

УДК 8:372.8
ISSN 2500-2953

Rhema. Рема

4.2025

Издается с 2002 г.

**Учредитель
и издатель:**

Московский
педагогический
государственный
университет

Выходит 4 раза в год

Журнал входит в Перечень ведущих
рецензируемых научных журналов
и изданий ВАК РФ:

Педагогика

5.8.2. Теория и методика обучения
и воспитания

Филология

5.9.1. Русская литература и литературы
народов Российской Федерации

5.9.3. Теория литературы

5.9.5. Русский язык. Языки народов России

5.9.6. Языки народов зарубежных стран

5.9.8. Теоретическая, прикладная
и сравнительно-сопоставительная
лингвистика

Когнитивные науки

5.12.3. Междисциплинарные исследования
языка (филологические науки)

Свидетельство
о регистрации СМИ:
ПИ № ФС 77–67769
от 17.11.2016

Адрес редакции:

109240, Москва,
ул. В. Радищевская,
д. 16–18, комн. 223

Сайт: rhema-journal.com

E-mail: rhema.pema@gmail.com

Подписной индекс журнала по Объединенному каталогу
«Пресса России» – **85006**

ISSN 2500-2953

Rhema. Pema

4.2025

**The Founder
and Publisher:**

Moscow Pedagogical
State University

Mass media
registration
certificate

ПИ № ФС 77-67769
as of 17.11.2016

Editorial office:

Moscow, Russia,
Verhnyaya
Radishchevskaya str.,
16–18, room 223,
109240

The journal is included in the list of the leading peer-reviewed scholarly journals the Higher Attestation Commission of the Ministry of Science and Higher Education of the Russian Federation recommended to PhD candidates and those working for their habilitation who wish to publish the results of their research.

The journal has been published since 2002

The journal is published 4 times a year

E-mail: rhema.pema@gmail.com

Information on journal can be accessed via: rhema-journal.com

Редакционная коллегия

Главный редактор

Антон Владимирович Циммерлинг – доктор филологических наук; ведущий лабораторией типологической и теоретической лингвистики Государственного института русского языка имени А.С. Пушкина, г. Москва; ведущий научный сотрудник отдела типологии и ареальной лингвистики Института языкознания РАН, г. Москва

Заместитель главного редактора

Екатерина Анатольевна Лютикова – доктор филологических наук, доцент; профессор кафедры теоретической и прикладной лингвистики филологического факультета Московского государственного университета имени М.В. Ломоносова

Ответственный секретарь

Анастасия Алексеевна Герасимова – кандидат филологических наук; научный сотрудник лаборатории автоматизированных лексикографических систем Научно-исследовательского вычислительного центра Московского государственного университета имени М.В. Ломоносова

Татьяна Михайловна Воителева – доктор педагогических наук, профессор; профессор кафедры методики преподавания русского языка и литературы Московского государственного областного университета

Наталья Дмитриевна Гальскова – доктор педагогических наук, профессор; профессор кафедры лингводидактики Государственного университета просвещения, г. Москва

Елена Валентиновна Гетманская – доктор педагогических наук; профессор кафедры методики преподавания литературы Института филологии Московского педагогического государственного университета

Павел Валерьевич Гращенков – доктор филологических наук; доцент кафедры теоретической и прикладной лингвистики филологического факультета Московского государственного университета имени М.В. Ломоносова; старший научный сотрудник отдела языков народов Азии и Африки Института востоковедения РАН, г. Москва

Атле Грённ – профессор кафедры литературы, страноведения и европейских языков Университета Осло, Норвегия

Сурен Тигранович Золян – доктор филологических наук, профессор; ведущий научный сотрудник отдела теоретической философии Института философии, социологии и права Национальной Академии наук Армении, г. Ереван, Армения; профессор Института гуманитарных наук Балтийского федерального университета имени Иммануила Канта, г. Калининград

Алексей Александрович Корнев – кандидат педагогических наук; доцент кафедры теории преподавания иностранных языков факультета иностранных языков и регионоведения Московского государственного университета имени М.В. Ломоносова

Анатолий Симонович Либерман – доктор филологических наук; профессор кафедры немецкого, нидерландского и скандинавских языков Университета Миннесоты, г. Миннеаполис, США

Сильвия Лураги – PhD (филология); профессор факультета гуманитарных наук, Университет Павии, Италия

Михаил Николаевич Михайлов – PhD (филология); профессор переводоведения (русский и финский языки) Института языков, перевода и литературы Университета Тампере, Финляндия

Нерея Мадарьяга Писано – PhD (филология); доцент кафедры классических языков (секция славянской филологии) Университета Страны Басков, г. Витория, Испания

Владимир Александрович Плунгян – доктор филологических наук, профессор, академик РАН; заместитель директора Института русского языка им. В.В. Виноградова РАН; заведующий сектором типологии Института языкознания РАН; профессор кафедры теоретической и прикладной лингвистики Московского государственного университета имени М.В. Ломоносова

Велка Александрова Попова – кандидат филологических наук; доцент кафедры болгарского языка, сотрудник лаборатории прикладной лингвистики факультета гуманитарных наук Шуменского университета имени Епископа Константина Преславского, Болгария

Наталья Вадимовна Сердобольская – кандидат филологических наук; доцент Учебно-научного центра лингвистической типологии Института лингвистики Российского государственного гуманитарного университета

Андрей Стоянович – доктор филологических наук, профессор; заведующий кафедрой иностранных языков Белградского университета, Сербия

Младен Ухлик – доктор филологических наук, профессор; заведующий кафедрой русского языка Отделения славянских языков философского факультета Университета Любляны, Словения

Александр Иосифович Федута – доктор филологических наук; редактор биографического альманаха «Асоба і час» («Личность и время»), г. Минск, Республика Беларусь

Любовь Георгиевна Чапаева – доктор филологических наук; профессор кафедры общего и прикладного языкознания Московского педагогического государственного университета

Editorial Board

Editor-in-Chief

Anton V. Zimmerling – Dr. Phil. Hab.; head of the Laboratory of Typological and Theoretical Linguistics, Pushkin State Russian Language Institute; leading researcher at the Department of Typology and Areal Linguistics, Institute of Linguistics, Russian Academy of Sciences, Moscow, Russian Federation

Deputy chief editor

Ekaterina A. Lyutikova – Dr. Phil. Hab.; professor at the Department of Theoretical and Applied Linguistics, Lomonosov Moscow State University, Russian Federation

Executive secretary

Anastasia A. Gerasimova – PhD in Linguistics; researcher at the Research Computing Center, Lomonosov Moscow State University, Moscow, Russian Federation

Lyubov G. Chapaeva – Dr. Phil. Hab.; professor at the Department of General and Applied Linguistics, Institute of Philology and Foreign languages, Moscow Pedagogical State University, Russian Federation

Aleksandr Feduta – Dr. Phil. Hab.; editor-in-chief of the biographical almanac «Personality and Time», Minsk, Belarus

Natalia D. Galskova – Dr. Ped. Hab.; professor at the Department of Linguodidactics, Federal State University of Education, Moscow, Russian Federation

Elena V. Getmanskaya – Dr. Ped. Hab.; professor at the Department of Methods of Teaching Literature, Moscow Pedagogical State University, Russian Federation

Pavel V. Grashchenkov – Dr. Phil. Hab.; associate professor at the Department of Theoretical and Applied Linguistics, Faculty of Philology, Lomonosov Moscow State University; research fellow at the Institute of Oriental Studies, Russian Academy of Sciences, Moscow, Russian Federation

Atle Grønn – Dr. Phil. Hab.; professor at ILOS – Department of Literature, Area Studies and European Languages, University of Oslo, Norway

Alexey A. Korenev – PhD in Pedagogy; associate professor at the Department of Foreign Language Teaching Theory, Faculty of Foreign Languages and Area Studies, Lomonosov Moscow State University, Russian Federation

Anatoly Liberman – Dr. Phil. Hab.; professor at the Department of German, Dutch and Scandinavian, University of Minnesota, USA

Silvia Luraghi – PhD in Philology; associate professor at the Department of Humanities, Section of General and Applied Linguistics, University of Pavia, Italy

Mikhail Mikhailov – PhD in Philology; professor at the School of Language, Translation and Literary Studies, University of Tampere, Finland

Nerea Madariaga Pisano – PhD in Philology; associate professor at the Department of Classical Studies (Section of Slavic Philology), University of the Basque Country, Vitoria, Spain

Vladimir A. Plungian – Dr. Phil. Hab., full member of the Russian Academy of Sciences; deputy director of the Vinogradov Institute of Russian Language, Russian Academy of Sciences; head of the Sector of Typology, Institute of linguistics, Russian Academy of Sciences; professor at the Department of Theoretical and Applied Linguistics, Lomonosov Moscow State University, Russian Federation

Velka A. Popova – PhD in Philology; associate professor at the Department of Bulgarian, research fellow at the Laboratory of Applied Linguistics, Faculty of Humanities, the Constantin of Preslav Bishop Shumen University, Bulgaria

Natalia V. Serdobolskaya – PhD in Philology; associate professor at the Training and Research Center for Linguistic Typology, Institute of Linguistics, Russian State University for the Humanities, Moscow, Russian Federation

Andrej Stojanović – Dr. Phil. Hab., full professor; head of the Department of Foreign Languages, University of Belgrade, Serbia

Mladen Uhlik – Dr. Phil. Hab., professor; head of the Department of Russian, Institute of Slavic Languages, Faculty of Philosophy, University of Ljubljana, Slovenia

Tatiana M. Voiteleva – Dr. Ped. Hab.; professor at the Department of the Methods of Teaching Russian Language and Literature, Moscow Region State University, Russian Federation

Suren T. Zolyan – Dr. Phil. Hab., professor; leading research fellow at the Department of Theoretical Philosophy, Institute of Philosophy, Sociology and Law Studies, National Academy of Sciences, Yerevan, Armenia; professor at the Institute for the Humanities, Immanuel Kant Baltic Federal University, Kaliningrad, Russian Federation

Содержание

ЛИНГВИСТИКА

С.С. Камаданова

Индоарийские инфинитивы
в сложных словах «бахуврихи»:
от древнеиндийских до среднеиндийских языков 9

П.М. Котта Рамузино

Биномальные конструкции с указанием
на малое количество в русском и итальянском языках 27

В.И. Лемаев, Н.В. Лукашевич

Влияние порядка слов на качество
межязыкового переноса моделей
для распознавания эмоций на русском языке 53

И. Мельчук

Фразы типов
Ивану не работается ~ Ивану хорошо работается:
два случая глагольной деривации 78

*Е.А. Миронова, Ш.Ш. Унарокова,
Ю.А. Ландер*

Абазинские и адыгейские демонстративы:
моноцентричность под видом полицентричности 106

Contents

LINGUISTICS

S. Kamadanova

Indo-Aryan infinitives n “bahuvrihi” compounds:
From ancient to Middle Indo-Aryan languages 9

P.M. Cotta Ramusino

Binominal constructions of small quantity
in Russian and Italian 27

V. Lemaev, N. Loukashevitch

The influence of word order on the quality
of cross-linguistic transfer
of emotion recognition models in Russian 53

I. Mel'čuk

Sentences of types *Ivanu ne rabotaetsja*
'Ivan is undisposed to work' ~ *Ivanu xorošo rabotaetsja*
'Ivan's work is going well': Two cases of verbal derivation 78

*E. Mironova, Sh. Unarokova,
Yu. Lander*

Abaza and West Circassian demonstratives:
Monocentricity under the guise of polycentricity 106

DOI: 10.31862/2500-2953-2025-4-9-26

УДК 811.211+.212+.213

С.С. Камаданова

Московский государственный университет
имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация

Индоарийские инфинитивы в сложных словах «бахуврихи»: от древнеиндийских до среднеиндийских языков

Статья посвящена изучению функционирования в истории индоарийского основосложений «бахуврихи» вида «инфинитив на *-(i)tum* + существительное *kāma-* ‘желание’ или *manas-* ‘мысль’». Ранее данные сложные слова никогда не становились объектом отдельного исследования ни в отношении синхронии, ни в плане диахронии. Во-первых, в работе рассматриваются структурные и функциональные особенности подобных грамматикализованных сложных слов; во-вторых, прослеживается история существования данной «конструкции» от ведийского – ранней фазы древнеиндоарийского через санскрит до среднеиндоарийских языков пра-критов. Как показало наше исследование, подобного рода примеры обнаруживаются не только в древнеиндоарийском и на ранней ступени среднеиндоарийских (пали), но и в более поздних среднеиндоарийских языках (ардхамагадхи и джайнском махараштри).

Ключевые слова: инфинитивы *-(i)tum*, сложные слова «бахуврихи», древнеиндийские языки, ведийский, санскрит, среднеиндийские языки, пали, ардхамагадхи, джайнский махараштри



ДЛЯ ЦИТИРОВАНИЯ: Камаданова С.С. Индоарийские инфинитивы в сложных словах «бахуврихи»: от древнеиндийских до среднеиндийских языков // Рема. Rhema. 2025. № 4. С. 9–26. DOI: 10.31862/2500-2953-2025-4-9-26

DOI: 10.31862/2500-2953-2025-4-9-26

S. Kamadanova

Lomonosov Moscow State University,
Moscow, 119991, Russian Federation

Indo-Aryan infinitives n “bahuvrihi” compounds: From ancient to Middle Indo-Aryan languages

The article is devoted to the study of how the “bahuvrihi” compounds function in the history of Indo-Aryan. They consist of infinitive in *-(i)tum* and noun *kāma-* ‘desire’ or *manas-* ‘thought’. Previously, the compounds of this type have never been the subject of a separate study, neither in terms of synchrony, nor of diachrony. Firstly, the paper examines the structural and functional features of such grammaticalized compounds; secondly, the history of this «construction» is also traced from the Vedic – the early phase of Old Indo-Aryan (OIA), through Sanskrit, to the Middle Indo-Aryan or Prakrit languages (MIA). According to research conducted, examples of these forms are found not only in OIA and at the early stage of MIA (Pali), but also at the later MIA phase (Ardhamagadhi and Jaina-Maharashtra).

Key words: infinitives in *-(i)tum*, “bahuvrihi” compounds, Ancient Indo-Aryan, Vedic, Sanskrit, Middle Indo-Aryan, Pali, Ardhamagadhi, Jaina-Maharashtra

FOR CITATION: Kamadanova S. Indo-Aryan infinitives in bahuvrihi compounds: From ancient to Middle Indo-Aryan languages. *Rhema*. 2025. No. 4. Pp. 9–26. DOI: 10.31862/2500-2953-2025-4-9-26

1. Введение

1.1. Инфинитивы на *-(i)tum*

В древнеиндийском (ведийском и санскрите)¹ существовало множество разных отглагольных производных, образующих разветвленную систему отглагольных имен. Некоторые из них традиционно определяются как «инфинитивы».

1.1.1. Отглагольные имена на *-(i)tum* в ведийском

В ведийском, как наиболее архаичной фазе древнеиндоарийского языка, такие неличные глагольные формы как инфинитивы, изначально представляли собой отглагольные имена от разных основ, маркированных теми или иными падежно-числовыми показателями.

1. Показатели датива:

- инфинитивы на $-e^2$ – исходно представляют собой отглагольные имена с окончанием дательного падежа единственного числа, присоединяемого непосредственно к корню (*vid-e* ‘знать’, *drś-e* ‘видеть’);
- инфинитивы на $-ase$ – восходят к основам на $-as$ с флексией дательного падежа (*jīv-ase* ‘жить’, *bhoj-ase* ‘поглощать’);
- инфинитивы на $-aye^3$ – исторически представляют собой отглагольные имена на $-i$ в форме дательного падежа единственного числа (*yudh-aye* ‘бороться’, *tuj-aye* – ‘ударять’);
- инфинитивы на $-taye^4$ – структурно омонимичны именам женского рода на $-ti$ в форме дательного падежа единственного числа [Елизаренкова, 1982, с. 362] (*vī-taye* ‘стремиться’ *iṣ-taye* ‘двигать’);

¹ Термины «древнеиндийский», «среднеиндийский» и «новоиндийский» относятся к периодизации истории развития индоарийских языков. К древнеиндийским языкам традиционно причисляют ведийский (примерно с XII в. до н.э. до сер. I тыс. до н.э.) и санскрит (примерно с V–IV вв. до н.э.). В среднеиндийский включаются многочисленные пракрыты – языки и диалекты, распространившиеся по Индостану с середины I тыс. до н.э. В частности, помимо прочих, к раннему среднеиндоарийскому относят пали, к среднему этапу среднеиндоарийских языков – ардхамагадхи, а поздний среднеиндоарийский представлен апабхрамша. Приблизительно с X в. начинается новоиндийский период, представленный несколькими десятками языков, к которым, помимо прочих, относятся хинди, панджаби, гуджарати, маратхи и т.д.

² У корней с исходом на $-ā$ при присоединении окончания $-e$ в соответствии с правилами сандхи образуется дифтонг $-ai$ ($-ā + -e \Rightarrow -ai$). Например, *prakhyaī* ‘смотреть’.

³ $-au$ является средней ступенью чередования (guṇa) корневой гласной $-i$.

⁴ Здесь и далее применительно к морфемам и словам на ведийском приведены формы без знаков ударения (/тона): например, $-ase$, $-aye$, $-taye$, а не $-áse$, $-áye$, $-táye$ и т.п.

- инфинитивы на *-tave* – диахронически представляют собой отглагольные имена с суффиксом *-tu* в форме дательного падежа единственного числа (*dhā-tave* ‘класть’, *kar-tave* ‘делать’);
- инфинитивы на *-mane* и *-vane* – восходят к именам в форме дательного падежа единственного числа, образованным от основ на *-man* и *-van* соответственно (*dā-mane* ‘давать’, *tur-vane* – ‘пересекать’).

2. Показатели аккузатива:

- инфинитивы на *-am* – образуются присоединением суффикса *-am* (застывшей флексии аккузатива) к корневым основам (*vr̥dh-am* ‘возрастать’, *ārabh-am* ‘начинать’);
- инфинитивы на *-(i)tum* – исторически представляют собой отглагольные имена с основой на *-tu*, к которой присоединена падежно-числовая флексия винительного падежа (*praṣ-ṭum* ‘спрашивать’, *han-tum* ‘убивать’).

3. Показатели генитива-аблатива:

- инфинитивы на *-as* – исторически представляют собой отглагольные имена, образующиеся путем присоединения к корню на согласный генитивно-аблативного окончания единственного числа (*avagam-as* ‘следовать’, *niḥsṛj-aḥ* ‘отпустить’);
- инфинитивы на *-tos* – так же, как и формы на *-tave* и *-tum*, диахронически соотносятся с отглагольными именами с основой на *-tu*, маркированной, в отличие от двух последних, флексией генитивно-аблативного падежа (*vad-itos* ‘говорить’, *car-itos* ‘идти’).

4. Показатели локатива:

- инфинитивы на *-i* – соотносятся с корневыми именами, маркированными окончанием локатива (*budh-i* ‘пробуждаться’, *vyuṣ-i* ‘восходить’);
- *-tari* – структурно представляют собой имена деятеля с суффиксом *-tar* в форме локатива единственного числа (*vidhar-tari* ‘поддерживать’, *so-tari* ‘выжимать’ и др.).

Однако ввиду того, что в ведийском данные формы уже не встроены в целостную парадигму склонения по падежам и числам, целесообразно рассматривать их не столько как отглагольные существительные, сколько как собственно инфинитивы, хотя граница между ними, как отмечают исследователи, не является достаточно четкой [Whitney, 1879, p. 185, 315; MacDonell, 1916, p. 167; Burrow, 1965, p. 122, 365; Барроу, 1976, с. 117, 340–341; Disterheft, 1980, p. 12].

Ведийские инфинитивы являются, по сути, изолированными формами отглагольных имен действия («*nomina actionis*»), которые не имеют

соответствующих коррелятов, образованных от тех же основ при помощи иной падежно-числовой флексии. В связи с этим «застывшие» окончания такого рода фиксированных ведийских форм следует рассматривать не как словоизменительные форманты, а как словообразовательные (деривационные) суффиксы [Елизаренкова, 1982, с. 358].

Среди всех указанных выше видов инфинитивов в ведийском⁵ одним из максимально редких по встречаемости является таковой на *-(i)tum*. Инфинитивы на *-(i)tum*, так же, как и инфинитивы на *-tave*⁶, исходно представляют собой отглагольные имена с основой на *-tu*, к которой присоединена падежно-числовая флексия. Однако, в отличие от инфинитивов на *-tave*, они присоединяют застывший маркер *-m* винительного, а не дательного падежа *-e* [МакДонелл, 1916, с. 194; Елизаренкова, 1982, с. 362–363]. Ср., например, формы *at-tave* и *at-tum*, образованные от корня *ad-* ‘есть’, ‘поедать’.

Образования «отглагольное имя / инфинитив на *-(i)tum + kāma-/manas-*», речь о которых пойдет далее, ни в самих ведах, ни в поздневедийских текстах нами обнаружены не были⁷. Самые ранние тексты, в которых засвидетельствованы подобные формы, принадлежат эпическому санскриту, представленному эпосами Махабхарата и Рамаяна.

1.1.2. Инфинитивы на *-(i)tum* в санскрите

Впоследствии (в эпическом и классическом санскрите) формы на *-(i)tum* остаются практически единственными активно и продуктивно порождаемыми инфинитивами [Whitney, 1879, р. 313; Disterheft, 1980, р. 14]. В санскрите данные инфинитивы образуются путем присоединения к корню в ступени «*guṇa*» единого деривационного суффикса *-(i)tum*, который уже перестал восприниматься как комбинация «суффикс отглагольного имени + флексия аккузатива» [Buttow, 1965, р. 364; Барроу, 1976, с. 339].

⁵ Об инфинитивах см.: в древнегреческом и латинском [Шантрен, 1953, с. 232–238; Тронский, 2001, с. 260–262; Sihler, 1995, р. 606–613], славянских языках [Горшкова, Хабургаев, 1981, с. 283–284, 286, 293, 295, 322–325, 344–351; Борковский, Кузнецов, 2006, с. 302; Изотов, 2010, с. 146], восточноиранских языках [Эдельман, 1990, с. 37–38].

⁶ Структура *-tave* образуется путем присоединения окончания *-e* (DAT.SG) к суффиксу *-tu*, принимающему в ступени «*guṇa*» в положении перед гласной вид *-av-*.

⁷ Такого рода выводы, дистрибутивно-статистического характера, сделаны при опоре на данные цифрового корпуса древнеиндийских текстов «Digital Corpus of Sanskrit» (см. Vedic Concordance). URL: <http://www.sanskrit-linguistics.org/dcs/index.php?contents=abfrage> (дата обращения: 11.11.2024).

(1) <i>śro-tum</i>	<i>icchāmahe</i>	<i>brahman-Ø</i>
услышать-INF ⁸	желать.PRS:1PL.MID	брахман-VOC.SG
<i>kiṃ</i>	<i>rāra-m</i>	
что-ACC.SG	дурное-ACC.SG	
<i>ḥtavān</i>	<i>asi </i>	
делать.PF.ACT.PART:NOM.M.SG.	быть.PRS:2SG	
‘О брахман, [мы] хотим услышать , что дурное [ты] совершил?’ ⁹		
(MBhT ¹⁰ -I.101.15.4)		

Инфинитивы на *-(i)tum* употребляются в древнеиндийском преимущественно в функции именной составляющей глагольно-именного сказуемого. См., например, предложение на эпическом санскрите (1), в котором инфинитив *śrotum* ‘услышать’ является инфинитивной частью составного глагольного сказуемого *śrotum icchāmahe* ‘хотим услышать’. Инфинитив здесь заполняет валентность прямого дополнения при переходном глаголе ‘хотеть’. Такого рода конструкции, демонстрирующие аналитический способ выражения соответствующих семо-синтаксических отношений, встречаются в индоарийском гораздо чаще, чем те редкие периферийные синтетические образования с существительными *kāta-* и *manas*, речь от которых и пойдет далее.

1.2. Синтетизм и аналитизм

В индоарийском (и в древнеиндоарийском, и в среднеиндоарийских языках) идею «желания/намерения [что-либо] сделать» можно было выразить не только аналитически (см. (1)), но и синтетически, как продемонстрировано в палийском примере (2). Ср. *śrotum icchāmahe* и *sotukāta*. Оба словосочетания выражают одну и ту же мысль ‘[мы] хотим услышать’ и включают в качестве первого компонента инфинитивы на *-(i)tum*, образованные от одного и того же корня *śru-* (древнеиндоарийский) > *su-* (среднеиндоарийский) ‘слышать’. Однако во втором случае инфинитив на *-(i)tum* препозиционно присоединяется в виде основы (без исторического маркера аккузатива *-m*) прямо к существительному *kāta*, образуя с последним неразделимое синтетическое единство.

⁸ В данной работе используется лейпцигская система глоссирования.

⁹ Переводы, приведенные здесь и далее, осуществлены автором статьи с языков оригинала.

¹⁰ Список сокращений: DaśKC – Daśa-kumāra-Carita; LañS – Lañkāvatārasūtra; MBhT – Mahābhārata; Rām – Rāmāyaṇa; Ṛv – Rīgveda; Vh – Vāsudevahiṇḍī. Все древнеиндийские тексты, кроме Vāsudevahiṇḍī, приведены по цифровому корпусу «Digital Corpus of Sanskrit» (<http://www.sanskrit-linguistics.org>). Текст Vāsudevahiṇḍī приведен по изданию: Vāsudevahiṇḍī. Part I. Bhavnagar: Vallabhadas Gandhi, Secretary Shri Jain Atmanand Sabda, 1930.

