качестве перспектив данной линии исследований можно назвать уточнение границы между реферемами и простыми упоминаниями, а также более детальный учет различных видов цитирования.

СПИСОК СОКРАЩЕНИЙ

ВИЖ — корпус «Веселые истории из жизни»
РВ — референциальный выбор
РС — референциальное средство

РЦ — референциальная цепочка
ЭДЕ — элементарная дискурсивная единица
VIP — Variable Importance in Projection

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Буденная 2018 — Буденная Е. В. *Эволюция субъектной референции в языках балтийского ареала*. Дис. ... канд. филол. наук. М.: ИЯз РАН, 2018. [Budennaya E. V. *Evolutsiya sub'ektnoi referentsii v yazykakh baltiiskogo areala* [Evolution of subject reference in languages of Baltic sprachbund]. Candidate diss. Moscow: Institute of Linguistics of the Russian Academy of Sciences, 2018.]
- Ефимова 2006 — Ефимова З. В. *Референциальная структура нарратива в японском языке (в сопоставлении с русским)*. Дис. ... канд. филол. наук. М.: РГГУ, 2006. [Efimova Z. V. *Referentsial'naya struktura narrativ v yaponskom yazyke (v sopostavlenii s russkim)* [Referential structure of Japanese narrative (in comparison with Russian)]. Candidate diss. Moscow: Russian State Univ. for the Humanities, 2006.]
- Залманов, А. А. Кибрик 2023 — Залманов Д. А., Кибрик А. А. Уровень активации и специальные ограничения при математическом моделировании референциального выбора. *Язык и искусственный интеллект: Сб. ст. по итогам конф. «Лингвистический форум 2020: Язык и искусственный интеллект», Москва, 12–14 ноября 2020 г.* Вдовиченко А. В. (ред.). М.: Языки славянских культур, 2023, 142–166. [Zalmanov D. A., Kibrik A. A. Activation level and special constraints in referential choice computational modeling. *Yazyk i iskusstvennyi intellekt. Coll. of papers from the conf. "Lingvisticheskii forum 2020: Yazyk i iskusstvennyi intellekt"*, Moscow, November 12–14, 2020. Vdovichenko A. V. (ed.). Moscow: Yazyki slavyanskikh kul'tur, 2023, 142–166.]
- А. А. Кибрик 1997 — Кибрик А. А. Моделирование многофакторного процесса: выбор референциального средства в русском дискурсе. *Вестник Московского университета. Сер. 9: Филология*, 1997, 4: 94–105. [Kibrik A. A. Multifactorial process modeling: Choosing referential devices in Russian discourse. *Lomonosov Philology Journal*, 1997, 4: 94–105.]

