
РОССИЙСКАЯ АКАДЕМИЯ НАУК

ВОПРОСЫ ЯЗЫКОЗНАНИЯ



2

2024



НАУКА

— 1727 —

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Отделение историко-филологических наук

ВОПРОСЫ ЯЗЫКОЗНАНИЯ

Журнал основан в январе 1952 года

Выходит 6 раз в год

2

МАРТ — АПРЕЛЬ

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

Ю. Д. Апресян д. ф. н., проф., академик РАН, Институт проблем передачи информации им. А. А. Харкевича РАН; Институт русского языка им. В. В. Виноградова 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; Стокгольмский университет, Швеция

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 — Норвежский арктический университет, 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); Brill Linguistic Bibliography (Online); Cambridge University Press Language Teaching (Online); De Gruyter Saur Dietrich's Index Philosophicus; EBSCOhost MLA International Bibliography (Modern Language Association); Elsevier BV Scopus; European Reference Index for the Humanities and Social Sciences (ERIH PLUS); Gale MLA International Bibliography (Modern Language Association); ProQuest Linguistics and Language Behavior Abstracts (Online), Core; ProQuest MLA International Bibliography (Modern Language Association); Russian Science Citation Index (RSCI); Web of Science Core Collection's Emerging Sources Citation Index (ESCI); Wiley-Blackwell Publishing Ltd. Linguistics Abstracts (Online).

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>

ISSN 0373-658X

© 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

2
MARCH — APRIL

Moscow
2024

Editor-in-chief:

Vladimir A. PLUNGIAN 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
Yury D. APRESJAN Kharkevich Institute for Information Transmission Problems (RAS); Vinogradov Russian Language Institute (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
Dmitrij O. DOBROVOĻ'SKIĬ Vinogradov Russian Language Institute (RAS); Institute of Linguistics (RAS), Moscow, Russia; Stockholm University, Sweden
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. PODGORNIAIA

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: Brill Linguistic Bibliography (Online); Cambridge University Press Language Teaching (Online); De Gruyter Saur Dietrich's Index Philosophicus; EBSCOhost MLA International Bibliography (Modern Language Association); Elsevier BV Scopus; European Reference Index for the Humanities and Social Sciences (ERIH PLUS); Gale MLA International Bibliography (Modern Language Association); ProQuest Linguistics and Language Behavior Abstracts (Online), Core; ProQuest MLA International Bibliography (Modern Language Association); Rossiiskii indeks nauchnogo tsitirovaniya (RINTs); Russian Science Citation Index (RSCI); Web of Science Core Collection's Emerging Sources Citation Index (ESCI); Wiley-Blackwell Publishing Ltd. Linguistics Abstracts (Online).

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>

ISSN 0373-658X

Содержание

С. О. Савчук, Т. А. Архангельский, А. А. Бонч-Осмоловская, О. В. Дониная, Ю. Н. Кузнецова, О. Н. Ляшевская, Б. В. Орехов, М. В. Подрядчикова. Национальный корпус русского языка 2.0: новые возможности и перспективы развития	7
N. A. ZEYAKHINA, A. A. SHCHIRKOVA. Asymmetry in Russian metalinguistic comparatives: Corpus and experimental evidence	35
В. И. Подлесская. Скажем так: дискурсивные маркеры, восходящие к глаголам речи в русском языке	52
М. А. Кронгауз. Чередование <i>а</i> и <i>о</i> в основе и прагматика распределения вариантов	83
Ю. А. Ландер, Ш. Ш. Унарокова. Принципы выделения местоименных серий (на материале адыгейских демонстративов)	93
И. А. Зибер. Чукотские сонорные согласные в типологической перспективе	122

Обзоры

Ю. В. Николаева. «Язык и семиотика тела». Исследования Г. Е. Крейдлина по воплощенной семиотике	143
---	-----

Рецензии

Т. А. Майсак [Рец. на:] L. Johanson. <i>Code Copying. The strength of languages in take-over and carry-over roles</i> . Leiden: Brill, 2023	159
---	-----

Contents

Svetlana O. SAVCHUK, Timofey ARKHANGELSKIY, Anastasiya A. BONCH-OSMOLOVSKAYA, Ol'ga V. DONINA, Yuliya N. KUZNETSOVA, Ol'ga N. LYASHEVSKAYA, Boris V. OREKHOV, Mariya V. PODRYADCHIKOVA. Russian National Corpus 2.0: New opportunities and development prospects	7
Natalia A. ZEVAKHINA, Alina A. SHCHIPKOVA. Asymmetry in Russian metalinguistic comparatives: Corpus and experimental evidence	35
Vera I. PODLESSKAYA. <i>Skažem tak</i> : Discourse markers originating from verbs of speech in Russian	52
Maxim A. KRONGAUZ. The alternation of stem-internal <i>a</i> and <i>o</i> and pragmatic mechanisms of alternant distribution	83
Yuri A. LANDER, Shamset Sh. UNAROKOVA. Defining pronominal series: Demonstratives in West Circassian	93
Inna A. SIEBER. Chukchi sonorant consonants in a typological perspective	122

Overviews

Yulia V. NIKOLAEVA. "Language and body semiotics". Grigory E. Kreidlin's research on embodied semiotics	143
---	-----

Reviews

Timur A. MAISAK [Review of:] L. Johanson. <i>Code Copying. The strength of languages in take-over and carry-over roles</i> . Leiden: Brill, 2023	159
--	-----

Национальный корпус русского языка 2.0: новые возможности и перспективы развития

© 2024

Светлана Олеговна Савчук

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

Тимофей Александрович Архангельский

Университет Гамбурга, Гамбург, Германия; timarkh@gmail.com

Анастасия Александровна Бонч-Осмоловская

Национальный исследовательский университет «Высшая школа экономики», Москва, Россия;

Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия;

abonch@gmail.com

Ольга Валерьевна Донина

Воронежский государственный университет, Воронеж, Россия; olga-donina@mail.ru

Юлия Николаевна Кузнецова

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

Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия;

kuznetsova.yn@gmail.com

Ольга Николаевна Ляшевская

Национальный исследовательский университет «Высшая школа экономики», Москва, Россия;

Институт русского языка им. В. В. Виноградова РАН, Москва, Россия; olesar@yandex.ru

Борис Валерьевич Орехов

Национальный исследовательский университет «Высшая школа экономики», Москва, Россия;

nevmenandr@gmail.com

Мария Владимировна Подрядчикова

независимый исследователь; mpodr2015@gmail.com

Аннотация: В статье подводятся итоги проекта фундаментальной реконструкции и модернизации платформы Национального корпуса русского языка, осуществленного в 2020–2023 гг. В фокусе статьи новые возможности, которые открываются для лингвистов и более широкой аудитории, в частности, улучшение репрезентативности имеющихся корпусов, создание новых корпусов, новая разметка, полученная с помощью применения нейросетевых моделей, новые интерфейсные решения. Более детально рассматриваются три ярких новых компонента: ресурсный — новый корпус «Социальные сети», поисковый — Панхронический корпус, объединяющий поиск по корпусам разных периодов, и аналитический — функциональный комплекс статистики и визуализации данных.

Ключевые слова: количественная лингвистика, корпусная лингвистика, русский язык

Благодарности: Исследование проводилось в рамках работ, поддержанных грантом Министерства науки и высшего образования № 075-15-2020-793. Авторы выражают свою признательность за ценную помощь и плодотворное сотрудничество Д. В. Сичинаве, А. Н. Дышканту, С. Ю. Толдовой, Н. С. Горбунову, Д. А. Фурсиной, А. А. Маховой, С. В. Пискуновой, Н. Н. Буйловой,

Д. Г. Бородиной, А. Д. Козеренко, И. И. Виноградовой, С. А. Гладилину, Д. А. Морозову, В. Г. Сизову, П. В. Дяченко, А. О. Казенникову, Н. А. Власовой, А. В. Глазковой, С. С. Столярову, Т. А. Гаринову, И. А. Смалю.

Для цитирования: Савчук С. О., Архангельский Т. А., Бонч-Осмоловская А. А., Донина О. В., Кузнецова Ю. Н., Ляшевская О. Н., Орехов Б. В., Подрядчикова М. В. Национальный корпус русского языка 2.0: новые возможности и перспективы развития. *Вопросы языкознания*, 2024, 2: 7–34.

DOI: 10.31857/0373-658X.2024.2.7-34

Russian National Corpus 2.0: New opportunities and development prospects

Svetlana O. Savchuk

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

Timofey Arkhangelskiy

Hamburg University, Hamburg, Germany; timarkh@gmail.com

Anastasiya A. Bonch-Osmolovskaya

HSE University, Moscow, Russia; Kharkevich Institute for Information Transmission Problems,
Russian Academy of Sciences, Moscow, Russia; abonch@gmail.com

Ol'ga V. Donina

Voronezh State University, Voronezh, Russia; olga-donina@mail.ru

Yuliya N. Kuznetsova

Lomonosov Moscow State University, Moscow, Russia; Kharkevich Institute for Information
Transmission Problems, Russian Academy of Sciences, Moscow, Russia; kuznetsova.yn@gmail.com

Ol'ga N. Lyashevskaya

HSE University, Moscow, Russia; Vinogradov Russian Language Institute,
Russian Academy of Sciences, Moscow, Russia; olesar@yandex.ru

Boris V. Orekhov

HSE University, Moscow, Russia; nevmenandr@gmail.com

Mariya V. Podryadchikova

independent researcher; mpodr2015@gmail.com

Abstract: The paper provides an overview of the results of the fundamental reconstruction and modernization project of the National Corpus of the Russian Language platform, carried out from 2020 to 2023. The focus of the paper is on the new opportunities that are opening up for linguists and a wider audience. This includes improving the representativeness of existing corpora, creating new corpora, new annotation obtained through the application of neural network models, and new interface solutions. Three notable new components are examined in more detail: a resource-related one, which is the new Social Networks corpus, a search-related one, which is the Panchronic corpus that combines searches across corpora from different periods, and an analytical one, which is the functional complex of statistics and data visualization.

Keywords: corpus linguistics, quantitative linguistics, Russian

Acknowledgements: The current research was a part of research work supported by the Ministry of Science and Higher Education grant No. 075-15-2020-793. The authors express their gratitude for the valuable assistance and fruitful cooperation to D. V. Sichinava, A. N. Dyshkant, S. Y. Toldova, N. S. Gorbunov, D. A. Fursina, A. A. Makhova, S. V. Piskunova, N. N. Builova, D. G. Borodina, A. D. Kozerenko, I. I. Vinogradova, S. A. Gladilin, D. A. Morozov, V. G. Sizov, P. V. Dyachenko, A. O. Kazennikov, N. A. Vlasova, A. V. Glazkova, S. S. Stolyarov, T. A. Garipov, I. A. Smal’.

For citation: Savchuk S. O., Arkhangelskiy T., Bonch-Osmolovskaya A. A., Donina O. V., Kuznetsova Yu. N., Lyashevskaya O. N., Orekhov B. V., Podryadchikova M. V. Russian National Corpus 2.0: New opportunities and development prospects. *Voprosy Jazykoznanija*, 2024, 2: 7–34.

DOI: 10.31857/0373-658X.2024.2.7-34

Введение

Национальный корпус русского языка был открыт для публичного использования в 2004 г. [Сичинава 2005]. За прошедшее время корпус увеличился многократно и по своим объемам, и по разнообразию представленных данных и типов аннотации, однако к 2020 г. стало понятно, что дальнейшее развитие Национального корпуса тормозится из-за накопившихся нерешенных технологических проблем: отсутствия универсальной платформы для развертывания лингвистических корпусов, медленной индексации и нерегулярных обновлений, ограниченности поисковых возможностей, негибкого, плохо адаптируемого интерфейса взаимодействия с корпусной базой. Кроме технологических сложностей, была еще и концептуальная сложность, заключавшаяся в изолированном существовании разных корпусов в составе НКРЯ, из-за этого процессы подготовки и загрузки данных требовали дополнительных усилий программистов. Накопилось и «содержательное» отставание, проявлявшееся в недостаточной репрезентативности диахронических данных основного корпуса, и явная недостаточность статистического функционала относительно современных стандартов корпусной лингвистики. Иными словами, в результате более чем 15 лет своего плодотворного развития, Национальный корпус русского языка оказался в той точке, когда критическим образом стали необходимы глубокие системные и концептуальные решения о том, как должна быть в дальнейшем организована экосистема корпуса, которая бы обеспечивала возможности по его поддержке, обновлению и внедрению современных корпусных технологий. В 2020 г. началась работа научного коллектива, созданного на базе консорциума пяти научных институтов и вузов и поддержанная субсидией Министерства науки и образования. Глобальной целью проекта стало создание компьютерно-лингвистической платформы нового поколения, ядром которой является Национальный корпус русского языка (НКРЯ). Работы по проекту были сгруппированы в три взаимосвязанных направления:

- инфраструктурное направление, в рамках которого была заложена программно-технологическая основа для создаваемой компьютерно-лингвистической платформы;
- ресурсное направление, обеспечивавшее развитие и пополнение всех корпусов платформы;
- научно-исследовательское направление, в котором созданная платформа была апробирована в конкретных лингвистических исследованиях.

Цель настоящей статьи — очертить наиболее значимые изменения, связанные с новым этапом развития Национального корпуса русского языка. В фокусе статьи будут в первую очередь новые возможности, которые открываются для лингвистов и для более широкой аудитории. В части 1 представлен общий краткий обзор модернизации НКРЯ, части 2, 3 и 4 посвящены наиболее интересным и многообещающим обновлениям: новому корпусу «Социальные сети» (2), Панхроническому корпусу (3) и инструментам статистического анализа и визуализации данных (4).

1. Общий обзор изменений Национального корпуса русского языка

Работа над так называемым «Корпусом 2.0» — технологически и концептуально обновленным (а в части архитектуры даже пересобранном) Национальным корпусом русского языка — велась в нескольких направлениях: технической модернизации программной архитектуры, ресурсного пополнения и обновления данных и разметки. Эти направления тесно взаимосвязаны: без технологической перестройки корпусной платформы невозможно было выстроить процесс оперативного пополнения корпусов (каждый корпус раньше существовал изолированно и требовал сложной настройки), устаревшая модель интерфейса ограничивала возможности по развитию разметки корпусов, добавлению новых типов метаданных и новых поисковых возможностей. Принципиально важна и обратная связь — данные, подготовленные с помощью нового инструмента автоматической разметки и снятия омонимии, позволили внедрить инструменты статистического анализа. Таким образом, Корпус 2.0 стал мощным аналитическим инструментом и на уровне отдельных лемм (портрет слова), и на уровне их сочетаний (простых частотностей и коллокаций), и на уровне отбираемых пользователем подкорпусов (портрет подкорпуса). Ниже мы кратко перечислим наиболее значимые изменения, произошедшие в Корпусе на этапе 2.0.

1.1. Пополнение корпусов, новые корпуса

К концу 2022 г. общий объем коллекций текстов НКРЯ превысил 2 млрд словоупотреблений. Благодаря увеличению базы текстов корпуса на сотни миллионов слов Корпус 2.0 теперь представляет собой наиболее полную и разнообразную коллекцию текстов на русском языке, собранную в одном месте и объединенную удобным интерфейсом.

Особое внимание было уделено улучшению показателей репрезентативности в области жанров и периодов, ранее мало представленных в Корпусе. Так, коллекции нехудожественной прозы и публицистики **Основного корпуса** были пополнены текстами научных трудов XVIII — начала XIX в., были добавлены документы, научные работы, описания путешествий, подборки мемуаров, публицистика (в том числе из журналов «Сын отечества», «Московский наблюдатель», «Вестник Европы»), тексты дневников и воспоминаний XIX–XXI вв. из проекта «Прожито» и др. С другой стороны, увеличена представленность образцов современного языка, в частности была добавлена проза XX — начала XXI в., коллекция современных путеводителей, собрание современных научных текстов разных жанров (тезисы, программы, учебные пособия, задачи, конспекты), коллекция производственно-технических инструкций и пособий. Коллекции **Газетного корпуса** расширили временной охват: теперь они включают хронологический диапазон с 1983 по 2021 г. Газетный корпус достиг более 800 млн словоупотреблений.

Существенно были расширены коллекции корпусов со специальной разметкой, причем принципиально важным является не только увеличение объема, но и качественные характеристики источников, расширение репрезентативности по параметрам, определяемым спецификой корпуса. Так, в **Диалектный корпус** были добавлены тексты из различных регионов и диалектных зон: северные говоры (Архангельская область), среднерусские (Тверская область), южнорусские (Смоленская, Тамбовская области), говоры позднего формирования (Поволжье, Урал, Сибирь). Расширение **Параллельных корпусов** коснулось не только существенного увеличения имеющихся коллекций и жанрового разнообразия представленных в них текстов, но и создания новых языковых пар, таких, например, как сербско-русская, румынско-русская, корейско-русская, хинди-русская. В **Древнерусский корпус** были добавлены знаменитые памятники русской литературы, такие как «Сказание

о Борисе и Глебе», «Поучение Владимира Мономаха», «Слово о полку Игореве» и др., а также коллекции пергаменных и бумажных деловых документов. В **Мультимедийный корпус** были добавлены большие коллекции текстов устной научной речи (доклады на конференциях, учебные и популярные лекции, теле- и радиопередачи), устной политической речи (интервью, пресс-конференции, выступления на митингах, собраниях и съездах, ток-шоу на радио и ТВ и мн. др.). Пополнение **Поэтического корпуса** включало в себя тексты поэтов второй половины XX в., а также большую коллекцию русских переводов античной поэзии: «Илиаду» Гомера в переводе Н. И. Гнедича, «Энеиду» Вергилия в переводах В. Я. Брюсова и С. М. Соловьева и сатиры Горация в переводе А. А. Фета.

Наиболее радикально обновлен **Обучающий корпус**, предназначенный для использования в школьном преподавании русского языка и литературы, в него было добавлено более 1000 новых текстов. Теперь в Обучающем корпусе есть все основные произведения из российской школьной программы по русской литературе, включая те, которые рекомендуются для внеклассного чтения. В дополнение к собственно пополнению коллекции Обучающего корпуса был разработан раздел «Упражнения на основе Корпуса», в котором представлены упражнения, составленные на материале Обучающего корпуса и других корпусов НКРЯ. Упражнения относятся к разным разделам школьного курса русского языка и предназначены для самостоятельной работы на уроке и дома, а также для контроля знаний.

Кроме пополнения уже имеющихся корпусов были подготовлены и опубликованы новые корпуса. Это, во-первых, **корпус «Русская классика»**. Корпус включает художественные, публицистические и эпистолярные произведения из собраний сочинений русских классических писателей. Произведения русских классических писателей имеют особый статус для истории русского литературного языка. Если считать, что литературный язык — такой, который «обработан мастерами», тексты этих мастеров и составляют ядро корпуса русского литературного языка. С таким корпусом можно сверяться как с нормативным, а не узусным источником, из него можно извлекать авторитетные примеры для академических грамматик, словарей и учебных пособий.

Во-вторых, это **корпус «От 2 до 15»**. Корпус включает в себя литературу на русском языке, которую читают современные дети и подростки. Тексты подобраны по результатам масштабных опросов детей, подростков, учителей и родителей. Каждый текст размечен в соответствии с возрастом, в котором его обычно наиболее интересно читать. Для автоматической разметки фрагментов текстов по минимальному возрасту, в котором они предположительно будут понятны читателям, была создана нейросетевая модель [Morozov et al. 2022].

В-третьих, был подготовлен **корпус «Восточнославянская эпиграфика»**. Основу коллекции эпиграфических текстов в составе НКРЯ составляют средневековые надписи, собранные в базе www.epigraphica.ru. Для корпуса отобраны только ранее опубликованные тексты XI–XV вв., написанные на славянских языках, а также содержащие азбуки или цифры. В состав первого релиза корпуса входят 663 надписи общим объемом более 5200 слов. Корпус снабжен пословной морфологической разметкой со снятой омонимией.

Наконец, еще одним новым корпусом в семействе корпусов НКРЯ стал **корпус «Социальные сети»**. Этот корпус, основанный на данных VK и Telegram, а также ряда других источников, отражает живые языковые изменения за пределами литературного языка. Такие языковые явления почти не фиксируются в текстах Основного и Газетного корпусов. Более подробно особенности и структурные характеристики корпуса социальных сетей будут представлены в разделе 2.

1.2. Использование нейросетевых моделей для разметки текстов корпуса

Еще одним важным шагом в развитии НКРЯ стал переход на использование нейросетевых моделей для подготовки и анализа данных. Наиболее значимым результатом стало

развитие модели Rubic [Lyashevskaya et al. 2023], которая использовалась для автоматического снятия омонимии и разметки новых данных Основного и газетных корпусов. Модель использует стандарт разметки CONLL-U [Straka et al. 2016], совмещающий морфологическую и синтаксическую разметку. На сегодняшний день этот формат наиболее широко используется для подготовки корпусных данных для задач NLP, а также для корпусной лингвистики в целом. В России распространение CONLL-U было во многом стимулировано соревнованием морфосинтаксических парсеров GRAMEVAL, прошедшим в 2020 г. [Lyashevskaya et al. 2020]. Ранее в Основном корпусе противопоставлялись тексты со снятой (вручную) морфологической омонимией (около 6 млн словоупотреблений) и основная масса текстов, где были представлены все возможные разборы; в большинстве корпусов со специальной разметкой присутствовала только неснятая омонимия. В результате использования нейросетевой модели для Основного и газетных корпусов это противопоставление фактически снято: действительно, при нынешних объемах Корпуса морфологическая омонимия в нем может быть разрешена только программными методами. Автоматическое снятие омонимии открывает возможности для развития новых статистических инструментов, таких как поиск по коллокациям, поиск семантически близких слов (подробнее об этих инструментах будет говориться в разделе 4). Совмещение морфологической и синтаксической разметки, характеризующее используемый стандарт CONLL-U, расширяет поисковый функционал, позволяя включить на больших объемах текстов инструменты синтаксического поиска (типы и направления синтаксических связей, синтаксические роли слов), а также выявить для каждого слова так называемые «скетчи» — устойчивые словосочетания с заданными синтаксическими отношениями (см. раздел 4). Эксперименты с нейросетевыми моделями проводились и для абсолютно новых задач, так, например, с помощью модели *rutermextract*¹ были размечены ключевые слова в текстах корпуса региональных СМИ. Одно ключевое слово (например, для газетной заметки о росте травматизма в зимние праздники) может состоять из однословного ключа (*праздник, переломы*) либо из двусловного сочетания (*таяние снега*). Разметка по жанрам корпуса социальных сетей также была сделана автоматически с помощью обучения нейросети (см. подробнее раздел 2). Наконец, еще одно новое экспериментальное направление приложения нейросетевых моделей связано со словообразовательным разбором. Автоматические разборы были сгенерированы нейросетевым алгоритмом (модель НейроКРЯ), ядром которого является сверточная нейросеть, архитектура которой была предложена в работе [Sorokin, Kravtsova 2018]. Нейросеть была обучена на двух источниках данных. Разметка морфем в Обучающем корпусе опирается на [Тихонов 2002], содержащий около 100 тыс. лексем. Для каждого слова указан список морфем, их тип (приставка, корень, интерфикс, суффикс, окончание или постфикс) и линейная позиция в слове. В основе разметки словообразовательной структуры в Основном корпусе лежит специально разработанный для корпуса словарь морфемного анализа, где по состоянию на май 2023 г. даны разборы для 75 тыс. лексем (310 тыс. неуникальных морфем) [Ляшевская и др. 2009]. Словарь создан на основе словаря морфем [Кузнецова, Ефремова 1986]. Применение модели НейроКРЯ в Основном корпусе позволяет объединять в словообразовательные гнезда даже те слова, которых нет в словаре, т. е. в исходных обучающих данных.

В целом экспериментальное направление приложения нейросетевых моделей для подготовки данных и их анализа является сегодня наиболее перспективным для развития всей экосистемы Национального корпуса русского языка, особенно с учетом объема и масштаба его пополнений. Ожидается, что с дальнейшим развитием возможностей нейросетевого моделирования процесс подготовки новых данных будет значительно упрощен, а аналитический функционал корпуса получит дополнительные возможности и для новых, и для старых данных.

¹ <https://github.com/igor-shevchenko/rutermextract>

1.3. Новые интерфейсные решения

Внешний облик Национального корпуса русского языка претерпел существенные изменения, причем эти изменения коснулись не только дизайна страницы в интернете, но и общего подхода к представлению информации о Корпусе. Обновление интерфейса началось с концептуально новой главной страницы, дающей сводную информацию о всех корпусах и их объеме, а также подробного пользовательского руководства. Поисковая строка на главной странице — это так называемый «обзор возможностей»: простой запрос, подобный обычному запросу в поиске, ведет пользователя на страницу агрегированных результатов, которые кроме собственно выдачи-конкорданса включают в себя график Панхронического корпуса (см. подробнее раздел 3), случайное стихотворение из Поэтического корпуса со словом из запроса, карточку свойств слова с его грамматическими и семантическими характеристиками и список коллокаций из Основного корпуса. Таким образом, это набор иллюстраций, «на лету» демонстрирующий, какую информацию можно найти в разных корпусах НКРЯ по простому запросу. Были усовершенствованы формы поиска и отбора подкорпуса, в частности, в форму поиска добавлена возможность поиска по словоформе (в ряде корпусов — также с использованием регулярных выражений), а при отборе подкорпуса появилась возможность выбрать диапазон дат обновления версий корпуса. Эта функция важна при воспроизведении результатов корпусного исследования, поскольку при пополнении корпуса показатели частотности могли измениться. Цель многочисленных изменений интерфейса — сделать работу с корпусом более понятной и удобной. Стоит отметить, что за этими изменениями стоит еще и глобальная перестройка всей внутренней архитектуры корпуса, в результате которой появилась принципиальная возможность гибкой настройки и адаптации интерфейса под специфические особенности корпусов [Гладилин, Козеренко 2022]. Постепенно все корпуса Национального корпуса русского языка были переведены на новый интерфейс.

Основные направления развития НКРЯ, очерченные выше, объединила общая задача — создание новой корпусной платформы, в которой, с одной стороны, интегрированы и оптимизированы уже имеющиеся корпуса и концептуальные решения, а с другой стороны, реализован модульный подход, обеспечивающий необходимую гибкость и доступность для изменения и дальнейшего развития. Для рассказа обо всех важных результатах, достигнутых в рамках проекта, объема настоящей статьи недостаточно. Среди наиболее интересных и значимых результатов проекта, которые мы бы хотели осветить подробнее, выделим три: новые корпусные данные, введенные в научный оборот с помощью корпуса «Социальные сети», объединение текстов исторических и современных корпусов в формате Панхронического корпуса, а также разнообразные инструменты статистики и визуализации, обеспечивающие возможности применения современных количественных методов для анализа корпусных данных.

2. Корпус «Социальных сетей»

2.1. Электронная коммуникация в НКРЯ

Сфера электронной коммуникации (или компьютерно-опосредованного общения, язык интернета) стала предметом лингвистического внимания с момента массового распространения интернета, см. [Иванов 2000; Бергельсон 2002; Кузьмина 2003; Трофимова 2004; Капанадзе 2005; Горошко 2007; Какорина 2008] и др. В исследованиях начала XXI в. поднимались вопросы о специфике этого модуса существования языка: гипертекстовой природе

текста в среде сети Интернет, его гибридном характере, о тенденциях в орфографии, языковой игре и др. При создании Национального корпуса русского языка сфера электронной коммуникации была указана среди основных сфер функционирования, и тексты электронной коммуникации были включены в состав Основного корпуса письменных текстов (самые ранние датированы 2002 г.). В корпусе представлены образцы существовавших в тот период типов компьютерно-опосредованной коммуникации: форумы, чаты в ICQ, записи в «Живом журнале», СМС-сообщения, комментарии в СМИ. В дальнейшем, по мере развития этой сферы общения, в корпус включались новые типы текстов — ленты новостей и групповые чаты, форумы по интересам, блоги на других платформах.

Учитывая, что тексты электронной коммуникации представляют собой особое подмножество в составе Основного корпуса письменных текстов, в среднем существенно сильнее нормированном орфографически, при подготовке текстов проводилась предварительная нормализация орфографии, при которой всем формам с отклонениями от стандартных написаний приписывалась правильная стандартная форма, которая в дальнейшем получала морфологическую аннотацию и участвовала в лексико-морфологическом поиске. Благодаря этой предварительной работе в подкорпусе электронной коммуникации по запросу какой-либо лексемы мы можем получить разные варианты ее написания, например: по запросу леммы *ничто* будут выдаваться контексты, среди которых встретятся формы *ничего, ничо, ничё, ниче*; по запросу *красавчик* среди контекстов встретим написания *красавчег, красавчеге*.

Помимо нормализации орфографии стандарт подготовки текстов предполагал ручное редактирование и метатекстовую разметку: удаление повторов, разметку ников и реплик в форумах и др. Все это делает поиск в корпусе более точным, однако существенный минус такого подхода — трудозатратность при подготовке текстов, которая по своей сложности приближается к оформлению транскриптов устной речи. Как следствие медленный процесс пополнения корпуса перестал соответствовать темпу развития электронной коммуникации, которая с распространением социальных сетей (в русскоязычном сегменте интернета примерно начиная с 2006 г.) испытала взрывной рост.

Для изучения сферы электронной коммуникации в ее современном состоянии создан корпус социальных сетей, опирающийся на иные принципы отбора, подготовки и организации материала.

2.2. Принципы создания корпуса «Социальные сети»

В разработке концепции корпуса принимали участие Б. В. Орехов, С. О. Савчук, Д. В. Сичинава. В отличие от небольшого подкорпуса электронной коммуникации (текущий объем не превышает 3,4 млн словоупотреблений), которому отведена определенная доля в общем массиве текстов Основного корпуса, корпус «Социальные сети» должен иметь **большой объем** (от 100 млн словоупотреблений) и регулярно пополняться, что позволит получать при его использовании статистически значимые результаты. Объем корпуса на конец 2023 г. составлял 157 млн словоупотреблений.

Представительность (репрезентативность) корпуса социальных сетей обеспечивается тем, что в его состав включены тексты разных типов интернет-коммуникации: интернет-форумы, записи в блогах, сообщения в мессенджерах, при этом ставится задача как можно более широкого охвата социальных сетей, популярных форумов и каналов известных блогеров. **Региональная репрезентативность** корпуса состоит в том, что наряду с общероссийскими социальными сетями в него включены текстовые коллекции локальных сетей — местные форумы, группы VK и Telegram, популярные в регионе блогеры. В настоящее время в корпусе присутствуют коллекции текстов локальных соцсетей Воронежа и Воронежской области (Большой Воронежский форум, Воронежский рыболовный

форум, группы VK «Типичный Воронеж», «Регион-36» и др.), подготовленные в Воронежском государственном университете Н. С. Горбуновым, Д. А. Фурсиной, П. Д. Есиповой, А. Ю. Луценко, А. С. Шудриковой и др. [Дониная и др. 2024 (в печати)]. В следующем году планируется расширить региональную коллекцию материалами популярных блогеров и форумов, собранную воронежскими коллегами в соцсетях соседних Курской, Ростовской и Тамбовской областей.

Тематическое разнообразие достигается за счет включения в корпус текстов, относящихся к разным тематическим областям, что в случае социальных сетей не представляет трудности, поскольку в настоящее время свои страницы в интернете имеют органы власти и управления, предприятия и организации, группы по интересам, профессиональные сообщества, а современные блогеры вышли далеко за рамки личных дневников — сегодня люди ведут не только персональные блоги, но и рассчитанные на широкую публичность: корпоративные, экспертные, тематические, новостные, кулинарные, туристические, музыкальные, игровые, автомобильные. Многие из этого тематического разнообразия представлено в текущей версии корпуса, например, «Чат для художников», «Ворон и Ёжка. Почти серьезный канал о жизни Воронежа» в Telegram, «kolokolschool. Гончарная школа», «Благотворительный фонд Милосердие», «Привет, Воронеж. Новости» и др. в VK.

2.3. Материал и способ представления данных

Планируемый большой объем корпуса социальных сетей может быть получен только при автоматическом сборе и автоматической обработке материала. Все тексты взяты из открытых источников: VK, Telegram, Livejournal, Liveinternet, Blogspot, Большой Воронежский форум и др. В сборе материала принимали участие Б. В. Орехов (сбор и обработка), Е. И. Пискунова, А. Б. Хазова, группой студентов и сотрудников ВГУ под руководством О. В. Дониной подготовлена воронежская коллекция.

Подготовка электронных версий текстов включает очистку их от html-разметки, элементов верстки веб-страниц и снабжение текстов xml-разметкой, используемой в корпусе; поиск и удаление дублей — повторяющихся текстов, которые неизбежно присутствуют в социальных сетях; поиск и удаление текстов на иностранных языках — вся эта работа выполняется с помощью программных средств. Подобные проблемы решались при создании газетных корпусов в составе НКРЯ, с ними сталкиваются все составители веб-корпусов.

В базе данных корпуса тексты организованы по-разному. Для основной массы текстов один документ включает один текст — пост за определенную дату. Для части текстов сохранена их диалоговая природа: один документ включает пару сообщений — исходный пост и комментарий к нему, они также имеют точную датировку. Часть текстов, в основном это многостраничные форумы из воронежской коллекции, представлена в одном документе без разделения на отдельные сообщения — они объединены общей темой, но точная датировка таких документов невозможна, она представляет собой интервал в несколько дней, месяцев и даже лет, если общение на форуме продолжается. Некоторый разноречивый в датировке может отражаться на точности статистической обработки, например, при отображении распределения результатов поиска по датам, поэтому в дальнейшем предполагается унифицировать способ подачи дат текстов.

Ручная орфографическая нормализация текстов не проводилась, так что все они представлены в оригинальной орфографии, см. примеры (1) и (2).

- (1) *Ребятки, 2.08. 15 в 18: 00 я жду всех-всех-всех на афтепати к fđwg и препати к моему дэрэ-))) буду рада видеть всех желающих сделать чин-чин во славу фитнеса и за мое здоровье☺☺регистрация не обязательно, количество мест не ограничено-))) про-сто приходим и.... место: pub daddy, ул. Солдатская, 6-А [vk (28.07.2015)]*

- (2) *Кароче, все ж уже понятно. Давайте так. Вы как настоящий школьник еще раз провозгласите, что я слился, и на этом мы прекратим душный разговор ни о чем. Второй вариант, вы возьмете себя в руки, перечитаете чат, и будете общаться по существу. Вариант со школьником, мне в вашем случае представляется наиболее вероятным* 😊😊 [Rozetked Discuss. telegram Rozetked Discuss (10.07.2021)]

Таким образом, при лемматизации формы слова в нормативном и ненормативном (нестандартном) написании не объединяются в одну единицу и не опознаются как ее варианты. Поэтому для того, чтобы получить полное представление обо всех вхождениях слова, нужно учитывать все варианты его написания и осуществлять поиск по каждому из них. Например, с помощью двух запросов по словоформе со звездочкой *короч** и *короч** получаем набор вариантов написания вводного слова, с которого начинается второй пример: *короче, короч, кароче, карочи, кароч, карочь, карочки, карочеееее, карочееееееее*. Как для автоматически определенных лемм, так и для словоформ при вводе с клавиатуры в форме запроса появляется всплывающая подсказка (suggest) — список присутствующих в корпусе последовательностей с таким началом; это позволяет пользователю ориентироваться в богатстве лексики и орфографии корпуса социальных сетей.

В дальнейшем предполагается опробовать на текстах социальных сетей механизм отождествления различных орфографических вариантов слова, который использовался при разработке Панхронического корпуса (см. раздел 3 настоящей статьи). В результате все возможные варианты написания слова можно будет получить по одному запросу на любой из вариантов.

2.4. Лингвистическая разметка и организация поиска

В корпусе используются два вида разметки — морфологическая разметка словоформ и метаразметка текстов. Морфологическая разметка выполнена на основе Mystem (<https://yandex.ru/dev/mystem/>) с неснятой морфологической омонимией. Это дает повышенную долю шума при поиске в сравнении с подкорпусом электронной коммуникации в составе основного корпуса, особенно в случае несловарных слов и нестандартных написаний. Однако, как уже говорилось, отказ от нормализации орфографии дает большой выигрыш во времени при подготовке текстов, а качество автоматической лемматизации и снятия морфологической неоднозначности в текстах с нестандартной орфографией и пунктуацией еще требует дополнительных исследований и оценки.

Метаразметка текстов включает минимум параметров: дата создания текста, интернет-платформа, на которой опубликован текст, регион охвата соцсети, тип текста, жанр текста, автор (или его условное имя) и название текста (последние два параметра определены для части текстов). Отбор подкорпуса возможен по дате, в том числе можно задать диапазон дат, по названию социальной сети или блог-платформы, по региону, по типу и жанру текста. Отбор подкорпуса по типу текста позволяет искать отдельно в исходных постах, то есть текстах автора блога, и в комментариях пользователей и подписчиков. В том случае, если выбран комментарий, текст исходного поста будет виден, но поиск по нему не ведется. Поиск по жанрам возможен благодаря экспериментальной разметке, выполненной автоматически для основного массива текстов корпуса, на чем необходимо остановиться подробнее.

2.5. Автоматическая разметка жанров

Социальные сети представляют большой интерес для исследователей речевых жанров, что отражено в многочисленных работах. Жанры интернета (или жанры 2.0, как их еще

называют) рассматриваются в сопоставлении с жанрами в других сферах коммуникации (прежде всего разговорной, публицистической, художественной); отмечается подвижность системы интернет-жанров, их зависимость от технологических платформ, вариативность, изменчивость; предлагаются различные классификации жанров, обсуждается роль корпусных методов в изучении речевых жанров, ср. [Щипицина 2009; Шмелева 2012; Дементьев 2016; Литвиненко 2016; Горошко, Землякова 2017; Кириллов 2017; Шилихина 2018; Карасик 2019; Егорова 2021] и др. Накоплен богатый материал, однако до общепринятой типологии, как признают исследователи, еще далеко.

Большинство исследователей сходятся в том, что форум, блог, канал в соцсетях представляют собой сложные образования — гипер- или мегажанры. Отдельные же сообщения, которыми обмениваются участники общения в форуме, авторские записи в блогах и комментарии к ним могут иметь разную жанровую природу: сообщать новость, привлекать внимание к какому-то событию, объявлять о каком-то мероприятии, рекомендовать какую-то покупку, давать полезные советы, и наоборот, содержать просьбу о помощи. В системе метатекстовой разметки НКРЯ определение жанровой принадлежности текста, как и его тематики, относится к базовым текстовым характеристикам и производится вручную. На материале корпуса социальных сетей было решено провести эксперимент по автоматическому определению жанровой принадлежности текстов с использованием нейросетевой модели. Для разметки использована модель RuRoBERTa, дообученная на текстах корпуса. Работы проводились группой нейроразметки (А. В. Глазкова, Д. А. Морозов, Н. А. Владова, Т. А. Гарипов, С. С. Столяров, И. А. Смаль).

Эксперимент проводился в два этапа. На первом этапе на основе анализа текстов корпуса социальных сетей и изучения исследовательской литературы была составлена номенклатура жанров, наиболее типичных для основного массива корпуса (всего девять жанров). Из основного и регионального газетного корпусов отобраны коллекции текстов, относящихся к этим жанрам (объемом не менее 200 документов каждая), составивших датасет для обучения модели. С помощью обученной модели был размечен основной массив текстов корпуса социальных сетей, а результаты разметки были протестированы экспертами. Эксперты оценивали корректность определения жанра, отмечая случаи ошибочных решений и предлагая альтернативный вариант. Особое внимание было уделено категории «неопределенный жанр», в которую попадали тексты, не относящиеся ни к одному из заданных жанров, либо не имеющие четких жанровых признаков (например, фрагменты диалога). Последующий анализ этих текстов позволил выявить еще целый ряд типичных для социальных сетей жанров, например, *инструкция/рекомендация/совет, цитаты, афоризмы, история, интернет-рейтинг, гороскоп*. Они были включены в итоговый список, используемый для разметки корпуса. Обучающий датасет также был расширен за счет текстов, относящихся к этим дополнительным жанрам.

На втором этапе модель была обучена на расширенном датасете и использована для повторной разметки массива корпуса. Полученные результаты в настоящее время представлены на сайте НКРЯ. Собираются данные об ошибках в определении жанровой принадлежности, которые в дальнейшем будут учтены при коррекции модели. В целом жанровое распределение текстов в текущей версии корпуса социальных сетей выглядит следующим образом. Наиболее велика доля текстов, относящихся к жанрам объявления/анонса (объединены в одну группу) — 20,6 % и информационного сообщения (новости) — 19 %. Большая группа жанров занимает нишу от 2 % до 10 %: инструкция/рекомендация/совет (6,1 %), отзыв/рецензия (5,1 %), рецепт (4 %), цитаты/афоризмы (3 %), поэзия (3 %), история (2,5 %), анекдот (2,3 %). К жанрам с долей менее 2 % в общей структуре относятся оценка (1,3 %), поздравление (1,1 %), интернет-рейтинг (0,4 %), гороскоп (0,4 %), вопрос (0,1 %), подпись к фото (0,06 %). К текстам блогов и форумов из воронежской коллекции, которые представлены большими файлами, не разделенными на отдельные посты, методику автоматического определения жанров применить оказалось невозможно, на их долю приходится, соответственно, 6,4 % и 6,1 % словоупотреблений.

Наконец тексты, жанровую принадлежность которых определить не удалось, отнесены к неопределенной категории и составляют 13 % от общего количества словоупотреблений. В дальнейшем предстоит повторный анализ текстов этой категории с последующим дообучением модели; результатом этого этапа будет уменьшение доли неопределенных в жанровом отношении текстов.

В целом эксперимент по автоматическому определению жанров электронной коммуникации следует признать успешным: накопленный богатый опыт ручной метатекстовой разметки и достаточно большие коллекции текстов разных жанров удалось продуктивно использовать для обучения нейросетевых моделей. На основе моделей в дальнейшем планируется проводить автоматическую разметку текстов по разным метатекстовым и текстовым признакам. В ближайшие планы развития корпуса входит разметка тематики текстов, их тональности, поиск ключевых слов, а также разметка невербальных компонентов (эмодиконов).

3. Панхронический корпус

Панхронический корпус — собрание текстов, охватывающее несколько общепризнанных диахронических периодов языка (в том числе и языка-предка вместе с языком-потомком; разумеется, решение вопроса о том, в каких случаях перед нами «древне-L язык», а в каких «язык M, являющийся предком языка L», зависит от традиции и формализовано быть не может). Такой корпус возможен только для языка (группы языков) с очень длинной более или менее непрерывной письменной историей. Среди таких ресурсов, например, французская база данных Frantext (<https://www.frantext.fr/>) или латинский Corpus Corporum Цюрихского университета [Roelli 2014]. Ведется работа над панхроническим корпусом чешского языка HiCKoK, который объединяет коллекции Чешского национального корпуса и корпусов, собиравшихся другими историками языка, на базе единой разметки в формате Universal Dependencies (<https://korpus.cz/hickok>). Он охватывает «исторический континуум» с XIII по XXI в.

Панхронические корпуса применяются как для исследования многовековых тенденций в развитии языка, проявляющихся непрерывно на протяжении долгого периода (таких как грамматикализация вида или одушевленности, изменение предложного управления и т. п.), так и для феноменов, фиксируемых в письменных источниках со значительными временными интервалами — так называемых «скрытых» (submerged) явлений (ср. [Adams, Vincent (eds.) 2016] для латинского языка). Как известно, значительная часть лексем, фразеологизмов, конструкций эпизодически фиксируется письменными источниками древнерусской эпохи, а затем проявляется лишь в текстах Нового времени, диалектных словарях и т. п. (ср. [Зализняк 2024: 159 и сл.]), что связано с жанровой и стилистической неполнотой дошедшего до нас закрытого корпуса древних текстов.

Вполне естественно, что панхронический корпус строится путем объединения уже существующих. Важнейший стимул при его создании — это именно стремление получить поиск «в одном окне», не повторяя запрос много раз в разных поисковых формах с разными орфографическими принципами, форматами запроса и т. п. Что касается корпусов разных исторических периодов, то над ними обычно работают разные команды специалистов, затем получающие возможность привести свои результаты к «общему знаменателю»: и латинский, и чешский панхронический корпус возникли именно после появления такого единого формата разметки, сделавшего возможным единый интерфейс поиска. Заметим, что команда НКРЯ приняла решение объединить разные корпуса на общей платформе независимо от проектов, посвященных латыни и чешскому (о последнем нам стало известно уже после публикации первого релиза нашего Панхронического корпуса), — оно вполне логично вытекает из общей задачи.

Важным моментом стал количественный и качественный скачок в методиках обработки большого объема данных при помощи механизмов машинного обучения — будь то морфологический анализ или распознавание древних рукописей. Если ранее подготовка такого анализа данных предполагала полностью ручной набор, разметку и нормализацию (особенно это касалось текстов на нестандартизированном языке с неустойчивой орфографией и т. п.), сейчас обучение на «золотом стандарте» размеченных вручную текстов с последующей экспертной пострекоррекцией обеспечивает статистику и быстрый поиск для куда большего количества текстов — и наличие в разметке, рассчитанной на пользователя-специалиста, того или иного числа ошибок искупается самим фактом наличия значительного массива текстов, доступного для поиска, разметка которого постоянно улучшается.

О перспективе Панхронического корпуса на базе НКРЯ Д. В. Сичинава упоминает в статье о развитии Старорусского корпуса [Сичинава 2016: 209–210], в то время еще не имевшего морфологической разметки и представлявшего собой «слабое звено» между полностью размеченными вручную Древнерусским корпусом и Корпусом берестяных грамот (оба представляют тексты до XV в.), с одной стороны, и также морфологически аннотированным (тогда в основном с неснятой омонимией) Основным корпусом (куда входят тексты начиная с 1700 г.), с другой стороны. Тогда же О. Н. Ляшевская возглавила работу по автоматической лемматизации и снятию омонимии не только в Старорусском, но и в Основном корпусе НКРЯ; изначально оба корпуса планировалось разметить именно при помощи совместимого набора признаков, допускающих сквозной поиск. Появилась также идея синтаксической разметки всего временного диапазона на базе Universal Dependencies (тоже независимо от чешского проекта), ее интеграция в Панхронический корпус — дело будущего, но версии Старорусского и Основного корпусов в соответствующем формате уже сформированы. Работы над созданием Панхронического корпуса начались в 2020 г. Первая версия опубликована в ноябре 2022 г.; второй релиз, включивший также новый корпус «Восточнославянская эпиграфика», вышел в конце 2023 г.

Панхронический корпус в составе НКРЯ объединяет четыре разработанных в составе НКРЯ исторических корпуса. Это Древнерусский корпус [Мишина, Пичхадзе 2015], Старорусский корпус [Гаврилова и др. 2016], корпус «Берестяные грамоты» [Сичинава 2022] и появившийся недавно корпус «Восточнославянская эпиграфика» [Sitchinava, Dyshkant 2021; Sitchinava 2023]. Пятой составной частью Панхронического корпуса стал Основной корпус, самые ранние тексты которого относятся к рубежу XVII и XVIII вв.; с этой же временной отметки начинается и Поэтический корпус, который также планируется включить в Панхронический. Нами принято решение не делать объем текстов в пределах каждого временного промежутка одинаковым (иногда такая количественная нормализация осуществляется в исторических корпусах, например, в американском корпусе СОНА [Davies 2010]). Основной корпус включен полностью, несмотря на то что его объем в 30 раз больше, чем пяти исторических вместе взятых. При этом 70 % хронологического диапазона приходится на исторические корпуса.

Перечисленные выше пять корпусов-компонентов продолжают независимое существование в рамках НКРЯ для исследователей, занимающихся именно этими массивами текстов, а не макродиахроническими сюжетами. Привычный пользователям язык запросов, глубина разметки, орфография и т. п. в них сохраняется; впрочем, в части из них теперь доступны для поиска поздние- и раннедревнерусские леммы (см. ниже), интерфейс между которыми разработан специально для Панхронического корпуса. Таким образом, Панхронический корпус и его корпуса-компоненты способствуют взаимному усовершенствованию и обогащению информацией.

Панхронический корпус планируется обновлять раз в год, после пополнения входящих в него корпусов-компонентов. Его разметка также обновляется с учетом встретившихся в новых текстах лексем.

3.1. Представление лемм

Корпуса-компоненты Панхронического корпуса размечались на базе разных стандартов разметки и разного морфологического описания. Стандарт подачи лемм зависел от словаря соответствующей эпохи (Словаря древнерусского языка XI–XIV вв., Словаря русского языка XI–XVII вв., Грамматического словаря русского языка А. А. Зализняка); особую проблему представляла лемматизация имен собственных и образованных от них прилагательных, названий жителей, этнонимов, которые в большинство исторических словарей не включаются. Корпус берестяных грамот был размечен на основании словоуказателя ко второму изданию «Древненовгородского диалекта» А. А. Зализняка [2004], стандарты которого несколько отличаются от стандартов главных древнерусских словарей. Этот словоуказатель ориентирован на позднерусский фонетический состав.

Кроме того, и внутри Древнерусского корпуса разные тексты, восходившие к изначально независимым базам данных, были размечены в соответствии с разными принципами (например, краткие прилагательные могли даваться как отдельные леммы или как словоизменительные формы). Внутри Основного корпуса, который также представлял собой с самого своего появления в 2004 г. объединение текстов, размеченных по двум разным морфологическим стандартам, похожие различия были к началу 2020-х гг. сглажены (усилиями О. Н. Ляшневской и Д. В. Сичинавы), но не в полной мере. Пятый корпус-компонент Панхронического корпуса — «Восточнославянская эпиграфика» — был добавлен позже, и нового разноморфного формата удалось избежать. Он размечался по тем же стандартам, что и Древнерусский корпус.

Ключевая задача объединения этих ресурсов — связать соответствиями (не обязательно одно-однозначными, но позволяющими сквозной поиск интересующих пользователя слов и контекстов) представлений слов на следующих трех этапах:

- Раннедревнерусский (до падения и прояснения редуцированных): *загърдѣтиса*,
- Позднедревнерусский и старорусский (после падения и прояснения редуцированных): *загордѣтиса*,
- Современный русский язык (русский язык Нового времени): *загордеться* (ныне устаревшее и просторечное, в отличие от более стандартного *загордиться*, ср.: *Загорделась, Санечка, загорделась, хоть бы разок приехала* [Чехов, «Иванов»]).

Кроме этих фонетических соответствий, возможны дополнительные, как регулярные (например, отверждение *ц*), так и нерегулярные, включающие в себя и морфологические замены (древнему *ужина* соответствует современное *ужинъ*, а наряду с устаревшим *загордеться* представлено более стандартное и более частотное в корпусе *загордиться*). Следует учитывать также явление расщепления лемм, связанного с фонетическими процессами, влиянием церковнославянской нормы и/или семантической дифференциацией (*сѣборъ* > *сбор*, *сборъ*; *грѣчьскыи* > *греческий*, *грецкий*), а также появления новых омонимов (*мочи* [глагол], *мочь* [существительное] > *мочь*).

Чтобы «связать» три исторических корпуса — Древнерусский, Старорусский и корпус «Берестяных грамот» (а затем и добавившийся позже четвертый, «Древнерусская эпиграфика») — генерировался полный список лемм в первых двух корпусах, а затем строилось их соответствие в представлении другой эпохи. Допустим, в Корпусе берестяных грамот встретила лемма *сгонити*, для нее порождалась раннедревнерусская лемма *съгонити*; в Древнерусском корпусе для формы *звѣнигородѣць* была построена поздняя лемма *звенигородець*. Для этого используется механизм регулярных выражений со сложными условиями и полным последующим ручным контролем.

Что касается аналогичной карты соответствия позднерусских лемм (полученных из всех трех исторических корпусов) и лемм современного русского корпуса, она строилась при помощи алгоритма, разработанного Т. А. Архангельским. Этот алгоритм учитывал как

регулярные правила, так и расстояния символов между леммами-кандидатами. Т. А. Архангельский предложил модификацию классического алгоритма расстояния между строками символов, известного как расстояние Левенштейна, где стоимость любой операции (замена, удаление или вставка символа) зависит от контекста, в котором находится символ в строке. Эти функции определяются с использованием нейронной модели, обученной на 50 тыс. пар слов, предварительно выровненных с учетом расстояния Левенштейна и выборочно проверенных. Получившаяся карта соответствий была усовершенствована Т. А. Архангельским и Д. В. Сичиновой вручную и в дальнейшем корректируется по мере работы с корпусом, для включения исправлений в новые релизы.

На последнем этапе проведения алгоритмического объединения карт А. Н. Дышкант собрал единые списки панхронических лемм, охватывающие все четыре интегрируемых корпуса. В итоговой версии панхронического поиска лемм в любом тексте каждая лемма может быть приведена к нормализованной форме раннедревнерусского периода (*възбранити*), позднедревнерусского / старорусского (*возбранити*) и современного русского (*возбранить*). Поиск осуществляется одновременно для всех вариантов лемм: таким образом, пользователь, вводя один и тот же запрос *возбранить*, сможет найти упоминания слова *возбранить* из XVIII в., *възбранити* из XI в. и *возбранити* из XV в. Эти формы могут быть найдены при выборе любого из фонетически-орфографических обликов. Важно отметить, что пользователь не обязан знать заранее, в какой именно форме он ищет интересующее его слово. В Панхроническом корпусе по мере ввода леммы появляется всплывающая подсказка (suggest; см. также выше о корпусе «Социальные сети»), в которой представлены начинающиеся с тех же букв леммы, и пользователь может выбрать нужный ему вариант.

Начиная с версии 2023 г. в Панхроническом корпусе введено также различие между междостережними омонимами/омографами, которым соответствуют отдельные строки соответствия лемм. Например, древнерусское *и* — местоимение (соответствующее совр. *он*) и союз; древнерусское и старорусское *дати* — глагол (современное *дать*) и союз 'чтобы' (утрачен современным языком), современный омограф *пропасть* — глагол *пропа́сть* (древнее *пропасти*) и существительное *пропа́сть* (не изменившее внешнего облика).

Полученные соответствия не являются однозначными: нашей целью было предоставить возможность найти все варианты данного слова. Например, при запросе *ивангородский* будут найдены словообразовательные варианты *ивангородский*, *иванегородский*, *иванигородский*, *иваногородский*, *иванягородецкий*. Отметим, что имена собственные и дериваты от них весьма избирательно охвачены историческими словарями, так что новая информация для исторической лексикографии очевидна уже на уровне разметки.

3.2. Грамматика

Грамматический набор признаков в Панхроническом корпусе отличается от разметки корпусов-компонентов. Обычно в этом наборе отсутствуют признаки, которые размечены только в части корпусов (например, управление предлогов, вид глагола, счетная форма), либо их интерпретация унифицирована. Например, информация об одушевленности представлена только в форме винительного падежа для слов *о*-склонения, форм множественного числа, местоимений и адъективов. Грамматические признаки, утраченные в современном языке (такие как аорист и имперфект), сохранены в исторических корпусах, поскольку они позволяют формировать информативные запросы, охватывающие большой временной диапазон из нескольких корпусов-компонентов.

Грамматические параметры влияют также и на выделение лемм в разных корпусах. Например, в Древнерусском корпусе компаратив *выше*, краткая форма *высокъ* и полная форма *высокийи* — разные лексемы; в других корпусах подача этой леммы унифицирована (хотя конкретная орфография может различаться).

В текстах, входящих в Панхронический корпус, снята лексическая и грамматическая омонимия. В Старорусском корпусе и большинстве текстов Основного корпуса приписывание лемм и грамматических характеристик сделано автоматически при помощи нейросетевых механизмов, причем в Старорусском корпусе предусмотрена ручная построчная коррекция ошибок нейросетевого разбора.

3.3. Семантика

Для построения запроса по семантической разметке в Панхроническом корпусе используется привязанное к каждому входящему в запрос слову поле «семантика», сопровождаемое формой с отдельными страницами по знаменательным частям речи и местоимениям. Эта форма устроена полностью аналогично форме лексико-семантического поиска в текстовых корпусах НКРЯ, в которых выступает современный русский язык. Соответственно, можно задавать, например, для предметных имен — названия частей тела или этнонимы, для абстрактных имен — названия физических свойств или эмоций, для прилагательных — цветообозначения, для глаголов — глаголы движения или физического воздействия на объект и т. п. Фасетная классификация, используемая в семантической разметке современных текстов НКРЯ, описана, в частности, в работе [Рахилина и др. 2009].

Как указывает Д. В. Сичинава [2024: 347], «ограничения принятого решения очевидны. Разумеется, семантически размечены те и только те древне- и старорусские слова, которые имеют вошедшие в современный семантический словарь НКРЯ когнаты. Поскольку лексическая семантика подвержена историческим изменениям, а ряд слов, включая частотные, вообще утрачен современным языком — или же маргинализован в нем — и в семантическом словаре НКРЯ отсутствует, в семантической разметке исторических текстов заведомо есть неточности и неполнота, на данном этапе сознательно заложенные в проекте, и к ней надо относиться с осторожностью». Например, помету «отрицательная оценка» получают слова *тварь* или *челядь*, совершенно нейтральные в древнерусском или церковнославянском. Случаи полисемии и омонимии (опять же, являющиеся таковыми в современном русском, например, *вода* в значении ‘бессодержательный текст’ во всех контекстах ищется на «отрицательную оценку»), не снятые в семантической разметке «современного» НКРЯ, унаследованы и семантической разметкой Панхронического корпуса.

«Тем не менее полученное при таком подходе высокое покрытие исторических текстов семантической разметкой и диахроническая стабильность семантических классов большинства лексем в значительной мере компенсируют неизбежные недостатки такой разметки и делают оправданным само это предприятие» [Сичинава 2024: 347]. Например, запрос «аорист от глаголов движения» дает на исторических текстах большую и корректную выборку (*поиде, придохъ, понесохъ, повезоша, поѣха, слѣзе* и мн. др.).

3.4. Применение

В Панхроническом корпусе можно строить релевантные для нескольких веков истории русского языка запросы типа «предлог *по* с предложным падежом», «история существительного *забава*», «одушевленность названий животных», «сочетаемость глаголов движения с абстрактным субъектом», «имена собственные на *-славъ*» на всем этом массиве текстов, не вводя каждый раз пять запросов в интерфейс всех корпусов поочередно.

На материале Панхронического корпуса могут быть построены нормализованные частотные графики (количество вхождений на миллион) на всем хронологическом диапазоне запроса. Разумеется, выводы на основании частотного графика (а в будущем — также полученных на корпусе коллокаций и частотных словарей, когда такой функционал будет доступен) должны делаться осторожно и с учетом специфики текстов, включенных в корпус в разные периоды. Указанная проблема относится к азбучным основам количественной лингвистики. О схожей опасности предостерегает А. А. Зализняк [2024: 297], анализируя работы исследователя, который «подсчитывает количество имперфектов по отношению не к общему числу словоформ прошедших времен, а к общему объему текста»: «Единственно, чем хороша такая статистика, — это тем, что так легче считать. В остальном она никуда не годится: какой смысл может иметь подсчет среднего числа имперфектов на страницу текста в условиях, когда в одном тексте описания событий в прошлом могут быть представлены, скажем, вдвое чаще, чем в другом?» В ряде случаев заметные пики частот — статистические выбросы — связаны с конкретными текстами, для которых характерна высокая частота той или иной конструкции, а также с общим объемом текстов, доступных для того или иного интервала.

Решению этих проблем могут поспособствовать такие подходы, как использование в Панхроническом корпусе обобщенных жанровых признаков для отбора подкорпуса, альтернативные алгоритмы подсчета частот и другие. В релизе 2023 г. уже задействован метатекстовый признак «категория текста» (можно отдельно рассматривать литературные, богословские, бытовые тексты и т. п.), а кроме того, построенный средствами визуализации НКРЯ график можно корректировать как по оси абсцисс, так и по оси ординат, выводя «за кадр» отдельные выбросы.

4. Статистика и визуализация

Современная корпусная лингвистика очень во многом опирается на количественные методы анализа данных: сравнения частотностей лексико-грамматических конструкций, анализ коллокационных метрик, сравнение частотных словарей корпусов. В отличие от традиционного конкорданса эти методы предполагают работу с обобщенными данными, а не с изолированными примерами, поэтому результаты анализа представляются в виде таблицы или в виде графиков и диаграмм, визуализирующих дистрибуцию лексико-грамматических единиц или метаданных корпусных текстов. Корпус 2.0 предлагает целый пакет таких возможностей, включающий в себя поиск по коллокациям, подсчет частотностей, представление данных в виде графиков и автоматическое построение диаграмм. Кроме этого, в корпусе появился новый функционал, связанный с исследованием свойств отдельных лексем — «портрет слова». Ниже будет представлен краткий обзор возможностей основных инструментов статистики и визуализации, внедренных в корпусную платформу 2.0.

4.1. Коллокации

Поиск по коллокациям является, пожалуй, наиболее значительным (и долгожданным) нововведением из всего блока преобразований, относящихся к статистике и визуализации. В корпусной лингвистике понятие коллокаций несколько отличается от понятия, принятого в лексикографии: к коллокациям не предъявляется требование некомпозиционности, а коллокат понимается просто как единица, встречающаяся с заданным словом (ключом) чаще, чем случайно. Для того чтобы определить меру такой неслучайности, в НКРЯ используются хорошо известные в корпусной лингвистике метрики: t -score, MI^3 , LogDice

и Loglikelihood². Все они основаны на частотности каждого из элементов коллокации и раз-
мере корпуса, однако дают несколько разное ранжирование. Например, такая классическая
метрика, как t-score, дает ранжирование, близкое простой частоте, а MI³ выводит на более вы-
сокие позиции более редкие коллокаты. Пользователь может выбрать метрику, наиболее под-
ходящую под его конкретные исследовательские задачи, и ранжировать результаты по ней.

Как уже говорилось выше, внедрение поиска по коллокациям потребовало доработки
поискового интерфейса: функция поиска по коллокациям теперь доступна в нем наравне
с поиском точных форм и лексико-грамматическим поиском. Интерфейс дает пользова-
телю широкие возможности, поскольку как ключ, так и коллокаты могут быть заданы
не просто как леммы, но и определены с помощью набора грамматических и/или семан-
тических признаков; кроме того, можно задать ограничения на тип и направление синтак-
сической связи между ключом и коллокатом. Синтаксические параметры поискового за-
проса по коллокациям позволяют искать не просто наиболее частотные ассоциаты слова,
но и наиболее типичные заполнители его активных и пассивных валентностей.

4.2. Частотность

Принципиально новые возможности дает такой новый формат выдачи результатов лек-
сико-грамматического поиска, который называется в интерфейсе НКРЯ «частотность». Благодаря его внедрению исследователь избавлен от необходимости составлять частот-
ный список лексем и их сочетаний, встретившихся в той или иной конструкции, вручную: теперь все данные о частотном распределении результатов поиска выводятся автомати-
чески в одном окне. Например, нас интересует сочетание «предлог с + форма родитель-
ного падежа на -у»: в современных текстах лидируют сочетания *с виду*, *с ходу*, *<сбить>*
с толку, *с голоду*, *с размаху*, *с глазу <на глаз>*, *с краю*. Как мы можем видеть на рис. 1, вы-
дача по частотности показывает нам распределение лексем в рассматриваемой конструк-
ции, а именно долю вхождений каждой лексемы в полученной выдаче.

Слово 1 Лемма	← [1..2] →	Слово 2 Лемма	Вхождения	Доля	Δ Доля	frq	Комментарии
с	1	голова	2857	13,49%	(13,03%, 13,95%)	7,63	Примеры
с	1	вид	2828	13,35%	(12,9%, 13,81%)	7,55	Примеры
с	1	толк	2296	10,84%	(10,49%, 11,26%)	6,13	Примеры
с	1	код	1589	7,5%	(7,15%, 7,86%)	4,24	Примеры
с	1	глаз	1490	7,03%	(6,7%, 7,39%)	3,98	Примеры
с	1	размах	1253	5,91%	(5,6%, 6,24%)	3,35	Примеры
с	1	бок	1010	4,77%	(4,49%, 5,06%)	2,7	Примеры
с	1	край	885	4,18%	(3,92%, 4,46%)	2,36	Примеры
с	1	лоп	586	2,77%	(2,55%, 3%)	1,56	Примеры
с	1	час	513	2,42%	(2,22%, 2,64%)	1,37	Примеры
с	1	разбег	475	2,24%	(2,05%, 2,45%)	1,27	Примеры
с	1	перепуг	462	2,18%	(1,99%, 2,39%)	1,23	Примеры
с	1	мак	310	1,46%	(1,31%, 1,63%)	0,83	Примеры
с	1	бой	276	1,3%	(1,16%, 1,46%)	0,74	Примеры
с	1	жир	253	1,19%	(1,06%, 1,35%)	0,68	Примеры
с	1	испуг	248	1,17%	(1,03%, 1,32%)	0,66	Примеры
с	1	верх	189	0,89%	(0,77%, 1,03%)	0,5	Примеры

Рис. 1. Распределение частотностей существительных,
которые встречаются в конструкции с предлогом с в родительном падеже на -у

² Для подсчета коллокаций были использованные формулы, приведенные в работе [Evert, Krenn 2003].

Потенциал этого статистического инструмента состоит в возможности разных группировок результата поиска: мы можем сгруппировать результаты (т. е. вывести общее количество найденных единиц) не только по более привычным для пользователей корпуса леммам и словоформам, но и по грамматическим признакам, обозначенным набором грамматических тэгов, причем при настройке выдачи возможна и группировка по расстояниям между элементами запроса. Таким образом, частотность как режим отображения результатов поиска обеспечивает пользователя данными не только о лексическом распределении слов запроса, но и о моделях управления, паттернах синтаксической упорядоченности, а также помогает выявлять лексико-грамматические конструкции.

4.3. Графики

Графики, показывающие хронологическое распределение слова по текстам корпуса, существовали и в предыдущей версии НКРЯ. Но в текущей версии корпуса внедрены два новшества, которые делают графики еще более полезным и достоверным инструментом.

Во-первых, появилась возможность сравнить на графике результаты нескольких сложных запросов (до пяти), в том числе и заданных на разных подкорпусах, причем сравнивать можно не только результаты поиска по точным формам, но и результаты лексико-грамматического поиска. Хронологическое распределение результатов нескольких запросов визуализировано на одном графике, где каждому запросу соответствует отдельная линия. Второе нововведение заключается в тепловой шкале, дополняющей собственно график и отражающей сравнительное количество найденных документов по запросу на каждый год. Количество документов визуализируется с помощью цветовой насыщенности. Благодаря тепловой шкале пользователь имеет возможность сразу соотнести частотность словоупотреблений с количеством найденных документов и скорректировать интерпретацию, см. рис. 2.

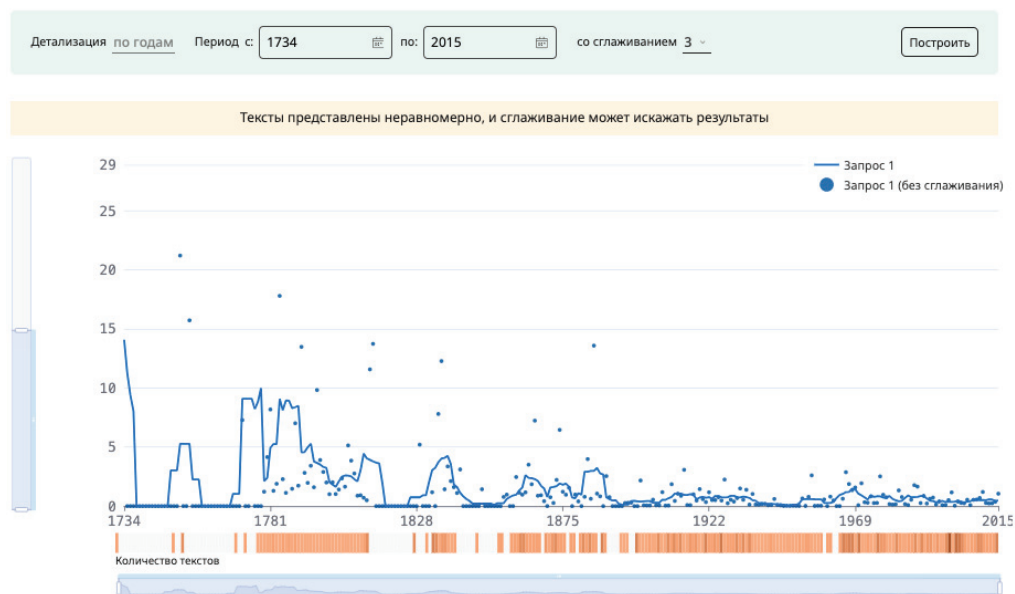


Рис. 2. Хронологический нормализованный график частотности леммы *кофий*. Тепловая шкала отражает количество найденных текстов по запросу за каждый год

Можно заметить, что некоторые колебания графика могут отражать не столько изменение частотности слова, сколько неравномерность хронологического распределения документов, таким образом, повышается надежность и достоверность выводов исследования.

4.4. Сравнение подкорпусов

К новым инструментам, построенным на подсчете распределения частотности, относятся и частотные словари. Первый частотный словарь на материале Основного и устного корпусов НКРЯ был создан еще 15 лет назад [Ляшевская, Шаров 2009], однако за последние годы принципиально изменился состав и объем корпуса, поэтому насущной стала задача обновления его частотного словаря. Реализация решения этой задачи стала скорее необходимым первым шагом для задачи более сложной и амбициозной: сравнения частотных словарей корпуса и подкорпуса. Простая, на первый взгляд, идея потребовала разработки быстросействующих алгоритмов подсчета частотностей «на лету»: поскольку возможности пользователя по созданию собственного подкорпуса практически не ограничены, соответственно, хранить частотные словари всех подкорпусов технически невозможно и требуется считать частотности лемм только после запроса пользователя. Теперь же благодаря новому алгоритму у нас в руках оказался простой инструмент, позволяющий на базовом уровне определить схожесть корпуса и подкорпуса и строить свои выводы с учетом этого знания. Например, одного взгляда на сопоставление частотного словаря Основного корпуса с подкорпусом, состоящим из документов, относящихся к официально-деловой сфере функционирования (см. рис. 3), достаточно, чтобы увидеть существенные отличия на уровне лексики.