(2) <...> *amhākam pi tāva kathetha,*
 мы.DAT/GEN EMPH стoль_много рассказывать.OPT:2PL.ACT
so-tu+kām' *amhā*" *'ti.*
 слушать-INF+желание.NOM.M.PL. мы.NOM.PL PRCL
 '...Нам тоже побольше [о них] расскажите, мы **желаем**
послушать (букв. '**обладающие желанием послушать**')!
 [Fausboll, 1879, p. 185] Garahita-Jātaka

Также см. санскритское предложение (3), в первой части которого инфинитив *kar-tum* является частью распространенной конструкции с глаголом, образованным от корня *iṣ-* 'желать'. Во второй же части употребляется альтернативное, гораздо более редкое образование «инфинитив на *-tum* + существительное *manas* 'мысль'»: в данном случае – *bhoktu-manas* 'обладающий мыслью поглотить':

(3) *kaccin na*
 что-то.ACC.SG. NEG
ta-syāpāpa-sya (<= *ta-sya* + *apāpa-sya*)¹¹ *papa-ṃ*
 тот-GEN.SG_беззлобный-GEN.SG дурное-ACC.SG
kar-tum iheccha-si |
 делать-INF здесь_желать.PRS-2SG.ACT
akaṅṭakaṃ bhok-tu+manā
 беспрепятственно поглощать-INF+мысль:NOM.M.SG.
rājya-ṃ ta-syānuja-sya (<= *ta-sya* + *anuja-sya*) *ca ||*
 царство-ACC.SG тот-GEN.SG_младший.брат-GEN.SG и
 'Неужели [ты] **желаешь сделать** ему, [Рама], беззлобному,
 что-то дурное [и] **намерен** беспрепятственно **присвоить** его
 и [его] младшего брата [Лакшманы] царство?'
 (Rām: AY-84.13.1-2)

В истории индоарийского (от древнеиндоарийского через среднеиндоарийские к новым индоарийским языкам) в целом наблюдается тенденция перехода от синтетизма к аналитизму, поэтому с эволюционной точки

¹¹ В древнеиндийском действуют фонетические правила «сандхи» (санскр. *saṃdhi* 'слияние', 'соединение') – регулярные правила изменения звуков на морфных швах и границах слов (соответственно внутренние и внешние «сандхи»). Это, в частности, зачастую влечет и слитное написание словоформ. В данном примере соседство конечной краткой гласной *a* (в *tasya*) и начальной краткой гласной *a* (в последующем слове *apāpasya*) предполагают их слияние в одну долгую гласную *ā* (*tasya* + *apāpasya* => *tasyāpāpasya*). Правила сандхи, пусть и с меньшей регулярностью и системностью, но до некоторой степени выполняются и в среднеиндийских языках. В связи с этим для облегчения понимания глосс и соответствующих им элементов оригинального текста в наиболее трудных случаях в скобках приводится вариант текста с разьясненными внешними сандхи. Например, *ta-syāpāpa-sya* (<= *ta-sya* + *apāpa-sya*).

зрения синтетические образования типа «инфинитив на *-(i)tum* + существительное *kāma-* ‘желание’ или *manas-* ‘мысль’» являются менее продвинутыми по сравнению с соответствующими аналитическими аналогами. В связи с этим объяснимым и оправданным представляется постепенное исчезновение такого рода сложных слов из системы поздних среднеиндоарийских языков и, затем, новых индоарийских языков, где подобные формы уже не встречаются.

2. О структуре образований с *kāma-* и *manas-*

Как отечественные, так и зарубежные исследователи не уделяли подобным образованиям пристального внимания. В лучшем случае они ограничивались кратким упоминанием существования таких «форм» в череде всех прочих конструкций с инфинитивами [Ashtadhyayi, 1897, p. 1105; Wackernagel, 1905, p. 31, 44; Pischel, 1957, p. 395; Burrow, 1965, p. 365; Барроу, 1976, p. 341; Ghatge, 1993, p. 198; Speijer, 1998, p. 302; Geiger, 2005, p. 193; Keydana, 2013, p. 49; Парибок, 2022, с. 127]¹².

Первые прямые упоминания о нестандартном устройстве сочетаний инфинитивов на *-(i)tum* с именами *kāma-* и *manas-* засвидетельствованы в Aṣṭādhyāyī (‘Восьмикнижие’)¹³ Панини (≈V в. до н.э.) и встречаются почти с самого начала панинийской комментаторской традиции, как минимум, обнаруживаясь у Патанджали (II–I вв. до н.э.) и, несколько позднее, у Ваманы и Джаядитьи (≈VII в.)¹⁴.

Обращаясь к наследию классической индологии, отметим, что о возможности употребления инфинитивов в составе сложных слов «*-(i)tum + kāma-/manas-*» в санскрите кратко упоминает Дж.С. Шпайер.

¹² В работе Т. Оберлиса в разных разделах «Эпического санскрита» упоминаются две формы *yoddhu-kāma* ‘обладающий желанием сражаться’ и *hantu-kāma* ‘обладающий желанием убивать’, однако они являются частью более расширенного контекста примеров, обсуждаемых в связи с иными вопросами, не имеющими отношения к проблематике инфинитивов в целом и образований с *kāma-* и *manas-* в частности [Oberlies, 2003, p. 42, 199]. Аналогично касательно среднеиндийского буддийского гибридного санскрита см. [Edgerton, 1953, vol. I, p. 179, 218, 220], где в грамматическом обзоре обнаруживается лишь краткое упоминание нескольких сложных слов с *kāma* без указания на их специфику и особый статус: *syotu-kāma* ‘обладающий желанием шить’, *dhovitu-kāma* ‘обладающий желанием очистить’, *pivāpayetu-kāma* ‘обладающий желанием напоить’; в словаре Ф. Эджертон дополнительно зафиксированы следующие формы с *kāma*: *avitu-kāma* ‘обладающий желанием отправиться’, *prativinudayitu-kāma* ‘обладающий желанием избавиться’, *pratisamlātu-kāma* ‘обладающий желанием уйти в уединение’, *vikurvitu-kāma* ‘обладающий желанием явить чудеса’ [Ibid., vol. II, p. 8, 368, 370, 481].

¹³ См. сутру VI.1.144.

¹⁴ Подробнее о комментариях, поясняющих сутру Панини VI.1.144., см. [Sarma, 1968, p. 130; Roodbergen, 2008, p. 207–208, 243, 259].

Он иллюстрирует это примерами *ākhyātu-kāma* ‘обладающий желанием сказать’, *dagdhu-kāma* ‘обладающий желанием сжечь’, *vaktu-manas* ‘обладающий намерением сказать’, в которых, как он справедливо отмечает, имеет место «bahuvrīhi»¹⁵ [Speijer, 1998, p. 302]. Т. Барроу отмечает, что инфинитивы на *-(i)tum*, включаясь в состав сложных слов с *kāma-/manas-* как именная основа, сохраняют таким образом след своего именного происхождения [Burgow, 1965, p. 365; Барроу, 1976, с. 341].

Я. Ваккернагель упоминает эти санскритские формы в контексте сложных слов, в которых второй член композита управляет первым. Определяя падежные отношения, исследователь указывает на аккумулятив¹⁶ (‘желание/намерение [сделать]’) и приводит следующие примеры: *darśayitu-kāma* ‘обладающий желанием показать’, *draṣṭu-manas* ‘обладающий намерением увидеть’ и *draṣṭu-kāmaḥ sabhām*¹⁷ ‘обладающий желанием увидеть собрание’ [Wackernagel, 1905, p. 31, 44].

Как отмечает Я. Ваккернагель, подобные формы восходят ко времени, когда основа на *-tu* еще не была ограничена¹⁸, как это будет позднее в классическом санскрите, инфинитивом на *-(i)tum*. Однако форма на *-tu* в качестве первого члена сложного слова в ведийском, по-видимому, ограничивается единственным композитом *śrotu-rāti* ‘обладающий милостью слышания/слышать’ (4) [Wackernagel, 1905, p. 44]:

¹⁵ Бахуврихи (санскр. *bahuvrīhi*, букв. ‘[тот, у кого] много риса’) – термин древнеиндийской лингвистической традиции. Это сложное слово со значением обладания, которое указывает на признак, принадлежащий внешнему по отношению к самому композиту объекту (носителю этого свойства). Например, *mahodara* – букв. ‘обладающий большим животом’, ‘толстобрюхий’ < *mahā* ‘большой’ + *udara* ‘живот’. Ср. с англ. *redhead* ‘рыжеволосый’ < *red* ‘красный’ + *head* ‘голова’; нем. *dickkopf* ‘[тот, у кого] «толстая» голова’, ‘упрямец’ < *dick* ‘толстый’ + *kopf* ‘голова’.

¹⁶ Дискуссию о падежном управлении см. в работе [Keydana, 2013, p. 49]. Г. Кейдана оспаривает тезис Лероя о том, что глагольное управление у слова *kāma* является древним индоиранским наследием. Согласно исследователю, имеющиеся тому свидетельства из авестийского оказываются несостоятельными, а глагольное управление в древнеиндийских композитах на *-kāma* предстает как собственная специфическая черта индоарийского, а не общее индоиранское наследие.

¹⁷ Я. Ваккернагель не отмечает это, но данный пример, почерпнутый из Махабхараты (II.34.4a), особенно примечателен тем, что инфинитив *draṣṭu(m)* также имеет при себе прямое дополнение – *sabhā-m* ‘собрание’ (естественно, маркированное падежно-числовым маркером аккумулятива *-m*). Это свидетельствует в пользу трактования форм на *-tu* в образованиях с *kāma-/manas-* как инфинитивов, а не имен действия, поскольку последние в подобном случае требовали бы иного падежного управления – генитивного (т.е. **draṣṭu-kāmaḥ sabhāyās* ‘желание видения собрания’).

¹⁸ Основа на *-tu* в ведийском используется для образования абстрактных имен действия (например, *gan-tu-* ‘движение’, *śru-tu-* ‘слушание’). Кроме того, как уже было отмечено ранее, от основы на *-tu* порождались инфинитивы на *-tave*.

(4) <i>śro-tu</i>	<i>naḥ</i>
слышать.IMP-3SG.АКТ	МЫ.ACC.PL
<i>śrotu+rāti-ḥ</i> <...>	<i>sindhu-r</i> <...>
слышание+милость-NOM.F.SG	Синдху-NOM.SG
‘Пусть услышит нас Синдху, [наделенная] милостью	
слышания ¹⁹ ...!’	
(Ṛv: I-122.6.2)	

Все перечисленные исследователи вслед за древнеиндийской лингвистической традицией определили особенности строения форм вида «инфинитив + *kāma-/manas-*», связанные с выпадением конечного *-m* в суффиксе инфинитива, но далее этого определения не продвинулись.

При этом представляется необходимым рассмотрение этих специфичных и типологически редких квази-конструкций в рамках проблематики сложных слов (т.н. *samāsa*).

Употребить термин «аналитическая конструкция» по отношению к рассматриваемому сочетанию нельзя. Данное образование можно отнести к подклассу сложных слов «*tatpuruṣa*», первый элемент которого (в данном случае инфинитив) находится в подчинительной связи по отношению к конечному компоненту (существительному *kāma-/manas-*). На основании же *tatpuruṣa* имеет место бахуврихи, в результате чего форма приобретает адъективное значение ‘обладающий желанием [сделать]’, ‘желающий [сделать]’.

Примечательно, что во всех проанализированных нами источниках, содержащих подробные сведения о каждом из классов сложных слов, не была зафиксирована возможность вхождения в качестве первого члена *tatpuruṣa* инфинитива на *-(i)tum* (ни в формулировке «инфинитив», ни, что было бы еще более ожидаемым, как «отглагольное имя»). Это, в свою очередь, вновь указывает на особую специфичность рассматриваемых образований с *kāma-/manas-* [Кнауэр, 1908, с. 92–96; Елизаренкова, 1982, с. 175–189; Кочергина, 1990, с. 6 и сл.; Goldman, 1999, p. 207–250; Scharf, 2003, p. 47–58; Deshpande, 2007, p. 261–279; Ruppel, 2017, p. 137–149]²⁰.

¹⁹ Или ‘желающая одарить слышанием’: о возможностях перевода и специфике функционирования слова *rāti* см. [MacDonell, 1916, p. 257].

²⁰ Классификация сложных слов была разработана еще для дренеиндоарийских языков, однако вполне может быть применима и для среднеиндоарийских, в связи с чем хотелось бы несколько поправить объяснение, предложенное А.В. Парибком в монографии «Система палийского глагола». Он обозначает данное образование как «причастие желания» или «дезидеративное причастие». Исследователь рассматривает *-tukāma-* как присоединяемый к основе смыслового глагола единый аффикс [Парибок, 2022, с. 16, 25], в то время как, очевидно, аффиксом (и тем более аффиксом причастия) это не является. Также не следует полагать, что сочетание инфинитива на *-(i)tu-* с *-kāma-* имеет грамму

Возможность членения суффикса *-(i)tum* на *-tu* и конечный *-m* отсылает нас к истории происхождения древнеиндийских инфинитивов на *-(i)tum*, которые в ведийском представляли собой отглагольные имена с суффиксом *-tu*, в аккузативе маркированные падежным окончанием *-m*. Однако причина, по которой в формах с *kāma-/manas-* инфинитив в качестве первого компонента предстает в виде основы, обуславливается общим механизмом образования сложных слов независимо от их типа²¹. Так, в древнеиндийских и среднеиндийских композитах только последнее из сочлененных слов может иметь при себе падежно-числовой маркер, в то время как все предыдущие звенья этой «цепи» предстают исключительно в виде основы. Например, ср. со сложным словом *svarga-kāma*²² ‘обладающий желанием [достигнуть] небес’ – композитом, также относящимся к типу «бахуврихи» на основе «татпуруша», но первая часть которого представлена не инфинитивом или отглагольным именем, а существительным (*svarga* ‘небо’).

Таким образом, в формах вида «инфинитив на *-(i)tum* + существительное *kāma-* ‘желание’ или *manas-* ‘мысль’» имеет место некоторая деграмматикализация инфинитива, возвращающего себе статус отглагольного имени на *-tu*.

Отдельно отметим, что сложные слова с *kāma-* могли «усложняться» ещё больше. Так, на материале текстов древнеиндоарийских и среднеиндоарийских языков нами были обнаружены редкие именные формы с историческим суффиксом абстрактных существительных *-tā*, надстраивающиеся над рассматриваемыми нами сложными словами и указывающие уже не на обладателя желания что-либо сделать, а на само это желание. Например, *pātu-kāma-tā* ‘желание пить’ (но не ‘обладающий желанием пить’) в санскрите (5) и *kattu-kāma-tā* ‘желание делать’ (но не ‘обладающий желанием делать’) (пали).

несовершенного вида, как и собственно вида как такового [Парибок, 2022, с. 28]. Безусловно, будучи не причастием, а сложным словом, функционирующим в высказывании как атрибутив, оно не может имплицировать глагольную категорию вида. Например, *passitu-kāmā* может иметь оба значения: и ‘обладающий желанием видеть’, и ‘обладающий желанием увидеть’ (ср. [Там же, с. 127]. При этом, если несколько абстрагироваться от диахронии и рассматривать подобное образование на уровне синхронии раннего среднедревнеарийского, когда уже произошла грамматикализация и перераспределение форм, не исключено, что в пали *tukāma* можно рассматривать как сложный суффикс со значением ‘желающий (с)делать X’. Подробнее о данных формах в среднедревнеарийских языках см. раздел 3.

²¹ Подробнее о разных типах сложных слов в древнеиндоарийских языках («татпуруша», «кармадхарая», «двандва» и др.) см. [Елизаренкова, 1982].

²² В данном случае первому компоненту композита *svarga* ‘небо’ соответствует генитив: ‘неба желание’. Аналитическим аналогом этого выражения было бы словосочетание *svarga-sya kāma-*, где *-sya* является окончанием генитива единственного числа для основы на *-a-*. В составе сложного слова это окончание закономерно опускается, а слово *svarga-* предстает в виде основы.

(5) ...*grīṣmābhīṭaptāḥ* (<= *grīṣma*+*abhīṭaptāḥ*)

зной_мучить.PAST.PF.PART:NOM.M.PL

pā-tu+kāmatayā

pradhāva-nti... |

пить-INF+желание.DAT.SG

бежать.PRS-3PL.ACT

‘[Они,] истерзанные зноем, бегут с **желанием испить** [воды]’

(LañS: 2.153.3)

3. Эволюция от древне- к среднеиндийскому

Завершение древнеиндийской и начало среднеиндийской стадии индоарийского языкового состояния, обнаруживаемого во «вторичных пракритах», относят приблизительно к V–IV вв. до н.э [Вертоградова, 2002, с. 6]²³.

Пали является классическим стандартизованным языком раннесреднеиндийского периода, в первую очередь отражающим культуру южноиндийской буддийской словесности [Geiger, 2005, р. XXII].

С лингвистической точки зрения пали можно рассматривать как промежуточное звено между поздним древнеиндоарийским и литературными пракритами среднеиндоарийских языков [Елизаренкова, Топоров, 2003, р. 7–8, 10].

Наиболее часто встречающимся суффиксом инфинитива в пали является наследуемый из древнеиндоарийского *-(i)tum* (исторически – показатель застывшего в аккумулятиве отглагольного имени), являвшийся также наиболее продуктивным способом порождения инфинитивов в классическом санскрите [Duroiselle, 1997, р. 109; Geiger, 2005, р. 191].

Ср. (6)–(7), иллюстрирующие употребление инфинитива в составе сложного слова со значением «обладающий желанием сразиться» *yoddhu-kāma-* (санск.) и *ujjjh-itu-kāma-* (пали). Первое выражение образовано от корня *yudh-* «сражаться», а второе – от порожденной от этого корня палийской основы *ujjjha-*. При этом (6) демонстрирует атрибутивное функционирование, а (7) – употребление в составе сложного именного сказуемого (с опущенной глагольной связкой ‘*быть*’).

²³ Термин «*ṛākṛta*» ‘естественный/природный’ противопоставлен термину «*saṃskṛta*» ‘очищенный/искусственный’, используемого для обозначения санскрита. Определение «вторичные» по отношению к пракритам среднеиндоарийских языков возникло в сопоставлении с «первичными пракритами» – живыми разговорными диалектами, видимо, также существовавшими в ведийскую эпоху.

основе [Jain, 1923, p. XXXIII (j); Vaidya, 1944, p. 53; Ghatge, 1993, p. 137]. По сравнению с инфинитивами на *-(i)ttae* формы на *-(i)um* в ардхамагадхи встречаются достаточно редко, а сфера их дистрибуции ограничена преимущественно поэтическими текстами [Pischel, 1967, p. 394].

В отношении форм, к которым диахронически восходят инфинитивы на *-(i)um*, Р. Пишель упоминает древнеиндийские инфинитивы на *-tum*, исторически представляющие собой имена на *-tu-* с застывшим показателем винительного падежа *-m* [Ibid., p. 393]. Ср., например, два инфинитива *kar-tum* и *kā-um*, образованные от одного и того же корня *kr-* ‘делать’ в древнеиндийском и ардхамагадхи соответственно.

В ходе работы мы не нашли в имевшихся в нашем распоряжении текстах на ардхамагадхи примеров использования указанных сложных слов, однако не вызывает сомнений сам факт их существования в этом языке. Например, *gīṅh-īu-kāma-* ‘обладающий желанием схватить’: *grh-* (др.-индоар.) > *grh-ṇa-* (о.н.в.) > *gīṅh-a-* (о.н.в.) (ср.-индоар.) > *gīṅh-īum* (ср. с повсеместно встречающейся в палийских джатаках форме *gīṅh-ītu-kāma-*) [Pischel, 1957, p. 395; Geiger, 2005, p. 193; AMD, 1927, 1930, 1932]²⁴.

При этом нам удалось найти пример атрибутивного употребления такого рода слова в еще одном языке средней фазы развития среднеиндийского – джайнском махараштри (языке неканонической литературы джайнской секты шветамбаров) [Pischel, 1967, p. 20]. См. фрагмент из текста «Странствия Васудэвы» V–IV вв. (10). Ср. с *tāḍīu-maṇā* ‘обладающий намерением избить’ [Ibid., p. 395].

(10) <...> *jammana+maṇa+bahula-m*²⁵

рождение+смерть+многий-ACC.M.SG

saṃsāra-m ***chid-īu+kāmā***

сансара-ACC.SG отсекать-INF+желание.NOM.SG

tato *ku-dhamma-m* *parica-ittā...*

затем дурная_дхарма-ACC.SG оставлять-AVS

‘[...те, кто] **хотят** (букв. ‘**обладают желанием**’) **пресечь** сансару, изобилующую [чередой] рождений [и] смертей’, отвергнув дурное воззрение,...’

(VH, p. 268)

²⁴ Среди языков, где засвидетельствованы подобные формы, помимо ардхамагадхи и джайнского махараштри, Р. Пишель также упоминает сценический пракрит шаураसेи (*śaurasēṇī*), приводя в качестве примеров следующие формы: *jīvidu-kāma* ‘обладающий желанием жить’, *tāḍīu-maṇā* ‘обладающий намерением ударить’ [Pischel, 1957, p. 395].

²⁵ Поскольку джайнский махараштри и ардхамагадхи имеют много общего, то ввиду отсутствия доступного специализированного словаря для джайнского махараштри для перевода наиболее трудных пракритских лексем мы обращались к ардхамагадхи-английскому словарю (с пояснениями на санскрите, гуджарати и хинди) [AMD, 1927, p. 838; 1930, p. 679; 1932, p. 407, 427, 497].

В текстах не более поздних, чем джайнский махараштри, праkritах анализируемые образования не были обнаружены. По-видимому, к началу формирования новоиндийских языков рассмотренная модель уже полностью перестала использоваться носителями.

Выводы

Проанализированный материал позволяет рассмотреть важнейшие характерные особенности структуры и эволюции образований вида «инфинитив на *-(i)tum* + существительное *kāta-* ‘желание’ / *manas-* ‘мысль’», а также определить их место в морфологической классификации.

1. Примечателен процесс частичной грамматикализации, произошедшей в отношении этих специфичных и редких сложных слов. Он заключается в следующем:

- особое функционирование слов *kāta-* и *manas-*, которые приобретают, по сути, характер квазисуффиксов;
- ограниченность выбора лексемы на второе («управляющее») место в сложном слове; т.е., другими словами, узость семо-синтаксической сочетаемости существительного и инфинитива; при этом сами лексемы с синонимичным значением ‘желание/намерение’ *kāta-* и *manas* находятся скорее в отношении свободного варьирования, чем в дополнительном распределении.

2. В терминах традиционной индийской классификации сложных слов подобные комплексы следует отнести к разновидности класса «таппуруша». Первая часть сложного слова (инфинитив) находится в подчинительной связи по отношению ко второму. На базе «таппуруша» путем конверсии выстраивается «бахуврихи», в результате чего сложное слово приобретает характерную для прилагательных функцию атрибутива «обладающий желанием/намерением сделать [что-либо]».

3. Была кратко прослежена вся история формирования и эволюции в индоарийском основосложения указанного типа: от раннего древнеиндоарийского (ведийского) через санскрит к раннему среднеиндоарийскому (пали) и «средним» среднеиндоарийским языкам (ардхамагадхи и джайнскому махараштри).

Библиографический список / References

Барроу, 1976 – Барроу Т. Санскрит / пер. с англ. Н. Лариной. М., 1976. [Burrow T. Sanskrit [The Sanskrit Language]. N. Larina (transl. from English). Moscow, 1976.]

Борковский, Кузнецов, 2006 – Борковский В.И., Кузнецов П.С. Историческая грамматика русского языка. изд. 3-е. М., 2006 [Borkovsky V.I., Kuznecov P.S.

Istoricheskaya grammatika russkogo yazyka [Historical grammar of the Russian language]. 3rd ed. Moscow, 2006.]

Вертоградова, 2002 – Вертоградова В.В. Праkritы. 2-е изд. М., 2002. [Vertogradova V.V. Prakrity [The Prakrits]. 2nd ed. Moscow, 2002.]

Горшкова, Хабургаев, 1981 – Горшкова К.В., Хабургаев Г.А. Историческая грамматика русского языка. М., 1981. [Gorshkova K.V., Haburgaev G.A. Istoricheskaya grammatika russkogo yazyka [Historical grammar of the Russian language]. Moscow, 1981.]

Елизаренкова, 1982 – Елизаренкова Т.Я. Грамматика ведийского языка. М., 1982. [Elizarenkova T.Ya. Grammatika vediyского yazyka [The Grammar of Vedic Language]. Moscow, 1982.]

Елизаренкова, Топоров, 2003 – Елизаренкова Т.Я., Топоров В.Н. Язык пали. 2-е изд. М., 2003. [Elizarenkova T.Ya., Toporov V.N. Yazyk pali [The Pali Language]. 2nd ed. Moscow, 2003.]

Изотов, 2010 – Изотов А.И. Старославянский язык в сравнительно-историческом освещении: учебное пособие. М., 2010. [Izotov A.I. Staroslavjanskiy yazyk v sravnitel'no-istoricheskom osveschenii [The Old Slavonic language in comparative historical coverage]. Textbook. Moscow, 2010.]

Кнауэр, 1908 – Кнауэр Ф.И. Учебник санскритского языка. Грамматика. Хрестоматия. Словарь. Лейпциг, 1908. [Knauer F.I. Uchebnik sanskritского yazyka. Grammatika. Khrestomatiya. Slovar [The Sanskrit Handbook. Grammar. Text. Dictionary]. Leipzig, 1908.]

Кочергина, 1990 – Кочергина В.А. Словообразование санскрита (префиксация и основослоложение). М., 1990. [Kohergina V.A. Slovoobrazovanie sanskrita (prefiksatsiya i osnovoslozhenie) [Derivation in Sanskrit (prefixes and compounds)]. Moscow, 1990.]