- А. А. Кибрик и др. 2009а — Кибрик А. А., Подлесская В. И., Коротаев Н. А. Структура устного дискурса: основные элементы и канонические явления. *Рассказы о сновидениях: корпусное исследование устного русского дискурса*. Кибрик А. А., Подлесская В. И. (ред.). М.: Языки славянских культур, 2009, 55–101. [Kibrik A. A., Podlesskaya V. I., Korotaev N. A. Spoken discourse structure: Main elements and canonical phenomena. *Rasskazy o snovideniyakh: korpusnoe issledovanie ustnogo russkogo diskursa*. Kibrik A. A., Podlesskaya V. I. (eds.). Moscow: Yazyki slavyanskikh kul'tur, 2009, 55–101.]
- А. А. Кибрик и др. 2009б — Кибрик А. А., Подлесская В. И., Коротаев Н. А. Неканонические явления. *Рассказы о сновидениях: корпусное исследование устного русского дискурса*. Кибрик А. А., Подлесская В. И. (ред.). М.: Языки славянских культур, 2009, 102–176. [Kibrik A. A., Podlesskaya V. I., Korotaev N. A. Non-canonical phenomena. *Rasskazy o snovideniyakh: korpusnoe issledovanie ustnogo russkogo diskursa*. Kibrik A. A., Podlesskaya V. I. (eds.). Moscow: Yazyki slavyanskikh kul'tur, 2009, 102–176.]
- Литвиненко и др. 2009 — Литвиненко А. О., Подлесская В. И., Кибрик А. А. Анализ рассказов о сновидениях с точки зрения иерархической структуры дискурса. *Рассказы о сновидениях: корпусное исследование устного русского дискурса*. Кибрик А. А., Подлесская В. И. (ред.). М.: Языки славянских культур, 2009, 431–463. [Litvinenko A. O., Podlesskaya V. I., Kibrik A. A. Dream stories analysis from the point of hierarchical discourse structure. *Rasskazy o snovideniyakh: korpusnoe issledovanie ustnogo russkogo diskursa*. Kibrik A. A., Podlesskaya V. I. (eds.). Moscow: Yazyki slavyanskikh kul'tur, 2009, 431–463.]
- Чейф 1982 — Чейф У. Данное, контрастивность, определенность, подлежащее, топики и точка зрения. *Новое в зарубежной лингвистике*. Кибрик А. Е. (ред.). М.: Прогресс, 1982, 277–316. [Chafe W. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. *Novoe v zarubezhnoi lingvistike*. Kibrik A. E. (ed.). Moscow: Progress, 1982, 277–316.]
- Ariel 1990 — Ariel M. *Accessing noun-phrase antecedents*. London: Routledge, 1990.
- Arnold 2008 — Arnold J. E. Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, 2008, 23: 495–527.
- Carlson et al. 2003 — Carlson L., Marcu D., Okurowski M. E. Building a discourse-tagged corpus in the framework of rhetorical structure theory. *Current and new directions in discourse and dialogue*. van Kuppevelt J., Smith R. (eds.). Dordrecht: Kluwer, 2003, 85–112.
- Chafe 1994 — Chafe W. *Discourse, consciousness, and time*. Chicago: Univ. of Chicago Press, 1994.
- Dahl 1973 — Dahl Ö. On so-called “sloppy identity”. *Synthese*, 1973, 26: 81–112.
- Fahnestock 2003 — Fahnestock J. Verbal and visual parallelism. *Written Communication*, 2003, 20(2): 123–152.
- Fillmore 1968 — Fillmore C. The case for case. *Universals of linguistic theory*. Bach E., Harms R. T. (eds.). New York: Holt, Rinehart & Winston, 1968, 1–88.
- Fox 1995 — Fox B. *Discourse structure and anaphora: Written and conversational English*. Cambridge: Cambridge Univ. Press, 1995.
- Grosz et al. 1986 — Grosz B., Weinstein S., Joshi A. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 1986, 12(3): 175–204.
- Gundel et al. 1993 — Gundel J. K., Hedberg N., Zacharski R. Cognitive status and the form of referring expressions in discourse. *Language*, 1993, 69: 274–307.
- А. А. Кибрик 1996 — Кибрик А. А. Anaphora in Russian narrative prose: A cognitive calculative account. *Studies in anaphora*. Fox B. (ed.). Amsterdam: John Benjamins, 1996, 255–303.
- А. Е. Кибрик 1997 — Кибрик А. Е. Beyond subject and object: Toward a comprehensive relational typology. *Linguistic Typology*, 1997, 1(3): 279–346.
- А. А. Кибрик 2011 — Кибрик А. А. *Reference in discourse*. Oxford: Oxford Univ. Press, 2011.
- А. А. Кибрик, Красавина 2005 — Кибрик А. А., Красавина О. Н. A corpus study of referential choice: The role of rhetorical structure. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”*, 2005, 4: 561–569.
- Loukachevitch et al. 2011 — Loukachevitch N. V., Dobrov G. B., Kibrik A. A., Khudyakova M. V., Linnik A. S. Factors of referential choice: computational modeling. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conf. “Dialogue”*, 2011, 10: 518–528.
- Mann, Thompson 1987 — Mann W. C., Thompson S. A. *Rhetorical structure theory: A theory of text organization*. Los Angeles: Univ. of Southern California, 1987.
- Molnar 2022 — Molnar C. *Interpretable machine learning. A guide for making black box models explainable*. 2022. <https://christophm.github.io/interpretable-ml-book/>.