Часть речи Глагол Скачать ?

Подкорпус					Корпус		
№	Разница	Лемма	фрт	Вхождения	№	фрт	Вхождения
1	=	быть	8801.01	44818	1	12930.39	4841784
2	=	мочь	2933.61	14939	2	3036.61	1137059
3	↑ 6	иметь	2664.38	13568	9	1052.63	394157
4	↑ 181	установить	1161.93	5917	185	127.97	47920
5	=	бы	1019.37	5191	5	2020.08	756419
6	↑ 37	находиться	1018.58	5187	43	369.96	138533
7	↑ 33	принять	864.82	4404	40	389.85	145980
8	↑ 18	получить	825.75	4205	26	556.75	208476
9	нов	предусмотреть	815.34	4152	>500	—	—
10	↑ 35	являться	753.68	3838	45	364.71	136566
11	↑ 182	указать	720.88	3671	193	122.57	45896
12	↑ 177	производить	699.09	3560	189	126.06	47203
13	↑ 4	дать	615.82	3136	17	731.6	273947
14	↑ 234	определять	605.22	3082	248	98.81	36998
15	нов	осуществлять	594.22	3026	>500	—	—

Рис. 3. Сравнение словарей Основного корпуса и подкорпуса, ограниченного официально-деловой сферой функционирования

Сравниваться могут абсолютные и относительные величины, а также ранги позиций слов в списках, при этом при сравнении подкорпуса и корпуса мы можем сразу увидеть «передвижения» рангов слов подкорпуса относительно списка корпуса и, таким образом, получить информацию о лексической специфике подкорпуса. Объем частотного словаря, доступного пользователю, составляет 500 лексем, при этом пользователь может посмотреть частотный список определенной части речи — существительных, глаголов, прилагательных, наречий. При помощи этого инструмента можно сравнивать с генеральной совокупностью лексикон определенной эпохи, определенного автора и т. п. Сравнение между несколькими корпусами пока не доступно, но со временем может также принести интересные результаты.

Еще одним способом «посмотреть сверху» на состав корпуса (или его подкорпуса) является сравнительный анализ метатрибутов текстов, реализованный в виде интерактивных диаграмм (в том числе диахронических). Эти внешние параметры, их равномерность или неравномерность распределения пользователь всегда должен принимать во внимание, чтобы учитывать их возможное воздействие на результаты исследования. Так, диаграмма распределения текстов по тематике в Основном корпусе (рис. 4) наглядно показывает, что документы, входящие в корпус, очень разнообразны по этому параметру. Важно учитывать, что на диаграммах выводится только 10 первых значений, а остальные объединяются в категорию «Прочее».

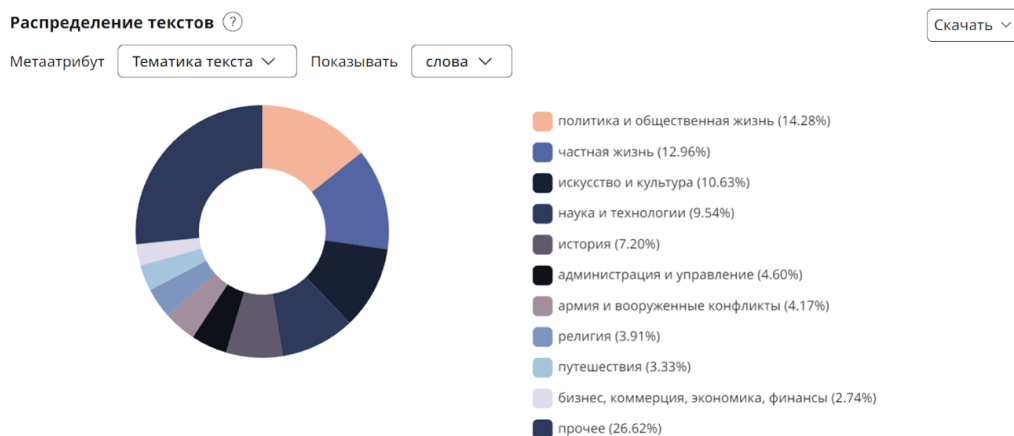


Рис. 4. Распределение текстов по тематике в Основном корпусе

4.5. Портрет слова

Портрет слова представляет собой интерфейс для комплексного анализа свойств отдельной лексемы на материале конкретного корпуса. При разработке концепции портрета слова мы вдохновлялись функционалом сервиса «Word at a glance», представленным в Национальном чешском корпусе (<https://www.korpus.cz/slovo-v-kostce/>), учитывался также опыт Цифрового словаря немецкого языка (<https://www.dwds.de/>).

На странице портрета слова собраны общие параметры слова, соответствующие его грамматической и семантической разметке в корпусе, а также в компактном виде представлена самая разная информация о функционировании слова в заданном корпусе. В своем максимальном разнообразии портрет слова реализован в Основном корпусе.

Кроме пяти случайных примеров из конкорданса, диахронического графика и диаграммы частотности слова по метаатрибутам текстов корпуса, в портрет слова включены и другие блоки, которые в комплексе дают пользователю разнообразную информацию о том, как слово существует и функционирует в языке. Во-первых, в портрет слова входят скетчи — предподсчитанные коллокации, распределенные по типу синтаксической связи. Каждый скетч представляет собой список из 10 самых частотных коллокаций в заданных синтаксических отношениях. Для каждой части речи был определен свой список таких отношений. Например, для существительных скетчи включают в себя прилагательные-определения к заданному существительному, глаголы, для которых это существительное является подлежащим, глаголы, для которых оно же является прямым дополнением или косвенным дополнением, другие существительные, связанные с заданным сочинительной связью. Скетчи позволяют пользователю быстро оценить сочетаемость выбранной лексики, прежде чем приступить к более детальному исследованию (см. рис. 5).

Скетчи ?

душа Существительное

Определения		Сказуемые		Глаголы с прямым дополнением		Глаголы с косвен	
1. мертвый	10,41	1. болеть	7,83	1. чаять	9,2	1. кривить	
2. человеческий	9,81	2. рваться	6,39	2. отвести	9,02	2. любить	
3. живой	8,93	3. жаждать	6,3	3. наполнять	8,89	3. болеть	
4. бессмертный	8,42	4. гореть	5,95	4. раздирать	8,62	4. отдохнуть	
5. чистый	8,05	5. искать	5,92	5. отдать	8,37	5. отдыхать	
6. грешный	7,89	6. просить	5,57	6. отводить	8,36	6. покривить	
7. добрый	7,89	7. стремиться	5,52	7. упокоить	8,04	7. ненавидеть	
8. детский	7,81	8. петь	5,51	8. погубить	7,95	8. чувствовать	
9. чужой	7,43	9. лежать	5,51	9. спасти	7,84	9. стремиться	
10. христианский	7,41	10. радоваться	5,44	10. продать	7,79	10. привязаться	

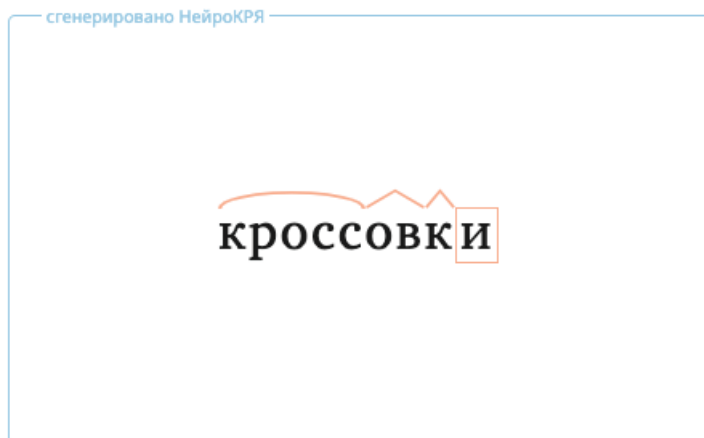
Показать все коллокации

Рис. 5. Скетчи для слова *душа* по основным синтаксическим связям существительных: определения к *душе*, глаголы-сказуемые, глаголы, которые управляют существительным *душа* как прямым дополнением

В портрет слова добавлен еще один совершенно новый блок информации о слове — морфемный разбор словообразовательной структуры слова (подробнее о разметке словообразовательной структуры слова говорилось выше в разделе 1.2, посвященном развитию нейросетевых моделей в проекте). Морфемный разбор строится не только для тех слов, которые присутствуют в базовом словаре морфемного анализа [Ляшевская и др. 2009], но и для тех слов, которые не находятся в этом словаре, в таком случае их морфемная структура строится автоматически с помощью нейросетевой модели НейроКРЯ. Например, слово *кроссовки* отсутствует в словаре морфемного анализа, так что его членение (*кросс-ов-к-и*) предсказано алгоритмом (см. рис. 6). Такие разборы снабжены специальным признаком «сгенерировано НейроКРЯ».

Морфемный разбор β ?

Оценить

Рис. 6. Автоматически предсказанный морфемный разбор слова *кроссовки*

Еще один сервис портрета слова, который обеспечивается моделью НейроКРЯ, — это автоматический подбор однокоренных слов. Пользователю выводятся 10 наиболее частотных однокоренных слов с заданным словом, встречающихся в корпусе (см. рис. 7).

Однокоренные слова β ?

Оценить

Рис. 7. Однокоренные слова к слову *гарантировать*, сгенерированные с помощью алгоритма НейроКРЯ

Дистрибуция форм слова в корпусе отображена с помощью таблицы «формы слова» (см. рис. 8). Важно подчеркнуть, что сравниваются именно орфографические варианты слова, так, например, мы можем увидеть формы слова в старой и новой орфографии, а также получить информацию о том, какие формы слова не встречаются в корпусе вовсе. Формы слова *плен* показывают не только наличие в корпусе этого слова в старой орфографии, но и конкуренцию форм предложного падежа и почти полное отсутствие парадигмы

множественного числа. Насыщенность цветового фона показывает частотность формы, относительная частотность *ipm* формы выдается во всплывающем окне.

Формы слова 

Падеж	единственное	множественное
именительный	плен	—
	плѣнь	
родительный	плена	<u>пленов</u>
	плѣна	
родительный 2	плѣну	—
	плѣну	
дательный	плѣну	—
	плѣну	
винительный	плен	—
	плѣнь	
	плѣн	
творительный	плѣном	—
	плѣномъ	
предложный	плѣне	—
	плѣну	
	плѣнь	
предложный 2	плѣну	—
	плѣну	

Рис. 8. Таблица распределения форм слова *плен* в корпусе

Кроме морфологически и синтаксически связанных слов портрет слова позволяет увидеть так называемые ассоциаты слова, или «похожие слова» — слова, которые наиболее часто встречаются в корпусе в одном и том же контексте. В основе представляемого ассоциативного ряда лежат предподсчитанные модели дистрибутивной семантики — матрицы взаимной встречаемости слов. Характерно, что разные корпуса дают разные списки «похожих слов» к одной и той же лексеме, потому что наиболее частотные контексты одного и того же слова в них будут разными (см. рис. 9).

В настоящий момент портрет слова имеется у всех корпусов НКРЯ, однако его наполненность различна в разных корпусах. Дальнейшее наполнение виджетами портретов слова связано с развитием нейросетевых технологий корпуса, таких как автоматическое снятие омонимии, модели дистрибутивной семантики, технологии морфемного анализа.

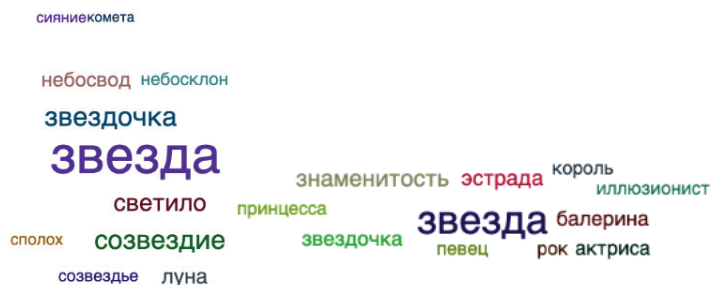


Рис. 9. «Похожие слова» для слова *звезда* в Основном корпусе (слева) и в корпусе Региональных СМИ (справа)

5. Дальнейшие направления развития корпуса

Выше был дан общий обзор основных изменений в составе корпусов и пользовательском функционале НКРЯ, которые стали результатом работы проектной группы в 2020–2023 гг. Следует подчеркнуть, что в основе всех этих изменений лежит фундаментальная и масштабная перестройка архитектуры корпуса. Результатом преобразований стало создание новой корпусной платформы, которая должна обеспечить технологические возможности для устойчивого развития корпусов НКРЯ, их текстового пополнения, подготовки данных для индексации, гибкой настройки интерфейсов.

Создание платформы стало необходимым условием развития технологий нейросетевого моделирования для разметки данных корпуса и создания новых аналитических инструментов: ведь сложность автоматических алгоритмов состоит не только в их экспериментальной настройке и оптимизации на не самых обычных (а часто весьма экзотических) данных, но и во внедрении в собственно технологическую базу корпуса. Очевидно, что решение проблемы интеграции экспериментальных автоматических методов открывает огромный горизонт для дальнейшего развития — экстенсивного, связанного с «выравниванием» функциональных возможностей разных корпусов, с автоматическим снятием омонимии, разметкой метаданных, разработкой статистических аналитических экспериментов, и интенсивного, углубленного, связанного с созданием совершенно новых типов разметок и аналитических инструментов.

Наконец, еще одним направлением развития, обеспеченным технологическими преимуществами новой платформы, является движение в сторону персонализации пользовательской активности. Уже сейчас регистрация на сайте корпуса дает пользователю расширение доступного функционала, такого, как, например, возможность сравнения хронологической дистрибуции нескольких запросов. В дальнейшем пользователи смогут иметь возможность выстроить в своем личном кабинете свой собственный персонифицированный профиль НКРЯ, отражающий личные настройки, историю запросов, сохраненные выдачи и отбор подкорпусов и т. д. Таким образом, перспективное развитие Национального корпуса русского языка 2.0 включает не только совершенствование инструментов и ресурсов для изучения русского языка в его историческом и современном многообразии, но и развитие пользовательской инфраструктуры лингвистических — и шире — общегуманитарных исследований.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Бергельсон 2002 — Бергельсон М. Б. Языковые аспекты виртуальной коммуникации (Языковое поведение в сети Интернет). *Вестник МГУ. Сер. 19. Лингвистика и межкультурная коммуникация*, 2002, 1: 55–67. [Bergel'son M. B. Linguistic aspects of virtual communication (Linguistic behavior on the Internet). *Vestnik MGU. Ser. 19. Lingvistika i mezhkul'turnaya kommunikatsiya*, 2002, 1: 55–67.]
- Гаврилова и др. 2016 — Гаврилова Т. С., Шалганова Т. А., Ляшевская О. Н. К задаче автоматической лексико-грамматической разметки старорусского корпуса XV–XVII вв. *Вестник ПСТГУ. Серия III: Филология*, 2016, 2(47): 7–25. [Gavrilova T. S., Shalganova T. A., Lyashevskaya O. N. On the problem of automatic lexical and grammatical markup in the Old Russian corpus of the XV–XVII centuries. *Vestnik PSTGU. Seriya III: Filologiya*, 2016, 2(47): 7–25.]
- Гладилин, Козеренко 2022 — Гладилин С., Козеренко А. Новый интерфейс поиска для НКРЯ: системное описание и реализация. *Информационные технологии и системы 2022 (ИТС 2022): материалы конференций*. Шилин Л. Ю. и др. (ред.). Минск: БГУИР, 2022, 113–121. [Gladilin S., Kozerenko A. The new search interface for the RNC: System description and implementation. *Informatsionnye tekhnologii i sistemy 2022 (ITS 2022)*. Conf. proc. Shilin L. Yu. et al. (eds.). Minsk: Belarusian State Univ. of Informatics and Radioelectronics, 2022, 113–121.]
- Горошко 2007 — Горошко Е. И. Теоретический анализ Интернет-жанров: к описанию проблемной области. *Жанры речи: Сб. науч. ст. Вып. 5. Жанр и культура*. Дементьев В. В. (ред.). Саратов: Наука, 2007, 119–127. [Goroshko E. I. Theoretical analysis of Internet genres: Towards a description of the problem area. *Zhanry rechi*. Coll. of papers. No. 5. *Zhanr i kul'tura*. Dement'ev V. V. (ed.). Saratov: Nauka, 2007, 119–127.]
- Горошко, Землякова 2017 — Горошко Е. И., Землякова Е. А. Полиформатный мессенджер как жанр 2.0 (на примере мессенджера мгновенных сообщений Telegram). *Жанры речи*, 2017, 1(15): 92–100. [Goroshko E. I., Zemlyakova E. A. A multi-format messenger as a genre 2.0 (on the example of the Telegram instant messenger). *Zhanry rechi*, 2017, 1(15): 92–100.]
- Дементьев 2016 — Дементьев В. В., Степанова Н. Б. Корпусная генеристика: проблема ключевых фраз. *Жанры речи*, 2016, 1(13): 24–41. [Dement'ev V. V., Stepanova N. B. Corpus generistics: The problem of key phrases. *Zhanry rechi*, 2016, 1(13): 24–41.]
- Донина и др. 2024 (в печати) — Донина О. В., Фурсина Д. А., Горбунов Н. С. Создание регионального подкорпуса: от идеи до воплощения. *Труды международной конференции «Корпусная лингвистика-2023»* (в печати). [Donina O. V., Fursina D. A., Gorbunov N. S. Creation of a regional subcorpus: From idea to implementation. *Trudy mezhdunarodnoi konferentsii «Korpusnaya lingvistika-2023»* (in print).]
- Егорова 2021 — Егорова В. И. Социальные сети и их речевые жанры. *Russian Linguistic Bulletin*, 2021, 3(27): 123–128. [Egorova V. I. Social networks and their speech genres. *Russian Linguistic Bulletin*, 2021, 3(27): 123–128.]
- Зализняк 2004 — Зализняк А. А. *Древненовгородский диалект*. М.: Языки славянской культуры, 2004. [Zaliznyak A. A. *Drevnenovgorodskii dialekt* [Old Novgorod dialect]. Moscow: Yazyki slavyanskoi kul'tury, 2004.]
- Зализняк 2024 — Зализняк А. А. *Слово о полку Игореве: взгляд лингвиста*. 4-е изд. М.: Альпина, 2024. [Zaliznyak A. A. *Slovo o polku Igoreve: vzglyad lingvista* [The Tale of Igor's Campaign: A linguist's view]. Moscow: Al'pina, 2024.]
- Иванов 2000 — Иванов Л. Ю. Язык интернета: заметки лингвиста. *Словарь и культура русской речи*. М.: Азбуковник, 2000, 131–147. [Ivanov L. Yu. The language of the Internet: Notes of a linguist. *Slovar' i kul'tura russkoi rechi*. Moscow: Azbukovnik, 2000, 131–147.] <http://faq-www.ru/lingv.htm>.
- Какорина 2008 — Какорина Е. В. СМИ и интернет-коммуникация (интернет-форум как новый коммуникативно-речевой жанр). *Современный русский язык: активные процессы на рубеже XX–XXI веков*. Крысин Л. П. (отв. ред.). М.: Языки славянских культур, 2008, 549–578. [Kakorina E. V. Mass media and Internet communication (Internet forum as a new communicative and speech genre). *Sovremennyyi russkii yazyk: aktivnye protsessy na rubezhe XX–XXI vekov*. Krysin L. P. (ed.). Moscow: Yazyki slavyanskikh kul'tur, 2008, 549–578.]
- Капанадзе 2005 — Капанадзе Л. А. На границе письменного и устного текста: структура и тенденции развития электронных жанров. *Голоса и смыслы. Избранные работы по русскому языку*. М.: ИРЯ РАН, 2005, 305–320. [Kapanadze L. A. On the border of written and oral text: The structure and

- trends in the development of electronic genres. *Golos i smysly. Izbrannyye raboty po russkomu yazyku*. Moscow: Vinogradov Russian Language Institute, 2005, 305–320.]
- Карасик 2019 — Карасик В. И. Жанры сетевого дискурса. *Жанры речи*, 2019, 1(21): 49–55. [Karasik V. I. Genres of online discourse. *Zhanry rechi*, 2019, 1(21): 49–55.]
- Кириллов 2017 — Кириллов А. Г. Трансформация жанра блога в программах обмена мгновенными сообщениями. *Жанры речи*, 2017, 2(16): 260–267. [Kirillov A. G. The transformation of the blog genre in instant messaging programs. *Zhanry rechi*, 2017, 2(16): 260–267.]
- Кузнецова, Ефремова 1986 — Кузнецова А. И., Ефремова Т. Ф. *Словарь морфем русского языка*. М.: Русский язык, 1986. [Kuznetsova A. I., Efremova T. F. *Slovar' morfem russkogo yazyka* [Dictionary of morphemes of the Russian language]. Moscow: Russkii yazyk, 1986.]
- Кузьмина 2003 — Кузьмина М. В. Компьютерный вид общения «чат» как жанр естественной письменной речи: основные характеристики. *Естественная письменная русская речь: исследовательский и образовательный аспекты: материалы конф. Ч. II: Теория и практика современной письменной речи*. Голев Н. Д. (ред.). Барнаул: Изд-во Алтайского ун-та, 2003, 86–91. [Kuz'mina M. V. Computer speech style “chat” as a genre of natural written speech: Basic features. *Estestvennaya pis'mennaya russkaya rech': issledovatel'skii i obrazovatel'nyi aspekty*. Conf. proc. P. II: *Teoriya i praktika sovremennoi pis'mennoi rechi*. Golev N. D. (ed.). Barnaul: Altai State Univ. Press, 2003, 86–91.]
- Литвиненко 2016 — Литвиненко Ж. М. Современная русистика о жанрах интернет-коммуникации: форум, блог, чат. *Вестник ТГПУ*, 2016, 3(168): 48–52. [Litvinenko Zh. M. Modern Russian studies on the genres of Internet communication: forum, blog, chat. *TSPU Bulletin*, 2016, 3 (168): 48–52.]
- Ляшевская и др. 2009 — Ляшевская О., Гришина Е., Тагабилева М., Иткин И. О задачах и методах словообразовательной разметки в корпусе текста. *Полярный вестник*, 2009, 12: 5–25. [Lyashevskaya O., Grishina E., Tagabileva M., Itkin I. On the tasks and methods of word-formation markup in a text corpus. *Polyarnyi vestnik*, 2009, 12: 5–25.]
- Ляшевская, Шаров 2009 — Ляшевская О. Н., Шаров С. А. *Частотный словарь современного русского языка (на материалах Национального корпуса русского языка)*. М.: Азбуковник, 2009. [Lyashevskaya O. N., Sharov S. A. *Chastotnyi slovar' sovremennogo russkogo yazyka (na materialakh Nacional'nogo korpusa russkogo yazyka)* [Frequency dictionary of contemporary Russian based on the Russian National Corpus data]. Moscow: Azbukovnik, 2009.]
- Мишина, Пичхадзе 2015 — Мишина Е. А., Пичхадзе А. А. Древнерусский подкорпус Национального корпуса русского языка. *Труды Института русского языка им. В. В. Виноградова*, 2015, 6: 99–115. [Mishina E. A., Pichkhadze A. A. The Old Russian subcorpus of the Russian National Corpus. *Proceedings of the V. V. Vinogradov Russian Language Institute*, 2015, 6: 99–115.]
- Рахилина и др. 2009 — Рахилина Е. В., Кустова Г. И., Ляшевская О. Н., Резникова Т. И., Шеманаева О. Ю. Задачи и принципы семантической разметки лексики в НКРЯ. *Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы*. Плунгян В. А. (отв. ред.). СПб.: Нестор-История, 2009, 215–239. [Rakhilina E. V., Kustova G. I., Lyashevskaya O. N., Reznikova T. I., Shemanaeva O. Yu. Tasks and principles of semantic markup of lexicon in the RNC. *Natsional'nyi korpus russkogo yazyka: 2006–2008. Novye rezul'taty i perspektivy*. Plungian V. A. (ed.). St. Petersburg: Nestor-Istoriya, 2009, 215–239.]
- Сичинава 2005 — Сичинава Д. В. Национальный корпус русского языка: очерк предистории. *Национальный корпус русского языка: 2003–2005*. Плунгян В. А. (отв. ред.). М.: Индрик, 2005, 21–30. [Sitchinava D. V. Russian National Corpus: An outline of the prehistory. *Natsional'nyi korpus russkogo yazyka: 2003–2005*. Plungian V. A. (ed.). Moscow: Indrik, 2005, 21–30.]
- Сичинава 2016 — Сичинава Д. В. Старорусские/среднерусские тексты в НКРЯ: от экстенсивной коллекции к корпусу. *Rašytinis palikimas ir skaitmeninė technologijos: VI tarptautinė mokslinė konferencija*, Vilnius, 2016 m. rugpjūčio 22–28 d. Vilnius: Lietuvos mokslo taryba, 2016, 208–210. [Sitchinava D. V. Old/Middle Russian texts in the RNC: from an extensive collection to a corpus. *Rašytinis palikimas ir skaitmeninė technologijos: VI tarptautinė mokslinė konferencija*, Vilnius, 2016 m. rugpjūčio 22–28 d. Vilnius: Lietuvos mokslo taryba, 2016, 208–210.]
- Сичинава 2022 — Сичинава Д. В. Корпус берестяных грамот как параллельный. *Труды Института русского языка им. В. В. Виноградова*, 2022, 2: 92–106. [Sitchinava D. V. The corpus of birch bark letters as a parallel corpus. *Proceedings of the V. V. Vinogradov Russian Language Institute*, 2022, 2: 92–106.]
- Сичинава 2024 — Сичинава Д. В. Панхронический корпус: интеграция исторических и современных корпусных ресурсов. *Труды Института русского языка им. В. В. Виноградова*, 2: 336–353. [Sitchinava D. V. A panchronic corpus: Integration of historical and contemporary corpus resources. *Proceedings of the V. V. Vinogradov Russian Language Institute*, 2: 336–353.]

- Тихонов 2002 — Тихонов А. Н. *Морфемно-орфографический словарь: около 100 000 слов*. М.: АСТ, 2002. [Tikhonov A. N. *Morfemno-orfograficheskii slovar': okolo 100 000 slov* [Morphemic and spelling dictionary: about 100,000 words]. Moscow: AST, 2002.]
- Трофимова 2004 — Трофимова Г. Н. *Функционирование русского языка в Интернете: концептуально-сущностные доминанты*. Автореф. дис. ... докт. филол. наук. М.: РУДН, 2004. [Trofimova G. N. *Funktsionirovanie russkogo yazyka v Internetе: kontseptual'no-sushchnostnye dominanty* [The functioning of the Russian language on the Internet: conceptual and essential dominants]. Abstract of cand. diss. Moscow: RUDN Univ., 2004.]
- Шилихина 2018 — Шилихина К. М. Лексические маркеры жанров интернет-коммуникации. *Жанры речи*, 2018, 3(19): 218–225. [Shilikhina K. M. Lexical markers of Internet communication genres. *Zhanyr rechi*, 2018, 3(19): 218–225.]
- Шмелева 2012 — Шмелева Т. В. Жанр в современной медиасфере. *Жанры речи: сб. науч. ст.* Вып. 8. *Жанр и творчество*. Дементьев В. В. (ред.). Саратов; М.: Лабиринт, 2012, 26–37. [Shmeleva T. V. Genre in the modern media sphere. *Zhanyr rechi*. Coll. of papers. No. 8. *Zhanyr i tvorchestvo*. Dement'ev V. V. (ed.). Saratov; Moscow: Labirint, 2012, 26–37.]
- Щипицина 2009 — Щипицина Л. Ю. *Жанры компьютерно-опосредованной коммуникации*. Архангельск: Поморский ун-т, 2009. [Shchipitsina L. Yu. *Zhanyr komp'yuterno-oposredovannoi kommunikatsii* [Genres of computer-mediated communication]. Arkhangelsk: Pomor State Univ., 2009.]
- Adams, Vincent (eds.) 2016 — Adams J. N., Vincent N. (eds.). *Early and Late Latin continuity or change?* Cambridge: Cambridge Univ. Press, 2016.
- Davies 2010 — Davies M. *The Corpus of Historical American English (COHA)*. Electronic resource, 2010. <https://www.english-corpora.org/coha/>.
- Evert, Krenn 2003 — Evert S., Krenn B. *Computational approaches to collocations*. Introductory course at the European Summer School on Logic, Language, and Information (ESSLLI 2003), Vienna. 2003. www.collocations.de.
- Lyashevskaya et al. 2020 — Lyashevskaya O. N., Shavrina T. O., Trofimov I. V., Vlasova N. A. GRAMEVAL 2020 shared task: Russian full morphology and universal dependencies parsing. *Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной международной конференции «Диалог»*, 2020, 19: 553–569. [Lyashevskaya O. N., Shavrina T. O., Trofimov I. V., Vlasova N. A. GRAMEVAL 2020 shared task: Russian full morphology and universal dependencies parsing. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conf. "Dialogue"*, 2020, 19: 553–569.]
- Lyashevskaya et al. 2023 — Lyashevskaya O., Afanasev I., Rebrikov S., Shishkina Y., Suleymanova E., Trofimov I., Vlasova N. Disambiguation in context in the Russian National Corpus: 20 years later. *Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной международной конференции «Диалог»*, 2023, 2: 307–318. [Lyashevskaya O., Afanasev I., Rebrikov S., Shishkina Y., Suleymanova E., Trofimov I., Vlasova N. Disambiguation in context in the Russian National Corpus: 20 years later. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conf. "Dialogue"*, 2023, 22: 307–318.]
- Morozov et al. 2022 — Morozov D. A., Glazkova A. V., Iomdin B. L. Text complexity and linguistic features: their correlation in English and Russian. *Russian Journal of Linguistics*, 2022, 2(26): 425–447.
- Roelli 2014 — Roelli Ph. The Corpus Corporum, a new open Latin text repository and tool. *Archivum Latinitatis Medii Aevi: Bulletin Du Cange*, 2014, 72: 289–304.
- Sitchinava, Dyshkant 2021 — Sitchinava D., Dyshkant A. Integration of the Old East Slavic epigraphical databases, corpora and indices. *Scripta & e-Scripta: The Journal of Interdisciplinary Medieval Studies*, 2021, 21: 93–106.
- Sitchinava 2023 — Sitchinava D. Multiple interpretation and fragmented texts within a historical corpus: the case of Old East Slavic vernacular writing. *Jazykovedný časopis*, 2023, 74(1): 266–274.
- Sorokin, Kravtsova 2018 — Sorokin A., Kravtsova A. Deep convolutional networks for supervised morpheme segmentation of Russian language. *Artificial Intelligence and Natural Language. AINL 2018. Communications in Computer and Information Science*. Ustalov D., Filchenkov A., Pivovarov L., Žižka J. (eds.). Springer: Cham, 2018, 3–10.
- Straka et al. 2016 — Straka M., Hajič J., Straková J. UDPipe: trainable pipeline for processing CoNLL-U files performing tokenization, morphological analysis, pos tagging and parsing. *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16)*. 4290–4297.

Asymmetry in Russian metalinguistic comparatives: Corpus and experimental evidence

© 2024

Natalia A. Zevakhina

HSE University, Moscow, Russia; nzevakhina@hse.ru

Alina A. Shchipkova

HSE University, Moscow, Russia; schipkovalina@gmail.com

Abstract: The paper reports a corpus and experimental study of three types of Russian asymmetric metalinguistic comparatives (*lučše*, *skoree*, and *bol'she* meta-comparatives), two parts of which belong to different syntactic categories and/or have different morphological forms. We discuss their morphological and syntactic properties and uncover diversity among their types and various grammatical patterns that they allow. From a morphological point of view, infinitives in meta-comparatives tend to be imperfective, and finite verbs in the subjunctive and indicative mood are generally perfective. From a syntactic point of view, asymmetric meta-comparatives are either bi-clausal, bi-phrasal, or mixed (clausal-phrasal). The bi-phrasal variety shows functional homogeneity. *Lučše* meta-comparatives are more frequent in the corpus and are evaluated as more grammatical than *skoree* meta-comparatives. Furthermore, the asymmetric pattern ‘noun phrase + infinitive clause’ is evaluated as more grammatical than the pattern ‘finite clause + infinitive clause’ for *lučše* meta-comparatives, while there is no such contrast between these patterns for *skoree* meta-comparatives. Our study sheds light on the grammatical portrait of Russian meta-comparatives as well as on the use and evaluation of asymmetric meta-comparatives in a natural language in general and thus contributes to the discussion of grammatical asymmetry in parallel syntactic structures, such as comparative and coordinated structures, among others.

Keywords: comparative, comparison, corpus linguistics, Russian, syntax

Acknowledgements: We express our sincere gratitude to the three anonymous reviewers of the paper and to the audience of the conference “Grammatical processes and systems in synchrony and diachrony” (Moscow, June 13–15, 2022) for their valuable and thought-provoking comments. This study is an output of the research project No. 23-18-00695 supported by RSF and entitled “Logical and cognitive approach to reasoning: Modeling the interplay between the normative and the descriptive”.

For citation: Zevakhina N. A., Shchipkova A. A. Asymmetry in Russian metalinguistic comparatives: Corpus and experimental evidence. *Voprosy Jazykoznanija*, 2024, 2: 35–51.

DOI: 10.31857/0373-658X.2024.2.35-51

Асимметрия в русских металингвистических компаративах: корпусное и экспериментальное исследование

Наталья Александровна Зевахина

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; nzevakhina@hse.ru

Алина Алексеевна Щипкова

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; schipkovalina@gmail.com

Аннотация: В статье излагаются результаты корпусного и экспериментального исследования трех видов русских асимметричных металингвистических компаративов (далее — метакомпаративов; метакомпаративы с *лучше*, *скорее* и *больше*), две части которых представляют собой различные

синтаксические группы и/или имеют различные морфологические формы. В работе обсуждаются их морфологические и синтаксические особенности. С морфологической точки зрения инфинитивы в металингвистических компаративах обычно несовершенного вида, в то время как финитные глагольные формы изъявительного и сослагательного наклонений обычно совершенного вида. С синтаксической точки зрения асимметричные металингвистические компаративы могут включать две клаузы, две группы или клаузу и группу, т. е. быть смешанными. Две группы в составе метакомпаративов функционально однородны. *Лучше*-метакомпаративы более частотны в корпусе и оцениваются как более грамматичные, чем *скорее*-метакомпаративы. Кроме того, асимметричная конструкция 'именная группа + инфинитивная клауза' оценивается как более грамматичная, чем асимметричная конструкция 'финитная клауза + инфинитивная клауза' в случае *лучше*-метакомпаративов, при этом такого контраста не наблюдается между аналогичными конструкциями в случае *скорее*-метакомпаративов. В статье выявляются грамматические черты русских асимметричных металингвистических компаративов, а также частотность их употребления и оценка носителей в отношении их грамматичности. Результаты статьи могут быть интересны с точки зрения грамматической асимметрии в параллельных синтаксических структурах, к которым, в частности, относятся сравнительные и сочинительные предложения.

Ключевые слова: компаратив, корпусная лингвистика, русский язык, синтаксис, сравнение

Благодарности: Мы выражаем искреннюю благодарность трем анонимным рецензентам и аудитории конференции «Грамматические процессы и системы в синхронии и диахронии» (Москва, 13–15 июня, 2022) за высказанные ценные замечания. Исследование выполнено за счет гранта Российского научного фонда № 23-18-00695, <https://rscf.ru/project/23-18-00695/>.

Для цитирования: Zevakhina N. A., Shchipkova A. A. Asymmetry in Russian metalinguistic comparatives: Corpus and experimental evidence. *Voprosy Jazykoznanija*, 2024, 2: 35–51.

DOI: 10.31857/0373-658X.2024.2.35-51

1. Introduction

Metalinguistic comparatives (or merely meta-comparatives, glossed as MC henceforth) involve a comparison between propositions according to the speaker's judgement of their prototypicality, probability, or preferability. To illustrate, in (1), the speaker believes that the sentence *His speech was eloquent* is a more appropriate/precise description of a given situation than the alternative *His speech was convincing*.

(1) *His speech was **more** [eloquent] **than** [convincing].* [Jespersen 1924: 252]

In this respect, meta-comparatives are distinguished from standard comparatives, which involve semantic comparison of degrees of some properties. Consider (2), in which it is stated that the degree to which Mary's speech was eloquent is greater than the degree to which Peter's speech was eloquent. Importantly, it is not dependent on the speaker's preference for a particular proposition or description of a given state of affairs.

(2) *Mary's speech was **more** eloquent **than** Peter's.*

Meta-comparatives have much in common with meta-linguistic negation [Horn 1989], which also reflects the speaker's preference for a particular utterance among alternatives. For example, (3) does not negate the fact that the water in the lake is cool. Rather, it conveys the idea that *cold* is a better word than *cool* to describe the situation.

(3) *The water in the lake is not cool, it is cold.*

A large body of literature has been accumulated, which provided cross-linguistic evidence for meta-comparatives as a separate grammar phenomenon, see [McCawley 1991; Morzycki 2011; Sassoon 2015] for English; [Sawada 2007] for Japanese; [Giannakidou, Stavrou 2009] for Greek; [Giannakidou, Yoon 2011] for Korean; [Herburger, Rubinstein 2014] for German; [Goncharov 2014; Zevakhina et al. 2017; Goncharov, Irímia 2018] for Russian; [Goncharov, Irímia 2018]

for Romanian, French; [Goncharov, Irimia 2020] for Bulgarian, among others. Since the study reported in this paper examines Russian meta-comparatives, in what follows, we primarily focus on Russian data and take a broader perspective for illustrative and argumentative purposes.

The Russian meta-comparatives are grammaticalized from the following three comparative forms: *skoree* ‘quicker, sooner’ (4), *lučše* ‘better’ (5), and *bol’she* ‘more’ (6). They originated from the positive forms of the adverbs *skoro* ‘quickly, soon’, *xorošo* ‘well’, and *mnogo* ‘a lot’.

- (4) *ix granicy sleduet skoree* [_{VP} *čuvstvovat*’],
 their borders.ACC should MC feel.INF
čem [_{VP} *opredeljat*’ *rassudkom*].
 than determine.INF mind.INST
 ‘(...) the borders [of theatrical plays] should be felt rather than determined by reason.’
 [M. Chekhov. *O texnike aktëra* (1945)]
- (5) *Lučše* [_{VP} *umeret*’ *stoja*], *čem* [_{VP} *žit*’ *na kolenjax*].
 MC die.INF stand.CONV than live.INF on knees.LOC
 ‘It is better to die standing than to live on your knees.’ [V. Shalamov. *Kolymskie rasskazy* (1954–1961)]
- (6) *Èto bol’she* [_{NP} *isključenie*], *čem* [_{NP} *pravilo*].
 this MC exception.NOM than rule.NOM
 ‘This is more an exception than a rule.’ [B. Bujanov. In ‘Volga’, 2013]

Adverbs of speed, quality, and quantity are cross-linguistically common sources of meta-comparatives. Adverbs of speed used in this function include French *plutôt*, Romanian *mai degrabă* [Goncharov 2014]; Greek *malon* [Giannakidou, Stavrou 2009; Giannakidou, Yoon 2011]; German *ehrer* [von Fintel, Kratzer 2014]; English *rather*, which goes back to *hræth* ‘quick’ and the comparative form *hrathor* ‘earlier, more quickly’ [Stevenson, Waite (eds.) 2011]. Greek additionally has meta-comparatives derived from adverbs of quality (*kalitera*) and quantity (*perissotero* and *pjo poly*), cf. also English *more* [Giannakidou, Stavrou 2009; Giannakidou, Yoon 2011].

Several semantic accounts were proposed for various types of meta-comparatives. Morzycki [2011] and Sassoon [2015] give two different accounts: ‘more’ meta-comparatives convey the idea either that, according to the speaker, one sentence is more precise than another to describe a given state of affairs or that, according to the speaker, one proposition is more prototypical than another to describe some property of an entity. For some other meta-comparatives (e.g., *lučše*), [Giannakidou, Stavrou 2009; Giannakidou, Yoon 2011] argue that the speaker believes that one proposition is more preferable than another one. As for the *skoree* meta-comparative, Inkova-Manzotti [2001] and Goncharov [2014] independently suggest that, according to the speaker, one proposition is more probable than another one (cf. also [Beaulieu-Masson, Inkova-Manzotti 2003]).

The syntactic and morphological properties of the three Russian meta-comparatives were reported in [Zevakhina et al. 2017]. In what follows, we concentrate only on such a syntactic property that is relevant to the present study, namely, on the cross-categorial distribution of meta-comparatives. According to [Zevakhina et al. 2017], *skoree* and *bol’she* meta-comparatives are combined with adjectival phrases, adverbial phrases, noun phrases in the Nominative case, numeral phrases, prepositional phrases, verb phrases, predicative clauses, and full clauses. *Lučše* meta-comparatives are used either with noun phrases in the Nominative case or verbal forms in the future indicative, subjunctive or imperative (2nd-person imperative and *pust’*-imperative) mood. All these predicate forms are possible both with verb phrases and full clauses in *lučše* meta-comparatives. As seen from this overview, *lučše* meta-comparatives reveal more grammatical restrictions than *skoree* and *bol’she* meta-comparatives.

Generally, all three meta-comparatives require symmetry between syntactic categories and morphological forms in their two parts, so that, for example, if the first part of a meta-comparative comprises a noun phrase in the Nominative case, then its second part also consists of a noun

phrase in the Nominative case, cf. (6). It cannot be a verb phrase or a prepositional phrase. Other types of phrases are also typically symmetric, see (4) and (5) for symmetric clauses with infinitives.

Morzycki [2011] demonstrated that English meta-comparatives do not always require symmetric syntactic and morphological properties of their parts. Consider (7). Its first part is an adjectival phrase, while its second part is a prepositional phrase. This suggests that syntactic asymmetry holds for meta-comparatives.

(7) *George is **more** [AP afraid of Dick] **than** [PP in love with him].* [Morzycki 2011: 41]

Coordinated sentences also provide evidence for syntactic asymmetry. Culicover and Jackendoff [1997] argue that coordination makes possible unparallelled, or asymmetric, syntactic structures. To illustrate, in (8), the first part of a coordinated sentence includes a noun phrase, while its second part is a full clause.

(8) [_{NP} *One more can of beer*] **and** [_{CP} *I'm leaving*]. [Culicover, Jackendoff 1997: 196]

What about Russian? Sannikov [2008] shows that Russian coordinated structures are asymmetric. To illustrate, in (9), a noun phrase and an adverbial phrase are coordinated, while in (10), an adjectival phrase and a noun phrase are coordinated.

(9) *Mne* [_{NP} *nikto*] **i** [_{AdvP} *nikogda*] *ne pomožet.*
I.DAT nobody.NOM and never not help.FUT.3SG
'Nobody will ever help me.' [Sannikov 2008: 402]

(10) *Knigi* [_{AdjP} *moi*] **i** [_{NP} *moego otca*].
books.NOM my and my father.GEN
'The books are mine and my father's.' [Ibid.: 401]

Sannikov points out that standard comparatives are symmetric in that the syntactic position, morphological coding, and lexical-semantic class of both their parts are identical. To illustrate, in (11), the subjects of two comparative clauses *Petja* and *Vasja* are marked with the Nominative case and belong to the lexical-semantic class of people. The verb in the second clause is elided.

(11) *Petja prygaet **vyše**, **čem** Vasja prygaet.*
Petja.NOM jump.PRS.3SG high.COMP than Vasja.NOM jump.PRS.3SG
'Petja jumps higher than Vasja [does].'

Besides this prototypical version of a standard comparative, Sannikov argues that Russian comparatives represent two more variants, which are viewed as deviations from the prototypical variant. Before moving on, let us introduce the notion of 'syntactic function' borrowed from [Sannikov 2008]. This notion incorporates, on the one hand, a syntactic status of a phrase in terms of the dichotomy between arguments vs. adjuncts, and, on the other hand, the semantic role of a phrase.

One deviation lacks identical morphological coding and lexical-semantic class but features an identical 'syntactic function' of both parts of a standard comparative. Consider (12). It compares a noun phrase in the Dative case and a prepositional phrase in the Accusative case. Their morphological coding and lexical-semantic classes are different. However, both are arguments of the verb *pisat'* 'to write' and have an identical semantic role, that is, addressee/direction.¹ Sannikov refers to this deviation as 'functional homogeneity'.

(12) *Ja **časće** *pišu* [_{NP} *tovarišču*], **čem** [_{PP} *v institut*].*
I often.COMP write.PRS.1SG friend.DAT than to institute.ACC
'I write more often to a friend than to the institute.' [Ibid.: 390]

¹ The set of semantic roles depends on the theoretical framework. At any rate, in (12), *v institut* 'to the institute' is used metonymically and means 'to some people who work at the institute', that is, to some addressees.

Another deviation lacks identical morphological marking and ‘syntactic function’ of two parts of a comparative but preserves their identical lexical-semantic class. To illustrate, in (13), prepositional phrases are arguments of the verb *govorit* ‘in two different meanings. Their morphological marking is also different: Instrumental case vs. Locative case.

- (13) *Ja oxotnee govorju [pp s xudožnikami],*
 I.NOM willingly.COMP talk.PRS.1SG with artists.INST
čem [pp o xudožnikax].
 than about artists.LOC

‘I talk more willingly with artists than about artists.’ [Ibid.: 391]

Regarding the symmetry in meta-comparatives, Inkova [2013] points out that two parts of the meta-comparative *skoree... čem* ‘MC ... than’ (‘substitutive connector’ in her terms) have to be symmetric. Presumably, this means that they need to have identical syntactic categories and morphological forms. *Skoree* and *lučše* meta-comparatives have not been investigated with respect to syntactic and morphological (a)symmetry yet. The present study addresses the question of grammatical (a)symmetry in the three Russian meta-comparatives with *skoree*, *lučše*, and *bol’še*.

The paper has the following structure. Section 2 presents the goals of the current study that has two parts: corpus and experimental. Section 3 reports a corpus study, while Section 4 documents an experimental study. Section 5 concludes the paper.

2. The present study

The aim of the present study is to show that Russian meta-comparatives do not require symmetry between their two parts either in syntactic categories or morphological coding. To justify this claim, we collected and analyzed naturally occurring sentences from the Russian National Corpus, as well as native speaker judgements of sentences constructed on the basis of the occurrences observed in the corpus.

The results reveal three asymmetrical patterns that are not dependent upon the type of a meta-comparative or syntactic categories. The evidence for the first asymmetry comes from corpus occurrences that are functionally homogeneous in terms of [Sannikov 2008]. This means that they lack identical syntactic categories or morphological coding between the phrases in comparison. The second asymmetry does not demonstrate functional homogeneity. Instead, the examples involve either two clauses or a clause and a phrase that are compared. These two asymmetries have not been paid attention to in the literature. We tested them additionally in a follow-up experiment and observed that, overall, they are semi-grammatical and demonstrate further distinctions in their grammaticality status.²

3. Corpus study

3.1. Search queries and their results

The corpus study was based on data from the Russian National Corpus (further referred to as RNC, www.ruscorpora.ru), more precisely, on data from the Main corpus that comprises genre and stylistically diverse texts from the beginning of the 18th century to the present. In April 2021, when the study was conducted, the Main corpus contained 322 million tokens. The goal

² The corpus data and experimental materials can be accessed via <https://osf.io/qb5ju/>.

of the corpus study was to reveal the contexts in which the first part of a meta-comparative differed from its second part in terms of its syntactic category or morphological form. Accordingly, the search query in the corpus was as given in (14).

(14) Meta-comparative {*skoree* / *lučše* / *bol'she*} + first part + *čem* + second part³

The distance between the components of a search query ranged from 1 to 5 words. In each search query, one of the three meta-comparatives was used. Parts of meta-comparative sentences were noun phrases, adjectival phrases, adverbial phrases, prepositional phrases, and clauses with various verbal forms (finite forms in the indicative, subjunctive, and imperative mood; infinitives; participles; converbs). Although imperatives belong to finite forms on a par with indicatives and subjunctives, they were considered separately since they showed patterns distinct from the patterns of the other two finite forms. Therefore, 9 patterns (noun phrases, adjectival phrases, adverbial phrases, prepositional phrases, finite verb phrases (indicative and subjunctive forms), imperative verb phrases, infinitive phrases, participle phrases, converb phrases) were combined with each other in different orders and with each of the three meta-comparatives. This resulted in 72 combinations, or coarse-grained patterns. Each coarse-grained pattern was tested for each of the three meta-comparatives, resulting in $72 \times 3 = 216$ fine-grained patterns.

The contexts found with the help of the queries were manually browsed in order to identify meta-comparatives. Not all fine-grained patterns were found. Table 1 presents the number of occurrences of meta-comparative patterns that were attested in the corpus. The patterns are sorted from the most frequent to the least frequent ones. Sentences from (15) to (25) illustrate uses of some of the meta-comparatives.

According to Table 1 (p. 41), among 72 coarse-grained patterns, 30 patterns (41.7%) occurred in meta-comparative contexts, while 42 patterns (58.3%) were absent. Moreover, 3 coarse-grained patterns (4.2%) were found with all three meta-comparatives, while 6 patterns (8.4%) were observed in two meta-comparatives, and 21 patterns (29.2%) were used only in one type of meta-comparative.

Regarding the individual portraits of the meta-comparatives, *lučše* meta-comparatives received 142 occurrences (70.3%) in 20 coarse-grained patterns (47.6%), *skoree* meta-comparatives received 46 occurrences (22.8%) in 14 patterns (33.3%), while *bol'she* meta-comparatives received 14 occurrences (6.9%) in 8 patterns (19%).

These descriptive statistics are summarized in Figure 1.

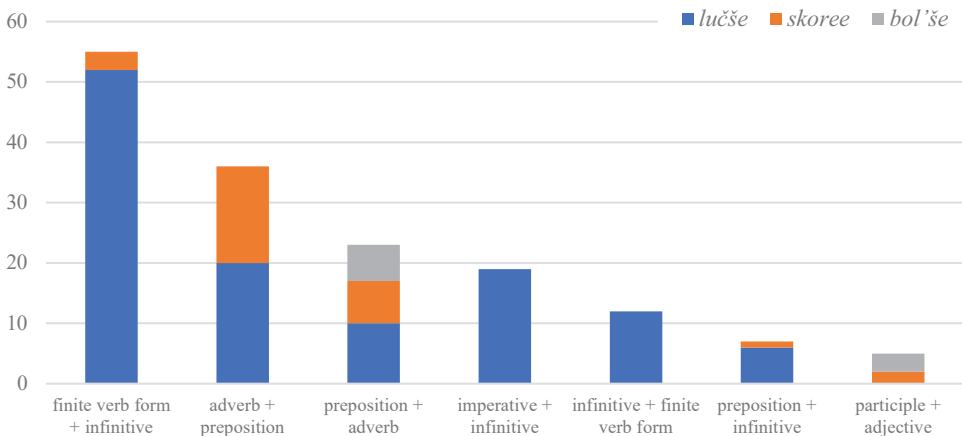


Fig. 1. The distribution of the most frequent asymmetrical patterns among the three meta-comparatives

³ Like standard comparatives, meta-comparatives allow not only the *čem* connector but also *neželi*. Since the former one is more frequent in meta-comparatives than the latter one [Zevakhina et al. 2017], the search queries used in this study include only the *čem* connector.

Table 1

Asymmetrical patterns of meta-comparatives attested in the corpus

#	Coarse-grained patterns in the search queries (with reference to some examples)	Number of <i>lučše</i> meta- comparative contexts	Number of <i>skoree</i> meta- comparative contexts	Number of <i>bol'she</i> meta- comparative contexts	Numbers of occurrences per combina- tion (in total)
1	finite form + infinitive, cf. ex. (15)	52	3	—	55
2	adverb + preposition, cf. ex. (18)	18	17	—	35
3	preposition + adverb	10	7	6	23
4	imperative + infinitive, cf. ex. (16)	19	—	—	19
5	infinitive + finite form	12	—	—	12
6	preposition + infinitive, cf. ex. (23)	6	1	—	7
7	adjective + noun, cf. ex. (19)	1	2	1	4
8	noun + infinitive, cf. ex. (21)	4	—	—	4
9	adjective + participle	—	4	—	4
10	participle + adjective	—	2	1	3
11	adjective + preposition	2	1	—	3
12	noun + adverb, cf. ex. (24)	1	1	1	3
13	infinitive + noun	3	—	—	3
14	infinitive + preposition	3	—	—	3
15	finite form + adverb	—	3	—	3
16	preposition + converb, cf. ex. (20)	—	2	1	3
17	preposition + noun	—	—	2	2
18	adverb + infinitive	2	—	—	2
19	finite form + noun, cf. ex. (22)	2	—	—	2
20	noun + preposition	2	—	—	2
21	converb + finite form	—	1	—	1
22	participle + finite form, cf. ex. (17)	—	—	1	1
23	adjective + finite form	—	—	1	1
24	adverb + noun, cf. ex. (25)	1	—	—	1
25	infinitive + adverb	1	—	—	1
26	preposition + participle	—	1	—	1
27	noun + participle	1	—	—	1
28	finite form + converb	—	1	—	1
29	converb + preposition	1	—	—	1
30	adverb + converb	1	—	—	1
Numbers of occurrences (in total)		142	46	14	202
Numbers of patterns (in total)		20	14	8	42
		(out of 72)	(out of 72)	(out of 72)	(out of 216)

As an anonymous reviewer noticed, three of the five leftmost columns show ‘some kind of verb form + some kind of verb form’, and the other two show ‘adverb + preposition’. Since

prepositional phrases and adverbial phrases tend to be functionally homogeneous (e.g., they tend to be adjuncts), what we see here is a preference to join similar things together: verb forms tend to go with verb forms, and adjunct-like phrases tend to go with adjunct-like phrases.

Regarding the grammatical features, infinitives were primarily imperfective (103 occurrences, 89.4%), and only a few of them were perfective (11 occurrences, 10.6%). Finite forms, that is, indicatives and subjunctives, were unequally distributed. In the coarse-grained pattern ‘finite form + infinitive’ in *lučše* meta-comparatives, there were 33 occurrences of subjunctive forms (64.7%) vs. 18 occurrences of future indicative forms (35.3%). In the coarse-grained pattern with the opposite order ‘infinitive + finite form’ in *lučše* meta-comparatives, the tendency was reversed: there were 9 occurrences of future indicative forms (81.8%) vs. 2 occurrences of subjunctives (18.2%). In both patterns, all future indicative forms were perfective. As for the subjunctive mood, let’s have a closer look at the coarse-grained pattern ‘finite form + infinitive’. Subjunctive forms in this pattern were perfective in 21 occurrences (63.6%) and imperfective in 12 occurrences (36.4%). This yields quite an interesting aspectual contrast between finite forms and infinitives. Finite forms tend to be perfective, and infinitives tend to be imperfective in meta-comparatives.

Imperatives (in the coarse-grained pattern ‘imperative + infinitive’) were almost equally distributed between 2nd-person imperatives (10 occurrences, 52.6%) and *pust’*-imperatives (9 occurrences, 47.4%).

There are three groups of meta-comparative sentences distinguished with respect to the clausal/phrasal status and the presence/absence of functional homogeneity. The first group, which is the most numerous (93 occurrences in our sample, that is 46%), comprises meta-comparative sentences with two clauses, in which morphological forms of verbs are distinct from each other. This is illustrated by examples (15)–(17). Functional homogeneity is not applicable to this group.

- (15) [_{CP} *My eě lučše poslušaem*], *čem* [_{TP} *samim vydumyvāt*].
we.NOM she.ACC MC listen.FUT.1PL than selves.INST think.up.INF

‘We would rather listen to her than invent [something] by ourselves.’

[A. Permjakov. In ‘Volga’, 2016]

- (16) *Lučše* [_{CP} *pojďemte k ozeru*], *čem* [_{TP} *zdes’ stojāt*].
MC go.IMP.1PL to lake.DAT than here stand.INF

‘We would rather go to the lake than stand here.’ [A. Skalon. *Strela letjaščaja* (1964)]

- (17) *Sil’no zametno, čto* [_{CP} *on bol’še progulivaem eju*]
very noticeable that he.NOM MC be.walked.PTCP.PRS she.INST
neželi.čem [_{CP} *on progulivaet eě*].
than he.NOM walk.PRS.3SG she.ACC

‘It seems that he is being walked by her rather than he is walking her.’

[V. Volodin. In ‘Volga’, 2000]

In (15), the first clause of a finite future indicative form is compared with the second clause with an infinitive. In (16), the first clause with a finite form in the imperative mood is compared with the second clause with an infinitive. In (17), the first clause with a participle is compared with the second clause with a finite form in the indicative mood and the present tense.

The second group, which is also relatively numerous (77 occurrences in the sample, 38%), is constituted by meta-comparative sentences with functionally homogeneous parts belonging to one clause, cf. examples (18)–(20).

- (18) *Skoree* [_{AdvP} *avtomatičeski*], *čem* [_{PP} *s interesom*],
MC automatically(ADV) than with(PREP) interest.INST
snjal odnu iz knig s šatkoj etažerki.
take.PST.SG.M one.ACC out.of books.GEN from wobbly bookcase.GEN

‘[He] took a book from the wobbly bookcase automatically rather than selectively.’

[V. Aksënov. *Novyj sladostnyj stil’* (1997)]

- (19) *Èto byli ljudi skoree* [_{AP} *bednye*],
 this be.PST.PL people.NOM MC poor(ADJ)
čem [_{NP} *srednego dostatka*].
 than average prosperity.GEN(N)
 ‘These were poor people rather than people of average prosperity.’ [S. Sergeev-Tsenskiy. Flot i krepost’ (1941)]

- (20) *Vsë budet xorošo, — skazal Foma,*
 everything be.FUT.3SG good say.PST.SG.M Foma.NOM
skoree [_{PP} *po privyčke*], *čem* [_{TP} *tak dumaja*].
 MC by(PREP) habit.DAT than so think.CONV
 ‘‘Everything will be good,’’ said Foma, by force of habit rather than thinking so.’
 [S. Osipov. Strasti po Fome... (1998)]

In (18), an adverbial phrase is compared with a prepositional phrase. Both are adjuncts and have the semantic role of manner. In (19), an adjectival phrase is compared with a noun phrase. Both are parts of a verb phrase with the copula *byli* ‘[they] were’ and are used predicatively. In (20), a prepositional phrase is compared with a converb clause. Both are adjuncts and have the semantic role of manner.

The third group (32 occurrences in the sample, 16 %) comprises meta-comparative contexts either with a clause and a phrase smaller than a clause or two phrases smaller than a clause. See examples (21)–(25). Functional homogeneity is not applicable to this group.

- (21) *Pust’ už lučše* <...> [_{NP} *vsemu konets*],
 IMP.PART PART MC everything.DAT end.NOM
čem [_{TP} *vsë vremja bojat’sja*].
 than all time.ACC be.afraid.INF
 ‘It’s better to finish everything than to be afraid all the time.’
 [V. Belousova. Žil na svete rycar’ bednyj (2000)]

- (22) *Lučše* [_{CP} *by my possorilis*],
 MC SUBJ we.NOM quarrel.PST.PL
čem [_{NP} *èto oskorbivšee menja nevnimanie*]!
 than this insult.PTCP.PST I.ACC inattention.NOM(N)
 ‘I would rather quarrel with you than [put up with] your being inattentive, which insulted me.’
 [V. Kaverin. Otkrytaja kniga (1949–1956)]

- (23) *Lučše* [_{PP} *na peredovuju*], *čem* [_{TP} *zdes’ korpet*].
 MC on(PREP) front.line.ACC than here persevere.INF
 ‘It is better [to go] to the front line than to persevere here.’
 [G. Baklanov. Yužnee glavnogo udara (1957)]

- (24) *Ja lučše* [_{NP} *pulju sebe v lob*], *čem* [_{AdvP} *tak-to*]!..
 I.NOM MC bullet.ACC(N) self.DAT in forehead.ACC than so(ADV)
 ‘I would rather [fire] a bullet in my forehead than [live] in this way!’
 [B. Vasil’ev. Byli i nebyli. Kniga 1 (1988)]

- (25) *Lučše* [_{AdvP} *ploxo*], *čem* [_{NP} *ničego*].
 MC badly(ADV) than nothing.GEN(PRON)
 ‘It is better [to speak] ill than [to say] nothing.’ [R. Frumkina. O nas — naikosok (1995)]

The sentences (21)–(23) include a clause and a phrase smaller than a clause. In (21), a noun phrase is compared with an infinitival clause, while in (22), a clause with a finite form in the subjunctive mood is compared with a noun phrase. In (23), a prepositional phrase is compared with an infinitive clause. The sentences (24) and (25) have two phrases smaller than a clause. In (24), a noun phrase is compared with an adverbial phrase, while in (25), an adverbial phrase is compared with a noun phrase with a pronoun.

However, it is hard to decide which syntactic analysis to propose for examples (21)–(25). Let us start with example (24). We can assume that the first part of a meta-comparative in (24) is a full clause with an elided verbal finite form (e.g., *pušču* ‘[I] will fire’) and its second part is a clause with an elided finite form and an infinitive that forms the analytic future tense (*budu žit* ‘[I] will live’), see (26a) repeated from (24) with added verbal forms. An alternative analysis for (24) is that the second part includes only an elided infinitival form (e.g., *žit* ‘to live’), see (26b).

- (26) a. *Ja lučše* [CP *pušču pulju sebe v lob*],
 I.NOM MC shoot.FUT.3SG bullet.ACC(N) self.DAT in forehead.ACC
čem [CP *tak-to budu žit*]!..
 than so(ADV) be.FUT.1SG live.INF
 ‘I would rather [fire] a bullet in my forehead than [will live] in this way!’
 [B. Vasil’ev. Byli i nebyli. Kniga 1 (1988)]

- b. *Ja lučše* [CP *pušču pulju sebe v lob*],
 I.NOM MC shoot. FUT.3SG bullet.ACC(N) self.DAT in forehead.ACC
čem [TP *tak-to žit*]!..
 than so(ADV) live.INF
 ‘I would rather [fire] a bullet in my forehead than [to live] in this way!’
 [B. Vasil’ev. Byli i nebyli. Kniga 1 (1988)]

Regarding examples (21)–(23), it is less clear which lexical items are elided. Below we repeat these examples adding verbal forms that can be assumed to be elided, cf. (27)–(29):

- (27) *Pust’ už lučše* <...> [CP *vsemu nastupit/pridět konets*],
 IMP.PART PART MC everything.DAT come.FUT.3SG end.NOM
čem [TP *vsě vremja bojat’sja*].
 than all time.ACC be.afraid.INF
 ‘It’s better to finish everything than to be afraid all the time.’
 [V. Belousova. Žil na svete rycar’ bednyj (2000)]

- (28) *Lučše* [CP *by my possorilis*], *čem* [TP *vyslušivat’/terpet’ èto*
 MC SUBJ we quarrel.PST.PL than listen.INF/put.up.with.INF this
oskorbivšee menja nevnimanie]!
 insult.PTCP.PST I.ACC inattention.NOM(N)
 ‘I would rather quarrel with you than [put up with] your being inattentive which insulted me.’
 [V. Kaverin. Otkrytaja kniga (1949–1956)]

- (29) *Lučše* [TP *poexat’/otpravit’sja na peredovuju*],
 MC go.INF/set.off.INF on(PREP) front.line.ACC
čem [TP *zdes’ korpet*].
 than here persevere.INF
 ‘It is better [to go] to the front line than to persevere here.’
 [G. Baklanov. Yužnee glavnogo udara (1957)]

Finally, consider (25). It is also not clear which lexical items are elided in this context. It can be *govorit* ‘to say’, *otzyvat’sja* ‘to speak [somehow of somebody]’, *dumat* ‘to think’ or some other speech or mental verbs.

3.2. Discussion

Let us summarize the main findings of the corpus study. Although the number of occurrences of asymmetric meta-comparatives shows that they represent a marginal grammatical

phenomenon, the corpus provided enough data to make the following generalizations about their syntactic and morphological properties.

Firstly, *lučše* is used in asymmetric meta-comparative sentences more often than *skoree* and *bol'se*. It is worth pointing out that asymmetric meta-comparative sentences with *lučše* demonstrate at least the same grammatical restrictions as symmetric meta-comparative sentences with *lučše*, see more in Section 1 with a reference to [Zevakhina et al. 2017]. The present corpus study added to this list more syntactic categories and more morphological forms which *lučše* is compatible with.

Secondly, asymmetric meta-comparatives show bias towards imperfective infinitives and perfective finite forms.

Thirdly, asymmetric meta-comparatives are either bi-clausal, bi-phrasal or mixed (clausal-phrasal). Their bi-phrasal variety shows functional homogeneity. The clausal/phrasal status of the parts of meta-comparatives is dependent upon verbal ellipsis, which is sometimes hard to determine and seems to be conventionalized.

The experiment reported in the next section tests bi-clausal and mixed (clausal and phrasal) groups of asymmetric meta-comparatives. For both groups, we test the grammaticality of patterns with infinitives. For the former group, we take the most frequent coarse-grained pattern 'finite form + infinitive', while for the latter group, we take the coarse-grained pattern 'noun + infinitive' that lacks a finite verbal form. Both patterns are tested with the *lučše* meta-comparative, which is the most frequent across all patterns in the corpus, and with the *skoree* meta-comparative.

4. Experimental study

4.1. Hypotheses

The first hypothesis (H1) is that the asymmetric meta-comparatives *lučše* and *skoree* would be judged as semi-grammatical in contrast to grammatical and ungrammatical sentences. This hypothesis is mainly motivated by the low number of occurrences of asymmetric meta-comparatives in the corpus. A plausible reason for their rarity is that they lie in the periphery of Russian grammar and seem to be neither fully grammatical nor entirely ungrammatical.

According to the second hypothesis (H2), *lučše* meta-comparatives would be evaluated as more grammatical than *skoree* meta-comparatives, since the former were more frequent in the corpus and revealed more patterns in the corpus study than the latter ones.

The third hypothesis (H3) can be divided into two parts.

On the one hand, infinitive clauses and noun phrases can take argument positions in a clause and be coordinated, cf. (30), and are also grammatical as parts of standard comparatives, cf. (31). Therefore, we hypothesize (H3a) that the pattern 'noun + infinitive' would be evaluated as more grammatical than the pattern 'finite form + infinitive'.

- (30) *Ja ljublju tantsy i pet'.*
 I.NOM like.PRS.1SG dances.ACC and sing.INF
 'I like dances and singing.'

- (31) *Ja bol'se ljublju tantsy, čem pet'.*
 I.NOM more like.PRS.1SG dances.ACC than sing.INF
 'I like dances more than singing.'

On the other hand, as seen from the corpus data, meta-comparatives with finite verbal forms and infinitives are more frequent than those comparing infinitive clauses and noun phrases. According to the hypothesis (H3b), the pattern 'finite form + infinitive' would be evaluated as more grammatical than the pattern 'noun + infinitive'.

4.2. Participants

A hundred and five people voluntarily participated in the experiment. Seventy-seven of them were women and twenty-eight were men. Their ages ranged from 20 to 48 years, with the mean age being 23 years.

4.3. Design and materials

Critical items included 24 sentences, which were similar to the meta-comparatives observed in the corpus study from Section 2. The experiment had a 2 x 2 factorial design: meta-comparative contexts (*lučše* vs. *skoree*) x asymmetric grammatical forms in two parts of a meta-comparative ('finite form + infinitive' vs. 'noun + infinitive'). Each of the 4 experimental conditions had 6 versions. This yielded 24 critical items illustrated by examples (32)–(35).

(32) *lučše* + finite form + *čem* + infinitive

[_{CP} *Alyona lučše fil'm posmotrit*], *čem* [_{TP} *delami zanimat'sja*].
Alyona.NOM MC film.ACC watch.FUT.3SG than work.INST.PL do.INF

'Alyona would rather watch a film than do work.'⁴

(33) *skoree* + finite form + *čem* + infinitive

[_{CP} *Ix roditeli skoree oslepnut*], *čem* [_{TP} *videt' ètot užas*].
their parents.NOM MC go.blind.FUT.3SG than see.INF this horror.ACC

'Their parents would rather go blind than see this terrifying thing.'

(34) *lučše* + noun + *čem* + infinitive

[_{NP} *Dlja nas lučše družnaja sem'ja*], *čem* [_{TP} *žit' bogato*].
for we.GEN MC close-knit family.NOM(N) than live.INF richly

'For us, it is better to be a close-knit family than to live richly.'

(35) *skoree* + noun + *čem* + infinitive

[_{NP} *Tut važno skoree umenie*], *čem* [_{TP} *iskat' lëgkie puti*].
here important MC skill.NOM(N) than search.INF easy ways.ACC

'In this case, it is more important to have a skill than to look for an easy way.'

All the critical items that constitute one condition resembled each other. Finite forms were in the indicative mood, perfective aspect, and future tense. The reason for choosing these forms was that, firstly, they were one of the most frequent verbal forms observed in the corpus study and, secondly, the patterns that included nouns and infinitives were only in the indicative mood in the corpus. All the infinitives were derived from imperfective verbs, since, according to the corpus study, imperfective infinitives were more frequent than the perfective ones.

The experiment involved 26 control items. Since the critical items were expected to be evaluated as semi-grammatical, one half of the control items were grammatically correct, and the other half was grammatically incorrect (13 vs. 13 respectively). Grammatically incorrect

⁴ An anonymous reviewer pointed out that the reverse clause order is also possible, cf.: *Čem* [_{TP} *delami zanimat'sja*], [_{CP} *Alyona lučše fil'm posmotrit*]. Moreover, it sounds more natural. It is worthy to note that *skoree... čem* is also possible with the reverse clause order. However, only the direct order for *bol'se... čem* is grammatical, whereas the reverse order is non-grammatical. It is an interesting issue, and we thank an anonymous reviewer for drawing our attention to it. We leave it for future research.

control items contained a preposition in an incorrect position. The sentences (36)–(37) illustrate control items.