Парибок, 2022 – Парибок А.В. Система палийского глагола (глагольные формы и их значение в диахронии). М., 2022. [Paribok A.V. Sistema paliyского glagola (glagolnye formy i ikh znachenie v diakhronii) [Verbal System in Pali (verbal forms and their meaning in diachrony)]. Moscow, 2022.]

Тронский, 2001 – Тронский И.М. Историческая грамматика латинского языка. Общеиндоевропейское языковое состояние (вопросы реконструкции). 2-е изд. М., 2001. [Tronsky I.M. Istoricheskaya grammatika latinskogo yazyka. Obscheindoeuropeyskoe yazykovoe sostoyanie (voprosy rekonstruktsii) [Historical grammar of the Latin language. Common Indo-European language status (reconstruction issues)]. 2nd ed. Moscow, 2001.]

Шантрэн, 1953 – Шантрэн П. Историческая морфология греческого языка / пер. со 2-го фр. изд. Я.М. Боровского. М., 1953. [Chantraine P. Istoricheskaya morfologiya grecheskogo yazyka [Morphologie Historique Du Grec]. Ya.M. Borovskiy (transl. from French). Moscow, 1953.]

Эдельман, 1990 – Эдельман Д.И. Сравнительная грамматика восточно-иранских языков. Морфология. Элементы синтаксиса. М., 1990. [Edelman D.I. Sravnitel'naya grammatika vostochnoiranskikh yazykov. Morfologiya. Elementy sintaksisa [Comparative grammar of the Eastern Iranian languages. Morphology. Syntax elements]. Moscow, 1990.]

AMD, 1927 – An illustrated Ardha-Magadhi dictionary. Vol. II. London, 1927.

AMD, 1930 – An illustrated Ardha-Magadhi dictionary. Vol. III. London, 1930.

AMD, 1932 – An illustrated Ardha-Magadhi dictionary. Vol. IV. London, 1932.

Ashtadhyayi, 1897 – The Ashtadhyayi of Panini. Book VI. Translated into English by Srisa Chandra Vasu, B.A. Benares: published by Sindhu Charan Bose, at the Panini Office, 1897.

Burrow, 1965 – Burrow T. The Sanskrit language. 2nd ed. London, 1965.

Deshpande, 1980 – Deshpande M.M. Evolution of syntactic theory in Sanskrit grammar: Syntax of the Sanskrit infinitive *-tumUN*. Ann Arbor, 1980.

Deshpande, 2007 – Deshpande M. Madhav. A Sanskrit primer. Michigan Papers on South and Southeast Asian Studies № 47. Centers for South and Southeast Asian Studies University of Michigan, 2007.

Disterheft, 1980 – Disterheft D. The syntactic development of the infinitive in Indo-European. Columbus, Ohio, 1980.

Duroiselle, 1997 – Duroiselle C. A practical grammar of the Pāli Language. 3rd ed. 1997 (1906).

Edgerton, 1953 – Edgerton F. Buddhist hybrid Sanskrit grammar and dictionary. Vol. I. Grammar. Vol. II. Dictionary. Yale University Press, 1953.

Fausbøll, 1879 – The Jātaka, Together with its Commentary, being Tales of the Anterior Births of Gotama Buddha, Vol. II. V. Fausbøll (ed.). London, 1879.

Fausbøll, 1883 –The Jātaka, Together with its Commentary, being Tales of the Anterior Births of Gotama Buddha, Vol. III. V. Fausbøll (ed.). London, 1883.

Fausbøll. 1891 – The Jātaka, Together with its Commentary, being Tales of the Anterior Births of Gotama Buddha, Vol. V. V. Fausbøll (ed.). London, 1891.

Geiger, 2005 – Geiger W. A Pali grammar. Translated into English by Batakrishna Ghosh. Revised and edited by K.R. Norman. Oxford, reprinted 2005 (1916).

Ghatge, 1993 – Ghatge A.M. Introduction to ArdhhaMagadhi. Pune, 1993 (1941).

Goldman et al., 1999 – Goldman P. et al. Devavāṇīpraveśikā. An introduction to the Sanskrit Language. 3rd ed. Berkeley, 1999.

Jain, 1923 – Jain B.D. Ardhha-Māgadhī reader. Lahore, 1923.

Kāśikā, 1969–1970 – Kāśikā. A commentary on Pāṇini's grammar. By Vāmana and Jayāditya. A. Sharma, Shri K. Deshpande (eds.). Part I–II. Hyderabad, 1969–1970.

Keydana, 2013 – Keydana G. Infinitive im Rgveda: Formen, Funktion, Diachronie. Brill's studies in Indo-European languages & linguistics; 9. Leiden, 2013.

MacDonell, 1916 – MacDonell A. A Vedic grammar for students. Oxford, 1916.

Oberlies, 2003 – Oberlies T. A grammar of Epic Sanskrit. Berlin, 2003.

Pischel, 1957 – Pischel R. Comparative grammar of the Prakrit Languages. S. Jha (transl. from the German). Motilal Banarsidas, 1957 (1900).

Roodbergen, 2008 – Roodbergen J.A.F. Dictionary of Pāṇinian Grammatical Terminology. Pune, 2008.

Ruppel, 2017 – Ruppel A.M. The Cambridge introduction to Sanskrit. Cambridge University Press, 2017.

Sarma, 1968 – Sarma K.M.K. Pāṇini Kātyāyana and Patanjali. Delhi, 1968.

Scharf, 2003 – Scharf P. Rāmopākhyāna – The story of Rāma in the Mahābhārata. An independent-study reader in Sanskrit. London, 2003.

Sgall, 1958 – Sgall P. Die Infinitive im Rgveda. Praha, 1958.

Sihler – Sihler A.L. New comparative grammar of Greek and Latin. Oxford University Press, 1995.

Speijer, 1998 – Speijer J.S. Sanskrit Syntax. Delhi, 1998 (1886).

Vaidya, 1944 – Vaidya P.L. A manual of Ardhhamagadhi grammar. 4rd ed. Poona, 1944.

VMBh, 1885 – The Vyākaraṇa-Mahābhāṣya of Patañjali. F. Kielhorn (ed.). Vol. III. Bombay, 1885.

Wackernagel, 1905 – Wackernagel J. Altindische Grammatik. Göttingen, 1905.

Whitney, 1879 – Whitney W.D. A Sanskrit grammar, including both the classical language, and the older dialects, of veda and brahmana. Leipzig; London, 1879.

Статья поступила в редакцию 20.02.2025

The article was received 20.02.2025

Сведения об авторе / About the author

Камаданова София Сергеевна – аспирант кафедры индийской филологии института стран Азии и Африки, Московский государственный университет имени М.В. Ломоносова; старший преподаватель кафедры индийской филологии Института стран Азии и Африки, Московский государственный университет имени М.В. Ломоносова

Sofia S. Kamadanova – PhD student at the Department of Indian Philology of the Institute of Asian and African Studies, Lomonosov Moscow State University; senior lecturer at the Department of Indian Philology of the Institute of Asian and African Studies, Lomonosov Moscow State University

ORCID: <https://orcid.org/0009-0005-9854-9230>

E-mail: kamadanovass@my.msu.ru

P.M. Cotta Ramusino

Milan State University,
Milan, 20099, Italy

Binominal constructions of small quantity in Russian and Italian

This paper offers a contrastive analysis of binominal constructions expressing small quantity in Italian and Russian, focusing on six core nouns (*goccia/kaplja*, *briciola/kroška*, *granello/krupica*). Through a combined quantitative and qualitative corpus-based approach, the study investigates how these nouns evolve from literal measure expressions to metaphorical and grammaticalized quantifiers or degree modifiers. Drawing on theoretical frameworks from grammaticalization theory and usage-based Construction Grammar, the analysis highlights the role of cognitive profiling, metaphorical shift, and negative polarity contexts in shaping these constructions. Findings reveal significant cross-linguistic asymmetries. In Russian, metaphorical uses are tightly linked to negative polarity and emotional intensification, especially in singular forms. Italian, by contrast, shows a broader functional diversification, including the emergence of masculine forms (*goccio* and *briciolo*) that have undergone structural decategorization, losing their morphological features of number, and acquired abstract quantificational or scalar functions. Particular attention is paid to the semantic features of N_1 elements, whose combination with N_2 's shapes metaphorical reinterpretation, often grounded in domains such as emotions (*briciola/briciolo*) or epistemic knowledge (*krupica*). The study demonstrates that while Italian and Russian share typological similarities in binominal quantification, their metaphorical paths and grammaticalization processes diverge markedly. These findings provide insight into the interplay of lexis, grammar, and discourse, and offer a framework for future diachronic and typological investigations of quantifier constructions across languages.



Key words: contrastive linguistics, binominal constructions, quantifier constructions, degree constructions, grammaticalization

Acknowledgements. I would like to express my sincere gratitude to the anonymous reviewers for their invaluable feedback and constructive comments on this manuscript. Their insightful suggestions and detailed critiques greatly enhanced the quality of this work. The time and effort they dedicated to reviewing this paper are deeply appreciated, and their contributions have been instrumental in refining and strengthening the final version of the manuscript. Any remaining errors or omissions are solely my responsibility.

FOR CITATION: Cotta M. Ramusino P. Binominal constructions of small quantity in Russian and Italian. *Rhema*. 2025. No. 4. Pp. 27–52. DOI: 10.31862/2500-2953-2025-4-27-52

DOI: 10.31862/2500-2953-2025-4-27-52

П.М. Котта Рамузино

Миланский государственный университет,
20099 г. Милан, Италия

Биноминальные конструкции с указанием на малое количество в русском и итальянском языках

На фоне исследований о грамматикализации биноминальных конструкций, обозначающих количество, в статье представлен сопоставительный анализ биноминальных конструкций, выражающих значение малого количества в русском и итальянском языках. Исследование сосредоточено на шести ключевых существительных (*goccia/капля*, *briciola/крошка*, *granello/крупница*) и основано на сочетании количественного и качественного анализа данных. В рамках теорий грамматикализации и грамматики конструкций автор рассматривает, как данные существительные переходят от буквального обозначения меры к метафорическим и грамматикализованным функциям квантификаторов или модификаторов степени. Результаты выявляют значительные межязыковые асимметрии. Анализ показывает, что при внешнем типологическом сходстве биноминальные конструкции

мало количества в русском и итальянском языках демонстрируют различные траектории метафоризации и грамматикализации. Работа вносит вклад в изучение взаимодействия лексики, грамматики и прагматики, а также открывает перспективы для дальнейших типологических исследований количественных конструкций.

Ключевые слова: контрастивная лингвистика, биномиальные конструкции, выражающие интенсивность признака квантификационные конструкции, грамматикализация

Благодарности. Я выражаю искреннюю благодарность анонимным рецензентам за их бесценные отзывы и конструктивные комментарии к этой рукописи. Их проницательные предложения и подробные критические замечания значительно повысили качество этой работы. Время и усилия, которые они посвятили рецензированию этой статьи, высоко оцениваются мною, их вклад сыграл решающую роль в доработке и улучшении окончательной версии рукописи. Все оставшиеся ошибки или упущения являются исключительно моей ответственностью.

ДЛЯ ЦИТИРОВАНИЯ: Котта М. Рамузино П. Биноминальные конструкции с указанием на малое количество в русском и итальянском языках // Рема. Rhema. 2025. № 4. С. 27–52. DOI: 10.31862/2500-2953-2025-4-27-52

1. Introduction

In recent decades, scholars have shown increasing interest in binominal constructions of the type $[N_1]$ of $[N_2]$ that serve a quantifying function, such as the English *a pile of commonplaces*, *loads of furniture*, or *a bunch of grapes* [Brems, 2003]. In their prototypical form, these constructions involve head nouns that do not explicitly denote units of measurement – such as *meter*, *acre*, or *ounce* – but instead refer to spatially continuous entities that either contain a portion of mass (e.g., *bucket of water*) or consist of it (e.g., *bunch of flowers*) [Ibid., p. 88]. According to R.W. Langacker, these constructions represent an emerging means of quantification and “constitute a diverse and open-ended class” [Ibid.]. In addition to their literal meaning, they have also developed figurative uses. This semantic extension is made possible by “a conception of their typical size, which is part of their encyclopedic characterization” [Langacker, 1991, p. 88]. Owing to their rich semantic content and cognitive associations, these constructions possess significant pragmatic potential. They have been examined within various theoretical frameworks, most notably grammaticalization theory and Construction and Cognitive Grammar. E.C. Traugott, for instance, has explored their diachronic development in English using a constructional approach, according to which changes in form and meaning take place simultaneously [Traugott, 2008, 2015]. S.Ch. Li and

E. Rachilina have investigated Russian binominal constructions – particularly those expressing large quantities – highlighting their grammaticalization process from a cognitive perspective [Li, 2005; Rachilina, Li, 2009]. L. Brems (2003, 2007, 2010) has provided an analysis of measure nouns, including those of small quantity, as quantifiers in English and Dutch, raising important questions about the grammaticalization of low-frequency constructions and the status of the measure noun in these constructions. The degree of grammaticalization of Italian measure constructions has been addressed more recently by A. Giacalone Ramat (2019). P.M. Cotta Ramusino et al. (2025) adopt a similar perspective to compare binomials denoting small quantities in three Slavic languages: Belarusian, Russian, and Ukrainian. Furthermore, binominal constructions of this type have also been noted to function as classifier-like elements and sometimes exhibit actional or verb-like properties [Benigni, 2022; Benigni et al., 2023], thus potentially representing a link between nominal and verbal categories.

Although binominal constructions have been extensively studied, their small-quantity subtypes have attracted much less scholarly interest. Based on rich empirical data, the present study aims to explore further the theoretical issues these constructions pose and to contribute to a better understanding of their structure and function. It focuses specifically on binominal constructions of the type $[N_1]$ of $[N_2]$ – and, in Russian, $[N_1 + N_{2GEN}]$ – as illustrated in examples (1) and (2), in Russian and Italian, respectively.

(1) Ru

В ней нет ни *капли* вины и ни *крошки* страха
(Timestamped JSI webcorpus 2014–2021 Russian, Sketch Engine)
'There isn't a drop of guilt or a shred of fear in her.'

(2) It.

<...> Ci sono subito due *briciole* di malinconia.
(Italian Trends 2014 – today, Sketch Engine)
'Here are immediately two crumbs of melancholy.'

To account for the patterns exemplified in (1)–(2), the analysis is conducted within the broad framework of a usage-based Construction Grammar approach. Following E.C. Traugott, these constructions are treated as individual micro-constructions, related to meso- and even macro-constructions expressing quantity [Traugott, 2008, p. 32] (e.g., in Russian, constructions with numerals followed by the genitive case). Given the synchronic focus of this study, it is possible to draw on the concepts of inheritance links and motivation, also understood in the context of quantifier constructions.

From a synchronic perspective, these constructions may be also interpreted as representing one stage in a broader grammaticalization pathway from quantifier phrases to degree modifiers.

The concepts of *base* and *profile* from Cognitive Grammar [Croft et al., 2004; Langacker, 2013] are also essential for understanding how different speakers cognitively represent and construct novel associations in binominal constructions of the form $[N_1]$ of $[N_2]$, as they not only express the small quantity but also the “quality” of the quantified objects. While the association between *drop* and a liquid substance is readily interpretable, the cognitive basis for its use with abstract nouns or solid substances requires further exploration.

In addition to their semantic and cognitive features, the pragmatic distribution of small-quantity binominals – particularly their frequent occurrence in negative polarity contexts – deserves attention. As thoroughly discussed by L. Brems (2007), sensitivity to negative polarity may account for the grammaticalization of these constructions even in the absence of high-frequency usage. Negative polarity contexts also activate two distinct pragmatic processes: they minimize the quantified reading and intensify the expression of minimal quantity [Brems, 2007; see also Giacalone Ramat, 2019].

Ultimately, there is a clear rationale for comparing two languages from the same linguistic family, yet belonging to different subgroups. Cross-linguistic comparison can shed light on processes and changes that may occur even in unrelated languages, as such phenomena are intrinsically tied to underlying cognitive and communicative mechanisms [Traugott, 2015].

The paper is structured as follows: Section 2 presents the data and methodology; Section 3 analyzes the data based on the theoretical framework outlined above; and Section 4 discusses the findings and draws some conclusions.

2. Data and methodology

To operationalize these theoretical assumptions, the present study examines binominal constructions of the type $[N_1]$ of $[N_2]$ (and in Russian, $[N_1 + N_{2\text{GEN}}]$). Generally, these constructions may feature different types of N_1 :

i) a concrete or abstract noun with no inherent size component, such as ru. *gora*, it. *montagna*, ‘mountain’, ru. *more*, it. *mare*, ‘sea’, ru. *reka*, it. *fiume*, ‘river’ and similar. These nouns undergo a semantic and metaphorical shift to function as metaphorical quantifiers and degree modifiers, as in ru. *more ljudej*, it. *un mare di gente*, ‘a sea of people’;

ii) a concrete noun that also functions as a measure noun, such as ru. *mešok*, it. *sacco*, ‘sack’, ru. *kaplja*, it. *goccia*, ‘drop’, which typically appears in literal quantifying contexts (e.g. ru. *mešok muki*, it. *sacco*

di farina, ‘sack of flour’). When used with non-prototypical N_2 collocates, however, they acquire a metaphorical, intensifying meaning, as illustrated in examples (1) and (2).

The binomials examined here belong to type (ii). They are size nouns that, in certain contexts and with specific collocates, undergo a metaphorical shift and a pragmaticalization process that transform them from measure nouns into metaphorical quantifiers and degree modifiers.

For each language, three nouns were selected from a larger set¹, beginning with the Russian data and identifying their equivalents using the Russian–Italian parallel corpus of the Russian National Corpus:

Ru. *kaplja, kroška, krupica*

It. *goccia, briciola, granello*

Eng. ‘drop’, ‘crumb’, ‘grain’

Two of the nouns – *drop* and *grain* – can be classified as strictly (small) size nouns, whereas *crumb*, although referring to a very small part of a dry substance, cannot *strictu sensu* be considered a size noun. The initial selection was based on the distinct semantic and cognitive characteristics of these three nouns. While the feature SMALL is shared by all three, each noun also possesses additional semantic properties that, it is argued, contribute to the metaphorical processes. *Drop* represents the minimal measurable unit of liquid substances and is thus associated with the semantic features FLUID and COHESIVE. *Crumb*, on the other hand, is characterized by being FORMLESS and FRAGMENTARY; it is a remainder of something larger, thus suggesting incompleteness, scarcity, and at the same time intrinsic value – especially when associated with bread. Finally, *grain* denotes the fundamental unit of certain cereals and is conceptually relevant as something small that can be gathered to form a larger quantity, thus carrying the features COMPLETE PART OF A WHOLE and VALUABLE. Although the numerous diminutive forms of these nouns in both languages fall outside the scope of this study, some attention will be devoted to two Italian nouns, *briciolo* and *goccio*² for the insight they can give into the grammaticalization process.

2.1. Dataset

The present study draws on data from two corpora available on Sketch Engine, which are comparable in both size and content:

¹ K.A. Bordjugova lists the following nouns: *edinicy* ‘a few’, *kaplja* ‘drop’, *kapel'ka* ‘tiny drop’, *kroška* ‘crumb’, *krochi* ‘crumbs’, *kučka* ‘small heap’, *gorstka* ‘tiny handful’, *iskra* ‘spark’, *krupica* ‘grain’ [Bordjugova, 2015, p. 42].

² *Briciolo* and *goccio* are the masculine forms of *briciola* and *goccia*, respectively, marked by the ending -o. While their literal meanings are the same, they are used in different contexts, as will be shown.

- Italian Trends Corpus (2014 – today) (10,031,703,651 tokens);
- Timestamped JSI Russian Web Corpus (7,218,536,772 tokens).

Both corpora consist of news articles and texts from sources that are updated on a daily basis. The data were retrieved using the Corpus Query Language (CQL).

The data have been analyzed using both quantitative and qualitative approaches. The quantitative analysis employed frequency and concordance tools applied to the entire datasets, while the qualitative analysis was performed on a random sample of 300 occurrences for each word. To ensure the validity of the chosen sample, the convergence and stabilization of empirical frequency was examined, as discussed in Section 3. A summary table will be provided for each word, presenting the overall results.

The qualitative analysis considers both the N_2 collocates and their contextual environments, as previous studies have shown that factors such as negative polarity sensitivity are closely associated with the use of these binominal constructions [Brems, 2007; Giacalone, 2019]. The qualitative analysis was also essential to distinguish metaphorical from literal uses of the nouns. Because the metaphorical shift may function differently for singular and plural forms of the nouns³, the queries were carried out separately for both forms. The analysis further considers the use of in/definite articles and adverbs that function as modifiers or accompany the constructions.

Considering that, as already mentioned, all three nouns represent size nouns, in the analysis the literal reading will be distinguished from the metaphorical, while the quantified reading will be considered to be on a continuum with the literal one.

3. Data analysis and discussion

3.1.1. *Kaplja* ‘drop’: Russian data

The noun *kaplja* was searched in the Russian corpus using the CQL tool⁴. The query was designed to identify the noun in both singular and plural

³ In Italian, for instance, the plural form of size nouns of big quantity does not support the metaphorical shift [see Giacalone Ramat, 2019, p. 3].

⁴ The search has been carried out using the following search strings: 1. [lemma=“капля” & number=“s”] [tag=“N...g.”] – to retrieve all singular forms of the lemma *капля* followed by a genitive singular noun.; 2. [word=“ни” | word=“не”] []? [word=“капли”] [tag=“N...g.”] – to search for instances where *капли* follows *ни* or *не*, possibly separated by any token (see also note 3 to Table 1); 3. [word=“без”] []? [word=“капли”] [tag=“N...g.”] – to retrieve occurrences of *капли* within a variable distance from the preposition *без* (‘without’); 4. [word=“лишь”] []? [lemma=“капля” & number=“s”] [tag=“N...g.”] – to identify instances in which the singular form of *капля* is preceded by *лишь*. A parallel search was conducted substituting *лишь* (‘only’) with the adverb *даже* (‘even’); 5. Finally, plural forms were queried using [lemma=“капля” & number=“p”] [tag=“N...g.”], followed by specific searches for each case as outlined above.

forms, as well as its most frequent occurrences in negative polarity contexts. The results are summarized in Table 1.

Table 1
Occurrences and frequency of the singular and plural forms of *kaplja*

	Occurrences	Ipm ¹	Negative polarity ²		
			<i>ne/ni</i>	<i>bez</i>	<i>liš/daže</i>
<i>Kaplja</i> (sg.)	10 377	1.44	2612 ³	923	94/90
<i>Kapli</i> (pl.) ⁴	13 098	1.81	22 000	10	52/103

¹ Number of hits per million tokens.

² The query has been aimed at identifying the contexts in which only *ne* or *ni* occur at a variable distance from the genitive forms of the word. This type of query catches also other possible modifiers, such as *odna* ('one') *edinaja* ('single')

³ Regarding the negations *ne* and *ni*, the data shows that the search for *ne* yields only 74 occurrences, while the double negation (*ne/ni* + *kapli*) appears 2612 times. Given this significant difference, only the results for the double negation are presented here.

⁴ The homonymy between the genitive singular of a feminine noun and its nominative plural form is a well-known issue in corpus queries for Russian language. To estimate the extent of this ambiguity, a sample of 300 occurrences, on the basis below discussed of the empirical frequency, was manually analyzed: 204 were plural forms (across various cases), 62 were genitive singular forms (mostly preceded by a negation), and 34 were genitive singular occurrences found in the idiom *dve kapli vody* ('two drops of water') or following the numerals 2, 3, or 4, with a percentage of 68% of real plural forms. If the 34 cases where the genitive follows numerals – interpretable as semantic plurals – are included, the overall percentage rises to 79.3%.

The words *ne*, *ni*, *bez*, *liš*, and *daže* have been identified as the most frequent triggers of negative polarity contexts, though they are certainly not the only ones. Nevertheless, the results can be considered reliable within a reasonable margin of approximation. The number of plural forms appears to be higher than that of singular forms, although the homonymy between the genitive singular and nominative plural forms must be taken into account. Overall, these discrepancies do not significantly affect the results, as the impact of homonymy on the automatic retrieval of negative polarity contexts is minimal. As shown in Table 1, the most notable difference concerns polarity sensitivity, which is markedly higher in the singular form than in the plural, where it is almost negligible⁵.

⁵ In discussing the sensitivity of SSNs to negative polarity (NP) contexts, L. Brems (2007) draws on M. Israel's work (1996; 2004), who mentions the role of "weakly negative form(s)" [Israel, 2004, p. 710]. The dataset presented in L. Brems (2007?) is significantly smaller than the one discussed in the present article. The role of NP contexts should therefore be revisited in a dedicated analysis to better evaluate their relationship with these binomials.

To assess the distribution of literal and metaphorical uses of the target word, we manually analyzed a random sample of 300 instances extracted from the larger dataset of 10 377 occurrences⁶. The chart in Figure 1 illustrates how the relative frequencies of both literal and metaphorical uses converge as the number of observations increases. The curve representing metaphorical usage stabilizes around 59%, while literal usage stabilizes around 41%. These values suggest a strong predominance of metaphorical usage in the data. The early fluctuations in both curves diminish after approximately 150 observations, indicating that the frequency estimates become increasingly stable and reliable. Such stabilization, a well-established criterion in empirical frequency methods, supports the representativeness of the sample and justifies the use of a reduced dataset for estimating overall trends in usage⁷.