- Podlesskaya 2010 — Podlesskaya V. I. Parameters for typological variation of placeholders. *Fillers, pauses and placeholders*. Amiridze N., Davis B., MacLagan M. (eds.). Amsterdam: John Benjamins, 2010, 11–32.
- Rosa, Arnold 2017 — Rosa E. C., Arnold J. E. Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language*, 2017, 94: 43–60.
- Same, van Deemter 2020 — Same F., van Deemter K. A linguistic perspective on reference: Choosing a feature set for generating referring expressions in context. *Proc. of the 28th International Conference on Computational Linguistics*. Scott D., Bel N., Zong C. Barcelona: International Committee on Computational Linguistics, 2020, 4575–4586.
- Strube, Wolters 2002 — Strube M., Wolters M. A Probabilistic genre-independent model of pronominalization. *Proc. of the 1st North American Chapter of the Association for Computational Linguistics Conference*. Stroudsburg: Association for Computational Linguistics, 2002, 18–25.
- van Vliet 2012 — van Vliet S. *Proper nouns and pronouns: The production of referential expressions in narrative discourse*. Utrecht: LOT, 2012.
- Zalmanov, Kibrik 2021 — Zalmanov D. A., Kibrik A. A. Short definite descriptions and referent activation. *Advances in cognitive research, artificial intelligence and neuroinformatics*. Velichkovsky B. M., Balaban P. M., Ushakov V. L. (eds.). Cham: Springer, 2021, 284–292.

Получено / received 26.01.2023

Принято / accepted 24.09.2024

[Рец. на: / Review of:] **R. Meyer. *Iranian syntax in Classical Armenian: The Armenian perfect and other cases of pattern replication***. Oxford: Oxford University Press, 2023. xx + 310 pp. (Oxford studies in diachronic and historical linguistics, 53.) ISBN 9780198851097.

Данил Андреевич Алексеев

Московский государственный университет имени М. В. Ломоносова, Москва, Россия; Институт языкознания РАН, Москва, Россия; dnl.alksv@gmail.com

Danil A. Alekseev

Lomonosov Moscow State University, Moscow, Russia; Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia; dnl.alksv@gmail.com

Благодарности: Работа поддержана грантом РФФИ № 24-18-00199 «Структура клаузы и позиционные феномены в языках SOV». Автор глубоко признателен редакции журнала за неоценимую помощь в подготовке этой рецензии.

Acknowledgements: This work was supported by an RSF grant No. 24-18-00199 “Clause structure and positional phenomena in SOV languages”. I am indebted to the editors of the journal for their help in the preparation of this review.

DOI: 10.31857/0373-658X.2024.6.160-166

Развитие синтаксического уровня языка вне рамок теории грамматикализации сравнительно редко становится предметом изучения в диахронических работах, тем более предполагающих контактное происхождение исследуемых явлений. На этом фоне выделяется рецензируемая монография Робина Мейера, доцента (assistant professor) Университета Лозанны, являющаяся кульминацией десяти лет его исследований, начиная со статьи [Meyer 2013].

Работа посвящена некоторым проблемным сюжетам исторического синтаксиса классического армянского, главным образом — маркированию ядерных актантов (alignment) в клаузах, вершиной которых выступает глагол в форме перифрастического перфекта. В отличие от синтетических времен, например аориста, в котором наблюдается стандартная для древнеармянского¹ номинативно-аккузативная стратегия маркирования глагольных актантов (1), в перифрастическом перфекте, состоящем из причастия на *-eal* и глагольной связки (которая, однако, может и опускаться), применяется иная стратегия, предполагающая генитивное кодирование агенса в переходных конструкциях (2):

- (1) *Du es ayn, or kotor-ec'-er z=Ari-s*
ты быть.PRS.3SG тот который разрушать-AOR-AOR.3SG OBJ=ариец-ACC.PL
aysč'ap' am-s...
так.много год-ACC.PL

‘[А, лиса, это ты тот, смутьян, который столько лет нас мучил,] это ты столько лет убивал арийцев?’² [ИА IV.54]³ (с. 80 рецензируемой монографии)

¹ В этой рецензии термины «древнеармянский» и «классический армянский» употребляются в качестве синонимов.

² Глоссирование во многих случаях изменено. О том, почему не используется глоссирование автора монографии, говорится в конце настоящей рецензии.

³ «История Армении Фавтоса Бузанда»; цитируемые издания см. в Списке источников.

- (2) ?oŋ' z=gir=n z=ayn ic'ē ant'erc'-eal jer
 NEG OBJ=Писание.ACC.SG=DEF OBJ=ТОТ быть.3SG.SBJV читать-PTCP 2PL.GEN

‘Неужели вы не читали сего в Писании: [камень, который отвергли строители, тот самый сделался главою угла;]’ [Мк. 12:10]⁴ (с. 83)

Монографию открывают две вводные главы. **Глава 1** вводит читателя в проблемы изучения классического варианта армянского языка и контактов его носителей с ираноязычными сообществами, а также описывает структуру работы. В **главе 2** представляется сумма известных обобщений (в частности, социоисторических) о характере армянско-иранских языковых и культурных контактов и приводятся примеры иранского влияния на фонологическую систему и лексический фонд древнеармянского (раздел 2.3), его морфологию и фразеологию (раздел 2.4)⁵.