- (36) Example of a grammatically correct control item

Progulka v solnečnyj den' podnimet vam nastroenie.
 walk.NOM on sunny day.ACC raise.FUT.3SG you.DAT mood.ACC
 'A walk on a sunny day will cheer you up.'

- (37) Example of a grammatically incorrect control item⁵

**Po ego mneniju, tebja est' u vse šansy postupit'.*
 by(PREP) his opinion.DAT you.GEN exist by(PREP) all chances.NOM enter.INF
 'You have every chance of being admitted to the university.'

Besides 24 critical items and 26 control items, the experiment included three control items always presented before the main part of the experiment. The control items consisted of two grammatically correct sentences and one grammatically incorrect one, in which a preposition also had a wrong position. To conclude, the experiment consisted of 53 sentences such that their length ranged from 7 to 11 words.

4.4. Procedure

The experiment was administered at the Ibex Farm (<https://ibex.spellout.net>), which was a web platform to host online psycholinguistic experiments until September 30, 2021. This Internet resource is designed in such a way that it makes it possible to randomly order critical and control items for each participant, to measure the reaction times of the participants' responses to given items, and to restrict the time for an answer.

The experiment was carried out as a grammaticality judgment task. This means that participants had to read a critical or control sentence presented on the computer screen, press the spacebar, and answer a question about whether the sentence was grammatically correct, either selecting the 'F' key for 'yes' or selecting the 'J' key for 'no'. The time to give an answer was limited to 5 seconds. This was a reasonable time interval not to be rushed and not to be distracted. The experiment started with a questionnaire in which participants were asked to fill in their personal information required for statistical purposes: name, age, gender, and first language. Furthermore, participants had to confirm their agreement to participate in the experiment and to give permission for data processing. The next step was to read the following instruction (here translated into English): 'Your task is to read sentences and decide whether they are grammatically correct, that is, whether they are constructed according to the rules of Russian grammar. To illustrate, the sentence *I am very grateful for the help* is correct from a grammatical point of view, while the sentence *I are very grateful the help for* is not, since it contains a grammatical error. To start the experiment, press the spacebar. After reading the sentence, you will need to answer the question whether it is grammatically correct. You have up to 5 seconds. Press the 'F' key to choose 'yes' and the 'J' key to choose 'no'. After that period, you will be presented with the next sentence. To move from one sentence or question to another, press the spacebar. The experiment takes around 10 minutes to complete. There are no correct or incorrect answers. Don't think too long and don't be in a hurry. The first 3 sentences will be training.'

⁵ The sentence (37) is ungrammatical since the preposition *u* selects a noun phrase in the Genitive case and is detached from its argument *tebja*.

4.5. Results

Responses from five participants were excluded from the statistical analysis. Three of these participants were not native Russian speakers. Two other participants gave incorrect responses to more than 50 % of the control items. Furthermore, nine critical and seven control items received from several participants (that is, 16 data points) were removed since they were given null answers. This left 2391 answers for the critical items and 2593 for the control items (in total, 4984 answers).

Figure 2 shows the distribution of *yes/no*-answers to the critical and (in)correct control items.

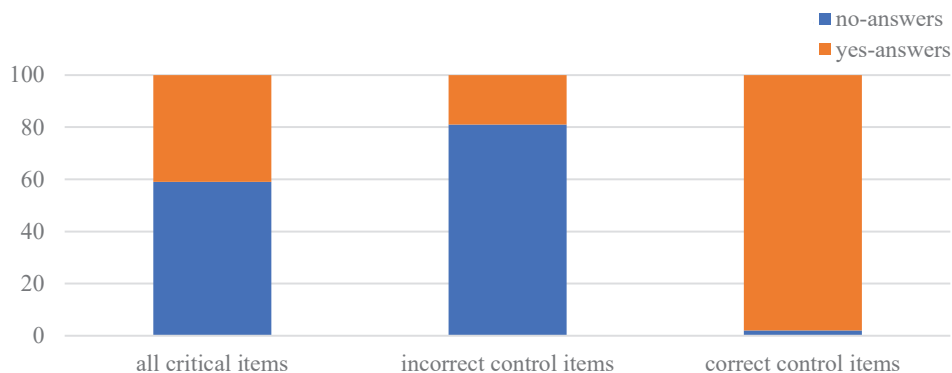


Fig. 2. The distribution of the answers to critical items and (in)correct control items

Using R [R Core Team 2020], we performed a generalized mixed-effects logistic regression model [Bates et al. 2015] with random intercepts included for subjects and items to test H1.

The correct control items received significantly more *yes*-answers than the incorrect control items ($\beta = -6.4918$, $SE = 0.5261$, $Z = -12.34$, $p < 0.0001$). The incorrect control items received around 20 % of *yes*-answers. For all of these cases, the participants commented on their decisions, saying that they considered the wrong position of prepositions as typos, rather than errors. This could have been avoided if the instruction had included something like ‘there are no typos in the sentences’. However, statistically this drawback of the procedure was not that important.

The critical items received significantly more *yes*-answers than the incorrect control items ($\beta = -1.7708$, $SE = 0.4286$, $Z = -4.132$, $p < 0.0001$), but significantly less *yes*-answers than the correct control items ($\beta = 4.7209$, $SE = 0.4624$, $Z = 10.210$, $p < 0.0001$). Moreover, as seen in Figure 2, the critical items were evaluated as semi-grammatical (around 42 % of *yes*-answers vs. around 58 % of *no*-answers). This confirms hypothesis H1.

Figure 3 shows the distribution of *yes/no*-answers for the following four meta-comparatives: *lučše* + finite form + *čem* + infinitive, *skoree* + finite form + *čem* + infinitive, *lučše* + noun + *čem* + infinitive, *skoree* + noun + *čem* + infinitive.

According to Figure 3, the participants evaluated the pattern ‘noun + infinitive’ with *lučše* meta-comparatives as the most grammatical (around 73 % *yes*-answers). The pattern ‘finite form + infinitive’ with *lučše* meta-comparatives was semi-grammatical (around 42 % *yes*-answers). Finally, the two patterns with *skoree* meta-comparatives were evaluated as the least grammatical, but still not as entirely ungrammatical (with 74 % and 71 % *no*-answers respectively).

To test H2 and H3, using R [R Core Team 2020], we performed a generalized mixed-effects logistic regression model [Bates et al. 2015], with random intercepts included for subjects and items.

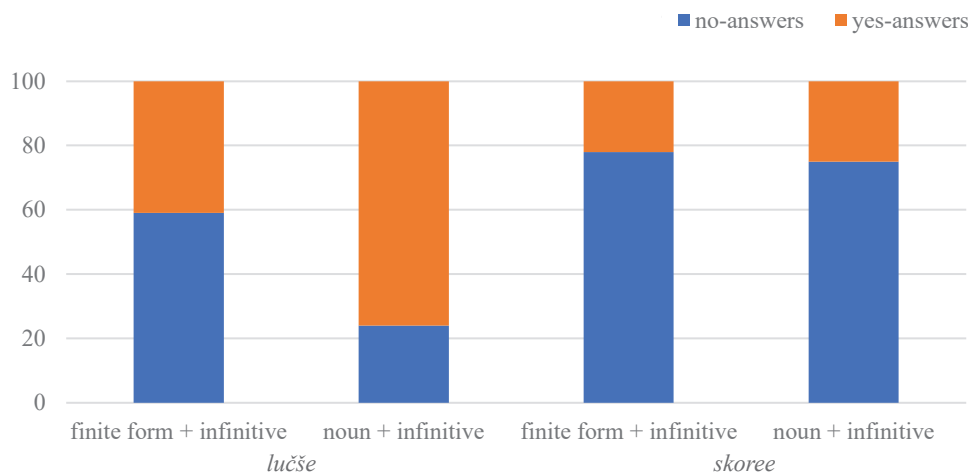


Fig. 3. The distribution of *yes/no*-answers for ‘finite form/noun + infinitive’ patterns in *lučše* and *skoree* meta-comparatives

Lučše meta-comparatives were evaluated as more grammatical than *skoree* meta-comparatives ($\beta = -1.9242$, $SE = 0.4344$, $Z = -4.429$, $p < 0.0001$). This confirms hypothesis H2. The pattern ‘noun + infinitive’ was judged to be more grammatical than the pattern ‘finite form + infinitive’ ($\beta = 1.1293$, $SE = 0.5393$, $Z = 2.094$, $p = 0.03625$). This confirms hypothesis H3a. The interaction of the factor of the meta-comparative type (with 2 levels: *lučše* and *skoree*) vs. the factor of pattern (with 2 levels: ‘noun + infinitive’ and ‘finite form + infinitive’) was also significant ($\beta = -1.8349$, $SE = 0.6290$, $Z = -2.917$, $p = 0.00353$).

Regarding the individual patterns, in *lučše* meta-comparatives, the pattern ‘noun + infinitive’ received significantly more *yes*-answers than the pattern ‘finite form + infinitive’ ($\beta = 2.0397$, $SE = 0.4461$, $Z = 4.573$, $p < 0.0001$). Furthermore, the pattern ‘noun + infinitive’ in *lučše* meta-comparatives received significantly more *yes*-answers than the pattern ‘finite form + infinitive’ in *skoree* meta-comparatives ($\beta = -3.0430$, $SE = 0.4507$, $Z = -6.752$, $p < 0.0001$), as well as the pattern ‘noun + infinitive’ in *skoree* meta-comparatives ($\beta = -2.8383$, $SE = 0.4496$, $Z = -6.313$, $p < 0.0001$). The pattern ‘finite form + infinitive’ in *lučše* meta-comparatives received significantly more *yes*-answers than the pattern ‘finite form + infinitive’ in *skoree* meta-comparatives ($\beta = -1.0034$, $SE = 0.4430$, $Z = -2.265$, $p = 0.0235$). Other comparisons did not reveal any statistically significant distinctions (all p ’s > 0.05).

4.6. Discussion

The results of the experimental study showed that meta-comparatives were evaluated as semi-grammatical. This agrees with the results of the corpus study that also demonstrated marginality of asymmetric meta-comparatives.

Furthermore, the results revealed that *lučše* meta-comparatives were judged as more grammatical. This also agrees with the results of the corpus study, where they were more frequent than meta-comparatives of other types.

Finally, the pattern ‘noun + infinitive’ was evaluated as more grammatical than the pattern ‘finite form + infinitive’ for *lučše* meta-comparatives. This is not expected from the corpus data that demonstrated a reverse contrast. A possible reason is that nouns and infinitives take argument positions in a clause, and, in this sense, they are functionally homogeneous.

5. Conclusion

To conclude, the paper presents certain morphological and syntactic properties and shows the diversity among their various grammatical patterns. Morphologically, infinitives tend to be imperfective, and finite verb forms (in the subjunctive and indicative mood) tend to be perfective in asymmetric meta-comparatives. Syntactically, asymmetric meta-comparatives are either bi-clausal, bi-phrasal, or mixed (clausal-phrasal). Their bi-phrasal variety shows functional homogeneity. The clausal/phrasal status of the parts of meta-comparatives is dependent upon verbal ellipsis, which is sometimes hard to determine and seems to be conventionalized. *Lučše* meta-comparatives are more frequent in the corpus and are evaluated as more grammatical than *skoree* meta-comparatives. The asymmetric pattern ‘noun phrase + infinitive phrase’ is evaluated as more grammatical than the pattern ‘verb phrase (finite form) + infinitive phrase’ for *lučše* meta-comparatives, while there is no such contrast between these patterns for *skoree* meta-comparatives. All these facts shed light on the grammatical portrait of Russian meta-comparatives as well as on the use and evaluation of asymmetric meta-comparatives in a natural language.

By providing corpus and experimental evidence, the paper contributes to our deeper understanding of asymmetric (meta-)comparative structures that represent a marginal though important phenomenon of language grammar, demonstrate morphological and syntactic properties, and show within-group diversity in terms of their grammaticality status. The study contributes to the discussion of grammatical asymmetry in parallel syntactic structures (comparative and coordinated structures, among others).

ABBREVIATIONS

1, 3 — 1, 3 person
ACC — accusative case
ADJ — adjective
ADV — adverb
COMP — standard comparative
CONV — converb
DAT — dative case
FUT — future tense
GEN — genitive
IMP — imperative mood
IMP.PART — imperative particle
INF — infinitive
INST — instrumental case

LOC — locative case
MC — meta-comparative
NOM — nominative case
N — noun
PART — particle
PL — plural
PREP — preposition
PRON — pronoun
PRS — present tense
PTCP — participle
PST — past tense
SG — singular
SUBJ — subjunctive

REFERENCES

- Bates et al. 2015 — Bates D., Mächler M., Bolker B., Walker S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 2015, 67(1): 1–48.
- Beaulieu-Masson, Inkova-Manzotti 2003 — Beaulieu-Masson A., Inkova-Manzotti O. “Plutôt que”: de la comparaison à la substitution. *L'Analisi Linguistica e Letteraria*, 2003, 11: 563–594.
- Culicover, Jackendoff 1997 — Culicover P. W., Jackendoff R. Semantic subordination despite syntactic coordination. *Linguistic Inquiry*, 1997, 28: 195–218.
- Giannakidou, Stavrou 2009 — Giannakidou A., Stavrou M. On metalinguistic comparatives and negation in Greek. *MIT Working Papers in Linguistics*, 2009, 57: 57–74.
- Giannakidou, Yoon 2011 — Giannakidou A., Yoon S. The subjective mode of comparison: Metalinguistic comparatives in Greek and Korean. *Natural Language and Linguistic Theory*, 2011, 29(3): 621–655.

- Goncharov 2014 — Goncharov J. Comparing propositional attitudes. *McGill Working Papers in Linguistics*, 2014, 24(1): 17–36.
- Goncharov, Irímia 2018 — Goncharov J., Irímia M. A. Epistemic comparatives: a cross-linguistic investigation. *West Coast Conference on Formal Linguistics*, 2018, 35: 169–176.
- Goncharov, Irímia 2020 — Goncharov J., Irímia M. A. Epistemic comparatives and other expressions of speaker's uncertainty. *Advances in formal Slavic linguistics 2017*. Marušić F., Mišmaš P., Žaucer R. (eds.). Berlin: Language Science Press, 2020, 75–95.
- Herburger, Rubinstein 2014 — Herburger E., Rubinstein A. Is 'more possible' more possible in German? *Semantics and Linguistic Theory*, 2014, 24: 555–576.
- Horn 1989 — Horn L. *A natural history of negation*. Chicago: Chicago Univ. Press, 1989.
- Inkova-Manzotti 2001 — Инькова-Манзотти О. Ю. *Коннекторы противопоставления во французском и русском языках (сопоставительное исследование)*. М.: Изд-во Московского ун-та, 2001. [Inkova-Manzotti O. Yu. *Konnektory protivopostavljeniia vo francuzskom i ruskom yazykakh (sopostavitel'noe issledovanie)* [Connectors of contrast in French and Russian (a comparative study)]. Moscow: Moscow Univ. Press, 2001.]
- Inkova 2013 — Инькова О. Ю. О семантике так называемых градационных союзов *не столько... сколько и скорее... чем*. *Вопросы языкознания*, 2013, 1: 38–52. [Inkova O. Yu. On the semantics of so-called gradational connectors *ne stol'ko... skol'ko* and *skoree... čem*. *Voprosy Jazykoznanija*, 2013, 1: 38–52.]
- Jespersen 1924 — Jespersen O. *The philosophy of grammar*. London: George Allen & Unwin LTD, 1924.
- McCawley 1991 — McCawley J. D. Contrastive negation and metalinguistic negation. *Chicago Linguistic Society*, 1991, 27(2): 189–206.
- Morzycki 2011 — Morzycki M. Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics*, 2011, 19(1): 39–86.
- R Core Team 2020 — R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2020. <https://www.R-project.org/>.
- Sannikov 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.]
- Sassoon 2015 — Sassoon G. W. Between-Noun Comparisons. *Logic, Language, and Computation, TbiLLC 2013, Lecture Notes in Computer Science*. Vol. 8984. Aher M., Hole D., Jeřábek E., Kupke C. (eds.). Berlin: Springer, 2015, 276–289.
- Sawada 2007 — Sawada O. Pragmatic properties of the Japanese scalar reversal adverbs. *Talk at the Midwest Workshop on Semantics*, 2007.
- Stevenson, Waite (eds.) 2011 — Stevenson A., Waite M. (eds.). *Concise Oxford English dictionary*. Oxford: Oxford Univ. Press, 2011.
- von Fintel, Kratzer 2014 — von Fintel K., Kratzer A. Modal comparisons: two dilettantes in search of an expert. *The art and craft of semantics: a festschrift for Irene Heim*. Crnič L., Sauerland U. (eds.). MIT Working Papers in Linguistics, 2014, 70(1), 175–179.
- Zevakhina et al. 2017 — Zevakhina N., Dzhakupova S., Vishenkova A. The grammar of Russian metalinguistic comparatives. *Scando-Slavica*, 2017, 63(2), 179–197.

Скажем так: дискурсивные маркеры, восходящие к глаголам речи в русском языке

© 2024

Вера Исааковна Подлеская

Институт языкознания РАН, Москва, Россия; podlesskaya@iling-ran.ru

Аннотация: На основании данных НКРЯ, и прежде всего, его мультимедийного подкорпуса (МУРКО), исследуются два представителя семейства дискурсивных маркеров, восходящих к личной форме глагола речи *скажем* — изолированное *скажем* и его комбинации с наречием, в том числе местоименным, — *скажем прямо*, *скажем так* и др. Показано, что эти маркеры демонстрируют типичный эффект прагматикализации: в отличие от исходного речевого глагола, они не отсылают к речевому событию, их основной функцией становится выражение нечеткой номинации, или хеджирования, т. е. эти маркеры используются тогда, когда говорящий снимает с себя ответственность за возможную неточность пропозиционального содержания или возможную неточность используемой номинации применительно к фиксированному пропозициональному содержанию. Продемонстрированы основные сдвиги в синтактике этих маркеров, которые происходят в связи с тем, что они выпадают из жесткой клаузуальной структуры и приобретают статус оператора, сферой действия которого являются сегменты дискурса различной (потенциально — любой) протяженности. Инструментально и перцептивно исследованы основные просодические свойства этих маркеров. Кроме того, исследовано употребление маркеров этого семейства совместно с другими средствами, сигнализирующими, что говорящему требуются дополнительные усилия при выборе адекватного средства выражения. Это могут быть другие лексические маркеры нечеткой номинации, лексические и просодические показатели открытого списка, самоисправления, долексические маркеры хезитации (заполненные паузы, нефонологические удлинения звуков) и пр. Полученные результаты убедительно показывают, (1) что даже внутри одного семейства маркеров сдвиг от единицы с пропозициональным значением к дискурсивному маркеру может проходить по различающимся траекториям и приводить к большим или меньшим сдвигам в значении и форме и (2) что исследование дискурсивных маркеров с привлечением данных о просодии и таких специфических феноменах устного дискурса, как речевые затруднения, позволяет сделать лексикографические портреты отдельной единицы и их семейств более многомерными.

Ключевые слова: глаголы речи, дискурсивные маркеры, русский язык, прагматикализация, просодия, устный дискурс

Благодарности: Исследование проводилось при поддержке гранта РНФ № 22-18-00120.

Для цитирования: Подлеская В. И. *Скажем так: дискурсивные маркеры, восходящие к глаголам речи в русском языке. Вопросы языкознания*, 2024, 2: 52–82.

DOI: 10.31857/0373-658X.2024.2.52-82

Skažem tak: Discourse markers originating from verbs of speech in Russian

Vera I. Podlesskaya

Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia; podlesskaya@iling-ran.ru

Abstract: Based on data from the Russian National Corpus, first and foremost, from its multimedia corpus, we address two members of the family of discourse markers developed from the finite

form of the verb of speech *skažem* (the first person plural present tense form of the verb ‘say’): isolated *skažem* and its combinations with an adverb or a pronominal adverb such as *skažem prjamo* ‘let’s put it bluntly’, *skažem tak* ‘let’s put it this way’, etc. We show that these markers demonstrate a typical pragmaticalization trajectory: unlike the original speech verb, they do not refer to a speech event, their main function is to express a fuzzy nomination, or hedging, i.e., these markers are used when the speaker disclaims responsibility for the possible inaccuracy of the propositional content or of the nomination used in relation to the fixed propositional content. We demonstrate the main shifts in the distribution of these markers, which occur due to the fact that they fall out of the rigid clausal structure and acquire the status of an operator which scopes over discourse segments of various (potentially any) length. We study the main prosodic properties of these markers instrumentally and perceptually and, in addition, discuss the use of markers of this family together with other means of signaling that the speaker needs additional efforts in choosing an adequate expression. These can be other lexical markers of fuzzy nomination, lexical and prosodic indicators of an open list, self-corrections, signals of hesitation (filled pauses, non-phonological lengthening of sounds), etc. The results obtained convincingly show: (1) that even within the same family of markers, the shift from a unit with propositional meaning to a discourse marker can follow different trajectories and lead to larger or smaller shifts in meaning and form; and (2) that the study of discourse markers that uses data on prosody and such specific phenomena of oral discourse as speech disfluencies makes it possible to get more comprehensive lexicographic portraits of single units and their families.

Keywords: discourse markers, oral discourse, pragmaticalization, prosody, Russian, verbs of speech

Acknowledgements: The current research was supported by the Russian Science Foundation, grant No. 22-18-00120.

For citation: Podlesskaya V. I. *Skažem tak*: Discourse markers originating from verbs of speech in Russian. *Voprosy Jazykoznanija*, 2024, 2: 52–82.

DOI: 10.31857/0373-658X.2024.2.52-82

1. Постановка задачи

Предметом данного исследования являются дискурсивные маркеры, восходящие к глаголам речи в русском языке. В статусе дискурсивного маркера глагол частично или полностью утрачивает способность к словоизменению — по существу, консервируется одна из словоформ парадигмы. Вместе с утратой «морфологических свобод» происходят сдвиги в значениях и синтактике. Значение сдвигается от пропозиционального (для глаголов речи — это обозначение и описание речевого события) к прагматическому — выражение эпистемической и других типов модальности, обеспечение когерентности, регулирование процесса коммуникации и пр.; в терминах У. Чейфа единица приобретает статус регуляторной [Chafe 1994: 63–69]. Изменения в синтактике могут быть описаны термином «кооптация» (ср. cooptation — термин, введенный в работах [Kaltenböck et al. 2011; Heine et al. 2021] и других публикациях этой группы авторов): единица выпадает из жесткой клаузуальной структуры и приобретает статус оператора, сферой действия которого являются сегменты дискурса различной (потенциально — любой) протяженности.

В центре нашего внимания будет дискурсивный маркер СКАЖЕМ, на примере которого мы постараемся проследить процессы, связанные со сдвигом от глагольной словоформы (в данном случае формы презенса 1 л. мн. ч.) к дискурсивному маркеру. Материалом служат данные [НКРЯ]: основного и (в отдельных случаях) поэтического корпуса для анализа значения и структурных свойств конструкций со СКАЖЕМ и мультимедийного корпуса (МУРКО) — для анализа просодии. Нас будут интересовать два варианта функционирования СКАЖЕМ: вариант, при котором СКАЖЕМ выступает изолированно, см. (1), и вариант, при котором СКАЖЕМ сопровождается наречием или местоименным наречием — СКАЖЕМ ПРЯМО / ЧЕСТНО / ТАК / ИНАЧЕ и др., см. (2):

- (1) Вот почему порой
художник оказывается заморожен горой
или, скажем, развалинами.
[И. А. Бродский. Посвящается Пиранези (1993–1995)]
- (2) Да, прямо скажем, этот край
Нельзя назвать дорогой в рай.
[К. М. Симонов. Мурманские дневники (1938)]

Глаголы речи в русском языке, в том числе флагманские *говорить* и *сказать*, прагматизируются очень широко. В результате возникают как дискурсивные маркеры, типа репортативов *грит/грю* (а в более далекой исторической перспективе — *мол от молвить*)¹, так и дискурсивные формулы — единицы, типа *кто бы говорил*, которые выступают как клишированные реакции на ту или иную речевую ситуацию и могут самостоятельно формировать реплику в диалоге [Рахилина и др. 2021]. В прагматизированном значении глаголы речи могут входить и в более сложно организованные конструкции, имеющие в толковании одну или более переменных. Такова, например, конструкция *У, чтобы не сказать Х*, которая используется для объяснения выбора одной из двух возможных номинаций — таких, что одна вынужденно менее категорична, чем другая, ср. *она не очень умная, чтобы не сказать дура*.

При сдвиге такого рода могут консервироваться разные формы глагольной парадигмы: инфинитив (*сказать по правде*), деепричастие (*откровенно говоря*), императив (*скажи на милость!*), презенс возвратного глагола (*как говорит*). Закрепленная в интересующем нас маркере СКАЖЕМ форма презенса 1 л. мн. ч., как кажется, не представлена больше ни в каких других маркерах и формулах, восходящих к глаголам речи, однако обнаруживается в смежных зонах — в частности, в единицах с эпистемическим значением, восходящих к глаголам мысли и восприятия: *посмотрим* (в значении ‘может быть’), *допустим, предположим* и др.

Путь прагматизации всегда индивидуален, лексически и структурно сходные единицы дают разный, нередко противоположный эффект. Например, отрицательная форма императива глагола *говорить* — *(и) не говори!* — становится формулой эмоционального подтверждения, а отрицательная форма императива глагола *сказать* — *(ну) не скажи!* — становится формулой эмоционального опровержения. Или другой пример. Сравним вышеупомянутую конструкцию *У, чтобы не сказать Х* со структурно сходной — отличающейся местом отрицания — конструкцией *не сказать чтобы Х: она не очень умная, чтобы не сказать дура* vs. *она не сказать чтобы очень умная*. Первая конструкция означает, что она ‘совсем не умна’, вторая — что она ‘не очень умна’.

В рамках данной работы мы исследуем, во-первых, значение и употребление маркера СКАЖЕМ — этому посвящен раздел 2, во-вторых, связь его просодических свойств с его линейным расположением — этому посвящен раздел 3, и в-третьих, связь сферы действия маркера с речевыми затруднениями и сбоями в его окрестностях — этот вопрос обсуждается в разделе 4. В разделе 5 мы подводим итоги.

2. Значение и употребление маркера СКАЖЕМ

Как было сказано выше, нас будут интересовать два варианта употребления маркера СКАЖЕМ: изолированное (далее — СКАЖЕМ1) и в сопровождении наречия или местоименного наречия (далее — СКАЖЕМ2). Оба варианта используются для выражения так

¹ Подробное описание синтаксических и морфологических сдвигов от полнозначных глаголов речи к репортативам в русском языке см. в [Копотев 2014]. (Автор использует термин «ренарративы» для этого класса единиц.)

называемой «нечеткой номинации», или хеджирования. Термины hedge ‘уклончивое выражение’, в буквальном переводе ‘прикрытие’, и hedging ‘использование уклончивых выражений’ вводятся в работе [Lakoff 1972]; ср. также определение в словаре [Баранов, Добровольский 1996: 265]: «Лексическая единица, размывающая границы экстенционального множества языкового выражения, и тем самым ограничивающая ответственность говорящего за сказанное».

Этому феномену посвящен внушительный корпус литературы (ср., например, объемные коллективные монографии [Markkanen, Schröder 1997; Kaltenböck et al. 2010]). В работе [Lakoff 1972] вводится (а в работе [Prince et al 1982] и ряде других публикаций разрабатывается) дальнейшее деление хеджей: на так называемые аппроксиматоры и «щиты» (shields). Оба типа используются, когда говорящий снимает с себя ответственность за точность сказанного, но аппроксиматоры указывают на возможную неточность пропозиционального содержания, а «щиты» указывают на возможную неточность используемой номинации применительно к фиксированному пропозициональному содержанию.

Дискурсивный маркер СКАЖЕМ1, т. е. при его изолированном использовании, обычно функционирует как аппроксиматор, или, более узко — как репрезентатив. Проиллюстрируем это примером (3):

- (3) ...большинство римских церквей закрывается на сиесту с полудня до четырех (и это еще по-божески, потому что некоторые из них работают, скажем, только в последнюю субботу месяца или по воскресеньям с 10 до 11...) [Ольга Гринкруг. Рим (2012)].

Здесь СКАЖЕМ1 вводит пример, иллюстрирующий общее утверждение, что церкви работают далеко не каждый день, причем в очень ограниченные часы. Это общее утверждение не формулируется, а репрезентируется примерами *в последнюю субботу месяца и по воскресеньям с 10 до 11*. Общее значение маркера СКАЖЕМ1 удачно сформулировано в базе данных «Русский конструктор» [Russian Constructicon 2021]: «Конструкция (со СКАЖЕМ1. — В. П.) вводит пояснение некоторого общего утверждения с помощью конкретного частного примера. В качестве примера, как правило, описывается целая ситуация или называется ее элемент. Конструкция часто вводит пример как гипотетически возможный».

Дискурсивный маркер СКАЖЕМ2, т. е. в сочетании с наречием, используется иначе — не как аппроксиматор, а как «щит». Его функция — аргументировать предпочтительный выбор формулировки, по возможности, предъявляя основание, на котором строится это предпочтение. При этом в тексте может эксплицитно присутствовать альтернативная формулировка, как в (4) и (5).

- (4) Мне иногда кажется, что своим искусством... ну ладно, скажем скромнее: ремеслом своим я обязана бабушкам [Дина Рубина. Медная шкатулка (2011–2015)].
- (5) В этом и есть ваше вражеское нутро, значит, вы должны быть уничтожены, или, скажем мягче — мы ведь гуманисты, единственные подлинные гуманисты! — изолированы! [Ю. О. Домбровский. Факультет ненужных вещей, часть 2 (1978)].

В (4) номинация *ремесло* предлагается как более предпочтительная, чем номинация *искусство* — будучи более *скромной*. В (5) номинация *изолированы* предлагается как более предпочтительная, чем номинация *уничтожены* — будучи более *мягкой*.

Альтернативная номинация/формулировка может и не присутствовать в тексте: говорящий просто декларирует, на каком основании сделал свой выбор из возможных — не названных — альтернатив. Так, в (6) номинация *глуповатое* выбрана как *честная* (хотя и не очень лестная для обладателя лица). В (7) формулировка *нам всем там будет нечего делать* выбрана как *откровенная* (хотя обсуждаемый исход и не самый благоприятный для участников ситуации):

- (6) Его доброе, но, скажем честно, глуповатое лицо выражало смущение [Сергей Лукьяненко. Новогодний Дозор (2014)].
- (7) Потому что если квиддич станет настоящим спортом, то, скажем откровенно, нам всем там будет нечего делать [Соня Беккина. Как сказку сделать былью // «Русский репортер», 2012].

И СКАЖЕМ1, и СКАЖЕМ2 демонстрируют существенный сдвиг грамматического значения: форма презенса совершенного вида уже не обозначает будущего времени — как ей положено в полнозначном употреблении. Но при этом СКАЖЕМ2 обладает гораздо большими грамматическими свободами, чем СКАЖЕМ1. СКАЖЕМ2 допускает не только форму множественного числа 1 лица, но и форму единственного: *скажем/скажу честно, скажем/скажу откровенно* и т. д. СКАЖЕМ1 такой вариативности не допускает: *пусть он придет скажем** *скажу в восемь*. Наречие может располагаться как слева, так и справа от СКАЖЕМ: *честно скажем / скажем честно*. В отличие от СКАЖЕМ1, СКАЖЕМ2 может быть обременен частицей, как в (8), и даже дееспричастным оборотом, как в (9):

- (8) Согласно новому критерию, семья считается бедной (скажем уж прямо, нищей), если семь из десяти — очевидных! — благ, которыми должен пользоваться человек, ей недоступны [Александр Волков. На полпути от Миллениума // «Знание — сила», 2011].
- (9) А недурственно бы (скажем прямо, не кривя неизвестными направлениями) обладать вот какой красивой и полезной весяма для страны и в быту должностью, которая мгновенно бы оправдала мое труднооправдываемое существование: уполномоченный по особым поручениям неизвестных направлений [Валерий Володин. Повесть временных лет // «Волга», 2011].

Более того, в единичных случаях обнаруживаются употребления СКАЖЕМ2 с комплементаризером:

- (10) Скажем прямо, что особого труда и ума для этого ныне не требуется [Мы все зависимы друг от друга (2002) // «Жизнь национальностей», 05.06.2002].
- (11) Скажем так, что в лучшем случае социализм верующих в него масс представляет неясный идеал лучшей жизни и большей социальной справедливости [Б. Г. Бажанов. Воспоминания бывшего секретаря Сталина (1980)].

Такие употребления можно рассматривать как «реликтовые», в которых маркер сохраняет синтактику глагола, к которому восходит. Конструкции такого рода действительно редки: *скажем так что* представлен в основном корпусе НКРЯ двумя примерами, *скажем прямо что* — пятью.

Наиболее часто со СКАЖЕМ2 сочетаются наречие *прямо* и местоименное наречие *так*. По запросу «словоформа *скажем* плюс наречие или местоименное наречие на расстоянии от -1 до +1» выдан 4021 пример, верхняя часть таблицы частотности по основному корпусу НКРЯ выглядит следующим образом:

Таблица 1

Частотность сочетаний СКАЖЕМ с наречием по данным основного корпуса НКРЯ

Слово 1	Словоформа	← [1..2] →	Слово 2	Лемма	Вхождения	Доля	Доля	ipm
скажем		1	так		942	22,84 %	[21,59 %, 24,15 %]	2,52
скажем		-1	прямо		654	15,86 %	[14,78 %, 17 %]	1,75
скажем		1	прямо		192	4,66 %	[4,05 %, 5,34 %]	0,51
скажем		-1	так		143	3,47 %	[2,95 %, 4,07 %]	0,38

Как видим, почти половина выдачи — это сочетания СКАЖЕМ с местоименным наречием *так* (справа от СКАЖЕМ 22,84 % случаев, слева от СКАЖЕМ — 3,47 %) и наречием *прямо* (слева от СКАЖЕМ — 15,86 %, справа от СКАЖЕМ — 4,66 % случаев). Абсолютными чемпионами являются, таким образом, *скажем так* и *прямо скажем*. Вариант с *так* — это дистиллированный хеджинг: говорящий предъявляет некоторую формулировку как предпочтительную, оговариваясь, что она одна из возможных, не уточняя, на каком основании был сделан выбор. В частности, *скажем так* широко применяется, когда более точная номинация оказывается неуместной по каким-то соображениям — этикетным или, например, цензурным. Так, в (12) формулировка *несколько другой комплекции* употреблена вместо *полная/толстая*, а в (13) — *космополит* вместо *еврей*:

- (12) И вот в том месте дуэта, где вступает Миловзор: «Я здесь, но скучен, томен, смотри, как похудал...», я решила обыграть контрастность наших фигур: Виктория — худенькая, тоненькая, как тростиночка, а я, скажем так, несколько другой комплекции [И. К. Архипова. Музыка жизни (1996)].
- (13) И если во время войны приходилось убедительно настаивать на еврейской половине, то в конце сороковых, когда били, скажем так, космополитов, Кио немало сил приложил, чтобы считаться немцем [И. Э. Кио. Иллюзии без иллюзий (1995–1999)].

По данным основного корпуса НКРЯ, СКАЖЕМ1 — многократно более частотный маркер, чем СКАЖЕМ2. Мы сравнили выдачу по следующим трем запросам: (1) словоформа СКАЖЕМ в запятых с двух сторон, (2) словоформа СКАЖЕМ после запятой плюс (местоименное) наречие перед запятой и (3) (местоименное) наречие после запятой плюс словоформа СКАЖЕМ перед запятой, см. табл. 2. Мы намеренно ограничились случаями «в запятых» — это позволило отсеять значительную часть ложной выдачи, связанной с употреблением СКАЖЕМ в качестве формы полнозначного глагола.

Таблица 2

**Частотность пунктуационно выделенных СКАЖЕМ1 и СКАЖЕМ2
по данным основного корпуса НКРЯ**

Пунктуационно выделенное СКАЖЕМ		Запрос	Число вхождений	В том числе:
СКАЖЕМ1 в запятых	—	{ <i>скажем</i> , <i>всomma</i> & <i>асomma</i> }	10 400	—
СКАЖЕМ2 в запятых	СКАЖЕМ плюс (местоименное) наречие в запятых	{ <i>скажем</i> , <i>асomma</i> ; ADV ADVPRO, <i>всomma</i> , на расстоянии от 1 до 1 от слова 1}	800	522 случая <i>скажем так</i>
	(Местоименное) наречие плюс СКАЖЕМ в запятых	{ADV ADVPRO, <i>асomma</i> ; <i>скажем</i> , <i>всomma</i> , на расстоянии от 1 до 1 от слова 1}	690	520 случаев <i>прямо скажем</i>

Как показывает табл. 2, на запрос {*скажем*, *всomma* & *асomma*} выдано 10 400 примеров; на запрос {*скажем*, *асomma*; ADV | ADVPRO, *всomma*, на расстоянии от 1 до 1 от слова 1}, выдано 800 примеров, из которых 522 примера на *скажем так* в запятых; на запрос {ADV | ADVPRO, *асomma*; *скажем*, *всomma*, на расстоянии от 1 до 1 от слова 1} выдано 690 примеров, из которых 520 на *прямо скажем* в запятых. Таким образом, с учетом ограничения «в запятых» СКАЖЕМ1 дает выдачу 10 400 случаев, СКАЖЕМ2 дает выдачу 800+690 = 1490 случаев на одном и том же объеме основного корпуса (374 449 975 слов), т. е. СКАЖЕМ1 встречается примерно в семь раз чаще, чем СКАЖЕМ2².

² Возможно, значительное преобладание СКАЖЕМ1 по сравнению со СКАЖЕМ2 частично связано с преобладанием в основном корпусе НКРЯ текстов с низким уровнем спонтанности —

Учитывая выше сказанное, было бы заманчиво рассматривать СКАЖЕМ1 и СКАЖЕМ2 как элементы единой эволюционной цепочки от полнозначного глагола к максимально прагматикализованному дискурсивному маркеру, однако, по-видимому, в истории языка эти варианты не были строго упорядочены во времени. Ниже приводятся две визуализации, предлагаемые НКРЯ, которые позволяют сравнить изменения частотности *скажем* и *скажем так*, выделенных запятыми, по годам, см. рис. 1 и рис. 2.



Рис. 1. Запрос {скажем, всомта & асомта}. Распределение по годам (частота на миллион словоформ) в основном корпусе с 1766 по 2021 со сглаживанием

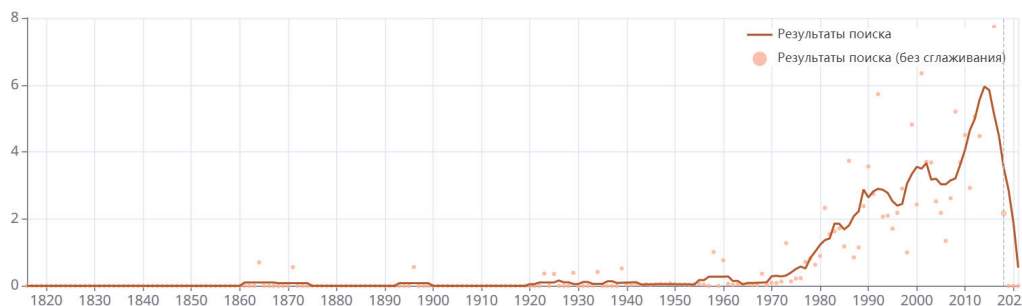


Рис. 2. Запрос {скажем, асомта так, всомта, на расстоянии от 1 до 1}. Распределение по годам (частота на миллион словоформ) в основном корпусе с 1766 по 2021 со сглаживанием

Графики показывают, что рост частотности обособленного одиночного *скажем* начинается с последней трети XIX в., а рост частотности обособленного *скажем так* начинается гораздо позже — в 1960–1970 гг. И пик частотности у *скажем так* тоже сдвинут вправо:

прежде всего, письменных. Об этом заставляют задуматься данные, опубликованные в [Богданова-Бегларян 2012]. В этой работе сравнивается частотность одиночного *скажем* (т. е. нашего СКАЖЕМ1) и *скажем* в составе сочетаний — сюда попадают и сочетания *скажем* с наречиями (т. е. нашего СКАЖЕМ2), и сочетания *скажем* с другими дискурсивными словами (*там, вот* и т. п.). Показано, что в основном корпусе НКРЯ, научной и публичной речи в составе устного подкорпуса НКРЯ, а также в текстах извлеченных монологов «Сбалансированной аннотированной библиотеки» [Богданова-Бегларян и др. 2019] действительно преобладает одиночное *скажем*, однако в спонтанной устной речи по данным корпуса «Один звуковой день» наблюдается значительное превосходство *скажем* в составе сочетаний. В частности, одно только сочетание *скажем так* составляет 40,5 % от всех вхождений *скажем*, а одиночное *скажем* — только 27 %.

у скажем он приходится на 1990-е гг., а у скажем так — на 2014–2015 гг.³. Как кажется, это обстоятельство мешает уверенно признать СКАЖЕМ1 «младшим братом» СКАЖЕМ2. Так или иначе, обоснованные суждения о совместной и раздельной истории этих маркеров нуждаются в серьезном диахроническом обосновании, что выходит за рамки настоящей статьи. Все дальнейшее изложение ограничивается строго синхронным подходом. В следующем разделе мы рассмотрим варианты линейного расположения маркеров СКАЖЕМ1 и СКАЖЕМ2 и способы их просодической реализации.

3. Линейное расположение и просодические свойства дискурсивного маркера СКАЖЕМ

Вопрос о просодии дискурсивных маркеров поднимается в литературе, как правило, в контексте исследований о парентезе [Dehé 2007]. Однако среди этих исследований крайне мало таких, которые основывались бы на систематическом инструментальном анализе данных естественного дискурса, особенно корпусных. Чаще, признавая важность просодической информации, исследователи ограничиваются данными о пунктуации как о косвенных письменных свидетельствах о возможной просодической реализации [Heine et al. 2021]⁴. Среди редких образцовых исключений назовем работы Гюнтера Кальтенбёка [Kaltenböck 2008; 2009], который описал связь сферы действия и просодии дискурсивных маркеров, восходящих к клаузам, типа *I think* и *it seems* в английском языке. Но на русском материале таких исследований пока нет, и мы попытаемся частично восполнить этот пробел.

Мы исследуем связь линейного расположения маркера и следующих просодических параметров:

- i. имеется ли просодический шов (слева от маркера, справа от маркера, с обеих сторон, ни с одной стороны);
- ii. произносится ли маркер с коммуникативно значимым акцентом или атонически;
- iii. если коммуникативно значимый акцент есть, то маркирует ли движение тона в акцентируемом сегменте просодическую завершенность (тот или иной вариант нисходящего тона) или просодическую незавершенность (тот или иной вариант восходящего тона).

Мы исследуем просодию как инструментально — с использованием компьютерного анализатора речи PRAAT [Boersma, Weenink 2021], так и перцептивно — опираясь на собственные суждения слушающего — носителя языка⁵.

³ Интересно, что оба графика демонстрируют резкое падение частотности обоих маркеров в последнюю пятилетку. У нас пока нет хорошего объяснения этому явлению, кроме общих соображений о сменяемости «моды» на дискурсивные маркеры.

⁴ Обращение к пунктуации само по себе является довольно информативным приемом (ср., например, использование сведений о пунктуационном выделении как свидетельстве интонационного — и далее, коммуникативного — выделения некоторых типов обстоятельств времени в истории русского языка в [Сай 2010]). Как видно из нашего изложения, мы тоже прибегаем к этому методу работы с корпусными массивами. Однако это не может быть единственным источником информации о просодической сегментации конструкции.

⁵ Вслед за [Kaltenböck 2009] и иными авторитетными авторами, к которым он апеллирует [Wichmann 2001; Peters 2006], мы не склонны ни приуменьшать роль перцептивного («импрессионистического», основанного на интроспекции исследователя) анализа просодии, ни, тем более, отказываться от него, поскольку в реальной коммуникации и адресат воспринимает поступающую на вход звуковую цепочку, прежде всего, «на слух». Компьютерные инструменты играют при этом важную роль, позволяя частично объективизировать перцептивное впечатление.

Ниже мы последовательно рассмотрим связь линейного расположения и просодических свойств дискурсивного маркера СКАЖЕМ в трех позициях: слева от опорного сегмента, справа от опорного сегмента и внутри опорного сегмента⁶. Прототипически опорным сегментом для дискурсивного маркера является клауза, однако это могут быть и мультиклаузальные последовательности (*Не можешь — научим, не хочешь — заставим, скажем так*), или наоборот фрагменты клаузы, например: парцелированные (*Купи немного сыра — скажем, грамм двести*).

Предварительно по данным основного корпуса НКРЯ попробуем оценить, как количественно распределяются эти три позиции. Обратимся вновь к метрикам, связанным с пунктуацией⁷. Данные сведены в таблицы 3а и 3б и прокомментированы ниже.

Таблица 3а

Частотность СКАЖЕМ1 и СКАЖЕМ2 в начале и в середине опорного сегмента по данным основного корпуса НКРЯ

Пунктуационно выделенное СКАЖЕМ		Запрос	Число вхождений	Доля числа случаев в начале опорного сегмента от числа случаев в середине опорного сегмента
СКАЖЕМ1	СКАЖЕМ в запятых	{скажем, bcomma & acomma}	10 400	—
	СКАЖЕМ с заглавной буквы перед запятой	{скажем, bcomma & capital}	2535	24,4 %
СКАЖЕМ2	СКАЖЕМ плюс (местоименное) наречие в запятых	{скажем, acomma; ADV ADVPRO, bcomma, на расстоянии от 1 до 1 от слова 1}	800	—
	СКАЖЕМ с заглавной буквы плюс (местоименное) наречие перед запятой	{скажем, capital; ADV ADVPRO, bcomma, расстоянии от 1 до 1 от слова 1}	158	19,8 %
	(Местоименное) наречие плюс СКАЖЕМ в запятых	{ADV ADVPRO, acomma; скажем, bcomma, на расстоянии от 1 до 1 от слова 1}	690	—
	(Местоименное) наречие с заглавной буквы плюс СКАЖЕМ перед запятой	{ADV ADVPRO, capital; скажем, bcomma, на расстоянии от 1 до 1 от слова 1}	203	29,4 %

⁶ В качестве английского эквивалента «опорного сегмента» в [Kaltenböck 2008; 2009] используется термин *host* ‘хозяин’, однако нам представляется, что этот термин имеет слишком сильные синтаксические коннотации, тогда как исследуемые нами единицы связаны с тем, что мы называем «опорным сегментом», дискурсивно — семантически и просодически, но не интегрированы в его синтаксическую структуру.

⁷ Поскольку в последующих разделах мы будем работать с материалом МУРКО, следовало бы и количественные оценки давать по этому материалу, но, по понятным причинам, транскрипты в МУРКО не используют стандартную пунктуацию. Так что с метриками, основанными на стандартной пунктуации, мы работаем с данными основного корпуса, полагая, что ввиду его большого объема и сбалансированности, экстраполяция такого рода результатов может считаться достаточно надежной.

Таблица 36

**Частотность СКАЖЕМ1 и СКАЖЕМ2 в конце и в середине опорного сегмента
по данным основного корпуса НКРЯ**

	Пунктуационно выделенное СКАЖЕМ	Запрос	Число вхождений	Доля числа случаев в конце опорного сегмента от числа случаев в середине опорного сегмента
СКАЖЕМ1	СКАЖЕМ в запятых	{ <i>скажем</i> , <i>бcomma</i> & <i>acomma</i> }	10 400	—
	СКАЖЕМ после запятой перед точкой	{ <i>скажем</i> , <i>bdot</i> & <i>acomma</i> }	96	0,92 %
СКАЖЕМ2	СКАЖЕМ плюс (местоименное) наречие в запятых	{ <i>скажем</i> , <i>acomma</i> ; ADV ADVPRO, <i>бcomma</i> , на расстоянии от 1 до 1 от слова 1}	800	—
	СКАЖЕМ после запятой плюс (местоименное) наречие перед точкой	{ <i>скажем</i> , <i>acomma</i> ; ADV ADVPRO, <i>bdot</i> , на расстоянии от 1 до 1 от слова 1}	54	6,7 %
	(Местоименное) наречие плюс СКАЖЕМ в запятых	{ADV ADVPRO, <i>acomma</i> ; <i>скажем</i> , <i>бcomma</i> , на расстоянии от 1 до 1 от слова 1}	690	—
	(Местоименное) наречие после запятой плюс СКАЖЕМ перед точкой	ADV ADVPRO, <i>acomma</i> ; <i>скажем</i> , <i>bdot</i> , на расстоянии от 1 до 1 от слова 1}	20	2,9 %

О чем говорят данные этих таблиц? Число случаев СКАЖЕМ в запятых примерно дает представление о его частотности внутри опорного сегмента. Как было сказано выше, поиск СКАЖЕМ1 в запятых по запросу {*скажем*, *бcomma* & *acomma*} дает 10 400 примеров. Проверим теперь запрос {*скажем*, *бcomma* & *capital*}, т. е. СКАЖЕМ1 с заглавной буквы перед запятой: выдача — 2535 примеров, что составляет 24,4 % от СКАЖЕМ1 в запятых. Примерно ту же пропорцию демонстрируют и два варианта СКАЖЕМ2: поиск последовательности «(местоименное) наречие + СКАЖЕМ2» по запросу {ADV | ADVPRO, *acomma*; *скажем*, *бcomma*, на расстоянии от 1 до 1 от слова 1} в запятых дает, как было сказано выше, 690 примеров, а та же последовательность перед запятой, но с заглавной буквы по запросу {ADV | ADVPRO, *capital*; *скажем*, *бcomma*, на расстоянии от 1 до 1 от слова 1} дает 203 примера (доля 29,4 %). Аналогичным образом, поиск последовательности «СКАЖЕМ2 + (местоименное) наречие» в запятых дает 800 примеров, а та же последовательность, но с заглавной буквы и перед запятой по запросу {*скажем*, *capital*; ADV | ADVPRO, *бcomma*, на расстоянии от 1 до 1 от слова 1} дает 158 примеров (доля 19,8 %). Таким образом, маркеры СКАЖЕМ1 и СКАЖЕМ2 появляются в начальной позиции (слева от опорного сегмента) в три-пять раз реже, чем внутри опорного сегмента.

Посмотрим теперь на частотность заключительной позиции маркера. Примерное представление может дать поиск интересующих нас последовательностей после запятой перед точкой. По запросу {*скажем*, *bdot* & *acomma*} выдается 96 примеров, по запросу {*скажем*, *acomma*; ADV | ADVPRO, *bdot*, на расстоянии от 1 до 1 от слова 1} выдается

54 примера; по запросу {ADV | ADVPRO, аcomma; скажем, bdot, на расстоянии от 1 до 1 от слова 1} — 20 примеров, Таким образом, число случаев СКАЖЕМ1 после запятой перед точкой составляет 0,92 % от числа случаев СКАЖЕМ1 в запятых (96:10400); число случаев последовательности «СКАЖЕМ2 + (местоименное) наречие» после запятой перед точкой составляет 6,7 % от случаев той же последовательности в запятых (54:800); число случаев последовательности «(местоименное) наречие + СКАЖЕМ2» после запятой перед точкой составляет 2,9 % от случаев той же последовательности в запятых (20:690). Т. е. мы видим, что маркеры СКАЖЕМ1 и СКАЖЕМ2 появляются в заключительной позиции (справа от опорного сегмента) многократно (более чем в 10 раз!) реже, чем внутри опорного сегмента.

Итак, метрики, связанные с пунктуацией, позволяют — пусть и не прямым образом — продемонстрировать, что для дискурсивного маркера СКАЖЕМ прототипической является позиция внутри опорного сегмента, а начальная и, особенно, заключительная позиции гораздо более редки, хотя и допустимы. При этом, как мы увидим из анализа данных МУРКО, в каждой из трех позиций у СКАЖЕМ имеются свои особенности семантики, синтактики и просодии.

3.1. СКАЖЕМ слева от опорного сегмента

Рассмотрим сначала поведение СКАЖЕМ1 в левой периферии опорного сегмента.

Дефолтным (но как мы увидим дальше — не единственным!) паттерном в таких случаях является полная просодическая интеграция маркера с правым контекстом: маркер произносится без акцента (по интонограмме может быть заметно слабое падение тона, в более узком, чем окружающий контекст тональном диапазоне), между маркером и правым контекстом нет просодического шва. В примере (14) нет шва между *скажем* и *доски*, в (15) — между *скажем* и *нанайцы*⁸:

- (14) Представляете/ в рыбинспекции нет даже приличной мебели. Скажем/ доски для забора/ уже проблема [Александр Серый, Григорий Горин. Ты — мне, я — тебе, к/ф (1976)].

/Представляете,

В /рыбинспекции нет даже приличной \мебели!

(0.20)

Скажем /доски для забора (ц 0.37) \уже проблема (конец строки взволнованным шепотом. — В. П.).

⁸ Примеры приводятся в полной графической форме, как она дана в МУРКО, при необходимости приводится также просодическая транскрипция части примера с разметкой движения тонов и локализацией фразовых акцентов и интонограмма в формате анализатора PRAAT. Для указания на направление движения тона иконически используются знаки «/», «\» и «-». Ударный слог слова — носителя рематического акцента подчеркивается. Длина пауз в секундах указывается в круглых скобках; пауза может быть абсолютной или вокализованной: *ц*, *?*, *э* и др. Внутри фигурных скобок приводятся обозначения неречевых вокальных действий: *gp* — глотывание, *hm* — хмыканье и др. О других деталях используемой системы просодической транскрипции см. [Кибрик, Подлесская (ред.) 2009]. Напомним, что в той версии примера, которая дается по МУРКО, знак «/» имеет другую интерпретацию — там он используется для членения речевого потока. В сегментное наполнение в просодической транскрипции вносятся при необходимости уточнения по сравнению с графическим вариантом МУРКО.

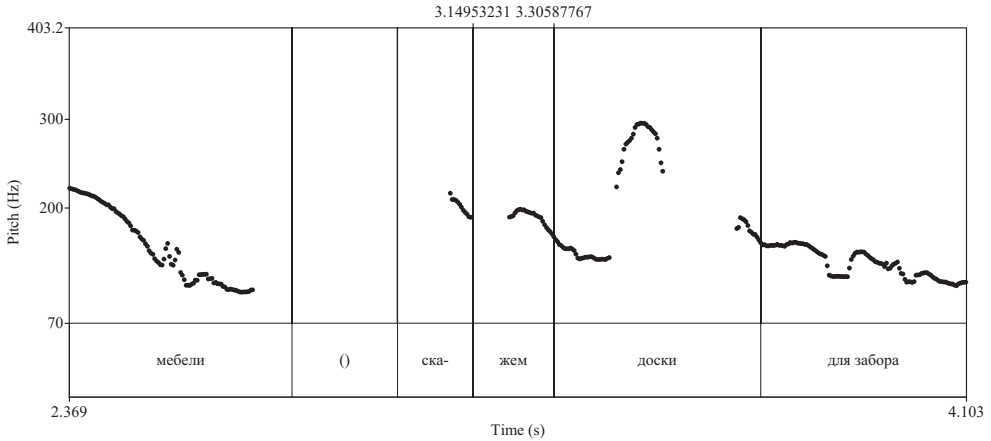


Рис. 3. Интонограмма к примеру (14)

- (15) В те времена это был пиджин/ потому что — и что очень важно — местное... местное население между собой контактировало на нём. Скажем/ нанайцы/ там/ предположим/ не знаю/ с удэгейцами или с кем-то ещё [Евгений Головкин. Как рождаются языки. Лекции Полит.ру (2011)].

В /те времена это был \пиджин.

Потому что —

(ц 0.36)

и ==

(? 0.16) что очень /важно,

(0.49)

— местное|| местное население между \собой контактировало на нём.

Скажем /\нана-айцы там-м,,,

(0.61)

предположим не знаю с \удэгейцами,

или с кем-то /ещё,

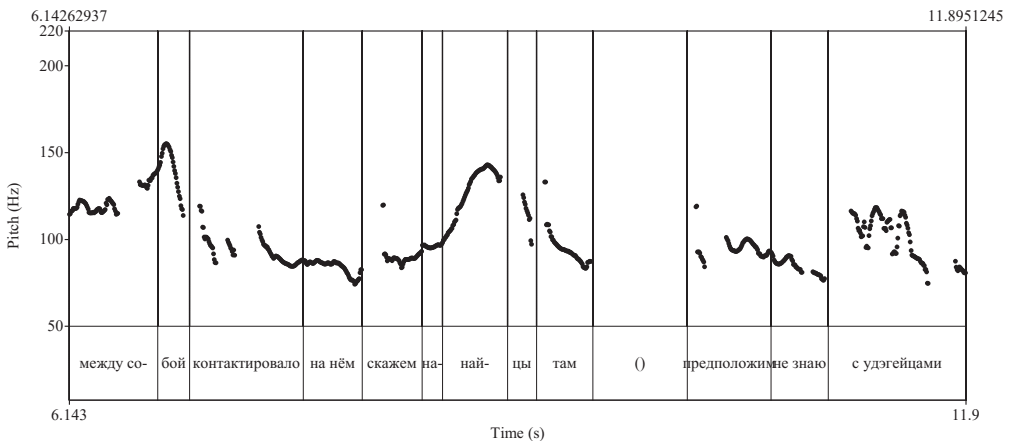


Рис. 4. Интонограмма к примеру (15)

Просодической интеграции не препятствует наличие паузы между маркером и правым контекстом, решающим остается отсутствие коммуникативно значимого движения тона на маркере, см. паузу между атоническим *скажем* и *половина* в (16):

- (16) Но и те звёзды/ которые кажутся нам одиночными/ в действительности/ как правило/ живут коллективами. Скажем/ половина всех звёзд при их детальном изучении в телескоп оказываются двойными [Жизнь звездных систем. Программа «Гордон» (НТВ) (2003)].

и \те звёзды которые кажутся нам /одиночными,

(0.49)

в /действительности как /—правило живут \коллективами.

(ц 0.42)

Скажем (0.58) \половина всех /звёзд —

(0.23)

(при их детальном \изучении),

— в /телескоп оказываются \двойными.

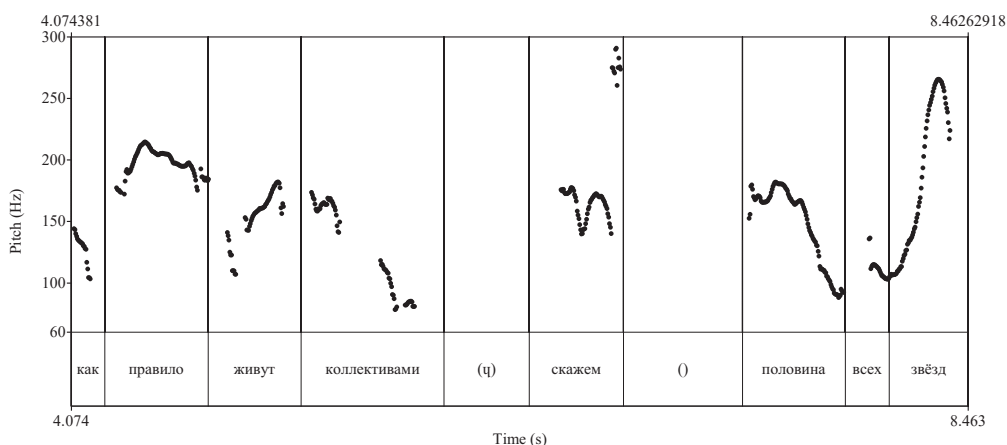


Рис. 5. Интонограмма к примеру (16)

Употребление начального СКАЖЕМ1 с коммуникативно значимым акцентом встречается гораздо реже атонического, но оно возможно. Для такого автономного употребления характерен нисходящий тон, т. е. маркер реализуется как просодически завершенная единица. Это, по существу, тот же паттерн, что и у авторской ремарки, предшествующей прямой речи — такая ремарка реализуется как просодически завершенная, а лексико-грамматически, наоборот, как проецирующая ожидаемое продолжение. В стандартной ситуации такому паттерну соответствует закрывающее двоеточие. По аналогии в соответствующей позиции мы используем двоеточие после СКАЖЕМ и в транскрипте, см. (17):

- (17) Скажем/ наша локальная группа содержит тридцать галактик/ туда входит Млечный Путь/ Андромеда/ Магеллановы Облака — то/ что я вам показывал/ но в некоторых кластерах более тысячи галактик/ а в некоторых супер-кластерах порядка сотни кластеров [Михаил Маров. Солнечная система. Проект Academia (ГТРК Культура) (2010)].

\Скажем:

(0.20)

наша локальная \↑группа (0.20) содержит \тридцать \→галактик,

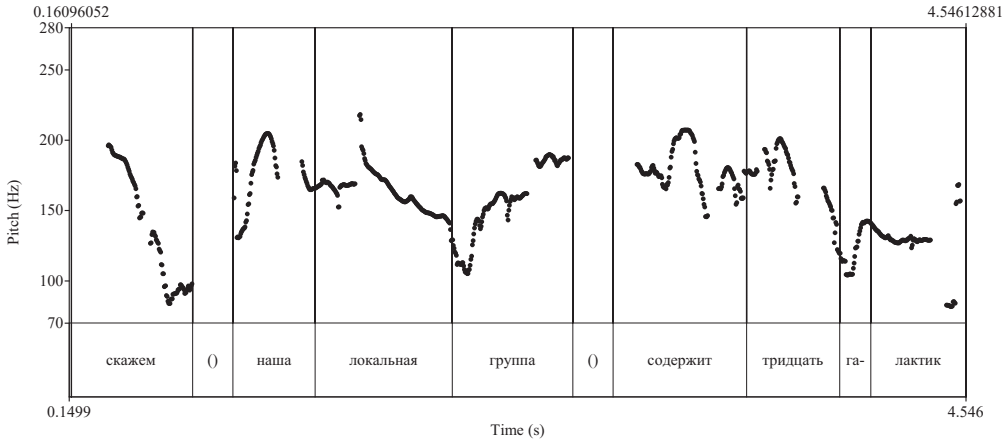


Рис. 6. Интонограмма к примеру (17)

В редких случаях обнаруживается употребление начального автономного ударного СКАЖЕМ1 с подъемом тона. В (18) акцентированное СКАЖЕМ реализуется с высоким быстрым подъемом на ударном слоге и резким падением на заударных, т. е. по типу ИК-3 в терминах интонационных конструкций Е. А. Брызгуновой [1980], а в (19) — с подъемом на ударном слоге с последующим продолжением подъема на заударном, демонстрируя один из вариантов ИК-4 (см. об этой просодической конфигурации в [Янко 2008]):

- (18) Скажем/ материнская агрессия. То есть это/ совершенно/ как бы/ святое дело/ потому что взаимодействие «мать/ дитя»/ это всё прекрасно и здорово [Вячеслав Дубынин. Мозг и агрессия (2017)].

(э 0.76) /Скажем,

\материнская агрессия.

(0.46)

То есть это-о как бы \святое /дело,

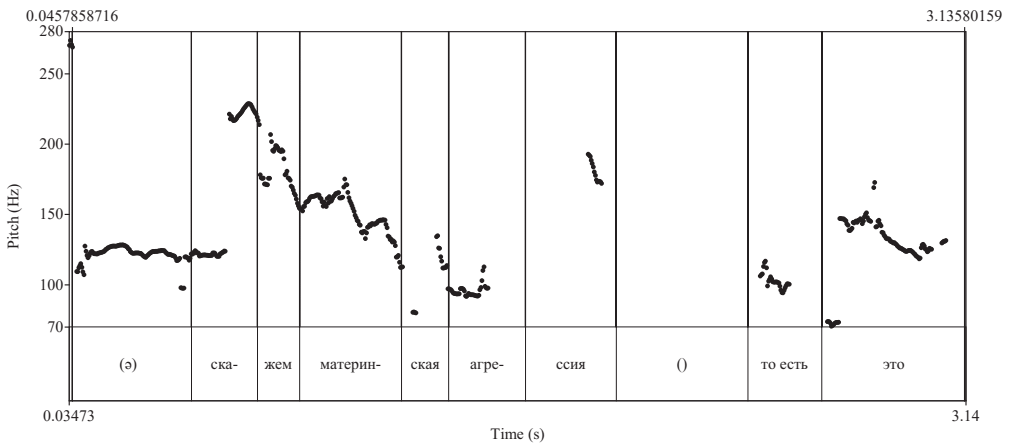


Рис. 7. Интонограмма к примеру (18)

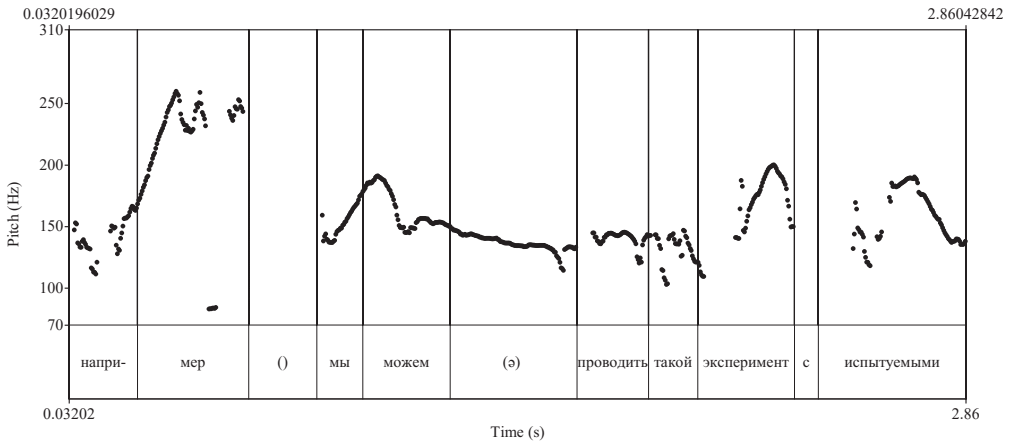


Рис. 9. Интонограмма к примеру (20)

начального СКАЖЕМ1: просодически интегрированное с правым контекстом атоническое произнесение (без просодического шва между маркером и опорным сегментом), просодически автономное акцентированное произнесение с падающим тоном, аналогичное препозитивной авторской ремарке, и просодически автономное акцентированное произнесение с восходящим тоном, аналогичное начальной адвербиальной группе. Однако — в отличие от СКАЖЕМ1 (!) — дефолтным для СКАЖЕМ2 является не атоническое употребление, а просодически автономное употребление с падающим тоном. Приведем в качестве иллюстрации пример (21) (напомним, что в транскрипте мы по аналогии с авторскими словами, предшествующими цитате, используем двоеточие в качестве закрывающего знака).

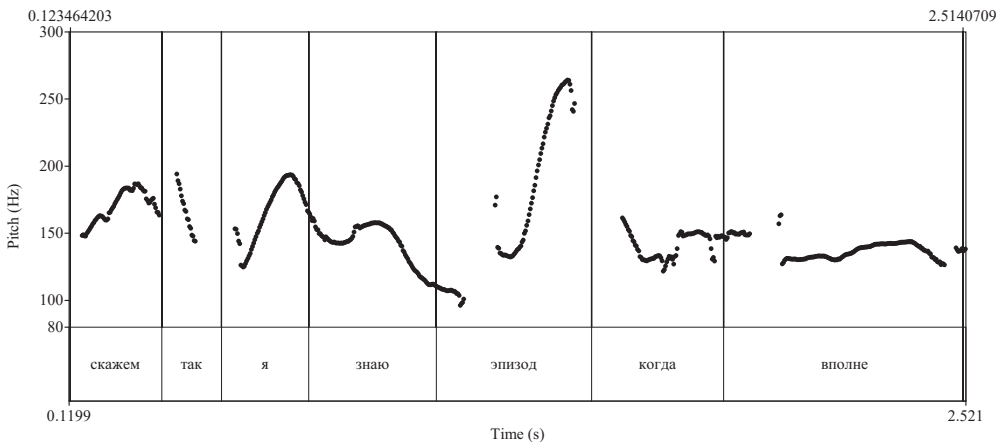


Рис. 10. Интонограмма к примеру (21)

- (21) Скажем так/ я знаю эпизод/ когда вполне грамотная женщина/ правда ботаник/ а не зоолог/ действительно попала в больницу/ сильно поеденная стаей собак в лесопарке [Андрей Тупикин. Бездомные собаки глазами биолога (2015)].

Скажем \так:

я \знаю /эпизод,

когда вполне-е (0.76) (э 0.84) грамотная /женщина —

правда /–ботаник,,,
 а не \зоолог,
 — (э 0.63) действительно-о (ʔ 0.71) (0.78) \попала в /больницу,
 /сильно поеденная стаей собак в \лесопарке.

Допустимый, но сравнительно более редкий вариант — автономное произнесение с восходящим тоном:

- (22) Да славится во веки веков имя его и имя его счастливых родителей. Прямо скажем/маловато самокритики. Но не в этом дело. Надо вывеску сменить [Геннадий Казанский, Лазарь Лагин. Старик Хоттабыч, к/ф (1956)].

... /имя его счастливых \родителей.
 (0.77)
 /Прямо скажем,
 \маловато самокритики!