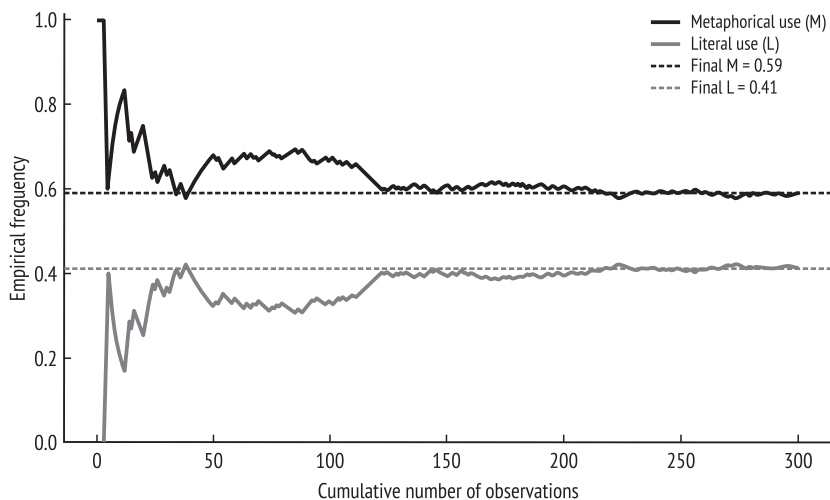


Fig. 1. Distribution of literal (41%) and metaphorical (59%) uses

The results obtained through the combination of quantitative and qualitative approaches yield several important insights. First, in Russian, the noun *kaplja*

⁶ To the best of our knowledge, the only reliable way to distinguish literal uses, such as “Достаточно капнуть туда каплю йода <...>” ‘It is enough to drop a single drop of iodine there’, from metaphorical ones (“И есть ли хоть капля смысла <...>?” ‘And is there even a drop of meaning?’), is to examine each occurrence in context. Although computational methods for metaphor detection exist, they are not yet sufficiently reliable for fine-grained corpus analysis of this kind. This manual screening was performed on 300 occurrences for each word, which, based on their empirical frequencies, can be considered sufficient to represent the entire dataset.

⁷ I would like to thank Luca Sala Gallini for his valuable suggestion to examine the empirical frequency and for his assistance in presenting the results in graphs.

‘drop’ is used metaphorically primarily in the singular form, while in the plural, the metaphorical shift appears to be marginal.⁸ Second, all occurrences in negative polarity contexts are metaphorical, as the manual examination of the 300 data has shown.⁹ These findings suggest, on the one hand, that the process of decategorization is nearly complete, as the new quantifier and degree meaning applies exclusively to the singular form; on the other hand, they confirm previous research – particularly L. Brems (2007) – which emphasizes the significant role of negative polarity contexts in facilitating metaphorical shift and, consequently, in processes of grammaticalization, whereby the expression comes to denote and emphasize a very small quantity. The second step of the analysis aims to identify the cognitive processes underlying the extension of usage to non-liquid and, subsequently, abstract collocates. To this end, the frequency tool was used to automatically extract collocates from the entire dataset, ranked by frequency. Table 2 provides an overview of the results.

Table 2

The type frequency of the first 15 collocates for *kaplja* ‘drop’

Collocate	Hits	Collocate	Hits
Воды ‘water’	2250	Сомнения ‘doubt’	103
Крови ‘blood’	1860	Жизни ‘life’	93
Масла ‘oil’	368	Совести ‘conscience’	93
Терпения ‘patience’	282	Никотина ‘nicotine’	89
Дождя ‘rain’	200	Стеснения ‘embarrassment’	56
Макияжа ‘maquillage’	157	Косметики ‘Cosmetics’	49
Пота ‘sweat’	137	Сожаления ‘regret’	34
Йода ‘Iodine’	128		

Apparently, the first three positions are well expected, as they are the prototypical liquid substances profiled by *drop*, however, at a closer look at the occurrences, some observations are due. First, both water and blood

⁸ This claim is based on the manual examination of the random sample of 300 occurrences. Out of the 204 genuine plural forms, 195 display a literal use, while only 9 show a metaphorical one.

⁹ More precisely, in the manual examination of 300 occurrences, 59 examples – belonging to the 59% of metaphorical uses – occurred in NP contexts, whereas in the 41% of literal uses, only 6 examples showed such a context. These findings confirm that NP contexts are more closely associated with metaphorical than with literal uses.

occur metaphorically, not just literally. One of the most frequent examples with water is the idiom ‘like two drops of water’ (3). Other occurrences, in negative polarity contexts, hint at signifying the complete absence of water. Likewise, a drop / some drops of blood occur in their literal and quantified meaning, but also very often with a metonymic and/or metaphorical shift, by which blood stands for conflict, war, death and similar (4).

(3) Герой этого фильма, как *две капли воды* похожий на Михаила Касьянова, <...>
‘The protagonist of this film looks exactly like Michail Kas’janov <...>’

(4) Они смогут войти сюда, только если всех нас убьют, мы будем драться *до последней капли крови*
‘They will only be able to get in here if they kill us all; we will fight to the *last drop of blood*’

In a similar way, *pot* ‘sweat’ and *makijaž* ‘maquillage’ are used predominantly in negative polarity contexts, and degree meaning prevails over the quantified. The abstract nouns in the first 15 positions may broadly be classified within the semantic field of emotions, in line with the conceptual metaphor EMOTIONS ARE LIQUIDS [Kövecses, 2000].

3.1.2. *Goccia* ‘drop’: Italian data

An identical query was performed for the Italian equivalent *goccia*, the results are shown in Table 3.

Table 3

Occurrences and frequency of the singular and plural forms of *goccia*

	Occurrences	Ipm	Negative polarity ¹
			<i>neanche, nemmeno, neppure</i>
<i>Goccia</i> (sg.)	14 541	1.33	528
<i>Gocce</i> ² (pl.)	15 381	1.47	4

¹ The query has been aimed at identifying the contexts in which only *neanche, nemmeno, neppure* ‘not even’ occur at a variable distance from the N₂. This type of query catch also other possible modifiers, such as *una* ‘one’, *unica* ‘unique’.

² The plural form of the feminine noun *goccia* is marked by the ending *-e*.

As shown in Table 3, the data unequivocally demonstrate that the discrepancy between the singular and plural forms is especially salient in negative polarity contexts while, overall, NP does not seem to be so relevant. In Table 4 the first 15 collocates of *goccia* are listed.

Table 4

The first 15 collocate types for *goccia* ‘drop’

Collocate	Hits	Collocate	Hits
Sangue ‘blood’	2677	Detersivo ‘laundry soap’	218
Pioggia ‘rain’	2008	Latte ‘milk’	204
Olio ‘oil’	1966	Energia ‘energy’	195
Sudore ‘sweat’	1120	Aria ‘air’	175
Limone ‘lemon’	618	Petrolio ‘petrol’	135
Succo ‘juice’	372	Carburante ‘fuel’	134
Acqua ‘water’	297	Sapone ‘soap’	131
Aceto ‘vinegar’	270		

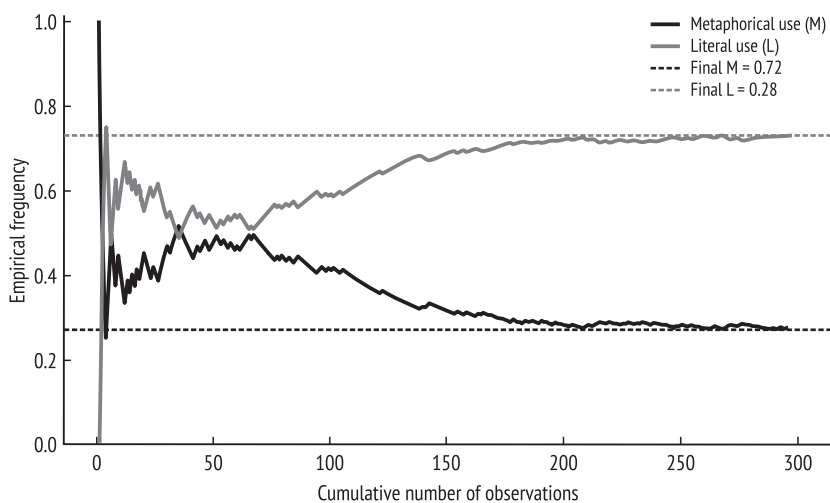


Fig. 2. Distribution of literal and metaphorical uses according to empirical frequency

The 15 most frequently occurring collocates (Table 5) reveal that no nouns with abstract meaning or referring to solid substances appear among the top entries, except for *energia*. A closer examination shows that non-liquid or abstract collocates are only marginally represented: *splendore*

‘splendor’ occurs 66 times, *speranza* ‘hope’ – 37, *vita* ‘life’ – 54, *talento* ‘talent’ – 24, *felicità* ‘happiness’ – 22, *luce* ‘light’ – 22, and *solidarietà* ‘solidarity’ – 20. Table 5, and Figure 2, based on the same principle as above, present the distribution of literal and metaphorical meanings. Interestingly, in the Italian language, *goccia* is predominantly used with its literal meaning also in the singular form. It should be noted that this use is nearly exclusive in the plural form. At the same time, most metaphorical uses occur in negative polarity contexts.

Table 5

**Distribution of literal and metaphorical uses
(a random sample of 300)**

	Literal, %	Metaphorical, %
<i>goccia</i>	72.0	28.0
<i>gocce</i>	97.6	2.3

3.1.3. *Goccio* ‘drop’: Italian data

The situation observed with the Italian *goccia* construction, which functions mainly with the literal-quantifying meaning, hints at one direction of development represented by the noun *goccio*, which appears to represent a further step in the grammaticalization process with its 2196 occurrences and an IPM of 0.21. Notably, according to a core principle of constructionalization theory [Traugott, Trousdale, 2013], which focuses on the emergence of new constructions – i.e., form–meaning pairings – during language change, both form and meaning have changed in this case: the masculine form *goccio* has replaced the feminine *goccia*, and its function has also shifted. Unlike *goccia*, which can be used literally to mean ‘a drop of’ or ‘some drops of’, *goccio* functions almost exclusively as a quantifier, denoting a small amount (5). The plural form *gocci*, although attested, is extremely rare (only three occurrences in the corpus). As already pointed out in relation to *briciolo* by A. Giacalone, “[t]he singular pattern of Italian may be interpreted as a sign of grammaticalization, because the head noun has lost the category feature of plurality” [Giacalone, 2019, p. 4]. Moreover, in line with the grammaticalization process, *goccio* can also function adverbially, occurring without any accompanying N_2 (6). These features point to an ongoing process of structural decategorization. The occurrence of *goccio* in negative polarity contexts is limited (approximately 5.3%), and its pragmatic function of intensification is less prominent. Its role as a quantifier is primarily evidenced by its frequent use in recipes. Use with metaphorical collocates is almost entirely limited to *hapax legomena* (7).

- (5) <...> quindi cuocetele sfumando con un *goccio* di vino bianco e poca acqua.
‘Then cook them, deglazing with a *splash* of white wine and a little water’.
- (6) Un film ben costruito, recitato e scritto, che ci pare un *goccio troppo* mainstream per la competizione veneziana.
‘A well-constructed, well-acted, and well-written film, which strikes us as just a *bit* too mainstream for the Venice competition’.
- (7) <...> che tampona con la tenerezza e un *goccio* di malinconia ogni piccola deriva, e non conosce spietatezza.
‘<...> who/which soothes every small drift with tenderness and a *hint* (drop) of melancholy, and knows no ruthlessness.’

3.2.1. *Kroška* ‘crumb’: Russian data

A crumb is the smallest part of a dry substance – a fragment of a whole – so the small quantity it represents is often perceived as *INSUFFICIENT*, *negligible*, but *VALUABLE*. The frequency in the corpus is summed up in Table 6.

Table 6

Occurrences and frequency of the singular and plural forms of *kroška*

	Occurrences	Ipm	Negative polarity		
			<i>ne, ni</i>	<i>bez</i>	<i>liš’/daže</i>
<i>Kroška</i> (sg.)	(617 ¹) 393	0.09	20	18	11
<i>Kroški</i> (pl.)	433	0.06	42	–	10

¹ In fact, upon excluding all instances in which *kroška* functions as an appositive (i.e., *kroška* followed by a proper noun), a total of 393 occurrences remain.

The only significant collocates from a frequency perspective, included NP contexts, are *bread* and, more generally, *food*. As shown in Table 7, most collocates are concrete nouns directly related to the literal meaning of *crumb*, while abstract nouns appear only as *hapax legomena*.

That said, it should be noted that in negative polarity contexts – and particularly in reference to *bread* and *food* – the use is metaphorical rather than literal, as illustrated in examples (8–9), conveying the complete absence of these substances and, by metonymy, of nourishment. This continued preference for combinations involving *bread* and this semantic field is likely due to the strong cultural and symbolic value attributed to bread.

Table 7

Most frequent collocate types for *kroška* ‘crumb’

Collocates (>10)	Hits	Collocates (>10)	Hits
Хлеба ‘Bread’	120	Other food	33
Печенья ‘Biscuits’	22	Stones and building materials	83
Еды ‘Food’	14	Abstract nouns	22

(8) Бывало и так, что в доме не было *ни крошки хлеба*, а дети донашивали друг за другом одежду.

‘There were times when there wasn’t *a single crumb of bread* in the house, and the children wore each other’s hand-me-down clothes’

(9) В холодильнике – *ни крошки еды*. В квартире ужасный бардак, тараканы, мухи.

‘There wasn’t *a single crumb of food* in the fridge. The apartment was a complete mess – cockroaches, flies, and filth everywhere’.

As far as the plural form is concerned, the collocates with a significant frequency are just a few: bread occurs 147, biscuits 42, food 33, and all the others, included materials, occur less than 10 times.

3.2.2. *Briciola* ‘crumb’: Italian data

In this case, the Italian data considerably differs from the Russian ones: not only is *briciola* used more frequently (see Table 8), especially in the plural, but it also co-occurs with a greater number of abstract nouns, as shown in Table 9 (where singular and plural forms are grouped together).

Table 8

Occurrences and frequency of the singular and plural forms of *briciola*

	Occurrences	Ipm	Negative polarity
			<i>non, neanche, nemmeno, neppure</i>
<i>Briciola</i> (sg.)	895	0.08	56 (0,1%)
<i>Briciole</i> ¹ (pl.)	1854	0.18	18

¹ As for *goccia* ‘drop’ and its plural *gocce*, the plural of the feminine noun *briciola* ‘crumb’ is formed with the ending *-e* (*briciole*).

Furthermore, as in Russian, even occurrences with food-related nouns – typically used literally – tend to take on a metaphorical meaning in negative polarity contexts, expressing the complete absence of something (10).

(10) Nessuno si degna di offrire a noi meridionali *nemmeno una briciola* di pane.

‘No one even deigns to offer us southerners so much as a *crumb* of bread.’

Table 9

Most frequent collocates types for *briciola* ‘crumb’

Collocates (>10) sg.	Hits	Collocates (>10) pl.	Hits
Pane ‘bread’	107	Pane ‘bread’	647
Potere ‘power’	38	Cibo ‘food’	48
Speranza ‘hope’	29	Biscotti ‘biscuits’	32
Energia ‘energy’	26	Pasta ‘pasta’	28
Umanità ‘humanity’	20	Potere ‘power’	27
Tempo ‘time’	17	Dignità ‘dignity’	19
Consenso ‘consensus’	17	Felicità ‘happiness’	18
Verità ‘truth’	15	Speranza ‘hope’	17
Dignità ‘dignity’	13	Impasto ‘dough’	17
Cibo ‘food’	13	Biscotto ‘biscuit’	17
Felicità ‘happiness’	12	Consenso ‘consensus’	17
Libertà ‘freedom’	11	Luce ‘light’	16
		Pan ‘bread’	15
		Vita ‘life’	14
		Panettone ‘panettone’	14

As observed, abstract nouns – although more commonly found in collocations with the singular form – also appear with the plural. These nouns do not belong to a single, well-defined semantic class, but they tend to cluster around the domain of feelings and emotions. Some refer specifically to emotional states, while others denote positively connoted concepts such as truth and humanity, or values like power and consensus – entities for which even a small amount (‘a crumb’) is perceived as highly valuable.

3.2.3. *Briciolo* ‘a bit’

As in the case of *goccio*, a comparable process can be observed with *briciolo*, the masculine (with ending *-o*) counterpart of *briciola*, which appears exclusively in the singular and occurs with notably high frequency.

Table 10

Occurrences and frequency of *briciolo* ‘a bit’

	Occurrences	Ipm	Negative polarity
			<i>non, neanche, nemmeno, neppure</i>
<i>Briciolo</i> (sg.)	9297	0.84	437 (4.7%)

In line with A. Giacalone (2019, p. 14), our data confirm that *briciolo* never denotes a small fragment of food and exclusively collocates with nouns expressing emotional or epistemic concepts (Table 11). *Briciolo* proves that a constructionalization process has occurred, as both gender and number have shifted – along with the meaning. The data also suggest that negative polarity contexts are relatively infrequent, which may indicate that, once the grammaticalization process is complete, the factors that initially facilitated it are no longer necessary.

Table 11

The first 15 collocates types for *briciolo* ‘a bit’

Collocates	Hits	Collocates	Hits
Speranza ‘hope’	514	Normalità ‘ordinariness’/‘normality’	125
Dignità ‘dignity’	474	Fiducia ‘trust’	104
Umanità ‘humanity’	376	Attenzione ‘attention’	104
Fortuna ‘chance’	208	Credibilità ‘credibility’	102
Allegria ‘joy’	192	Rispetto ‘respect’	96
Serenità ‘peace’	182	Testa ‘head’	96
Verità ‘truth’	151	Follia ‘folly’	96
Energia ‘energy’	128	Ottimismo ‘Optimism’	94

Briciolo builds upon the semantic features of *briciola* – namely fragmentariness, insufficiency, and value – and is used to convey both degree and evaluative functions (11).

(11) Mi son chiesto: la coerenza e un *briciolo di dignità* dove sono andati a finire?

‘I asked myself: consistency and an *ounce of dignity*, where did they go?’

3.3.1. *Krupica* ‘grain’: Russian data

As already stated in Section 2, *grain* denotes the fundamental unit of certain cereals or of some material as sand, gold and others, and is conceptually relevant as something small that can be gathered to form a larger quantity, the features that can be selected in the conceptualization process, apart from the small size, are its being defined, independent and able to interact with other bits concurring to form something bigger. The semantic component INDIVIDUAL, SINGLE is effectively reflected in the frequent use of modifiers such as indefinite pronouns (*каждая, какая-то, кое-какая, любая*), or deictics and determined modifiers (*эта, та, из тех, одна*).

(12) Наш долг – представителей более молодых поколений – бережно хранить *каждую крупицу* правды.

‘It is our duty, as representatives of younger generations, to carefully preserve *every grain* of truth.’

Table 12

Occurrences and frequency of the singular and plural forms of *krupica* ‘grain’

	Occurrences	Ipm	Negative polarity		
			<i>ne, ni</i>	<i>bez</i>	<i>liš’/daže</i>
<i>Krupica</i> (sg.)	382	0.5	6 + 31 ¹	7	3 + 10 ²
<i>Krupicy</i> (pl.)	854	0.12	5	–	30

¹ Manual screening of the “supposed” plural forms (see the cell below) preceded by *ne* or *ni* revealed that, out of 36 occurrences, only 5 were true plural forms (nominative or genitive cases), while the remaining 31 were in fact genitive singular. These 31 instances are therefore reported here.

² The same as explained above.

As will be shown below, these nouns do not have comparable frequency in the two languages.

In all the cases of negative polarity, *krupica* collocates with abstract nouns, except for one case in which the grain’s small size is used to visualize the tiniest parts of computer (12).

- (12) ...самоорганизующиеся микроскопические компьютеры, размеры которых не превышают *крупницу* песка.
 '...self-organising microscopic computers no bigger than a *grain of sand*'.

Krupica is frequently used to conceptualize information as in (13):

- (13) ...к журналистам *The Guardian* попала лишь *крупница информации* о финансовых вливаниях в имидж Украины.
 'The journalists at *The Guardian* received only a *grain of information* about the financial investments in Ukraine's image.'

The collocates, in both singular and plural forms, are abstract nouns denoting epistemic concepts (Table 13) such as knowledge, information, news, and history – entities that are cognitively perceived as aggregations of numerous small units. This cognitive association may explain the noun's marked preference for this specific semantic field.

Table 13

**The collocates of the singular
and plural forms ≥ 10 for *krupica* 'grain'**

Collocates (sg.)	Hits	Collocates (pl)	Hits
Правды 'truth'	60	Информации 'information'	206
Информации 'information'	49	Правды 'truth'	54
Творчества 'art'	21	Золота 'gold'	48
Соли 'salt'	20	Знаний 'knowledge'	24
Золота 'gold'	19	Соли 'salt'	19
Истины 'truth'	18	Истины 'truth'	18
Знаний 'knowledge'	12	Истории 'history'	17
		Сведений 'showing'	17
		Данных 'data'	13
		Мудрости 'wisdom'	11
		Опыта 'experience'	11
		Времени 'time'	10
		Пыли 'dust'	10

- (14) Но из скудных официальных сообщений и *собранных крупиц* информации от нескольких наших украинских источников <....>
 ‘But from the meager official reports and the *collected bits* of information from a few of our Ukrainian sources <....>’

In these contexts it is frequently encountered after the verbs *собирамы/собрать* ‘gather’, and *соединямы/соединять* ‘unite’. This meaning is made evident also by the frequent use of the Dative plural form *по крупицам* ‘bit by bit’ as in (15).

- (15) Сложно переоценить роль поисковых отрядов в деле восстановления по *крупницам* судеб погибших во время Великой Отечественной войны солдат.
 ‘It is difficult to overestimate the role of search teams in the restoration *bit by bit* of the fate of soldiers who died during the Great Patriotic War.’

3.3.2. *Granello* ‘grain’: Italian data

Granello occurs significantly more frequently in the Italian corpus than in the Russian one (see Table 14).

A look at the collocates (Table 15) clearly reveals a preference for concrete nouns in Italian.

Despite the apparent predominance of concrete collocates, a closer analysis reveals that they are often used metaphorically (16), not only literally (17).

Table 14

Occurrences and frequency of the singular and plural forms of *granello*

	Occurrences	Ipm	Negative polarity
			<i>non, neanche, nemmeno, neppure</i>
<i>Granello</i> (sg.)	3050	0.27	101
<i>Granelli</i> (pl.)	2936	0.27	4

- (16) Ma gli incentivi sono un *granello* di sabbia
 ‘But the incentives are just a *grain* of sand.’
- (17) Diversi i sintomi con cui la persona reagisce all’inalazione di un *granello* di polline.
 ‘A person may react to the inhalation of a single *grain* of pollen with a range of symptoms.’

The collocate types ≥ 10 for *granello* 'grain'

Collocates	Hits	Collocates	Hits
<i>Sabbia</i> 'sand'	1742	<i>Terra</i> 'earth'	20
<i>Polvere</i> 'dust'	510	<i>Colpa</i> 'guilt'	17
<i>Sale</i> 'salt'	190	<i>Speranza</i> 'chance'	17
<i>Pepe</i> 'pepper'	60	<i>Sporco</i> 'dirty'	16
<i>Zucchero</i> 'sugar'	38	<i>Ghiaccio</i> 'ice'	14
<i>Verità</i> 'truth'	36	<i>Polline</i> 'pollen'	14
<i>Riso</i> 'rice'	29	<i>Materia</i> 'matter'	10
<i>Senapa</i> 'mustard'	27		

Due to the high number of occurrences, it was necessary to refer to empirical frequency, as illustrated in the chart in Figure 4, based on the analysis of 300 occurrences.

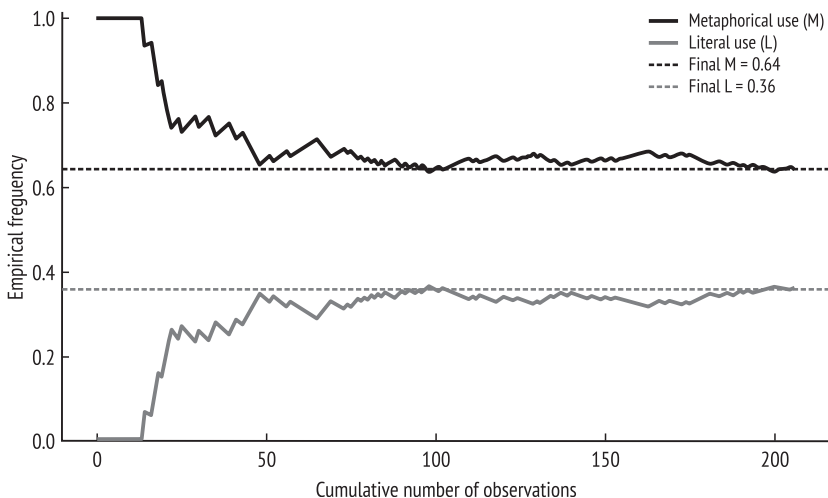


Fig. 3. Distribution of literal and metaphorical uses of *granello*

In negative polarity contexts, *granello* functions as a degree intensifier, while in many other contexts it is perceived as a disruptive element – something that interferes with the functioning of a mechanism (e.g., *ingranaggi* 'gears' (12)).

(12) Il *granello* di sabbia che rischia di inceppare gli ingranaggi della ripresa economica.

‘The *grain* of sand that risks jamming the gears of economic recovery.’

Despite the predominantly metaphorical meaning of the word *granello*, it is noteworthy that its use in negative polarity contexts is relatively infrequent.

4. Conclusions

Against the background of ongoing scholarly research on binominal constructions expressing quantity, this study has offered a contrastive analysis of binominal constructions denoting small quantity in Italian and Russian. It focused on six core nouns (*goccia/kaplja*, *briciola/kroška*, *granello/krupica*), with special attention to their metaphorical development as well as to their grammaticalization and pragmaticalization processes.