Стоит отметить, что из иранских идиомов рассматриваемого периода наибольшее значение, как показано в разделе 2.3, в силу своего преимущественного влияния на классический армянский, для монографии представляет **парфянский язык**, являвшийся государственным в Парфянском царстве (III в. до н. э. — III в. н. э.), которым правила династия Аршакидов [Молчанова 1999: 14]. Одна из аршакидских ветвей пришла к власти и в Великой Армении.

Таким образом, вторая глава подготавливает почву для обсуждения двух главных лакун предметной области монографии: влияния иранских языков на синтаксис классического армянского (раздел 2.6) и конкретной динамики иранско-армянских контактов (раздел 2.7). К «заполнению лакун» автор монографии подходит, ставя следующие вопросы:

- (3) а. «Есть ли в [древне]армянском синтаксические шаблоны, которые возникли на основе схожих моделей в контактном языке одного из ираноязычных народов или, по крайней мере, испытали значительное влияние этих моделей?»⁶
 б. «Что существование таких шаблонов означает для иранско-армянских языковых контактов? Иными словами, могут ли лингвистические данные предложить новый взгляд на характер взаимодействия между этими языками и его силу?»⁷

Таким образом, повествование приобретает спиралевидный характер.

Глава 3 излагает известные данные первичных исторических источников о взаимодействии армянского и иранских языковых сообществ, а также об армянских и иранских государствах в целом. Список рассматриваемых источников действительно широк: древнеперсидские и среднеиранские надписи (раздел 3.1), армянские литературные источники (раздел 3.2), работы античных авторов и новозаветная традиция (раздел 3.3), свидетельства историков китайской династии Хан (раздел 3.4).

С **главы 4** начинается (с точки зрения наличия оригинального вклада) лингвистически содержательная часть монографии, в которой Р. Мейер подходит к рассмотрению вопроса о структуре и происхождении древнеармянского перифрастического перфекта. В подготовительном разделе 4.1 с опорой на [Bickel, Nichols 2008] определяется суть и различные стратегии маркирования глагольных актантов, далее в разделе 4.2 выделяются внешнесинтаксические (залоговые и валентностные) типы контекстов, в которых засвидетельствовано

⁴ Мф. — Евангелие от Матфея; Мк. — Евангелие от Марка; Лк. — Евангелие от Луки; Ин. — Евангелие от Иоанна. Здесь и далее новозаветные отрывки цитируются в Синодальном переводе.

⁵ Совместное рассмотрение данных языковых подсистем объясняется большей изученностью фонологии и лексики, в отличие от морфологии и фразеологии, на что также указывает автор.

⁶ “Have any Armenian syntactic patterns been modelled on (or significantly influenced by) similar patterns in an Iranian contact language?” (с. 31), перевод мой — Д. А.

⁷ “What does the existence of such patterns mean for Iranian–Armenian language contact? That is, can the linguistic data provide any insights into the kind and degree of interaction between these languages?” (с. 32), перевод мой — Д. А.

использование перифрастического перфекта (символом «*» обозначена более маргинальная для конкретного контекста стратегия кодирования актантов):

(4) **Непереходные + S_{NOM} (α)**

or ust-ek' ust-ek' ek-eal has-eal
 который откуда-ADV откуда-ADV приходить-PTCP прибывать-PTCP
ēin i t'ikun-s
 быть.PRS.3SG В помощь.PL-ACC.PL

‘[Но Хосров поднялся со своим огромным войском и теми,] кто прибыл с разных мест, дабы стать в войне его соратником’. [Аг I.8]⁸ (с. 81)

(5) **Непереходные + S_{GEN} (α*)**

yaɣn-žam matuc'-eal ašakert-ac'n nora, ase-n c'=na
 это-время прибывать-PTCP ученик-GEN.PL 3SG.GEN говорить-PRS.3PL IO=3SG

‘Тогда ученики Его, приступив, сказали Ему: [знаешь ли, что фарисеи, услышав слово сие, соблазнились?].’ [Мф. 15:22] (с. 82)