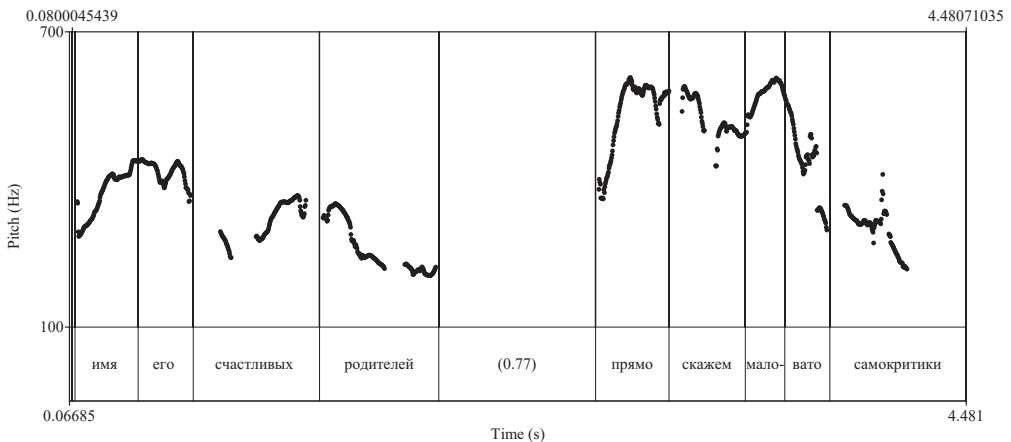


Рис. 11. Интонограмма к примеру (22)

И уж совсем редко обнаруживается атоническое просодически интегрированное произнесение. С некоторой натяжкой такой паттерн можно усмотреть в (23), где на *прям(о)* есть заметный и перцептивно, и на интонограмме (см. рисунок 12, с. 69) подъем тона, но акцентирование явно недостаточное, чтобы маркер воспринимался как просодически автономный:

- (23) Но только ход поезда удивительно медленный. Прямо даже оскорбительно ехать. Ну и/ конечно/ через такой ход в вагоне бывает ужасно как скучно. Прямо скажем/ делать нечего [М. М. Зощенко. Веселенькая история (исп. Ю. Кириллова) (1927, 2017)].

...в /вагоне бывает /ужасно как \скучно.
 (0.45)
 Прямо= скажем (0.18) \делать нечего!

Итак, по данным МУРКО, в начальной позиции для СКАЖЕМ1 прототипическим является атоническое употребление, интегрированное с опорным сегментом, тогда как для СКАЖЕМ2 прототипическим является автономное употребление с просодическим швом между маркером и опорным сегментом. Для того чтобы количественно объективировать это наблюдение, обратимся снова к пунктуационным метрикам, применимым к массиву основного корпуса НКРЯ. Косвенным симптомом автономного акцентированного употребления маркера может служить закрывающее двоеточие, см. примеры ниже.

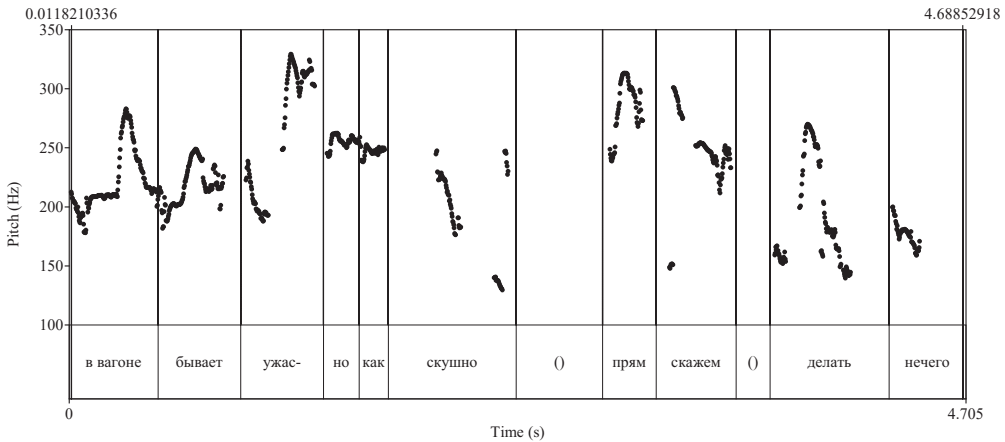


Рис. 12. Интонограмма к примеру (23)

- (24) Сразу скажем: создать сад, которому не требуется уход, невозможно! [Татьяна Ефимова. Скажи: легко! // «Сад своими руками», 2003.09.15].
- (25) Скажем проще: ей свойственно пристальное внимание к самым главным вопросам бытия [Е. В. Сафронова. Союз Евтерпы и Харона. Ольга Дернова. Человец.-- Борис Кутенков. Неразрешённые вещи // «Волга», 2015].
- (26) Скажем: надо ли дать в начале бумаги строгое определение понятия «культура» или не надо? [Александр Привалов. О благопотребном многоглаголании // «Эксперт», 2014].

Сравним частотность начального СКАЖЕМ перед запятой и перед двоеточием. Данные сведены в табл. 4 (с. 70) и далее комментируются.

Сначала для СКАЖЕМ2 посмотрим раздельно на случаи с наречием, предшествующим маркеру, как в (24), и с наречием, следующим за маркером, как в (25).

По запросу {ADV | ADVPRO, capital; скажем, bcolon, на расстоянии от 1 до 1 от слова 1} (последовательность «наречие с заглавной буквы + СКАЖЕМ2 перед двоеточием») выдается 17 примеров, а по запросу {ADV | ADVPRO, capital; скажем, bcomma, на расстоянии от 1 до 1 от слова 1} (та же последовательность, но перед запятой) выдается 203 примера. Таким образом, для наречия, предшествующего маркеру, полученное соотношение 17:203, т. е. случаи с двоеточием составляют 8,4 % от числа случаев с запятой.

По запросу {скажем, capital; ADV | ADVPRO, bcolon, на расстоянии от 1 до 1 от слова 1} (последовательность «СКАЖЕМ2 с заглавной буквы + наречие перед двоеточием») выдается 187 примеров, а по запросу {скажем, capital; ADV | ADVPRO, bcomma, на расстоянии от 1 до 1 от слова 1} (та же последовательность, но перед запятой) выдается 158 примеров. Таким образом, для наречия, следующего за маркером, полученное соотношение 187:158, т. е. случаи с двоеточием составляют 118,4 % от числа случаев с запятой.

Посмотрим теперь, как распределяются случаи перед двоеточием и перед запятой для начального СКАЖЕМ1. По запросу {скажем, bcolon & capital} (СКАЖЕМ1 с заглавной буквы перед двоеточием) выдается 48 примеров, по запросу {скажем, bcomma & capital} (СКАЖЕМ1 с заглавной буквы перед запятой) выдается 2535 примеров. Таким образом, полученное соотношение 48:2535, т. е. случаи с двоеточием составляют 1,9 % от числа случаев с запятой.

Как видим, результаты весьма красноречиво свидетельствуют о том, что для начального СКАЖЕМ1 автономное употребление (косвенным свидетельством которого является закрывающее двоеточие) совершенно не характерно. Полной противоположностью оказывается начальное СКАЖЕМ2 с наречием в постпозиции (*скажем прямо / честно / так*),

Таблица 4

Частотность СКАЖЕМ1 и СКАЖЕМ2 перед запятой и перед двоеточием по данным основного корпуса НКРЯ

	Пунктуационно выделенное СКАЖЕМ	Запрос	Число вхождений	Доля числа случаев с двоеточием от числа случаев с запятой
СКАЖЕМ1	СКАЖЕМ с заглавной буквы перед запятой	{ <i>скажем</i> , bcomma & capital}	2535	—
	СКАЖЕМ с заглавной буквы перед двоеточием	{ <i>скажем</i> , bcolon & capital}	48	1,9 %
СКАЖЕМ2	СКАЖЕМ с заглавной буквы плюс (местоименное) наречие перед запятой	{ <i>скажем</i> , capital; ADV ADVPRO, bcomma, на расстоянии от 1 до 1 от слова 1}	158	—
	СКАЖЕМ с заглавной буквы плюс (местоименное) наречие перед двоеточием	{ <i>скажем</i> , capital; ADV ADVPRO, bcolon, на расстоянии от 1 до 1 от слова 1}	187	118,4 %
	(Местоименное) наречие с заглавной буквы плюс СКАЖЕМ перед запятой	{ADV ADVPRO, capital; <i>скажем</i> , bcomma, на расстоянии от 1 до 1 от слова 1}	203	—
	(Местоименное) наречие с заглавной буквы плюс СКАЖЕМ перед двоеточием	{ADV ADVPRO, capital; <i>скажем</i> , bcolon, на расстоянии от 1 до 1 от слова 1}	17	8,4 %

которое автономно выступает чаще, чем не автономно. Начальное СКАЖЕМ2 с наречием в препозиции (*прямо / честно / так скажем*) занимает в этой тройке промежуточное положение, но и у него доля автономных употреблений почти в 10 раз выше, чем у СКАЖЕМ1.

Тем самым результаты качественного просодического анализа по данным МУРКО подтверждаются данными количественного анализа по данным основного корпуса НКРЯ: начальное СКАЖЕМ1 тяготеет к интеграции с правым контекстом, а начальное СКАЖЕМ2 тяготеет к автономному употреблению.

3.2. СКАЖЕМ справа от опорного сегмента

Как мы писали выше, по косвенным данным, полученным из основного корпуса с применением пунктуационных метрик, СКАЖЕМ справа от опорного сегмента встречается значительно реже, чем в двух других линейных позициях — начальной (слева от опорного сегмента) и, особенно, срединной.

Для заключительного СКАЖЕМ имеются две просодические опции: (1) атоническое произнесение, интегрированное с левым контекстом (т. е. без просодического шва между опорным сегментом и маркером) и (2) автономное произнесение. Для СКАЖЕМ1, как и в случае с начальной позицией, более характерно атоническое употребление. Ср. пример (27), где на СКАЖЕМ1 нет акцента, нет движения тона, интонограмма его «не видит»:

- (27) [Александр Милитарев, муж, 60, 1943, научный сотрудник] Выявляем и выкидываем из списка. То есть получится список не на сто слов/ а на девяносто пять/ скажем. Потому что пять заимствований убираются просто [Древо языков. Программа «Гордон» (НТВ) (2003)].

То есть получится список не на /сто слов,
а на девяносто \пять скажем.

(0.16)

И пять \↑заимствований \убираются просто.

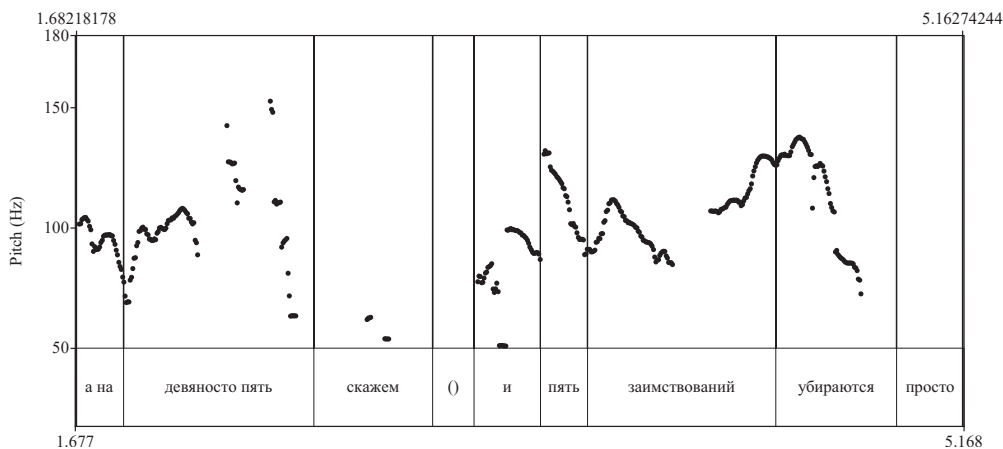


Рис. 13. Интонограмма к примеру (27)

Автономное употребление заключительного СКАЖЕМ1 встречается крайне редко; но даже и в этих случаях можно наблюдать частичные симптомы утраты просодической автономности. Рассмотрим следующий пример (см. рисунок 14, с. 72):

- (28) Ну вы знаете/ с одной стороны я готов с вами согласиться/ с другой стороны/ тоже немаленькая страна — Соединённые Штаты Америки/ скажем. Где есть две страны внутри одной/ как минимум [Экономическое пространство будущего. Программа «Гордон» (НТВ) (2003)].

с /другой стороны \тоже \немаленькая /страна Соединенные Штаты \Америки,
\скажем.

Здесь заключительное СКАЖЕМ1 реализуется с акцентом, с падающим тоном. Перед ним есть явный просодический шов — предшествующая порция дискурса реализуется как самостоятельная, просодически завершенная, с падающим тоном в рематическом акценте (акцентоноситель — словоформа *Америки*), не проецирующая продолжения. Это дает основания говорить о просодической автономности СКАЖЕМ1. Вместе с тем маркер реализуется в низком регистре в узком тоновом диапазоне — по сравнению с предшествующим фрагментом. Кроме того, как видно на интонограмме, маркер встроен в общий рисунок так называемой деклинации — постепенного понижения частоты основного тона в пределах порции дискурса, обладающей цельностью. Т. е. автономия маркера не абсолютная.

СКАЖЕМ2 в заключительной позиции несколько более склонен к автономизации, чем СКАЖЕМ1, в том смысле, что СКАЖЕМ2, как правило, акцентируется и чаще, чем СКАЖЕМ1, отделен от левого контекста паузой. Однако акцент обычно не сильный, маркер реализуется в суженном регистре, встроен в цепочку деклинации. Т. е., как и в случае со СКАЖЕМ1, полной автономии нет — таковы два следующих примера. В (29) слабо автономное

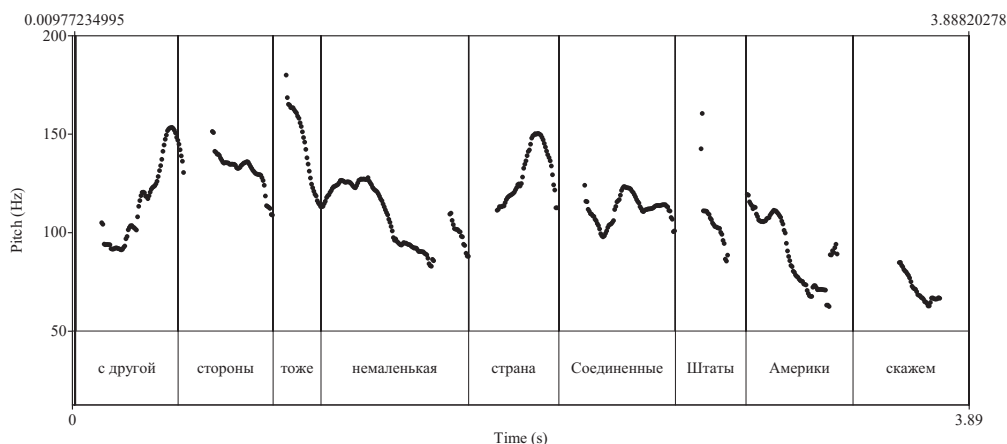


Рис. 14. Интонограмма к примеру (28)

СКАЖЕМ ТАК с нисходящим тоном следует после просодически заверщенного парцеллированного фрагмента *на улице* (рематический акцент, нисходящий тон), в (30) слабо автономное ПРЯМО СКАЖЕМ с нисходящим тоном следует после просодически заверщенного фрагмента *не ход интеллектуального человека* (рематический акцент на *ход*, нисходящий тон):

(29) Но животных/ переваливших за десять лет/ я знаю много на улице/ скажем так [Андрей Тупикин. Бездомные собаки глазами биолога (2015)].

Но-о /животных (э 0.18) переваливших за десять /лет,
я знаю \много.

(0.54)

На \улице.

(0.42)

Скажем \так.

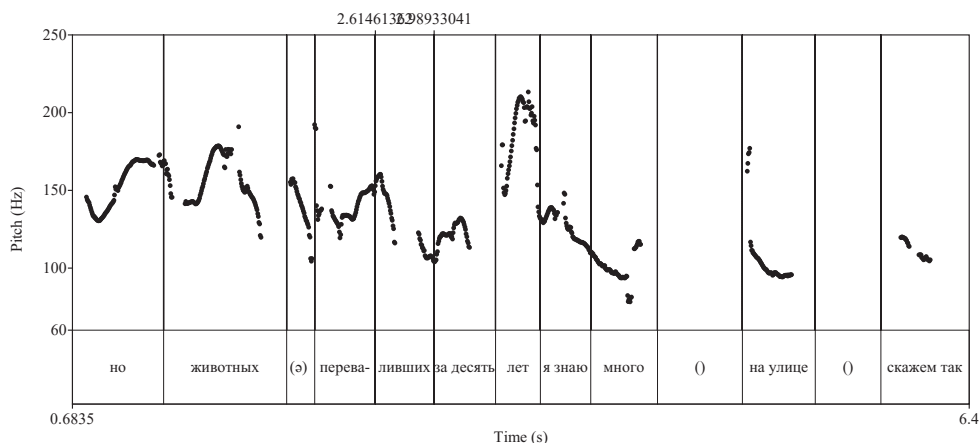


Рис. 15. Интонограмма к примеру (29)

(30) Помню/ что книжка лежала справа/ и я на неё кофе пролила. Не ход интеллектуального человека/ прямо скажем [Татьяна Черниговская. Язык и сознание: Что делает нас людьми? Лекция Полит.Ру (2008)].

/He-e \ход (0.32) интеллектуального человека,
 \прямо скажем.

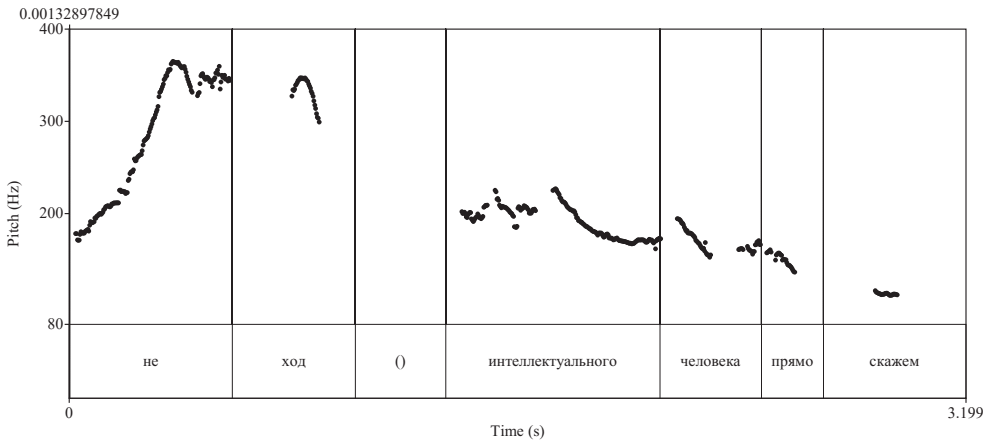


Рис. 16. Интонограмма к примеру (30)

Таким образом, и для СКАЖЕМ1, и для СКАЖЕМ2 заключительная позиция менее характерна, чем начальная и срединная, и автономизация в заключительной позиции затруднена. Это интересный результат, который согласуется с имеющимися в литературе данными о дискурсивных маркерах и иных прагматизированных единицах, выпадающих из синтаксической структуры клаузы, в русском языке и в ряде других языков. В частности, мы обнаружили ту же тенденцию при анализе конструкций типа *как говорил/писал Лев Толстой* [Подлесская, Пожилов 2021]: они редко употребляются справа от своей сферы действия, т. е. от цитаты, которую они вводят, и если все-таки употребляются, то их просодическая автономизация сильно затруднена. Сходные наблюдения о некоторых дискурсивных маркерах в английском и корейском приводятся в [Haselow 2015: 163; Heine et al. 2021: 52, 90, 97, 198]. Возможно, мы имеем дело с универсальной тенденцией, но это, разумеется, нуждается в проверке на широкой выборке языков. Посмотрим теперь, как ведет себя СКАЖЕМ внутри опорного сегмента — как мы видели выше, по данным основного корпуса НКРЯ, это наиболее частотная линейная позиция для данного маркера.

3.3. СКАЖЕМ внутри опорного сегмента

Внутри опорного сегмента СКАЖЕМ может занимать любую позицию. Он легко разрывает любые синтаксические связи, располагаясь между глаголом и его дополнением, определяемым и согласованным определением, предлогом и подчиненной предложу группой и т. д. При этом в зоне семантической сферы действия маркера может оказаться как часть опорного сегмента справа от маркера, так и часть слева от него. Иначе говоря, СКАЖЕМ выступает как оператор «хеджирования» не в отношении всего опорного сегмента, а только части его. Именно эта часть — до или после маркера — объявляется как зона нечеткой номинации. В соответствии с этим подключаются те же просодические паттерны, которые мы наблюдали для начальных и заключительных позиций маркера.

Основные просодические конфигурации, которые представлены в нашем материале, связаны с расположением маркера относительно главного релативного акцента опорного сегмента. В общем случае маркер просодически «работает» как начальный, если

основной рематический акцент опорного сегмента находится справа от маркера, и как заключительный, если основной рематический акцент находится слева.

В подавляющем большинстве случаев рематический акцент опорного сегмента расположен справа от СКАЖЕМ. Как и в случае с начальным маркером, СКАЖЕМ1 реализуется в таких условиях как атонический, а СКАЖЕМ2 — с акцентом. Так, в следующем примере контрастный рематический акцент расположен на наречии *временно* справа от СКАЖЕМ. Именно номинация *временно проживают* попадает в сферу действия маркера и подается как нечеткая номинация, как условное именование — по-видимому, от нежелания прямого упоминания трудовых мигрантов. Трудности выбора этой номинации дополнительно подчеркиваются целым кластером сигналов речевого затруднения — слглатыванием, заполненными паузами, частичней «поиска» ну (см. об этом подробнее в следующем разделе). СКАЖЕМ1 реализуется атонически:

- (31) И особенно ээ зачастую критике Москву подвергают те/ кто ээ здесь ээ/ ну скажем/ временно проживают. А вот те/ кто здесь ээ живут и работают давно и долго/ то и каждому любому позитивному шагу они его видят и радуются [Круглый стол, посвященный прямым выборам мэра Москвы (2013)].

...подвергают /те/

кто {gr 0.28} (э 0.19) здесь (0.15) (э 0.26) ну скажем /временно проживают.

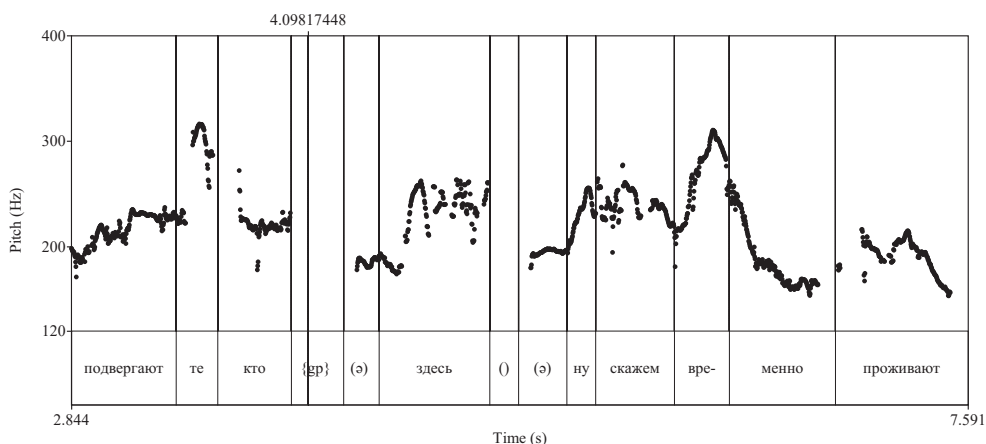


Рис. 17. Интонограмма к примеру (31)

Для СКАЖЕМ2 в интерпозиции при расположении рематического акцента опорного сегмента справа от маркера возможны те же реализации, которые мы обнаружили для начальной позиции: с нисходящим тоном (по модели препозитивной авторской ремарки) и с восходящим или слабо восходящим тоном (по модели начальной адвербиальной группы). Вариант с нисходящим тоном (по модели препозитивной авторской ремарки, ср. (21)) реализован в (32), маркер (*скажем так*) появляется между предлогом *от* и управляемой им именной группой (*угрозы продовольственной безопасности в отношении свинины*). Именно эта группа (акцентовоситель — *свинины*, справа от маркера) и попадает в сферу действия маркера и преподносится как не вполне прямая номинация — с некоторым сарказмом ввиду ее избыточной пафосности:

- (32) ...мы абсолютно аа избавились от/ скажем так/ аа угрозы национальной приро... продовольственной безопасности в отношении свинины [Конец нефтяной эпохи. Программа «Правда на ОТР» (2016)].

избавились от —

(ə 0.22) {hm} скажем \так:

(ц)

— (ə 0.59) /угрозы приро-о= || продовольственной /безопасности в отношении \сви-
нины.

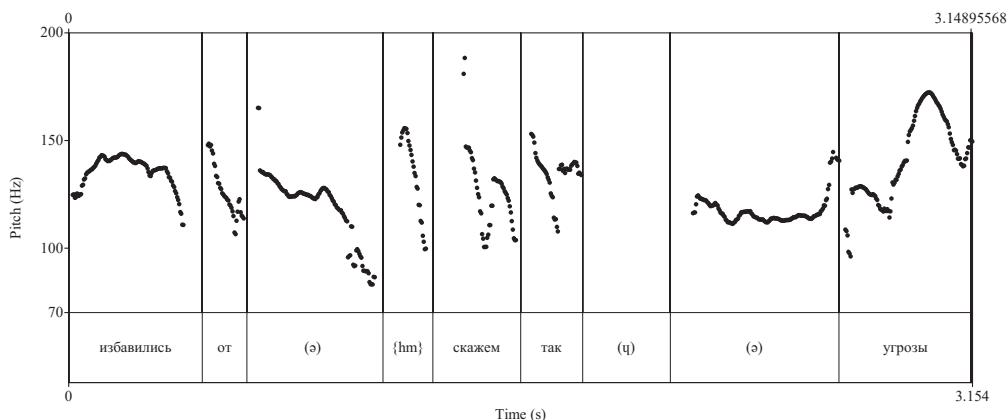


Рис. 18. Интонограмма к примеру (32)

Вариант со слабо восходящим тоном (по модели начальной адвербиальной группы, ср. (23)) реализован в (33) (см. рисунок 19, с. 76). Маркер (*прямо скажем*) появляется между определяемым (*советы*) и группой согласованного определения (*очень хорошие*). Именно эта группа (акцентоноситель — *хорошие*, справа от маркера) попадает в сферу действия маркера и подается как оценочная номинация, которая выбрана после некоторого обдумывания, что подтверждается и дополнительным маркером хезитации, так называемым плейсхолдером *такие*⁹ (напомним, что роли хезитации в сочетании со СКАЖЕМ посвящен следующий раздел):

- (33) Искусственный интеллект начал ему давать какие-то советы *такие/* прямо скажем/ очень хорошие [Илья Утехин. Как устроен диалог? Лекция Полит.Ру (2017)].

Искусственный /интеллект начал ему /давать какие-то /советы такие == —

(1.30)

/прямо скажем,

— очень \хорошие!

Конфигурации, где в сфере действия маркера оказывается часть опорного сегмента слева от СКАЖЕМ, встречаются гораздо реже. Такое становится возможным, если по каким-то причинам — обычно связанным с контрастом/эмфазой — происходит сдвиг главного акцента опорного сегмента влево (о других возможных причинах сдвига рематического акцента см. [Янко 2021: 218–221]). Таковы следующие два примера, где маркер *скажем так* расположен между определением и определяемым. Слева от маркера акцентированные определения — согласованное *этнокультурные* в (34) и несогласованное *нехорошего качества* в (35), и именно они попадают в сферу действия маркера, другими словами, подаются с хеджированием, т. е. как номинации, требовавшие раздумья, отобранные

⁹ Об использовании маркера *такой* в контексте хезитации см. [Сатюкова, Воейкова 2010]. В случаях, для которых мы используем термин «плейсхолдер», авторы говорят о поисковой функции местоимения *такой* и о функции грамматического оформителя.

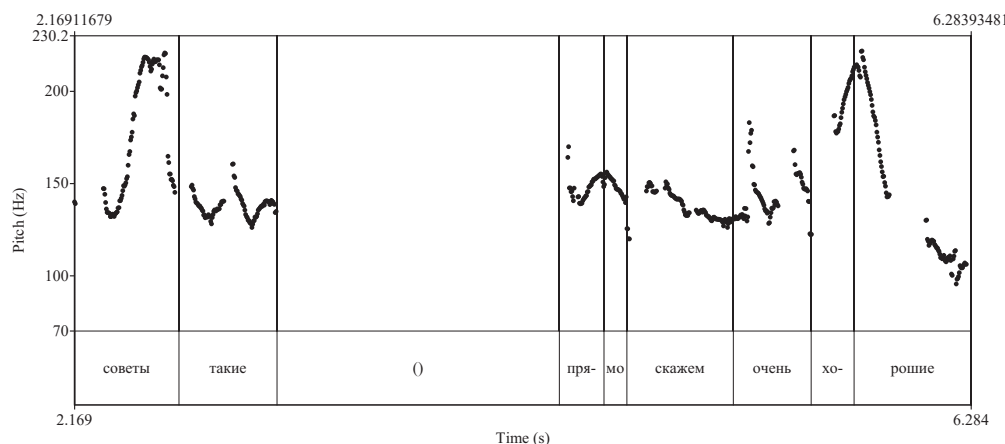


Рис. 19. Интонограмма к примеру (33)

из набора возможных. В (34) сфера действия подчеркивается дополнительным маркером хезитации — плейсхолдером *такие*. И в (34), и в (35) маркер реализуется с акцентом, с нисходящим тоном — следуя паттерну, типичному для заключительного маркера. Дополнительный акцент с восходящим тоном, который возникает в части опорного сегмента справа от маркера — на словах *особенности* в (34) и *видео* в (35) не рематический, он сигнализирует о незавершенности, о том, что слушающему следует ожидать продолжения:

- (34) ...где национальные виды спорта/ национальные виды деятельности/ такие этнокультурные/ скажем так/ особенности наши [Алексей Цыденов, Анна Атанова. Врио Главы Бурятии Алексей Цыденов в программе «От первого лица» (2017)].

где-е (0.46) национальные виды /-спорта,,,
 (э 0.35) национальные виды /-деятельности,,,
 (1.04)
 такие ==
 (? 0.30) \этнокультурные —
 (0.55)
 скажем \так,
 — /особенности наши-и,,,

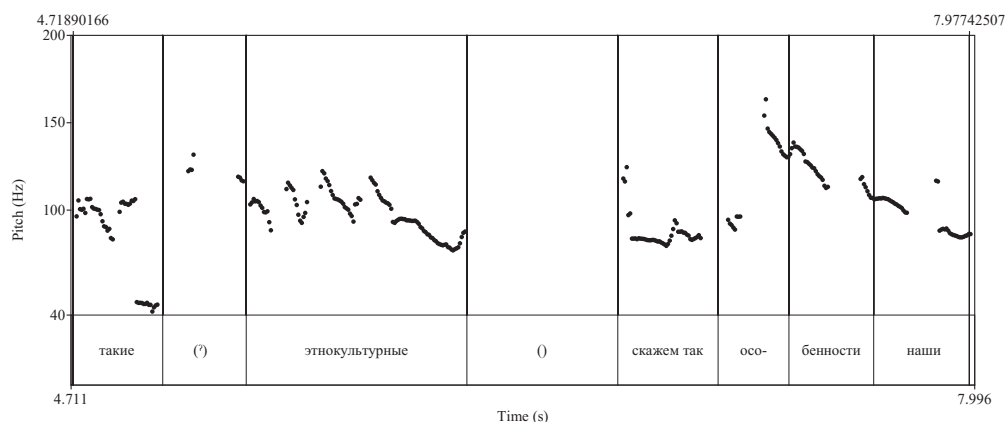


Рис. 20. Интонограмма к примеру (34)

- (35) ...но это не широкоэкранное кино/ а часто очень нехорошего качества/ скажем так/ видео... [Илья Утехин. Как устроен диалог? Лекция Полит.Ру (2017)].

часто-о (э 0.22) очен-нь н-нехорошего \качества —

(ц 0.37) скажем \так°,

— (э 0.28) (эшэ 0.42) /видео,

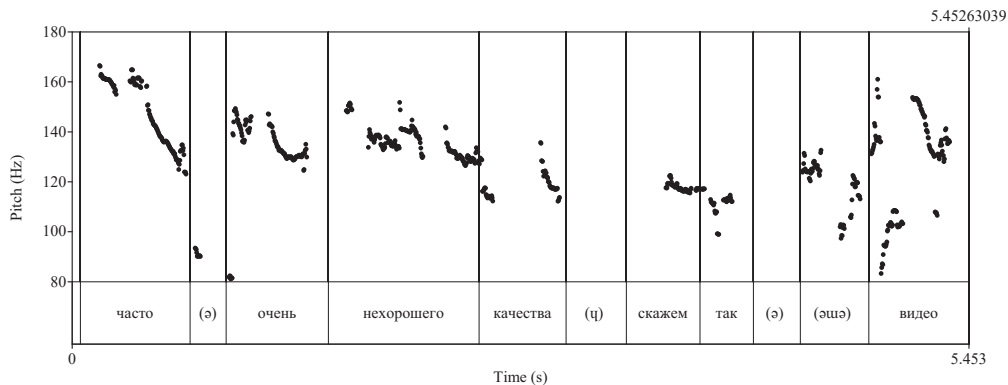


Рис. 21. Интонограмма к примеру (35)

Таковы основные просодические конфигурации, представленные в конструкциях с маркером СКАЖЕМ. Покажем теперь, какой вклад в целостный портрет маркера СКАЖЕМ вносят речевые сбои в его окрестностях.

4. Сфера действия дискурсивного маркера и речевые сбои

Помимо просодии, есть еще один круг явлений, который становится доступным исследователю при обращении к аудио- и видеоданным. Это сигналы речевых затруднений и сбоев. К сожалению, при поиске по корпусу опереться здесь только на транскрипты НКРЯ оказывается недостаточным. При тех объемах данных, с которыми имеют дело разработчики НКРЯ, они не могут (и не должны!) полностью отражать в транскрипте детали явлений такого рода. Здесь нужен прицельный инструментальный и перцептивный анализ. Такого рода анализ позволил нам сделать следующее важное наблюдение: маркер СКАЖЕМ регулярно сопровождается такими симптомами речевых затруднений, как самоисправления, повторы, лексические маркеры — гезитации (в т. ч. так называемые плейсхолдеры — *это самое, как его, такой* и т. п.) и реформулирования (*то есть, вернее* и пр.), долексические маркеры гезитации (мэканья, нефонологические удлинения звуков и т. п.), неречевые вокальные явления (сглатывание, хмыканье, чмоканье и пр.). СКАЖЕМ может — даже в пределах одного и того же опорного сегмента — дублироваться другими маркерами нечеткой номинации (в том числе аппроксиматорами/репрезентативами: *там, допустим, типа* и т. п.) Мы считаем, что такое взаимное тяготение СКАЖЕМ и речевых затруднений не является случайным: поскольку центральная функция данного маркера — предъявление номинации как нечеткой, одной из возможных, то на поверхность всплывают симптомы того, что выбор номинации требовал определенной работы, затраты дискурсивных усилий. Продемонстрируем это на примерах.

В (36) опорный фрагмент *это длительность жеста в словах* — это та часть высказывания, та формулировка, над которой работает говорящая. Фрагмент замыкается маркером

СКАЖЕМ2 (*скажем так*), а открывается другим хеджем, близким к первому по смыслу, — дискурсивным маркером *грубо говоря*, тоже восходящим к глаголу речи. При этом внутри опорного сегмента имеется повтор предлога с частью слова (*в сло=*), что дополнительно свидетельствует о преодолеваемом затруднении:

- (36) Так вот/ зоной действия жеста я считаю сочетание ударной части и постударного удержания. Грубо говоря/ это длительность жеста в словах/ скажем так [Е. А. Гришина. Доклад на конференции «Диалог 2012» (2012)].

зоной действия /жеста /я' считаю сочетание /ударной части (0.68) и-и (э 0.21) постударного \удержания.

(0.58)

/Грубо говоря,

это \длительность /жеста в сло= || в \словах,

(0.34)

скажем \так.

Мы видели уже в предыдущем разделе в примерах (33), (34), что плеисхолдеры (например, *такой/такие*) могут указывать на границу семантической сферы действия маркера СКАЖЕМ. В (37) ниже плеисхолдер *таких* выполняет ту же роль: сигнализирует о том, что сферой действия СКАЖЕМ является именно формулировка *потасканных собак*: плеисхолдер проецирует, во-первых, падежно-числовую форму отложенной именной группы (винительный падеж, множественное число), а во-вторых, ее просодическую реализацию, дублируя акцентированное произнесение с нисходящим тоном на акцентоносителе — словоформе *собак*. Имеется еще и риторический аргумент в пользу того, что говорящий работает именно над формулировкой *потасканных собак* — это фигура речи, которая употреблена для усиления образности, поскольку повествование ведется про оленей, и после заключительного *так скажем* говорящий переходит к прямому их описанию:

- (37) То есть поначалу они такие сильные/ все с такими рогами/ упитанные. А через месяц они/ м-м/ гона/ уже превращаются в таких потасканных собак/ так скажем. Очень худые все/ еле несут эти рога на голове/ прям вот так вот к земле клонятся. Бывает/ погибают даже [Саами (Ловозеро Мурманская область). Д/ф из цикла «Письма из провинции» (ТК «Культура») (2017)].

через /месяц со днь= || (? 0.31) /гона (0.20) уже превращаются в \таких || (0.48) потасканных \собак,

\так скажем!

Очень /-худы-ые,,,

все еле \несут эти-и /-рога,,,

на \голове,

Разнообразные сигналы речевых затруднений нередко группируются вокруг СКАЖЕМ, образуя кластеры. Хорошей иллюстрацией служат следующие два примера. В (38) автор ищет адекватную номинацию для значения 'один или чуть почаше'. О том, что на пути поиска говорящий испытывает серьезные затруднения, говорят — помимо интересующего нас СКАЖЕМ2 (*примерно скажем*) — самоисправления (отмененное *из ты=* — это забракованное *из тысячи*, вместо нужного *из ста*); повторы (несколько *по=* — это попытки подступиться к слову *почаше*); паузы, заполненные эканием (э), мэканием (ш) и скрипом (?); удлинения (*на-а, по-о=*), лексические маркеры хезитации (*там, ну, так*):

- (38) Более того/ там/ мы с обезьяной отличаемся последовательностью нуклеотидов на один. Ну/ так/ примерно скажем. Из ты... из ста по чуть-чуть почаше. И вот что же остальные девяносто девять из ста? [Николай Янковский. Генетика и геномика. Проект Academia (ГТРК Культура) (2010)].

мы с /обезьяной *отличаемся последовательностью* /нуклеотидов (ц 0.23) на-а (ц 0.44) (э 0.28) \один —
там \ну,
\так,
(ш 0.42) 0.26) \примерно скажем,
— по-о= || из /ты= || из || из /ста,
(ц 0.41)
(э 0.35) по-о= || (? 0.29) чуть-чуть почаше.

В (15), повторенном ниже как (39), помимо интересующего нас СКАЖЕМ1, имеется три лексических хеджа (*там, предположим, не знаю*), а также лексико-грамматический маркер открытого списка (конструкция *или с кем-то еще*) и нефонологические удлинения звуков (*нана-айцы, там-м*). Плюс к этому, в строке *Скажем нана-айцы там-м* акцентоноситель словоформа *нанайцы* произносится растянуто с восходяще-нисходящим акцентом типа ИК-6 в терминологии интонационных конструкций по Е. А. Брызгуновой [1980]. Эта конфигурация, которую условно можно назвать интонацией многоточия, широко используется для выражения незавершенности при описании череды событий, открытого списка, неуверенного припоминания [Янко 2008: 109–113, 166–167]. Т. е., помимо лексико-грамматического маркера нечеткой номинации, имеется и просодический маркер:

(39)=(15) В те времена это был пиджин/ потому что — и что очень важно — местное... местное население между собой контактировало на нём. Скажем/ нанайцы/ там/ предположим/ не знаю/ с удэгейцами или с кем-то ещё [Евгений Головкин. Как рождаются языки. Лекции Полит.ру (2011)].

В /те времена это был \пиджин.
Потому что —
(ц 0.36)
и ==
(? 0.16) что очень /важно,
(0.49)
— местное|| местное население между \собой контактировало на нём.
Скажем /\нана-айцы там-м,,
(0.61)
предположим не знаю с \удэгейцами,
или с кем-то /ещё,

Таким образом, мы видим, что СКАЖЕМ не оказывается одиноким дискурсивным маркером, а встраивается в сложную, многоуровневую систему, подчиненную единой задаче — выразить значение нечеткой номинации. Эту задачу решают совокупно лексические, грамматические, просодические и риторические средства.

5. Итоги

На основании данных НКРЯ, и прежде всего его мультимедийного подкорпуса, нам удалось получить представление об использовании в живой русской речи семейства дискурсивных маркеров, восходящих к личной форме глагола речи СКАЖЕМ. Мы исследовали два представителя этого семейства — изолированное СКАЖЕМ1 и комбинацию с наречием (в том числе, местоименным) — СКАЖЕМ2. Мы обнаружили как свойства, общие для всего семейства, так и свойства отдельных его представителей.

1. Маркеры семейства СКАЖЕМ демонстрируют типичный эффект прагматикализации по сравнению с исходным значением речевого глагола: они не отсылают к речевому событию, их основной функцией становится выражение нечеткой номинации (хеджирования), т. е. эти маркеры используются тогда, когда говорящий снимает с себя ответственность за возможную неточность пропозиционального содержания или возможную неточность используемой номинации применительно к фиксированному пропозициональному содержанию. Как таковые, маркеры этого семейства часто употребляются совместно с другими средствами, демонстрирующими, что говорящему требуются дополнительные усилия при выборе адекватного средства выражения. Это могут быть другие лексические маркеры нечеткой номинации, лексические и просодические показатели открытого списка, самоисправления, долексические маркеры хезитации (заполненные паузы, нефонологические удлинения звуков) и пр.
2. Маркеры семейства СКАЖЕМ демонстрируют типичный эффект так называемой «кооптации»: они выпадают из жесткой клаузуальной структуры и приобретают статус оператора, сферой действия которого являются сегменты дискурса различной (потенциально — любой) протяженности.
3. Маркеры семейства СКАЖЕМ демонстрируют существенный сдвиг грамматического значения: форма презенса совершенного вида уже не обозначает будущее время — как при полнозначном употреблении. Однако разные представители этого семейства затронуты этим сдвигом в разной степени: СКАЖЕМ2 обладает гораздо большими грамматическими свободами, чем СКАЖЕМ1. СКАЖЕМ2 допускает не только форму множественного числа 1 лица, но и форму единственного, СКАЖЕМ1 такой вариативности не допускает. Наречие может располагаться как слева, так и справа от СКАЖЕМ 2, кроме того, в отличие от СКАЖЕМ1, СКАЖЕМ2 может быть обременен частицей и даже деепричастным оборотом.
4. Маркеры семейства СКАЖЕМ могут произноситься как атонически, так и с акцентом и значимым движением тона в акцентоносителе, однако для СКАЖЕМ1 атоническое употребление значительно более характерно, чем для СКАЖЕМ2. Иными словами, СКАЖЕМ1 в большей степени, чем СКАЖЕМ2 утратил способность к просодической автономии. Т. е. СКАЖЕМ1 имеет не только меньше грамматических свобод, чем СКАЖЕМ2, но и меньше просодических свобод. Можно говорить о том, что СКАЖЕМ1 дальше ушел от своего полнозначного прародителя в сторону дискурсивного маркера.
5. Маркеры семейства СКАЖЕМ чаще всего располагаются внутри своего опорного сегмента, реже — слева от него и совсем редко — справа. Если маркер расположен слева или справа от опорного сегмента, то, как правило, его семантическая сфера действия, т. е. фрагмент дискурса, который объявляется нечеткой номинацией, совпадает с опорным сегментом. Однако в тех случаях, когда маркер располагается внутри опорного сегмента, его семантической сферой действия оказывается не весь опорный сегмент, а его часть до или — гораздо чаще! — после маркера. Получается, что маркер преимущественно располагается слева от своей семантической сферы действия (которая при этом либо совпадает с опорным сегментом, либо является частью опорного сегмента, следующей за маркером).
6. Если маркер семейства СКАЖЕМ располагается в своей прототипической позиции слева от своей сферы действия и при этом произносится с коммуникативно-значимым акцентом, вся конструкция следует одному из двух просодических паттернов: просодически автономное акцентированное произнесение с падающим тоном, аналогичное препозитивной авторской ремарке, и просодически автономное акцентированное произнесение с восходящим тоном, аналогичное начальной адвербиальной группе. В этом можно усмотреть просодический рудимент исходного употребления СКАЖЕМ в качестве самостоятельного речевого глагола.

Полученные результаты поучительны в ряде отношений. Во-первых, становится очевидным, что даже внутри одного семейства маркеров сдвиг от единицы с пропозициональным значением к дискурсивному маркеру может проходить по различающимся траекториям и приводить к большим или меньшим сдвигам в значении и форме. Во-вторых, исследование дискурсивных маркеров с привлечением данных о просодии и таких специфических феноменах устного дискурса, как речевые затруднения, позволяет сделать лексикографические портреты каждой отдельной единицы и их семейств более многомерными.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Баранов, Добровольский 1996 — Баранов А. Н., Добровольский Д. О. *Англо-русский словарь по лингвистике и семиотике*. М.: Помовский и партнеры, 1996. [Baranov A. N., Dobrovolskii D. O. *Anglo-russkii slovar' po lingvistike i semiotike* [English-Russian dictionary of linguistics and semiotics]. Moscow: Pomovskii i partnery, 1996.]
- Богданова-Бегларян 2012 — Богданова-Бегларян Н. В. Конструкция (...) скажем (...) в повседневной русской речи. *Вестник Калмыцкого института гуманитарных исследований РАН*, 2012, 2: 153–157. [Bogdanova-Beglaryan N. V. The (...) skazhem (...) construction in everyday Russian speech. *Vestnik Kalmytskogo instituta gumanitarnykh issledovaniy RAN*, 2012, 2: 153–157.]
- Богданова-Бегларян и др. 2019 — Богданова-Бегларян Н. В., Блинова О. В., Зайдес К. Д., Шерстинова Т. Ю. Корпус «Сбалансированная аннотированная текстовая» (SAT): изучение специфики русской монологической речи. *Труды Института русского языка им. В. В. Виноградова*, 2019, 3(21): 110–126. [Bogdanova-Beglaryan N. V., Blinova O. V., Zaides K. D., Sherstinova T. Yu. Corpus “Balanced annotated text collection (Textotec)” (SAT): Studying the specificity of Russian monological speech. *Proceedings of the V. V. Vinogradov Russian Language Institute*, 2019, 3(21): 110–126.]
- Брызгунова 1980 — Брызгунова Е. А. Интонация. *Русская грамматика*. Т. 1: Фонетика, фонология. Ударение. Интонация. Словообразование. Морфология. Шведова Н. Ю. (ред.). М.: Наука, 1980, 103–118. [Bryzgunova E. A. Intonation. *Russkaya grammatika*. Vol. 1: *Fonetika, fonologiya. Udarenie. Intonatsiya. Slovoobrazovanie. Morfologiya*. Shvedova N. Yu. (ed.). Moscow: Nauka, 1980, 103–118.]
- Кибрик, Подлеская (ред.) 2009 — Кибрик А. А., Подлеская В. И. (ред.). *Рассказы о сновидениях: Корпусное исследование устного русского дискурса*. М.: Языки славянских культур, 2009. [Kibrik A. A., Podlesskaya V. I. (eds.). *Rasskazy o snovideniyakh: Korpusnoe issledovanie ustnogo russkogo diskursa*. [Night Dream Stories: A corpus study of spoken Russian discourse]. Moscow: Yazyki Slavyanskikh Kul'tur, 2009.]
- Копотев 2014 — Копотев М. В. Эволюция русских маркеров ренарратива: синтаксис или лексика? *Acta Linguistica Petropolitana*, 2014, X(2): 712–740. [Kopotev M. V. The ways to re-narrate: development of the reported speech markers in Russian. *Acta Linguistica Petropolitana*, 2014, X(2): 712–740.]
- НКРЯ — Национальный корпус русского языка [Russian National Corpus]. <http://www.ruscorpora.ru>.
- Подлеская, Пожилов 2021 — Подлеская В. И., Пожилов Ю. М. Семантика, грамматика и просодия вводно-союзных конструкций по данным мультимедийного подкорпуса НКРЯ. *Компьютерная лингвистика и интеллектуальные технологии: по материалам международной конференции «Диалог 2021»*, 2021, 20(27): 546–559. [Podlesskaya V. I., Pozhilov Yu. M. Semantics, grammar and prosody of parentheticals introduced by the subordinator *kak 'as'*. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: po materialam mezhdunarodnoi konferentsii «Dialog 2021»*, 2021, 20(27): 546–559.]
- Рахилина и др. 2021 — Рахилина Е. В., Бычкова П. А., Жукова С. Ю. Речевые акты как лингвистическая категория: дискурсивные формулы. *Вопросы языкознания*, 2021, 2: 7–27. [Rakhilina E. V., Bychkova P. A., Zhukova S. Yu. Speech acts as a linguistic category: The case of discourse formulae. *Voprosy Jazykoznanija*, 2021, 2: 7–27.]
- Сай 2010 — Сай С. С. Динамика развития обстоятельств времени со значением предшествования на интервал в русском языке. *Acta Linguistica Petropolitana*, 2010, VI(2): 131–183. [Sai S. S. Development of temporal adverbials with the meaning of “preceding interval” in Russian. *Acta Linguistica Petropolitana*, 2010, VI(2): 131–183.]
- Сатюкова, Воейкова 2010 — Сатюкова Д. Н., Воейкова М. Д. Особенности функционирования местоимения *такой* в устной разговорной речи. *Acta Linguistica Petropolitana*, 2010, VI(2): 184–224. [Satyukova D. N., Voeikova M. D. Specific functions of the pronoun *takoj* ‘such’ in spoken Russian. *Acta Linguistica Petropolitana*, 2010, VI(2): 184–224.]

- Янко 2008 — Янко Т. Е. *Интонационные стратегии русской речи в сопоставительном аспекте*. М.: Языки славянских культур, 2008. [Yanko T. E. *Intonatsionnye strategii russkoi rechi v sopostavitel'nom aspekte* [Intonation strategies of Russian speech in a comparative aspect]. Moscow: Yazyki slavyanskikh kul'tur, 2008.]
- Янко 2021 — Янко Т. Е. Коммуникативная структура как отражение точки зрения говорящего. *Проблемы функциональной грамматики. Отношение к говорящему в семантике грамматических категорий*. Казаковская В. В., Воейкова М. Д. (ред.). 2-е изд. М.: Языки славянских культур, 2021, 191–222. [Yanko T. E. Communicative structure as a reflection of the speaker's point of view. *Problemy funktsional'noi grammatiki. Otnoshenie k govoryashchemu v semantike grammaticheskikh kategorii*. Kazakovskaya V. V., Voeikova M. D. (eds.). 2nd edn. Moscow: Yazyki slavyanskikh kul'tur, 2021, 191–222.]
- Boersma, Weenink 2021 — Boersma P., Weenink D. *Praat: doing phonetics by computer* [Computer program]. Version 6.1.38, 2021. <http://www.praat.org/>.
- Chafe 1994 — Chafe W. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: Univ. of Chicago Press, 1994.
- Dehé 2007 — Dehé N. The relation between syntactic and prosodic parenthesis. *Parentheticals*. Dehé N., Kavalova Y. (eds.). Amsterdam: John Benjamins, 2007, 261–284.
- Haselow 2015 — Haselow A. Left vs. right periphery in grammaticalization: The case of *anyway*. *New directions in grammaticalization research*. Smith A. D. M., Trousdale G., Waltereit R. (eds.). Amsterdam: John Benjamins, 2015, 157–186.
- Heine et al. 2021 — Heine B., Kaltenböck G., Kuteva T., Long H. *The rise of discourse markers*. Cambridge: Cambridge Univ. Press, 2021.
- Kaltenböck 2008 — Kaltenböck G. Prosody and function of English comment clauses. *Folia Linguistica*, 2008, 42(1): 83–134.
- Kaltenböck 2009 — Kaltenböck G. English comment clauses: Positions, prosody, and scope. *Arbeiten aus Anglistik und Amerikanistik*, 2009, 34(1): 51–77.
- Kaltenböck et al. 2010 — Kaltenböck G., Mihatsch W., Schneider S. (eds.). *New approaches to hedging*. Bingley: Emerald, 2010.
- Kaltenböck et al. 2011 — Kaltenböck G., Heine B., Kuteva T. On thetical grammar. *Studies in Language*, 2011, 35(4): 852–897.
- Lakoff 1972 — Lakoff G. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Chicago Linguistic Society Papers*, 1972, 8: 183–228.
- Markkanen, Schröder 1997 — Markkanen R., Schröder R. (eds.). *Hedging and discourse approaches to the analysis of a pragmatic phenomenon in academic texts*. Berlin: Walter de Gruyter, 1997.
- Peters 2006 — Peters J. Syntactic and prosodic parenthesis. *Proc. of the International Conf. on Speech Prosody. 2–5 May 2006, Dresden. 2006*. <http://www.isca-speech.org/archive>.
- Prince et al. 1982 — Prince E. F., Frader J., Bosk C. On hedging in physician-physician discourse. *Linguistics and the professions. Proc. of the Second Annual Delaware Symposium on Language Studies*. Pietro R. J. (ed.). Norwood (NJ): Ablex, 1982, 83–97.
- Russian Constructicon 2021 — Bast R., Endresen A., Janda L. A., Lund M., Lyashevskaya O., Mordashova D., Nessel T., Rakhilina E., Tyers F. M., Zhukova V. *The Russian Constructicon. An electronic database of the Russian grammatical constructions*. 2021. <https://constructicon.github.io/russian/>.
- Wichmann 2001 — Wichmann A. Spoken parentheticals. *A wealth of English. Studies in honour of Göran Kjellmer*. Aijmer K. (ed.). Göteborg: Acta Universitatis Gothoburgensis, 2001, 177–193.

Чередование *a* и *o* в основе и прагматика распределения вариантов

© 2024

Максим Анисимович Кронгауз

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; mkronhaus@yandex.ru

Аннотация: В статье рассматриваются чередования букв *a* и *o* в основе, пропущенные в школьных и академических грамматиках, например *шпаргалка* — *шпора*, *башка* — *бошки*, *симпатичный* — *симпотный*, *платит* — *плотит*, *налит* — *нолит*. Предпринимается попытка объединить ряд разнородных (в том числе и уникальных) явлений в рамках особого чередования, а конкретнее — перехода *a* в *o* в некодифицированной области языка: в профессиональных и молодежных жаргонах, просторечии, подчеркнуто разговорной речи. Таким образом, можно говорить об относительно регулярных чередованиях *a* и *o* на стыке литературного и нелитературного языка. Буква *o* передает фонему [o], которая появляется под ударением в не вполне нормативных или неформальных словоформах и словах (просторечных, жаргонных и др.), то есть подчеркивает эту неформальность. Появление *o* вместо привычного *a* и фонетически и графически маркирует неформальность слова, а часто также его экспрессивность.

Ключевые слова: вокализм, морфология, разговорная речь, русский язык, чередования

Для цитирования: Кронгауз М. А. Чередование *a* и *o* в основе и прагматика распределения вариантов. *Вопросы языкознания*, 2024, 2: 83–92.

DOI: 10.31857/0373-658X.2024.2.83-92

The alternation of stem-internal *a* and *o* and pragmatic mechanisms of alternant distribution

Maxim A. Krongauz

HSE University, Moscow, Russia; rmkronhaus@yandex.ru

Abstract: The article considers the alternation of letters *a* and *o* in the stem, missing in school and academic grammars, for example, *shpargalka* — *shpora*, *boshka* — *boshki*, *simpatichnyj* — *simpotnyj*, *platit* — *plotit*, *nalit* — *nolito*. We attempt to combine a number of heterogeneous (including unique) phenomena within the framework of a special alternation, and more specifically, the transition from *a* to *o* in an uncoded area of the language: in professional and youth jargons, popular language, emphatically colloquial speech. Thus, one can speak of relatively regular alternations of *a* and *o* at the junction of literary and non-literary language. The letter *o* conveys the phoneme [o], which appears under stress in not quite normative or informal word forms and words (colloquial, slang, etc.), that is, it emphasizes this informality of a word and often also its expressiveness.

Keywords: alternations, colloquial speech, morphonology, Russian, vowels

For citation: Krongauz M. A. The alternation of stem-internal *a* and *o* and pragmatic mechanisms of alternant distribution. *Voprosy Jazykoznanija*, 2024, 2: 83–92.

DOI: 10.31857/0373-658X.2024.2.83-92

Казус как источник проблемы

В интернете из года в год публикуются статьи о том, как падает грамотность и насколько безграмотны современные школьники. В одной из статей в качестве примера безграмотности школьников фигурировало слово *шпаргалка*. Безграмотность школьников подчеркивалась не только тем, что они пишут это слово с буквой *о*, но и тем, что они проверяют эту безударную гласную словом *шпора*, а она на самом деле непроверяемая.

Вопросы о том, как писать слово *шпаргалка* и можно ли его проверить, встречаются в интернете. Вот, например, на сайте [Большой вопрос.ру] спрашивают: «Как проверить слово *шпаргалка*?». Один из ответов таков:

«Шпар-гал-ка, три слога в слове, ударная гласная во втором слоге.

Слово имеет польские корни, найти в русском языке проверочное слово для проверки безударной гласной *А* в первом слоге нам не удастся, поэтому написание слова остается только запомнить.

Слово *шпаргалка* включено в список словарных слов».

Здесь просто игнорируется жаргонное слово *шпора*, а *шпаргалка* объявляется словарным словом. В двух других ответах жаргонизм упоминается:

«Слово *Шпаргалка* оказывается существительным женского рода и первого склонения, в котором находим окончание *-А*: *Шпаргалка-Шпаргалкой-Шпаргалку*.

Ударение в нем падает на второй слог: *шпаргАлка*.

Корнем слова оказывается морфема *ШПАРГАЛК-*: *Шпаргалочка-Шпаргалочный*.

Обратим внимание, что в этом корне имеется безударная гласная *А*, а само слово можно ошибочно написать как *шпОргалка*.

Проверять эту безударную гласную словом *Шпора* ни в коем случае нельзя, потому что эти слова ничем не связаны, кроме обозначаемого предмета и само слово *Шпаргалка* следует запомнить. Ведь проверочных слов для него нет, а само оно пришло из польского языка».

«Слово *ШПАРГАЛКА* пришло из польского языка, *szpargalka* — исписанная бумажка, а оно в свою очередь восходит к греческому *sparganon* — пленка, лохмотья от *spargo* — завертываю, пеленаю.

Несмотря на то, что в разговорном языке есть слово *ШПОРА*, слово *ШПАРГАЛКА* пишем через букву *А*, это нужно запомнить».

В одном из ответов даже предпринимается попытка объяснить, почему *шпора* не может быть проверочным словом:

«Проверять эту безударную гласную словом *Шпора* ни в коем случае нельзя, потому что эти слова ничем не связаны, кроме обозначаемого предмета».

Трудно понять, что это означает с лингвистической точки зрения, возможно, это отрицание того, что эти слова являются однокоренными.

Несколько иначе объясняется данная ситуация на образовательном портале [Мел]:

«Интересно, ещё остались те, кто пишет *шпаргалки* от руки, а не печатает их на принтере? Народная мудрость гласит (а личный опыт подтверждает), что это один из самых эффективных способов подготовки к экзамену — пока делаешь письменную “шпору”, в голове точно что-то отложится.

Главное — не проверять этим жаргонным словечком слово “шпаргалка”. В русский язык оно пришло из украинского через польский, где *szpargal* означает “старая исписанная бумажка”. Так что в середине там “а”, а не “о”. И “шпора” — отнюдь не проверочное слово».

Здесь также остается не вполне понятным последнее утверждение. *Шпора* — не проверочное слово, но почему?

Все-таки графическое/фонетическое сходство *шпаргалки* и *шпоры* настолько велико, что хочется признать их однокоренными, и тогда приходится говорить о чередовании букв *а* и *о*¹. Другая теоретическая возможность состоит в том, чтобы развести их по разным системам. Слово *шпаргалка* относится к литературному языку, а слово *шпора* в этой языковой системе просто не существует и относится к школьному жаргону.

В любом случае интересно проверить, как это чередование соотносится с другими чередованиями *а* и *о* в русском языке.

Чередования *а* и *о*, описанные в лингвистической литературе

Одно из них, включающее целый ряд корней, известно еще по школьной программе:

Написание корней с чередующимися гласными А и О может зависеть от места ударения в слове, от значения слова и от тех букв, которые следуют за чередующейся гласной.

1. От места ударения в слове зависит написание корней ГОР/ГАР, КЛОН/КЛАН, ТВОР/ТВАР, ПЛОВ/ПЛАВ, ЗОР/ЗАР. Под ударением в этих корнях пишется то, что слышится, там ошибку допустить невозможно, а вот гласные, пишущиеся в безударном положении, следует выучить наизусть: ГОР, КЛОН, ТВОР, ПЛАВ, ЗАР. Исключения: ВЫГАРКИ, ИЗГАРЬ, ПРИГАРЬ, ПЛОВЕЦ, ПЛОВЧИХА.
2. От значения слова зависит выбор буквы в корнях МАК/МОК и РАВН/РОВН. Слово со значением «погружать в жидкость пишется с буквой А (МАКАТЬ КИСТЬ В КРАСКУ), если же его значение «постепенно пропитываться», то следует писать О (ОБУВЬ ПРОМОКАЕТ). Слово, имеющее значение «гладкий, горизонтальный, прямой», пишется с буквой О (РАЗРОВНЯТЬ ЗЕМЛЮ НА КЛУМБЕ), а слово со значением «одинаковый» следует писать с буквой А (РАВНЕНИЕ ПО РОСТУ, УРАВНОВЕШЕННЫЙ ЧЕЛОВЕК). Исключения: РАВНИНА, РОВЕСНИК, ПОРОВНУ, УРОВЕНЬ.
3. От следующих после чередующихся гласных букв зависит написание корней ЛАГ/ЛОЖ, РАСТ/РАЩ/РОС, КАС/КОСН, СКАК/СКОЧ. Рассмотрите примеры: ПРЕДЛАГАТЬ — ПРЕДЛОЖИТЬ, РАСТЕНИЕ — ВЫРАЩИВАТЬ — ПОДРОСЛИ, КАСАТЬСЯ — КОСНУТЬСЯ, СКАКАТЬ — ВСКОЧИТЬ. Исключения: ОТРАСЛЬ, РОСТОК, РОСТОВИЩИК, РОСТОВ, РОСТИСЛАВ и СКАЧОК, СКАЧУ [Грамота.ру].

Это вполне традиционное школьное описание вызывает целый ряд вопросов. В частности, должны ли входить имена собственные *Ростов* и *Ростислав* в список исключений, то есть можно ли их на синхронном уровне относить к корню *рос*? Можно ли на синхронном уровне говорить о сохранении единого корня *твор/твар*? Вот что об этом написано в [Кронгауз и др. 2018: 29]:

«В школьном курсе русского языка обычно в этой группе рассматривают ещё корень *твор* // *твар*. Исторически это, действительно, один корень, имевший два варианта. Однако со временем варианты очень сильно разошлись по значению и стали самостоятельными корнями. Первый вариант связан с идеей творчества и творения

¹ Здесь и далее речь пойдет прежде всего о графических, а не фонетических чередованиях, хотя и фонетические чередования упоминаются в ряде работ: [РГ 1980, 1; Трубецкой 1987]. В дальнейшем это различие не комментируется, поскольку оно вытекает из примеров и контекста.

и пишется с буквой *о* не только в ударной, но и в безударной позиции: *твори́ть, творческий, творе́ц, творе́ние, болезнетво́рный*. Второй вариант встречается в нескольких словах, изначально связанных с идеей творения, но по существу потерявших или теряющих эту связь. Он также может встречаться и в ударной, и в безударной позиции: *тварь, тварю́га, у́тварь*».

Однако в любом случае очевидно, что чередование *шпаргалка / шпора* в этот ряд чередований не вписывается. Есть еще один хорошо известный ряд регулярных чередований *а* и *о*. Он подробно описан в [РГ 1980, 1: 428–429]:

«§ 1076. Неконечные чередования (кроме беглости гласной), как правило, непродуктивны и выступают только в определенных корнях, главным образом глагольных. Единственным продуктивным из относящихся сюда чередований является чередование гласных *о — а* после парно-твердой согласной в глагольных корнях перед суф. морфом *-ива-*: *заморозить — замораживать, исхлопотать — исхлопатывать, звонить — перезванивать* (о морфонологических условиях этого чередования см. § 845). В образованиях, мотивированных глаголами на *-ивать* (орфогр. также *-ывать*), содержащими результат чередования *о — а*, регулярно выступает противоположное ему по направлению чередование *а — о*, т. е. восстанавливается первоначальная корневая гласная словообразовательной цепочки: *перезванивать — перезвон- Ø, ухаживать* (например, за больным) — *уход- Ø, позванивать — позвонок* (колокольчик), *разговаривать — разговор- Ø и разговорный*».

Переход *о* в *а* в этой позиции относится к возможным, но не обязательным (ср.: *сосредоточивать, отсрочивать...*), однако тенденция настолько сильна, что она модифицирует и новые корни. Так, именной корень *тамож*, появляясь в глаголе, подчиняется этой тенденции: *Тамо́жня — растамо́жить — растама́живать*.

Более редкие и даже уникальные чередования отмечены в [Трубецкой 1987: 99]:

«§ 40. Чередование *о ~ а* представлено: А. В *именном основообразовании* единственным изолированным примером служит *л'і́б'о́ткă* (лебедка) ~ *л'і́б'ắžj* (лебяжий)³⁴. — В. В *словоизменении глагола* это чередование имеет место только в двух изолированных нерегулярных парадигмах: наст. (буд. вр. *л'ăgŭ* (лягу) ~ прош. вр. *л'оК* (лег) и причастие прош. вр. страдат. залога *ză|přăžjŋ* (запряжен) ~ прош. вр. м. р. *ză|přòК* (запряг)».

Н. С. Трубецкой также отмечает чередование в смешанном основообразовании, о котором будет сказано ниже.

Чередования *а* и *о*, не описанные в лингвистической литературе

Очевидно, что среди описанных чередований нет таких, в которые можно было бы вписать чередование *шпаргалка / шпора*. Однако некоторое сходство с ним можно усмотреть в чередованиях, которые также оказались вне лингвистических описаний. Здесь придется вернуться на шаг назад к упомянутому выше хорошо известному чередованию *зар/зор*, которое описывается следующим образом:

Под ударением пишется то, что слышится, а в безударном положении пишется *а*:
за́рево, лучеза́рный;
зо́ри, зо́рька, зо́ренька;
за́ря, зарни́ца, озаре́ние, озаре́нность, озари́ть, за́рянка.

Более редкое и даже потенциальное словообразование подчиняется этому же правилу: *зёрюшка, зариюшечка...*

Следует обратить особое внимание на пару *заря*—*зори*, поскольку чередование происходит внутри лексемы существительного: в единственном числе с безударным корнем пишется *а*, а во множественном числе под ударением *о*.

С точно таким же распределением реализуется чередование в таких парах, где второй член заведомо находится вне литературного языка:

пальто—*пóльта*;
башка—*бóшки*²;
балда—*бóлды* (обычно в значении «голова»);
табло—*тóбла* (в значении «лицо»);
фата—*фóты*...

Ср. примеры из [НКРЯ]:

Если он в конце вечера бывал сильно пьян, то, сидя у барьера и мельком доброжелательно взглядывая на подходящих, тут же указывал пальцем своему напарнику, где висит то или иное пальто. По словам Афони, его бил еще Есенин. — За что? — полюбопытствовал я. — Да ни за что. Загонит за пóльта и дерется. Но не больно. Зато уж и платил!.. [Константин Ваншенкин. Писательский клуб (1998)]³.

Но вот после того, как пальто стало очень распространенной одеждой, его название сделалось общенародным, а когда народ ощутил это слово таким же своим, чисто русским, как, скажем, яйцо, колесо, молоко, толокно, он стал склонять его по правилам русской грамматики: пальто, пальту, пальтом и даже пальта [К. И. Чуковский. Живой как жизнь (разговор о русском языке) (1962)].

Взяли, конечно, билеты по рубль тридцать. Поднялись по лестнице. Вдруг назад кличут. Велят раздеваться. — Пóльта, говорят, сымайте. Локтев, конечно, с дамой моментально скинули пальта. А я, конечно, стою в раздумьи. Пальто у меня было в тот вечер прямо на ночную рубашку надето. Пиджака не было [М. М. Зощенко. Прелести культуры (1926)].

А вам мы еще бошки поотрываем! — пригрозил Диман братьям [Алексей Слаповский. Синдром Феникса // «Знамя», 2006].

— Льнов, скажи по секрету, на кого ты собираешься охотиться? Мамонтам бошки сносить? Так они вымерли [Михаил Елизаров. Pasternak (2003)].

Живут в лесу и на торфяных болотах. Люди, а бошки пёсьи. В тридцать седьмом начальник зоны полковник Рытов их как-то приручил, что ли... [Алексей Иванов (Алексей Маврин). Псоглавцы. Гл. 1–20 (2011)].

Какие хорошие!!!! Но дурные болды — все в хозяйку... :)))))) [Livejournal, комментарий (2009)].

Для форм *тобла* и *фоты* примеров ни в корпусе, ни в интернете не найдено, их существование основано только на устных свидетельствах носителей языка.

Вместе с этим списком хочется упомянуть и русское слово *говно*, имеющее в словарях пометы *простореч.*, *груб.* Формально в нем нет чередования, однако в единственном числе в безударном корне ошибочное написание *а* встречается довольно часто:

² Эта пара выделяется из ряда тем, что *о* может появляться и в единственном числе: *бóшка*. Подробнее об этом далее.