Because previous research has highlighted the preference of such constructions for NP contexts, the search strings were designed to capture these contexts. Although the search for negative contexts may not have been exhaustive, it nonetheless yielded additional data relevant to this issue, which needs to be explored further. For instance, all non-assertive contexts – e.g., *У него есть хоть капля разума* ‘He has at least a drop of reason’, where the nouns not only quantify but also intensify, expressing absence, scarcity, or emotional intensity – identified during the manual screening should be systematically retrieved through automatic searches. In addition to the quantitative retrieval of data – such as NP contexts and collocates – and given the well-known homonymy between the genitive singular and the nominative plural of feminine Russian nouns, a qualitative examination of random samples of 300 occurrences was conducted with two main objectives: first, to assess the significance of NP contexts for singular and plural forms; and second, to determine the proportion of metaphorical and literal uses of the nouns in both languages. The reliability of the selected sample was verified by observing the convergence and stabilization pattern of empirical frequency. A further issue that had to be taken into account was the different behaviour of singular and plural forms in these binomials.

With regard to the criteria outlined above, Section 3 discussed the data and yielded the following results.

1. Concerning NP contexts – a factor already discussed in the literature – it was observed that across both languages, singular forms are more strongly associated with metaphorical uses and NP contexts, partially confirming

the role of such environments in facilitating semantic shift, as suggested by previous studies (e.g., [Brems, 2007; Giacalone Ramat, 2019]). It must be stressed, though, that only the singular form of *kaplja* – and, to a lesser extent, *kroška* – shows NP contexts that can be considered relevant factors in metaphorical development or grammaticalization, while in all other cases, in both languages, the percentage of such occurrences does not appear to be so significant. This is therefore a direction in which further research is required, in order to investigate all possible NP contexts, including non-assertive ones, and to extend the analysis to a larger set of nouns. A weaker connection between NP co-occurrence and plural forms can likewise be confirmed. Moreover, the two Italian nouns that appear to have undergone a full process of constructionalization – *goccio* and *briciolo* – interestingly do not display a strong link with NP contexts, although *briciolo* occurs slightly more frequently in such environments.

2. With regard to the metaphorical processes undergone by the nouns and the consequent paths of conceptualization, the data retrieved through the automatic search of collocates, combined with the manual screening of at least 300 occurrences for each noun, yielded the following results.

First, all nouns – with the exception of the Italian plural *briciole* – undergo metaphorical extension predominantly in the singular form rather than in the plural. The metaphorical use does not display parallel behaviour in the two languages: *kaplja* is frequently used metaphorically, whereas its Italian counterpart *goccia* is not. The Russian *kaplja* is most often conceptualized as a small quantity of emotion, feeling, or a similar abstract notion, thereby confirming the conceptual metaphor EMOTIONS ARE LIQUID.

Kroška shows a preference for concrete collocates, mostly referring to food, yet closer examination reveals that it is used predominantly metaphorically – and in NP contexts – to indicate the complete absence of something, thus pointing to the development of the noun into a degree modifier rather than a quantifier. The inherent features of SCARCITY and VALUE associated with the concept of “crumb” likely serve as a semantic basis for metaphorical extensions. In Russian, the strong link with a limited set of collocates suggests that we are dealing with a process of idiomatization rather than grammaticalization.

Italian *briciola* and *briciole*, by contrast, tend to collocate with abstract nouns belonging to the semantic field of emotions, suggesting a completely different path of conceptualizing emotions in Italian. While *briciola* marks an initial stage – combining with nouns that denote emotions or positively connoted abstract qualities – the process culminates in *briciolo*.

The comparison between *granello* and *krupica* further reinforces this asymmetry. While Russian *krupica* is primarily used in epistemic and informational contexts – often emphasizing gradual accumulation (“bit by bit”) – Italian *granello* shows a strong preference for concrete collocates, although it is also frequently used metaphorically to denote disruption or minimal impact, still retaining its primary concrete features (e.g., *un granello di sabbia negli ingranaggi*).

3. As far as the grammaticalization and/or constructionalization processes are concerned, the empirical data yield particularly interesting results. Almost all nouns show a preference for the singular form in the process of metaphoricization, thus suggesting a certain degree of decategorization. Although NP contexts are not always significant, the manual examination shows that they occur almost exclusively in metaphorical environments, and even in the case of *kroška*, where concrete collocates predominate, they display metaphorical and degree-related meanings.

In Italian, a broader functional split can be observed. In particular, the emergence of the masculine forms *goccio* and *briciolo* indicates a more advanced stage of constructionalization: both nouns have undergone a shift in gender, number, and even category, and are now used primarily with quantifying (*goccio*) or intensifying (*briciolo*) functions. *Briciolo*, one could argue, has become a full scalar modifier, gradually losing its lexical core in favour of an abstract quantificational and degree meaning.

These findings lead to further observations concerning the semantic and cognitively salient features of N_1 in its combination with N_2 and the relevance of NP contexts, which needs further investigation.

Altogether, the results of this research suggest that although both languages share the typological basis for binominal constructions of small quantity, the processes of metaphorical shift, constructionalization, and semantic and functional extension vary significantly. Italian appears to offer a wider range of structural and pragmatic flexibility, particularly through the grammaticalization of masculine variants. Russian, on the other hand, reveals a tighter correlation between negative polarity and metaphorical intensification, particularly within emotional and evaluative domains. Overall, the value of a comparative approach in studying binomials expressing small quantity is clearly confirmed, as the comparison highlights the limits of some theoretical issues and suggest new perspective for research.

Future research may explore the diachronic paths of these constructions, their potential parallels in other Indo-European languages, and their implications for the interface between lexis, grammar, and discourse pragmatics.

References

Benigni, 2022 – Benigni V. Binominal constructions with metaphorical quantifiers in Russian: “vsplesk emocij i volna pozitivna”. *Studi Slavistici*. 2022. No. 2. Pp. 169–192.

Benigni et al., 2023 – Benigni V., Latos A. Una montagna di errori: costruzioni binominali con classificatori metaforici. *Studia universitatis hereditati znanstvena revija za raziskave in teorijo kulturne dediščine*. 2023. No. 11. Pp. 11–31.

Bordjugova, 2015 – Bordjugova K.A. Osobennosti grammatikalizacii predmetnykh imen-kvantifikatorov neopredelennogo malogo količestva (na materiale francuzskogo i russkogo jazykov). *Vestnik Čerepoveckogo gosudarstvennogo universiteta*. 2015. Vol. 8. No. 69. Pp. 41–44. (In Rus.) [Бордюгова К.А. Особенности грамматикализации предметных имен-квантификаторов неопределенно-малого количества (на материале французского и русского языков) // Вестник Череповецкого государственного университета. 2015. Т. 8. № 69. С. 41–44.]

Brems, 2003 – Brems L. Measure Noun constructions: An instance of semantically-driven grammaticalization. *International Journal of Corpus Linguistics*. 2003. Vol. 8. No. 2. Pp. 283–312. DOI: 10.1075/ijcl.8.2.05bre

Brems, 2007 – Brems L. The grammaticalization of small size nouns. Reconsidering frequency and analogy. *Journal of English Linguistics*. 2007. Vol. 34. No. 4. Pp. 293–324. DOI: 10.1177/0075424207307597

Brems, 2010 – Brems L. Size nouns constructions as collocationally constrained constructions: Lexical and grammaticalized uses. *English Language and Linguistics*. 2010. Vol. 14. No.1. Pp. 83–109. DOI: 10.1017/S1360674309990372.

Cotta Ramusino et al., 2025 – Cotta Ramusino P.M., Maiko T. Binominal constructions of small quantity in Russian, Belarusian, and Ukrainian. *Scando-Slavica*. 2025. Vol. 71. No. 1. Pp. 1–24. DOI: 10.1080/00806765.2025.2507025

Croft et al., 2004 – Croft W., Cruse D.A. *Cognitive Linguistics*. Cambridge University Press, 2004.

Giocalone Ramat, 2019 – Giocalone Ramat A. Degrees of grammaticalization and measure constructions in Italian. *Revue Romane. Langue et littérature. International Journal of Romance Languages and Literatures*. 2019. Vol. 54. Pp. 257–277. DOI: 10.1075/rro.16015.gia

Kövecses, 2000 – Kövecses Z. *Metaphor and emotion: Language, culture, and body in human feeling*. Cambridge University Press, 2000.

Israel, 1996 – Israel M. Polarity sensitivity as lexical semantics. *Linguistics and Philosophy*. 1996. Vol. 19. Pp. 619–666.

Israel, 2004 – Israel M. The pragmatics of polarity. *The Handbook of Pragmatics*. L.R. Horn, G. Ward (eds.). Oxford/Malden, MA, 2004. Pp. 701–23.

Langacker, 1991 – Langacker R.W. *Foundations of Cognitive Grammar*. Vol. 2. Descriptive applications. Stanford University Press, 1991.

Langacker, 2013 – Langacker R.W. *Essentials of cognitive grammar*. Oxford University Press, 2013.

Li, 2005 – Li S.Ch. Kognitivnyj analiz russkich konstrukcij s imennymi kvantifikatorami [Cognitive analysis of Russian constructions with nominal quantifiers]. PhD dis. Moscow, 2005. [Ли С.Х. Когнитивный анализ русских конструкций с именными квантификаторами: дис. ... канд. филол. наук. М., 2005.]

Rachilina et al., 2009 – Rachilina E., Li S.Ch. Semantika leksičeskoj množestvenosti v ruskom jazyke. *Voprosy jazykoznanija*. 2009. No. 4. Pp. 13–40. (In Rus.) [Рахилина Е., Ли С.Х. Семантика лексической множественности в русском языке // Вопросы языкознания. 2009. № 4. С. 13–40.]

Traugott, 2008 – Traugott E.C. The grammaticalization of NP of NP constructions. *Constructions and Language Change*. A. Bergs, G. Diewald (eds.). New York, 2008. Pp. 21–43.

Traugott, 2015 – Traugott E.C. Toward a coherent account of grammatical constructionalization. *Diachronic Construction Grammar*. J. Barðdal, E. Smirnova, L. Sommerer, S. Gildea (eds.). Amsterdam, 2015. Pp. 51–79.

Traugott et al., 2013 – Traugott E.C., Trousdale G. Constructionalization and constructional changes. Vol. 6. Oxford University Press, 2013.

The article was received on 06.07.2025

Статья поступила в редакцию 06.07.2025

About the author / Сведения об авторе

Paola Maria Cotta Ramusino – PhD in Slavistics; Full Professor at the Department of Languages, Literatures, Cultures and Mediations, University of Milan (Universita degli Studi di Milano La Statale), Milan, Italy

Котта Рамузино Паола Мария – PhD (славистика); профессор кафедры языков, литератур, культур и межязыковой коммуникации, Миланский государственный университет, г. Милан, Италия

ORCID: <https://orcid.org/0000-0002-5016-7568>

E-mail: paola.cottaramusino@unimi.it

В.И. Лемаев, Н.В. Лукашевич

Московский государственный университет
имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация

Влияние порядка слов на качество межъязыкового переноса моделей для распознавания эмоций на русском языке

В данной работе рассмотрен перенос обученных языковых моделей с одного языка на другой для задачи автоматической классификации эмоций в тексте и кратко описаны основные современные подходы для решения этой задачи. Основной целью было исследование влияния порядка слов на качество классификации перенесенных моделей на основе данных соревнования SemEval-2025 Task 11. Для исследования была использована модель XLM-RoBERTa, отдельно обученная на немецком, испанском, китайском, хинди и татарском языках и затем протестированная на данных русского языка. Для полученных результатов были проведены анализ и подсчет их статистической значимости, на основе которых были описаны основные факторы, влияющие на качество переноса языковых моделей.

Ключевые слова: обработка естественного языка, межъязыковой перенос, распознавание эмоций, трансформеры, XLM-RoBERTa

Для ЦИТИРОВАНИЯ: Лемаев В.И., Лукашевич Н.В. Влияние порядка слов на качество межъязыкового переноса моделей для распознавания эмоций на русском языке // Рема. Rhema. 2025. № 4. С. 53–77. DOI: 10.31862/2500-2953-2025-4-53-77



DOI: 10.31862/2500-2953-2025-4-53-77

V. Lemaev, N. LoukashevitchLomonosov Moscow State University,
Moscow, 119991, Russian Federation

The influence of word order on the quality of cross-linguistic transfer of emotion recognition models in Russian

This study explores the transfer of pretrained language models from one language to another for the task of automatic emotion classification in text and provides a brief overview of current state-of-the-art approaches to this problem. The primary objective was to investigate the impact of word order on the classification performance of transferred models using the dataset from the SemEval-2025 Task 11 competition. For the experiments, the XLM-RoBERTa model was employed, separately fine-tuned on German, Spanish, Chinese, Hindi, and Tatar, and subsequently evaluated on Russian-language data. The obtained results were analyzed for statistical significance, allowing the identification of key factors influencing the effectiveness of cross-lingual transfer in language models.

Key words: natural language processing, cross-lingual transfer, emotion recognition, transformers, XLM-RoBERTa

FOR CITATION: Lemaev V., Loukashevitch N. The influence of word order on the quality of cross-linguistic transfer of emotion recognition models in Russian. *Rhema*. 2025. No. 4. Pp. 53–77. DOI: 10.31862/2500-2953-2025-4-53-77

1. Введение

Автоматическое распознавание эмоций является одной из популярных задач в сфере обработки естественного языка и состоит в приписывании всему тексту или его отрывку определенной эмоции. На практике алгоритмы распознавания эмоций могут применяться как для улучшения популярных в настоящее время виртуальных собеседников (чат-ботов), позволяя им выстраивать более естественный и понятный

диалог с человеком, так и для более целенаправленных задач, например, для постановки диагноза в области здравоохранения [Hansen et al., 2022] или для обеспечения общественной безопасности [Lefter, Jonker, 2017].

Современные языковые модели, такие как BERT и GPT, предобучались на большом объеме обучающих данных, в которых превалирует английский язык. При этом многие языки мира имеют ограниченный объем данных, доступный для обучения языковых моделей. В таких условиях альтернативой часто выступает перенос на целевой язык моделей, заранее обученных на другом языке, но данный подход пока не достиг тех же результатов, на которые способен человек [Pikuliak, 2021]. Применимо к задаче автоматического распознавания эмоций межъязыковой перенос также осложняет разница в подходах исследователей к составлению датасетов, в том числе разница в используемых наборах эмоций.

Целью данной работы является исследование влияния порядка слов в предложении (в первую очередь влияние базового порядка слов, т.е. взаимного расположения субъекта, объекта и глагола) на качество классификации эмоций в тексте при переносе текстовых языковых моделей с одного языка на другой на основе реальных данных, используемых на практике.

В исследовании использовалась мультязычная модель XLM-RoBERTa [Conneau et al., 2020], дообученная на эмоциональных текстах немецкого, испанского, китайского, хинди и татарского языков и затем протестированная на русском языке. При обучении и тестировании были использованы данные из датасета BRIGHTER [Muhammad et al., 2025a], собранные в рамках соревнования SemEval-2025 Task 11 [Muhammad et al., 2025b]. В соответствии с правилами подзадания Track C, которое в данном соревновании состояло в межъязыковом переносе моделей, обученных для решения задачи распознавания эмоций, в текущем исследовании при обучении модели также не были использованы данные целевого (русского) языка (подход zero-shot).

Структура данной работы оформлена следующим образом: в разделе 2 представлен разбор предыдущих работ и описание лингвистических факторов и их влияния на качество классификации при переносе обученной языковой модели с одного языка на другой. В разделах 3 и 4 содержатся краткое описание применяемых в исследованиях наборов эмоций и современных подходов к классификации эмоций соответственно. В разделе 5 содержится описание использованных в данной работе обучающих и тестовых данных, а также языковой модели. Наконец, в разделе 6 представлены полученные в рамках данного исследования результаты и их описание.

2. Влияние лингвистических факторов на качество переноса обученных языковых моделей

На данный момент существует достаточное количество исследований о влиянии лингвистических признаков на качество результатов работы обученных языковых моделей при их межъязыковом переносе. Однако большинство таких исследований рассматривает только наиболее общие задачи обработки языка, такие как распознавание именованных сущностей, частеречная разметка, ответы на вопросы и т.д., почти не затрагивая задачу распознавания эмоций. Так, в работе [Philippy et al., 2023] собрано подробное описание факторов, влияющих на качество переноса языковых моделей. Всего авторы выделяют пять таких факторов: структурное сходство языков, пересечение лексики между ними, данные для обучения, архитектура языковой модели и гиперпараметры ее обучения. В настоящей работе в первую очередь рассматривается влияние лингвистических факторов, поэтому мы не будем подробно останавливаться на архитектуре моделей и их гиперпараметрах, а сосредоточимся только на первых трех факторах.

К *структурному сходству* языков в первую очередь относятся *синтаксис* и *порядок слов*. Хотя они традиционно считаются одними из самых важных признаков, исследования показывают, что при переносе языковых моделей они не всегда оказывают значимое влияние. Например, в работах [Deshpande et al., 2022; Wu et al., 2023] авторы проводят сравнение между случайными перестановками слов и целенаправленной адаптацией порядка слов под другой язык. Как было обнаружено, если в предложениях первого языка порядок слов адаптирован под порядок слов второго языка, то качество классификации языковой модели при ее переносе снижается меньше, чем в случае случайного порядка слов в первом языке.

Среди основных лингвистических признаков также выделяют *близость языков*, в частности, генетическую, географическую и фонологическую. Согласно исследованиям [Lin et al., 2019; Lauscher et al., 2020], все типы близости значительно влияют на качество классификации при межъязыковом переносе языковых моделей, при этом фонологическая близость является более значимым критерием для задач на уровне обработки токенов (например, для распознавания именованных сущностей), чем для задач на уровне обработки всего предложения. Определение структурного сходства языков обычно происходит на основе типологических баз данных, например, WALIS [Dryer, Haspelmath, 2013], и векторов лингвистических признаков языков – чаще всего для этого применяется библиотека lang2vec [Littell et al., 2017].

Под *пересечением лексики* языков подразумевается доля слов, присутствующих в обоих исследуемых языках, которая рассчитывается как процент общих для двух языков слов от суммы всех уникальных слов, встречающихся в этих языках. Для подсчета пересекающейся лексики также часто используется расстояние Левенштейна [Wichmann et al., 2010] и метрика ezGlott [Kovacevic et al., 2022]. Как показывают исследования, хотя пересечение лексики и оказывает влияние на качество межъязыкового переноса языковых моделей, это влияние почти никогда не является изолированным. Так, в работе [Patil et al., 2022] на материале близких языков было установлено, что большее количество пересекающейся лексики значительно сильнее улучшает качество переноса в тех случаях, когда количество обучающих данных исходного языка мало. Работа [Deshpande et al., 2022] же демонстрирует, что пересечение лексики значительнее всего влияет на качество работы перенесенной модели тогда, когда порядки слов у языков различаются.

Влияние *обучающих данных* на качество межъязыкового переноса моделей в первую очередь зависит от *размера обучающего корпуса* исходного языка. Так, увеличение количества обучающих данных на этапе предобучения значительно улучшает качество переноса современных языковых моделей на основе трансформеров, таких как mBERT [Devlin et al. 2019], особенно в сравнении с более старыми нейросетевыми подходами, у которых рост качества при увеличении объема обучающих данных в определенный момент замедлялся вплоть до остановки [Liu et al., 2020].

Помимо этого, для мультиязычных моделей на этапе предобучения также важно *соответствие между корпусами* всех используемых языков, причем в первую очередь более важна принадлежность корпусов к одной и той же области, чем их параллельность [Deshpande et al., 2022].

3. Составление инвентаря эмоций

Одной из частых проблем, возникающих при переносе обученных языковых моделей, является различие в подходах к классификации данных. Для области автоматического распознавания эмоций эта проблема стоит особенно остро, т.к. эмоции имеют высокую вариативность и значимые различия в их выражении в зависимости от языка и культурного бэкграунда говорящего [Jackson et al., 2019]. Типичные наборы эмоций, используемые в различных областях психологии, могут насчитывать около 300 разнообразных эмоций и их вариаций. Однако для задач автоматического распознавания исследователи обычно выделяют 6 эмоций, которые являются универсальными для большинства языков и называются

«архетипическими»: это гнев, отвращение, страх, радость, грусть и удивление [Cowie et al., 2001]. Эти 6 эмоций и образуют базовый набор, используемый в большинстве исследований.

Другой подход, который часто применяется наравне с небольшим набором эмоций, – это классификация данных на основе двумерного пространства. В качестве первого измерения используется понятие «возбуждение», которое описывает степень интенсивности выражаемой эмоции. Возбуждение часто коррелирует с физиологическими процессами, происходящими в теле человека при выражении определенных эмоций. Например, повышенный сердечный ритм и быстрая речь характерны для эмоций гнева, страха и радости, в то время как, наоборот, пониженный сердечный ритм и медленная речь характерны для эмоции грусти. В текстах степень возбуждения можно часто оценить по наречиям, например, *очень* и *слегка*.

В качестве второго измерения выступает понятие «валентность», отражающее степень положительности или, наоборот, отрицательности выражаемой эмоции. Валентность помогает дифференцировать эмоции гнева и радости, для которых часто характерна одинаково высокая степень интенсивности, и в текстах валентность часто соотносится с положительно или отрицательно окрашенными словами. Преимуществом подхода на основе возбуждения/валентности является то, что обученная модель получает более четкое представление о соотношении эмоций между собой и даже в случае ошибки может с большей вероятностью предсказать эмоцию, близкую к истинной [Mitsios et al., 2024].

4. Подходы к классификации

Наиболее популярными в настоящее время являются подходы на основе самообучения, позволяющие алгоритмам самостоятельно находить закономерности в исходных данных и тем самым обрабатывать большие объемы данных с минимальной предварительной экспертной разметкой. Это стало возможно в первую очередь благодаря появлению архитектуры трансформера [Vaswani et al., 2017], в основе которой лежит модифицированная идея механизма внимания, состоящая в том, что при решении конкретной задачи языковая модель может обращать больше внимания на определенные части данных в зависимости от контекста и требований самой задачи. Данная модель смогла не только улучшить предыдущие результаты в различных задачах, но и также стала способна одновременно обрабатывать большее количество данных, тем самым ускорив этап обучения. На архитектуре трансформера базируются как большие языковые модели (англ. LLM), например, GPT [Radford et al.,

2018], так и более классические BERT [Devlin et al., 2019] и RoBERTa [Liu et al., 2019]. В ходе соревнования SemEval-2025 Task 11, на котором были поставлены задачи как автоматического распознавания эмоций предложений в рамках одного языка, так и переноса обученных моделей на другие языки, наилучшие результаты показали подходы на основе ансамблей LLM-моделей и написанных для них промт-запросов.

В рамках данной статьи для задачи переноса используется мультиязычная модель XLM-RoBERTa-base¹, обученная на текстовых данных 100 языков общим объемом 2,5 терабайт, представленных в открытой библиотеке CommonCrawl² (большинство текстов написаны на английском языке). Модель XLM-RoBERTa основана на модели RoBERTa, представляющей собой оптимизированную архитектуру BERT, которая улучшила качество ее работы за счет обучения на большем объеме данных с увеличенным количеством шагов самого обучения, а также других небольших изменений. Предобучение модели XLM-RoBERTa производилось на задаче MLM (англ. Masked Language Modeling), состоящей в предсказании предварительно замаскированных в предложении слов за счет других слов, стоящих по соседству.

5. Данные и модель

В качестве данных был взят мультиязычный датасет BRIGHTER [Muhammad et al., 2025a], содержащий письменные эмоциональные тексты для 28 языков. В качестве источников для данного датасета выступили посты в социальных сетях, личные блоги, новостные источники и вручную или автоматически сгенерированные предложения. В датасете представлены 6 эмоций для каждого языка: гнев/anger, отвращение/disgust, страх/fear, радость/joy, грусть/sadness, удивление/surprise. Для каждого языка тексты внутри датасета разбиты на обучающую, валидационную и тестовую выборки. Из-за небольшого объема имеющихся данных, а также для воспроизводимости результатов исследования, в данной работе использовались исходные выборки без удаления или добавления дополнительных примеров за исключением следующих случаев:

- 1) в данной работе использовались только те примеры, которые выражают как минимум одну эмоцию;
- 2) в случае выражения нескольких эмоций одновременно истинной была принята первая по списку.

¹ URL: <https://huggingface.co/FacebookAI/xlm-roberta-base>

² URL: <https://commoncrawl.org/>

В качестве алгоритма классификации была использована модель XLM-RoBERTa-base. В рамках текущей работы данная модель была отдельно дообучена на данных каждого исходного языка, представленного в датасете BRIGHTER (использовались данные обучающей и валидационной выборок), и затем перенесена и протестирована на тестовой выборке русского языка без дополнительного обучения. В качестве основной метрики качества классификации во всех примерах использовалась взвешенная F-мера. На основе результатов переноса модели было отобрано 5 языков с наилучшим общим качеством классификации: немецкий, испанский, китайский, хинди и татарский.

В качестве зависимой переменной для исследования выступил базовый порядок слов, в частности, порядки SVO, SOV и VSO. Рассмотрим базовый порядок слов в данных языках более подробно, приведенные сведения опираются на работу [Dryer, Haspelmath, 2013].

Немецкий: имеет комплексную систему порядка слов, которая выражается в использовании порядка V2 (verb second) в главной клаузе – при этом в таких случаях преобладает порядок SVO – и порядка SOV в зависимых клаузах. В то же время, несмотря на предпочтительное расположение S перед O в предложении, каких-то реальных ограничений на это нет, и их порядок является относительно свободным.