(6) **Непереходные пассивные + S_{NOM} + (A_{i+ABL}) (β)**

erkir ew mardkan, or i nmanē
 Земля.NOM.SG и человечество.NOM.SG который PREP 3SG.ABL
en stelc-eal
 быть.3PL.PRS создать-PTCP

‘[Ибо лишь Он творец неба и ангелов, прославляющих Его величие, и] земли и людей, созданных им...’ [Аг V.9] (с. 82)

(7) **Переходные + A_{GEN} + O_{NOM/ACC} (γ)**

z=or ar-eal t'agawor-i=n, z=amenesean
 OBJ=который принять-PTCP царь-GEN=DEF, OBJ=всё.ACC.PL
astuac-a-kard lcoyn hnazand-ec'uc'-anēr
 Бог-ер-предписанный иго.DAT.DEF подчиненный-CAUS-AOR.3SG

‘И царь подчинил всех установленному Богом игу, [тем более, что причиной был не он, а воля всемогущего Христа...].’ [Аг 0.35] (с. 83)

(8) **Переходные + A_{NOM} + O_{ACC} (γ*)⁹**

pok'a ar-eal tane-i-n z=na
 3PL.NOM схватить-PTCP уводить-PST-3PL OBJ=3SG.ACC

‘Тогда наконец он предал Его им на распятие. И взяли Иисуса и повели’. [Ин. 19:16/17] (с. 83)¹⁰

В качестве отдельного типа δ выделяются также безличные предикации, однако, по предположению Р. Мейера, при дальнейшем анализе они могут быть отнесены к одному из ядерных типов.

После рассмотрения возможных подходов к объяснению маркирования актантов в классическом армянском, которые предлагались предыдущими исследователями (раздел 4.2.2), Мейер выдвигает собственную гипотезу (раздел 4.2.3): в перифрастическом перфекте наблюдается трехчастная стратегия кодирования глагольных актантов — агенс переходных конструкций (γ; пример 7) маркируется генитивом, пациенс получает двойное

⁸ «История Армении», приписываемая Агатангелосу.

⁹ В этом случае, однако, зачастую нельзя определить, от какого из предикатов объект получает аккузатив — от матричного глагола или от причастия, выполняющего конвербную функцию (с. 83, сн. 38; 153–155).

¹⁰ В переводе этой главы Евангелия от Иоанна на классический армянский отрывок, соответствующий последнему предложению стиха 16 в Синодальном переводе, включен в состав стиха 17.

маркирование морфологическим аккузативом и проклитикой $z=$, а подлежащие непереходных предикаций получают номинатив (α ; пример 4).

Такая ситуация сложилась, по предположению автора, при копировании иранской эргативной модели и конкуренции новой стратегии с номинативно-аккузативной. С морфосинтаксической точки зрения среднеиранский перифрастический перфект также состоит из причастия (на $-d/t$) и бытийной связки. Ср. парфянские примеры:

(9) **Непереходные + S_{DIR} (α)**

'wd 'z 'g-d hym kw 'c bz-kr bwj- 'n
и 1SG.DIR идти-PTCP.PST быть.PRS.1SG чтобы от зло-делать спасти-1SG.SBJV

‘И я пришел, чтобы спасти (тебя) от Грешника, [исцелить (тебя) от боли и принести радость в твоё сердце]’¹¹. [Boyce 1954: 150–151] (с. 117)

(10) **Переходные + A_{ENCL} + O_{DIR} (γ)**

bys 'w's cy=m dyd 'yy 'w=m
но теперь когда=1SG.ENCL увидеть.PTCP.PST быть.PRS.2SG и=1SG.ENCL

tw sxwn 'šnw-d
2SG речь услышать-PTCP.PST

‘Но теперь, когда я увидел тебя и слышал твою речь, [я знаю, что твоя мудрость превосходит мою]’¹². [Sundermann 1981: 89] (с. 118)

Высказанная выше гипотеза подтверждается данными корпусного исследования, результаты которого излагаются в отдельной **главе 5**. Анализ затрагивает пять произведений древнеармянской литературы (перечислены в предполагаемом хронологическом порядке создания): «Житие Маштоца» Корюна (440-е гг. н. э.); «История Армении», приписываемая Агатангелосу (460-е гг. н. э.); упомянутая выше «История Армении Фавтоса Бузанда» (470-е гг. н. э.); «История Армении» Лазаря Парпеци (конец V в. н. э.); «О Вардане и войне армянской» историка и богослова Егише (конец V — начало VI вв. н. э.).