³ В статье используются примеры из НКРЯ и из интернета, полученные с помощью поисковой системы Яндекс. Обращение к НКРЯ и интернету — 19 августа 2023 г.

в НКРЯ — 12 текстов и 14 примеров, что может отражать параллельную орфографическую реальность. Однако во множественном числе вариативности нет, встречается только «правильный»⁴ вариант: *гòвна*.

*Я їх осмàтривàю, слòвно
покинутые мною гòвна
в лесу, в лугàх, среди цветòв.*

[Н. Байтов. Нескончаемые сетования (2000)]

Особенность слов из упомянутого выше списка заключается в том, что в литературном языке это либо несклоняемые, либо имеющие в словарном описании помету «мн. затрудн.»⁵ существительные. Все особые формы множественного числа оказываются за пределами литературного языка и могут быть отнесены к просторечию.

Отдельно надо сказать о слове *баишкà*. В толковых словарях, в частности в [МАС], оно имеет помету «простореч.». Однако также существует и слово *бóишка* с формами и во множественном, и в единственном числе. Оно не встречается в нормативных толковых словарях, но присутствует, например, в «Словаре русского арго» В. С. Елистратова [Елистратов 2000]: «БО'ШКА, -и, ж. Башка, голова». Десятки примеров обнаруживаются и в НКРЯ:

Тут ему на бoишку кирпич с крыши бац! [Алексей Иванов. Земля — Сортировочная (1990–1991)].

Читал, наверное, но из моей дырявой бoишки и более крупные самоцветы вываливаются [Александр Щербаков. Пах антилопы // «Октябрь», 2002].

Чё то у меня под вечер бoишка дырывается [Автогонки-3 (форум) (2005)].

Таким образом, ударное *о* в этом случае маркирует не особое просторечное множественное число, а целое слово («еще менее литературный»⁶ вариант), что похоже на противопоставление в паре *разлив* — *розлив*, которая будет рассмотрена ниже.

Оставаясь в рамках имени существительного, можно найти еще одно чередование, стоящее обособленно. Существует лексема *штоники*, которую также можно отнести к просторечию⁷. Она отсутствует в НКРЯ, но встречается в интернете:

...в вскр начала шить себе штоники новые... [Diary.ru 2003].

...для мальчиков юбочки и штоники по фигуре — да-да-да! всякие там растянутые оверсайсы и провисающие бойфренды — нет-нет-нет! [Livejournal 2015].

Муж и его семья говорят всегда штАники, впервые от них услышала это слово в таком звучании. Все кого я знаю, говорят штОники. Муж утверждает, что штАны, значит штАники [Baby.ru 2020].

В последнем примере отражена лингвистическая дискуссия о проверке написания гласных с помощью однокоренных слов (с точки зрения мужа) и возможном чередовании (с точки зрения жены). Мнение мужа интересно еще и тем, что ударная гласная проверяется безударной.

⁴ Этот вариант правильный с точки зрения орфографии, но с точки зрения грамматики образование множественного числа у этого слова затруднено.

⁵ В соответствии с [Зализняк 1987].

⁶ Конечно, строго говоря, мы не можем сравнивать более или менее просторечные варианты, но отсутствие *бoишки* в нормативных словарях, значительно более редкая встречаемость (в частности, если судить по НКРЯ) сообщает нам о большей удаленности этого варианта от границ литературного языка.

⁷ Не исключено влияние белорусского, где есть слово *штонікі* ‘штанишки’.

Наконец, надо отметить, что рассматриваемое чередование встречается в неформальном словообразовании личных имен:

*Лариса — Лора, Лорка...;
Матвей — Мотя, Мотька...;
Тамара — Тома, Томочка...*

*Из того, что она на секунду забыла про **Лору** (она сразу знала, что будет **Лора**, **Лариса**, имя для дочери придумала еще в детстве), почувствовала себя такой виноватой, что заплакала [Галина Щербакова. Реалисты и жлобы (1997)].*

*— Вот, слушай, **Мотька**, — меня **Матвеем** окрестили, — слушай! [Максим Горький. Исповедь (1908)].*

*— А как же святая царица **Тамара**? — напоминала **Тома**, которой еще прежде **Василиса** говорила о ее заступнице царице **Тамаре**, в честь которой она была крещена [Людмила Улицкая. Казус Кукоцкого [Путешествие в седьмую сторону света] // «Новый Мир», 2000].*

Конечно, в этом случае нельзя говорить о просторечии, но, очевидно, речь идет о неформальных разговорных формах личных имен.

К личным именам примыкает и еще один пример, связанный с типичной кличкой для барана — **Борька**. Строго говоря, здесь нельзя говорить о чередовании, но, скорее, о поэтическом сближении названия животного и имени, которое подтверждается многочисленными видеороликами в интернете, а также стихотворением А. Барто «Баран Борька»:

*Всюду ходит за Дуняшей
Борька, маленький **барашек**...*

Следует отметить также два обособленных чередования, связанных не с просторечием, а с жаргонами.

Во-первых, это прилагательное из молодежного сленга **симпóтный** (от *симпатичный*). Оно встречается в «Словаре русского арго» [Елистратов 2000] и в некоторых других словарях.

*Но Леонид никак не проявлял своего внимания — «ему не до этого» — ему всегда не до этого!.. А ведь **симпóтная** девчонка! Просто не может он с ними... [Алексей Шепелев. Кгышф Вшытундф-ТВ (2008) // «Волга», 2011].*

*А иной раз у таких уродливых родителей бывают очень **симпóтные** детки, прямо диву даёшься [Женщина + мужчина: Брак (форум) (2004)].*

Во-вторых, интересное чередование представлено в существительном в приставке **раз-**: **разлив** — **ро́злив**. Здесь соблюдаются правила чередования, характерного для приставки **рас/рос/раз/роз**, но некоторая особенность заключается в том, что здесь фактически представлена дублетная пара, то есть синонимы, которые формально различаются только вариантом приставки (и ударением, естественно). Впрочем, они различаются и стилистически: второе слово в паре (**ро́злив**) маркировано как профессиональное: в МАС имеется помета **спец.** Кстати, примеры показывают, что это слово характерно не только для профессионального жаргона, но и для неформального стиля:

*Налей водочки, Илюша, — просит Марфа. Я беру **розлив** в свои руки. Меня переполняет счастье [Ашот Аршакаян. Шведский дебют Ивана Денисовича // «Сибирские огни», 2012].*

*Замызанную подмосковную платформу с кафе-тошннловкой «Ветерок», где портвейн «Иверия» в **розлив** под яйца вкрутую, он назвал именем легендарного поселка старателей Клондайка [Алексей Моторов. Преступление доктора Паровозова (2013)].*

Пиво в розлив по кружкам, водка в розлив по стаканам, ну и портвешок на смазку [Леонид Хейфец. Всполохи // «Октябрь», 2013].

Спрос на недорогое пиво фактически сформировал новый сегмент рынка — продажу так называемого живого пива в розлив [Кира Ремнева. Забродивший продукт // «Однако», 2010].

Если же, наконец, обратиться к глаголу, то мы увидим просторечные формы настоящего времени у целого ряда глаголов:

*дарить, дарит — дорит;
катиться, катится — котится;
платить, платит — плотит;
растить, растит — ростит;
(по)садить, (по)садит — содит.*

...какой подарок вам подорит Бог когда вы попадете к нему? [<https://otvet.mail.ru> 2008].

И без перерыва за антисоветскими частушками следовали частушки «красных»: Эх, яблочко, куда **котишься**? К Колчаку **попадешь** — не **воротишься**! [И. М. Дьяконов. Книга воспоминаний. Глава третья (1926–1928) (1995)].

— *Мы ухаживаем, в магазин ходим, стираем, он нам **плотит** за это* [Владимир Войнович. Монументальная пропаганда // «Знамя», 2000].

*Отец-то меня **ростил** — это верно, а кто твою ораву **ростит**?* [Федор Абрамов. Две зимы и три лета (1968)].

[Эля Смирнова (жен)] *Как-то раз засиделся, метро не ходит, в трамвай не **содят*** [Л. С. Петрушевская. День рождения Смирновой (1977)].

*«У нас в селе один перевозил через речку, баркас знаешь какой у него был? Человек пятнадцать **посодит**, и еще место останется»* [Борис Хазанов. Я воскресение и жизнь (1976)].

Можно привести и пример лингвистической рефлексии по этому поводу на одном из форумов:

«Все говорят слова звОнит, плОтит, дОрит, брОется — все с ударением на первый слог. Ну это звонит, платит, дарит и бреется» [mujchin.net 2012].

Тут необходимо вернуться к [Трубецкой 1987: 99], где описывается нечто подобное:

«С. В смешанном (девербативно-именном, деноминативно-глагольном) основообразовании примеров с чередованием о ~ а оказывается уже несколько больше (в целом их 6), но все же они довольно разрозненны: наст. вр. plòt'ĭT (графич. платит)³⁵ ~ plàtā (плата), наст. вр. dòrĭT (графич. дарит)³⁵ ~ dar (дар), наст./буд. вр. pā|sòd'ĭT (графич. посадит)³⁵ — pā|sāTkā (посадка)...

В сноске 35 Н. С. Трубецкой отмечает:

«Произношение plàt'ĭT, dāfĭT, sād'ĭT совершенно чуждо москвичам (по крайней мере, моего поколения) и воспринимается либо как провинциализм, либо как нарочитое уподобление графическому облику; и напротив, пока еще употребляется Skāt'ĭT, а параллельная форма Skòt'ĭT считается просторечной».

Иначе говоря, ситуация для Н. С. Трубецкого в определенном смысле противоположна современной, поскольку, не [o], а именно [a] воспринимается как признак ненормативности, маркированности.

Трубецкой также говорит еще о трех парах, которые, впрочем, не имеют прямого отношения к рассматриваемым чередованиям:

«...прош. вр. м. р. *tfoS* (графич. тряс, впрочем, параллельно существует и произношение *tfaS*) ~ *třaŠkă* (тряска), также *ză|pfòK* (запряг) ~ *ză|pfăŠkă* (запряжка), *žaT'* (жать, наст. вр. *žmju*) ~ *žom* (жом)» [Трубецкой 1987: 99].

Важно подчеркнуть, что в современном русском языке просторечные формы типа *плотит* существуют наряду с литературными формами *платит*, где корень также находится под ударением, то есть здесь чередование оказывается единственным показателем нелитературности данного ряда форм.

Говоря о глаголе, нельзя не упомянуть и обособленное чередование, связанное с маркированной формой *нолито* (от *налить*). Оно уникально и тем, что замена на *о* происходит в приставке, имеющей единственный графический облик — *на*⁸, и тем, что оно обозначает особую «алкогольно-свойскую ситуацию», а именно возможность выпить, обеспеченную наполненностью сосудов. Ср.:

...парни, ну сколько можно, у нас давно нолито каждую секунду испаряется двести молекулярных слоев [Екатерина Завершнева. Высотка (2012)].

Сейчас все напишем. — У меня уже нолито. — Володя зубами открыл бутылку пепси и выплюнул пробку [Дмитрий Каралис. Дорогая Мирей Матье (1990)].

Экспрессивное *нолито* также существует наряду с «близнецом» *нали́то* из литературного языка.

Заключение

Итог этой статьи можно подвести разными способами.

Наиболее скромный состоит в констатации того, что в тексте приведена подборка примеров, иногда довольно специфических, которая демонстрирует явление, похожее на чередование букв в корнях и приставках, однако ранее пропущенное в школьных и академических грамматиках. Это явление неоднородно, тут нельзя говорить о какой-то единой фонетической, грамматической или семантической позиции, хотя можно выделить частные позиции и «аналогию», обеспечивающие скопление примеров. Например, для существительных речь идет о множественном числе слов женского рода акцентной парадигмы *d*: *грозá* — *грозы́*, *сосна́* — *со́сны*, *сноха́* — *сно́хи*..., а для глаголов это формы настоящего времени⁹, относящиеся к акцентной парадигме 4с: *ходить*.

Можно, однако, исходить и из более сильных исследовательских амбиций, а именно объединить ряд разнородных (в том числе и уникальных) явлений в рамках особого чередования, а конкретнее — перехода *а* в *о* в некодифицированной области языка: в профессиональных и молодежных жаргонах, просторечии, подчеркнуто разговорной речи.

Таким образом, можно говорить об относительно регулярных чередованиях *а* и *о* на стыке литературного и нелитературного языка. Буква *о* передает фонему [о], которая появляется под ударением в не вполне нормативных или неформальных словоформах и словах (просторечных, жаргонных...), то есть подчеркивает эту неформальность.

Возможно, точнее говорить не о регулярных чередованиях, а об определенной продуктивности модификации *а* в *о* с некой прагматической и стилистической целью. Появление

⁸ То есть приставка *на* отличается от приставки *рас*, где возможно и описано чередование *а* и *о*.

⁹ Точнее говоря, пять словоформ настоящего времени (кроме формы 1 лица единственного числа), где корень находится под ударением.

о вместо привычного *а* и фонетически, и графически маркирует неформальность слова. Кроме признака неформальности можно отметить и определенную экспрессивность форм с *о*, особенно заметную в том случае, когда в литературном языке существует полный грамматический аналог с *а*.

Возвращаясь к слову *шпаргалка*, с которого начиналась эта статья, можно сказать, что «жаргонность» формы слова *шпора* обеспечивается не только ударным [о], но и сокращением корня [šparga] до [špor], что также характерно для перехода слова в молодежный и другие жаргоны: *чел* от *человек* (молодежный жаргон) или *комп* от *компьютер* (компьютерный жаргон).

Результатом этого исследования можно считать не только описание некоего нового «странного» чередования, существование которого, скорее всего, вызовет дискуссию и будет подвергнуто сомнению, но и обнаружение определенных лакун в морфонологической теории и неточностей в формулировании школьных правил, предлагающих проверку безударных гласных. Наконец, это еще одно подтверждение важности установления границ между литературным языком и субстандартом и необходимости привлечения субстандартного языкового материала для формулирования общих языковых закономерностей.

СПИСОК ИСТОЧНИКОВ

- Большой вопрос.ру — Как проверить слово «шпаргалка»? [How to check the spelling of *shpargalka*?] <http://www.bolshoyvopros.ru/questions/2101583-kak-proverit-slovo-shpargalka.html#answer6487542>
- Грамота.ру — Правписание корней с чередующимися гласными [The spelling of roots with alternating vowels]. http://gramota.ru/class/coach/tbgramota/45_93
- Мел — Как правильно: «шпоргалка» или «шпаргалка» [*Shporgalka* or *shpargalka*: Which is correct?]. https://mel.fm/gramotnost/gramotny-otvet/1936408-cheat_sheet
- НКРЯ — Национальный корпус русского языка [Russian National Corpus]. <http://www.ruscorp.ru>

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Елистратов 2000 — Елистратов В. С. *Словарь русского арго: Материалы 1980–1990 гг.* М.: Русские словари, 2000. [Elistratov V. S. *Slovar' russkogo argo. Materialy 1980–1990 gg.* [A dictionary of Russian slang: Data from 1980–1990]. Moscow: Russkie slovari, 2000.]
- Зализняк 1987 — Зализняк А. А. *Грамматический словарь русского языка: Словоизменение.* 3-е изд. М.: Русский язык, 1987. [Zalizniak A. A. *Grammaticheskii slovar' russkogo yazyka: Slovoizmenenie* [A grammatical dictionary of Russian. Inflection]. 3rd edn. Moscow: Russkii yazyk, 1987.]
- Кронгауз и др. 2018 — Кронгауз М., Арутюнова Е., Панов Б. *Неучебник по русскому языку. Орфография.* Ч. 1: *О корнях и не только.* М.: Клевер-Медиа-Групп, 2018. [Krongauz M., Arutyunova E., Panov B. *Neuchebnik po russkomu yazyku. Orfografiya.* Ch. 1: *O kornyakh i ne tol'ko* [“Not-a-Textbook” of Russian. Spelling. P. 1: On roots and beyond]. Moscow: Klever-Media-Grupp, 2018.]
- МАС — *Словарь русского языка.* В 4 т. Евгеньева А. П. (ред.). М.: Русский язык, 1981–1984. [*Slovar' russkogo yazyka* [A dictionary of Russian]. In 4 vols. Evgen'eva A. P. (ed.). Moscow: Russkii yazyk, 1981–1984.]
- РГ 1980 — Шведова Н. Ю. (гл. ред.). *Русская грамматика.* В 2 т. М.: Наука, 1980. [Shvedova N. Yu. (ed.). *Russkaya grammatika* [Russian grammar]. In 2 vols. Moscow: Nauka, 1980.]
- Трубецкой 1987 — Трубецкой Н. С. *Морфонологическая система русского языка.* Трубецкой Н. С. *Избранные труды по филологии.* Виноградов В. А., Нерознак В. П. (сост.). М.: Прогресс, 1987. [Trubetzkoy N. S. *Morphonological system of Russian.* Trubetskoi N. S. *Izbrannye trudy po filologii.* Vinogradov V. A., Neroznak V. P. (comps.). Moscow: Progress, 1987.]

Принципы выделения местоименных серий (на материале адыгейских демонстративов)

© 2024

Юрий Александрович Ландер

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; yulander@yandex.ru

Шамсет Шабановна Унарокова

Адыгейский государственный университет, Майкоп, Россия; unarokova.shamset@yandex.ru

Аннотация: В статье обсуждается понятие «местоименная серия» и сложности, которые могут возникать при выделении местоименных серий в конкретных языках. Основным материалом послужила исключительно богатая и типологически интересная система демонстративных выражений адыгейского языка, в котором выделяются не менее одиннадцати серий демонстративов. Мы формулируем два принципа, которыми предлагается руководствоваться при определении местоименных серий, и затем, применив эти принципы на адыгейском материале, предлагаем еще несколько дополнительных положений, связанных с выделением серий.

Ключевые слова: абхазо-адыгские языки, адыгейский язык, дейксис, демонстративы, местоимения, морфология

Благодарности: Мы выражаем признательность И. Г. Багировой и Ф. Д. Унароковой за обсуждение материала, А. Д. Ландер, написавшей программу, позволяющую предварительно вычленить местоименные серии на материале корпуса адыгейского языка, анонимным рецензентам и редакторам за ценные комментарии к черновым вариантам статьи. За любые ошибки несем ответственность только мы. Исследование Ю. А. Ландера осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ.

Для цитирования: Ландер Ю. А., Унарокова Ш. Ш. Принципы выделения местоименных серий (на материале адыгейских демонстративов). *Вопросы языкознания*, 2024, 2: 93–121.

DOI: 10.31857/0373-658X.2024.2.93-121

Defining pronominal series: Demonstratives in West Circassian

Yury A. Lander

HSE University, Moscow, Russia; yulander@yandex.ru

Shamset Sh. Unarokova

Adyghe State University, Maykop, Russia; unarokova.shamset@yandex.ru

Abstract: We discuss the concept of “pronominal series” and some principles that can be used for distinguishing pronominal series from similar phenomena. These principles are further applied to demonstratives in West Circassian (a polysynthetic language of the West Caucasian family) as represented in a large corpus including texts of various genres. We show that West Circassian displays at least eleven series of demonstratives but also make additional observations on how the principles defining pronominal series can be used in complex cases.

Keywords: deixis, demonstratives, morphology, pronouns, West Caucasian, West Circassian

Acknowledgements: We are grateful to Irina Bagirokova, Anna D. Lander, Fatima Unarokova, two anonymous reviewers and editors for improving our understanding of the data and correcting its presentation. All errors are ours. Yury Lander's study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

For citation: Lander Yu. A., Unarokova Sh. Sh. Defining pronominal series: Demonstratives in West Circassian. *Voprosy Jazykoznanija*, 2024, 2: 93–121.

DOI: 10.31857/0373-658X.2024.2.93-121

1. Введение

В лингвистических описаниях нередко встречается понятие местоименной серии. Оно активно используется, например, при описании неопределенных местоимений, особенно после известной работы М. Хаспельмата [Haspelmath 1997], в которой он показал, как разные местоименные корни (вроде русских *кто*, *что*, *где* и т. д.) сочетаются с разными показателями (вроде русских *-то*, *-нибудь*, *ни-*). То же понятие появляется и в описаниях других местоимений, в том числе личных, как, например, в исследованиях по языкам майя [Nida, Moisés 1950] и манде [Vydrine 2006].

В этой работе мы обсудим понятие местоименной серии и возникающие при выделении серий сложности на примере указательных местоимений (демонстративов), хотя наши выводы применимы и к иным типам местоимений. Указательные местоимения ассоциируются в первую очередь с дейктической референцией к конкретным объектам, но выражения, образованные от указательных (демонстративных) основ, часто имеют и другие функции (см., в частности, [Himmelfmann 1996; 1997], а также сборники [Coniglio et al. (eds.) 2018; Næss et al. (eds.) 2020]). Самая известная из не собственно дейктических функций — анафорическая: нередко демонстративы используются как анафорические местоимения 3-го лица или развиваются в такие местоимения (см., например, [Heine, Song 2011]). Хотя некоторые функции (эндофорические, связанные с текстом, а не с контекстом речевого акта) можно описать как «внутритекстовый дейксис» (textual deixis) [Lyons 1977: 667–668] или «дискурсивный дейксис» [Levinson 2004: 118–119], мы следуем более широкому пониманию демонстративов как единиц, обеспечивающих совместное внимание (joint attention) адресата и адресанта [Diesel 2006] и при этом образующих закрытый класс, характеризующийся особым морфосинтаксическим поведением [Levinson 2018: 4].

Указательные основы могут образовывать единицы, описывающие не только собственно объекты (ср. *сей*, *тот*), но и качество (*так-сяк*, *такой-сякой*), место (*там-сям*) и т. д.; см. обсуждение функциональной и синтаксической типологии демонстративов в [Diesel 1999a; 1999b; Dixon 2003; König, Umbach 2018; Killian 2022]. В этом они аналогичны многим другим нелокуторным местоименным элементам, построение которых описывается как сочетание указания на функцию и указания на онтологический тип [Bhat 2004: 153]. Подобные сочетания регулярно формируют то, что описывается как местоименные серии, но насколько нам известно, точные критерии выделения серий отсутствуют, а авторы грамматик при описании соответствующих явлений опираются на традицию и/или собственную интуицию. Ниже мы попытаемся сформулировать некоторые критерии такого рода, учитывая при этом, что многие ситуации допускают разные описательные решения, и соответственно, намеренно оставляя некоторые вопросы, связанные с постулированием местоименных серий, без ответа.

Основным материалом для нас послужила исключительно богатая и типологически интересная система демонстративных выражений адыгейского языка, в котором, даже используя простейшие критерии, можно выделить не менее одиннадцати серий демонстративов. Впрочем, детальное описание сферы употребления любой из них достойно отдельного исследования, и мы не ставим такую задачу, ограничиваясь лишь набросками их семантики.

Работа построена следующим образом. В разделе 2 мы излагаем общий подход к выделению местоименных серий. В разделе 3 даются основные сведения о грамматике и указательных корнях адыгейского языка. Разделы 4–14 обсуждают потенциальные местоименные серии, включающие эти корни. В разделе 15 подводятся итоги выделения серий указательных местоимений в адыгейском языке. Наконец, раздел 16 возвращается к общим вопросам и формулирует положения, учитывающие выявленные сложности выделения серий на основе простейших критериев, сформулированных в разделе 2.

2. Формы, сочетания и серии

В качестве отправной точки возьмем следующее неформальное понимание: местоименные серии — наборы единиц с одной и той же функцией, образующихся от разных местоимений, причем в каноническом случае по одной и той же модели¹.

Такого определения, однако, недостаточно: исходя из него, отдельными сериями признавались бы, например, формы любого падежа. Но серии — не простое словоизменение. Во-первых, одни местоимения в серии могут быть представлены множеством форм (например, разные падежные формы местоимений *-нибудь*-серии, образованных от *кто* и *какой*), а другие — лишь одной (*-нибудь*-местоимение, образованное от *где*); см. таблицу 1. Во-вторых, серии местоимений — не всегда морфологическое явление. Сочетания местоименных форм с *угодно* в таблице 2 скорее являются единицами более высокого уровня, хотя описание их в соответствии с синтаксическими правилами тоже чревато сложностями (см., например, [Testelefs, Bylinina 2005]).

Таблица 1

Фрагмент описания форм серии *-нибудь*

<i>кто-нибудь</i>	<i>какой-нибудь</i>	<i>где-нибудь</i>
<i>кто-нибудь</i>	<i>какой-нибудь</i>	<i>где-нибудь</i>
<i>кого-нибудь</i>	<i>какого-нибудь</i>	
<i>кому-нибудь</i>	<i>какому-нибудь</i>	
...	...	

Таблица 2

Фрагмент описания серии *угодно*

<i>кто угодно</i>	<i>какой угодно</i>	<i>где угодно</i>
<i>кто угодно</i>	<i>какой угодно</i>	<i>где угодно</i>
<i>кого угодно</i>	<i>какого угодно</i>	
<i>кому угодно</i>	<i>какому угодно</i>	
...	...	

¹ Последняя оговорка связана с тем, что среди местоимений особенно распространен супплетивизм, поэтому отождествление разных способов образования порой сопряжено с трудностями. С другой стороны, можно допустить существование синонимичных местоименных серий, отличающихся моделями образования, — их противопоставление основано как раз на представлении о том, что канонически при выделении серии речь идет о паре «функция—модель образования». В последнем случае теоретически допустимо и объединение разных способов образования в единую модель: это естественно, если способы образования местоимений отличаются лишь незначительно (ср. наше описание адыгейских демонстративов образа действия в разделе 9).

Противопоставление серий каноническому словоизменению (см. о его типичных свойствах, например, [Перцов 2001: 74–90]) не представляет трудности. Местоимения (и шире проформы) часто заимствуют основные грамматические категории у других частей речи. При выделении местоименных серий можно отсечь словоизменение, доступное другим частям речи, оставив регулярное образование форм, свойственное исключительно местоимениям. Так, формант *-нибудь* образует серию, поскольку он не сочетается с неместоименными единицами. При этом, если в языке выражение какой-либо категории (например, падежа в английском языке) ограничено местоимениями, разные формы этой категории придется признать местоименными сериями. Реальная практика этому не противоречит: например, английские личные местоимения, в отличие от существительных различающие падежи, во многом аналогичны местоимениям языков манде, где грамматические отношения используются для выделения разных серий [Vydrine 2006]².

Таким же образом мы решаем проблему противопоставления серий обычным синтаксическим сочетаниям: если конструкция строится по правилам, не ограниченным местоимениями, трактовать ее как серию — значит создавать лишний элемент описания. Однако и здесь встречаются дискуссионные случаи. Например, в германских языках единицы, которые обычно функционируют как предлоги, в сочетаниях с некоторыми местоимениями выступают в качестве послелогов, причем орфография предписывает слитное написание подобных сочетаний; ср. англ. *herewith* ‘посредством этого’ (< *here* ‘здесь’, *with* ‘с’), нем. *damit* ‘тем самым’ (< *da* ‘там’, *mit* ‘с’) и т. д. Мы считаем возможным описание таких послелогов как образующих местоименные серии, хотя и допускаем, что, если исследователь на каком-то уровне отождествляет эту конструкцию и сочетания с предлогами, он может описывать ее иначе.

Впрочем, описывать через местоименные серии все особенности синтаксического поведения местоимений избыточно. Например, особенность германских образований, приведенных выше, представима и как свойство местоимений сочетаться с послелогом, а не с предлогом, в какой-то мере аналогичное, например, способности русских вопросительных местоимений предшествовать фокусным показателям вроде *именно* и *это*, а не следовать за ними, как это делают другие именные группы (ср. *Я ищу именно твою книгу* / *Это твою книгу я ищу* vs. *Что именно ты ищешь?* / *Что это ты ищешь?*). Отчасти эту проблему решает требование наличия у местоимений серии общей функции, хотя, конечно, и само понятие функции не является достаточно четким.

Фактически для нас принципиально, что серии — это особые сочетания морфем или слов, плохо соотносимые с обычными синтаксическими сочетаниями. Мы вводим следующее требование к выделению местоименных серий:

(1) Принцип исключительности:

В местоименные серии включаются сочетания, которые образованы по семантическим и/или морфосинтаксическим правилам, ассоциируемым исключительно с местоимениями.

Отдельная проблема связана с лексикализацией — появлением фиксированных сочетаний. С одной стороны, принцип исключительности не позволяет признать сериями даже сильно лексикализованные выражения (порою орфографически представляемые как единые слова), если они могут быть описаны регулярными правилами, не ограниченными местоимениями, при том что в силу частотности такие местоименные выражения могут восприниматься как фиксированные. С другой стороны, лексикализация и фразеологизация иногда приводят к появлению сложных форм, не сводимых к регулярным правилам

² Использование термина «серия» для форм падежа/числа/рода для языков, в которых те же категории выделяются и для неместоименных единиц, тоже встречается (см., например, [Беликов 1972]). Тем не менее очевидно, что в таких случаях бессмысленно говорить об особой местоименной серии.

и для местоимений, — своего рода «островков» в местоименных конструкциях (ср., например, русское фразеологическое выражение *тот еще*, не имеющее аналогов с демонстративом *этот*). При выделении серий необходимо исключить из рассмотрения такие нерегулярные явления. Соответственно, мы формулируем еще одно пожелание к выделению местоименных серий:

(2) Принцип регулярности:

В каноническом случае серия должна образовываться от всех основ местоимений или релевантного подкласса местоимений.

Продуктивность серий и естественность образуемых форм могут различаться. Поскольку наше исследование ориентировано в первую очередь на корпус, оно, с одной стороны, рассматривает и редкие формы, которые носители оценивают как маловероятные, а с другой стороны, может пропустить редкие формы, которые не зафиксированы в корпусе. В связи с этим далее мы различаем незафиксированные формы (отсутствующие в корпусе, но допускаемые опрошенными нами носителями; в приводимых парадигмах такие формы даются в квадратных скобках) и недопустимые формы (встретившиеся в корпусе, но оцениваемые носителями как невозможные; в парадигмах помечаются тремя знаками вопроса).

Несмотря на сложности, возникающие при исследовании, ориентированном на корпус, принципы исключительности и регулярности кажутся соответствующими практике описания местоименных систем. Однако материал адыгейского языка показывает, что опора исключительно на них может дать достаточно пеструю картину.

3. Базовые сведения об адыгейской грамматике и адыгейских демонстративах

Адыгейский язык вместе с близкородственным кабардино-черкесским языком образует адыгскую/черкесскую ветвь абхазо-адыгской (западнокавказской) семьи. Описан адыгейский сравнительно неплохо: у него есть несколько подробных общих грамматических описаний [Яковлев, Ашхамаф 1941; Рогава, Керашева 1966; Smeets 1984; Paris 1989; Зекох 2002; Зекох и др. 2003; Paris 1989; Аркадьев и др. 2009; Konuk 2022], помимо большого числа монографий и статей по отдельным темам.

Типологически адыгейский язык описывается как агглютинативный и полисинтетический, т. е. легко образующий исключительно сложные морфологические выражения (как префиксацией, так и суффиксацией, реже — апофонией), а также характеризующийся другими признаками, ассоциируемыми с полисинтетизмом [Кумахов 1964; Lander, Testelelets 2017]. С агглютинативностью в сочетании с чрезвычайной продуктивностью аффиксов (что свойственно и другим полисинтетическим языкам [de Reuse 2009]), возможно, связано слабое противопоставление знаменательных частей речи: одни и те же основы присоединяют морфологические показатели сказуемого, аргумента и обстоятельства [Lander, Testelelets 2017; Bagirokova, Lander 2019]. В качестве ядра именной группы часто выступает сложный именной комплекс (обычно комбинация имени и некоторых его определений), который по многим параметрам ведет себя как единое слово [Lander 2017]. Порядок слов преимущественно левоветвящийся (зависимые элементы предшествуют главным), хотя допускается и следование зависимых за сказуемым.

И в морфологии, и в синтаксисе (но не в порядке слов) адыгейский проявляет более или менее последовательную эргативность [Гишев 1985; Lander 2009; Letuchiy 2012; Ershova 2019] (но см. [Аркадьев и др. 2009: 80–83; Lander et al. 2021] о некоторых исключениях). Грамматические отношения выражаются как с помощью вершинного маркирования (так называемых местоименных аффиксов в синтаксических вершинах: сказуемых, вершинах

именных и послеложных групп), так и с помощью зависимостного маркирования (падежей). Индексация участников местоименными аффиксами затрагивает эргативный и абсолютивный аргументы, неотторжимый посессор, объект послелога, а также не прямые объекты, роль которых обычно уточняется сопровождающим их аппликативным показателем (бенефактивным, комитативным, посессивным, одним из многочисленных локативных «превербов» и т. д.). Традиционно выделяется четыре падежа, показатели которых появляются на вершинах именных групп:

- абсолютивный (в других описаниях «именительный»), маркирующий непереходный субъект и переходный пациент, а также иногда вынесенный топик;
- косвенный («эргативный»), маркирующий в основном именные группы, описывающие участников, которые проиндексированы местоименными префиксами в вершине конструкции (в сказуемом, в вершине посессивной конструкции, в послелоге), регулярно встречается как на группах, указывающих на основных участников ситуации, так и на группах, описывающих время и место;
- инструментальный («творительный»), имеющий множество функций, в том числе локативных; может присоединяться к показателю косвенного падежа [Зекон 1969; Сердобольская, Кузнецова 2009; Serdobolskaya 2011];
- адвербиальный («превратительный»), маркирующий, в частности, вторичную предикацию [Vydrin 2008] и предикатные актаны при матричных предикатах вроде ‘становиться (кем-то)’, ‘считать (кем-то)’ и т. д.

Мы будем называть абсолютивный и косвенный падежи ядерными, а инструментальный и адвербиальный — периферийными. Показатели ядерных падежей могут отсутствовать в личных местоимениях, именах собственных, посессивах (индексирующих посессора вершинах приименных посессивных конструкций) и нереферентных формах; последние к тому же не определены в отношении числа [Arkadiev, Testelets 2019; Bagirokova et al. 2022]. Базовая именная парадигма проиллюстрирована в таблице 3³.

Таблица 3

Базовая именная парадигма существительного *матэ* ‘корзина’

	Нереферентная форма	Единственное число	Множественное число
Абсолютивный	<i>матэ</i> <i>mate</i>	<i>матэ-р</i> <i>mate-r</i>	<i>матэ-хэ-р</i> <i>mate-xe-r</i>
Косвенный	<i>матэ</i> <i>mate</i>	<i>матэ-м</i> <i>mate-m</i>	<i>матэ-хэ-м, матэ-мэ, матэ-хэ-мэ</i> <i>mate-xe-m, mate-me, mate-xe-me</i>
Инструментальный	<i>матэ-клэ</i> <i>mate-č'e</i>	<i>матэ-м-клэ</i> <i>mate-m-č'e</i>	<i>матэ-хэ-м-клэ</i> <i>mate-xe-m-č'e</i>
Адвербиальный	<i>мат-эу</i> <i>mat-ew</i>	<i>мат-эу</i> <i>mat-ew</i>	<i>матэ-х-эу</i> <i>mate-x-ew</i>

В современном адыгейском языке выделяется четыре основных корня, образующих указательные местоимения: *тэ*, *тве*, *а*, *ž'a* (в двух адыгейских диалектах — бжедугском и шапсугском — вместо *тве* используется указательное местоимение *wə/we*)⁴.

³ Представляя адыгейские данные, мы используем традиционную кавказоведческую транскрипцию (см. [Коряков 2006]). В таблицах и примерах приводится параллельное написание в адыгейской орфографии.

⁴ Еще один кандидат в дейктические указательные местоимения — *ž'a* ‘теперь’. Это слово действительно обладает некоторыми свойствами демонстративов. Так, в одном фольклорном тексте в корпусе зафиксирована форма косвенного падежа *ž'a-š'*. Кроме того, от *ž'a* образуется

Семантика этих корней изучена недостаточно. Большая часть описаний адыгейского языка вообще касается лишь трех указательных местоимений: *тə*, *тwе* и *а*. Традиционно утверждается, что они различаются степенью удаленности объекта от говорящего. Например, Н. Ф. Яковлев и Д. А. Ашхамаф [1941: 289] описывают образуемый ими контраст как противопоставление между обозначениями близкого видимого, далекого видимого и невидимого соответственно; таким же образом представляет тройку *тə* — *тwе* — *а* К. Парис в своем очерке абадзехского диалекта [Paris 1989: 177]. Однако уже М. А. Кумахов [1964: 206] отметил, что, хотя эти местоимения противопоставляются по признаку «близкий — далекий», «[м]естоимение *а* указывает на видимый или находящийся поблизости предмет, или на невидимый, но известный, определенный предмет». Согласно грамматике Г. Рогавы и З. И. Керашевой [1966: 85], *тə* указывает на объект около говорящего, *тwе* — на объект несколько дальше от говорящего (причем исторически, по мнению авторов, оно указывало на объект около слушающего), *а* — на объект, расположенный «дальше от первого и второго лиц». Наконец, согласно экспериментальному исследованию Д. В. Барыльниковой [2013], посвященному одному из говоров темиргоевского диалекта, *тə* указывает на близость к говорящему, *тwе* — на доступность (но не близость) и в основном употреблении часто сопровождается указательным жестом, *ǰ'a* — на отсутствие видимости говорящему, а *а* преимущественно употребляется анафорически.

Ни в одной из упомянутых работ не приводится инвентарь единиц, образуемых упомянутыми выше корнями. Наша работа призвана восполнить этот пробел, используя данные Адыгейского корпуса [Архангельский и др. 2018–2023], включающего разнообразные тексты на адыгейском языке.

4. Простые употребления демонстративов

Под простыми употреблениями демонстративов мы понимаем использование форм, включающих демонстративные корни, но не «модифицирующую» их морфологию, либо при имени, либо в сочетании с показателями числа и ядерных падежей. Объединение этих двух типов употреблений противоречит проводимому в литературе жесткому противопоставлению демонстративов-детерминаторов (приименных демонстративов) и демонстративов-местоимений (самостоятельных демонстративов, образующих именные группы); см. [Diessel 1999a; 1999b]. Следуя этому противопоставлению, приименные и самостоятельные местоимения в адыгейском языке пришлось бы выделить в разные серии, что противоречит интуиции: естественно считать появление именных категорий при самостоятельном употреблении местоимения следствием того, что выражение образует полную именную группу (как это делается в отношении существительных), которая и принимает соответствующие показатели, а не того, что перед нами особый класс демонстративов.

В приименном употреблении⁵, проиллюстрированном для *тə* и *а* в (3)–(4), демонстративы не сочетаются с какими-либо грамматическими показателями и, по-видимому, представляют собой морфологически связанные элементы. Иногда они даже описываются — исторически [Яковлев, Ашхамаф 1941: 290] или синхронно [Dumézil 1932: 34; Зекох 2002: 164; Зекох и др. 2003: 235] — как префиксы. Этому, однако, противоречит то, что они могут быть оторваны от имени (4).

предикативная форма *ǰ'a-rə* 'это теперь', аналогичная предикативным формам указательных и личных местоимений (см. о них раздел 5). Однако в силу нерегулярности этих выражений далее корень *ǰ'a* не рассматривается.

⁵ Строго говоря, демонстративы могут выступать в качестве модификаторов не только при именах, но и при сказуемых относительных клаузах в отсутствие внешней вершины [Ландер 2012].

- (3) *Мы гъогум шъутетэу занклэу шъучъ.* [Цукько Нэфсэт. Гъулацый]
 мэ к'wэг'э-м ш'wэ-тје-т-ew занч'ew ш'wэ-ч
 DEM дорога-OBL 2PL.ABS-LOC-стоять-ADV прямой-ADV 2PL.ABS-бежать
 'По этой дороге (букв. «находясь на этой дороге») езжайте прямо'.
- (4) *⟨...⟩ а уздаклорэ нюом кышлэцтэн.* [Бабыхъуко Щорэ-цлык'урэ Нарт Хъашхъ-онэжърэ]
 а wэ-z-da-k'we-re нэwe-м q-э-шэ-с'т-ер
 DEM 2SG.ABS-REL.IO-LOC-идти-DYN бабушка-OBL DIR-3SG.ERG-знать-FUT-NEG
 '⟨...⟩ та бабушка, к которой ты идёшь, не узнает это'.

Образуя самостоятельные именные группы, указательные местоимения присоединяют показатели падежей (5), а также могут принимать морфологию множественного числа — суффикс *-хе* и маркер косвенного падежа множественного числа *-те* (последний в указательных местоимениях, но не в существительных требует и присутствия *-хе*) (6)⁶. Формы ядерных падежей адыгейских демонстративов показаны в таблице 4. Самостоятельные указательные местоимения единственного числа получают суффикс косвенного падежа *-š'/-j*, отличный от суффикса существительных *-т* и обнаруживаемый также у личных местоимений при послелогах; ср. постулирование отдельного «сравнительного падежа» в [Зекох 1969: 66; 2002: 94]⁷.

- (5) *Адыгэлыба мыр?! [Цуяко Юныс. Къэзгъэзжърэ ижърэ шыухэр]*
 adəge-λə-ba мэ-г
 адыг-мужчина-EMP DEM-ABS
 'Разве этот не адыгский мужчина?!'
- (6) *Пцэкъэнтфыклэ хъун ахэлыба мыхэмэ? [Къуекко Налбий. Псым ыхъырэ луашхъ. Едзыгъуитг'у хъурэ драм]*
 pceqentfə-č'e χ'wə-n а-хе-λ-ba мэ-хе-me
 удочка-новый случаться-MOD 3PL.IO-LOC-лежать-EMP DEM-PL-OBL.PL
 'Разве среди них не найдется новой удочки?'

⁶ Как показано в [Bagirokova et al. 2022: 296–297] (см. также [Ситимова 2004: 77–78]), на указательных местоимениях — в отличие от групп, образованных существительными, — маркирование множественности факультативно.

⁷ Фольклорный подкорпус включает также употребления демонстративов с алломорфом показателя косвенного падежа *-š'ə* (i), в единичных случаях даже в сочетании с именным суффиксом косвенного падежа *-т* (ii) (а также, несколько чаще, с суффиксом косвенного падежа множественного числа *-те*):

- (i) *Джы дэдэм о мыцы ыплъэгъугъэр зэк'э а нью улы баджэм ил'шлагъ.* [Саусырыко ил'лак]
 ž'ə-dede-m we мэ-š'ə š'ə-p-лэк'wə-ке-г zeč'e
 теперь-именно-OBL ты DEM-OBL LOC-2SG.ERG-видеть-PST-ABS все
 а нэwe-wədə-bzaž'e-m jə-ʔe-ša-к
 DEM старуха-ведьма-зловредный-OBL POSS-рука-делать-PST
 'Все то, что ты сейчас здесь увидел(а), дело рук злой колдуньи'.
- (ii) *Ацым лофыр зынэсыджэ плакъохэр пил'оклодыцтых.* [Шхъак'лумэдэ Ким. Псэйт'уку кл'энэкл'элъэ орэд]
 а-š'ə-м ʔ'wefə-r зə-nesə-ž'e p-лаq'we-хе-г
 DEM-OBL-OBL дело-ABS REL.TMP-достигать-INS 2SG.PR-нога-PL-ABS
 p-š'we-k'wedə-š'tə-x
 2SG.IO-MAL-исчезать-FUT-PL
 'Если дело дойдет до этого, ты потеряешь ноги'.

Ср. также форму *ʔa-šə(-m)-g'e* [DEM-OBL(-OBL)-INS], упоминаемую в [Smeets 1992: 307] для адыгейского говора шапсугов Дюздже (Турция).

Таблица 4

Формы ядерных падежей указательных местоимений

Число	Падеж	мы тэ	мо mwe	джа ǰ'a	а а
Немаркированные формы	Абсолютив	мы-р тэ-г	мо-р mwe-г	джа-р ǰ'a-г	а-р а-г
	Косвенный	мы-иц(ы), мы-й тэ-š'(э), тэ-j	мо-иц(ы), мо-й mwe-š'(э), mwe-j	джа-иц(ы), джа-й ǰ'a-š'(э), ǰ'a-j	а-иц(ы), а-й а-š'(э), а-j
Множественное число	Абсолютив	мы-хэ-р тэ-хе-г	мо-хэ-р mwe-хе-г	джа-хэ-р ǰ'a-хе-г	а-хэ-р а-хе-г
	Косвенный	мы-хэ-м, мы-хэ-мэ тэ-хе-м, тэ-хе-me	мо-хэ-м, мо-хэ-мэ mwe-хе-м, mwe-хе-me	джа-хэ-м, джа-хэ-мэ ǰ'a-хе-м, ǰ'a-хе-me	а-хэ-м, а-хэ-мэ а-хе-м, а-хе-me

Употребление ядерных падежных форм указательных местоимений в целом соответствует употреблению именных форм. Отметим только, что косвенные формы демонстративов, как и обычные косвенные именные группы, могут описывать место:

- (7) *Мыцц ылэ дахэ ыцыфаусыгь — Маркиза.* [Матыжева Аминэт. Маркизэ]
тэ-š' **се-дахэ** **š'ə-f-a-wəsa-в** **mark'əza**
 DEM-OBL имя-красивое LOC-BEN-3PL.ERG-сочинять-PST Маркиза
 'Здесь её назвали красивым именем — Маркиза'.

Хотя в [Рогава, Керашева 1966: 332] такие формы представлены как наречия места, при их появлении место обычно вводится в сказуемом с помощью локативных аппликативов (в приведенном примере — с помощью общелокативного аппликатива *š'ə-*), то есть ведет себя как участник ситуации. В то же время, как мы увидим в разделах 7 и 12, в этой функции простые косвенные формы демонстративов конкурируют с представителями других демонстративных серий.

5. Предикативные употребления демонстративов и предикативная серия

Адыгейские местоимения не могут выступать в функции сказуемого в немаркированной форме. Определенные местоимения (как личные, так и указательные) в функции сказуемого принимают особый предикативный показатель *-рə* (см. обсуждение в [Smeets 1992]); ср. (8). Предикативные формы указательных местоимений *тэ-* и *мо-* могут при этом вводить в рассмотрение новых участников (9) подобно указательным идентификаторам (см. раздел 6).

- (8) *Ары, ылуагь зы купым. Хьау, арэн, аиц фэд нылэн, алуагь нэмыкхэм. Сэры ар, ылуагь ежьым.* [Библия (перевод)]
а-гə, **ə-ʔ^{wa}-в** **зə** **к^{wə}рə-м.**
 DEM-PRED 3SG.ERG-говорить-PST один группа-OBL

haw, a-r-ep, a-š' fe.d nә.ʔep, a-ʔwa-к
 нет DEM-PRED-NEG DEM-OBL подобный лишь 3PL.ERG-говорить-PST
 nemәč' -xe-m. se-rә a-r, ә-ʔwa-к jež' ә-m
 другой-PL-OBL я-PRED DEM-ABS 3SG.ERG-говорить-PST сам-OBL

‘Иные говорили: это он, а иные: нет, не он, только похож на него. Он же говорил: это я’.

- (9) *Мыры мыщ тетхагъэр — «Мэрэтыкъу, 1962-рә илъэс».* [Адыгэ макъ]

mә-rә мә-š' tje-txa-ke-r «meretaq', 1962-re jәles»
 DEM-PRED DEM-OBL LOC-писать-PST-ABS Меретук 1962-ADJ год

‘Вот, что здесь написано: «Меретук, 1962 год»’.

Предикативная форма *arә*, образованная от корня *a-*, функционирует как связка в предложениях тождества [Сумбатова 2009] (10). Мы рассматриваем связку наравне с другими предикативными формами: подобно им, она способна образовывать сказуемое и самостоятельно, поскольку вводимая ей часть всегда может быть опущена (например, если в (10) опустить *webzek' әr* ‘топор’, мы получим предложение со сказуемым ‘это то’).

- (10) *lašә nlonәу алыгъ закъор — обзэгъур ары.* [Шәшә Аслән. Аслән плъыжъ, Аслән джәф]

ʔaše p-ʔwe-n-ew a-ʔәк zaq'we-r
 оружие 2SG.ERG-говорить-MOD-ADV 3PL.ERG-держатъ единственный-ABS
 webzek' ә-r a-rә
 топор-ABS DEM-PRED

‘А вместо оружия у них топор’. (букв. «Единственное, что они держат как, мол, оружие, — это топор».)

Помимо простых демонстративных форм с предикативным суффиксом в текстах регулярно наблюдаются и орфографические словоформы типа *мырары*, *морары*, *джарары*, *арары* (11), в том числе с грамматическими показателями. По-видимому, эти формы представляют собой сочетания указательного местоимения в абсолютном падеже и связки⁸:

- (11) *Лъапсәу иләр джарары.* [Адыгэ макъ]

laps-ew jә-ʔe-r ʔ'a-r-a-rә
 корень-ADV POSS-быть-ABS DEM-ABS-DEM-PRED

‘Вот какой корень имеет’. (букв. «Корень, который у него есть, — это тот».)

Поскольку эти конструкции композициональны, а их синтаксическая структура очевидна, оснований постулировать отдельную серию указательных местоимений для них нет, несмотря на слитное написание. Впрочем, в разделе 6 мы увидим, что в двух случаях такие местоимения все же развились в отдельную серию.

⁸ Возможный контраргумент состоит в том, что встречаются примеры, где предполагаемая связка *arә* присоединяется к личному местоимению с дополнительным согласным *-r*, при том что абсолютный показатель *-r* к личным местоимениям не присоединяется:

- (i) *О слыфые-слыфыц зыфалорәр оары!* [Ушгый Клыщыкъу. Нарт Тәтәршәо натыфрә хьо блатхьэрә кьызэрхыгъэр]

we şəfəje~şəfəš' zә-f-a-ʔwe-re-r we-r-a-rә
 ты ужасный.человек REL.IO-BEN-3PL.ERG-говорить-DYN-ABS ты-ʔ-DEM-PRED

‘Тот, о ком говорят «ужасный человек», — это ты!’

Однако подобные примеры — не норма. Обычно в таких контекстах предикативный суффикс присоединяется к личным местоимениям напрямую. Возможно, форма *werarә* возникает в результате применения правила, вставляющего *-r-* в интервокальной позиции.

Предикативные формы присоединяют основные суффиксальные показатели сказуемых (времени, подчинения, отрицания) и часть префиксальной морфологии (например, префиксальное отрицание, присоединение которого, однако, связано с морфологически нерегулярными формами; ср. краткое обсуждение в [Герасимов 2010]). Простые предикативные формы адыгейских демонстративов показаны в таблице 5.

Таблица 5

Простые предикативные формы указательных местоимений

	<i>мы</i> тэ	<i>мо</i> mwе	<i>джа</i> ž'a	<i>а</i> а
Базовая предикативная форма	<i>мыры</i> тэгэ	<i>моры</i> mwегэ	<i>джары</i> ž'агэ	<i>ары</i> агэ

Аналогично образуются формы личных местоимений вроде *se-rə* [я-PRED] 'это я'.

6. Указательные идентификаторы

Помимо форм, упомянутых в разделе 5, встречаются еще два местоимения, содержащие предикативный суффикс *-rə*: местоимения *marə* и *mwaraə*. Они используются как указательные идентификаторы, соответствующие (по крайней мере на первый взгляд) русским *вот* и *вон*, и даже могут функционировать в качестве так называемых сентенциальных демонстративов [Fillmore 1982: 47], вводя целые клаузы (12)–(14). Согласно грамматике [Рогова, Керашева 1966: 87], *marə* происходит из *mərarə*, которое встречается в аналогичной функции (ср. (28) в разделе 10); *mwaraə* в таком случае естественно возводить к *mwərarə*. Тем не менее синхронно это малочленимые формы с особой функцией.

- (12) *Мары темэу кыгьгэнэфэгьэхэр...* [Адыгэ макъ]

ma.rə tjem-ew q-ə-ʁe-nəfa-ʁe-xe-r
 вот тема-ADV DIR-3SG.ERG-CAUS-ясный-PST-PL-ABS
 'Вот тема, которую он обозначил...'

- (13) *Мары Барони мэгулэжьы.* [Цуякьо Юныс. Къэзгъэзэжьрэ ижьырэ шыухэр]

ma.rə baron-jə me-g^wəʔe-ž'ə
 вот Барон-ADD DYN-переживать-RE
 'Вот и Барон переживает'.

- (14) *Муары джыри зыгорэ изакъоу къэчъэ.* [Библия (перевод)]

mwa.rə ž'ə-tjə zə-g^were jə-zaq^w-ew q-e-ʁe
 вон теперь-ADD один-некий POSS-единственный-ADV DIR-DYN-бежать
 'Вон еще бежит один человек'.

Как указательный идентификатор используется и *ž'arə* (от демонстратива *ž'a*). В этом случае можно говорить о форме, омонимичной простой предикативной форме, но образованной подобно *marə* и *mwaraə*. Все вместе они, соответственно, должны описываться как отдельная серия (таблица 6). Дополнительный аргумент в пользу такого решения состоит в том, что эти местоимения морфологически ущербны: никакие другие морфологические формы от них не образуются.

Таблица 6

Указательные идентификаторы				
	мы тэ	мо mwe	джа ǰ'a	а а
Указательный идентификатор	мары maɾə	муары mwaɾə	джары ǰ'aɾə	—

7. Инструменталис
и инструментальная серия

Показатель инструментального падежа -ǰ'e, присоединяясь к референтным именным группам (кроме личных местоимений, имен собственных и посессивов единственного числа), следует за показателем косвенного падежа; ср. *waʂ^we-m-ǰ'e* [небо-OBL-INS] 'к небу'. Поскольку группы, образуемые указательными местоимениями, всегда являются семантически определенными, при присоединении показателя инструменталиса на них ожидаем и суффикс косвенного падежа; ср. форму единственного числа в (15) и форму множественного числа в (16):

(15) *Джащкǰэ цуакъэ цэфы*. [Къуекъо Налбий. Перестройкэр къэкǰуа къэмыкǰуа?]
ǰ'a-ǰ'-ǰ'e c^waqe ʂ'efə
DEM-OBL-INS обувь покупать
'На это купи обувь'.

(16) *Сыда мыхэмкǰэ тхылъеджэхэм япломэ нилонгъуагъэр?* [Адыгэ макъ]
səd-a mə-xe-m-ǰ'e txəl-je.ǰ'e-xe-m ja-p-ɽ^we-me
что-Q DEM-PL-OBL-INS книга-читать-PL-OBL 3PL.IO+DAT-2SG.ERG-говорить-COND
p-ǰ^we-jə.k^wa-ke-r
2SG.IO-MAL-время-PST-ABS
'Что ты этими вещами хотел донести до читателя?'

Тем не менее регулярно встречаются и демонстративы, представляющие собой простые сочетания корня указательного корня и инструментального суффикса:

(17) *Мокǰэ шукъэтлыс*. [Къуекъо Налбий. Псым ыхъырэ ǰуашъхъ]
mwe-ǰ'e ʂ^wə-qe-ɬəs
DEM-INS 2PL.ABS-DIR-садиться
'Сядьте туда'.

В отличие от демонстративов в инструменталисе с косвенным показателем, эти местоимения используются исключительно в локативной функции — в первую очередь для выражения направления и местонахождения. Кроме того, подобные указательные местоимения не допускают выражения множественного числа⁹. Эти ограничения отличают их от обычных именных групп (в том числе образованных указательными местоимениями, ср. раздел 4). Соответственно, мы постулируем отдельную инструментальную серию с локативным значением; см. таблицу 7.

⁹ Один из рецензентов предположил, что место здесь понимается не как конкретный референт, а как более размытый, менее жестко заданный участок пространства, с чем и связано отсутствие падежа (также маркирующего референтность) и числа.

Таблица 7

Демонстративы с инструментальным показателем

	мы mә	мо mwe	джа ǰ'a	а a
Простые инструментальные формы	мыцк/э, [мыйк/э] mәs'č'e, [mәjč'e]	моцк/э, [мойк/э] mweš'č'e, [mwejč'e]	джацк/э ǰ'aš'č'e	ацк/э, айк/э aš'č'e, ajč'e
Инструментальная локативная серия	мык/э mәč'e	мок/э mweč'e	джак/э ǰ'ač'e	ак/э ač'e

В разговорной речи косвенный падеж окказионально отсутствует и в других функциях вроде *a-č'e wә-mә-š* [DEM-INS 2SG.ERG-NEG-делать] 'Этим не делай!' В отличие от локативных употреблений, такое использование оценивается как ненормативное: речь идет об опущении падежного показателя, свойственном разговорной речи, в то время как в стилистически нейтральном контексте (17) представлена форма, для которой восстановление косвенного падежа не требуется.

Показатель инструменталиса появляется также на сказуемых некоторых сентенциальных актантов [Сердобольская, Мотлохов 2009], например при предикате *je-neg^wja-* 'предполагать' и его производных. В таком употреблении суффикс *-č'e* присоединяется к предикативным формам демонстративов, причем предикативный маркер *-rә* регулярно усекается до *-r* (ср. (18)). Эти формы — результат присоединения регулярной морфологии сказуемого к предикативным формам, и постулирование отдельной серии для них не требуется.

(18) *Енэгуягъо къэбзэныгъэр арыклэ.* [Адыгэ макъ]

je.neg^wja-кwe qebze.nә.кe-r a-rә-č'e
предполагать-NTIME чистота-ABS DEM-PRED-INS

'Я считаю, что дело в чистоте'. (букв. «Вероятно, что чистота — это то».)

8. Адвербиалис и неосложненная адвербиальная серия

Адвербиальный показатель *-ew* в сочетании с именами и местоимениями во всех грамматических описаниях описывается как показатель отдельного падежа (см. историю вопроса в [Зехох 1969]), хотя тот же суффикс образует и типичные наречия образа действия (ср. *psәnč'-ew* 'быстро' < *psәnč'e* 'быстрый'), и так называемые деепричастные формы. Адвербиальный маркер присоединяется к демонстративным корням напрямую или появляется на их предикативных вариантах; см. таблицу 8 и примеры (19)–(20). Возникающие таким образом формы мы будем называть неосложненными, в отличие от сходных с ними адвербиальных демонстративов образа действия, содержащих специальный показатель (см. раздел 9).

Неосложненные формы *ǰ'a-w* и *a-w* образуются нестандартно: у адвербиального показателя *-ew* в сочетании с именами нет алломорфа *-w*. В то же время указательные основы — единственные незаимствованные основы, которые можно описать как исходно (до применения морфонологических правил) содержащие конечный гласный /a/, и появление алломорфа *-w* скорее всего связано именно с этим.

(19) *Куп цыклой джау тпэмчыжъэу, къэхэлъэ гъунэм дэжъ цытых.* [Цуякъо Юныс. Къэзгъэзжърэ ижырэ шыухэр]

к'wәр-сәк'-ew ǰ'a-w t-pe-mә-č'әž'-ew,
группа-маленький-ADV DEM-ADV 1PL.IO-LOC-NEG-далекий-ADV

qexelɛ-ʷəne-m deʒ' š'ə-tə-x
кладбище-граница-OVL возле LOC-стоять-PL

‘Они стояли маленькой группой там, недалеко от нас, у края кладбища’.

(20) *Сыда мырэу узфэчэфынчээр?* [Щэшлэ Аслъан. Сае зыщыгъ нысхъап]

səd-a mǝ-r-ew wə-z-fe-č'efǝ-nče-r
что-Q DEM-PRED-ADV 2SG.ABS-REL.IO-BEN-веселый-CAR-ABS

‘Что же ты так дуешься?’ (букв. «Что есть то, из-за чего ты так невесел?»)

Таблица 8

Неосложненные адвербиальные демонстративы

	мы mǝ	мо mwe	джа ʒ'a	а a
Неосложненная адвербиальная серия	мэу mew	моу mwew	джау ʒ'aw	ау aw
Адвербиальные предикативные формы	мырэу mǝrew	морэу mwereu	джарэу ʒ'areu	арэу areu

Неосложненные адвербиальные демонстративы без предикативного суффикса иногда используются в локативном значении, не типичном для адвербиальных форм; ср. (19) выше. Кроме того, у *mweu* и *aw* развились новые функции, намного более частотные: *mwew* используется как частица призыва ‘ну-ка’, а *aw* — как адверсативный союз ‘но’ (см. [Степанянц 2023] о развитии таких союзов):

(21) *Ау, моу къысэдэлү, сипиъашъ, мы лофыр къызэрыкло дэдэу ыцтэн.* [Щэшлэ Аслъан. Аслъан плъыжъ, Аслъан джэф]

aw, mwew qə-s-e-deʔw, s-jə-pšaš, mǝ ʔwefǝ-r
но ну.ка DIR-1SG.IO-DAT-слушать 1SG.IO-POSS-девушка DEM дело-ABS
qə.ze.rǝ.kʷe-ded-ew š'ə-t-ep
простой-очень-ADV LOC-стоять-NEG

‘Но послушай, дочка, это дело не очень простое’.

В силу относительной нерегулярности образования и некомпозициональности можно считать формы без предикативного суффикса особой серией. Напротив, формы, включающие предикативный суффикс, — формы предикативной серии, поскольку их функции аналогичны формам обычных адвербиальных сказуемых (‘так’ < ‘будучи тем/этим’).

Неосложненные формы регулярно наблюдаются в текстах со своего рода вспомогательным стативным предикатом *š'ə-t-* ‘стоять’ (с общелокативным провербом *š'ə-*). В таком употреблении они выполняют нелокативную функцию, замещая образ действия или даже целую ситуацию:

(22) *Арышъ, зэхэтицэрэ мэфэкхэр ау ыцтми реклохъху nlon плъэклыцтэн.* [Адыгэ макъ]

a-rǝ-š, ze-xe-t-š'e-re mefeč' -xe-r a-w š'ə-t-m-jǝ
DEM-PRED-CS REC.IO-LOC-1PL.ERG-вести-DYN праздник-PL-ABS DEM-ADV LOC-стоять-COND-ADD
r-je-kʷe-č'ə-x-ew p-ʔwə-n p-лeč'ə-š't-ep
LOC-DAT-идти-выходить-PL-ADV 2SG.ERG-говорить-MOD 2SG.ERG-мочь-FUT-NEG

‘Итак, нельзя сказать, что праздники, которые мы устраиваем, несмотря на это (букв. «хотя так стоят»), проходят’.

В следующем разделе показано, что из этой конструкции развилась дополнительная серия демонстративов, которая, хотя в большинстве случаев и конкурирует с адвербиальными

формами (и выигрывает эту конкуренцию), тем не менее, не ограничена адвербиальными контекстами.

9. Демонстративы образа действия

Подвергшись дальнейшей грамматикализации, статив *š'ə-t-* 'стоять' в сочетаниях с демонстративами стал формантом *-š't*, образующим демонстративы образа действия, см. таблицу 9. Предикативный суффикс *-rə* в таких демонстративах может как присутствовать (23), так и отсутствовать (24). Зафиксированы и примеры, в которых адвербиальный маркер не вычленяется (25), хотя от него остается след в виде гласного /e/. Все такие местоимения отсылают к характеристике ситуации.

- (23) *Аццыц горэм мырэуцтэу кьело джы* {...}. [Щэшлэ Аслъан. Аслъан плтыжъ, Аслъан джэф]
a-š'ə-š'-g^{were}-m **mə-r-ew-š't-ew**
 3PL.IO-LOC-быть.частью-некий-OBL DEM-PRED-ADV-PMAN-ADV
q-j-e-ŋ^{wə} *ž'ə*
 DIR-3SG.ERG-DYN-говорить теперь
 'Один из них теперь говорит мне так {...}'.
- (24) *Зымышлэрэмэ мэуцтэу ало* {...}. [Къурлан (перевод)]
zə-mə-š-e-re-me **m-ew-š't-ew** *a-ŋ^{wə}*
 REL.ERG-NEG-знать-DYN-OBL.PL DEM-ADV-PMAN-ADV 3PL.ERG-говорить
 'Так говорят те, которые не знают {...}'.
- (25) *Хъау, тэрэзэу шъузэрэубытыгъэн, мырэуцтэу шъузэрэубытын фæе!* [Бэрэздж Мэрзан. Къолэжъым ыкъо Зезэрыхъэ итаурыхъэ]
haw, terez-ew *š^{wə}-ze-re-wəbətə-k-ep,*
 нет правильный-ADV 2PL.ABS-REC.IO-RCP-хватать-PST-NEG
mə-r-e-š't-ew *š^{wə}-ze-re-wəbətə-n* *fəje*
 DEM-PRED-[ADV]-PMAN-ADV 2PL.ABS-REC.IO-RCP-хватать-MOD должный
 'Нет, вы неправильно друг за друга взялись, вот так вы должны взяться!'

Таблица 9

Морфологические модели демонстративов образа действия

	<i>мы мэ</i>	<i>мо mwe</i>	<i>джа ž'a</i>	<i>а a</i>
X-ADV-š't-	—	<i>моуцт- mwewš't-</i>	<i>джауцт- ž'awš't-</i>	<i>ауцт- awš't-</i>
X-š't-	<i>мыцт- məš't-</i>	<i>моуцт- mwewš't-</i>	<i>джауцт- ž'aš't-</i>	<i>ауцт- aš't-</i>
X-PRED-ADV-š't-	<i>мырэуцт- mərewš't-</i>	<i>морэуцт- mwerewš't-</i>	<i>джарэуцт- ž'arewš't-</i>	<i>арэуцт- arewš't-</i>
X-PRED-e-š't-	<i>мырэцт- məreš't-</i> ¹⁰	[<i>морэцт- mwereš't-</i>]	[<i>джарэцт- ž'areš't-</i>]	<i>арэцт- areš't-</i>

¹⁰ Несмотря на присутствие таких форм в корпусе (два вхождения), наши консультанты оценивают их как недопустимые.

Мы исходим из того, что в формах с *-š't* этот показатель уже функционирует как особая морфема. Поскольку его сочетаемость в данной функции ограничена местоимениями¹¹, он трактуется как показатель серии, хотя данная серия имеет большое число вариантов.

Есть соблазн считать демонстративы с *-š't* аналогами указательных местоимений с адвербиальным суффиксом, поскольку появление этого показателя для них является нормой. Однако демонстративы образа действия встречаются и в контекстах, в которых адвербиальные формы недопустимы, в частности, в качестве независимых сказуемых:

- (26) *Арэуцтэн ар! — Арэуцт!* [Элицитировано]

a-r-ew-š't-e-p a-г **a-r-ew-š't**
DEM-PRED-ADV-PMAN-NEG DEM-ABS DEM-PRED-ADV-PMAN

‘Это не так! — Это так!’

Присоединение прочей морфологии — помимо предикативного суффикса, форманта *-š't* и показателя адвербиалиса — для адвербиальных демонстративов почти не зафиксировано. Исключение составляют единичные случаи отрыва адвербиального суффикса от *-š't* экспрессивным суффиксоидом *-gʷəš'e* ‘к сожалению’ и суффиксоидом *-šəpqe* ‘действительно’:

- (27) *Ау джарэуцтэ шыпкьэу ары цыфхэр зэрэлэхэрэр.* [Вильгельм Буш. Исус — тэ тинасып (перевод)]

aw **š'a-r-ew-š'te-šəp-q-ew** a-гə cəf-xe-г
ho DEM-PRED-ADV-PMAN-действительно-ADV DEM-PRED человек-PL-ABS
ze.ge-λe-xe-ge-г
REL.FACT-умирать-PL-DYN-ABS

‘Но именно так люди и умирают’.

Морфологически демонстративы образа действия объединяются с вопросительным местоимением *səd-ew-š't* [что-ADV-PMAN-] ‘как’ и с неопределенным местоимением *zə.gʷer-e-š't* / *zə.gʷer-ew-š't* [некий-ADV-PMAN-] ‘как-то’.

10. Демонстративы тождества

Особняком стоят сочетания демонстративных корней с интенсифицирующей морфемой *dede*, которую в разных контекстах можно перевести как ‘именно’, ‘очень’ и т. д. Сочетания *dede* с указательными местоимениями, для которых в корпусе встречается как слитное, так и раздельное написание, обычно интерпретируются как ‘(э)тот же самый’, ‘именно (э)тот’ (28). В таком употреблении указательные местоимения и *dede* сочетаются и с существительными (29).

- (28) *Мырары, Бэркьукъи джадэдэр кьылуи унэм икльыгъ.* [Чэтао Инал. Шыухэр]

mə-г-a-гə, berqʷəq-jə **š'a-dede-r** q-ə-2ʷə-jə
DEM-ABS-DEM-PRED Беркук-ADD DEM-именно-ABS DIR-3SG.ERG-говорить-ADD
wəne-m jə-š'e-ə-ə
дом-OBL LOC-выходить-PST

‘Вот Беркук то же самое сказал и вышел из дома’.

- (29) *⟨...⟩ джа пчыхъэ дэдэм тыщым кьыгъэзэжыгъ.* [Адыгэ макъ]

š'a **pč'əhe-dede-m** təš'ə-m
DEM вечер-именно-OBL дом.невесты-OBL

¹¹ Морфема *š't* встречается также при выражении будущего времени, имперфекта и эпистемической модальности. В этих случаях, вероятно, она тоже восходит как *š'ə-t* ‘стоять’ (ср. [Arkadijev, Maisak 2018]), но имеет иную функцию.

q-ə-keze-ž'ə-к

DIR-3SG.ERG-поворачивать-RE-PST

‘(...) в тот же вечер домой к родителям она вернулась’.

Эти сочетания композициональны и не требуют особого описания, но в той же функции встречаются и сочетания, в которых корень указательного местоимения сопровождается морфемой *-r* (30). Мы предполагаем, что ее можно отождествить с предикативным суффиксом *-rə*, что подтверждается окказиональными примерами, в которых демонстративный корень сопровождается таким алломорфом (31).