Испанский: базовым порядком считается порядок SVO. В то же время испанский имеет достаточно богатую морфологию, что на практике выражается в относительно свободном порядке слов.

Китайский: базовым порядком является порядок SVO.

Хинди: базовым порядком официально считается порядок SOV, при этом порядок может меняться с целью изменения прагматического смысла высказывания.

Татарский: базовым порядком слов является порядок SOV.

Русский: базовым порядком слов традиционно считается порядок SVO, но при этом, благодаря богатой морфологии, на практике порядок слов является в значительной степени свободным.

В таблице 1 кратко приведены данные о базовом порядке слов для каждого из упомянутых языков.

В таблице 2 для каждого языка представлено распределение обучающих данных по эмоциям, в таблице 3 – общее количество примеров в русской тестовой выборке. Объем валидационных выборок для каждого языка составил около 15% от объема соответствующей обучающей выборки (около 10% для хинди и около 20% для татарского языка): использовались валидационные выборки, представленные в датасете.

Таблица 1

**Базовый порядок слов в рассматриваемых
в данной статье языках**
[Basic word order in the languages used in this paper]

Язык [Language]	Порядок слов [Word order]
Немецкий [German]	V2 в главной клаузе, SOV в зависимой клаузе [V2 in the main clause, SOV in the subordinate clause]
Испанский [Spanish]	SVO
Китайский [Chinese]	SVO
Хинди [Hindi]	SOV
Татарский [Tatar]	SOV
Русский [Russian]	SVO

Таблица 2

**Количество обучающих примеров для каждой эмоции
в разных языках**
**[Number of training examples for each emotion
in each language used for training]**

Эмоция [Emotion]	Язык [Language]				
	Немецкий [German]	Испанский [Spanish]	Китайский [Chinese]	Хинди [Hindi]	Татарский [Tatar]
Гнев [Anger]	768	492	1178	422	137
Отвращение [Disgust]	294	257	56	181	108
Страх [Fear]	159	295	46	369	48
Радость [Joy]	481	579	523	442	182
Грусть [Sadness]	201	192	189	346	214
Удивление [Surprise]	55	181	47	191	126
Всего [Total]	1958	1996	2039	2556	815

Таблица 3

**Количество примеров в русской тестовой выборке
для каждой эмоции**
[Number of examples in the Russian test dataset for each emotion]

Эмоция [Emotion]	Количество [Number of examples]
Гнев [Anger]	452
Отвращение [Disgust]	172
Страх [Fear]	212
Радость [Joy]	386
Грусть [Sadness]	206
Удивление [Surprise]	144
Всего [Total]	1572

В таблицах 4–8 для каждого языка, данные которого использовались для обучения модели, указано количество обучающих примеров для каждого порядка слов в зависимости от эмоции.

Таблица 4

**Количество примеров для каждой эмоции
в немецкой обучающей выборке в зависимости
от порядка слов в предложении**
[Number of examples for each emotion in the German training dataset
by the word order in the sentence]

Эмоция [Emotion]	Порядок слов [Word order]		
	SVO	SOV	VSO
Гнев [Anger]	221	147	50
Отвращение [Disgust]	79	51	22
Страх [Fear]	29	39	11
Радость [Joy]	107	85	33
Грусть [Sadness]	44	36	12
Удивление [Surprise]	19	9	3
Всего [Total]	499	367	131

Таблица 5

**Количество примеров для каждой эмоции
в испанской обучающей выборке в зависимости
от порядка слов в предложении**
[Number of examples for each emotion in the Spanish training dataset
by the word order in the sentence]

Эмоция [Emotion]	Порядок слов [Word order]		
	SVO	SOV	VSO
Гнев [Anger]	86	52	22
Отвращение [Disgust]	34	26	6
Страх [Fear]	48	43	13
Радость [Joy]	69	63	20
Грусть [Sadness]	29	24	11
Удивление [Surprise]	16	26	7
Всего [Total]	282	234	79

Таблица 6

**Количество примеров для каждой эмоции
в китайской обучающей выборке в зависимости
от порядка слов в предложении**
[Number of examples for each emotion in the Chinese training dataset
by the word order in the sentence]

Эмоция [Emotion]	Порядок слов [Word order]		
	SVO	SOV	VSO
Гнев [Anger]	534	11	193
Отвращение [Disgust]	30	0	3
Страх [Fear]	22	0	15
Радость [Joy]	225	1	85
Грусть [Sadness]	79	3	29
Удивление [Surprise]	17	0	12
Всего [Total]	907	15	337

Таблица 7

**Количество примеров для каждой эмоции
в обучающей выборке на языке хинди
в зависимости от порядка слов в предложении**
[Number of examples for each emotion in the Hindi training dataset
by the word order in the sentence]

Эмоция [Emotion]	Порядок слов [Word order]		
	SVO	SOV	VSO
Гнев [Anger]	19	23	32
Отвращение [Disgust]	10	7	4
Страх [Fear]	7	19	20
Радость [Joy]	8	19	13
Грусть [Sadness]	15	16	13
Удивление [Surprise]	6	12	18
Всего [Total]	65	96	100

Таблица 8

**Количество примеров для каждой эмоции
в татарской обучающей выборке
в зависимости от порядка слов в предложении**
[Number of examples for each emotion in the Tatar training dataset
by the word order in the sentence]

Эмоция [Emotion]	Порядок слов [Word order]		
	SVO	SOV	VSO
Гнев [Anger]	21	24	5
Отвращение [Disgust]	15	29	11
Страх [Fear]	5	15	2
Радость [Joy]	27	47	16
Грусть [Sadness]	37	69	22
Удивление [Surprise]	21	23	5
Всего [Total]	126	207	61

В таблице 9 для русской тестовой выборки указано количество примеров для каждого порядка слов в зависимости от указанной эмоции. Для определения порядка слов была использована библиотека для обработки естественного языка SpaCy. Были выделены порядки SVO, SOV и VSO, при этом учитывались только примеры, где все три члена предложения (подлежащее, сказуемое и прямое дополнение) были выражены эксплицитно.

Таблица 9

**Количество примеров для каждой эмоции
в русской тестовой выборке
в зависимости от порядка слов в предложении
[Number of examples for each emotion in the Russian test dataset
by the word order in the sentence]**

Эмоция [Emotion]	Порядок слов [Word order]		
	SVO	SOV	VSO
Гнев [Anger]	86	30	19
Отвращение [Disgust]	36	13	18
Страх [Fear]	54	14	6
Радость [Joy]	66	6	22
Грусть [Sadness]	16	10	10
Удивление [Surprise]	14	12	6
Всего [Total]	272	85	81

Статистика порядка слов в указанных языках в целом совпадает с традиционно принятой классификацией базовых порядков. У немецкого и испанского языков количество примеров для порядков SVO и SOV находится на сравнительно равном уровне с небольшим преобладанием примеров с порядком SVO. В китайской обучающей выборке почти не наблюдается примеров с порядком SOV, а преобладающим порядком является SVO. Тестовая выборка на языке хинди имеет схожее количество примеров с порядками SOV и VSO, а также значительное количество примеров с порядком SVO. Наконец, татарская обучающая и русская тестовая выборки имеют ожидаемое преобладание примеров с порядком SOV и SVO соответственно.

6. Результаты и анализ

В данном разделе рассмотрим качество переноса модели, обученной на разных языках. Общее качество переноса языковых моделей на русский язык представлено в таблице 10, в таблице 11 для каждой модели представлено качество классификации отдельных эмоций.

Таблица 10

Общее качество классификации модели XLM-RoBERTa-base при ее дообучении на соответствующем исходном языке и последующем тестировании на русской выборке
[Overall classification performance of the XLM-RoBERTa-base model fine-tuned on each respective source language and evaluated on the Russian dataset]

Язык обучения [Language used for training]	Взвешенная F-мера [Weighted F-score]
Немецкий [German]	0,43
Испанский [Spanish]	0,63
Китайский [Chinese]	0,39
Хинди [Hindi]	0,55
Татарский [Tatar]	0,51

Таблица 11

Качество классификации отдельных эмоций в русских тестовых примерах в зависимости от исходного языка обучения модели
[Classification performance for individual emotions in the Russian test dataset by the source language used for model training]

Эмоция [Emotion]	Язык обучения [Language used for training]				
	Немецкий [German]	Испанский [Spanish]	Китайский [Chinese]	Хинди [Hindi]	Татарский [Tatar]
Гнев [Anger]	0,57	0,56	0,57	0,57	0,55
Отвращение [Disgust]	0,26	0,20	0,00	0,08	0,11
Страх [Fear]	0,43	0,81	0,00	0,71	0,27
Радость [Joy]	0,68	0,77	0,69	0,75	0,68
Грусть [Sadness]	0,09	0,69	0,39	0,39	0,68
Удивление [Surprise]	0,00	0,65	0,00	0,56	0,47

Наибольшее качество переноса показывают модели, дообученные на испанском языке и хинди. Скорее всего, модели, обученные на немецком и китайском языке, показывают более плохое качество в первую очередь из-за меньшего количества обучающих примеров для отдельных эмоций. Однако, несмотря на малое количество обучающих примеров для татарского языка, обученная на нем модель показала сравнительно высокое качество классификации, что, вероятнее всего, вызвано более близкой связью татарского языка с русским, а также схожестью их алфавитов.

Основное исследование состояло в проверке влияния порядка слов различных языков на качество переноса моделей, обученных на этих языках, на русские тексты. Для каждой перенесенной языковой модели были подсчитано ее качество классификации в зависимости от базового порядка слов. Полученные результаты приведены в таблице 12, на рисунке 1 представлен их график.

Таблица 12

**Общее качество классификации русских тестовых примеров
в зависимости от исходного языка обучения модели
и порядка слов в тексте**
[Overall classification performance on the Russian test dataset
by the source language and the word order in the text]

Язык обучения [Language used for training]	Порядок слов [Word order]		
	SVO	SOV	VSO
Немецкий [German]	0,55	0,41	0,43
Испанский [Spanish]	0,64	0,61	0,56
Китайский [Chinese]	0,34	0,28	0,32
Хинди [Hindi]	0,60	0,50	0,46
Татарский [Tatar]	0,41	0,47	0,38

Для каждого языка была оценена статистическая значимость порядка слов в предложении. В таблице 13 представлены значения *p-value*, показывающие значимость различия результатов классификации в зависимости от порядка слов в тестовых данных. В качестве метода использовался тест PERMANOVA [Anderson, 2001], который является непараметрическим пермутационным тестом оценки статистической значимости распределения в случае, когда имеет больше двух групп значений.

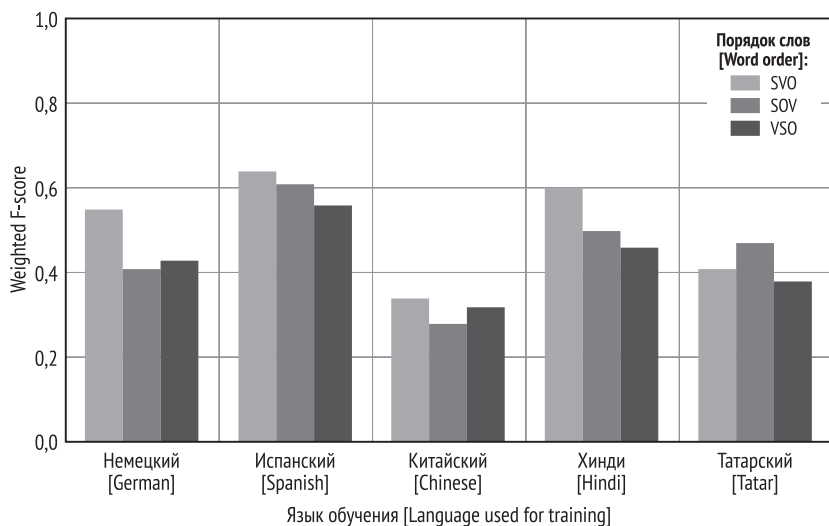


Рис. 1. График общего качества классификации русских тестовых примеров в зависимости от исходного языка обучения модели и порядка слов в тексте

Fig. 1. Graph showing the overall classification performance on the Russian test dataset by the source language and the word order in the text

Таблица 13

Результаты теста PERMANOVA для вычисления статистической значимости порядка слов при переносе обученной на соответствующем исходном языке модели на русский язык [PERMANOVA test results for assessing the statistical significance of word order when transferring a model trained on each respective source language to Russian]

Язык обучения [Language used for training]	p-value
Немецкий [German]	0,233
Испанский [Spanish]	0,444
Китайский [Chinese]	0,787
Хинди [Hindi]	0,108
Татарский [Tatar]	0,24

В качестве критического значения p -value здесь и во всех дальнейших подсчетах использовалось значение 0,05. Полученные в данном случае результаты указывают на отсутствие статистической значимости порядка слов в данных (p -value > 0,05 во всех случаях) для общей классификации эмоций. Модели, обученные на немецком, испанском и китайском языках, показывают наилучшее качество классификации для предложений, имеющих порядок SVO (см. рис. 1). В то же время у модели, обученной на татарском языке, можно увидеть увеличение качества при классификации предложений с порядком SOV. Из этого можно сделать предположение, что, хотя базовый порядок слов и не имеет достаточной статистической значимости при межъязыковом переносе моделей, в случае татарского языка базовый порядок слов все равно оказывает влияние на классифицирующую способность модели. Однако такой картины не наблюдается при обучении модели на языке хинди, тоже имеющим базовый порядок SOV. Вероятно, в данном случае на качество переноса также влияет близость языка-источника и языка-цели. Для более четкой картины необходимо провести дополнительные исследования с увеличенным объемом данных.

Для составления более четкой картины дополнительно было протестировано качество классификации для отдельных порядков слов в зависимости от эмоции, выраженной в предложении. В таблице 14 для каждого языка обучения представлено качество классификации примеров в зависимости от двух параметров: порядка слов и выражаемой эмоции. На основе этих данных на рисунке 2 представлены диаграммы для каждой отдельной эмоции.

Таблица 14

**Качество классификации отдельных эмоций
в русских тестовых примерах в зависимости
от исходного языка обучения модели и порядка слов в тексте**
[Classification performance for individual emotions in the Russian test
dataset by the source language and the word order in the text]

Эмоция [Emotion]	Язык обучения [Language used for training]	Порядок слов [Word order]		
		SVO	SOV	VSO
Гнев [Anger]	Немецкий [German]	0,75	0,97	0,73
	Испанский [Spanish]	0,69	0,70	0,88

Продолжение табл. 14

Эмоция [Emotion]	Язык обучения [Language used for training]	Порядок слов [Word order]		
		SVO	SOV	VSO
	Китайский [Chinese]	0,78	0,93	0,81
	Хинди [Hindi]	0,85	0,93	0,88
	Татарский [Tatar]	0,74	0,91	0,81
Отвращение [Disgust]	Немецкий [German]	0,78	0,63	0,62
	Испанский [Spanish]	0,24	0,70	0,20
	Китайский [Chinese]	0,00	0,00	0,00
	Хинди [Hindi]	0,11	0,00	0,00
	Татарский [Tatar]	0,00	0,00	0,00
Страх [Fear]	Немецкий [German]	0,62	0,44	0,50
	Испанский [Spanish]	0,94	0,73	1,00
	Китайский [Chinese]	0,00	0,00	0,00
	Хинди [Hindi]	0,77	0,60	0,80
	Татарский [Tatar]	0,31	0,25	0,50
Радость [Joy]	Немецкий [German]	0,90	0,80	0,90
	Испанский [Spanish]	0,94	1,00	0,90
	Китайский [Chinese]	0,86	0,80	0,71
	Хинди [Hindi]	0,94	1,00	0,84
	Татарский [Tatar]	0,71	1,00	0,62

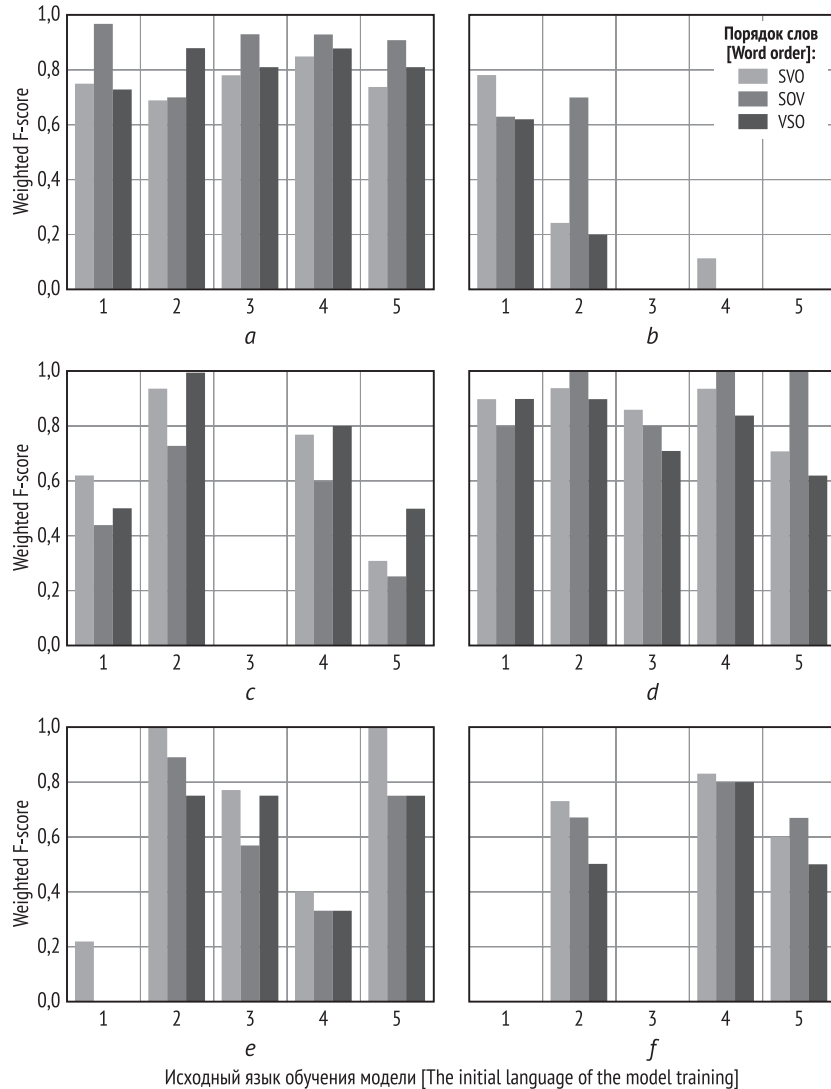
Эмоция [Emotion]	Язык обучения [Language used for training]	Порядок слов [Word order]		
		SVO	SOV	VSO
Грусть [Sadness]	Немецкий [German]	0,22	0,00	0,00
	Испанский [Spanish]	1,00	0,89	0,75
	Китайский [Chinese]	0,77	0,57	0,75
	Хинди [Hindi]	0,40	0,33	0,33
	Татарский [Tatar]	1,00	0,75	0,75
Удивление [Surprise]	Немецкий [German]	0,00	0,00	0,00
	Испанский [Spanish]	0,73	0,67	0,50
	Китайский [Chinese]	0,00	0,00	0,00
	Хинди [Hindi]	0,83	0,80	0,80
	Татарский [Tatar]	0,60	0,67	0,50

В целом, на основании полученных результатов сложно сделать однозначные выводы по причине малого количества примеров для некоторых эмоций, например, для эмоций отвращения и удивления. Однако рассмотрим основные факторы, влияющие на работу модели в данных условиях. На общей картине видно, что наилучшее качество предсказания наблюдается у эмоций гнева и радости, для которых имеется наибольшее количество примеров, а также эти эмоции обычно наиболее ярко выражены в тексте за счет лексики и пунктуации. Для других же эмоций картина более неоднозначная.

Для эмоции гнева хорошее качество классификации наблюдается у всех моделей, при этом более высокое качество наблюдается у предложений с порядком SOV. При более близком рассмотрении результатов одним из факторов является позиция окончания предложения, в которой при таком порядке оказывается глагол, часто несущий эмоциональную оценку.

Представляется, что на эту позицию модель обращает наибольшее внимание, ср. следующие предложения:

- 1) *Ублюдки бесят меня на протяжении всего дня!!!!;*
- 2) *Ты меня бесишь!!! Очень сильно!!!;*
- 3) *Всё, этот говнюк меня раззадорил. Меня еще никто так не бесил.*



Первые два предложения правильно классифицируются только двумя из пяти моделей, тогда как третье предсказывается правильно уже четырьмя моделями, за исключением испанской. Все предложения имеют один и тот же глагол *бесить*, 1-е и 3-е предложения имеют также оценочные слова *ублюдки* и *говнюк*, и при этом только 3-е предложение не имеет восклицательных знаков, часто свойственных для эмоции гнева, но все равно распознается лучше двух остальных за счет финальной позиции глагола.

Эмоция радости стабильно хорошо распознается всеми моделями, однако качество классификации у татарской модели заметно хуже других. В данном случае татарская модель предсказывает неправильную эмоцию гнева даже для таких сравнительно простых предложений, как *Надо же, они подняли настроение!* и *Я сделала это!*, с которыми справляются все остальные модели. К сожалению, однозначную причину этого установить сложно, и, скорее всего, это вызвано особенностью обучающих данных.

Для эмоций отвращения, страха и удивления высокое качество классификации наблюдается у языков, для которых в обучающей выборке было представлено большое количество примеров с соответствующей эмоцией. Так, качество классификации эмоции отвращения является наивысшим у немецкой модели, содержащей в обучающей выборке 294 примера с данной эмоцией, тогда как качество классификации эмоции страха выше всего у испанской модели, содержащей в обучающей выборке 295 примеров с данной эмоцией. В то же время на качество переноса здесь также влияет близость языков, например, испанская модель единственная правильно классифицировала предложение *Я стал виртуальным человеком, что меня и стало пугать*, в котором единственным показателем эмоции является глагол *пугать*, при том, что для модели на языке хинди имеется большее количество обучающих примеров для эмоции страха.

Рис. 2. Графики качества классификации отдельных эмоций в русских тестовых примерах в зависимости от исходного языка обучения модели и порядка слов в тексте:

a – гнев; *b* – отвращение; *c* – страх; *d* – радость; *e* – грусть; *f* – удивление
1 – немецкий; 2 – испанский; 3 – китайский; 4 – хинди; 5 – татарский

Fig. 2. Graphs showing the classification performance for individual emotions in the Russian test dataset by the source language and the word order in the text:

a – anger; *b* – disgust; *c* – fear; *d* – joy; *e* – sadness; *f* – surprise
1 – German; 2 – Spanish; 3 – Chinese; 4 – Hindi; 5 – Tatar

Порядок слов здесь опять же может оказывать влияние в зависимости от того, что находится на конце предложения, например, предложение *Самое страшное в жизни – это когда твоя семья тебя не понимает*(((испанская модель классифицирует как содержащее эмоцию «грусть» из-за наличия на конце грустных скобок (ср. предложение *Блин! могу сидеть только в твиттере, а так хочется с тобой поговорить*(((или *чтобы ты мне написал!!!*, которое содержит на конце три восклицательных знака и которое та же испанская модель классифицирует с эмоцией радости).

Наконец, для эмоции грусти наблюдается несколько иная картина. Хотя все языки имеют достаточное большое количество обучающих примеров для данной эмоции, некоторые модели при переносе показывают крайне плохие результаты классификации, особенно немецкая модель. Это обусловлено в первую очередь тем, что основным способом выражения грусти в русских тестовых предложениях – это несколько идущих подряд круглых скобок: (((. Это приводит к тому, что лучше всего такие примеры классифицируют модели, обученные на более близких языках или языках, имеющих схожую традицию передачи эмоций в тексте. Что интересно, среди данных языков только татарский имеет обучающие примеры, также содержащие несколько идущих подряд круглых скобок, но при этом испанская модель все равно способна их распознавать даже с более высоким результатом. Основным средством выражения эмоции грусти в испанской выборке является такой смайл: :(. Он встречается в большинстве примеров с эмоцией грусти (в отличие от других языков, где не наблюдается специфических смайлов или эмодзи). Высокое качество переноса для испанского языка в данном случае указывает на высокую способность модели к генерализации.

Полученные результаты, таким образом, указывают на отсутствие значимого влияния базового порядка слов языка-источника на качество переноса обученных моделей, но при этом порядок слов все еще оказывает влияние на классификацию: в первую очередь на это влияют слова и эмоциональные маркеры, стоящие в конце предложения. Тем не менее, основное влияние на качество переноса языковых моделей по-прежнему оказывает количество и качество обучающих данных моделей, а также близость языка-источника и языка-цели.

8. Заключение

В данной работе был рассмотрен перенос языковых моделей, предварительно обученных на одном языке, на другой язык в рамках задачи автоматической классификации эмоции в текстах на основе данных, представленных на соревновании SemEval-2025.

В рамках практического исследования был проведен анализ влияния базового порядка слов на качество классификации языковых моделей, обученных на немецком, испанском, китайском, хинди и татарском языках и затем протестированных на данных русского языка. Для полученных результатов также была подсчитана статистическая значимость.

На основе анализа работы языковых моделей и имеющихся данных были описаны основные факторы, влияющие на качество переноса рассмотренных моделей, в том числе и влияние порядка слов. Одним из основных выводов является отсутствие значимого влияния базового порядка слов в языке-источнике, но наблюдается влияние слов и/или знаков пунктуации, стоящих в финальной позиции предложения.

Несмотря на это, имеющихся эмоциональных текстов недостаточно для проведения полноценного всестороннего исследования влияния базового порядка слов на качество переноса языковых моделей, и дальнейшим направлением работы представляется сбор большего объема данных, а также потенциально более систематизированный подход к их отбору в зависимости от порядка слов и наличия в них тех или иных эмоциональных маркеров.