Несмотря на указываемые автором (с. 128–129) неизбежные недостатки подобного корпуса (сравнительно малый объем — около 200 тыс. словоупотреблений, жанровость — в корпусе представлены практически исключительно историографические работы ранних армянских авторов), предположения Мейера об историческом функционировании перифрастического перфекта, высказанные в главе 4, подтверждаются. Маргинальные стратегии маркирования актантов — генитивный субъект в непереходных предикациях (α^* ; пример 5) и номинативный агенс в переходных (γ^* ; пример 7) — показывают четкую диахроническую динамику — падение и рост частотности употребления соответственно, что свидетельствует о происходящей в рассматриваемый исторический период «деэргативизации» перифрастического перфекта (раздел 5.3.2).

Отметим, что в главе 5 также проводится анализ причастий на $-eal$ в атрибутивной (раздел 5.3.1) и конвербной (раздел 5.3.3) позиции. Кроме того, крайне любопытен экскурс в вопросы порядка следования составляющих, представленный в разделе 5.3.4.2.

Так, порядок слов в древнеармянском в целом имеет намного более свободный характер по сравнению с довольно фиксированным порядком AOV в среднеиранских языках, однако данные корпуса показывают, что О-участник переходных предикаций с опущенным субъектом и глаголом в форме перифрастического перфекта оказывается на первом месте в предложении примерно так же часто, как и единственный актант непереходного глагола (208 из 334 употреблений (62 %) для OV/VO, 596 из 804 употреблений (74 %) для

¹¹ Манихейский цикл гимнов «Ангад Рошнан» (Angad Rōšnān), 64a; “And I am come forth to save (thee) from the Sinner, to make (thee) whole from pain, and to bring gladness to thy heart”, цитируется по [Boyce 1954], перевод мой — Д. А.

¹² “Aber nun, da ich dich gesehen hat und deine Rede gehört habe, da weiß ich, dass deine Weisheit der meinen überlegen ist”, цитируется по [Sundermann 1983], перевод мой — Д. А.

SV/VS), что не подтверждается для А-участников при опущении объекта (24 из 51 употреблений, или 47%). Если же в клаузе выражены и А-, и О-участник, наиболее частотными порядками составляющих оказываются AOV, AVO и OVA.

В главе 6 Р. Мейер обращается к другим аспектам синтаксиса классического армянского для получения дополнительных подтверждений гипотезы о значительном контактном влиянии иранских языков на древнеармянский. В разделе 6.2 рассматриваются сходные употребления древнеармянского местоимения *ink'n* и среднеиранского *xwd* (< **xud*) в качестве рефлексива, интенсификатора и анафора, причем способного как к резумптивным употреблениям, так и к использованию для разрешения референциального конфликта (ср. русск. *том*). Последнее использование опционально и в классическом армянском, и в среднеиранских языках.

Подобное совмещение функций одним элементом не зафиксировано в других индоевропейских языках. Раздел 6.3 посвящен древнеармянскому комплементаризму (*e)t'e* и его среднеиранскому аналогу *kw* (< **ku*). Нетривиальное сочетание контекстов, в которых употребляются эти союзы: прямая и косвенная речь, введение sententialных актантов при глаголах восприятия, оформление причинных и пояснительных (clarificatory, ср. англ. *namely*) клауз, — также приводит Р. Мейера к мысли о контактном влиянии на функционирование (*e)t'e*.

Остановимся подробнее на разделе 6.1. В нем рассматриваются употребления древнеармянского относительного местоимения *or*, внешне структурно сходные с изафетной конструкцией в иранских языках:

- (11) *anun Astuac-oy or tear-n=n ararac-oc'*
 имя Бог-GEN.SG который хозяин-GEN=DEF создание-GEN.PL
 '[И] имя бога — владыки созданий [потонуло среди многочисленных имен богов, и не видно было его]'. [ЕК IV.1.21]¹³ (с. 173)
- (12) ПАРФЯНСКИЙ
m'd cy dyw-n
 мать EZ демон-PL
 'мать демонов' [Sundermann 1973: 63] (с. 176)

Тем не менее для полноты картины не хватает древнеармянских примеров, в которых бы содержались несколько модификаторов, присоединяемых сериально при помощи *or*. Как показано в [Haider, Zwanziger 1984: 158–159], в среднеиранских языках фиксируется несколько стратегий присоединения множественных модификаторов (изафетная частица появляется перед каждым модификатором; цепочка модификаторов присоединяется без какого-либо показателя, однако последний из них оформляется сочинительным союзом; изафетный показатель находится перед началом цепочки, а внутри используется сочинительный союз; изафет и сочинение используются одновременно). Такая вариативность, по мнению Г. Хайдера и Р. Цванцигера, лучше всего показывает нестабильность синхронного функционирования изафетного показателя в среднеиранских языках на его пути к грамматикализации, сопровождающаяся потерей его компонентами клаузуальных свойств. В монографии, тем не менее, такие примеры для классического армянского, как и возможность их существования, не упоминаются.