- (30) *Гъаулэ бгъэшлэныр псынклэн, ау ардэдэм етлани ухэтми шлыкэ уегъаулэ.* [Адыгэ макъ]

ka.ʃe b-keʃe-nə-r psənç'-ep, aw a-r-dede-m
жизнь 2SG.ERG-прожить-MOD-ABS легкий-NEG но DEM-PRED-именно-OBL

jetan-jə wə-xe-t-m-jə ʃə-č'e wə-j-e-ka-ʃe
потом-ADD 2SG.ABS-LOC-стоять-COND-ADD делать-NMNR 2SG.IO-3SG.ERG-DYN-CAUS-знать

‘Прожить жизнь нелегко, но в ней же находясь, ты учишься, как себя вести’.

- (31) *Арыдэдэм цтагъэм зэрэхэтэу ибоярмэ мырэу арыуагъ...* [Мэщбэшлэ Исхъакъ (пер.). Игорь идзэ фэгъэхьыгъэ псалъ]

a-rə-dede-m ʃ'ta-ke-m ze.re-xe-t-ew
DEM-PRED-именно-OBL пугаться-PST-OBL REL.FACT-LOC-стоять-ADV

jə-bojar-me mə-r-ew a-r-jə-ʔwa-к
POSS-боярин-OBL.PL DEM-PRED-ADV 3PL.IO-DAT-3SG.ERG-говорить-PST

‘В тот же момент находясь в испуге он его боярам вот так сказал...’

И все же, даже если *-rə* здесь предикативный суффикс, эта конструкция не композициональна, так как его появление не получается описать через обычные правила его употребления. Соответственно, мы постулируем еще одну серию указательных местоимений тождества, см. таблицу 10.

Таблица 10

Демонстративные выражения тождества

	мы мэ	мо mwe	джа ž'a	а а
Простые сочетания с <i>dede</i>	мы дэдэ- мэ dede-	[мо дэдэ-] [mwe dede-]	джа дэдэ- ž'a dede-	а дэдэ- a dede-
Местоимения тождества	мыр(ы)дэдэ- мэр(э)dede-	[мор(ы)дэдэ-] [mwer(э)dede-]	джар(ы)дэдэ- ž'ar(э)dede-	ар(ы)дэдэ- ar(э)dede-

11. Демонстративы подобия

Кроме адвербиальных демонстративов для указания на качество используются сочетания указательных корней с *fede-* ‘подобный, как’ (исходно сочетание бенефактивного/направительного аппликатива с корнем *de* ‘соответствовать’) вроде *мэ fede* ‘такой’, *мэ fedew* ‘так’ (с адвербиальным показателем):

- (32) *Нартмэ арыуагъ «мы фэдэ клалэр къэсыгъ», — алуи.* [Ацумыжъ Юсыф. Нарт орзмэджь укль хасэм кызырэрэращыжыгъэр]

part-me a-r-a-ʔwa-к «мэ fe.de č'ale-r
нарт-OBL.PL 3PL.IO-DAT-3PL.ERG-говорить-PST DEM подобный парень-AVS

qe-sə-к», а-ʔ^w-jə
DIR-достигать-PST 3PL.ERG-говорить-ADD
‘Нартам сказали: «Вот такой парень прибыл»’.

- (33) *А фэдэ зылоными taloklə, — кыыcelo.* [Нэпсэу Батчэрый. Борэжъ итхъаусых]
a fe.de zə-ʔ^we-nə-m-jə
DEM подобный REL.ERG-говорить-MOD-OBL-ADD
t-a-ʔ^w-e-č’e, qə-s-j-e-ʔ^we
1PL.ABS-3PL.IO-LOC-DYN-встречать DIR-1SG.IO-3SG.ERG-DYN-говорить
‘«Мы встречаем таких, кто так скажет», — говорит мне’.

В литературе выражения вроде русск. *такой*, англ. *such, so* именуются демонстративами подобия (similarity demonstratives) [Umbach, Gust 2014]. Однако в адыгейском языке соответствующая семантика возникает благодаря слову ‘подобный’, так что, на первый взгляд, постулирование специальной серии здесь не требуется. При этом в композиционной конструкции с таким значением демонстративы выступают в качестве непрямого объекта слова *fed(e)* и должны оформляться косвенным падежом, как в (34):

- (34) *Мыщ фэдэ фестиваль Улапэ цызэхэтицэнэу бэилагъэу сыфэягъ, — ело Любэ.* [Адыгэ макъ]
mə-š’ fe.de f’est’əval’ wəlapə š’ə-ze-xe-t-š’e-n-ew
DEM-OBL подобный фестиваль Уляп LOC-REC.IO-LOC-1PL.ERG-вести-MOD-ADV
be.šə.к-ew sə-fe-ja-к, j-e-ʔ^we l’ube.
давно-ADV 1SG.ABS-BEN-хотеть-PST 3SG.ERG-DYN-говорить Люба
‘«Я давно хотела, чтобы мы организовали в Уляпе такой (букв. «подобный этому») фестиваль», — говорит Люба’.

В формах типа *a fede* падежный показатель отсутствует, поэтому они не могут интерпретироваться как словосочетания, образованные по общему правилу, и претендуют на статус отдельной серии, см. таблицу 11¹². В отсутствие выражения падежа недопустима и суффиксация демонстративных корней показателем множественного числа; ср. невозможность форм вроде **mə-xe (a-)fe.de* [DEM-PL (3PL.IO-)подобный] (ожид. ‘подобный этим’) при допустимом *mə-xe-m a-fe.de* [DEM-PL-OBL.PL 3PL.IO-подобный].

Таблица 11

Демонстративные выражения подобия

	<i>мы</i> тə	<i>мо</i> mwe	<i>джа</i> ž’a	<i>а</i> а
Простые сочетания с <i>fede</i> -	<i>мыщ/мый фэдэ-</i> məš’/məj fede-	[<i>моу/мой фэдэ-</i>] [mweš’/mwej fede-]	[<i>джау/джай фэдэ-</i>] [ž’aš’/ž’aj fede-]	<i>ауц/[ай] фэдэ-</i> aš’/[aj] fede-
Морфологизованные сочетания с <i>fede</i> -	<i>мы фэдэ-</i> mə fede-	[<i>мо фэдэ-</i>] [mwe fede-]	[<i>джа фэдэ-</i>] [ž’a fede-]	<i>а фэдэ-</i> a fede-

При установлении подобия по количеству в адыгейском языке используется выражение *fedjəz*, образованное от *fede* прибавлением числительного ‘один’ (*fe.d-jə-zə* [подобный-LNK-один]). Как и в описанной выше конструкции, в сочетании с *fedjəz* показатель

¹² Примечательно, что в кабардино-черкесском языке аналогичная конструкция, включающая демонстративы с падежным маркированием, тоже претерпела изменения: показатель падежа редуцировался и оглушился, что в других контекстах, насколько нам известно, не встречается; ср. *арх^wəde* ‘такой’ < *а-бэ х^wəде* [DEM-OBL подобный]. Таким образом, в кабардино-черкесском и конструкция с косвенным падежом развилась в особую серию.

падежа при демонстративных корнях факультативен. В примере (35) числительное ‘один’ в составе *fedjəz* выступает в кратной форме ‘один раз’:

- (35) *А фəдизэ нэтыныя?! [Пщыдатэкьо Сахьид. Нарт Шэбатныкьо ипщыналь]*

a **fe.d-jə-z-e** pe-tə-nə-j-a
DEM подобный-LNK-один-MULT LOC-стоять-MOD-EMP-Q

‘Будет ли столько стоять перед ним?!’

В данном случае постулирование особой серии зависит от того, считать ли конструкцию с *fedjəz* ‘сколько’ подтипом аналогичной конструкции с *fed-* или независимой конструкцией.

12. Локативная серия

Особая локативная серия образуется от демонстративных корней с помощью суффикса *-de* (который окказионально усекается, а также имеет алломорф *-də*; см. ниже)¹³. Такие формы описывают место и обычно употребляются в функции обстоятельства вроде русских *здесь/сюда, там/туда*. В корпусе зафиксированы формы с *-de* от всех указательных корней, кроме *ʒ'a-*¹⁴ (хотя согласно нашим консультантам, последний все же образует некоторые локативные формы; см. таблицу 12 ниже), ср.:

- (36) *Адə джыдэдэмэ тыгъаклоба. [Щэшлэ Аслъан. Чыжьэмрэ благъэмрэ]*

a-de ʒ'a-dede-m-jə tə-ka-kʷe-ba
DEM-PLOC теперь-именно-OBL-ADD 1PL.ABS-CAUS-идти-EMP

‘Поехали туда и сейчас!’

Локативные демонстративы *ade* и *məde* также развились далее в дискурсивные частицы: первая обеспечивает связь с предыдущим контекстом (схоже с русским *ну*), вторая привлекает внимание адресата (в первую очередь в сочетании с императивами, но не только, ср. русск. *ну-ка*):

- (37) *Адə мы унэм сыд ихъухъагъэр? [Хъаклэмызэ Биболэт. Анахыклэр анахь клас]*

a.de mə wəne-m səd jə-χʷə-ha-ke-r
ну DEM дом-OBL что LOC-случаться-входить-PST-ABS

‘Так что в этом доме произошло?’

- (38) *Мыдэ пчѳиф чыгъхэр ори кыыздэуплэкълх. [Гутлэ Саныет. Хьэм ыцлагъэр Куплѳ]*

mə.de pč'ejəf-čəg-xe-r we-tjə qə-z-de-wəpłekʷə-x
ну.ка берега-дерево-PL-ABS ты-ADD DIR-1SG.IO-COM-проверять-PL

‘А ну-ка и ты проверь со мной берега’.

- (39) *Мыдэ мыц зызыригъэатэкъэтицырэр... [Мэщбэшлэ Исхъакъ. Чыгу-огу зэнэсым сыда шылэр?]*

mə.de mə-š' zə-ze.r-jə-ke-ateqe-pš'ə-re-r
ну.ка DEM-OBL RFL.ABS-REL.MNR-3SG.ERG-CAUS-петух-князь-DYN-ABS

‘Ну-ка [посмотрим], как он из себя петуха делает...’

Хотя в литературе формы на *-de* описываются как производные от указательных местоимений наречия [Рогава, Керашева 1966: 332], в корпусе зафиксированы употребления местоимений этой серии с показателем косвенного падежа (40), которые, однако, не принимаются нашими консультантами. Локативные демонстративы также регулярно встречаются

¹³ Тот же суффикс обнаруживается в вопросительном местоимении *təde* ‘где’.

¹⁴ Слово *ʒ'ade*, совпадающее с ожидаемым дериватом, имеет значение ‘огромный’.

с показателем инструментального падежа, но вместо *-de* в таком случае представлена его усеченная форма (41) или алломорф *-də* (42).

(40) *Мыдэм укьеблагъэмэ цыгъу-пластэр ослон*. [Ушгый Клыщыкыу. Нарт Щэбатыныкыо иппчыналь]

mə-de-m wə-q-je-blake-me š'əkwə-ṗaste-r
DEM-PLOC-OBL 2SG.ABS-DIR-DAT-заходить-COND соль-пастэ¹⁵-ABS
w-e-s-ṗwe-n
2SG.IO-DAT-1SG.ERG-говорить-MOD

‘Если ты сюда пожадуешь, скажу тебе «соль да пастэ»’.

(41) *Жьаур адклэ макло, тыгъэр мыдклэ къэкло* {...}. [Цурымыт Музэфэр. Сабий джэ-гуклэхэр]

ž'awə-r **a-d-č'e** ma-kwe, təkə-r mə-d-č'e q-e-kwe
тень-ABS DEM-PLOC-INS ДЫН-идти солнце-ABS DEM-PLOC-INS DIR-DYN-идти

‘Тень в ту сторону идет, солнце в эту сторону идет {...}’.

(42) *Яцанэрэм верандэр мыдыклэ рехьакл, ылагъ*. [Къуекыо Налбий. Шьэоцлыкыу икьэ-бархэр. Зы нэбгырэм пае къэшлыгъохэр]

ja.š'.a.ne-re-m v'erande-r **mə-də-č'e** r-je-ha-č',
третий-ADJ-OBL веранда-ABS DEM-PLOC-INS LOC-DAT-нести-выходить
ə-ṗwa-κ
3SG.ERG-говорить-PST

‘Третий сказал: «Веранду с этой стороны сделай (букв. “в эту сторону отнеси”)»’.

Основные зафиксированные формы локативных местоимений даны в таблице 12.

Таблица 12

Пространственные демонстративы

	<i>мы</i> mə	<i>мо</i> mwe	<i>джа</i> ž'a	<i>а</i> a
Простая форма	<i>мыдэ</i> mədə	<i>модэ</i> mwede	—	<i>адэ</i> ade
OBL	<i>??мыдэм</i> <i>??madəm</i>	—	—	—
INS	<i>мыд(ы)клэ</i> məd(ə)č'e	<i>мод(ы)клэ</i> mwed(ə)č'e	[<i>джад(ы)клэ</i>] [ž'ad(ə)č'e]	<i>ад(ы)клэ</i> ad(ə)č'e

В разделах 7 и 8 показано, что как серии с локативной функцией могут быть описаны и нестандартные сочетания демонстративных корней с инструментальным и адвербиальным показателями. Кроме того, в корпусе зафиксировано слитное написание указательных местоимений с корнем *бк'э* ‘бок, сторона’, что указывает на развитие серии, основанной на этом сочетании, хотя еще не соответствует предложенным в разделе 2 критериям:

(43) *Ардэдэми мыбгъумклэ кыызэрэклуагъэм фыклэгъожьыгъ*. [Чэтао Инал. Шыухэр]

a-r-dede-m-jə **mə-bk'ə-m-č'e** qə-ze.re-kwa-κe-m
DEM-PRED-именно-OBL-ADD DEM-сторона-OBL-INS DIR-REL.ФАКТ-идти-PST-OBL
fə-č'eκwe-ž'ə-κ
DEM-сожалеть-RE-PST

‘В то же время он пожалел, что пришел с этой стороны’.

¹⁵ Национальное адыгское блюдо, аналог мамалыги.

13. Временные демонстративы

Собственно временные формы, образованные от демонстративов, — это в первую очередь сочетания соответствующих корней с $-š'əwə$ — морфологически прозрачного сегмента, включающего общелокативный преверб $š'$ и связанный корень $-wə$ ‘время’ (также используемый как показатель номинализаций со значением времени; см. [Урусов 1980: 235–236; Lander 2016: 3515–3516]). Несмотря на прозрачность структуры, формы с указательными местоимениями обычно (хотя и не всегда) пишутся слитно и, по-видимому, в значительной степени лексикализированы. Самостоятельно $š'əwə$ как слово не функционирует¹⁶, что дает основание выделить еще одну серию демонстративов, см. таблицу 13.

Таблица 13

Временные демонстративы

	мы тə	мо mwə	джа ž'a	а а
Временная серия	??? <i>мыщыгъу-</i> ??? <i>məš'əwə-</i>	—	<i>джащыгъу-</i> <i>ž'aš'əwə-</i>	<i>ащыгъу-</i> <i>aš'əwə-</i>

Из временных местоимений в корпусе встречаются почти исключительно образованные от корней $a-$ и $ž'a-$. Наиболее частотная форма таких местоимений — это косвенный падеж, регулярно используемый при описании времени (см. раздел 3 выше); ср. (44). Тем не менее в корпусе встречаются и формы абсолютива (в конструкциях псевдоклефты, в которых соответствующая абсолютная именная группа оказывается в фокусе и образует именное сказуемое, сопровождаемое связкой) (45). Форм множественного числа у временных демонстративов нет.

- (44) *Адэ, кысфэгъэгъу ащыгъум, сьд уфаеми пфэсишэн.* [Хъуажь Фахъри. Орыгъэемэ сьд ебгъэшлэны?] ‘Ну тогда прости меня, что захочешь, то я и сделаю’.

a.de qə-s-fe-ke.kwə a-š'əwə-m,
 ну DIR-1SG.IO-BEN-простить DEM-PTIME-OBL
 səd wə-fa-je-m-jə p-fe-s-še-n
 что 2SG.ABS-BEN-хотеть-COND-ADD 2SG.IO-BEN-1SG.ERG-делать-MOD

‘Ну тогда прости меня, что захочешь, то я и сделаю’.

- (45) *Ащыгъур ары нылэн зэрагъащлэрэр хъульфэгъэм кызыгурьуагъэр.* [Адыгэ макъ]

a-š'əwə-r a-rə nə.ʔ.ep ze.r-a-ka-pse-re-r
 DEM-PTIME-ABS DEM-PRED лишь REL.FACT-3PL.ERG-CAUS-ложь-DYN-ABS
 χwəlfəke-m qə-zə-gwəraʔwə-ke-r
 мужчина-OBL DIR-REL.TMP-понимать-PST-ABS

‘Именно в тот момент мужчина понял, что его обманывают’. (букв. «Это только тогда — то, когда мужчина понял, что его обманывают».)

Помимо временных форм с $-š'əwə$, встречаются и другие лексикализированные сочетания с демонстративными корнями, которые принято писать слитно: сочетания с корнями ke ‘год, летний сезон’ и $č'a$ ‘зимний сезон’ вроде *таке* ‘в этот год’, *таč'a* ‘этой зимой’. Возможно, к сочетанию с демонстративом *тə* восходит слово *пəč'er* ‘этой ночью’ (ср. *č'eš'a* ‘ночь’), но синхронно такой связи явно нет. Наконец, в адыгейском языке есть слово *ате*

¹⁶ В адыгейском языке есть омонимичное слово $š'əwə$ ‘соль’, которое мы, разумеется, не учитываем.

‘тогда’, которое можно было бы связать с демонстративом *a*, но аналогичные формы с другими указательными корнями отсутствуют.

14. Контрастивная серия

Еще одна серия указательных местоимений подразумевает обязательное существование по меньшей мере одного сходного объекта наряду с описываемым. Эта серия образуется прибавлением к демонстративному корню *-drel/-darel/-dere*, см. таблицу 14. С точки зрения морфологии *-re* отождествляется с суффиксом «относительных прилагательных» (наблюдаемым в небольшом числе образований вроде *njepere* ‘сегодняшний’ < *njepe* ‘сегодня’, а также в порядковых числительных; см. [Урусов 1980: 249]), а *-d/-dә/-de* может соотноситься с локативной серией, описанной в разделе 12. Как и в случае с локативной серией, формы, образованные от *ž’a-*, не зафиксированы в корпусе, хотя их существование допускается носителями.

Контрастивные демонстративы появляются как в приименной позиции (46), так и самостоятельно (47), и в последнем случае к ним присоединяются как падежные показатели (причем в качестве показателя косвенного падежа выступает именной *-m*, а не местоименный *-š’*), так и показатели множественного числа¹⁷.

(46) *Модрэ пчъэмкІэ сищыжъ сэ, Щылэхъан.* [КІэрэщэ Тембот. Ны Іушым ыпхъу]
mwe-d.re pče-m-č’e s-jә-š’a-ž’ se, š’әlexan
DEM-CONTR дверь-OBL-INS 1SG.ABS-LOC-вести-RE я Щыলেখан
‘Щыলেখан, выведи меня через ту дверь’.

(47) *Зыр чэтыум ышыхыгъ, адрэхэр быбыжыгыгъэх.* [ЩэшІэ Аслъан. Бзыутиф щысыгъ]
zә-г č’etәwә-m ә-šxә-к, **a-d.re-xe-r** bәbә-ž’a-кe-x
один-ABS кошка-OBL 3SG.ERG-есть-PST DEM-CONTR-PL-ABS лететь-RE-PST-PL
‘Одну съела кошка, остальные улетели’.

Таблица 14

Контрастивные демонстративы

	<i>мы</i> <i>mә</i>	<i>мо</i> <i>mwe</i>	<i>джа</i> <i>ž’a</i>	<i>а</i> <i>a</i>
Контрастивная серия	<i>мыдрэ, мыдырэ,</i> <i>???мыдэрэ</i> <i>mәdre, мәdere,</i> <i>???mәdere</i>	<i>модрэ,</i> <i>модырэ</i> <i>mwedre,</i> <i>mwedәre</i>	[<i>джадрэ</i>], [<i>джадырэ</i>] [<i>ž’adre</i>], [<i>ž’adәre</i>]	<i>адрэ, адырэ,</i> <i>???адэрэ</i> <i>adre, adәre,</i> <i>???adere</i>

По крайней мере контрастивный демонстратив, образованный от корня *a-*, регулярно переводится на русский словами *другой, остальной* [Водождоков (ред.) 1960: 199; Тхаркахо 1991: 29]. Примечательно, что формы со значением ‘другой’, образованные от указательных местоимений, обнаруживаются также в турецком и азербайджанском языках (ср. *o bir-i* [тот один-POSS] ‘другой’) и в некоторых нахско-дагестанских языках [Lander, Maisak 2021; 2022], но насколько нам известно, не распространены в других ареалах (см. [Черемисинова, Ландер 2018]).

¹⁷ В [Ярославцева 2017] для бжедугского диалекта отмечается возможность присоединения к таким формам показателя ассоциативной множественности *-tхе* (*mә-d.re-tхе-r* [DEM-CONTR-APL-AVS]), хотя и маркирующего аддитивное множественное число (см. также [Bagirokova et al. 2022: 275]).

15. Итог: серии адыгейских указательных местоимений

В соответствии с выделенными в разделе 2 принципами, в адыгейском языке выделяются по меньшей мере одиннадцать серий демонстративов:

- серия простых указательных местоимений (*mə* ‘этот’, *mə-r* [этот-ABS] и т. д.),
- предикативная серия (*mə-rə* ‘это есть’),
- серия указательных идентификаторов (*ma.rə* ‘вот’),
- инструментальная локативная серия (*mə-č'e* ‘здесь, сюда’),
- неосложненная адвербиальная серия (*m-ew* ‘так’),
- демонстративы образа действия (со множеством вариантов; *m-ew-š't-*, *mə-r-ew-š't-*, *mə-š't-*, *mə-r-e-š't-* ‘так’),
- демонстративы тождества (*mə-r-dede-* ‘этот же’),
- демонстративы подобия (*mə-fed-* ‘такой’),
- локативные демонстративы (*mə-de* ‘здесь’),
- временные демонстративы (*mə-š'əwə* ‘в это время’),
- контрастивные демонстративы (*mə-d(ə)re*, *mə-d(e)re* ‘этот другой’).

Впрочем, равноценность этих серий не очевидна. Во-первых, в то время как выделение некоторых серий основано на упомянутом в разделе 1 противопоставлении различных функций и онтологических типов (место, время, образ действия и т. д.), в иных случаях сходная семантика (например, место) обнаруживается в разных сериях. Во-вторых, некоторые из серий образуются особенными маркерами и имеют полную ожидаемую парадигму форм, другие — показателями, обычно трактуемыми как словоизменительные, но в данном случае выступающими в особой функции, третьи всего лишь имеют некоторые морфологические особенности. Наконец, не во всех сериях указательные корни выступают одинаково регулярно. В заключительном разделе мы обсудим эти проблемы с точки зрения общего подхода к выделению местоименных серий.

16. Принципы выделения местоименных серий

В чем причина неоднородности местоименных серий, выделяемых в результате применения сформулированных в разделе 2 принципов? На наш взгляд, дело в нечеткости самих принципов, повторенных в (48)–(49):

(48) Принцип исключительности:

В местоименные серии не должны включаться сочетания (пусть и частотные), если они образованы по обычным семантическим и синтаксическим или морфологическим правилам, не ассоциируемым исключительно с местоимениями.

(49) Принцип регулярности:

В каноническом случае серия должна образовываться от всех основ местоимений или релевантного подкласса местоимений.

Каноническая местоименная серия должна иметь и формальные и семантические особенности и образовываться от всех основ. Но это соблюдается не всегда: помимо канонических местоименных серий существуют кандидаты в местоименные серии, которые отклоняются от канона. Встретившись с подобным явлением, исследователь может руководствоваться дополнительными соображениями, определяющими, выберет ли он описание через понятие местоименной серии или какое-либо альтернативное описание. В связи с этим ниже мы формулируем четыре положения, связанные с уточнением наших принципов.

(50) Положение 1:

Семантические особенности являются основанием для выделения местоименной серии, только если рассматриваемые формы некомпозициональны.

При этом, если конструкция отличается только семантикой, у нас меньше оснований считать ее серией, поскольку велика вероятность того, что особенности этой конструкции могут определяться не самой конструкцией, а только местоименной семантикой, и тогда их можно описать, исходя из предположения о композициональности. Так, вопросительные местоимения часто образуют весьма специфические конструкции (например, в адыгейском языке присоединя основные аффиксы сказуемого, в отличие от прочих местоимений), но маловероятно, что их будут описывать как серии.

(51) Положение 2:

Особенности, связанные со специфическим для местоименных выражений изменением формы, скорее являются основанием для выделения серии, если они касаются синтаксических конструкций, чем если они касаются морфологических конструкций.

Если у конструкции наблюдаются только формальные особенности, а семантически она аналогична неместоименной, выделение серии не очевидно. Такая конструкция часто может быть описана через супплетивизм (часто наблюдаемый в местоименных системах) или как реализация особенной местоименной словоизменительной модели. Синтаксические конструкции обычно более последовательны в отношении поведения их частей, чем слова, так что описание через механизмы вроде супплетивизма для специфических синтаксических конструкций более спорно, а постулирование особой серии более естественно. Поэтому, например, для адыгейского языка мы допускаем выделение демонстративов тождества и демонстративов качества, которые отличаются от аналогичных синтаксических конструкций нестандартным оформлением.

В то же время сложные местоименные выражения нередко претерпевают морфологизацию и находятся на границе синтаксиса и морфологии (см., например, [Аркадьев 2016] о русских сложных неопределенных местоимениях, которые, несмотря на возможность разрыва предлогами, предлагается считать словами). Не исключение здесь и упомянутые адыгейские серии: демонстративы тождества обычно образуют единый именной комплекс, а несамостоятельные указательные местоимения в демонстративах подобия, возможно, должны анализироваться как связанные морфологические элементы.

Формальный признак, который отличает местоименную форму, претендующую на то, чтобы представлять местоименную серию, от неместоименных форм, может быть связан не только с присутствием какого-то особого элемента, но и с отсутствием какого-то элемента. Адыгейская серия местоимений подобия выделяется на основании отсутствия косвенного падежного показателя на указательном корне; на том же основании мы выделяем самостоятельную инструментальную локативную серию, которая, впрочем, имеет и семантические особенности.

(52) Положение 3:

Факультативность специфической формальной местоименной черты свидетельствует о начале формирования серии, но допускает и альтернативное описание без постулирования местоименной серии.

Адыгейские серии демонстративов тождества и подобия поднимают и вопрос о трактовке ситуации, когда релевантная специфическая для конструкции черта оказывается факультативной. Должны ли мы, скажем, считать композициональное сочетание *mə-dede* 'этот самый' и морфологически неожиданное сочетание *mə-r-dede* (см. раздел 10) вариантами? Как кажется, факультативность специфической черты свидетельствует о начале формирования особой серии, но на этом этапе возможен и альтернативный подход, в рамках

которого в качестве основной при описании выбирается более регулярная конструкция, а менее регулярная представляется как ее вариант.

(53) Положение 4:

Местоименная серия не должна охватывать только одно местоимение из рассматриваемого класса, но охват лишь части местоимений местоименной серией допустим, особенно в тех случаях, когда он допускает ту или иную мотивацию.

Само понятие «серия» подразумевает некоторое множество и регулярность образования. Поскольку регулярность градуируема, принятие решения о ее наличии в условиях, когда какая-то конструкция охватывает не все местоимения рассматриваемого класса, не однозначно. Решение в пользу постулирования серии, которая не включает часть местоименных основ, может быть принято, если у такого невключения есть ясная мотивация. Например, поскольку исходное значение указательных местоимений — пространственный дейксис, метафорический перенос некоторых пространственных противопоставлений в другие семантические поля иногда затруднен. В адыгейском языке это объясняет более ограниченную сочетаемость указательного корня «среднего дейксиса» *mwe* (для которого в корпусе не встречаются производные выражения тождества, подобия и временные демонстративы) и, вероятно, не должно само по себе считаться аргументом против выделения серий, его не включающих.

Мы не предлагаем универсального рецепта выделения местоименных серий. Более того, за рамками нашего обсуждения остался сложный вопрос о противопоставлении и потенциальном объединении разных серий. Важно, однако, что явления, претендующие на описание в качестве местоименных серий, нередко допускают и альтернативные трактовки. Возможно, причина этого лежит в определенной диахронической нестабильности местоименных серий: в продолжающейся их морфологизации, с одной стороны, и конкуренции между разными сериями, с другой. Впрочем, подтверждение этого тезиса требует исследования истории местоименных серий, не доступной для многих языков. С дескриптивной точки зрения, однако, исследователю важно сознавать степень близости описываемого явления к каноническому образу местоименной серии и по возможности ее эксплицировать.

СПИСОК СОКРАЩЕНИЙ / ABBREVIATIONS

ABS — абсолютив
ADD — аддитив
ADJ — суффикс «относительных прилагательных»
ADV — адвербиальный показатель
BEN — бенефактив
CAR — каритив
CAUS — каузатив
COM — комитатив
COMP — компаратив
COND — кондиционал
CONTR — контрастивное местоимение
CS — консекутив
DAT — дативный аппликатив
DEM — демонстратив
DIR — директивный преверб
DYN — динамичность
EMP — эмфаза
ERG — эргативный участник
FACT — фактивная релятивизация ('то, что')

FUT — будущее время
INS — инструменталис
INT — интенсив
IO — не прямой объект
LNK — линкер
LOC — локативный преверб
MAL — малефактив
MNR — релятивизация образа действия ('то, как')
MOD — модальный суффикс
MULT — кратное числительное
NEG — отрицание
NMNR — номинализация образа действия
NTIME — номинализация времени
OBL — косвенный падеж
PL — множественное число
PLOC — локативное местоимение
PMAN — местоимение образа действия
PO — объект преверба
POSS — посессив

PR — неотторжимый посессор
 PRED — предикативный суффикс
 PST — прошедшее время
 PTIME — местоимение времени
 Q — вопрос
 RE — рефактив/реверсив

RCP — реципрокальный аппликатив
 REC — реципрок
 REL — релятивизация
 SG — единственное число
 TMP — темпоральная релятивизация ('когда')

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Аркадьев 2016 — Аркадьев П. М. К вопросу об эндоклитиках в русском языке. *Архитектура клаузы в параметрических моделях: Синтаксис, информационная структура, порядок слов*. Циммерлинг А. В., Лютикова Е. А. (ред.). М.: Языки славянских культур, 2016, 325–331. [Arkadiev P. M. On endoclitics in the Russian language. *Arkhtektura klauzy v parametricheskikh modelyakh: Sintaksis, informatsionnaya struktura, poryadok slov*. Tsimmerling A. V., Lyutikova E. A. (eds.). Moscow: Yazyki slavyanskikh kul'tur, 2016, 325–331.]
- Аркадьев и др. 2009 — Аркадьев П. М., Ландер Ю. А., Летучий А. Б., Сумбатова Н. Р., Тестелец Я. Г. Введение. Основные сведения об адыгейском языке. *Аспекты полисинтетизма: Очерки по грамматике адыгейского языка*. Тестелец Я. Г. (отв. ред.). М.: РГГУ, 2009, 17–120. [Arkadiev P. M., Lander Yu. A., Letuchiy A. B., Sumbatova N. R., Testeleets Ya. G. An introduction. Basic features of West Circassian. *Aspekty polisintetizma: Ocherki po grammatike adygeiskogo yazyka*. Testeleets Ya. G. (ed.). Moscow: Russian State Univ. for the Humanities, 2009, 17–120.]
- Архангельский и др. 2018–2023 — Архангельский Т. А., Багирова И. Г., Ландер А. Д., Ландер Ю. А. *Адыгейский корпус*. 2018–2023. <http://adyghe.web-corpora.net/>. [Arkhangelskiy T. A., Bagirokova I. G., Lander A. D., Lander Yu. A. *West Circassian Corpus*. 2018–2023.]
- Барыльникова 2013 — Барыльникова Д. В. *Пространственный дейксис в темиргоевском диалекте адыгейского языка*. Рук., 2013. [Baryl'nikova D. V. *Prostranstvennyi deiksiss v temirgoevskom dialekte adygeiskogo yazyka* [Spatial deixis in Temirgoy West Circassian]. Ms., 2013.]
- Беликов 1972 — Беликов В. И. К употреблению указательных местоимений в шугнанском языке. *Забонҳои помирӣ ва фольклор. Памятные языки и фольклор*. Т. I. Душанбе: Дониш, 1972, 74–77. [Belikov V. I. On the use of demonstrative pronouns in Shughni. *Zabonhoi pomirui va fol'klor. Pamirskie yazyki i fol'klor*. Vol. I. Dushanbe: Donish, 1972, 74–77.]
- Водождоков (ред.) 1960 — *Русско-адыгейский словарь*. Водождоков Х. Д. (ред.). М.: ГИИНС, 1960. [*Russko-adygeiskii slovar'* [Russian-West Circassian dictionary]. Vodozhdokov Kh. D. (ed.). Moscow: State Press of Foreign and National Dictionaries, 1960.]
- Герасимов 2010 — Герасимов Д. В. Предикативные формы личных местоимений в адыгейском языке. *Типологически редкие и уникальные явления на языковой карте России. Тезисы докладов международной научной конференции. Институт лингвистических исследований РАН*. СПб.: Нестор-История, 2010, 18–20. [Gerasimov D. V. Predicative forms of personal pronouns in West Circassian. *Tipologicheski redkie i unikal'nye yavleniya na yazykovoi karte Rossii. Conf. proc. Institute for Linguistic Studies (RAS)*. St. Petersburg: Nestor-Istoriya, 2010, 18–20.]
- Гишев 1985 — Гишев Н. Т. *Вопросы эргативного строя адыгских языков*. Майкоп: Адыгейское отд. Краснодарского книжного издательства, 1985. [Gishev N. T. *Voprosy ergativnogo stroya adygskikh yazykov* [Topics in ergativity of the Circassian languages]. Maykop: Adyghea Department of Krasnodar Publ. House, 1985.]
- Зекох 1969 — Зекох У. С. *Система склонения в адыгейском языке*. Майкоп: Адыгейское отд. Краснодарского книжного изд-ва, 1969. [Zekokh U. S. *Sistema skloneniya v adygeiskom yazyke* [Declension system in West Circassian]. Maykop: Adyghea Department of Krasnodar Publ., 1969.]
- Зекох 2002 — Зекох У. С. *Адыгейская грамматика*. Майкоп: Качество, 2002. [Zekokh U. S. *Adygeiskaya grammatika* [West Circassian grammar]. Maykop: Kachestvo, 2002.]
- Зекох и др. 2003 — Зекох У. С., Меретуков К. Х., Тхаркахо Ю. А. *Структурные особенности адыгейского языка*. Майкоп: Качество, 2003. [Zekokh U. S., Meretukov K. Kh., Tkharakho Yu. A. *Strukturnye osobennosti adygeiskogo yazyka* [Structural features of West Circassian]. Maykop: Kachestvo, 2003.]
- Коряков 2006 — Коряков Ю. Б. *Атлас кавказских языков*. М.: Пилигрим, 2006. [Koryakov Yu. B. *Atlas kavkazskikh yazykov* [An atlas of Caucasian languages]. Moscow: Piligrim, 2006.]
- Кумахов 1964 — Кумахов М. А. *Морфология адыгских языков. Синхронно-диахронная характеристика*. Т. I. Нальчик: Кабардино-Балкарское книжное изд-во, 1964. [Kumakhov M. A. *Morfologiya*

- adygskikh yazykov. Sinkhronno-diakhronnaya kharakteristika* [Morphology of the Circassian languages. A synchronic and diachronic description]. Vol. I. Nalchik: Kadardino-Balkarian Publ., 1964.]
- Ландер 2012 — Ландер Ю. А. *Релятивизация в полисинтетическом языке: адыгейские относительные конструкции в типологической перспективе*. Дисс. ... канд. филол. наук. М.: Институт востоковедения РАН, 2012. [Lander Yu. A. *Relyativizaciya v polisinteticheskom yazyke: adygeyskie otnositel'nye konstrukcii v tipologicheskoi perspektive* [Relativization in a polysynthetic language: West Circassian relative constructions in a typological perspective]. Cand. diss. Moscow: Institute of Oriental Studies, 2012.]
- Перцов 2001 — Перцов Н. В. *Инварианты в русском словоизменении*. М: Языки русской культуры, 2001. [Pertsov N. V. *Invarianty v russkom slovoizmenenii* [Invariants of Russian nominal declension]. Moscow: Yazyki russkoi kul'tury, 2001.]
- Рогав, Керашева 1966 — Рогав Г. В., Керашева З. И. *Грамматика адыгейского языка*. Краснодар; Майкоп: Краснодарское книжное изд-во, 1966. [Rogava G. V., Kerasheva Z. I. *Grammatika adygeiskogo yazyka* [Grammar of West Circassian]. Krasnodar; Maykop: Krasnodar Publ., 1966.]
- Сердобольская, Кузнецова 2009 — Сердобольская Н. В., Кузнецова Ю. Л. Двойное падежное маркирование: уникальный случай адыгейского языка. *Аспекты полисинтетизма: Очерки по грамматике адыгейского языка*. Тестелец Я. Г. (отв. ред.). М.: РГГУ, 2009, 166–200. [Serdobol'skaya N. V., Kuznetsova Yu. L. Double case marking: A unique case of West Circassian. *Aspekty polisintetizma: Ocherki po grammatike adygeiskogo yazyka*. Testeleets Ya. G. (ed.). Moscow: Russian State Univ. for the Humanities, 2009, 166–200.]
- Сердобольская, Мотлохов 2009 — Сердобольская Н. В., Мотлохов А. В. Семантика конструкций с сентенциальными актантами в адыгейском языке. *Аспекты полисинтетизма: Очерки по грамматике адыгейского языка*. Тестелец Я. Г. (отв. ред.). М.: РГГУ, 2009, 498–558. [Serdobol'skaya N. V., Motlokhov A. V. Semantics of constructions with sentential arguments in West Circassian. *Aspekty polisintetizma: Ocherki po grammatike adygeiskogo yazyka*. Testeleets Ya. G. (ed.). Moscow: Russian State Univ. for the Humanities, 2009, 498–558.]
- Ситимова 2004 — Ситимова С. С. *Бжедугский диалект в адыгейском языковом мире*. Майкоп: Качество, 2004. [Sitimova S. S. *Bzhehudskii dialekt v adygeiskom yazykovom mire* [The Bzhehud dialect in the Circassian linguistic world]. Maykop: Kachestvo, 2004.]
- Степанянц 2023 — Степанянц М. Г. К типологии диахронических источников адверсативных показателей. *Известия Российской академии наук. Сер. литературы и языка*, 2023, 82(1): 94–101. [Stepanyants M. G. Towards a typology of diachronic sources of adversative markers. *Izvestiya Rossiyskoi akademii nauk. Ser. literatury i yazyka*, 2023, 82(1): 94–101.]
- Сумбатова 2009 — Сумбатова Н. Р. Коммуникативная структура адыгейского предложения: перспектива и фокус. *Аспекты полисинтетизма: Очерки по грамматике адыгейского языка*. Тестелец Я. Г. (отв. ред.). М.: РГГУ, 2009, 559–611. [Sumbatova N. R. Information structure of West Circassian sentences: Perspective and focus. *Aspekty polisintetizma: Ocherki po grammatike adygeiskogo yazyka*. Testeleets Ya. G. (ed.). Moscow: Russian State Univ. for the Humanities, 2009, 559–611.]
- Тхаркахо 1991 — Тхаркахо Ю. А. *Адыгейско-русский словарь*. Майкоп: Адыгейское книжное изд-во, 1991. [Tkhar'kakh'o Yu. A. *Adygeisko-russkii slovar'* [West Circassian-Russian dictionary]. Maykop: Adygea Publ., 1991.]
- Урусов 1980 — Урусов Х. Ш. *Морфемика адыгских языков*. Нальчик: Эльбрус, 1980. [Urusov Kh. Sh. *Morfemika adygskikh yazykov* [Morphemics of the Circassian languages]. Nalchik: El'brus, 1980.]
- Черемисинова, Ландер 2018 — Черемисинова М. О., Ландер Ю. А. ДРУГОЙ: типологические наблюдения. *Рема*, 2018, 4: 109–124. [Cheremisinova M. O., Lander Yu. A. 'Other': typological observations. *Rhema*, 2018, 4: 109–124.]
- Яковлев, Ашхамаф 1941 — Яковлев Н. Ф., Ашхамаф Д. А. *Грамматика адыгейского литературного языка*. М.; Л.: Изд-во АН СССР, 1941. [Yakovlev N. F., Ashkhamaf D. A. *Grammatika adygeiskogo literaturnogo yazyka* [Grammar of West Circassian standard language]. Moscow; Leningrad: Academy of Sciences of the USSR Press, 1941.]
- Ярославцева 2017 — Ярославцева М. В. *Вопросы выражения множественного числа в адыгейском языке*. Курсовая работа. М.: НИУ ВШЭ, 2017. [Yaroslavtseva M. V. *Voprosy vyrazheniya mnozhestvennogo chisla v adygeyskom yazyke* [Aspects of the expression of plural in West Circassian]. Term paper. Moscow: HSE Univ., 2017.]
- Arkadiev, Maisak 2018 — Arkadiev P., Maisak T. Grammaticalization in the North Caucasian languages. *Grammaticalization from a typological perspective*. Narrog H., Heine B. (eds.). Oxford: Oxford Univ. Press, 2018, 116–145.

- Arkadiyev, Testeleys 2019 — Arkadiyev P., Testeleys Y. Differential nominal marking in Circassian. *Studies in Language*, 2019, 43(3): 715–751.
- Bagirokova, Lander 2019 — Bagirokova I., Lander Yu. *Asymmetric word class systems and noun primacy: West Circassian and beyond*. Ms., 2019.
- Bagirokova et al. 2022 — Bagirokova I., Lander Yu., Phelan P. Number in West Circassian. *Number in the world's languages*. Acquaviva P., Daniel M. (eds.). Berlin: De Gruyter Mouton, 2022, 261–304.
- Bhat 2004 — Bhat D. N. S. *Pronouns*. Oxford: Oxford Univ. Press, 2004.
- Coniglio et al. (eds.) 2018 — Coniglio M., Murphy A., Schlachter E., Veenstra T. (eds.). *Atypical demonstratives: Syntax, semantics and pragmatics*. Berlin: De Gruyter, 2018.
- de Reuse 2009 — de Reuse W. J. Polysynthesis as a typological feature: An attempt at a characterization from Eskimo and Athabaskan perspectives. *Variations on polysynthesis: The Eskaleut languages*. Mahieu M.-A., Tersi N. (eds.). Amsterdam: John Benjamins, 2009, 19–34.
- Diessel 1999a — Diessel H. *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins, 1999.
- Diessel 1999b — Diessel H. The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology*, 1999, 3(1): 1–50.
- Diessel 2006 — Diessel H. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics*, 2006, 17(4): 463–489.
- Dixon 2003 — Dixon R. M. W. Demonstratives: A cross-linguistic typology. *Studies in Language*, 2003, 27(1): 61–112.
- Dumézil 1932 — Dumézil G. *Études comparatives sur les langues caucasiennes du nord-ouest*. Paris: Adrien-Maisonneuve, 1932.
- Ershova 2019 — Ershova K. A. *Syntactic ergativity in West Circassian*. Ph.D. diss. Chicago: Univ. of Chicago, 2019.
- Fillmore 1982 — Fillmore C. J. Towards a descriptive framework for spatial deixis. *Speech, place, and action*. Jarvella R. J., Klein W. (eds.). Chichester: John Wiley, 1982, 31–59.
- Haspelmath 1997 — Haspelmath M. *Indefinite pronouns*. Oxford: Oxford Univ. Press, 1997.
- Heine, Song 2011 — Heine B., Song K. A. On the grammaticalization of personal pronouns. *Journal of Linguistics*, 2011, 47(3): 587–630.
- Himmelmann 1996 — Himmelmann N. P. Demonstratives in narrative discourse: A taxonomy of universal uses. *Studies in anaphora*. Fox B. A. (ed.). Amsterdam: John Benjamins, 1996, 205–254.
- Himmelmann 1997 — Himmelmann N. P. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Berlin: Max Niemeyer, 1997.
- Killian 2022 — Killian D. Towards a typology of predicative demonstratives. *Linguistic Typology*, 2022, 26(1): 1–41.
- Konuk 2022 — Konuk N. M. *Description grammaticale de l'abzakh — tcherkesse occidental*. Ph.D. diss. Paris: Université de la Sorbonne Nouvelle — Paris III, 2022.
- König, Umbach 2018 — König E., Umbach C. Demonstratives of manner, of quality and of degree. *Atypical demonstratives: Syntax, semantics and pragmatics*. Coniglio M., Murphy A., Schlachter E., Veenstra T. (eds.). Berlin: De Gruyter, 2018, 285–328.
- Lander 2009 — Lander Yu. *Subject properties of the Adyghe absolutive: evidence from relatives*. Ms., 2009.
- Lander 2016 — Lander Yu. Adyghe. *Word-Formation. An international handbook of the languages of Europe*. Vol. 5. Müller P. O., Ohnheiser I., Olsen S., Rainer F. (eds.). Berlin: De Gruyter Mouton, 2016, 3508–3527.
- Lander 2017 — Lander Yu. Nominal complex in West Circassian: between morphology and syntax. *Studies in Language*, 2017, 41(1): 76–98.
- Lander, Maisak 2021 — Lander Yu., Maisak T. “Other” strategies in the Eastern Caucasus (Part I): Data from Udi. *Iran and the Caucasus*, 2021, 25(3): 272–283.
- Lander, Maisak 2022 — Lander Yu., Maisak T. “Other” strategies in the Eastern Caucasus (Part II): Typology. *Iran and the Caucasus*, 2022, 26(3): 272–288.
- Lander, Testeleys 2017 — Lander Yu., Testeleys Y. Adyghe. *The Oxford handbook of polysynthesis*. Fortescue M., Mithun M., Evans N. (eds.). Oxford: Oxford Univ. Press, 2017, 948–970.
- Lander et al. 2021 — Lander Yu., Belyaev O., Bagirokova I. (Almost) everything is oblique in West Circassian. *Proceedings of the LFG'21 Conference*. Batt M., Findlay J. Y., Toivonen I. (eds.). Stanford: CSLI Publications, 2021, 223–242.
- Letuchiy 2012 — Letuchiy A. Ergativity in the Adyghe system of valency-changing derivations. *Ergativity, valency and voice*. Authier G., Haude K. (eds.). Berlin: De Gruyter Mouton, 2012, 323–354.

- Levinson 2004 — Levinson S. C. Deixis. *The handbook of pragmatics*. Horn L., Ward G. (eds.). Oxford: Blackwell, 2004, 97–121.
- Levinson 2018 — Levinson S. C. Introduction: Demonstratives: Patterns in diversity. *Demonstratives in cross-linguistic perspective*. Levinson S., Cutfield S., Dunn M. J., Enfield N. J., Meira S. (eds.). Cambridge: Cambridge Univ. Press, 2018, 1–42.
- Lyons 1977 — Lyons J. *Semantics*. Vol. 2. Cambridge: Cambridge Univ. Press, 1977.
- Næss et al. (eds.) 2020 — Næss Å., Margetts A., Treis Y. (eds.). *Demonstratives in discourse*. Berlin: Language Science Press, 2020.
- Nida, Moisés 1950 — Nida E. A., Moisés R. C. The pronominal series in Maya (Yucatec). *International Journal of American Linguistics*, 1950, 16(4): 193–197.
- Paris 1989 — Paris C. West Circassian (Adyghe: Abzakh dialect). *The indigenous languages of the Caucasus*. Vol. 2: *North West Caucasus*. Hewitt B. G. (ed.). Delmar: Caravan Books, 1989, 157–260.
- Serdobolskaya 2011 — Serdobolskaya N. Grammaticalization patterns of the Adyghe instrumental case. *Languages and cultures in the Caucasus. Papers from the International conference “Current advances in Caucasian studies”*, Macerata, Jan. 21–23, 2010. Tomelleri V. S. et al. (eds.). München: Otto Sagner, 2011, 515–539.
- Smeets 1984 — Smeets R. *Studies in West Circassian phonology and morphology*. Leiden: Hakuchi Press, 1984.
- Smeets 1992 — Smeets R. Towards an analysis of pronominal *rə* in Circassian. *Caucasologie et mythologie comparée. Actes du colloque international du CNRS, IVe colloque de caucasologie (Sèvres, 27–29 juin 1988)*. Paris C. (ed.). Paris: Peeters, 1992, 301–320.
- Testeleets, Bylinina 2005 — Testeleets Ya. G., Bylinina E. Sluicing-based indefinites in Russian. *Formal approaches to Slavic linguistics* # 13. Franks S., Gladney F. Y., Tasseva-Kurkchieva M. (eds.). Ann Arbor: Michigan Slavic Publications, 2005, 355–364.
- Umbach, Gust 2014 — Umbach C., Gust H. Similarity demonstratives. *Lingua*, 2014, 149: 74–93.
- Vydrin 2008 — Vydrin A. Are there depictives in Adyghe? *Secondary predicates in Eastern European languages and beyond*. Schroeder C., Hentschel G., Boeder W. (eds.). Oldenburg: BIS, 2008, 424–445.
- Vydrine 2006 — Vydrine V. Emergence of morphological cases in South Mande. *Case, valency, and transitivity*. Kulikov L., Malchukov A. L., de Swart P. (eds.). Amsterdam: John Benjamins, 2006, 49–64.

Получено / received 11.05.2023

Принято / accepted 26.09.2023

Чукотские сонорные согласные в типологической перспективе

© 2024

Инна Арнольдовна Зибер

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; innasieber@gmail.com

Аннотация: В статье описываются сонорные согласные чукотского языка. Особое внимание уделено согласным, демонстрирующим значительную фонетическую вариативность: аппроксимантам (*w, j* и *ɥ*) и *r*. Эта вариативность осмысливается с точки зрения акустики и артикуляции, типологии, истории и диалектологии чукотско-корякских идиомов, социолингвистики и влияния русского языка. Привлечение нескольких подходов позволяет связать разрозненные факты о чукотском языке и объяснить некоторые реализации согласных фонем из класса сонорных, а также привлечь внимание к межъязыковым сравнениям сонорных. В исследовании используются полевые записи на амгуэмском говоре чукотского языка.

Ключевые слова: чукотский язык, чукотско-камчатские языки, консонантизм, сонорные согласные, фонетика, фонетическая типология

Благодарности: Автор от всего сердца благодарит жительниц и жителей села Амгуэма за многолетнюю самоотверженную лингвистическую работу, за внимание, понимание и помощь во всем. Автор также низко кланяется редакции журнала и уважаемым анонимным рецензентам за комментарии и советы, благодаря которым типологическая перспектива чукотского материала стала полнее и богаче, а сама автор обрела уверенность смелее указывать на те параллели, которые кажутся ей интересными. Исследование выполнено в рамках Программы фундаментальных исследований НИУ ВШЭ в 2023 г.

Для цитирования: Зибер И. А. Чукотские сонорные согласные в типологической перспективе. *Вопросы языкознания*, 2024, 2: 122–142.

DOI: 10.31857/0373-658X.2024.2.122-142

Chukchi sonorant consonants in a typological perspective

Inna A. Sieber

HSE University, Moscow, Russia; innasieber@gmail.com

Abstract: The article focuses on the sonorant consonants of the Chukchi language. Particular attention is paid to the consonants that show considerable phonetic variation: approximants (*w, j* and *ɥ*) and *r*. This variation is described in terms of acoustics and articulation, typology, history and dialectology of the Chukchi-Koryak languages, sociolinguistics and the influence of Russian. We use several approaches to link facts about the Chukchi language and explain some realisations of sonorants, as well as draw attention to cross-linguistic comparisons of approximants and rhotics. The study uses field recordings in the Amguema dialect of Chukchi.

Keywords: Chukchi, Chukotko-Kamchatkan, consonants, phonetic typology, phonetics, sonorants

Acknowledgements: This study was carried out in 2023 in the framework of the Basic Research Program at the HSE University.

For citation: Sieber I. A. Chukchi sonorant consonants in a typological perspective. *Voprosy Jazykoznanija*, 2024, 2: 122–142.

DOI: 10.31857/0373-658X.2024.2.122-142

Введение

В отличие от грамматики чукотского языка, чукотская фонетика нечасто становится предметом обсуждения в сравнительных и типологических работах. Это можно объяснить тем, что система фонем чукотского языка невелика по объему и просто устроена. Однако за фонологической простотой при более внимательном рассмотрении кроются фонетическое богатство и типологические редкости. В этой статье отдельные аспекты артикуляционного разнообразия чукотских сонорных согласных встраиваются в контекст артикуляционной базы языка в целом и связываются с рядом фонетических явлений в других подсистемах. Это становится возможным благодаря привлечению данных акустики, типологии, истории и диалектологии чукотско-корякских идиомов, социолингвистики и влияния русского языка.

В исследовании используются полевые записи на амгуэмском говоре чукотского языка, сделанные летом 2018 и 2021 гг. в национальном селе Амгуэма Иультинского района Чукотского автономного округа. Село Амгуэма расположено в каменистой тундре в глубине Чукотского полуострова, на 91-м километре Иультинского тракта, за полярным кругом. В поселке живут около 500 человек, многие жители занимаются или занимались в молодости оленеводством.

Статья состоит из трех разделов. В разделе 1 приводятся общие сведения о чукотском языке и его фонетической системе, представленные в существующей литературе по теме, и с точки зрения типологии оцениваются объем и состав чукотского консонантного инвентаря. В разделе 2 рассматриваются реализации *r* в контексте артикуляционной базы говора, в разделе 3 — способ образования губных и велярных неносовых сонантов и отношения между ними.

1. Обзор чукотской фонетической системы

Чукотский язык относится к чукотско-корякским языкам и обычно включается в чукотско-камчатскую семью вместе с ительменским языком (ительменскими языками) [Fortescue 2005]. Носители чукотского языка живут на Чукотке, а также на западе Республики Саха. По данным переписи 2010 г., чукчей на территории России немногим менее 16 тысяч, а чукотским владеет 5095 человек.

Наиболее полных описаний языка существует три: сопоставительная грамматика чукотского, корякского и ительменского языков [Bogoras 1922], академическая двухтомная грамматика [Скорик 1961; 1977], а также англоязычная диссертация [Dunn 1999]. По фонетике, фонологии и морфонологии есть специальные работы, многие исследования посвящены сопоставлению чукотско-корякских и, шире, чукотско-камчатских языков, в том числе [Муравьева 1979; Krause 1980; Kenstowicz 1986; Асиновский 2003].

Диалектные различия в чукотском языке неоднократно описывались как незначительные [Bogoras 1922; Володин, Скорик 1997; Dunn 1999]. В [Богораз 1934] отмечается, что сколько-нибудь значимые различия наблюдаются между говорами приморских (оседлых) и оленных (кочевых) чукчей, однако в более поздней работе [Богораз 1937] приводится разделение на западную группу, включающую говоры большинства оленных чукчей, и восточную группу, в которую входят говоры приморских чукчей, а также оленных чукчей, живущих вблизи приморских поселков. На основе восточного (уэленского) говора, описанного в [Скорик 1961], была сформирована литературная норма; тем не менее, в [Dunn 1999: 23] утверждается, что описанный П. Я. Скориком идиом не соотносится с данными живой речи чукчей. Попытка уточнить диалектное членение чукотского языка была предпринята М. Ю. Пупыниной. На основании полевых материалов и данных

об ареалах чукотских кочевий были выделены некоторые «этнические и, возможно, лингвистические общности»: анадырская, хатырско-ваежская, чаунская, илирнейская, нижне-колымская и, возможно, омонская [Пупынина 2013: 258]. На основе лингвистических данных исследовательница выделяет имеющие существенные отличия от других хатырский (возможно, хатыро-ваежский) и колымский говоры [Там же: 251], таким образом отделяя от восточных говоров южные и западные, но их фонетическим особенностям уделяет немного внимания.

М. Данн создал грамматику на основе говора, носители которого живут к югу от Анадыря (“Telqer variety”) и относятся к хатырской/хатыро-ваежской общности по классификации М. Ю. Пупыниной. М. Данн неоднократно отмечает отличия южного «телькепского» говора, который он описывает, от восточного (уэленского) чукотского, на который опирается грамматика П. Я. Скорика. В настоящей работе привлекается материал говора села Амгуэма, которое находится в 300 км к северо-востоку от Анадыря (в 620 км от Хатырки в ту же сторону) и, несмотря на принадлежность к восточной диалектной зоне, более чем в 400 км к западу от Уэлена по прямой, а с учетом сложной транспортной доступности на Чукотке практически это расстояние намного больше. Комплексного описания амгуэмского говора пока не существует, сейчас ему посвящены отдельные работы участников экспедиций из НИУ ВШЭ.

Гендерные диалекты — различия между мужской и женской речью — упоминаются во всех основных описаниях чукотского языка и существуют на лексическом и фонетическом уровнях. На уровне фонетики они создаются:

- 1) качеством единственного сибиланта («мужской» чаще описывается как шипящий, а «женский» — как свистящий);
- 2) «чередованием» /r/~/ts/, см. ниже;
- 3) возможно, выпадением согласных в интервокальном положении в женском гендерном диалекте (описано только в [Bogoras 1922: 665] и другими исследованиями не подтверждалось [Dunn 1999: 26–31]).

«Чередование» согласных можно проиллюстрировать следующими примерами из [Dunn 1999: 27], см. таблицу 1.

Таблица 1

Некоторые слова чукотского языка в произношении мужчин и женщин
(примеры из южного говора, грамматика [Dunn 1999])

	Есть чередование		Нет чередования	
Перевод	‘комар’	‘морж’	‘чайник’	‘олень’
Женский чукотский	mtsen	tsətsətsə	tsajkok	qoŋaŋə
Мужской чукотский	mŋen	rərkə	sajkok	qoŋaŋə

В работах М. Данна, в том числе в грамматике [Dunn 1999] и статье [Dunn 2014], показано, что на синхронном уровне указанные фонетические соотношения непредсказуемы и обусловлены тем, что в мужском и женском чукотском дали разные рефлексy прачукотско-корякские фонемы *s, *r и *ð. Фонема *ð перешла в /r/ в мужском чукотском и в /ts/ в женском, слившись в них с рефlekсами *r и *s соответственно.

Несмотря на то, что в основных описаниях чукотского языка постулируется существование гендерных диалектов, это разделение, по всей видимости, сохранилось не на всей территории его распространения. Так, в амгуэмской тундре, по нашим данным, и женщины, и мужчины, говорящие на чукотском языке¹, придерживаются сейчас почти исключительно

¹ Имеются в виду носители чукотского старше 40 лет; носителей младшего возраста, насколько нам известно, в Амгуэме и близлежащей тундре нет.

мужского произношения, хотя еще их мамы и бабушки по-женски «цекали». Это можно связать с тем, что литературная норма, престиж которой в советское время поднимался через радиовещание и школьное образование, была сформирована на основе мужского диалекта. Кроме того, свидетельства наших консультанток позволяют предположить, что и в некоторых прибрежных поселках к северу от Амгуэмы женское произношение в поколениях их мам и бабушек распространено не было.

Чукотский вокализм включает шесть гласных: /i e a o u ə/. Шва ə во многих случаях используется для приведения слога к разрешенной структуре и потому имеет спорный фонемный статус [Володин, Скорик 1997; Dunn 1999]. В чукотском языке, в отличие от других чукотско-камчатских идиомов, сохранилась регулярная гармония гласных по подъему. Чукотский относится к квантовым языкам по классификации С. В. Кодзасова: его характеризует «довольно четкая» ритмико-слоговая организация [Муравьева 1979: 62]; слог внутри словоформы тяготеет к структуре (C)V(C). (Морфо)фонологический статус ударения спорен. В одних работах утверждается, что фонологического ударения в чукотском нет [Dunn 1999], в других описываются сложные правила его постановки, связанные с сегментным и морфологическим составом слова [Володин, Скорик 1997]. Фонетическая реализация ударения не изучена; в [Муравьева 1979] утверждается, что ударные гласные выделяются с помощью длительности.

Консонантная система чукотского языка включает 14 единиц шести мест образования. Сегментный статус гортанной смычки оспаривается некоторыми исследователями, в частности в [Dunn 1999] она трактуется как просодическая глоттализация. Инвентарь, зафиксированный нами в амгуэмском говоре в ходе полевой работы, представлен в таблице 2.

Таблица 2

Консонантный инвентарь амгуэмского говора (по нашим данным, нотация МФА)

	Губные	(Дентально-) альвеолярные	Палатальные	Велярные	Увулярные	Гортанные
Взрывные	<i>p</i>	<i>t</i>		<i>k</i>	<i>q</i>	<i>ʔ</i>
Фрикативные		<i>l</i> <i>s</i>				
Носовые	<i>m</i>	<i>n</i>		<i>ŋ</i>		
Аппроксиманты	<i>w</i>		<i>j</i>	<i>ɰ</i> ²		
Вибранты		<i>r</i>				

Чукотские согласные не противопоставлены по звонкости. Все звонкие согласные сонорные, все глухие — шумные. Контрасты по дополнительным артикуляциям в чукотском также отсутствуют.

С точки зрения типологии многие особенности чукотского консонантизма могут быть признаны редкими или неожиданными. Обратимся к некоторым таким особенностям.

В соответствии с закономерностью, выявленной в [Lindblom, Maddieson 1988], малым консонантным инвентарям свойственны простые артикуляции. К ним относятся самые частотные в языках мира согласные, при образовании которых взаимодействуют части речевого аппарата, близкие в состоянии покоя. С возрастанием числа согласных в инвентарях языков мы находим более нестандартно локализованные т. н. «усложненные» (elaborated) артикуляции, требующие дополнительного произносительного усилия. К ним

² Чукотский велярный аппроксимант в литературе и в практической транскрипции часто обозначается символом *y*, который в МФА предназначен для велярного фрикативного. В статье для чукотского велярного несмычного используется знак *ɰ* в соответствии с МФА. Транскрипции и орфографические записи для других языков и языков-предков приводятся по источникам.

относятся, в частности, из-за дальнейшей локализации увулярные согласные, а из-за дополнительной артикуляции — палатализованные, веляризованные или лабиализованные согласные. Артикуляции глухих носовых, латеральных и вибраторов также считаются «усложненными» за счет противоречия между отсутствием голоса и способом образования, типичным для сонорных и способствующим избеганию шума. Кроме того, выделяются «сложные» (complex) артикуляции, комбинирующие несколько «неудобных» жестов. Исследование [Lindblom, Maddieson 1988] показало, что чем больше в языке согласных, тем больше «усложненных» и «сложных» артикуляций.

Онлайн-атлас языковых структур WALS, составленный на основе взвешенной выборки из нескольких сотен языков мира, классифицирует языки в соответствии с числом смысло-различительных согласных сегментов. Выделяются языки с малым (6–14 согласных), умеренно малым (15–18 согл.), средним (22 ± 3 согл.), умеренно большим (26–33 согл.) и большим (34 и более согл.) инвентарями. Русский язык, к примеру, имеет умеренно большой (33 согл.) или большой (до 37) инвентарь в зависимости от решения о фонематическом статусе мягких заднеязычных и [ж:']; английский — средний. Чукотский язык (13 или 14 согл.) является по этой классификации языком с малым консонантным инвентарем — таких языков 16 % (89 языков из 563 в выборке). Несмотря на малый объем инвентаря, в чукотском языке мы находим не только ожидаемые базовые артикуляции, но и «усложненные»: увулярный смычный *q*, велярный аппроксимант *ɥ* и глухой шумный (по некоторым описаниям, и палатализованный, см., например, [Скорик 1961]) латеральный *l*. Эти согласные не только нарушают типичное соотношение объема инвентаря и артикуляционной сложности в понимании [Lindblom, Maddieson 1988], но и сами по себе нечасто встречаются в языках мира.

В выборке фонетической базы данных PHOIBLE увулярный глухой взрывной *q* входит в инвентарь всего 8 % языков (256 из 2186). В выборке фонетической базы данных LAPSyD (760 языков) увулярный глухой взрывной есть в 12,9 % языков, но таких языков, в которых при этом нет других увулярных, как в чукотском, всего 3,3 %.

В выборке PHOIBLE альвеолярный глухой латеральный фрикативный *l* входит в инвентарь всего 5 % языков (149 из 2186). В выборке LAPSyD (760 языков) этот согласный есть в 7,9 % языков, а таких языков, в которых при этом нет звонких латеральных, как в чукотском, всего пять, что составляет меньше процента.

Велярный аппроксимант, частотный в чукотском языке, — еще более редкое явление. В то время как *w* и *j* есть в большинстве фонетических систем, аппроксиманты других мест образования встречаются только в 2 % языков выборки WALS. В выборке LAPSyD велярный аппроксимант зафиксирован в 3 % языков, в выборке PHOIBLE — в 2 %.

Шумные и сонорные согласные выделяются по артикуляционным, акустическим и фонологическим свойствам в консонантном инвентаре любого языка. Универсальное соотношение классов — 30 % сонорных к 70 % шумных, или от двух третей до трех четвертей шумных от общего числа согласных [Lindblom, Maddieson 1988: 66]. В чукотском языке сонорных больше ожидаемого: из 14 согласных семь являются шумными, семь — сонорными. Это еще одно типологически нетривиальное свойство консонантной системы чукотского.

Сонорных согласных в чукотском языке семь: носовые *m*, *n* и *ŋ*, вибратор *r* и аппроксиманты *w*, *j* и *ɥ*. Составляя половину консонантного инвентаря, сонорные в речи даже более частотны, чем шумные. В спонтанном нарративе сонорные составляют 55,8 % всех согласных, шумные — 44,2 %³. Необычно высокая доля сонорных согласных в чукотской консонантной системе и в живой речи, их высокая фонетическая вариативность вместе с нестандартным составом инвентаря в целом делают чукотские сонорные интереснейшим предметом для исследования.

³ Для предварительной оценки использовались четыре спонтанных текста на амгуэмском говоре общим объемом 60 предложений (232 слова, 1757 звуков, из которых 1023 согласных). Употребление согласных в корневых и некорневых морфемах суммировалось.

Перейдем к двум сюжетам, которые касаются фонетической и акустической вариативности сонорных и представляют интерес для типологии и общей фонетики. В статье рассматриваются особенности поведения неносовых сонорных всех мест образования: альвеолярного вибранта (раздел 2), губного, палатального и велярного аппроксимантов (раздел 3).

Для анализа были записаны изолированные произнесения более сотни чукотских слов. В большинстве случаев слова совпадали со словарными входами из [Венстен 2015] и были выбраны так, чтобы разнообразие фонетических позиций для согласных было максимальным. В других случаях были записаны отдельные формы этих слов. В записи участвовали восемь женщин и один мужчина, носители чукотского языка, живущие в Амгуэме, в возрасте от 40 до 75 лет. Дикторы опознавали и переводили по 80–120 слов и произносили для записи каждое слово трижды. Слова записывались на портативный цифровой рекордер Zoom H2n с использованием встроенных микрофонов в формате WAV (частота дискретизации 48 кГц, 16 разрядов) в помещениях с минимальным уровнем реверберации, возможным в полевых условиях. Всего в фонетическую базу вошло 882 слова, т. е. более 2,5 тыс. изолированных произнесений. База не собиралась специализированно для исследования сонорных согласных, поэтому данных для количественного и статистического анализа пока недостаточно, однако собранный материал позволяет сформулировать некоторые наблюдения и обобщения.

2. Реализации *r* и изменчивость артикуляционной базы

В литературе отмечена высокая вариативность способа реализации чукотского альвеолярного *r*. П. Я. Скорик и А. С. Асиновский подробно описывают разнообразие способов образования *r* для нескольких фонетических позиций [Скорик 1961: 30–31; Асиновский 2003: 120–123], при этом их наблюдения различаются и плохо соотносятся с описанием М. Данна [Dunn 1999]. В амгуэмском говоре, по нашим данным, распространены все упомянутые авторами реализации *r*:

- двуударный вибрант;
- многоударный вибрант (3–4 размыкания смычки);
- одноударный звук, который может быть по классификации П. Ладефогед и И. Мэддисона [Ladefoged, Maddieson 1996] двух типов: «flap» (ср. *t* в амер. англ. *letter*) и «tap» (ср. *r* в исп. *caro*), в том числе глухой;
- фрикативный звук, в том числе глухой.

В начале слова и между гласными, например, в словах *rəswat* ‘личинки’, *rəkuətək* ‘застрять’, *jaranə* ‘дом’, *etleuqritqərīt* ‘сладкая каша’, *emiqsiretkəlʔin* ‘безработный’, встречаются все реализации, кроме многоударных, которые вообще нечасты. Однако некоторые носители предпочитают в этих позициях фрикативную реализацию и придерживаются почти исключительно ее. На конце слова фрикативный предпочтителен для всех говорящих, речь которых исследовалась, например, в словах *miqsir* ‘работа’, *kəqor* ‘пестрый олень’, *ikrer* ‘затравленный дикими зверями олень’. Если в позиции конца слова вибрант не реализуется как многоударный, он обычно оглушается. Об оглушении в сочетании с глухими согласными и на конце пишет и П. Я. Скорик применительно к говору села Уэлен [1961]. О способности чукотских сонантов приобретать свойства шумных в соседстве с шумными уже говорилось в разделе 1, это возможно и для вибранта, хотя и не так регулярно.