Библиографический список / References

Anderson, 2001 – Anderson M.J. A new method for non-parametric multivariate analysis of variance. *Austral Ecology*. 2001. Vol. 26. No. 1. Pp. 32–46.

Conneau et al., 2020 – Conneau A., Khandelwal K., Goyal N. et al. Unsupervised cross-lingual representation learning at scale. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. D. Jurafsky, J. Chai, N. Schlueter, J. Tetreault (eds.). Association for Computational Linguistics, 2020. Pp. 8440–8451.

Cowie et al., 2001 – Cowie R., Douglas-Cowie E., Tsapatsoulis N. et al. Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*. 2001. Vol. 18. No. 1. Pp. 32–80.

Deshpande et al., 2022 – Deshpande A., Talukdar P., Narasimhan K. When is BERT multilingual? Isolating crucial ingredients for cross-lingual transfer. *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. M.C. de Marnette, I.V. Meza Ruiz (eds.). Seattle, USA, 2022. Pp. 3610–3623.

Devlin et al., 2019 – Devlin J., Chang M.W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. J. Burstein, C. Doran, T. Solorio (eds.). Minneapolis, Minnesota, 2019. Pp. 4171–4186.

Dryer, Haspelmath, 2013 – Dryer M.S., Haspelmath M. The world atlas of language structures online. *Zenodo*. Leipzig, 2024. DOI: 10.5281/zenodo.13950591.

Hansen et al., 2022 – Hansen L., Zhang Y.P., Wolf D. et al. A generalizable speech emotion recognition model reveals depression and remission. *Acta Psychiatrica Scandinavica*. 2022. Vol. 145. No. 2. Pp. 186–199.

Jackson et al., 2019 – Jackson J.K., Watts J., Henry T.R. et al. Emotion semantics show both cultural variation and universal structure. *Science*. 2019. Vol. 366. No. 6472. Pp. 1517–1522.

Kovacevic et al., 2022 – Kovacevic L., Bradic V., Melo G. et al. EZGlot. 2022. URL: <https://ezglot.com> (date accessed: 18.08.2025).

Lauscher et al., 2020 – Lauscher A., Ravishankar V., Vulić I., Glavaš G. From zero to hero: On the limitations of zero-shot language transfer with multilingual transformers. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. B. Webber, T. Cohn, Y. He, Y. Liu (eds.). Association for Computational Linguistics, 2020. Pp. 4483–4499.

Lefter, Jonker, 2017 – Lefter I., Jonker C.M. Aggression recognition using overlapping speech. *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. San Antonio, TX, 2017. Pp. 299–304.

Lin et al., 2019 – Lin Y.H., Chen C.Y., Lee J. et al. Choosing transfer languages for cross-lingual learning. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. A. Korhonen, D. Traum, L. Márquez (eds.). Florence, Italy, 2019. Pp. 3125–3135.

Littell et al., 2017 – Littell P., Mortensen D.R., Lin K. et al. URIEL and lang2vec: Representing languages as typological, geographical, and phylogenetic vectors. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. M. Lapata, P. Blunsom, A. Koller (eds.). Valencia, Spain, 2017. Pp. 8–14.

Liu et al., 2019 – Liu Y., Ott M., Goyal N. et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv.org*. 2020. DOI: 10.48550/arXiv.1907.11692.

Liu et al., 2020 – Liu C.L., Hsu T.Y., Chuang Y.S., Lee H.Y. A study of cross-lingual ability and language-specific information in multilingual BERT. *arXiv.org*. 2020. DOI: 10.48550/arXiv.2004.09205.

Mitsios et al., 2024 – Mitsios M., Vamvoukakis G., Maniati G. et al. Improved text emotion prediction using combined valence and arousal ordinal classification. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*. K. Duh, H. Gomez, S. Bethard (eds.). Mexico City, Mexico, 2024. Pp. 808–813.

Muhammad et al., 2025a – Muhammad S.H., Ousidhoum N., Abdulmumin I. et al. BRIGHTER: BRIDging the Gap in Human-Annotated Textual Emotion Recognition datasets for 28 languages. *arXiv.org*. 2025. DOI: 10.48550/arXiv.2502.11926.

Muhammad et al., 2025b – Muhammad S.H., Ousidhoum N., Abdulmumin I. et al. SemEval-2025 Task 11: Bridging the gap in text-based emotion detection. *arXiv.org*. 2025. DOI: 10.48550/arXiv.2503.07269.

Patil et al., 2022 – Patil V., Talukdar P., Sarawagi S. Overlap-based vocabulary generation improves cross-lingual transfer among related languages. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. S. Muresan, P. Nakov, A. Villavicencio (eds.). Dublin, Ireland, 2022. Pp. 219–233.

Philippy et al., 2023 – Philippy F., Guo S., Haddadan S. Towards a common understanding of contributing factors for cross-lingual transfer in multilingual language models: A review. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. A. Rogers, J. Boyd-Graber, N. Okazaki (eds.). Toronto, Canada, 2023. Pp. 5877–5891.

Pikuliak et al., 2021 – Pikuliak M., Šimko M., Bieliková M. Cross-lingual learning for text processing: A survey. *Expert Systems with Applications*. 2021. Vol. 165. No. 113765. DOI: 10.1016/j.eswa.2020.113765.

Radford et al., 2018 – Radford A., Narasimhan K., Salimans T., Sutskever I. Improving language understanding by generative pre-training. 2018. URL: <https://gwern.net/doc/www/s3-us-west-2.amazonaws.com/d73fdc5ffa8627bce44dcda2fc012da638ffb158.pdf> (date accessed: 18.08.2025).

Vaswani et al., 2017 – Vaswani A., Shazeer N., Parmar N. et al. Attention is all you need. *arXiv.org*. 2017. DOI: 10.48550/arXiv.1706.03762.

Wichmann et al., 2010 – Wichmann S., Holman E.W., Bakker D., Brown C.H. Evaluating linguistic distance measures. *Physica A: Statistical Mechanics and its Applications*. 2010. Vol. 389. No. 17. Pp. 3632–3639.

Wu et al., 2023 – Wu Z., Tamkin A., Papadimitriou I. Oolong: Investigating what makes transfer learning hard with controlled studies. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. H. Bouamor, J. Pino, K. Bali (eds.). Singapore, 2023. Pp. 3280–3289.

Статья поступила в редакцию 10.06.2025

The article was received 10.06.2025

Сведения об авторах / About the authors

Лемаев Владислав Игоревич – аспирант кафедры теоретической и прикладной лингвистики филологического факультета, Московский государственный университет имени М.В. Ломоносова

Vladislav I. Lemaev – PhD student at the Department of Theoretical and Applied Linguistics of the Faculty of Philology, Lomonosov Moscow State University, Moscow, Russian Federation

E-mail: vladzhkv98@mail.ru

Наталья Валентиновна Лукашевич – доктор технических наук, кандидат физико-математических наук; профессор кафедры теоретической и прикладной лингвистики филологического факультета, Московский государственный университет имени М.В. Ломоносова

Natalia V. Loukashevitch – Dr. Hab. in Technical Sciences, PhD in Mathematics; Professor at the Department of Theoretical and Applied Linguistics of the Faculty of Philology, Lomonosov Moscow State University, Moscow, Russian Federation

E-mail: louk_nat@mail.ru

DOI: 10.31862/2500-2953-2025-4-78-105

УДК 81.114.2:811.161.1

И.А. МельчукМонреальский университет,
Н3С3J7 г. Монреаль, Канада**Фразы типов**
Ивану не работается ~
Ивану хорошо работается:
два случая глагольной деривации

Русские фразы типов (1) *Ивану не работается* 'Пытаясь работать, Иван находится в таком внутреннем состоянии, которое препятствует тому, чтобы Иван работал' и (2) *Ивану хорошо работается* 'Работа Ивана идет хорошо' не являются реализациями некоей лингвоспецифической «конструкции»: в них выступают два разных производных глагола – *работаться*¹ и *работаться*², выражающих две различные словообразовательные единицы – дериватемы: *БЫТЬ_НЕРАСПОЛОЖЕН* и *ИДТИ_Z-ОВО*. Языковое явление, наблюдаемое в этих фразах, есть не что иное, как синхронная (= регулярная) деривация. Описываются семантические, синтаксические и морфологические свойства фраз типов (1) и (2) и демонстрируются их существенные различия. Приводятся формальные структуры фраз (1) и (2) на семантическом, глубинно-синтаксическом, поверхностно-синтаксическом и глубинно-морфологическом уровнях языкового представления. Определяется понятие дериватемы и предлагается описание обеих вышеуказанных дериватем в виде словарных статей, опирающееся на формальные структуры фраз (1) и (2); толкование дериватемы *ИДТИ_Z-ОВО* оказывается дизъюнктивным. Характеризуются некоторые типы фраз, которые семантически и формально похожи на фразы типов (1) и (2), но по существу отличны от них. Обсуждаются два существенных понятия (безличный глагол и критерий объединяющей сочетаемости для дизъюнктивных лексикографических толкований).

Ключевые слова: русский язык, деривация (= словообразование), синхронная деривация, дериватема, словарная статья дериватемы, безличный глагол, дизъюнктивное толкование

Благодарности. Первые варианты данной статьи читали и комментировали Валентина Апресян, Игорь Богуславский, Лидия Иорданская, Светлана Крылосова, Ясмينا Миличевич, Елена Саввина, Валентин Томашпольский и Антон Циммерлинг, а окончательный вариант – два анонимные рецензента журнала. Ваши замечания и предложения помогли мне улучшить изложение. Спасибо вам от всей души!

Для ЦИТИРОВАНИЯ: Мельчук И.А. Фразы типов *Ивану не работается* ~ *Ивану хорошо работается*: два случая глагольной деривации // Рема. Rhema. 2025. № 4. С. 78–105. DOI: 10.31862/2500-2953-2025-4-78-105

DOI: 10.31862/2500-2953-2025-4-78-105

I. Mel'čuk

University of Montreal,
Montreal, H3C3J7, Canada

Sentences of types *Ivanu ne rabotaetsja* 'Ivan is undisposed to work' ~ *Ivanu xorošo rabotaetsja* 'Ivan's work is going well': Two cases of verbal derivation

Russian sentences of types *Ivanu ne rabotaetsja* 'Ivan is undisposed to work' = 'Trying to work, Ivan is in such an internal state that is impeding his working' ~ *Ivanu xorošo rabotaetsja* 'Ivan's work is going well' do not represent language-specific constructions but feature two different derived verbs – *rabotat'sja1* '[to] work.itself1' and *rabotat'sja2* '[to] work.itself2', which express special derivational units – the derivatememes *BYT_NERASPOLOŽEN* '[to] be undisposed' and *IDTI_Z-OVO* '[to] go Z-ily'. The linguistic phenomenon observed in these sentences is nothing else but synchronic (= regular) derivation. The paper offers a description of the semantic, syntactic and

morphological properties of sentences of types (1) and (2), while establishing their essential differences. The formal structures of sentences (1) and (2) are given on the semantic, deep-syntactic, surface-syntactic and deep-morphological levels of linguistic representation. A definition of the notion 'derivateme' is presented, followed by a rigorous description of both above derivatemes in the form of lexicographic entries, based on the formal structures of sentences (1) and (2); the lexicographic definition of the derivateme *IDTI_Z-OVO* '[to] go Z-ily' is disjunctive. Several sentence types of Russian are characterized which semantically and formally resemble sentences of types (1) and (2), but are essentially different. Two relevant notions are discussed: impersonal verb and criterion of unifying cooccurrence for disjunctive lexicographic definitions.

Key words: Russian language, derivation, synchronic derivation, derivateme, lexical entry of a derivateme, impersonal verb, disjunctive lexicographic definition

Acknowledgments. Early drafts of this article were read and commented on by Valentina Apresjan, Igor Boguslavsky, Lidija Iordanskaja, Svetlana Krylosova, Jasmina Miličević, Elena Savvina, Valentin Tomašpolskij, and Anton Zimmerling, and the final version was scrutinized by two anonymous reviewers. Your comments and suggestions helped me improve the presentation. Thank you from the bottom of my heart!

FOR CITATION: Melčuk I. Sentences of types *Ivanu ne rabotaetsja* 'Ivan is undisposed to work' ~ *Ivanu xorošo rabotaetsja* 'Ivan's work is going well': Two cases of verbal derivation. *Rhema*. 2025. No. 4. Pp. 78–105. DOI: 10.31862/2500-2953-2025-4-78-105

Грете не естся! Грете не спится!

С. Горелик

1. Объект, задачи и теоретические рамки статьи

В статье рассматриваются фразы типа (1) и типа (2):

(1) *Ивану не работается.*

(2) *Ивану хорошо работается.*

В современной русистике такие фразы традиционно считаются реализациями особой лингвоспецифической конструкции, которой посвящена обширная литература. Настоящая статья опирается, главным образом, на следующие работы: [Wierzbicka, 1992, pp. 424–428; Булыгина, Шмелёв, 1997, с. 105–108; Апресян, 2005, с. 8–10; Фичи, 2011; Фичи,

Жукова, 2012; Летучий, 2014]; в них указываются предшествующие публикации, приводятся языковые факты, характеризующие эту конструкцию, и высказываются важные соображения по поводу их интерпретации. Однако – насколько мне известно – формального описания данной конструкции дано не было. Более того, сам термин «конструкция», употребляемый для ссылки на фразы указанных типов, не имеет строгого определения, так что непонятно даже, представляют ли фразы (1) и (2) одну и ту же «конструкцию» или две разные. Поэтому в этой статье ставятся три следующие задачи.

(А) Показать, что во фразах (1) и (2) не имеет места некая особая языковая единица – «конструкция», которую можно было бы строго определить и которая точно идентифицировала бы лингвистическую природу этих фраз. В действительности, в них выступают два разных производных глагола – РАБОТАТЬСЯ¹ и РАБОТАТЬСЯ², выражающих две различные словообразовательные единицы – дериватемы: *БЫТЬ_НЕРАСПОЛОЖЕН* и *ИДТИ_З-ОВО*¹. Для демонстрации этого приводится лингвистическая характеристика данных фраз, а также определяется понятие дериватемы (раздел 2).

(В) Дать формальные структуры фраз (1) и (2) на четырех уровнях языкового представления – на семантическом [Сем-], глубинно-синтаксическом [ГСинт-], поверхностно-синтаксическом [ПСинт-] и глубинно-морфологическом [ГМорф-] (раздел 3).

(С) Предложить строгое описание обеих названных дериватем, опирающееся на формальные структуры фраз (1) и (2), – в виде словарных статей (раздел 4).

Затем бегло характеризуются некоторые типы фраз, которые семантически и формально похожи на фразы типов (1) и (2), а потому могут быть легко спутаны с ними (раздел 5).

Результаты данного исследования кратко суммируются в Заключение (раздел 6).

Два существенных понятия (безличный глагол и критерий объединяющей сочетаемости для дизъюнктивных лексикографических толкований) обсуждаются в Приложении.

Указанные задачи решаются в рамках подхода «Смысл – Текст». Это означает следующее:

- языковые факты трактуются в направлении от смысла к тексту, т.е. обсуждается, как некий заданный смысл выражается соответствующим текстом, а не как извлечь смысл из некоего заданного текста;
- формальные представления высказываний строятся в терминах семантических и синтаксических зависимостей; различаются семь уровней

¹ Разрядкой набираются особо релевантные термины при первом упоминании, а шрифтом *PT SANS ITALICS* – имена дериватем.

этих представлений, четыре из которых упомянуты выше (три остальные – поверхностно-морфологический, глубинно-фонологический и поверхностно-фонологический – для данной статьи нерелевантны); – инструментом описания языкового знака является его словарная статья.

Основные понятия и формализмы подхода «Смысл-Текст» предполагаются известными.

В статье используются результаты работы [Mel'čuk, Milićević, to appear].

2. Фразы типов (1) и (2)

Прежде всего, дается систематическое описание языковых свойств фраз (1) и (2) и подобных им (раздел 2.1). На основании этого описания делается вывод, что в таких фразах наблюдается два типа продуктивного глагольного словообразования, или, как предлагается говорить в дальнейшем, синхронной деривации (раздел 2.2). Затем рассматривается центральное понятие синхронной деривации – дериватема (раздел 2.3).

2.1. Фразы типов (1) и (2): языковые факты

Описание фраз обоих типов основано на известных фактах, почерпнутых, в частности, в вышеупомянутых работах.

Тип фразы задается описанием ее следующих свойств:

- (i) Сказуемое: глагол V-ся
- (ii) Семантика: значение глагола V-ся
- (iii) Синтаксис: актанты глагола V-ся
- (iv) Морфология: флективные формы глагола V-ся

2.1.1. Фразы типа (1)

(i) **Сказуемое** (= вершинный глагол) фразы типа (1) – безличный возвратный глагол V-ся. Такой глагол не имеет ГСинт-актанта **I**, так что на ПСинт-уровне фраза получает нулевое подлежащее-«пустышку» \emptyset _{(N, СРЕД)ЕД.}^{ПУСТОЕ}. (О термине «безличный глагол» см. Приложение.)

(ii) **Семантика**. Во фразе типа (1) говорится о **внутреннем состоянии** человека X. Глагол V-ся, обозначаемый как Y, имеет здесь следующее значение:

‘X-у не Y-ается’ = ‘[[Пытаясь Y-ать,]]² X находится в таком внутреннем состоянии, которое препятствует тому, чтобы X Y-ал’

² Скобки [[...]] выделяют пресуппозицию.

Семантические компоненты в предлагаемых толкованиях принимаются как интуитивно достаточно очевидные (т.е. не разлагаются); это касается и нашего центрального понятия – ‘внутреннее состояние [человека]’; имеется в виду психологическое и/или физи(ологи)ческое состояние³.

Семантический компонент ‘препятствует’ в разложении смысла ‘Х-у не Y-ается’ существенен. Он не может быть заменен, например, на ‘не позволяет’, т.к. *Мне не работается* не имплицитно ни того, что я не работаю, ни того, что я работаю:

(3) а. *Мне совсем не работалось, так что я даже и не приступил к делу.*

б. *Мне совсем не работалось, но я заставил себя сесть за стол и поработал три часа.*

А семантический компонент ‘не позволяет Х-у Y-ать’⁴ имплицитновал бы, что Y X-а не имеет места. При таком компоненте в смысле глагола V-ся фраза (3б) была бы противоречива, тогда как в действительности она совершенно нормальна.

Пресуппозиция в рассматриваемом толковании отражает два факта.

Во-первых, то, что Y – это действие или состояние, контролируемое X-ом (хотя бы в некоторой степени). Так, фраза (4) естественна в устах актера, старающегося изобразить хромоту/заикание:

(4) *Нет, что-то сегодня мне не хромается / не заикается.*

А во-вторых, то, что фраза *Мне не работается*, как только что сказано, не сообщает, работал ли я хоть немного или нет. (Смысл ‘X пытается Y-ать’ не имплицитно ни того, что X Y-ает, ни того, что X не Y-ает: предложение *Иван попытался подняться... можно продолжить и ...но не смог, и ...и встал на ноги.*)

Особенность смысла ‘Х-у не Y-ается’ состоит в следующем: положительное утверждение – «X находится в определенном состоянии» – выражается с участием отрицания *не* (*Ивану не работается*), а соответствующее отрицательное утверждение – «X не находится в таком состоянии» – должно выражаться посредством **неупотребления** частицы *не* (*Ивану работается*); такое состояние дел весьма анти-интуитивно.

³ В книге [Зализняк, 1992] детально анализируются значения большого числа русских лексем, выражающих внутренние состояния человека; однако состояния того типа, о котором идет речь здесь, не рассматриваются.

⁴ ‘Не позволять’ имеет здесь значение, представленное во фразе *Трение не позволяет объекту сдвинуться.*

Это одна из причин того, что фразы типа (1) так плохо поддаются семантическому анализу.

(iii) **Синтаксис.** Прототипически, во фразе типа (1) глагол *V-ся* имеет один актанта: Сем-актанта 'X' и его синтаксические воплощения. 'X' может обозначать только человека или же выражать персонификацию. (Это следует из значения глагола.)

В ряде случаев глагол *V-ся* может перенять от исходного глагола *V* некоторый актанта этого последнего, отличный от 'X', но лишь такой, который реализуется в ПСинт-структуре косвенным дополнением или обстоятельством⁵; прямое дополнение у *V-ся* совершенно исключается (5а), но и некоторые косвенные дополнения затруднены (5б):

- (5) а. *Не пишется мне этой ручкой / о здешней жизни.* ~
 **Не пишется мне эту статью.*
 б. *Не говорится мне сегодня по-французски.* ~
 ?*Не говорится мне сегодня с Иваном.*

Сем-актанта 1 'X' глагола *V-ся* выражается как ГСинт-актанта II / ПСинт-актанта «непрямой объект N_{ДАТ}». Этим нарушается стандартное и более естественное соответствие между актантами [-А] трех уровней, доминирующее в языках номинативно-аккузативного строя: [Сема 'X₁' ⇔ ГСинтА I ⇔ ПСинтА «подлежащее»]. При таком соответствии центральный Сем-актанта 'X₁' выражается центральными же Синт-актантами. А во фразе типа (1) центральный Сем-актанта глагола *V-ся* выражается посредством второстепенных ГСинт- и ПСинт-актантов.

NB Указанное расхождение семантики и синтаксиса в подобных конструкциях с возвратным глаголом типично для русского и других славянских языков, см. Заключение, раздел 6.

(iv) **Морфология.** Во фразе типа (1) глагол *V-ся* допускает только несов вид; будучи безличным, он имеет только формы 3-го лица ед. числа и инфинитив:

- (6) а. **Ивану не поработалось.*
 б. *Может же человеку иногда не работаться!*

2.1.2. Фразы типа (2)

(i) **Сказуемое** такой фразы, как и в случае фраз типа (1), – безличный возвратный глагол *V-ся*.

⁵ Напомню, что ГСинт-актанта может выражаться на ПСинт-уровне как дополнением, так и обстоятельством/модификатором.

(ii) **Семантика.** Во фразе типа (2) либо (в большинстве случаев) дается оценка – хорошо/плохо – **протекания действия Y, совершаемого X-ом**, либо (в ограниченном числе случаев) сообщается о **внутреннем состоянии X-а** в связи с его действием Y. Глагол V-ся, обозначающий действие Y, имеет здесь следующее значение:

‘X-у Y-ается Z-ово’ = ‘Y-ание X-а идет Z₁-ово или так, что Y-ать X-у – Z₂-ово’. (О дизъюнктивном толковании глагола V-ся [= Y] см. в конце данного подраздела.)

(iii) **Синтаксис.** Прототипически, во фразе типа (2) глагол V-ся имеет два актанта: Сем-актант ‘X’, который может либо обозначать человека, либо выражать персонификацию, и Сем-актант ‘Z’, о котором см. ниже (и, разумеется, синтаксические воплощения этих Сем-актантов).

Как и во фразах типа (1), во фразах типа (2) глагол V-ся может унаследовать от исходного глагола V некоторый «дополнительный» актант – на тех же самых условиях:

(7) а. *Этой ручкой / О здешней жизни мне пишется плохо. ~*
**Эту статью мне пишется плохо.*

б. *Сегодня мне говорится по-французски легко. ~*
?*Сегодня мне говорится с Иваном легко.*

Центральный Сем-актант ‘X₁’ глагола V-ся во фразах типа (2) также выражается как второстепенный ГСинт-актант II / ПСинт-актант «непрямой объект N_{ДАТ}», т.е. и в таких фразах нарушается обычное семантико-синтаксическое соответствие между актантами разных уровней.

Сем-актант ‘X’ во фразах типа (2) также может обозначать только человека или же выражать персонификацию. Но у V-ся в этих фразах имеется, как указано, еще и второй Сем-актант – ‘Z’, с которым связано следующее осложнение.

Либо ‘Z’ – это ‘Z₁’, некий оценочный (= эвалюативный) смысл, который характеризует смысл глагола V (т.е. смысл ‘Y’) безотносительно к Агенсу: ‘Z₁’ сообщает, как именно протекает факт Y. В таком случае ‘Z₁’ может быть каким угодно: этот актант реализуется любым наречием или любым наречиеобразным словосочетанием (в том числе – свободным), подходящим по смыслу (‘хорошо’ ~ ‘плохо’). Например:

(8) *Ивану работалось чудесно_{Z₁} <со скрипом_{Z₁}; лучше_{Z₁}, чем в Тимбукту; так мерзко_{Z₁}, что и вообразить невозможно>.*

Либо ‘Z’ – это ‘Z₂’, некий смысл, который характеризует смысл глагола V (т.е. ‘Y’) с точки зрения того, какво Агенсу X делать Y:

(9) а. *Ивану работалось весело* <*скучно; легко, тяжело, с трудом/*трудно, с удовольствием>.

б. *Ивану гулялось приятно* <уютно>.

Набор наречий, несущих смысл ‘Z₂’, т.е. наречий внутреннего состояния, ограничен лексически: те наречия, которые могут сочетаться с V-ся во фразе типа (2), должны задаваться списком (в словарной статье данной дериватемы, см. раздел 4.2).

Некоторые из этих наречий (далеко не все!) семантически эквивалентны предикативам внутреннего состояния:

(10) *Ивану работалось весело* ^{наречие} <с трудом>. ≡
Ивану было весело _{пред.сост} <трудно> *работать*.

Следует особо отметить, что оценочные наречия и наречия внутреннего состояния как зависимые глагола V-ся могут сочиняться:

(11) а. *Ивану работалось хорошо и легко* <хорошо, хотя и с трудом>.

б. *Ивану работалось плохо, но хотя бы легко*.