Не приводится и никаких количественных данных о частотности «изафетной» конструкции в классическом армянском. Кроме того, категориальная принадлежность адъюнктированного материала достаточно ограничена — это либо предложные группы (наиболее частый тип), либо именные (но, судя по всему, не прилагательные). Приводимый в монографии (с. 174) пример с причастием, по указанию автора, допускает и клаузуальную трактовку с использованием причастия без глагольной связки в качестве ядра предикации.

¹³ «Книга опровержений» Езника Кохбацци.

Клаузальный анализ, как отмечает сам Р. Мейер, возможен и для примеров вроде (11). К нему подталкивает и наличие сходных придаточных в древнегреческом, классической латыни, хеттском и других индоевропейских языках рассматриваемого периода.

В связи с этим, а также ввиду недостатка языковых данных, автор монографии предлагает компромиссное решение: неглагольные придаточные (т. е. содержащие имя, но не глагол-связку) были унаследованы в древнеармянский из праиндоевропейского, однако тесное контактное взаимодействие со среднеиранскими языками сохранило эту конструкцию в употреблении (в отличие от других индоевропейских языков этого периода, где она перестала использоваться).

Последняя глава 7 возвращается к вопросу (3b) и дает на него следующий ответ: представители династии Аршакидов, правившие в Великой Армении и являвшиеся носителями парфянского языка, перешли на древнеармянский. Иными словами, произошло **изменение суперстрата**.

Новая идентичность была «привязана» к классическому армянскому и принята аршакидскими правителями Армении после падения Парфянского царства и прихода в нем к власти Сасанидов в 224 г. н. э. Отношения с Сасанидами были враждебными, что привело к вытеснению парфянского, становлению армянского в качестве языка при аршакидском дворе и последующему переходу других носителей парфянского на древнеармянский. Таким образом, армянско-иранская контактная ситуация аналогична, например, сложившейся в Англии в XI в. н. э. при Нормандском завоевании, когда правящий нормандский класс постепенно перешел на среднеанглийский (см. раздел 7.3.1).

Конечно, подобное объяснение в отсутствие прямых исторических свидетельств для упомянутого языкового перехода сталкивается с проблемой недоказуемости и неverifiedности, что отмечает и сам Р. Мейер (с. 241), однако позволяет дать более стройное (в сравнении с предыдущими выдвинутыми гипотезами) объяснение возникновению и функционированию армянского перифрастического перфекта.

В заключение подчеркнем, что рецензируемая монография удачно сочетает в себе умелую работу с языковыми и социоисторическими данными, успешно следует целям, поставленным автором, и показывает, что факты языка могут сказать об истории, а факты истории — о языке. Тем не менее отметим, как нам кажется, неудачное решение относительно оформления примеров.

Автор очень вольно относится к межстрочному глоссированию, почти нигде не выделяя морфемы и указывая все граммы «через точку» вместе с переводом слова. Безусловно, не всегда и не в каждой лингвистической работе глоссирование должно быть максимально строгим, но в данном случае при погружении в материал (как минимум) двух потенциально незнакомых читателю языков остается легкое ощущение не только неудовлетворения от не столь полного понимания структуры форм, но и общего неудобства при чтении.

Так, например, в длинных примерах с древнеармянским перифрастическим перфектом, включающим причастие на *-eal*, выделение суффикса причастия могло бы помочь в визуальной идентификации данной глагольной формы. Кроме того, зачастую опускается даже минимальный левый и правый контексты перевода. Неудобство при чтении представляют и многочисленные цитирования неанглоязычных авторов, перевод к которым дается в сносках, но не в основном тексте (а иногда и отсутствует, например, на с. 173).

Это, однако, не портит общего впечатления о монографии, и тем более о ее содержательной части. Надеемся, что ее появление способствует возникновению новых работ в области исторического синтаксиса древнеармянского и среднеиранских языков.