По описанию А. С. Асиновского [2003], способ образования *r* в начале слова зависит от ряда следующего за ним гласного. Перед непередними гласными (*a*, *o*, *u*, *ə*), по его наблюдениям, вероятнее вибрант, перед передними (*i*, *e*) — щелевой. Это замечание заслуживает доверия в том числе по типологическим причинам: вибрантам вообще свойственно сопротивляться дополнительной артикуляции палатализации. В работе [Hall 2000],

посвященной различительной мягкости вибрантов и невибрантов, на основании межъязыковых сравнений формулируются следующие закономерности: ни в каких языках не палатализуются только вибранты, и палатализованные вибранты есть только в тех языках, в которых есть и другие палатализованные согласные [Ibid.: 10, 12]. В работе, обобщающей сведения о палатализации в языках мира [Bateman 2007], отмечается, что вибранты палатализуются реже согласных других способов образования [Ibid.: 56]. Смягчению вибрантов препятствует не только способ образования (см. об аэродинамических трудностях, связанных с палатализацией вибрантов [Solé 2002]), но и то, что они все произносятся апикально, т. е. кончиком языка. Отмечено, что и вибранты, и другие апикальные смычные склонны или не палатализоваться вовсе, или, приобретая палатализацию, значительно изменять место и/или способ образования [Hall 2000]. С. В. Князев в статье о связи апикального произношения и консонантных противопоставлений в русских говорах [1991] также указывает на физиологическое противоречие между апикальностью и палатализацией согласного. Он отмечает, что апикальные взрывные смягчаются меньше, чем дорсальные (произносимые с опущенным кончиком языка), и на слух кажутся полумягкими или даже твердыми [Кузнецова 1969: 61, 69; Князев 1991].

По описаниям [Скорик 1961] и [Асиновский 2003], чукотский *r* апикальный⁴. Апикально в чукотском языке произносится не только *r*, но и другие альвеолярные согласные: *n*, *l* и *tf*. Для носового П. Я. Скорик также отмечает артикуляционную и акустическую близость английскому *n*, так же произносимому апикально, а позиционное смягчение *t* описывает как незначительное. По-видимому, именно апикальный характер артикуляции этих согласных обуславливает их превращение друг в друга в некоторых позициях.

Так, в амгуэмском чукотском вибрант, следуя за *t*, в абсолютном большинстве случаев реализуется как *t*, а после *n* — как соответствующий полувзвонкий или звонкий звук *d*, но всегда апикальный: *piŋwətrən* [piŋwəttən] ‘мука’, *jaatran* [jaattan] ‘крайняя яранга’, *ʔomrenatrək* [ʔomrenattək] ‘припереть (например, камнем)’; *ənrarərow* [əndarərow] ‘поодаль’, *wajənre* [wajənde] ‘вон туда’. Представляется, что это не строгое фонологическое или морфонологическое правило, а фонетический процесс артикуляционной природы: в речи некоторых носителей это явление встречается реже, а те говорящие, которые сами всегда произносят *nr* и *tr* как [nd] и [tt], поясняют, что такое произношение необязательно («можно и так и так», «кто как говорит», «иногда съедается буква»). Кроме того, в этом положении изредка на месте *r* встречается и короткий фриктивный звук.

Есть и множество других переходов-взаимодействий с участием альвеолярных, которые описываются в литературе как фонологические или морфонологические, т. е. по сути обязательные [Муравьева 1979]:

r + t > tt;
r + tf > tf tf;
r + l > ll;
r + n > nn;
r + j > tj;
l + r > rr;
tf + t > tt;
tf + r > rr,

а также «диссимилятивные» [Там же]:

tf + tf' > t tf;
l + l' > tl;
r + r' > tr.

⁴ П. Я. Скорик уточняет, что чукотский *r* произносится с «немного загнутым вверх» кончиком языка, так что «язык получает ложкообразную форму» [Скорик 1961: 30].

Во многих из этих случаев невозможно понять, действительно ли происходит изменение способа образования звука или же одноударный согласный, предшествуя другому согласному, был воспринят и описан как взрывной. На эти вопросы еще предстоит ответить. Так или иначе, по-видимому, все эти процессы основаны на общей установке чукотских альвеолярных согласных на апикальное произношение.

Парадигматическому взаимодействию между *r* и *t*, по-видимому, благоприятствует не только апикальность вибранта и взрывного сама по себе, но и способность вибранта реализоваться одноударно. Интересна в связи с этим позиция конца слова. В литературе описана вариативность конечных *r* и *t* в чукотском языке: *iɣər~iɣət* ‘сейчас’; *huur~huut* ‘вдруг’; *janor~janot* ‘первый’ [Скорик 1961; Dunn 1999]. Эта вариативность существует и в амгуэмском говоре. В некоторых словах вариант с *r* ассоциируется у носителей с мужским гендерлектом, а вариант с *t* — с женским, но в настоящее время это различие стирается: в Амгуэме классический женский вариант чукотского уже совсем не используется и вытеснен литературным мужским, см. раздел 1, а вариативность *r~t* в конце слова чаще воспринимается как не зависящая от гендера говорящего. В близкородственном чукотскому чавчувенском корякском существует вариативность между *t* и *tʃ* в конце слов, а сам *t* описывается как апикальный [Жукова 1972: 31]. Соответствия между женскими и мужскими вариантами слов показывают близость женского чукотского к корякским диалектам [Dunn 2014]. Таким образом, можно предположить, что в женском гендерном диалекте апикальное произношение *t* было более распространено, чем в мужском, что благоприятствовало переходу конечного вибранта в его одноударном произношении во взрывной из-за перцептивной и артикуляционной близости этих жестов.

Переход одноударного вибранта во взрывной на конце слова засвидетельствован также в других языках мира. Так, в языке терибе (< чибчанские языки, Панама, Коста-Рика) [Quesada 2000: 27] и языке дага (< семья даган, Новая Гвинея) [Murane 1974: 6] описан переход *r > t* на конце слов. Параллели можно найти и в речевой патологии на материале русской речи. Так, при произношении одноударного вибранта типа «тар» вместо нормативного многоударного или одноударного типа «флар», обусловленном дислалией, наблюдается следующая вариативность: в словах *корова*, *ворона*, *огород*, *тир* произносятся одноударные типа «тар», в слове *мухомор* — фрикативный согласный, а вместо *кома[p]*, *самова[p]*, *пова[p]*, *помидо[p]*, *бо[p]*, *офице[p]* — отчетливые *кома[m]*, *самова[m]*, *пова[m]*, *помидо[m]*, *бо[m]*, *офице[m]* (мальчик пяти лет, по нашим материалам). Мы уже говорили выше о том, что дрожачий способ образования плохо совместим с палатализацией. В неродственных языках карибском (< карибские языки, Венесуэла, Гайана, Бразилия) и ягуа (< семья пеба-ягуа, Перу) *r* палатализовался в *dʲ* [Bateman 2007: 58], и логично предположить в этом процессе промежуточную стадию в виде одноударного.

Выше мы постарались сопоставить описанные в литературе факты чукотско-корякских и других языков и показать, что в чукотском языке из-за общей установки на апикальные артикуляции и из-за способности вибранта реализоваться одноударно оказываются связанными самые разные процессы, затрагивающие альвеолярный ряд. Однако, если мы обратимся к данным амгуэмского чукотского, мы увидим, что амгуэмский говор демонстрирует отличия от других описанных говоров, противоречащие этой картине. Таких отличий два: 1) частотное позиционное смягчение амгуэмского *t* и 2) реализация единственного сибиланта не как *tʃ*, а как *s*.

По данным слухового анализа, в амгуэмском говоре альвеолярный согласный *t* допускает вариативное положение кончика языка. Наблюдается явная тенденция к произношению согласного не только кончиком языка, как при апикальной артикуляции, но всей передней поверхностью языка (т. е. ламинально), подобно тому, как артикулируется *m* в русском языке (см., например, [Князев 2009]). При этом некоторые носители все еще предпочитают апикальное произношение *t*, то есть смыкание с альвеолами только кончика языка, а не всей его передней части. Более того, в ряде позиций именно апикальное произношение является, по-видимому, наиболее распространенным у абсолютного большинства носителей.

Если же предпочитается ламинальная артикуляция *t*, лучше всего она слышна перед гласными переднего ряда. В этой позиции *t* коартикуляционно палатализуется, причем, в противоположность описанному П. Я. Скориком [1961], в современном амгуэмском говоре это смягчение бывает довольно сильным. Многие носители произносят палатализованный *t* аффрицированно, подобно тому, как в русском литературном языке *m'* и *ð'* аффрицируются, по длительности смычки и шумовой части и по структуре шума сближаясь с аффрикатами [Кузнецова 1969; Князев 2009]. Сама по себе палатализация перед гласными переднего ряда является типологически ожидаемой [Bateman 2007], но она не обязательно сопровождается аффрицированностью. В чукотском языке нет корреляции по твердости-мягкости, и даже у носителей амгуэмского чукотского, в речи которых часто встречается палатализованный *t*, это изменение может быть нерегулярным перед *e* или лексически обусловленным и может не приводить к появлению фрикативного шума. Различие между палатализованным и непалатализованным *t* в целом с трудом осознается носителями и слабо контролируется. Даже если говорящие знают о том, что различие существует, и приписывают себе ту или иную стратегию (например, уверены, что всегда говорят «твердо», в том числе в русских словах типа *термометр*, или считают, что нужно говорить «мягко» в именах, но «твердо» в глаголах), они не следуют одной стратегии и произносят согласный вариативно.

Кроме позиции перед передними гласными, дополнительную палатализацию *t* регулярно вызывают следующие за ним согласные *w* и *ɥ* (*we[tʰ]ɥaw* ‘слово’, *wulqə[tʰ]win* ‘вечер’). Смягчение в этих контекстах также необязательно, но, по нашим наблюдениям, палатализация происходит в большинстве случаев. В сочетании *tw* при отсутствии палатализации у всех говорящих взрывной перед аппроксимантом реализуется как апикальный, что приводит к увеличению шумовой составляющей и дает эффект, сходный с аффрикатизацией. Таким образом функционально сближаются небольшое придыхание апикального альвеолярного согласного и аффрицированность, характерная скорее для зубного ламинального. Это сближение засвидетельствовано типологически (см., например, об аффрикатизации придыхательных в стандартном британском английском [Buizza, Plug 2012] и андалусийском испанском [Del Saz 2019]).

Мы уже приводили свидетельства в пользу того, что первоначально артикуляционной базе чукотского языка и других чукотско-корякских идиомов было свойственно апикальное произношение всех альвеолярных согласных. И если одного сочетания *tw*, по-видимому, недостаточно, чтобы утверждать, что придыхательное произношение апикальных взрывных стало причиной аффрикатизации чукотского *t* при палатализации, влияние русского языка здесь кажется более убедительным объяснением.

Вторым фактом, не вписывающимся в общую апикальную направленность чукотской артикуляционной базы, является качество единственного сибиллянта амгуэмского говора. В литературе, посвященной восточным говорам чукотского, сибиллянт в мужском чукотском описывается скорее как постальвеолярный (или альвеоло-палатальный) *tʃ* и обозначается в нормативной кириллической орфографии буквой *ч*. В грамматике [Dunn 1999] и других работах автора, привлекающих данные говоров Анадыря и поселков к югу от него, сибиллянт мужского чукотского реализуется в свободной вариативности как альвеолярный и постальвеолярный. Единственный сибиллянт женского чукотского во всех источниках описывается как зубная аффриката *ts*. В амгуэмском говоре, утратившем гендерные диалекты, единственный сибиллянт реализуется исключительно как альвеолярный. При этом положение кончика языка не специфицировано, возможны и ламинальные, и апикальные артикуляции, хотя, по-видимому, более «свистящие» ламинальные артикуляции частотнее (ср. оценку акустической вариативности *s* в разных языках [Зибер, Мороз 2019]). Ламинальные артикуляции у фрикативного, как и у соответствующего взрывного, более заметны перед гласными переднего ряда, где он палатализуется.

Мы не знаем, что послужило толчком к распространению в некоторых чукотских говорах альвеолярного звука вместо постальвеолярного/альвеолопалатального, но само

по себе это предпочтение типологически ожидаемо. По данным базы PHOIBLE, альвеолярный *s* (или неспецифицированный переднеязычный, обозначенный как *s*) встречается в языках чаще всех других сибилантов (67 % языков), в то время как *tʃ* имеется в меньшем количестве языков (40 %). По данным базы LAPSyD, включающей 701 язык, альвеолярный *s* (или неспецифицированный переднеязычный, обозначенный как *s*) имеется в 70 % языков, при этом в 40 % языков базы это единственный сибилант. А вот *tʃ* (40 % языков) бывает единственным сибилантом исключительно редко — в базе LAPSyD таких языков всего 2,4 % (17 языков из 701). Таким образом, для языка, в котором есть только один сибилант, типологически ожидаема альвеолярная реализация. Эта единственность *s* могла быть поддержана утратой гендерных диалектов, традиционно характеризовавшихся сибилантами разных локальных рядов и создававших вместо фонологического противопоставления внутри одной языковой системы социокультурное противопоставление внутри сообщества.

Неожиданные фонетические варианты *r* подсказывают необходимость более широкого взгляда на альвеолярное место образования чукотских согласных. Ламинальные произнесения взрывного, особенно заметные при палатализации, в том числе с аффрикатизацией; сдвиг в альвеолярную область и палатализация фрикативного; необязательность произношения *tr* как [tʃ] (при допустимости и даже нормативности апикальных произнесений всех альвеолярных) — эти и некоторые другие факты современного амгуэмского чукотского позволяют говорить о том, что в этом идиоме происходит или даже уже произошел сдвиг артикуляционной базы от традиционно чукотской апикальной доминанты к ламинальной, свойственной русскому языку. Перечисленные явления в амгуэмском или других чукотских идиомах не были описаны в более ранней литературе и, по-видимому, стали распространены не ранее второй половины XX в. В это же время можно говорить об усилении влияния русского языка: монолингвальные дети, выросшие в тундре, увозились, зачастую против воли, в поселковые школы-интернаты, где после освоения русского говорить по-чукотски вне уроков родного языка, по свидетельствам наших консультантов, запрещалось, из-за чего использование чукотского уже с детства часто было искусственно ограничено у поколения носителей 1950-60-х годов рождения. Одновременно в этом поколении амгуэмских чукчей был утерян женский гендерный диалект, который наши консультанты сейчас восстанавливают по воспоминаниям о речи матерей и даже бабушек и в котором, судя по всему, обязательность апикального произношения альвеолярных согласных поддерживалась рядом фонетических правил. Не исключено, что сдвиг к ламинальности, возможно, в меньшей степени проявляющийся в других вариантах чукотского, со временем мог бы привести к большему, чем сейчас, различиям в фонетике чукотских идиомов. Это еще предстоит исследовать.

3. Вариативность аппроксимантов и стабильность системы

3.1. Реализации аппроксимантов в амгуэмском чукотском

Исследователи единодушны в том, что в чукотском языке различаются губной, палатальный и велярный неносовые сонорные [Скорик 1961: 29–31; Dunn 1999: 43–44]. П. Я. Скорик [1961: 29] отмечает разнообразие реализаций этих согласных и приводит их распределение для восточного (уэленского) говора. Описание П. Я. Скорика во многом не соответствует тому, что фиксируется в Амгуэме. Первая возможная причина этого состоит в том, что в говоре, который лег в основу грамматики, аппроксиманты произносились иначе. Вторая — в том, что от фундаментальной грамматики ожидается упорядоченность, и те обобщения, которые делает П. Я. Скорик, могут намеренно обходить совсем уж хаотичную

вариативность, даже если она имела место в реальности. Третья причина может быть в том, что за шесть с лишним десятков лет, прошедших со времени фиксации чукотского языка П. Я. Скориком, язык значительно изменился, в том числе под влиянием русского языка и в целом в условиях языкового сдвига. Многие носительницы и носители чукотского языка, с которыми нам повезло работать с 2016 г., были совсем малы или только начали учиться в интернате в то время, когда вышла грамматика П. Я. Скорика, и как раз в их поколении, по-видимому, и начался языковой сдвиг, последствия которого мы наблюдаем сейчас.

В современном амгуэмском говоре и в потоке речи, и при изолированном произнесении слов можно услышать множество вариантов и губного, и велярного сонантов. Палатальный *j* в этом отношении более предсказуем, поэтому дальше речь пойдет прежде всего о первых двух. Для систематизации вариантов ниже используется нотация МФА [2020], но за тем, какие символы выбраны для каждого произнесения и какие артикуляции ему поставлены в соответствие, не стоит автоматической обработки и вычисления уровня шума — это исключительно слуховой анализ записей и визуальный анализ их спектрограмм.

Для губно-губного сонорного выделены следующие реализации (непалатализованный и палатализованный варианты даются через запятую):

- губно-губное место образования:
 - губно-губной аппроксимант *w*, *ɥ*;
 - губно-губной глухой фрикативный или аппроксимант *ɸ*, *ɸʲ*.
- губно-зубное место образования:
 - губно-зубной фрикативный *v*, *vʲ*;
 - губно-зубной аппроксимант *ʋ*, *ʋʲ*;
 - губно-зубной глухой фрикативный *f*, *fʲ*.
- вокализованные реализации:
 - сочетание *iw*;
 - сочетание *iv*;
 - слоговой *w*, неотличимый от гласного *и*.
- *∅*.

Для велярного сонорного выделены реализации (непалатализованный и палатализованный варианты даются через запятую):

- велярный аппроксимант *ɥ*, *ɥʲ~j*;
- велярный звонкий фрикативный *ɣ*, *ɣʲ~j*;
- велярный глухой фрикативный *x*;
- велярный звонкий взрывной *g* (редко);
- велярная глухая аффриката *kx* (редко);
- гортанный звонкий спираント *ɦ* (редко; возможно, только в междометиях типа *taham* ‘давай!’);
- *∅*.

Несмотря на дробность классификации, перечисленные реализации в большинстве случаев хорошо различаются на слух. Воспринимаемым различиям закономерно соответствуют различия в спектральной картине, связанные с четырьмя основными параметрами: наличие / отсутствие частоты основного тона (ЧОТ), наличие / отсутствие формантной структуры, наличие / отсутствие шума и движение формант соседних гласных. Глухим реализациям свойственно отсутствие ЧОТ, в отличие от звонких. Шумным реализациям (фрикативным) свойственны отсутствие формантной структуры и фрикативный шум, а сонорным (аппроксимантным) — отчетливая формантная структура и отсутствие или незначительное присутствие фрикативного шума. От гласных аппроксимант отличается приглушенными верхними формантами. Палатализованные реализации вызывают движение формант соседних гласных в направлении выше 2 кГц. Ниже представлены примеры реализаций, которые отличаются друг от друга по этим параметрам.

Спектрограммы двух произнесений слова *wəjeɥəɥɕən* ‘дыхание’ одной носительницы демонстрируют различие между фрикативным *v* (рис. 1) и аппроксимантом *w* (рис. 2).

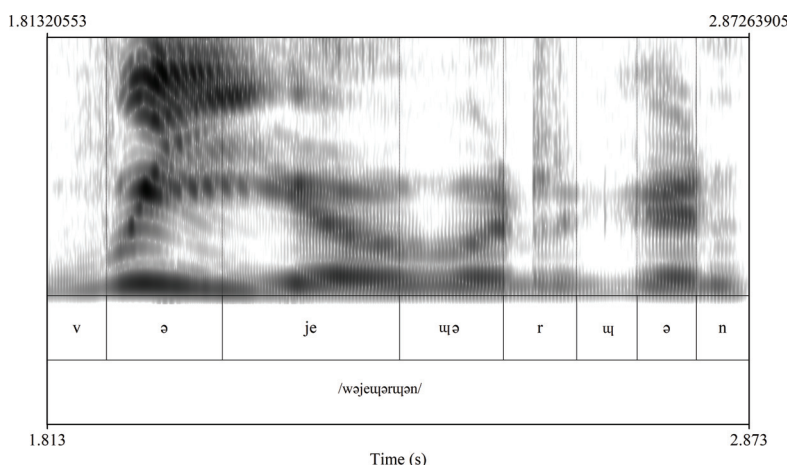


Рис. 1. Динамическая спектрограмма слова *wəjeɥəɥɕən* ‘дыхание’ (масштаб 6000 Гц)⁵

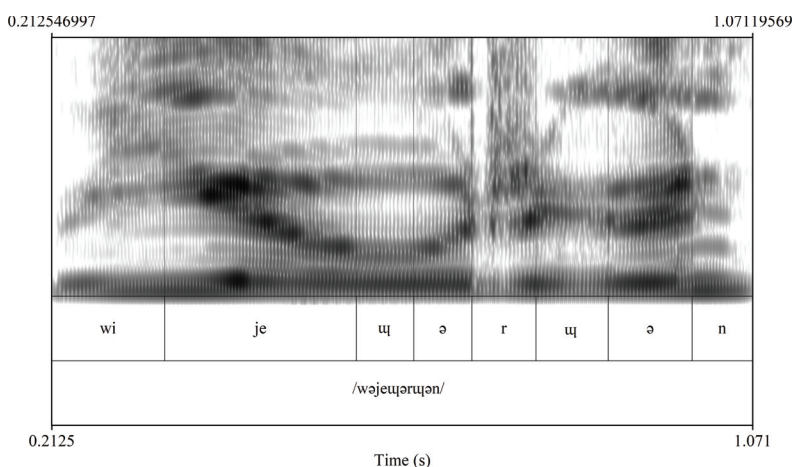


Рис. 2. Динамическая спектрограмма слова *wəjeɥəɥɕən* ‘дыхание’ (масштаб 6000 Гц)

На рисунке 1 можно различить слабый шум на первом сегменте, на рисунке 2 видимый шум отсутствует. На рисунке 1 граница между согласным и последующим гласным видна отчетливо: согласный не имеет формантной структуры, а у гласного она ярко выражена. Это звонкий шумный *v*. На рисунке 2 первый согласный имеет выраженную формантную структуру, подобную структуре последующего гласного, но с ослабленными высокими формантами. Это реализация губно-губного аппроксиманта *w*, имеющего, кроме того, дополнительную артикуляцию палатализации, что видно по форме второй форманты следующего гласного *i*. Палатальный аппроксимант *j* вызвал продвижение шва в передний ряд, и передний гласный вызвал палатализацию аппроксиманта *w*.

Сонорный, оглушенный под влиянием предшествующего фрикативного, можно сравнить со звонким на динамической спектрограмме слова *akwawarat* ‘иностранцы’ (рис. 3).

⁵ Здесь и далее спектрограммы созданы в программе Praat [Boersma, Weenink 2023].

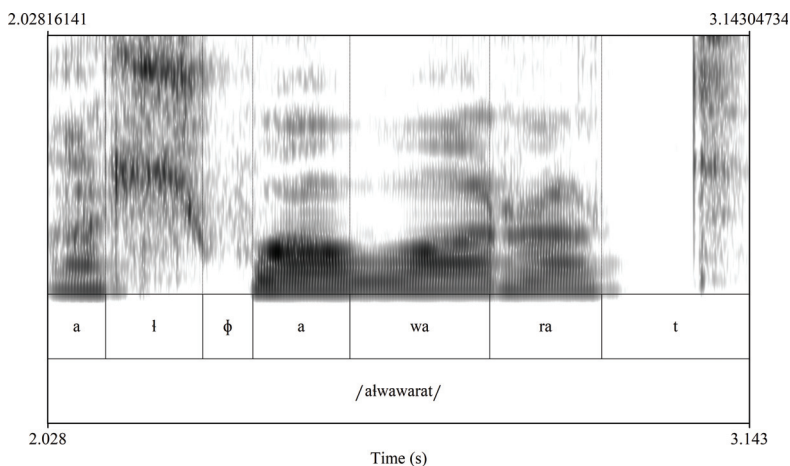


Рис. 3. Динамическая спектрограмма слова *alwawarat* 'иностранцы' (масштаб 6000 Гц)

На рисунке 3 у согласного, который обозначен как *φ*, отсутствует частота основного тона: он оглушился после глухого фрикативного *l*. Хотя на спектрограмме видно, что участок максимальной интенсивности латерального фрикативного понижается в направлении губно-губного, граница между *l* и *φ* отчетлива: для *φ* характерен шум слабой интенсивности без выраженных максимумов, тогда как *l* по уровню интенсивности сближается с сибилантами. Глухую реализацию губно-губного можно на этой спектрограмме сравнить и со звонкой аппроксимантной в следующем слове.

Для иллюстрации основных вариантов велярного аппроксиманта можно привести динамическую спектрограмму слова *liu 'яйцо', в котором корень редуцируется (рис. 4).*

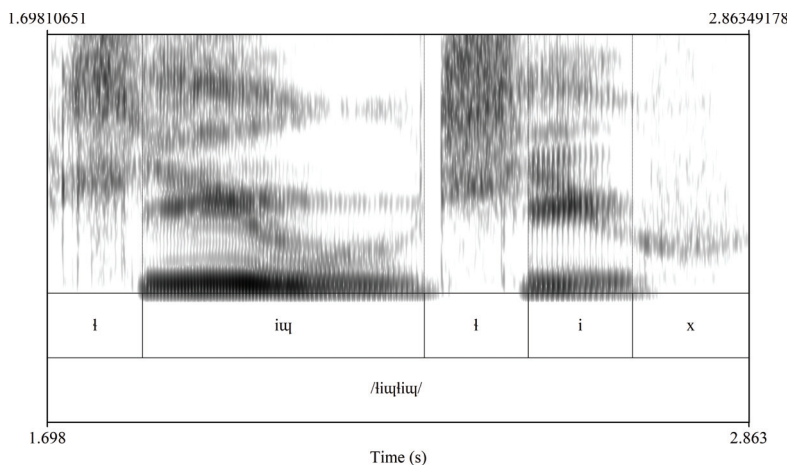


Рис. 4. Динамическая спектрограмма слова *liu 'яйцо' (масштаб 6000 Гц)*

Третий и шестой сегменты слова, конечные в повторениях корня, реализуются разными способами (см. рис. 4). Первый из них звучит и выглядит на спектрограмме как типичный аппроксимант с выраженной формантной структурой, отличающийся от соседнего гласного только интенсивностью формант. Второй велярный согласный в слове полностью оглушен: основной тон отсутствует (мы видим только слабоинтенсивное эхо основного тона предыдущего гласного в начале сегмента согласного), слабый фрикативный

шум сконцентрирован на тех же участках спектра, что и форманты соответствующего аппроксиманта.

Рассмотрим, от чего зависит выбор реализации в конкретном слове. Факторы выбора различаются для губно-губного и велярного.

Прежде всего прокомментируем место образования губного сонанта. Губно-губной аппроксимант (с фонацией или без) является основной, а для некоторых носителей и единственной, реализацией губного сонанта и встречается в записях всех носителей, чью речь мы анализировали, и во всех позициях. Другие носители предпочитают губно-зубное произношение. К тому или другому полюсу могут тяготеть говорящие, в речи которых встречаются реализации обоих типов. При этом почти все говорящие допускают губно-зубное произношение; его широкое распространение представляется следствием влияния русского языка. Можно выделить и фонетические условия, благоприятствующие появлению *v*. К таким условиям относится позиция после (глухих) фрикативных согласных *s* и *t*, а также соседство передних гласных *i*, *e* и гласных верхнего подъема *u*, *i*. К ним примыкает и шва, являющийся в амгуэмском чукотском гласным скорее средне-верхнего подъема. Есть и такие позиции, в которых колеблющиеся носители явно предпочитают аппроксимантный вариант фрикативному. Это позиция в соседстве с (глухими) взрывными (кроме предшествующего *p*) и сонорными, а также позиция конца слова. По-видимому, в исходной системе наиболее звучные реализации были привязаны к конечнослоговой позиции, а наименее звучные — к начальнослоговой (ср. [Асиновский 2003]), но в амгуэмском говоре в его современном состоянии аппроксимант возможен и в них.

Регулярное оглушение сонорного *w* в соседстве с шумным, чаще всего фрикативным, происходит только в направлении от предшествующего сегмента, т. е. прогрессивно: *a[tʃ]aŋqas* ‘в другой стороне’, но почти всегда *ku[wʃ]ewət* ‘чир (рыба)’. У носителей, предпочитающих губно-зубное произношение, согласный может оглушаться на конце слова, что является еще одним аргументом в пользу заимствования такой артикуляции из русского языка. В целом можно с уверенностью сказать, что в амгуэмском чукотском нет позиций, в которых оглушение аппроксимантов, в частности *w*, обязательно. Звонкий аппроксимант возможен и является основным вариантом в любой позиции. В положении между согласными, как в слове *nikwʔew* ‘высоко’, возможны реализации очень малой длительности или же полное выпадение согласного с сохранением лабиализации на предшествующем согласном: *[nikʔew]*. Аппроксимант взаимодействует с соседними оглушенными гласными, а также шва, который рядом с *w* становится оглушенным, сливается с ними: *kuwlewət* [ku:leut] ‘чир (рыба)’, *rərow* [rəro:] ‘далеко’.

Главное отличие велярного аппроксиманта от губного состоит в том, что распределение реализаций *ɰ* скорее зависит от фонетической позиции, чем от индивидуальных предпочтений носителей. Тем не менее все же можно выделить тех говорящих, которые предпочитают более шумные реализации и фрикативизацию, и тех, кто тяготеет к более звучному произношению, к звонким сонорным вариантам вплоть до слияния с соседними гласными. Аппроксимантная реализация наиболее частотна для *ɰ* и возможна в любой позиции. Она стандартна для конца слога (перед согласными), тогда как в начале слога (после согласных) чаще всего встречается фрикативный. Более звучная слоговая финаль соответствует принципу дисперсии сонорности — универсальной закономерности, в соответствии с которой во многих языках возрастание звучности в инициали должно быть резким, а спад звучности в финали — постепенным.

Оглушение во всех позициях необязательно; звук может оглушаться почти во всех позициях, но такой вариант никогда не бывает единственно возможным. Хотя для позиции конца слога характерна аппроксимантная реализация, в конце слова возможны почти все варианты (кроме редких: взрывного и аффрикаты). Выбор между глухими и звонкими зависит, по-видимому, в том числе от фразовой просодии: по нашим данным, в конце фразы безударные части слов могут оглушаться. В соседстве с *ə*, к которому *ɰ* близок по спектру, согласный иногда реализуется нулем, а между согласными часто выпадает.

Обобщить особенности произношения аппроксимантов в (амгуэмском) чукотском можно следующим образом. Во-первых, ориентация чукотского языка на слог, заметная в фонотактике и, возможно, в ударении, проявляется и в зависимости реализации *ɥ* и *w* от слоговой позиции. Во-вторых, сонорные согласные, которых в чукотском инвентаре столько же, сколько шумных, а в текстах даже больше, см. раздел 1, регулярно оглушаются и произносятся как шумные фрикативные и даже смычные, таким образом фонетически смещая системное равновесие в типологически ожидаемую сторону. На реализацию чукотских сонорных как фрикативных также можно смотреть как на частный случай типологически частого явления фортиции, или усиления, см. исследование [Bybee, Easterday 2019]⁶. В-третьих, артикуляция близких согласных русского языка может влиять на место образования чукотских согласных.

Анонимный рецензент статьи предлагает анализировать фонетическую вариативность чукотских сонорных с учетом важнейшего свойства чукотского консонантного инвентаря — отсутствия фонационных противопоставлений. Нам кажется это направление рассуждений продуктивным. Как уже отмечалось выше, все шумные согласные в чукотском глухие, а все звонкие — сонорные, то есть отличаются от глухих всегда и по фонации, и по способу образования. Однако и оглушенные сонорные, реализующиеся как фрикативные, не совпадают ни с какими чукотскими шумными, потому что в инвентаре согласных отсутствуют фрикативные соответствующих мест образования. Получается, что у несмычных согласных альвеолярное место образования закреплено за шумными согласными (всегда глухими и фрикативными), а губное, палатальное и велярное — за «сонорными» согласными (чаще звонкими, но обязательно). При отсутствии противопоставлений по способу среди несмычных в этих местах образования оказывается возможной широкая вариативность реализации согласных. Рецензент предлагает сравнить эту ситуацию с противоположной ситуацией австралийских языков без фонационных контрастов типа барди (< нью-ньюлские, Австралия) [Bowern et al. 2012; Kakadelis 2018], в которых сонорные стабильны, а вариативность характерна для т. н. «взрывных». Таким образом, возможно интерпретировать чукотские «аппроксиманты» без подробной спецификации по способу образования.

3.2. Лабиально-(палатально-)велярное взаимодействие в чукотско-камчатских и в неродственных языках

Несмотря на то что губной и велярный аппроксиманты в чукотском языке представляют разные фонемы, в некоторых позициях контраст между ними нейтрализуется. Рядом с огубленными гласными велярный аппроксимант лабиализуется и становится неотличим от губного. По-видимому, то же имеет в виду П. Я. Скорик [1961: 31–32], говоря об «ощутимых изменениях» качества сонанта в этой позиции. Заметим, что губно-губной аппроксимант *w*, в отличие от шумного губно-губного фрикативного *β*, обычно описывается как двухфокусный, то есть имеющий такую же сильную веляризацию, как и лабиализацию (ср. описание «лабиовелярный аппроксимант» в нотации МФА, в которой *w* вынесен из ряда как губных, так и велярных согласных). Это неудивительно: *w* артикуляционно ближе всего к гласному *u*, для которого сочетание заднеязычного и губного сужений формирует ротовой резонатор максимального объема, что создает аналогичный акустический эффект. Таким образом, чукотский губной аппроксимант отличается от велярного только огубленностью, которая может возникать коартикуляционно.

Некоторые такие случаи зафиксированы в словарях: для слова ‘вечер’ приводятся варианты *wuɫqətwin* и *ɥuɫqətwin* [Венстен 2015]. Одна носительница чукотского из Амгуэмы, начав произносить слово ‘вечер’ на диктофон, сама обратила внимание на то, что

⁶ Автор благодарит анонимного рецензента, обратившего внимание на данную интерпретацию.

в начале слова определить качество сегмента невозможно: «*wi* или *ци*? Черт его знает... *wi*... *ци*... хаха! <пытается написать слово в воздухе и так вспомнить, с какой “буквы” оно начинается> И *ц*, и *w*, все там вместе! Я не могу сказать [что там]». М. Фортескую, сравнивая чукотско-камчатские диалекты, для многих из них постулирует переход велярной прафонемы в губную рядом с огубленными гласными [Fortescue 2005: 8–9]. Во всех чукотско-корякских идиомах, кроме северо-восточного чукотского, начальный велярный **у* «может переходить» в *w* перед *и* или *о* [Ibid.]. В керекском же велярный **у* перед *и* в начале слова регулярно выпал, и можно предположить, что этому выпадению предшествовал переход *ц* в *w*, сблизивший согласный со следующим гласным до полного неразличения.

Переход велярного в губной происходит не только перед огубленными гласными, но и перед губными согласными. И. А. Муравьева приводит случаи типа *ətləɕ + poɕiɕən* > *ətlə[w]poɕiɕən* ‘отцовское копье’ в качестве примеров регулярного перехода велярных несмычных в губные перед губными согласными: *ɕp* > *wp*, *ɕm* > *wm*, *ɕw* > *ww* [Муравьева 1979: 129–131]. Этот переход, последовавший за диссимиляцией по способу образования, обусловил и более сложные преобразования: *kp* > *wp*, *km* > *wm* [Там же]. В амгуэмском чукотском в изолированном произнесении можно услышать в этой позиции и велярный, и губной: *n[əɕ]piɕin* и *n[əw]piɕin* ‘трудолюбивый, работающий’, при этом лабиализация распространяется и дальше на предшествующие сегменты, окрашивая гласный шва: *n[ɪw]piɕin*. Не только перед губным согласным, но и после него губной и велярный аппроксиманты иногда могут звучать и произноситься одинаково: *kottap[w]etɕaw*, *kottap[x]etɕaw*, *kottap[xw]etɕaw* ‘остроумное, меткое слово’ при отчетливом *w* в начале *wetɕaw* ‘слово’ в других контекстах.

Во всех перечисленных случаях мы наблюдаем распространение огубленности с одного сегмента на соседний, а иногда и дальше. Для огубленности вообще бывает характерна такая инертность и независимость от основных артикуляционных свойств соседних звуков. Они позволяют огубленности характеризовать сразу несколько сегментов и переходить через слоговые границы. Так, в русском литературном языке огубленный гласный первого предударного слога вызывает коартикуляционную лабиализацию предшествующего согласного, который, в свою очередь, влияет на качество предшествующего гласного: [нʲ] ‘пулярный’, [нʲ] ‘угад’, [зʲ] ‘рубешный’ [Пауфошима 1980], а в просторечии этот процесс может быть и прогрессивным: му[зʲ]кант [Реформатский 2008: 205; Пауфошима 1980: 63]. В скандинавских диалектах распространен лабиальный умлаут, ассимилятивная лабиализация корневого гласного под влиянием *и* или *у* следующего слога, ср. древнеисландские преобразования *land* ‘страна’ — *laend* (<**landu*) ‘страны’; *kalla* ‘звать’ — *kaellum* ‘мы зовем’, *a:r* ‘год’ — *o:rum* ‘годам’ [Ершова 1997: 65].

Лабиально-велярное взаимодействие свойственно разным диалектам чукотско-камчатской семьи. По М. Фортескую, в западном ительменском рефlekсами велярной прафонемы рядом со звонкими согласными выступают *v* или *w*, рядом с глухими — *f*, то есть губные согласные, в то время как губная **w* перешла не только в *f*, но и в велярный *x* (не в седанкинском говоре) [Fortescue 2005: 8–9]. При этом не всегда можно сформулировать надежные закономерности, которые помогали бы предсказывать внешний вид современных слов и морфем в ительменском, опираясь на данные родственных языков, поскольку ительменский язык отстоит значительно дальше от чукотско-корякских диалектов, чем они друг от друга. Так или иначе, в современном ительменском, по описанию А. П. Володина [1976], наблюдается следующая картина. Губные и велярные несмычные согласные — и глухие, и звонкие — различают слова: *фалч* ‘нож’ и *халч* ‘пора’ (служебное слово), *челфчелф* ‘брусника’ и *челхчелх* ‘шерсть (зверя), перо (птицы)’, *лвилх* ‘дикий олень’ и *луилх* ‘яйцо’. При этом многие слова зафиксированы в двух вариантах: *ветва/уетва* ‘прямо’, *пилвэпил/пилуэпил* ‘голод’, *вирник/уирник* ‘зверь’, *тизвин/тизуин* ‘ваш’, *хеневен/хенуен* ‘он сказал’, *чуфчуф/чухчух* ‘дождь’, *инфсенк/инхсенк* ‘иногда’, *фйал/хйал* ‘пурга’ и т. д. Вариативность, как показывают примеры, затрагивает и начальные, и конечные, и срединные позиции в слове перед гласным, после гласного, между гласными

и согласными. А. П. Володин пишет, что звонкие несмычные губного и веларного места образования артикулируются нечетко, если не различают смысла, и плохо различаются на слух, особенно звонкие.

Выше уже приводились примеры из чукотского, корякского и алюторского языков, в которых лабиальные и веларные переходят друг в друга в соседстве с губными артикуляциями. В ительменском мы наблюдаем смещения во всех позициях. Но и в чукотско-корякских диалектах можно найти свидетельства лабиально-веларного взаимодействия не только под влиянием ближайшего контекста. И. А. Муравьева [1979: 130] приводит примеры перехода веларного сонанта в губной в конечнослоговой позиции в корякском и алюторском, ср. алют. *ragatʔul* ‘куропаточье мясо’, *roro* (< *rawraw* < *ragrag*) ‘куропатка’. По наблюдениям И. А. Муравьевой, в конце редуцированных основ переход факультативен, и в целом сложно установить, насколько это регулярный переход и фонологическое правило, а не синхронная вариативность из-за артикуляционной и особенно акустической близости этих согласных, которая поддерживается особым статусом конечнослоговой позиции, о котором уже говорилось выше. В амгуэмском чукотском можно встретить разные произнесения перед передними гласными, что явно вызвано палатализацией, артикуляционно, акустически и перцептивно сближающей сонорные, ср. [ʉʲ] *esewperak* и [wʲ] *esewperak* ‘казаться веселым’, а у отдельных говорящих также и в словах, в которых трудно предположить влияние ближайшего контекста, ср. *pe[ʉʲ]litək* и *pe[wʲ]litək* ‘класть в рот; кусать; глотать’.

Данные некоторых языков указывают на особую связь между губными и веларными согласными. Подобные переходы объясняются тем, что при образовании и губных, и веларных согласных ротовая полость не разделяется препятствием, и это, в частности, влияет на акустическую картину соседнего гласного [Ohala 1993: 256]. Так, в некоторых западных говорах эвенского языка веларный фрикативный *y* может чередоваться с губно-губным *w*: [aʉʲi] — [awʲi] ‘гром’ [Дуткин 1995], ср. также примеры из говора села Себян-Кюёль (Якутия, Кобяйский улус), где [hawʲi] ‘старый’ соответствует [hagʲi] в стандартном эвенском [Aralova 2015: 17]. Предполагается, что пракрышчакский веларный несмычный согласный в кыпчакских языках между гласными подвергался их сильному воздействию и переходил в *w* или *j* [СИГТЯ 2002: 279]. В юго-западных и северо-восточных говорах русского языка, в отличие от говоров центра, не сформировалось последовательной корреляции по глухости-звонкости у губных согласных, и при усвоении новых слов на основании акустической близости регулярно происходили замены *f* > *p*, *f* > *xv*, *p* > *x* [Галинская 2009: 217]. В современном нестандартном русском языке, например, в речи изучающих русский как иностранный и в оговорках, по нашим наблюдениям, встречаются смещения падежных окончаний с *ф* и *х*: *жирафов/жирафах*, *пожаров/пожарах*; то, что смешиваются именно безударные окончания, может быть аргументом в пользу фонетической природы таких ошибок. Нейтрализация противопоставления губных и веларных под влиянием огубленности также фиксируется и в других языках. Так, для периода развития классического греческого языка из протоиндоевропейского реконструируется изменение огубленных веларных в губные [Ohala 1993: 242], а в современном японском языке *h* и *f* дополнительно распределены: губной выступает перед *и*, веларный — в остальных случаях (в поздних заимствованиях из западных языков ограничения менее строгие) [Алпатов и др. 2008: 42].

Переход типа *ʉe* > *je* артикуляционно прозрачен: место образования веларных, палатализованных за счет аккомодации передним гласным, значительно сдвигается вперед, так что накладывается на место образования палатального аппроксиманта *j*. Такие случаи фиксируются и в амгуэмском чукотском, ср. [ʉe] *ʉʉe*- и [je] *ʉʉe*- ‘каждый’ (префикс). Однако в исследуемом говоре *j* взаимодействует также с губным аппроксимантом. В позиции между *w* и *o*, палатальный аппроксимант исчезает: *arʔaw[jo]ʉʉən* > *arʔaw[uo]ʉʉən* ‘заварной чайник’, *ʔaw[jo]ʉ* > *ʔaw[uo]ʉ* ~ *ʔaw[o]ʉ* ‘двоюродная сестра’. Можно предположить, что артикуляция губно-губного веларизованного аппроксиманта и палатального аппроксиманта вступили в противоречие, и это противоречие разрешилось в пользу губного,

а не палатального. В отличие от согласных других мест образования (за исключением гортанной смычки), губные согласные не задействуют при образовании основного сужения языка. Таким образом, когда на основную губную артикуляцию накладывается дополнительная переднегубная артикуляция палатализации, согласный приобретает еще один фокус. Такая сложная артикуляция редко сохраняется даже при фонологическом статусе палатализации в языке. Так, в истории русского языка многократно имело место отверждение губных согласных на конце, то есть в позициях, где палатализация не поддерживалась вокалическим контекстом [Галинская 2009: 99–100]. Трудность усвоения палатализованных губных в неродном языке затрагивает не только смычные губные, но и фрикативные и аппроксиманты, см. [Зибер 2018] о влиянии родного чукотского на фонетику русской речи. Тенденцию поддерживает и относительная частотность аппроксимантов в языках мира. При том что губно-губной (веляризованный) *w* имеется в трех четвертях языков мира, его палатализованный вариант *ɥ* (ср. фран. *lui* ‘его’, *huit* ‘восемь’) исключительно редок [Ladefoged, Maddieson 1996: 322]. Итак, лабиализация и губное место образования плохо сочетаются с дополнительной артикуляцией палатализации, и согласные избегают совмещения этих артикуляционных движений. В случае чукотского $\eta aw[jo]t > \eta aw[(w)o]t$, по-видимому, это происходит за счет уподобления *j* предшествующему аппроксиманту и одновременно последующему огубленному гласному по ряду.

3.3. Некоторые обобщения

Губные и велярные аппроксиманты и фрикативные, глухие и звонкие, в разных языках смешиваются и переходят друг в друга. Это наблюдается и в амгуэмском чукотском, для чего есть акустические и артикуляционные причины. В статье приводятся в основном свидетельства лабиально-велярного взаимодействия, но и эти аппроксиманты в других позициях, и палатальный аппроксимант представляются не очень стабильными единицами. Они могут и вокализоваться до полного слияния с гласными, и оглушаться, становясь шумными.

Можно предположить, что различия между аппроксимантами вообще наследуют различия между соответствующими гласными, которые часто являются, как кажется, их диакроническими источниками. Гласные *i* и *u* частотны и есть почти во всех языках (за исключением линейных вокалических систем), а в малых (треугольных) вокалических системах ряд связан с огубленностью. Носители получают навык различения гласных *i* и *u* и успешно различают и соответствующие аппроксиманты, акустически сходные с ними. Аппроксимант *ɥ* же соответствует неогубленному гласному заднего ряда, который мы находим либо в четырехугольных системах типа тюркских, либо, менее часто, в очень расширенных треугольных системах, где раздвоение переднего ряда по огубленности могло дополниться точечным раздвоением заднего без формирования полноценных осей, как в четырехугольных системах. Если же в языке не различаются огубленные и неогубленные задние гласные, отсутствует и такой гласный, которому соответствует велярный аппроксимант *ɥ*. В таком случае различительная способность согласного не поддерживается гласным, а сам согласный, как показывают свидетельства самых разных языков, ее реализует не очень эффективно. Это может отчасти объяснять и типологическую редкость велярного аппроксиманта, и его неустойчивость в конкретных системах. В чукотском языке велярный аппроксимант удержался, но мы видим в реализации других согласных возможные пути разрешения этой нестабильности. Появление губно-зубных и потому преимущественно шумных реализаций *w*, расширение контекстов, в которых губно-зубная реализация допустима, со временем могут привести к более надежному различению губных как более шумных и велярных как более сонорных. Таким образом, появление новых по месту образования реализаций может рассматриваться не только как следствие

внешнего влияния русского языка, но и как процесс, обусловленный внутренним устройством фонетической системы чукотского языка.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Алпатов и др. 2008 — Алпатов В. М., Аркадьев П. М., Подлесская В. И. *Теоретическая грамматика японского языка*. В 2-х т. М.: Наталис, 2008. [Alpatov V. M., Arkadiev P. M., Podlesskaya V. I. *Teoreticheskaya grammatika yaponskogo yazyka* [A theoretical grammar of Japanese]. In 2 vols. Moscow: Natalis, 2008.]
- Асиновский 2003 — Асиновский А. С. *Сопоставительная фонетика чукотско-камчатских языков*. Дис. ... докт. филол. наук. СПб.: ИЛИ РАН, 2003. [Asinovskii A. S. *Sopostavitel'naya fonetika chukotsko-kamchatskikh yazykov* [Comparative phonetics of the Chukotko-Kamchatkan languages]. Cand. diss. St. Petersburg: Institute for Linguistic Studies, 2003.]
- Богораз 1934 — Богораз В. Г. Луораветланский (чукотский) язык. *Языки и письменность народов севера*. Ч. 3. *Языки и письменность палеоазиатских народов*. Крейнович Е. А. (ред.). М.; Л.: Учпедгиз, 1934. [Bogoraz V. G. The Luoravetlan (Chukchi) language. *Yazyki i pis'mennost' narodov severa*. P. 3. *Yazyki i pis'mennost' paleoaziatskikh narodov*. Kreinovich E. A. (ed.). Moscow; Leningrad: Uchpedgiz, 1934.]
- Богораз 1937 — Богораз В. Г. *Луораветланско-русский (чукотско-русский) словарь*. М.; Л.: Учпедгиз, 1937. [Bogoraz V. G. *Luoravetlansko-russkii (chukotsko-russkii) slovar'* [Luoravetlan-Russian (Chukchi-Russian) dictionary]. Moscow; Leningrad: Uchpedgiz, 1937.]
- Венстен 2015 — Венстен Ш. Чукотско-французско-англо-русский словарь. В 3-х т. Анадырь: Лема, 2015. [Vensten Sh. *Chukotsko-frantsuzsko-anglo-russkii slovar'* [Chukchi-French-English-Russian dictionary]. Anadyr: Lema, 2015.]
- Володин 1976 — Володин А. П. *Ительменский язык*. Л.: Наука, 1976. [Volodin A. P. *Itel'menskii yazyk* [The Itelmen language]. Leningrad: Nauka, 1976.]
- Володин 1997 — Володин А. П. Чукотско-камчатские языки. *Языки мира. Палеоазиатские языки*. Володин А. П., Вахтин Н. Б., Кибрик А. А. (ред.). М.: Индрик, 1997, 12–22. [Volodin A. P. The Chukotko-Kamchatkan languages. *Yazyki mira. Paleoaziatskie yazyki*. Volodin A. P., Vakhtin N. B., Kibrik A. A. (eds.). Moscow: Indrik, 1997, 12–22.]
- Володин, Скорик 1997 — Володин А. П., Скорик П. Я. Чукотский язык. *Языки мира. Палеоазиатские языки*. Володин А. П., Вахтин Н. Б., Кибрик А. А. (ред.). М.: Индрик, 1997, 23–39. [Volodin A. P., Skorik P. Ya. The Chukchi language. *Yazyki mira. Paleoaziatskie yazyki*. Volodin A. P., Vakhtin N. B., Kibrik A. A. (eds.). Moscow: Indrik, 1997, 23–39.]
- Галинская 2009 — Галинская Е. А. Историческая фонетика русского языка. М.: Изд-во Московского ун-та, 2009. [Galinskaya E. A. *Istoricheskaya fonetika russkogo yazyka* [Historical phonetics of Russian]. Moscow: Moscow Univ. Press, 2009.]
- Дуткин 1995 — Дуткин К. И. Аллаховский говор эвенов Якутии. СПб.: Наука, 1995. [Dutkin K. I. *Allaikhovskii govor evenov Yakutii* [The Allaikhov idiom of the Yakut Evens]. Saint Petersburg: Nauka, 1995.]
- Ершова 1997 — Ершова И. А. *Введение в германскую филологию*. Вып. 1. *Фонетика древнегерманских языков*. М.: Изд-во Московского ун-та, 1997. [Ershova I. A. *Vvedenie v germanskuyu filologiyu* [Introduction to Germanic philology]. Iss. 1. *Fonetika drevnegermanskikh yazykov* [Phonetics of the Old Germanic languages]. Moscow: Moscow Univ. Press, 1997.]
- Жукова 1972 — Жукова А. Н. *Грамматика корякского языка: фонетика, морфология*. Л.: Наука, 1972. [Zhukova A. N. *Grammatika koryakskogo yazyka: fonetika, morfologiya* [A grammar of Koryak: Phonetics, morphology]. Leningrad: Nauka, 1972.]
- Зибер 2018 — Зибер И. А. Фонетическая интерференция в русской речи чукчей (консонантизм). *Томский журнал лингвистических и антропологических исследований*, 2018, 1(19): 9–19. [Sieber I. A. Chukchi-Russian phonetic interference: Focus on consonants. *Tomsk Journal of Linguistics and Anthropology*, 2018, 1(19): 9–19.]
- Зибер, Мороз 2019 — Зибер И. А., Мороз Г. А. Исследование акустической вариативности с методом главных компонент. *Вестник Новосибирского государственного университета. Серия «Лингвистика и межкультурная коммуникация»*, 2019, 17(1): 49–64. [Sieber I. A., Moroz G. A. Estimating the acoustic variation of *s* via principal component analysis. *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2019, 17(1): 49–64.]

- Князев 1991 — Князев С. В. О связи особенностей артикуляционной базы говора с характером противопоставления глухих/звонких и твердых/мягких согласных. *Современные русские говоры*. Азарх Ю. С. (отв. ред.). М.: Наука, 1991, 28–35. [Knyazev S. V. On the connection of the basis of articulation of an idiom and the type of contrast between voiceless/voiced and hard/soft consonants. *Sovremennye russkie govory*. Azarkh Yu. S. (ed.). Moscow: Nauka, 1991, 28–35.]
- Князев 2009 — Князев С. В. О мягкости необычайной (заметки и загадки о русской фонетике). *Вопросы русского языкознания*. Вып. 13. *Фонетика и грамматика: настоящее, прошедшее, будущее*. Князев С. В. (ред.). М.: Изд-во Московского ун-та, 2009, 71–91. [Knyazev S. V. On exceptional softness (notes and riddles about Russian phonetics). *Voprosy russkogo yazykoznaniiya*. Iss. 13. Knyazev S. V. (ed.). Moscow: Moscow Univ. Press, 2009, 71–91.]
- Кузнецова 1969 — Кузнецова А. М. Некоторые вопросы фонетической характеристики явления твердости/мягкости согласных в русских говорах. *Экспериментально-фонетическое изучение русских говоров*. Высотский С. С., Кузнецова А. М., Пауфوشима Р. Ф. (ред.). М.: Наука, 1969, 35–137. [Kuznetsova A. M. Some issues of phonetic description of hardness/softness in Russian dialects. *Ekspperimental'no-foneticheskoe izuchenie russkikh govorov*. Vysotskii S. S., Kuznetsova A. M., Paufoshima R. F. (eds.). Moscow: Nauka, 1969, 35–137.]
- Муравьева 1979 — Муравьева И. А. *Сопоставительное исследование морфологии чукотского, корякского и алыторского языков*. Дис. ... канд. филол. наук. М.: МГУ, 1979. [Muravyova I. A. *Sopostavitel'noe issledovanie morfonologii chukotskogo, koryakskogo i alyutorskogo yazykov* [A comparative study of Chukchi, Koryak and Alyutor morphology]. Cand. diss. Moscow: Moscow State Univ., 1979.]
- Пауфوشима 1980 — Пауфوشима Р. Ф. Активные процессы в современном русском литературном языке: ассимилятивные изменения безударных гласных. *Известия Академии Наук СССР. Серия литературы и языка*, 1980, 39(1): 61–68. [Paufoshima R. F. Active processes in Modern Standard Russian: Assimilative changes in unstressed vowels. *Izvestiya Akademii Nauk SSSR. Seriya literatury i yazyka*, 1980, 39(1): 61–68.]
- Пупынина 2013 — Пупынина М. Ю. Чукотский язык: география, говоры и представления носителей о членении своей языковой общности. *Acta Linguistica Petropolitana*, 2013, 9(3): 245–260. [Pupynina M. Yu. The Chukchi language: Geography, idioms and the speakers' perception of distinctions in their linguistic community. *Acta Linguistica Petropolitana*, 2013, 9(3): 245–260.]
- Реформатский 2008 — Реформатский А. А. *Введение в языковедение*. Виноградов В. А. (ред.). 5-е изд., исп. М.: Аспект-Пресс, 2008. [Reformatskii A. A. *Vvedenie v yazykovedenie* [Introduction to the study of language]. Vinogradov V. A. (ed.). 5th edn., corr. Moscow: Aspekt-Press, 2008.]
- СИГТЯ 2002 — Тенишев Э. Р. (отв. ред.). *Сравнительно-историческая грамматика тюркских языков. Региональные реконструкции*. М.: Наука, 2002. [Tenishev E. R. (ed.). *Sravnitel'no-istoricheskaya grammatika tyurkskikh yazykov. Regional'nye rekonstruktsii* [A comparative and historical grammar of the Turkic languages. Regional reconstructions]. Moscow: Nauka, 2002.]
- Скорик 1961 — Скорик П. Я. *Грамматика чукотского языка*. Ч. 1. *Фонетика и морфология именных частей речи*. М.; Л.: Изд-во Академии Наук СССР, 1961. [Skorik P. Ya. *Grammatika chukotskogo yazyka* [Chukchi grammar]. P. 1. *Fonetika i morfologiya imennykh chastei rechi* [Phonetics and morphology of nominal categories]. Moscow; Leningrad: Academy of Sciences of the USSR Publ., 1961.]
- Скорик 1977 — Скорик П. Я. *Грамматика чукотского языка*. Ч. 2. *Глагол, наречие, служебные слова*. Л.: Наука, 1977. [Skorik P. Ya. *Grammatika chukotskogo yazyka* [Chukchi grammar]. P. 2. *Glagol, narechiye, sluzhebnye slova*. [Verb, adverb, functional words]. Leningrad: Nauka, 1977.]
- Aralova 2015 — Aralova N. *Vowel harmony in two Even dialects: Production and perception*. Ph.D. diss., Univ. of Amsterdam. Utrecht: LOT, 2015.
- Bateman 2007 — Bateman N. *A crosslinguistic investigation of palatalization*. Ph.D. diss. San Diego: Univ. of California San Diego, 2007.
- Boersma, Weenink 2023 — Boersma P., Weenink D. *Praat: doing phonetics by computer* [Computer program]. 2023. <http://www.praat.org/>.
- Bogoras 1922 — Bogoras W. Chukchee. *Handbook of American Indian languages*. Vol. 2. Boas F. (ed.). Washington: Government Printing Office, 1922, 631–903.
- Bowern et al. 2012 — Bowern C., McDonough J., Kelliher K. Bardi. *Journal of the International Phonetic Association*, 2012, 42(3): 333–351.
- Buizza, Plug 2012 — Buizza E., Plug L. Lenition, fortition and the status of plosive affrication: The case of spontaneous RP English /t/. *Phonology*, 2012, 29(1): 1–38.
- Bybee, Easterday 2019 — Bybee J., Easterday S. Consonant strengthening: A crosslinguistic survey and articulatory proposal. *Linguistic Typology*, 2019, 23(2): 263–302.

- Del Saz 2019 — Del Saz M. From postaspiration to affrication: New phonetic contexts in Western Andalusian Spanish. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia, 2019*. Calhoun S., Escudero P., Tabain M., Warren P. (eds.). Canberra: Australasian Speech Science and Technology Association Inc., 2019, 760–764.
- Dunn 1999 — Dunn M. *A Grammar of Chukchi*. Ph.D. diss. Canberra: Australian National Univ., 1999.
- Dunn 2014 — Dunn M. Gender determined dialect variation. *The expression of gender*. Corbett G. G. (ed.). Berlin: Mouton de Gruyter, 2014, 39–67.
- Fortescue 2005 — Fortescue M. D. *Comparative Chukotko-Kamchatkan dictionary*. Berlin: Mouton de Gruyter, 2005.
- Glottolog — Hammarström H., Forkel R., Haspelmath M., Bank S. *Glottolog 4.7*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2022. <http://glottolog.org>.
- Hall 2000 — Hall T. A. Typological generalizations concerning secondary palatalization. *Lingua*, 2000, 110: 1–25.
- Kakadelis 2018 — Kakadelis S. M. *Phonetic properties of oral stops in three languages with no voicing distinction*. Ph.D. diss. New York: City Univ. of New York, 2018.
- Kenstowicz 1986 — Kenstowicz M. The phonology of Chukchee consonants. *Studies in Linguistic Sciences*, 1986, 16(1): 79–96.
- Krause 1980 — Krause S. R. *Topics in Chukchee phonology and morphology*. Ph.D. diss. Champaign: Univ. of Illinois at Urbana-Champaign, 1980.
- Ladefoged, Maddieson 1996 — Ladefoged P., Maddieson I. *The sounds of the world's languages*. Cambridge (MA): Blackwell, 1996.
- LAPSyD — Maddieson I., Flavier S., Marsico E., Pellegrino F. *LAPSyD: Lyon-Albuquerque Phonological Systems Databases, Version 1.0*. 2014–2016. <http://www.lapsyd.ddl.ish-lyon.cnrs.fr/lapsyd/>.
- Lindblom, Maddieson 1988 — Lindblom B., Maddieson I. Phonetic Universals in Consonant Systems. *Language, speech and mind*. Li C., Hyman L. M. (eds.). London: Routledge, 1988, 62–78.
- Murane 1974 — Murane E. *Daga grammar: From morpheme to discourse*. Norman: Summer Institute of Linguistics, 1974.
- Ohala 1993 — Ohala J. J. The phonetics of sound change. *Historical linguistics: Problems and perspectives*. Jones C. (ed.). London; New York: Longman, 1993, 237–278.
- PHOIBLE — PHOIBLE 2.0. Moran S., McCloy D. (eds.). Jena: Max Planck Institute for the Science of Human History, 2019. <http://phoible.org>.
- Solé 2002 — Solé M. J. Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics*, 2002, 30(4): 655–588.
- Quesada 2000 — Quesada J. D. *A grammar of Teribe*. München: Lincom Europa, 2000.
- WALS — *The World Atlas of Language Structures Online*. Dryer M. S., Haspelmath M. (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. <http://wals.info>.

Получено / received 17.05.2023

Принято / accepted 26.09.2023

«Язык и семиотика тела».**Исследования Г. Е. Крейдлина по воплощенной семиотике**

© 2024

Юлия Владимировна Николаева

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

Национальный исследовательский университет «Высшая школа экономики»,

Москва, Россия; julianikk@gmail.com

Аннотация: Взаимодействие тела и среды посредством невербальных знаков и присвоение этим знакам лингвистических обозначений определяет переход от простого существования к семиотическому уровню. Один из основных принципов этого подхода заключается в следующем: тело как инструмент взаимодействия с окружающей средой и другими субъектами служит основой для оценки всех контекстов и событий прежде всего как благоприятных (для существования и развития организма), неблагоприятных и нейтральных, а также для определения значимости отдельных событий. Таким образом, тело становится источником смыслов, заложенных в знаках. Воплощенный подход к исследованию символических значений можно встретить в самых разных областях начиная с психоанализа, философии и работ по теории искусства и заканчивая исследованиями в области антропологии и когнитивистики. Накопленные на современном этапе знания позволяют более подробно изучить, как движения тела и в целом осознание тела и его частей в коммуникативном пространстве встраиваются в знаковую коммуникацию. В двухтомнике «Язык и семиотика тела», написанном коллективом авторов под руководством Г. Е. Крейдлина (2020), движения, части тела и их языковые обозначения рассматриваются как целостная и непротиворечивая знаковая система, что соответствует подходу, принятому в рамках воплощенной семиотики, в которой коммуникативные события приобретают смысл на основе опыта их существования в окружающем мире и телесного взаимодействия с ним. В статье рассматривается текст коллективной монографии «Язык и семиотика тела» и современные подходы к изучению символической репрезентации телесного как культурно-специфического феномена. Хотя сам Г. Е. Крейдлин говорит о своей области как о невербальной семиотике или мультимодальной лингвистике, можно сказать, что выход книги «Язык и семиотика тела» знаменует собой значительный вклад в представление воплощенной семиотики в русскоязычном лингвистическом пространстве.

Ключевые слова: жестикуляция, обзор, семиотика, телесность

Благодарности: Исследование выполнено в рамках Программы развития Междисциплинарной научно-образовательной школы Московского университета «Сохранение мирового культурно-исторического наследия».

Для цитирования: Николаева Ю. В. «Язык и семиотика тела». Исследования Г. Е. Крейдлина по воплощенной семиотике. *Вопросы языкознания*, 2024, 2: 143–158.

DOI: 10.31857/0373-658X.2024.2.143-158

“Language and body semiotics”.**Grigory E. Kreidlin’s research on embodied semiotics****Yulia V. Nikolaeva**

Lomonosov Moscow State University, Moscow, Russia; HSE University, Moscow, Russia;

julianikk@gmail.com

Abstract: The interaction between the body and the environment through non-verbal signs and the assignment of linguistic designations to these signs determines the transition from mere existence to the semiotic level. One of the basic principles of this approach is the following: the body as a tool in interaction with the environment and other subjects serves as a basis for evaluating all contexts and events primarily as favorable (for the existence and development of the body), unfavorable and neutral, as well as for determining the importance of individual events. In this way, the body becomes the source of the meanings inherent in signs. The body as a source of symbolic meanings has been considered in a wide range of fields, from psychoanalysis, philosophy, and art theory to studies in anthropology and cognitive science. The knowledge accumulated at the present stage allows us to study in more detail how body movements and, in general, the presence of the body and its parts in the communicative space are embedded in symbolic communication. The two-volume “Language and semiotics of the body”, written by a team of authors led by G. E. Kreidlin (2020), considers movements, body parts and their linguistic designations as a coherent and consistent sign system, which corresponds to the approach adopted within embodied semiotics, in which communicative events acquire their meaning based on the experience of their existence in the surrounding world and bodily interaction with it. This article considers the text of the collective monograph “Language and semiotics of the body” and modern approaches to the study of the symbolic representation of the body language as a culturally specific phenomenon. Although G. E. Kreidlin himself speaks of his field as non-verbal semiotics or multimodal linguistics, the publication of the book “Language and Semiotics of the Body” can be said to mark a significant contribution to the representation of embodied semiotics in the Russian-speaking linguistic space.

Keywords: embodiment, gestures, semiotics, survey

Acknowledgements: This research was performed according to the Development Program of the Interdisciplinary Scientific and Educational School of Moscow University “Preservation of the World Cultural and Historical Heritage”.

For citation: Nikolaeva Yu. V. “Language and body semiotics”. Grigory E. Kreidlin’s research on embodied semiotics. *Voprosy Jazykoznanija*, 2024, 2: 143–158.

DOI: 10.31857/0373-658X.2024.2.143-158

Введение

Телесная (или воплощенная) семиотика (также семиотика тела) — область, изучающая значение сообщений, передаваемых через естественные движения и указания на части тела. Внимание к этой сфере можно увидеть в самых разных областях начиная с литературоведческих работ, изучающих, как писатели выражали свой замысел через внешность и физическое действие и взаимодействие персонажей. Изучение мифов также открывало символические значения частей тела и физических действий персонажей в мире. Другие сферы, изучавшие разные области искусства, такие как театр, танец или живопись, обращали внимание на то, как посредством движений или даже знакового присутствия в коммуникации тело могло выражать разные значения, в том числе абстрактные. С другой стороны, психология, в частности психоанализ, еще с работ З. Фрейда учитывает соматические проявления и интерпретирует их как некое сообщение, не обязательно какому-то определенному адресату.

Исследования невербальной коммуникации, которые начались с работ антропологов, также большое внимание уделяли тому, как люди ведут себя в естественной обстановке, как проявляют себя и как эти их действия интерпретируются окружающими, которым доступен этот код.

В последние десятилетия с развитием воплощенного подхода в когнитивной психологии область невербальной семиотики приобрела новое измерение, включающее не только передачу информации, но и формирование концептов в процессе взаимодействия организма со средой.

Вклад лингвистики в эту междисциплинарную область можно обозначить как попытку описать значения невербальных знаков вне связи с конкретными участниками

коммуникации и ситуацией, в которой происходит взаимодействие. При этом лингвистика получает возможность исследовать, как визуально-кинестические знаки (такие как жесты, позы, выражение лица и т. п.), с одной стороны, выступают как единицы, аналогичные словам естественного языка, т. е. обладают собственным значением, а с другой — остаются обусловлены опытом взаимодействия говорящего с окружающим миром и тем, что это взаимодействие приносит человеку как субъекту, обладающему телом с его возможностями и ограничениями.

Выход в свет двухтомника «Язык и семиотика тела», написанного коллективом авторов под руководством Г. Е. Крейдлина [2020], рассматривает движения, части тела и их языковые обозначения как связную и последовательную знаковую систему, что соответствует подходу, принятому в рамках воплощенной семиотики (*embodied semiotics*), в котором коммуникативные события приобретают свое значение, исходя из опыта их существования в окружающем мире и телесного взаимодействия с ним. В данной статье сопоставляются текст упомянутой коллективной монографии и современные подходы к изучению того, как в языке отражается символическое представление о телесном как культурно-специфическом феномене. Осмысление взаимодействия тела и среды через невербальные знаки и присвоение этим знакам языковых обозначений определяет переход от простого существования к семиотическому уровню. Одним из базовых принципов такого подхода можно назвать следующий: тело как инструмент при взаимодействии с окружающей средой и другими субъектами служит основанием для того, чтобы оценивать все контексты и события в первую очередь как благоприятные (для существования и развития организма), неблагоприятные и нейтральные, а также определить важность отдельных событий. Таким образом, тело становится источником смыслов, присущих знакам.

1. Телесный опыт как предмет семиотики

1.1. Телесность сквозь призму феноменологии

Идея о том, что познание зависит от того, как человек взаимодействует со средой и другими людьми, подробно исследовалась в рамках феноменологического направления в философии в начале и середине XX в. Так, Э. Гуссерль пишет о том, что мы воспринимаем мир через наше материально-физическое тело; опыт столкновения с другими материальными вещами позволяет нам переживать, наряду с физическими ощущениями (осознания, движения, а также тепло, холод или боль), и чувства [Гуссерль 1913/1999].

М. Мерло-Понти [1945/1999], развивая идеи Гуссерля, пишет: «Все, что я знаю о мире, пусть даже и через науку, я знаю исходя из моего видения или того жизненного опыта, без которого символы науки были бы пустым местом»; физическое существование собеседников, наличие у них тела и внешности, согласно Мерло-Понти, — условие для диалога, а язык вырастает из телесности. Поскольку человек рождается в определенной культурной среде, его взаимодействие с миром оказывается одновременно природно и культурно обусловленными. Объекты воспринимаются с учетом того, какие возможные действия они предполагают. Мерло-Понти, помимо этого, говорит про язык как «телесный жест», считая речь такой же естественной, как, например, непроизвольное движение руки. Восприятие других живых существ происходит через призму собственного опыта: они также представляются в первую очередь воспринимающими телами и лишь затем — как сознающие существа, однако возможность сознания и определяется самим фактом восприятия.

М. де Серто, продолжая феноменологическую традицию, выводит понимание воплощенной природы познания на следующий уровень, обозначая тело как источник текста, смысла или дискурса, т. е. говорит о семиотизации тела, «превращении его в текст»

и «конвертации пространственной организации тела в семантическую организацию словаря (и наоборот)» [де Серто 1975/2001].