СПИСОК СОКРАЩЕНИЙ

1, 2, 3 — 1, 2, 3 лицо

AVL — отложительный падеж

ACC — винительный падеж

ADV — наречие

AOR — аорист

CAUS — каузатив

DAT — дательный падеж

DEF — определенность

DIR — директив (прямой падеж)

EP — эпентетический сегмент
 EZ — изафетная частица
 ENCL — энклитика
 GEN — родительный падеж
 IO — непрямой объект

NEG — отрицание
 NOM — именительный падеж
 OBJ — прямой объект
 PL — множественное число
 PRER — предлог

PRS — настоящее время
 PST — прошедшее время
 PTCP — причастие
 SUBJ — условное наклонение
 SG — единственное число

СПИСОК ИСТОЧНИКОВ

- Аг — Агатангелос. *История Армении*. Тер-Давтян К. С., Аревшатын С. С. (пер. с древнеарм.). Ереван: Наири, 2004.
 ЕК — Езник Кохбаци. *Книга опровержений (О добре и зле)*. Чалоян В. К. (пер. с древнеарм.). Ереван: Изд-во АН Армянской ССР, 1968.
 ИА — *История Армении Фавстоса Бузанда*. Геворгян М. А. (пер. с древнеарм.), Еремян С. Т. (ред.). Ереван: Изд-во АН Армянской ССР, 1953.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Молчанова 1999 — Молчанова Е. К. Парфянский язык. *Языки мира. Иранские языки. II. Северо-западные иранские языки*. Расторгуева В. С. (ред.). М.: Индрик, 1999, 14–27. [Molchanova E. K. The Parthian language. *Yazyki mira*. Vol. 2. *Severo-zapadnye iranskie yazyki*. Rastorgueva V. S. (ed.). Moscow: Indrik, 1999, 14–27.]
 Bickel, Nichols 2008 — Bickel B., Nichols J. Case marking and alignment. *The Oxford handbook of case*. Malchukov A., Spencer A. (eds.). Oxford Univ. Press, 2008, 304–321.
 Boyce 1954 — Boyce M. *The Manichaean hymn-cycle in Parthian*. Oxford: Oxford Univ. Press, 1954.
 Haider, Zwanziger 1984 — Haider H., Zwanziger R. Relatively attributive: The ‘ezāfe’-construction from Old Iranian to Modern Persian. *Historical syntax*. Fisiak J. (ed.). Berlin: Walter de Gruyter, 1984, 137–172.
 Meyer 2013 — Meyer R. Armeno-Iranian structural interaction: The case of Parthian *wxd*, Armenian *inkʻn*. *Iran and the Caucasus*, 2013, 17(4): 401–425.
 Sundermann 1973 — Sundermann W. *Mittelpersische und parthische kosmogonische und Parabeltexte der Manichäer*. Berlin: Akademie-Verlag, 1973.
 Sundermann 1981 — Sundermann W. *Mitteliranische manichäische Texte kirchengeschichtlichen Inhalts*. Berlin: Akademie-Verlag, 1981.

Получено / received 03.08.2024

Принято / accepted 24.09.2024

ВОПРОСЫ ЯЗЫКОЗНАНИЯ

научный журнал Российской академии наук
 (свидетельство о СМИ ПИ № ФС77-77284 от 10.12.2019 г.)

Оригинал-макет подготовлен С. С. Белоусовым

Адрес редакции: 119019, Москва, ул. Волхонка, 18/2,
 Институт русского языка им. В. В. Виноградова РАН, редакция журнала «Вопросы языкознания»,
 тел.: +7 495 637-25-16, e-mail: voprosy@mail.ru

Подписано к печати 99.99.2024. Дата выхода в свет 99.99.2024. Формат 70×100 1/16.
 Усл. печ. л.: 99,9. Уч.-изд. л. 99,9. Тираж 999 экз. Заказ 999. Цена свободная

Учредители: Российская академия наук, Институт русского языка им. В. В. Виноградова РАН

Издатель: Российская академия наук, 119991 Москва, Ленинский просп., 14

Исполнитель по контракту № 4У-ЕП-03 -24 ФГБУ «Издательство «Наука»

121099, г. Москва, Шубинский пер., д. 6, стр. 1

Отпечатано в ФГБУ «Издательство «Наука»

121099, г. Москва, Шубинский пер., д. 6, стр. 1