1.2. «Воплощенный разворот» в семиотике

Воплощенный подход в семиотике [Violi 2003] стал набирать популярность с конца 1980-х годов под влиянием направления в философии и психологии, которое получило название «воплощенное познание» (*embodied cognition*, см., например, [Shapiro 2007; Adams, Aizawa 2019; Varela et al. 1991; Varela и др. 1991/2023]).

Развитие когнитивной лингвистики в последние десятилетия опиралось на открытия в психологии, антропологии, нейрофизиологии и т. д., что привело к пониманию обусловленности когнитивных, семантических и языковых структур нашим опытом физического взаимодействия с миром; такие структуры также называют «воплощенными» [Stjernfelt 2006]. Этот подход представлял альтернативу формалистской традиции, которая изучала мышление и язык без учета тела как их инструмента и способа реализации. Известная метафора, сопоставляющая мозг и компьютер, предполагала, что свойства вычислительной машины не влияют на работу «программного обеспечения», в то время парадигма воплощенного сознания исходит из того, что устройство нашего тела и мозга определяет наше мышление и язык.

Дж. Лакофф сформулировал тезис об отражении телесного опыта в языке следующим образом: «Свойства определенных категорий являются следствием биологических возможностей человека и его опыта существования в некотором физическом и социальном окружении» [Lakoff 1987: 12] (пер. мой. — Ю. Н.). Среди основополагающих лингвистических работ в этом направлении можно также назвать статью М. Джонсона [Johnson 1987], который на примере предлогов показал, что семантическое содержание многих слов восходит к физическим действиям человека в мире.

Многие исследователи обращаются к миру животных, сравнивая их коммуникативные системы с человеческой. Организмы (включая насекомых, птиц и других животных), взаимодействуя с миром, уже имеют «встроенный» первичный семиотический механизм, не требующий обязательно каких-либо когнитивных процессов; этот механизм позволяет интерпретировать мир через непосредственное восприятие [Clark 1997]. Организм на этом этапе категоризирует ощущения как хорошие, плохие и не имеющие значения для выживания. Этот процесс определяется сенсорными / перцептивными возможностями организма (а далее — опытом и способами взаимодействия со средой и другими организмами, которые позже образуют устойчивые сочетания и системы [Deely 1990]). Таким образом, сознание, тело и мир неразделимы и вместе участвуют в процессе осмысленного существования.

В результате можно сказать, что сознание оперирует образами, определяемыми существованием и действиями тела — об этом писали также Лакофф и Джонсон [Lakoff, Johnson 1980]. Таким образом, сенсомоторные возможности человеческого тела служат основой не только для восприятия, но и для концептуализации, а мышление — это непрерывный процесс, когда мы сравниваем наш опыт в данный момент с предшествующим опытом в разных сферах через посредство метафор. Благодаря телесному опыту мы непосредственно воспринимаем такие понятия, как *вверху / внизу, внутри / снаружи, близко / далеко*, что создает основу для семиотических примитивов и мышления.

Дж. Дили развивает эту мысль дальше в [Deely 2013]: хотя первоначально наша семиотическая система основана на биологических предпосылках, мы выражаем наши идеи в форме знаков (язык, живопись и пр.) и, опираясь на них, строим наше восприятие и деятельность в окружающем мире. В результате телесный опыт переносится в другие области, позволяя осмыслять абстрактные понятия, например, радость как движение вверх или печаль как движение вниз.

Метафоризацию также можно рассматривать не как перенос сенсомоторного опыта в область абстрактных идей, а как смешение концептов (*conceptual blend*) [Fauconnier, Turner 1998], показывающее сходство двух областей, когда общие черты позволяют создать новое пространство (что может служить, например, основой для загадок и анекдотов или для создания вымышленных животных, таких как крылатый конь Пегас). Используя аналогии из материального мира, мы можем объединить «корабль» и «нож», поскольку они имеют схожие свойства — движение по некоторой поверхности, по прямой, оставляющее след, в результате можно сказать, что *корабль, как нож, рассекает воду*. В основе этих рассуждений также лежит идея о том, что воплощенные схемы служат основой для высших когнитивных функций, определяющих, в свою очередь, бесконечные семиотические возможности [O'Neil 2008].

Другой стороной воплощенного подхода в современной когнитивистике, наравне с признанием сенсорного опыта, является понимание социального взаимодействия как источника значений [Sinha, Jensen de López 2000]. Развивая приведенный выше в этом разделе тезис Дж. Лакоффа, К. Синха и К. Йенсен де Лопес формулируют свою позицию следующим образом: «Человеческое тело (и нервная система) взаимодействует с физическим (и социальным) окружением как универсальным источником образных схем (а также схем для отображения событий, движения и столкновения разнонаправленных импульсов)», и далее: «Метафорическое расширение является инструментом для того, чтобы воплощенный опыт помог структурировать многие другие не-физические (а, например, психологические и социально-межличностные) сферы» (пер. мой. — Ю. Н.). В качестве примера реализации таких схем в языке можно привести следующие выражения: «*преступить закон*», «*их отношения стали ближе*» или «*уклоняться от исполнения своих обязанностей*». Влияние социального окружения проявляется в том, что языковые практики по использованию воплощенных метафор, принятые в социуме, в котором растет ребенок, также служат основанием для его концептуальных схем, наравне с его собственным эмпирическим опытом. В таком случае культурные репрезентации будут проявляться двояко: и как социальные конструкты, влияющие на восприятие и категоризацию окружающего мира, и как когнитивные паттерны. Материальная культура также является выражением или воплощением культурного знания (в качестве примера можно привести осмысление маленькими детьми уже с возраста девяти месяцев чашки как контейнера, у которого есть прототипическая ориентация, даже если экспериментатор протягивал им чашку, перевернутую вверх дном). Соответственно, «язык является основным средством, благодаря которому тело входит в социокультурную сферу как объект семиозиса, и это происходит с помощью двух инструментов: через культурные дискурсы о теле и через языковые практики контроля и управления телом» [Bucholtz, Hall 2016] (пер. мой. — Ю. Н.).

Д. Лидов [Lidov 2018] подчеркивает, что семиозис строится на стирании лишнего, ситуативного, индивидуализирующих признаков (см. также [Terrado-Pablo 2017]); в результате физические источники знака забываются, что Лидов связывает с нашей способностью отделять означаемое от означающего. Приведенный выше эксперимент с чашкой как контейнером, у которого есть «правильное» положение, служит примером абстрагирования — создания схем, не привязанных к конкретному перцептивному контексту, расположению данной чашки в данный момент.

Воплощенная семиотика смыкается с когнитивной семантикой [Langacker 1999; Talmy 2011; Geeraerts 2021], когнитивной семиотикой [Zlatev 2012; Paolucci 2021] и социокогнитивной лингвистикой (см. обзор в [Ирисханова 2014]). Все эти области объединяет понимание того, что сенсомоторный опыт играет важную роль в формировании языковых репрезентаций и социальных практик. О. К. Ирисханова [2014] отмечает полимодальность как важное свойство моделей, создаваемых в рамках социокогнитивного подхода в лингвистике, подразумевая под этим «способность реализовываться через разные семиотические системы и каналы восприятия и обработки информации».

Важность подхода, учитывающего присутствие собеседников в едином физическом пространстве в момент коммуникации и возможность невербального взаимодействия между ними, можно показать на следующих примерах. Жесты и даже совсем небольшие по амплитуде движения — такие как изменение направления взгляда или движение бровей — могут существенно повлиять на ход коммуникации, при этом все участники одновременно являются активными действующими лицами, в визуально-кинестических каналах они не разделяются на активных (говорящих) и пассивно воспринимающих (слушающих); реакция адресата может определить ход беседы и повлиять на речь говорящего, даже если она выражена невербально. Помимо этого, невербальная сторона коммуникации не бывает нейтральной, незначимой; появление или отсутствие определенных знаков — таких как, к примеру, кивки слушающего или направление его взгляда — будет так или иначе проинтерпретировано его собеседниками и наблюдателями.

Рассмотрение телесного взаимодействия и воплощенной семиотики в когнитивной лингвистике, как предполагается, не только объяснит некоторые феномены, недоступные пониманию в рамках лингвоцентрического подхода, но изменит само понимание семиотических систем, возможно, выделяя тело и телесный опыт как основной и определяющий элемент [McDonald 2013].

Семиозис как процесс формирования знака всегда встроен в чувственные материальные процессы и по этой причине имеет, в терминах Пирса, динамическую (индексальную) сторону, которая позволяет протекать коммуникативному процессу [Пирс 2000]. Эту сторону семиотического процесса Хоффмейер [Hoffmeyer 2008] называет семиотической причинностью, поскольку она ограничивает возможные интерпретации знака теми, которые уместны в локальном контексте. Неизбежная укорененность в контексте знака, который используют участники коммуникации в какой-то определенный момент, обуславливает существование биосемиотики — науки о знаковых процессах в живых системах. Так, кожа позволяет нам ощущать внешний мир, но факт того, что мир можно ощутить, требует не только наличия кожных рецепторов, для него требуется еще биологический смысл — функциональные интерпретации множества сенсорных стимуляций для того, чтобы тело могло интерпретировать эти сигналы и реагировать соответствующим образом: импульсы от одного конкретного нервного окончания сравниваются с импульсами от близлежащих точек кожи, а также с импульсами от мышц и суставов. Осмысленная интеграция этих чувственных данных образует стереогностическое чувство, которое затем будет объединено со всеми остальными органами чувств, чтобы создать оптимальную картину происходящего.

С другой стороны, предмет семиотики начинает разрастаться и даже иногда включать явления, изначально считавшиеся чисто биологическими, такие как обменные процессы в клетках живого организма, см., например, [Hoffmeyer 2008; Copley 2016; Barbieri 2008; Kull 1999]. Изменение объекта семиотики выглядит как попытка отойти от структурализма, когда знаками считались только единицы речи, а в первую очередь слова, к коммуникации реальных (а не абстрактных) людей в естественных обстоятельствах. Однако это похоже на отклонение маятника в противоположную сторону, когда семиотическая интерпретация накладывается на химические и биологические процессы (и даже включает их противоположность в биологическом смысле, а именно объединения организмов, которые иногда называют суперорганизмами, такие как пчелиный улей или муравейник см., например, [Turner 2016; Favareau 2009]). Возможность такой интерпретации объясняют фактом передачи сигнала, однако тут можно возразить, что в языковой коммуникации присутствуют отправитель сообщения и адресат, в прототипическом случае их общение является намеренным и осознанным.

Другой взгляд, также влияющий на понимание предмета семиотики, подразумевает такую интерпретацию когнитивных процессов, которая включает эмоции, при этом предлагается следующая аргументация: природа построила кортикальные механизмы, отвечающие за рациональное поведение, не просто поверх субкортикальных структур, связанных

с биологической регуляцией, но как следствие последней и с ее участием; а эмоции являются мостом между рациональными и биологически обусловленными реакциями, т. е. между кортикальными и субкортикальными структурами [Damasio 1994: 128]. У людей мы можем в результате наблюдать формирование вторичных эмоций, которые возникают, «как только мы начинаем испытывать чувства и формировать систематические связи между категориями объектов и ситуаций, с одной стороны, и первичными эмоциями, с другой» [Ibid.: 134] (пер. мой. — Ю. Н.). Формирование вторичных эмоций не только зависит уже от лимбической системы, но и задействует префронтальную и соматосенсорную кору головного мозга.

В результате таких изменений воображаемая граница между воспринимаемой реальностью и ее концептуализацией становится прозрачнее, поскольку абстрактные представления определяются теми структурами, которые отвечают за деятельность и восприятие, а телесный опыт становится отправной точкой для абстрактных мыслительных процессов: они опираются на метафорический перенос понятий, присущих сенсомоторной сфере, в не связанные с ней области. В качестве примера можно привести метафору «знать — значит, видеть», которая присутствует в таких словах, как «просвещение» или выражении «давайте взглянем на этот вопрос с другой стороны» [Stjernfelt 2006].

В связи с новым теоретическим подходом к изучению знаков, когда встает вопрос других форм сознания, и, соответственно, других типов семиотик (например, предлагаются семиотики геномного кода [Мечковская 2004] или вопрос о существовании знаковых систем для разных типов живых организмов, действия которых могут быть описаны в терминах когнитивных схем, таких как движение к цели), предлагаемая на данный момент граница связана с идеей ментальных репрезентаций, которые должны быть промежуточным звеном между действием или ощущением, с одной стороны, и знаком — с другой [Stjernfelt 2006].

2. Лингвистический подход к описанию телесных знаков

2.1. Знаковое vs. незнаковое поведение

В семиотике давно ведется дискуссия о разграничении знаков и не-знаков. В частности, можно упомянуть У. Эко, который очерчивал границы семиотики, исключая те объекты и явления, которые не являются предметом социальных конвенций, т. е. оставляя семиотике социокультурную сферу; за ее пределами оказывались такие объекты, которые могут быть использованы как инструменты, ср. [Носачев 2008]. Интересно отметить, что граница между знаковостью и не-знаковостью проницаема: если некий объект, до этого выступавший как инструмент, начинает рассматриваться как знак, он переходит в область интересов семиотики, и обратное движение также возможно [Ipsen 2006]. Таким образом, границу между семиотической и несемиотической сферой можно обозначить как противопоставление явлений, которые воспринимаются в данной культуре как знаки, и тех, которые (еще) не вошли в семиосферу данного сообщества (если пользоваться терминами Ю. М. Лотмана [2000]) или вербальный репертуар некоторого пространственного или социального объединения (в понимании Гамперца [Gumperz 1964]). Семиосферу как систему характеризуют такие свойства, как сохранение и передача информации во времени и набор правил, по которым организована данная семиосфера. Таким образом, в поисках знаковых явлений среди невербальных сигналов можно ожидать от них сохранность на протяжении по крайней мере какого-то времени и системную организацию таких знаков.

Буш [Busch 2021] рассматривает вербальный репертуар Гамперца через призму телесного переживания языка субъектами — с одной стороны, в процессе взаимодействия (и в устоявшихся языковых формулах или текстах), а с другой — на уровне воплощенного

опыта участников коммуникации, который также влияет на обусловленные культурой знаки, используемые при коммуникации. Буш связывает эти три понятия через образ тела, понимаемый как воображаемое (чаще всего неосознанное) представление воплощенного Я во взаимодействии с другими людьми и миром. В подтверждение своей позиции она приводит наблюдения Дуранти [Duranti 1992], который отмечает, описывая самоанские ритуалы приветствий, что тело (например, положение тела, жесты, взгляд) обеспечивает контекст для интерпретации лингвистических единиц (слов, морфем и т. д.), а также помогает формировать новые, иногда дополняющие, иногда противоречивые сообщения. И даже более того, тело человека не просто семиотический инструмент, помогающий кодировать и декодировать языковые сообщения, но само по себе мощное средство передачи информации, особенно когда коммуникантам доступна богатая культурная традиция, связанная с невербальным взаимодействием. В качестве еще одного подтверждения важности невербальной составляющей общения можно привести работы Ч. Гудвина [Goodwin 2000; 2003]. Он подробно разбирает примеры из обычных диалогов и показывает, как участники коммуникации учитывают физическое расположение друг друга, предыдущий контекст и среду, в которой происходит общение. Их положение, действия и совместное внимание имеют решающее значение для того, как адресат понимает происходящее и согласует свои реплики и действия; таким образом Гудвин предлагает включать в лингвистический анализ не только речь и жесты, но и элементы внешнего контекста, с которыми они связаны.

А. Кендон [Kendon 2004: 8] предлагает свое определение для знаковых явлений, говоря о жестах следующим образом: это заметное движение, реализуемое как явное коммуникативное действие, обычно адресованное другому человеку и воспринимаемое адресатом (и жестикулирующим) как следствие явного и осознанного намерения, и предполагающее передачу некоторого значения, помимо собственно движения (пер. мой. — Ю. Н.). Жесты, сопровождающие речь, можно разделить на два основных типа: иллюстраторы (связанные с содержанием слов и передаваемым сообщением) и регуляторы (отвечающие за управление коммуникацией, т. е. установление, поддержание и завершение контакта; среди самых заметных функций регуляторов — управление сменой ролей в диалоге). Пограничными с жестикуляцией классами с точки зрения отнесения их к знакам являются выражения эмоций и адаптеры. Первые, если появляются одновременно со словами, могут выражать отношения к сообщаемым фактам («*Моя жена выиграла главный приз в лотерею*») или указывать на модальные значения («*Здесь, должно быть, какая-то ошибка*»). Вторые определяются как движения, направленные на восстановление и поддержание физического и эмоционального комфорта и могут встречаться без речи. Если адаптеры появляются одновременно со словами, они с большой вероятностью выражают оценочные или эмоциональные значения [Hübler 2006]. Из уже приведенных примеров видно, что знаковые движения в диалоге не ограничиваются жестами рук, хотя движения рук наиболее заметны среди всех невербальных средств как средство передачи информации, дополняющей или иллюстрирующей речь. Однако в коммуникации выражение эмоций (отношение говорящего к предмету речи, собеседникам и т. п.) и регуляторы не менее важны, хотя гораздо меньше изучены. Если попытаться учесть все существенные для вербальной коммуникации значения этих типов, задача разграничения знаковых и незнаковых движений и поз становится еще сложнее.

Дж. Мартин и М. Заппавинья в [Martin, Zappavigna 2019] указывают на два признака при разграничении знакового и соматического (в их терминах — незнакового) поведения: 1) у знака есть значение (не уточняя, однако, что имеется в виду; обычно в литературе это понимается как возможность интерпретации адресатом или наблюдателем определенного поведения как знакового), 2) этот знак повторяется в разных контекстах.

Решая задачу разграничения знаковых и незнаковых явлений (или, в терминах автора, знаковых движений от симптоматических [Крейдлин 2002]), Г. Е. Крейдлин и его соавторы опираются на языковые номинации, особенно такие, которые образуют устойчивые языковые метафоры (ср. [Lakoff, Johnson 1980]), и на семиотический анализ в целом.

Такая стратегия представляется самой надежной, поскольку позволяет выбрать именно знаки в лингвистическом понимании; ее ограничения, однако, состоят в том, что не нашедшие языкового обозначения телесные знаки останутся для дальнейших исследований. Тем не менее можно сказать, что изучение жестов через их языковые номинации выглядит вполне оправданным, поскольку у языковых обозначений движений и частей тела часто есть небуквальные (метафоричные) значения и коннотации, так что определенно значение (и сам знак) существуют в языке независимо от контекста.

Следует, однако, учитывать, что природа кинетических (т. е. жестовых) знаков может отличаться от того, что мы видим как их языковые обозначения (это разграничение, по-видимому, не вполне прослеживается в тексте). Можно сравнить эту ситуацию с тем, как начинались исследования жестикуляции в современной лингвистике. Первыми были изучены и описаны эмблематические жесты, такие как «знак ОК» или «покрутить пальцем у виска» [Григорьева и др. 2001], поскольку они наиболее близки к языковым единицам: у таких жестов есть устойчивая узнаваемая форма, есть неизменное значение, правила употребления знака и часто устойчивые названия или стандартное (конвенциональное) речевое сопровождение. Однако эмблематические жесты встречаются на несколько порядков реже, чем спонтанные движения рук, которые сопровождают речь (сейчас чаще используются термины «речевые жесты (co-speech gestures)» или «жестикуляция», на том этапе лингвистических исследований невербальной коммуникации они назывались иллюстративными). Очевидно, что знаковые движения рук в коммуникации как система не могут быть изучены, если обращать внимание только на эмблемы.

Тут следует оговориться, что сравнивать знаки разных типов по частоте, с которой они встречаются в некотором массиве данных, — это корпусный подход. Авторы книги «Язык и семиотика тела», скорее, составляют словарь (который может в дальнейшем использоваться для создания, разметки и описания корпусов). Такой подход, который можно было бы назвать лексикографическим, имеет свои неоспоримые преимущества.

В книге обозначены следующие цели: семиотическая концептуализация тела и исследование фразеологизмов и устойчивых сочетаний с обозначениями частей тела через признаковый подход. Объектами изучения в таком случае оказываются сами части тела, их языковые описания, жесты и комплексные жестово-речевые единицы (с. 42). В качестве метода используется анализ «вербальных и невербальных аспектов коммуникации, которые определяются контекстными факторами», а результатом являются «правила употребления отдельных слов и жестов» и «условия их совместной встречаемости» (с. 41).

В центре исследований Г. Е. Крейдлина, по его собственному определению, находится знак, означающим которого является движение тела (или даже часть тела как источник коннотаций, в первую очередь через устойчивые языковые выражения). Ему особенно интересны случаи, когда этот знак теряет свою изначальную иконичность и обретает символический смысл, во многом определяемый, однако, функциями этой части тела и движения.

Г. Е. Крейдлин и его соавторы собрали интересующие их обозначения движений и частей тела, которые они называют соматизмами (чаще всего отталкиваясь от их языковой номинации) в виде энциклопедии, насколько возможно полной, с систематическим описанием на созданном для этой цели метаязыке. В число соматизмов, упоминаемых в книге «Язык и семиотика тела», входят также воображаемые органы, такие как душа (см. [Руссо 2013]). В разных языках могут различаться свойства, приписываемые реальным и воображаемым органам, соответственно, их символизация тоже будет различаться (так, метафорическим вместилищем *гнева* или *храбрости* может быть сердце или печень, см., например, [Рубец 2009; Usmonova 2020]).

Можно считать, что описание значений частей тела и движений через призму языка и языковых сообщений — надежный способ остаться в рамках лингвистики, но расширить ее предмет за счет включения жестов и других движений, также участвующих в коммуникации лицом к лицу, наравне с вербальными; в частности, в этом аспекте упоминаются сообщения, в которых объединены жестовый и вербальный каналы.

Что характерно, в качестве примеров жестово-вербальных единиц в книге приводятся случаи, когда жестовые знаки синтаксически встроены в языковое выражение, как в примере «Он поймал вот такую рыбу!». Это самые заметные случаи сочетания жестов и речи, однако не самые частые; далеко не всегда синтаксическая структура высказывания и вербальные единицы так же эксплицитно, как в примере про рыбу, указывают на жест как на часть единого мультимодального сообщения. Другой пример, когда жест бывает обязательно встроен в вербальное общение, — этикетные жесты. Здесь опять-таки можно вспомнить историю изучения невербального общения в лингвистике. Первые мультимодальные (полимодальные) исследования также обращались к эмблематическим жестам и этикетным жестовым формулам [Creider 1977; Sparhawk 1976] и к жестам, замещающим какой-то элемент в синтаксической структуре [Poyatos 1977; Капанадзе, Красильникова 1973]. Можно предположить, что в данном случае переход к новым явлениям (от движений рук к любым движениям тела и даже значениям его частей) закономерно вызвал возвращение к прозрачным примерам с опорой на лексические и синтаксические маркеры коммуникативной природы жеста.

2.2. Семиотический подход к изучению жестов

Поскольку жестикуляция и в целом все движения тела, сопровождающие речь, традиционно считаются спонтанными (в подавляющем большинстве случаев), не образующими устойчивой связи между означаемым и означающим, то возможность правильной интерпретации таких движений определяется, во-первых, их синхронизацией с речью — символическим кодом, а во-вторых, иконичностью самого знака.

Г. Е. Крейдлин и его соавторы рассматривают части тела и движения как элементы знаковой системы, отраженной в языке и встроеной в язык. Таким образом, авторы исследуют условия употребления и сочетания лексических единиц для описания «соматических объектов». Среди задач книги упоминается выявление и описание (переносных) значений устойчивых выражений и сочетаний с наименованиями частей тела (*ноги не ходят, голова раскалывается*) и закономерностей невербального поведения. При этом самая масштабная цель авторов книги обозначена следующим образом: «рано или поздно (...) описать все семантические и культурно выделенные значения у каждого телесного признака» (с. 98). Под культурной выделенностью понимается «важность этого значения для осмысления культурного контекста, в котором употребляется языковое выражение этого значения», или, другими словами, представление о том, что определенный признак характеризует его обладателя и при этом «передает стереотипные представления о нем и его теле, закрепленные в данной культуре» (с. 95). То есть цель этой работы в максимальной перспективе — перечислить все признаки, их значения, контексты использования и интерпретации, в т. ч. метафорические употребления.

Очевидно, что тут можно найти как культурно-специфические, так и универсальные признаки, и это тоже учитывается авторами и исследуется в соответствующих разделах. Таким образом, объект исследования в данной работе — не единицы в коммуникации (ср. мультимодальные кластеры [Гришина 2017], полимодальный дискурс [Ирисханова (ред.) 2022]), а единицы языка как системы, если его не сводить к вербальным знакам.

В рамках такого подхода авторы выстраивают многоуровневую иерархию «часть тела — часть части тела — часть части части тела...» (например, руки — пальцы — ногти), уточняя, что в разных языках можно насчитать до шести таких уровней. Эта иерархия для рассматриваемой семиотической системы проявляется, в частности, в том, что по мере продвижения по ней уменьшается разнообразие функций, приписываемых этой части тела (если судить по устойчивым языковым выражениям) и нарастает определенность аномалий или дисфункций, присущих этой части.

Помимо этого, подход, использованный в книге, потенциально позволяет изучить, как признаки соматического объекта связаны с его семиотизацией. Авторы берут обе стороны континуума, на одном конце которого физическое проявление (например, размер части тела) и его языковое / жестовое осмысление, а на другом — абстрактные и метафорические значения («показать нос» в значении ‘дразнить’; использование уменьшительных суффиксов при описании внешности женщин и детей), и добавляют к этому разработанную ими систему описания соматических объектов, содержащую список всех потенциально релевантных параметров (например, геометрические признаки vs. сравнение с растениями и животными при описании внешности: «круглое лицо» — «нос картошкой» — «лебединая шея»).

2.3. Особенности книги «Язык и семиотика тела»

Во введении отмечается, что монография посвящена языку тела и языку жестов, таким образом «язык» понимается максимально расширительно (включая манеры, т. е. особенности вербального и невербального поведения). Однако в повседневном общении визуально-кинестическая составляющая (визуальная для адресата, кинестическая — для автора сообщения) почти всегда подчинена вербальной и не может служить самостоятельным средством коммуникации (очевидно, что жестовые языки представляют собой особую систему, которая принципиально отличается по своим возможностям и свойствам от жестикуляции, сопровождающей речь, даже если рассматривать ее в широком смысле, как движения и положение всех частей тела и тела целиком). В монографии также упоминаются как сопоставимые явления акциональный («язык знаковых действий») и ольфакторный («язык запахов») язык, что подводит к мысли о необходимости разграничения терминов и об обоснованности слова «язык» в этих контекстах, по крайней мере в лингвистических работах.

Помимо этого, хотя предметом исследования заявлен язык тела, скорее можно сказать, что в первом томе говорится о языковых описаниях частей тела типичных коммуникативных движений.

Попытка создания модели, которая в перспективе объясняла бы, почему можно сказать «руки-крюки», но нельзя *«ладони-крюки», и обосновывала бы другие подобные языковые феномены за счет создания правильной классификации (как это следует из логики первого тома) — сильная сторона книги и в целом очень масштабная и амбициозная цель. В дополнение к ней поиск способов создать «настоящую классификацию» с «непересекающимися классами» для «соматических объектов» (с. 89, сноска) выглядит как научный детектив.

Огромный объем рассматриваемого языкового материала (упоминается, что в его сборе помогали школьники — слушатели курсов Г. Е. Крейдлина в гимназии № 1514) и попытки учесть как можно больше существующих теорий и подходов в этой области впечатляют. С другой стороны, выход за рамки хорошо изученной тематики может быть причиной не всегда убедительных выводов, см. пример из подглавы 3.6. «Покровы» в первой главе: «Об особой культурной символизации ногтей говорит и фразеологический оборот «узнать (всю) подноготную», означающий ‘проникнуть силой в тайну, тщательно хранимую человеком, как бы загоняя ему для этого нечто острое под ногти’» (с. 79). Как представляется, это не самая вероятная интерпретация, ср. обсуждение этого фразеологизма в [Володина 2018]. Также вызывает вопросы отнесение хрящей к костям (с. 73), а сухожилий — к «сосудам» (вместе с кровеносными сосудами, с. 205).

Книга представляет собой результат масштабного и очень тщательного отбора и анализа всего, что касается отражения в языке и культуре частей тела и жестов, в этом отношении она развивает предыдущий знаковый труд Г. Е. Крейдлина — книгу «Невербальная семиотика» [2002]. Это в определенной степени новаторское исследование в малоосвоенной сфере, поэтому закономерна некоторая эклектичность и мозаичность в списке подразделов

(например, в том, что касается списка языков, для которых рассматриваются некоторые особенности невербального поведения, или стилей коммуникации). В этом случае, на мой субъективный взгляд, участие разных авторов (не указано, кто из них был автором какого раздела) является достоинством этой работы, поскольку цель — охват новой области настолько полный, чтобы сделать очевидной и простой последующую систематизацию этих данных и встраивание в уже сложившиеся сферы (во введении отмечается, что книга как междисциплинарное исследование может быть интересна «культурологам, психологам, литературоведам, философам, педагогам и искусствоведам»).

Если считать эту работу скорее лексикологической, чем мультимодально-лингвистической, представляются более обоснованными такие решения, как введение примеров исключительно через описание значения единицы или иногда с кратким вербальным описанием движения, так что фактически особенности реализации конкретного типа движений остаются за кадром, как и отличие встречаемых в реальности от тех, которые можно восстановить из литературных примеров. Следует, однако, учитывать, что не все соматические знаковые единицы находят отражение в языке. Диссонанс или противоречие между разными семиотическими системами и подсистемами также может служить частью общего сообщения, и в этом случае воплощенные знаки определяют значение языковых, а не наоборот [Bucholtz, Hall 2016]; структуру и закономерности использования таких сложных единиц еще предстоит изучить.

Семиотический подход позволяет увидеть важные взаимосвязи и предложить гипотезы для изучения другими методами (в частности, корпусными). Так, отмечается, что «в жестах из класса касаний ведущая роль отведена рукам и некоторым частям рук, в жестах из класса “выражение согласия / несогласия” — голове, а в молчаливых вопросах и ответах на них — глазам» (с. 39). Также заслуживает упоминания такое интересное наблюдение, как оценка степени «освоенности» соматического объекта через разнообразие его наименований. В дальнейшем эти наблюдения можно будет проверить количественными методами, например, на корпусном материале.

Очевидно, что не все предположения получают подтверждение. В книге транслируется часто встречающееся мнение, что языковые единицы количественно преобладают в диалоге над невербальными и что жестовые знаки более «наглядные», т. е. используются как дополнительные средства, если надо проиллюстрировать информацию, передаваемую словами. Как представляется, это связано с тем, что лингвисты традиционно смотрят на невербальные каналы в коммуникации через призму языка. Для сравнения можно вспомнить просодические средства, гораздо лучше изученные, чем жесты, для которых известно, что они в большой степени абстрактны и многозначны [Янко 2008], т. е. абстрактность значений противоречит тезису о «наглядности». Что касается доли кинетических (жестовых) единиц в коммуникации по сравнению с вербальными, этот вопрос еще предстоит исследовать; однако данные корпуса RUPEX [Кибрик 2018а; 2018б] позволяют предположить, что даже с учетом только движений рук и головы объем визуальной информации, передаваемой в ситуации общения лицом к лицу, по крайней мере сопоставим с вербальной, особенно если учесть прагматические и интерактивные значения жестов.

Заключение

Семиотические работы, следующие концепции воплощенного познания, помогают увидеть «многочисленные слои значений, определяемых нашей телесной природой, на которые опираются процессы мышления и коммуникации» [Pelkey 2023] (пер. мой. — Ю. Н.). Лингвистические исследования помогают проверить философские концепции в этой области и продемонстрировать, как в языке отражается связь тела и мышления, соединенных в непосредственно проживаемом опыте. Работы современных когнитивных лингвистов

показали, что такие абстрактные понятия, как «время», «познание» и «идеи» так же опираются на опыт взаимодействия с миром, как «бегать» или «видеть». Помимо укорененности в телесных ощущениях, воплощенный подход в современной лингвистике учитывает, что человек не просто воспринимает (и осмысляет) свои ощущения: мы как человеческие существа не только **обладаем** языком, опирающимся на наш телесный опыт, мы сами **являемся** такими воплощенными агентами в коммуникации, так что язык может быть не просто «способностью», а еще одним способом реализовывать наше бытие в мире [Di Paolo et al. 2018]. И даже более того, тело можно назвать необходимым условием существования языка [Bucholtz, Hall 2016]. Тот факт, что единицы и структуры языка опираются на перцептивную и моторную информацию, непосредственно подводит нас к пониманию того, как движения и жесты становятся важным элементом в коммуникации.

Благодаря когнитивно-лингвистическим исследованиям становятся понятными причины распространенности языковых метафор, опирающихся на телесные ощущения (такие как «счастье — вверх» или «плохое дурно пахнет»), а также как именно наш кинестетический (сенсорный) опыт определяет структуру абстрактных областей и взаимосвязь концептов. Телесные проекции можно наблюдать в словосочетаниях типа «*ножка кровати*», «*устье и рукава реки*», «*подножье горы*» или даже «*лобовое стекло*» автомобиля, когда отношения «часть-целое» переносятся с человеческого тела на другие, часто гораздо более крупные сущности. Воплощенные метафоры могут существенно различаться в разных языках, например, «высокий» или «низкий» музыкальные тоны в языке фарси будут названы как «толстый» и «тонкий», в кпелле — «легкий» и «тяжелый», а суи — «молодой» и «старый» соответственно; несмотря на различия, эмпирическая основа таких обозначений обеспечивает возможность перевода и взаимную понятность несовпадающих концептуальных систем, которые в итоге включают не только противопоставления, но также иконические схемы и разного рода маркированность понятий [Pelkey 2023].

Инструменты воплощенной семиотики позволяют продемонстрировать, как телесный опыт проявляется и на уровне абстрактного мышления (через пространственные и телесные метафоры для абстрактных понятий), и при каждом взаимодействии людей (например, помогая выразить их отношение к предмету речи и точку зрения).

Хотя сам Г. Е. Крейдлин говорит о своей сфере как о невербальной семиотике либо мультимодальной лингвистике, можно сказать, что выход книги «Язык и семиотика тела» отмечает появление в русскоязычном лингвистическом пространстве такой области, как воплощенная (или «отелесненная»), как предлагается в новом переводе [Варела и др. 1991/2023]) семиотика. Авторы, опираясь на концептуальный аппарат, разработанный в этой сфере для звучащего языка, показывают, как в русском языке развивались концептуальные схемы, источником которых является непосредственный телесный опыт.

В книге, вышедшей к юбилею Г. Е. Крейдина [Переверзева (ред.) 2023], это направление получило дальнейшее осмысление и развитие. Во-первых, идеи, описанные в книге «Язык и семиотика тела», находят свое отражение в таких сферах, как литературоведение, типология и описание языков, лексикология. Во-вторых, новое понимание невербальной составляющей языка и коммуникации дает возможность точнее описать и осознать разные стороны речевого взаимодействия. Наконец, намечаются новые направления в рамках лингвистики, связанные с идеями Г. Е. Крейдина.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Варела и др. 1991/2023 — Варела Ф., Томпсон Э., Рош Э. *Отелесненный ум. Когнитивная наука и человеческий опыт*. Тулупова К. (пер.), Лысенко В. (отв. ред.). М.: Фонд «Сохраним Тибет!», 2023 (1-я публ. 1991 г.). [Varela F., Thompson E., Rosch E. *Otelesnennyi um. Kognitivnaya nauka i chelovecheskii opyt* [The embodied mind. Cognitive science and human experience]. Tulupova K. (trans.), Lysenko V. (ed.). Moscow: “Save Tibet!” Fund, 2023 (first publ. in 1991).]

- Володина 2018 — Володина Т. В. Фразеология и магические практики в контексте антропо- и космогонических легенд (белорусская традиция на европейском фоне). *Миссия выполнима. Перспективы изучения фольклора*. Кийва М., Володина Т. (ред.-сост.). Тарту: ЭЛМ, 2018, 181–206. [Volodina T. V. Phraseology and magical practices in the context of anthropo- and cosmogonic legends (Belarusian tradition against the European backdrop). *Missiya vypolnima. Perspektivy izucheniya fol'klora*. Kyjva M., Volodina T. (eds.). Tartu: ELM, 2018, 181–206.]
- Григорьева и др. 2001 — Григорьева С. А., Григорьев Н. В., Крейдлин Г. Е. *Словарь языка русских жестов*. М.: Языки русской культуры, 2001. [Grigor'eva S. A., Grigor'ev N. V., Kreidlin G. E. *Slovar' yazyka russkikh zhestov* [A dictionary of Russian gestures]. Moscow: Yazyki russkoi kul'tury, 2001.]
- Гришина 2017 — Гришина Е. А. *Русская жестикulyация с лингвистической точки зрения. Корпусные исследования*. М.: Языки славянской культуры, 2017. [Grishina E. A. *Russkaya zhestikulyaciya s lingvisticheskoi tochki zreniya. Korpusnye issledovaniya* [Russian gesticulation from a linguistic point of view. Corpus studies]. Moscow: Yazyki slavyanskoi kul'tury, 2017.]
- Гуссерль 1913/1999 — Гуссерль Э. *Идеи к чистой феноменологии и феноменологической философии*. Т. I. Михайлов А. В. (пер.). М.: ДИК, 1999 (1-я публ. 1913 г.). [Husserl E. *Idei k chistoi fenomenologii i fenomenologicheskoi filosofii* [Ideas pertaining to a pure phenomenology and to a phenomenological philosophy]. Vol. 1. Mikhailov A. V. (trans.). Moscow: DIK, 1999 (first publ. in 1913).]
- Ирисханова 2014 — Ирисханова О. К. Социокогнитивная лингвистика: структуры социальных знаний и проблемы их описания. *Вопросы когнитивной лингвистики*, 2014, 4: 5–17. [Iriskhanova O. K. Sociocognitive linguistics: Social knowledge structures and problems of their description. *Voprosy kognitivnoi lingvistiki*, 2014, 4: 5–17.]
- Ирисханова (ред.) 2022 — Ирисханова О. К. (ред.). *Полимодальные измерения дискурса*. М.: Языки славянской культуры, 2022. [Iriskhanova O. K. (ed.). *Polimodal'nye izmereniya diskursa* [Multimodal dimensions of discourse]. Moscow: Yazyki slavyanskoi kul'tury, 2022.]
- Капанадзе, Красильникова 1973 — Капанадзе Л. А., Красильникова Е. В. Жест в разговорной речи. *Русская разговорная речь. Приложение*. Земская Е. А. (ред.). М.: Наука, 1973, 464–481. [Kapanadze L. A., Krasilnikova E. V. Gesture in colloquial speech. *Russkaya razgovornaya rech'. Prilozhenie*. Zemskaya E. A. (ed.). Moscow: Nauka, 1973, 464–481.]
- Кирик 2018a — Кирик А. А. Русский мультиканальный дискурс. Ч. I. Постановка проблемы. *Психологический журнал*, 2018, 39(1): 70–80. [Kibrik A. A. Russian multichannel discourse. Part I. Problem statement. *Psikhologicheskii zhurnal*, 2018, 39(1): 70–80.]
- Кирик 2018b — Кирик А. А. Русский мультиканальный дискурс. Ч. II. Разработка корпуса и направления исследований. *Психологический журнал*, 2018, 39(2): 79–90. [Kibrik A. A. Russian multichannel discourse. Part II. Corpus development and research directions. *Psikhologicheskii zhurnal*, 2018, 39(2): 79–90.]
- Крейдлин 2002 — Крейдлин Г. Е. *Невербальная семиотика*. М.: Новое литературное обозрение, 2002. [Kreidlin G. E. *Neverbal'naya semiotika* [Non-verbal semiotics]. Moscow: Novoe literaturnoe obozrenie, 2002.]
- Крейдлин 2020 — Крейдлин Г. Е. (отв. рук.). *Язык и семиотика тела*. М.: Новое литературное обозрение, 2020. [Kreidlin G. E. (ed.). *Yazyk i semiotika tela* [Language and body semiotics]. Moscow: Novoe literaturnoe obozrenie, 2020.]
- Лотман 2000 — Лотман Ю. М. *Семiosфера*. СПб.: Искусство–СПБ, 2000. [Lotman Yu. M. *Semiosfera* [Semiosphere]. St. Petersburg: Iskustvo–SPB, 2000.]
- Мерло-Понти 1945/1999 — Мерло-Понти М. *Феноменология восприятия*. Вдовина И. С., Фокин С. Л. (ред.). СПб.: Ювента; Наука, 1999 (1-я публ. 1945 г.). [Merleau-Ponty M. *Fenomenologiya vospriyatiya* [Phenomenology of perception]. Vdovina I. S., Fokin S. L. (eds.). St. Petersburg: Yuventa, Nauka, 1999 (first publ. in 1945).]
- Мечковская 2004 — Мечковская Н. Б. *Семиотика: Язык. Природа. Культура*. М.: Академия, 2004. [Mechkovskaya N. B. *Semiotika: Yazyk. Priroda. Kul'tura* [Semiotics: Language. Nature. Culture]. Moscow: Akademiya, 2004.]
- Носачев 2008 — Носачев П. Г. У. Эко: путь от семиотики к онтологии. *Вестник ПСТГУ. Серия 1: Богословие. Философия*, 2008, 23: 33–44. [Nosachev P. G. U. Eco: The path from semiotics to ontology. *Vestnik PSTGU. Seriya 1: Bogoslovie. Filosofiya*, 2008, 23: 33–44.]
- Переверзева (ред.) 2023 — Переверзева С. И. (ред.). *Грани естественного языка и кинесики: Сборник статей к 75-летию Григория Ефимовича Крейдлина*. М.: Дискурс, 2023. [Pereverzeva S. I. (ed.). *Grani estestvennogo yazyka i kinesiki: Sbornik statei k 75-letiyu Grigoriya Efimovicha Kreidlina*

- [Aspects of natural language and kinesics: A collection of papers on the 75th anniversary of Grigoriï Efimovich Kreidlin]. Moscow: Diskurs, 2023.]
- Пирс 2000 — Пирс Ч. С. *Избранные философские произведения*. Голубович К., Чухрукидзе К., Дмитриев Т. (пер.). М.: Логос, 2000. [Peirce Ch. S. *Izbrannnye filosofskie proizvedeniya* [Selected philosophical papers]. Golubovich K., Chukhruidze K., Dmitriev T. (trans.). Moscow: Logos, 2000.]
- Рубец 2009 — Рубец М. В. Восприятие тела в китайской традиции. *Телесность как эпистемологический феномен*. Бескова И. А. (отв. ред.). М.: ИФ РАН, 2009, 147–161. [Rubets M. V. Perception of the body in the Chinese tradition. *Telesnost' kak epistemologicheskij fenomen*. Beskova I. A. (ed.). Moscow: Institute of Philosophy, 2009, 147–161.]
- Руссо 2013 — Руссо М. М. Исчезающие и возникающие органы. *Лингвистика и методика преподавания иностранных языков*, 2013, 5: 173–199. [Rousseau M. M. Disappearing and emerging organs. *Lingvistika i metodika prepodavaniya inostrannykh yazykov*, 2013, 5: 173–199.]
- де Серто 1975/2001 — де Серто М. Разновидности письма, разновидности истории. Дубин Б. (пер.). *Логос*, 2001, 4: 4–18 (1-я публ. 1975 г.). [de Certeau M. Writings and histories. Dubin B. (trans.). *Logos*, 2001, 4: 4–18 (first publ. in 1975).]
- Янко 2008 — Янко Т. Е. *Интонационные стратегии русской речи в сопоставительном аспекте*. М.: Языки славянских культур, 2008. [Yanko T. E. *Intonatsionnye strategii russkoi rechi v сопоставительном аспекте*. Moscow: Yazyki slavyanskikh kul'tur, 2008.]
- Adams, Aizawa 2019 — Adams F., Aizawa K. Embodied cognition and the extended mind. *The Routledge companion to philosophy of psychology*. Robins S., Symons J., Calvo P. (eds.). London: Routledge, 2019, 193–213.
- Barbieri 2008 — Barbieri M. Biosemiotics: A new understanding of life. *Naturwissenschaften*, 2008, 95(7): 577–599.
- Bucholtz, Hall 2016 — Bucholtz M., Hall K. Embodied sociolinguistics. *Sociolinguistics: Theoretical Debates*, 2016, 1(1): 173–200.
- Busch 2021 — Busch B. The body image: Taking an evaluative stance towards semiotic resources. *International Journal of Multilingualism*, 2021, 18(2): 190–205.
- Clark 1997 — Clark A. *Being there: Putting brain, body, and world together again*. Cambridge (MA): MIT Press, 1997.
- Cobley 2016 — Cobley P. *Cultural implications of biosemiotics*. Dordrecht: Springer, 2016.
- Creider 1977 — Creider C. A. Towards a description of East African gestures. *Sign Language Studies*, 1977, 14: 1–20.
- Damasio 1994 — Damasio A. R. *Descartes' error: Emotion, reason, and the human brain*. New York: Avon Books, 1994.
- Deely 1990 — Deely J. *Basics of semiotics*. Bloomington: Indiana Univ. Press, 1990.
- Deely 2013 — Deely J. Objective reality and the physical world: Relation as key to understanding semiosis. *Semiotics*, 2013, 317–379.
- Di Paolo et al. 2018 — Di Paolo E. A., Cuffari E. C., De Jaegher H. *Linguistic bodies: The continuity between life and language*. Cambridge (MA): The MIT Press, 2018.
- Duranti 1992 — Duranti A. Language and bodies in social space: Samoan ceremonial greetings. *American Anthropologist*, 1992, 94(3): 657–691.
- Fauconnier, Turner 1998 — Fauconnier G., Turner M. Conceptual integration networks. *Cognitive Science*, 1998, 22(2): 133–187.
- Favareau 2009 — Favareau D. The semiotics of nature: Code-duality. *Essential Readings in Biosemiotics: Anthology and Commentary*. Dordrecht: Springer, 2009, 583–628.
- Geeraerts 2021 — Geeraerts D. Cognitive semantics. *The Routledge handbook of cognitive linguistics*. Wen X., Taylor J. R. (eds.). London: Routledge, 2021, 19–29.
- Goodwin 2000 — Goodwin C. Action and embodiment within situated human interaction. *Journal of Pragmatics*, 2000, 32(10): 1489–1522.
- Goodwin 2003 — Goodwin C. The semiotic body in its environment. *Discourses of the body*. Coupland J., Gwyn R. (eds.). London: Palgrave Macmillan, 2003, 19–42.
- Gumperz 1964 — Gumperz J. J. Linguistic and social interaction in two communities. *American Anthropologist*, 1964, 66(6.2): 137–153.
- Hoffmeyer 2008 — Hoffmeyer J. The Semiotic Body. *Biosemiotics*, 2008, 1: 169–190.
- Hübler 2006 — Hübler A. Assessing the body's share in conversation. *Semiotic bodies, aesthetic embodiments, and cyberbodies*. Nöth W. (ed.). Kassel: Kassel Univ. Press, 2006, 49–76.

- Ipsen 2006 — Ipsen G. Hybridization and extensions of the human body, or: The conquest of nature by culture. *Semiotic bodies, aesthetic embodiments, and cyberbodies*. Nöth W. (ed.). Kassel: Kassel Univ. Press, 2006, 77–99.
- Johnson 1987 — Johnson M. *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: Univ. of Chicago Press, 1987.
- Kendon 2004 — Kendon A. *Gesture: Visible action as utterance*. New York: Cambridge Univ. Press, 2004.
- Kull 1999 — Kull K. Biosemiotics in the twentieth century: A view from biology. *Semiotica*, 1999, 127(1–4): 385–414.
- Lakoff 1987 — Lakoff G. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: Univ. of Chicago Press, 1987.
- Lakoff, Johnson 1980 — Lakoff G., Johnson M. *Metaphors we live by*. Chicago: Univ. of Chicago Press, 1980.
- Langacker 1999 — Langacker R. W. The contextual basis of cognitive semantics. *Language and conceptualization*, 1999, 1: 229–253.
- Lidov 2018 — Lidov D. Embodiment and disembodiment in general semiotics. *Language and Semiotic Studies*, 2018, 4(3): 1–21.
- Martin, Zappavigna 2019 — Martin J. R., Zappavigna M. Embodied meaning: A systemic functional perspective on paralanguage. *Functional Linguistics*, 2019, 6(1): 1–33.
- McDonald 2013 — McDonald E. Embodiment and meaning: moving beyond linguistic imperialism in social semiotics. *Social Semiotics*, 2013, 23(3): 318–334.
- O'Neill 2008 — O'Neill S. *Interactive media: the semiotics of embodied interaction*. London: Springer, 2008.
- Paolucci 2021 — Paolucci C. *Cognitive semiotics*. Cham: Springer, 2021.
- Pelkey 2023 — Pelkey J. Embodiment and language. *WIREs Cognitive Science*, 2023, 14(5): e1649.
- Poyatos 1977 — Poyatos F. The morphological and functional approach to kinesics in the context of interaction and culture. *Semiotica*, 1977, 20(3–4): 197–228.
- Shapiro 2007 — Shapiro L. The embodied cognition research programme. *Philosophy Compass*, 2007, 2(2): 338–346.
- Sinha, Jensen de López 2000 — Sinha C., Jensen de López K. Language, culture and the embodiment of spatial cognition. *Cognitive Linguistics*, 2000, 11: 17–41.
- Sparhawk 1976 — Sparhawk C. M. P. *Linguistics and gesture: an application of linguistic theory to the study of Persian emblems*. Ph.D. diss. Ann Arbor: Univ. of Michigan, 1976.
- Stjernfelt 2006 — Stjernfelt F. The semiotic body. A semiotic concept of embodiment. *Semiotic bodies, aesthetic embodiments, and cyberbodies*. Nöth W. (ed.). Kassel: Kassel Univ. Press, 2006, 13–48.
- Talmy 2011 — Talmy L. Cognitive Semantics: An overview. *Semantics. An international handbook of natural language meaning*. Vol. 1. Maienborn C., Heusinger K., Portner P. (eds.). Berlin: De Gruyter Mouton, 2011, 622–642.
- Terrado-Pablo 2017 — Terrado-Pablo J. Toward a history of cross-cultural written symbols. *Language and Semiotic Studies*, 2017, 3(1): 62–78.
- Turner 2016 — Turner J. S. Semiotics of a superorganism. *Biosemiotics*, 2016, 9(1): 85–102.
- Usmonova 2020 — Usmonova D. S. A typological analysis of body parts names in English as part of somatic phraseology. *Проблемы современной науки и образования*, 2020, 2 (147): 32–34. [Usmonova D. S. A typological analysis of body parts names in English as part of somatic phraseology. *Problemy sovremennoi nauki i obrazovaniya*, 2020, 2 (147): 32–34.]
- Varela et al. 1991 — Varela F., Thompson E., Rosch E. *The embodied mind*. Cambridge (MA): MIT Press, 1991.
- Violi 2003 — Violi P. Embodiment at the crossroads between cognition and semiosis. *Recherches en Communication*, 2003, 19: 199–217.
- Zlatev 2012 — Zlatev J. Cognitive semiotics: An emerging field for the transdisciplinary study of meaning. *Public Journal of Semiotics*, 2012, 4(1): 2–24.

[Рец. на: / Review of:] **L. Johanson. *Code Copying. The strength of languages in take-over and carry-over roles***. Leiden: Brill, 2023. xvi + 149 p. (Brill Studies in Language, Cognition and Culture 38.) ISBN: 978-90-04-54843-5.

Тимур Анатольевич Майсак

Институт языкознания РАН, Москва, Россия;
Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия;
timur.maisak@gmail.com

Timur A. Maisak

Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia; HSE University, Moscow, Russia; timur.maisak@gmail.com

DOI: 10.31857/0373-658X.2024.2.159-164

Книга Ларса Юхансона, название которой можно перевести как «Копирование кода: сила языков в приобретении и передаче», посвящена процессам, происходящим в ситуации языкового контакта. Ларс Юхансон (р. 1936) — известный шведский тюрколог, автор многочисленных работ по структуре тюркских языков, в том числе создатель оригинальной концепции описания аспектуальности в тюркских и других языках мира, — в последние 30 лет разрабатывает собственную типологию контактных изменений. Его подход, названный им моделью копирования кода (Code-Copying Model), ранее был представлен в многочисленных статьях, теперь же основные его положения суммированы в рецензируемой монографии.

Стоит сразу сказать, что, несмотря на монографический характер работы, концепция языковых контактов Юхансона представлена в ней в относительно кратком и, можно сказать, «тезисном» виде. Книга имеет небольшой объем (на основной текст в ней приходится 125 страниц), и сам автор в первом предложении предисловия называет ее «эссе» (с. ix). Таким образом, книга может служить общим введением в модель копирования кода для тех, кто еще не знаком с ней, или же своеобразным резюме основных положений для тех, кто уже имеет какое-то представление о данной теории, однако знакомство с книгой вряд ли отменяет обращение к предшествующим статьям Юхансона, в которых содержится более подробный разбор конкретных случаев (заметим, что в списке литературы к книге упоминается более 40 работ автора с 1971 по 2002 г.).

Структура книги, в особенности для издания сравнительно небольшого объема, необычна: в ней 11 глав, в большинстве своем коротких (от 5 до 15 страниц) и предлагающих обсуждение основных понятий в рамках описываемого подхода. Последняя глава, посвященная языковым контактам письменных языков, стоит особняком как по объему (более 50 страниц), так и по содержанию; к ней мы вернемся ниже. Помимо основного текста, в книгу входит большой список литературы, три указателя (предметный, языковой и именной), а также краткое предисловие и ряд служебных разделов, касающихся оформления книги.

Концепция контактных явлений, предлагаемая Юхансоном, сравнительно проста и опирается на ряд ключевых понятий, для которых автор предлагает, как правило, свою оригинальную терминологию. Знакомство с ней необходимо уже для понимания заглавия книги, в том числе ключевого сочетания «копирование кода». Вся основная терминология вводится в первой главе. Код (code) в подходе Юхансона — это общее обозначение языка или его разновидности; копирование (copying) — это имитация одного кода другим. Копирование кода (code copying) состоит в том, что определенные черты исходного кода, т. е. языка-донора (Model Code или Donor Code), используются в принимающем коде, т. е. языке-реципиенте (Basic Code или Recipient Code), что может происходить

в различных языковых подсистемах: лексике, сегментной фонологии, просодии, морфологии или синтаксисе. Копирование может быть полным (глобальным) или частичным (селективным). В случае глобального копирования элементы языка-донора переносятся в язык-реципиент полностью, как цельный «блок» материальных, семантических, сочетаемостных и частотных свойств (ср. немецкое слово *Bahnhof* ‘железнодорожный вокзал’, скопированное в турецкий язык Германии). При селективном копировании лишь некоторые из материальных, семантических, сочетаемостных или частотных свойств переносятся в язык-реципиент. Так, примером селективного семантического копирования является приобретение значения ‘Айпад’ у шведского слова *padda* ‘жаба’ под влиянием фонетически близкого английского *iPad*. Примером селективного копирования сочетаемостной близости изменение порядка слов в генитивной конструкции караимского языка тюркской семьи под славянским влиянием (условно, от *Karinanin atası* ‘Карины отец’ к *atası Karina'nin* ‘отец Карины’). Собственно материальным копированием можно считать заимствование определенных фонем или интонационных моделей (так, в некоторых разновидностях караимского под влиянием славянских языков развились палатализованные фонемы). Возможно и копирование на уровне частотности распространения той или иной грамматической модели. Например, в случае вариативности в языке базовых порядков слов SOV и SVO контакт с языком, в котором доминирует один из этих порядков, может привести к повышению его частотности и в языке-реципиенте (что также имело место в караимском, где порядок SVO утратил свой маркированный характер). Как правило, при взаимодействии языков первыми появляются глобальные копии, и лишь позже — копии селективные. При этом автор подчеркивает, что копия никогда не идентична оригиналу (в этом смысле она не является «заимствованием»): процесс копирования создает нечто новое, что постепенно приспосабливается к принимающему коду и далее развивается уже в рамках него. Более подробно с примерами, в основном из тюркских языков, понятия глобального и селективного копирования разъясняются в главах 2 и 3.

Еще одна пара понятий, фундаментальных для подхода Юхансона, — это копирование посредством приобретения (*take-over*) или передачи (*carry-over*). Для пояснения этого различия автор использует также дополнительное различие между первичным кодом (*primary code*), т. е., условно говоря, первым или основным языком носителя, и вторичным кодом (*secondary code*), т. е. его вторым языком. Приобретение происходит в том случае, если носители первичного кода копируют в него что-то из вторичного кода: такой процесс характерен для ситуаций асимметричного контакта, когда в миноритарный язык попадают копии из доминирующего языка. Передача же — это, напротив, вставка элементов первичного кода во вторичный, в результате чего происходит изменение того варианта языка, который используется носителями как неосновной. В исторической лингвистике традиционно рассматривается несколько ситуаций, в которых задействованы механизмы приобретения и передачи: так, суперстрат является результатом влияния доминирующего идиома, субстрат — напротив, результатом воздействия более второстепенного идиома на доминирующий, а адстрат — результатом взаимодействия двух примерно равноценных идиомов. В качестве примера в главе 1 рассмотрены типы взаимодействий в языках Волжско-Камского языкового союза, где встречается как приобретение, так и передача языковых черт, в том числе в одной паре языков (например, в чувашском языке тюркской семьи и марийском языке уральской семьи).

Итак, в трех первых главах сформулированы и проиллюстрированы основные понятия модели копирования кода; небольшие главы с 4-й по 10-ю посвящены различным аспектам, связанным с контактным взаимодействием языков. В главе 4 рассмотрена грамматикализация в условиях контакта. Автор подчеркивает, что копированию подвергается не процесс грамматикализации как таковой, а только его результат. Носители языка-реципиента устанавливают, осознанно или нет, отношение эквивалентности между определенным элементом своего кода и элементом языка-донора, перенося его свойства на единицу собственного кода. Так, при контактно-обусловленном развитии двойственного числа

носители языка-реципиента могут использовать слово своего языка со значением 'два' в качестве показателя дуалиса. Это, однако, не будет означать, что данное числительное проходит тот же градуальный процесс грамматикализации, который прошел его аналог в языке-доноре: по мнению автора, превращение в грамматический показатель происходит одновременно. Зачастую показатель, появившийся в результате подобного аналогического развития, менее частотен и обязателен, чем исходный показатель в языке-доноре. В дальнейшем, однако, уже в рамках языка-реципиента скопированный показатель может претерпевать собственную эволюцию, достигая более продвинутых стадий грамматикализации. Эта тема продолжена в главе 5, посвященной переструктурированию (re-modelling) языков, в которой Юхансон пишет о том, что на протяжении столетий в языке могут накапливаться слои копирования, последующей конвенционализации и дальнейшей, уже внутриязыковой, грамматикализации языковых структур. Если таких, изначально скопированных, структур становится очень много, типологический профиль базового кода может значительно измениться. Однако и в этой главе, как и в других местах книги, автор подчеркивает, что даже меняющиеся подобным образом коды не меняют свою «наследственность» и не «смешиваются» с кодами-донорами: «Как кажется, ни один даже в высшей степени копирующий реципиентный код не превратился в тот донорский код, из которого он многое скопировал» (с. 51)¹. Копирование, впрочем, может оказаться и неуспешным, так что скопированный элемент подвергнется деконвенционализации и утрате.

Следующие две главы в основном посвящены разбору частных случаев. В главе 6 рассмотрены примеры взаимодействия тюркских языков с языками других семей, главным образом различных ветвей индоевропейской семьи, а также уральскими и монгольскими языками. Глава 7 посвящена краткому обзору контактов некоторых из крупных языков мира, а именно: английского, китайского, арабского и русского. Далее в главе 8, самой краткой в книге, автор останавливается на проблеме стабильности кодов, обладающих малым престижем и зачастую активно использующих копирование: рассмотрены некоторые проблемы отношения к родному языку внутри сообществ, а также документация такого рода идиомов. Глава 9 продолжает эту тему и содержит разбор примеров активно копирующих языков (high-copying codes), прежде всего близких автору северо-западного караимского, находящегося в славянском и балтийском окружении, и гагаузского на Балканах. Оба языка дают примеры всех видов копирования и в результате контакта с неродственными языками значительно изменили свой типологический профиль по сравнению с другими тюркскими языками. Юхансон вновь повторяет здесь свой тезис о том, что даже интенсивное копирование не ведет непременно к утрате языка: языковой сдвиг происходит лишь в том случае, когда первичный код перестает осуществлять коммуникативную функцию. При этом в мире находится множество примеров активно копирующих языков (от английского до того же гагаузского), которые вполне витальны и не являются угрожаемыми. Таким образом, отношение к такого рода языкам, их усвоению и использованию ни в коем случае не должно быть пренебрежительным.

В главе 10 рассмотрено противопоставление копий и когнатов — и те и другие являются примерами языковых сходств, но если первые появляются в результате контакта, часто между неродственными языками, то вторые наследуются родственными языками-потомками от общего предка. Юхансон рассматривает некоторые признаки, по которым, в случае отсутствия исторических свидетельств, можно предположить большую вероятность одного или другого сценария. Так, случаи «глобального» сходства на разных уровнях чаще наблюдаются при наследовании. О когнатности говорят и такие особые типы сходства, как чередования в основах или аффиксах. Напротив, ограниченность аффиксов заимствованными корнями скорее говорит о том, что мы имеем дело с заимствованными же

¹ Ранее в книге Юхансон упоминает языковой сдвиг как третий тип контактного изменения базового кода наряду с приобретением и передачей (с. 8), однако не говорит о сдвиге как о результате двух других процессов.

аффиксами. Как правило, если копирование на уровне аффиксов в принципе происходит, оно охватывает сразу несколько морфем. В главе обсуждаются некоторые аспекты сравнения языковых семей, входящих в гипотетическую алтайскую — или, как ее сейчас чаще принято называть, трансевразийскую, — макросемью. Заметим, что Юхансон не высказывает однозначного мнения о том, является ли наличие данной макросемьи как генеалогического единства доказанным или опровергнутым, однако и в данной главе, и в других местах книги он предостерегает от поспешных суждений о сходствах между языками входящих в нее семей как непременно проистекающих из контакта, а не наследования. В частности, Юхансон обращает внимание на то, что сходство между определенными типами словообразовательных показателей, проявляющееся как на материальном, так и на семантическом уровне, может более естественно объясняться развитием из общего праязыка, нежели копированием².

Наконец, глава 11 («Типы копирования в письменных языках»), как уже упоминалось, стоит несколько особняком. Насколько можно понять, она в наименьшей степени основана на предыдущих работах автора, а скорее, напротив, представляет собой оригинальное исследование, а именно общий обзор языковых ситуаций и языковых явлений, связанных с копированием в языках с письменной традицией, т. е. с копированием на уровне письменных текстов в большей степени, чем на уровне языков как таковых. В отличие от остальной части книги, в которой изложение более абстрактно, в данной главе рассмотрен конкретный материал из большого числа языковых семей и ареалов. Как отмечает Юхансон, поскольку литературные языки нередко имеют относительно «искусственный» характер, в них могут наблюдаться немного иные типологические, в том числе контактные, черты, чем в «естественных» языках (с. 80). Важным параметром является относительный престиж контактирующих языков, и всего в главе обсуждается пять ситуаций взаимодействия между более престижным кодом (*higher-ranking code*) и менее престижным кодом (*lower-ranking code*); под обоими кодами здесь подразумевается письменный язык. В первом случае происходит приобретение за счет копирования: копии из более престижного языка вставляются в тексты на менее престижном. Наблюдать это можно, в частности, применительно к аккадскому влиянию на западносемитские языки, арабскому влиянию на новоперсидский и, в свою очередь, новоперсидскому влиянию на такие тюркские языки, как османский или чагатайский. Автор подчеркивает, что даже значительное число копий не делает принимающий язык «смешанным», и он сохраняет принадлежность к своей языковой группе. Так, письменный английский, несмотря на значительную долю романских лексических копий, остается германским языком³. Точно так же для новоперсидского, который стал первым в истории «исламизированным» языком со словарем, почти наполовину скопированным из арабского языка, арабский не стал «родителем», а может считаться в лучшем случае «приемным родителем» (т. е. скорее свойственником, чем родственником «по крови») или даже, может быть, более корректно, — «учителем» или «наставником» (с. 90).

Второй случай — это копирование в результате передачи из менее престижного в более престижный литературный язык, примерами чего может служить влияние пракритов на неканонические варианты санскрита, а также японское, монгольское и маньчжурское

² Напомним, что несколько лет назад был издан сборник статей, посвященный противопоставлению копий и когнатов в морфологии, одним из редакторов которого был Юхансон [Johanson, Robbeets (eds.) 2012]; в одной из статей этого сборника [Robbeets 2012] приводится аргументация в пользу унаследованного характера ряда алтайских глагольных аффиксов. См. также рецензию на данный сборник [Майсак 2013].

³ В целом Юхансон подчеркивает, что «лексикон не является надежным диагностическим инструментом генеалогической классификации» и что распространенное представление о том, что базисная лексика не подвержена заимствованию, далеко не бесспорно, поскольку «даже большие части базисных словарей могут копироваться» (с. 83).

влияние на некоторые разновидности китайского. В третьем случае элементы более престижного и менее престижного кодов сосуществуют в одном тексте: здесь мы имеем дело не с копированием, а со своего рода «переключением кода» на письменном уровне. Зачастую это происходит, в том числе с целью сохранить и продемонстрировать свою идентичность, в текстах, написанных носителями менее престижного языка на более престижном идиоме. Примерами могут служить двуязычная («макароническая») поэзия Европы и Азии, двуязычные средневековые еврейско-романские тексты или использование среднеанглийского, англо-нормандского и латинского языков в деловой переписке средневековой Англии. Еще один, четвертый случай взаимодействия, — это использование менее престижного языка для экспликации текстов на более престижном. Сюда относится японская традиция аннотирования для чтения и понимания китайских текстов (значки *кунтэн*), бирманоязычные и сингалоязычные пояснения для чтения текстов на пали, а также караимские пословные переводы, использовавшиеся для понимания текстов на иврите. Наконец, в пятом случае элементы более престижного кода используются в тексте для графической репрезентации менее престижного кода (явление, известное как аллоглоттография). Так, в частности, применялась клинопись для записи семитских языков, эламская письменность для записи древнеперсидского, арамейская — для записи среднеперсидского, а китайские иероглифы — для передачи японского языка. Каждая из упомянутых контактных ситуаций в главе 11 кратко описывается на одной-двух страницах, как правило, с упоминанием большого количества ссылок на литературу по соответствующему вопросу.

Говоря о стиле книги, повторим, что он носит во многом конспективный, «тезисный» характер; далеко не всегда формулировки автора развернуты в полноценный текст. Данную работу Юхансона трудно назвать фундаментальной — это скорее очерки на соответствующую тему. Иногда в тексте встречаются назывные предложения, повторы предложений, опечатки и небольшие ошибки: так, в имени известного нидерландского специалиста по языковым контактам Питера Мейскена (Pieter Muysken) на с. 3, 138, 149 сделаны сразу две ошибки, тохарский язык в одном месте назван иранским (с. 56), а белорусский обозначен как White Russian (с. 116), тогда как в современной англоязычной славистике этот термин давно не употребляется. Начиная со с. 18 в книге многократно упоминаются «ирано-тюркские разновидности» (Irano-Turkic varieties или просто Irano-Turkic), при этом по сути нигде ни разу не поясняется, что под этим подразумевается. В целом ощущается недостаточная редакторская работа над книгой.

Содержательно, однако, новая и при этом итоговая книга Ларса Юхансона представляет безусловный интерес и может быть рекомендована всем, кто интересуется проблематикой, связанной с языковыми контактами. Можно согласиться с тем, что, как утверждает в предисловии сам автор, используемая в книге терминология интуитивна и проста, а изложение ведется ясным и аргументированным образом. Юхансон отмечает, что его модель подойдет и для описания копирования других процессов, помимо языковых, например культурных. Было бы интересно в будущем увидеть и такое применение теории «копирования кода». Завершить же нам бы хотелось одной из основных мыслей автора, которая высказывается на протяжении изложения не раз, а также фигурирует в самом названии (в слове *strength* ‘сила’). Ларс Юхансон подчеркивает, что копирование в языках не представляет собой «опасность» и не ведет к утрате языков; напротив, это креативный и подчиняющийся определенным правилам процесс, который может увеличить мощь копирующего кода за счет облегчения коммуникации между многоязычными носителями (с. x).

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

Майсак 2013 — Майсак Т. А. [Рец. на:] L. Johanson, M. Robbeets (eds.). *Copies versus cognates in bound morphology*. Leiden: Brill, 2012. *Вопросы языкознания*, 2013, 5: 130–135. [Maisak T. A. [Review of:]

- L. Johanson, M. Robbeets (eds). *Copies versus cognates in bound morphology*. Leiden: Brill, 2012. *Voprosy Jazykoznanija*, 2013, 5: 130–135.]
- Johanson, Robbeets (eds.) 2012 — Johanson L., Robbeets M. (eds.). *Copies vs. cognates in bound morphology*. Leiden; Boston: Brill, 2012.
- Robbeets 2012 — Robbeets M. Shared verb morphology in the Transeurasian languages: Copy or cognate? *Copies vs. cognates in bound morphology*. Johanson L., Robbeets M. (eds.). Leiden; Boston: Brill, 2012, 427–446.

Получено / received 13.01.2024

Принято / accepted 30.01.2024

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	Дата выхода в свет	
Формат 70×100 ¹ / ₆ Уч.-изд. л.	Тираж экз. Зак.	Цена свободная

Учредители: Российская академия наук, Институт русского языка им. В. В. Виноградова РАН

16+

Издатель: Российская академия наук
20 экземпляров распространяется бесплатно

Исполнитель: 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