NB Более того, в ряде случаев оценочное наречие и наречие внутреннего состояния совпадают по форме, так что в реальной фразе – вне достаточного контекста – различить их невозможно: *Ивану там работалось хорошо* может значить, и что работа Ивана там шла хорошо, и что ему там было хорошо работать; к тому же эти оба смысла очень близки и легко перетекают один в другой.

Из этого – в соответствии с лексикографическим критерием объединяющей сочетаемости (см. Приложение) – следует, что толкование глагола V-ся должно быть дизъюнктивным.

(iv) **Морфология.** Во фразах типа (2) глагол V-ся допускает оба вида: см. (17b).

2.1.3. Различия между фразами типов (1) и (2)

Как показано в разделах 2.1.1 и 2.1.2, фразы типа (1) и фразы типа (2) воплощают так называемую «дательно-безлично-возвратную конструкцию» и имеют значительные семантические и синтактико-морфологические сходства:

- их вершинный глагол-сказуемое – безличный возвратный глагол V-ся;
- во фразах типа (1) всегда, а во фразах типа (2) во многих случаях – с Сем-актантом ‘Z₂’ – характеризуется внутреннее состояние человека X: его расположенность/нерасположенность к действию Y;
- глагол V-ся имеет в качестве ПСинт-актанта существительное в дательном падеже, обозначающее человека – носителя состояния или деятеля – носителя состояния.

В то же время между фразами типов (1) и (2) наблюдаются следующие восемь серьезных различий.

Семантические различия

1. Во фразе (1) говорится о **внутреннем состоянии** Ивана, которое препятствует тому, чтобы Иван работал, а фраза (2) сообщает, **как протекает работа** Ивана:

‘Ивану не работается’ = ‘[[Пытаясь работать,]] Иван находится в таком внутреннем состоянии, которое препятствует тому, чтобы Иван работал’

‘Ивану хорошо работается’ = ‘Работа Ивана идет хорошо’⁶

2. Фраза (2) имплицитно, что Иван хотя бы какое-то время работал, а фраза (1) этой импликацией не имеет: она подходит и для ситуации, когда Иван работает, и для ситуации, когда Иван не работает.

Семантико-синтаксические различия

3. Во фразе (1) глагол имеет один Сем-актант – Экспериментер (‘Иван’), а во фразе (2) актантов у глагола – два: Агенса (‘Иван’) и Оценки протекания действия/состояния Агенса (‘хорошо’/‘весело’).

4. Во фразе (1) может быть опущено отрицание *не*, а во фразе (2) – наречие *хорошо*; в результате получается как бы одна и та же фраза (12):

(12) *Ивану работается.*

Однако под этим поверхностным «опущением» скрываются два языковых явления различной природы.

Фраза (1) без *не* выражает прямо обратный смысл, т.е. исходную Сем-структуру с отрицанием: ‘Иван **не** находится во внутреннем состоянии, которое препятствует тому, чтобы Иван работал’. И этот смысл может быть выражен естественным образом – обычно в контексте контраста:

(13) {– *Мне не работается...*} – *А Ивану работается!*
 <– *А Ивану работается, что ли?*>

Здесь глагол *не работается* ведет себя как типичный элемент с отрицательной полярностью, ср. **не пошевелить** пальцем, **гроша ломаного не стоит**, **не идти в сравнение** и т.п. (указание И. Богуславского).

Фраза же (2) без *хорошо* имеет все тот же смысл, что и с ним, т.е. смысл ‘хорошо’ по-прежнему наличествует в исходной Сем-структуре, но выражающее его наречие подвергается эллипсису на поверхности.

Таким образом, выражение в (12) является означающим двух разных фраз, соответствующих двум разным Сем-структурам.

⁶ Глагол *идти* здесь – это лексема *идти*_{13.3} ‘имеет место ситуация А1, развитие которой оценивается Говорящим как А2’ (*Бизнес*_{А1} *идет неважно*_{А2}), описанная в [Апресян, 2023, с. 53].

- (17) а. *Ивану не работалось.* ~ **Ивану не поработалось.*
 б. *Ивану хорошо работалось.* ~ *Ивану хорошо поработалось.*

Вышеперечисленные различия ясно показывают, что во фразах (1) и (2) – и, разумеется, во всех фразах этих типов – имеют место два разных языковых феномена. Иначе говоря, в них представлены не одна, а две разные «конструкции», которые должны рассматриваться по отдельности⁸. Как будет показано, эти «конструкции» суть производные глаголы РАБОТАТЬСЯ₁ и РАБОТАТЬСЯ₂ со своими актантами.

2.2. Фразы (1) и (2): не «конструкция», а деривация

Термин «конструкция» употребляется в сегодняшней лингвистике в очень широком, расплывчатом смысле: автор «Грамматики конструкций» Адель Гольдберг понимает под этим термином любые единицы, которые должны быть как-то где-то эксплицитно зафиксированы в описании языка: простые слова, сложные и производные слова, полнозначные префиксы (типа анти- или мини-), идиомы и всевозможные синтаксические образования – такие, как дитранзитивные и пассивные конструкции, конструкции топикализации, конструкции придаточных предложений и т.д. [Goldberg, Suttle, 2010]. Тем самым, сказать про какое-то выражение, что это конструкция, – значит ничего не сказать о его природе, т.е. никак не охарактеризовать его как определенную лингвистическую сущность: любое многокомпонентное выражение есть конструкция. В назывании фраз типов (1) и (2) *конструкциями* нет ничего ошибочного, но такое название абсолютно неинформативно.

Из сказанного в разделе 2.1 вытекает, что в рассматриваемых фразах выступают разные производные глаголы V-ся₁ и V-ся₂, образованные от исходного глагола V посредством суффиксов -ся₁ и -ся₂⁹. Здесь наблюдается продуктивное словообразование, или п р о д у к т и в н а я д е р и в а ц и я. Эта идея не нова; так, в статье Ю.Д. Апресяна читаем: «Как известно, <эта конструкция> формируется глаголами несов на -ся, **производными** (выделено мною. – И.М.) от нетранзитивных глаголов

⁸ Эта точка зрения проведена в [Wierzbicka, 1992, pp. 424–426] и последовательно отстаивается в [Фичи, 2011; Фичи, Жукова, 2012]; ср. также обсуждение «несимметричности положительных и отрицательных форм» в [Иванова, 2022, с. 438–439].

⁹ Суффикс -ся функционирует в русском языке, прежде всего, как флективный (= словоизменятельный) показатель: так, он выражает граммы пассива и рефлексива. Однако использование флективных аффиксов – а тем самым, граммем – для деривационных целей хорошо известно. Например, в русском языке граммема мн (число существительных) присоединяется к мужской фамилии L для выражения смысла ‘семья мужчины L’, тем самым производя новую лексему: Ивановы, Петровы, Сидоровы; точно так же обстоит дело во многих других языках. Еще один пример: в испанском граммема мн при названии продукта L плюс имя собственное человека X выражает сложный смысл ‘компания X-а, поставяющая на рынок продукт L’: например, *Carnes_L Cesáreo Gómez_X* букв. ‘Мясa Cесарео Гомес’ = ‘Мясопродукты С. Гомеса’.

или абсолютивных употреблений транзитивных глаголов, обычно в сочетании с дательным падежом экспериенсера-агенса» [Апресян, 2005, с. 8]. Нововведением представляется перенос центра тяжести обсуждения на проблемы деривации. Более точно: описание рассматриваемых конструкций сводится к описанию специальных языковых знаков, на которых зиждется деривация, а именно – к описанию дериватем.

Оба рассматриваемые производные глагола характеризуют внутреннюю расположенность/нерасположенность человека к совершению определенного действия. Их удобно называть по-английски *dispositionals* (*disposition* ≈ ‘расположенность’), а по-русски (за неимением лучшего термина) – *диспозиционалами*. Диспозиционалы существуют во всех славянских языках (хотя и с разными значениями и разными синтактико-морфологическими свойствами); в силу этого они представляют значительный интерес для лингвистики. Ниже рассматриваются только два русских диспозиционала, точнее – две русские диспозициональные дериватемы.

2.3. Дериватема

Дериватемы занимают в данной статье центральное место, но само понятие дериватемы не является широко употребляемым; поэтому необходимо дать его определение. (Здесь приводится отредактированный вариант определения, предложенного в [Mel’čuk, 1993, p. 288, Définition I.32; Мельчук, 1997, с. 272, Определение I.32]. Суть исходного определения полностью сохранена, но его формулировка улучшена.)

Определение: дериватема

Дериватемой называется языковой знак, такой что:

- его означаемое не является флективным (= словоизменительным), т.е. обязательно выражаемым¹⁰;
- его означающее – одно из тех формальных средств, какие используются для маркировки флективных означаемых, т.е., чаще всего, аффикс или грамматическая лексема («служебное слово»).

Так как в данной статье речь идет исключительно о синхронных дериватемах, фигурирующих в рамках синхронной деривации, прилагательное *синхронная* может опускаться: термин «дериватема» употребляется здесь только в смысле ‘синхронная дериватема’¹¹.

¹⁰ Флективное означаемое – это означаемое грамеммы.

¹¹ Диахронная, т.е. непродуктивная, деривация – с фразеологизацией производных лексем – оставлена в стороне. Строго говоря, с синхронической точки зрения непродуктивное производное является простой лексемой, помещаемой в словарь, независимо от того, насколько говорящие ощущают ее историческую производность – эту, так сказать, живую этимологию. Например, глагол *икаться* (*Х-у икается* значит ‘Х испытывает позыв к иканию’) есть синхронически простой безличный возвратный глагол. См. раздел 5.

Всякая дериватема *DER* есть исходный элемент регулярной деривации (= словообразования). Термин *регулярный* в данном случае имеет следующее значение.

1. С точки зрения смысла, все лексемы $DER \oplus L_i$, т.е. производные лексемы с дериватемой *DER*, композиционны: регулярные производные не фразеологизуются.

Символ \oplus обозначает (мета)операцию языкового объединения, которая соединяет языковые знаки по компонентам: означаемое знака А с означаемым знака В, означающее знака А с означающим знака В, и синтактику знака А с синтактикой знака В. В каждом языке операция \oplus воплощает общие правила, т.е. грамматику, этого языка. (Об операции \oplus см. [Мельчук, 1997, с. 129–136].)

2. С точки зрения формы, все лексемы вида $DER \oplus L_i$ строятся по общим правилам¹².

Такие дериватемы, как *ЖИТЕЛЬ* и *ДЕТЕНЬШ*, должны задаваться посредством нестандартных лексических функций [Mel'čuk, 2023, р. 227–228] в словарных статьях базовых лексем. Здесь, однако, нет возможности обсуждать вопрос о соотношении дериватем с парадигматическими лексическими функциями.

3. С точки зрения сочетаемости, дериватема *DER* присоединяется к любой лексеме подходящего класса (заданного в описании самой дериватемы), включая потенциальные неологизмы; *DER* продуктивна по определению.

Продуктивность дериватемы означает, что занести в словарь **все** продуктивно производные лексемы в принципе невозможно: их множество открыто. Разумеется, многие из них могут или даже должны храниться в словаре. Один из наиболее очевидных резонов для этого – наличие у производной лексемы $DER \oplus L$ фразеологизованного кополисема¹³

¹² Как это типично для естественного языка, и здесь тоже дело не обходится без серьезного осложнения. Продуктивная дериватема может иметь целый ряд формально нерегулярных выражений. Точнее говоря, в достаточно многочисленных случаях либо (i) распределение ее показателей не описывается общими правилами, либо (ii) ее производная лексема образуется от супплетивной основы базовой лексемы.

– Случай (i) иллюстрируется дериватемой *ЖИТЕЛЬ* (города): Тула ~ тул+як, Москва ~ москв+ич, Севастополь ~ севастопол+ец, Неаполь ~ неапол+итанец, Киев ~ киев+лянин, Ржев ~ ржев+итянин, и т.д. Соответствующий суффикс необходимо указывать в словарной статье базовой лексемы (т.е. при названии города).

– Случай (ii) соответствует дериватеме *ДЕТЕНЬШ*. Она выражается суффиксом -ёнок (лев \Leftarrow лев+ёнок, медведь \Leftarrow медвеж+ёнок и т.д.); однако названия детенышей многих домашних животных используют супплетивные основы базовых лексем: корова \Leftarrow тел+ёнок (*коров+ёнок), собака \Leftarrow щ+ёнок, курица \Leftarrow цыпл+ёнок и т.д.

¹³ Кополисемами называются лексемы, принадлежащие к одной полисемичной вокабуле, т.е. лексемы, означаемые которых имеют общие существенные семантические компоненты: например, БУМАГА 'материал ...' и БУМАГА 'официальный документ ...' суть кополисемы.

и/или омонима. Так, например, производная лексема куколка¹_I, регулярное уменьшительно-ласкательное производное от кукла, имеет кополисем куколка¹_{II} ‘миниатюрная красивая девочка или молодая женщина’ (и омоним куколка² ‘личинка насекомого в коконе’); поэтому лексема куколка¹_I должна указываться в вокабуле куколка¹. Однако в общем случае продуктивно производная лексема в словарь не заносится.

NB Тем самым, продуктивно производная лексема $DER \oplus L$, вообще говоря, не имеет обычной словарной статьи. Она получает всю релевантную словарную информацию путем регулярного комбинирования – посредством операции \oplus – сведений из словарной статьи ее базовой (= производящей) лексемы L и сведений из словарной статьи дериватемы DER .

В описании языка дериватема DER обозначается одним из двух способов – в зависимости от ее выражения в данном языке.

1. Дериватема DER выражается в данном языке таким показателем, в частности – аффиксом, который закреплен исключительно за ней и легко идентифицируется носителем языка. Подобным аффиксом в русском является, например, глагольный префикс недо- ‘не закончить полностью’ (*недочитать рассказ, недопить чай, недоплыть до мыса*). Соответствующая дериватема может быть названа просто *недо-*.

2. Дериватема DER не выражается в данном языке одним «удобным» аффиксом. Хороший пример – название жителя города: лондон+**ец**, париж+**анин**, москв+**ич**, хив+**инец**, одесс+**ит** (см. сноску 12). Для такой дериватемы необходимо ввести искусственное имя: например, *ЖИТЕЛЬ*.

В ГСинт-структуре дериватема DER помещается на том же узле, что и производящая лексема $L_{(V)}$; дериватема присоединяется к производящей лексеме посредством символа \oplus : *ЖИТЕЛЬ* \oplus Лондон, *ЖИТЕЛЬ* \oplus Париж и т.д.

Вот три определяющих свойства дериватем:

1) они не **обязательны** (= не навязываются языком), т.е. в этом отношении они трактуются языком как лексемы (а не как граммемы);

2) они выражаются **теми же формальными средствами, что и граммемы** (аффиксами и грамматическими лексемами) – т.е. в этом отношении они трактуются языком как граммемы (а не как лексемы);

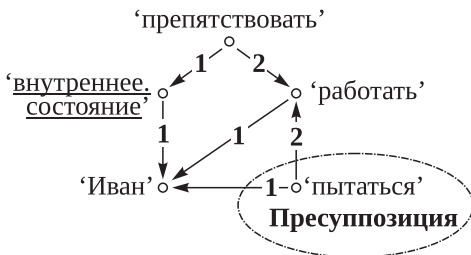
3) их выражение **квазирегулярно**: оно описывается достаточно общими правилами, но с большим количеством отклонений, каковые – в отличие от ситуации с граммемами – более целесообразно задавать не в виде списков исключений при правилах, а в виде указаний при соответствующих производящих лексемах, т.е. в словарных статьях этих последних. В этом проявляется родство дериватем с парадигматическими лексическими функциями (отмеченное в сноске 12).

Дериватемы представляют собой промежуточный случай между граммемами и лексемами.

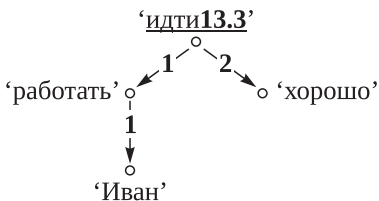
3. Формальные структуры фраз (1) и (2)

СемС

(1) *Ивану не работается.*

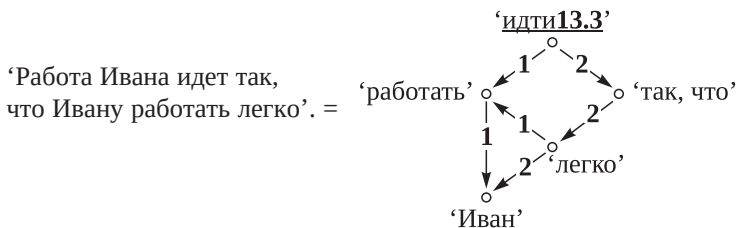


(2) *Ивану хорошо работается.*



- NB** 1. Ради простоты данные Сем-структуры не отражают семантические граммы (число существительных; вид, наклонение и время глагола).
 2. Семантема 'внутреннее.состояние' является сокращением для конфигурации семантем 'находиться в таком внутреннем состоянии, которое...'
 3. Фразы типа (2) могут иметь и другую Сем-структуру – если во фразе употреблено не оценочное наречие, а наречие-предикатив состояния, которое характеризует внутреннее состояние Агенса, связанное с его деятельностью (раздел 2.1.2, пункт (iii)); ниже представлена эта альтернативная структура.

Сем-структура фразы *Ивану работается легко*:

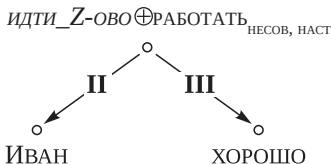


ГСинтС

(1) *Ивану не работается.*



(2) *Ивану хорошо работается.*



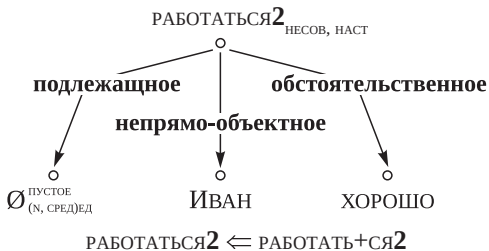
У обоих производных глаголов нет ГСинт-актанта **I**, т.е. они безличные.

ПСинтС

(1) *Ивану не работается.*



(2) *Ивану хорошо работается.*



ГМорфС

ИВАН_{ДАТ} НЕ РАБОТАТЬСЯ₁_{НЕСОВ, НАСТ, 3, ЕД}

ИВАН_{ДАТ} ХОРОШО РАБОТАТЬСЯ₂_{НЕСОВ, НАСТ, 3, ЕД}

4. Словарные статьи дериватем

БЫТЬ_НЕРАСПОЛОЖЕН и *ИДТИ_З-ОВО*

4.1. Словарная статья дериватемы *БЫТЬ_НЕРАСПОЛОЖЕН*

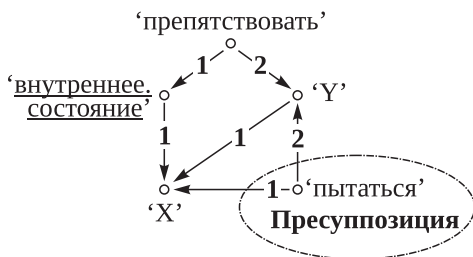
БЫТЬ_НЕРАСПОЛОЖЕН ⊕ L_(V)

– Означаемое

Текстовая версия

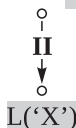
‘[[Пытаясь Y-ать,]] X находится в таком внутреннем состоянии, которое препятствует тому, чтобы X Y-ал’

Формальная версия



– Означающее (на двух уровнях: в ГСинт-структуре и в ПСинт-структуре)

• В ГСинт-структуре: *БЫТЬ_НЕРАСПОЛОЖЕН* ⊕ L_(V) | L_(V) = L('Y')(V)



• В ПСинт-структуре: L_(V) сЯ1 | L_(V) + сЯ1 = L_(V) сЯ1
 ○ — рестриктивное —> ○ НЕ

Как и в правилах соответствия между языковыми представлениями смежных уровней, в словарных статьях дериватем серая заливка обозначает контекст: элемент, который не является компонентом означающего дериватемы, а задает ее структурную позицию по отношению к основе производящей леммы и/или к ее актантам.

– Синтактика

• Самой дериватемы *БЫТЬ_НЕРАСПОЛОЖЕН*:

L_(V), т.е. мишень дериватемы, – преимущественно глагол действия в НЕСОВ виде¹⁴.

¹⁴ Это условие нуждается в серьезных уточнениях; данное замечание относится и к соответствующему условию в словарной статье дериватемы *ИДТИ_З-ОВО*.

• **Производного глагола** $L_{(V)}\text{СЯ1}$:

Флективные формы

Глагол $L_{(V)}\text{СЯ1}$ имеет только несов, только 3-е лицо ед. числа и инфинитив.

Модель управления

—	‘X’ ↔ II
1. Ø ^{ПУСТОЕ} _{(N, СРЕД.)ЕД}	1. N _{ДАТ} обязательно

- 1) C_{II} : **Если** ‘X’ = ‘я’
или ‘X’ идентифицирован в ближайшем контексте,
то ГСинт-актант II может быть элидирован:
Сегодня не работается [подразумевается – мне].
Иван поморщился и встал. Сегодня не работается
 [подразумевается – Ивану].

Синтаксические свойства

1. Глагол $L_{(V)}\text{СЯ1}$ употребляется в основном с отрицанием не; однако он может быть употреблен и без не, в особенности в контексте контраста (см. раздел 2.1.3, пункт 4).

2. Глагол $L_{(V)}\text{СЯ1}$ получает синтаксический признак «возвр».

3. Из актантов исходного глагола $L_{(V)}$, отличных от ‘X’, глагол $L_{(V)}\text{СЯ1}$ может наследовать только те, которые выражаются в ПСинт-структуре косвенными дополнениями или обстоятельствами.

4. В будущем времени частица не может зависеть и от вспомогательного глагола *будет*, и от инфинитива глагола $L_{(V)}\text{СЯ1}$ (см. раздел 2.1.3, пункт 6).

Лексические функции

$\text{Магн}_{[‘\text{препятствовать}’]}$: **редк. абсолютно, разг. никак, совершенно, совсем**
 | ← **рестр** – не

Запись «| ← **рестр** – не» означает, что все выражения лексической функции $\text{Магн}_{[‘\text{препятствовать}’]}$ синтаксически подчиняются частице не посредством рестриктивного ПСинт-отношения. В нижнем индексе при Магн задан семантический компонент смысла нашей дериватемы, который «усиливается» этим Магн.

Примеры

И не пишется, и не поется [В. Высоцкий].

Но и там им совершенно не отдыхалось.

Ивану опять не будет работатьсья. ≡ ...опять будет не работатьсья.

Комментарии

1. Различительные номера при суффиксах **-ся1** и **-ся2**, разумеется, условны: имеется целый ряд других **-ся**, т.е. **-ся** несет еще и другие значения, так что нумеровать суффиксы **-ся** следует соответственно; я же позволяю себе отвлечься от этой задачи. То же касается лексикографических номеров производных глаголов. (О возвратности в русском языке и, в частности, о многозначности суффикса **-ся** см. [Летучий, 2016].)

2. Модель управления производного глагола с дериватемой **НЕ_РАСПОЛОЖЕН**⊕**L_(V)** дана здесь только для самого простого случая: исходный глагол **L_(V)** имеет ровно один ГСинт-актант – **L('X')**, который становится ГСинт-актантом **II** производного глагола **L_(V)ся1**. Однако если **L_(V)** имеет несколько ГСинт-актантов и производным глаголом наследуется из них не один, а *n* (где *n* > 1), то ГСинт-актант **I_X** становится не ГСинт-актантом **II_X** глагола **НЕ_РАСПОЛОЖЕН**⊕**L_(V)**, а его самым последним по счету, т.е. (*n*+1)-ым, ГСинт-актантом. Так, глагол **мечтать** имеет два ГСинт-актанта: **'X_I мечтает о Y_{II}-е'**, и актант **Y_{II}** может наследоваться глаголом **мечтаться**; в таком случае ГСинт-актант **I** глагола **мечтать** становится ГСинт-актантом **III** глагола **мечтаться**:

(18) *Ивану_{III} больше не мечталось о путешествиях_{II}.*

4.2. Словарная статья дериватемы **идти_Z-ово**

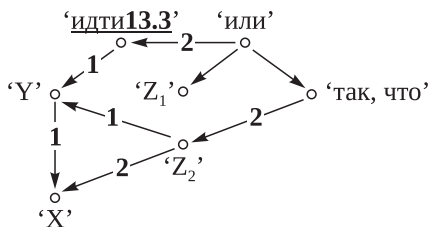
идти_Z-ово⊕**L_(V)**

– Означаемое

Текстовая версия

'Y-ание X-а идет^{13.3} Z₁-ово или так, что X-у Z₂-ово Y-ать'

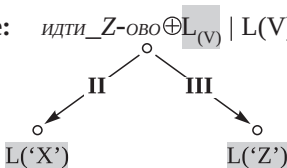
Формальная версия¹⁵



¹⁵ Семантические актанты сочинительных союзов 'и' и 'или' не нумеруются, т.к. их число при данном союзе теоретически не ограничено, а семантически они равноправны.

– **Означающее (на двух уровнях: в ГСинт-структуре и в ПСинт-структуре)**

- В ГСинт-структуре: $идти_Z-ово \oplus L_{(V)} \mid L(V) = L('Y')_{(V)}$



- В ПСинт-структуре: $L_{(V)}ся2 \mid L_{(V)}+ся2 = L_{(V)}ся2$
- **Синтактика**
- **Самой дериватемы** $идти_Z-ово$: $L_{(V)}$, т.е. мишень дериватемы, – преимущественно глагол действия в НЕСОВ ВИДЕ.
- **Производного глагола** $L_{(V)}ся2$:

Флективные формы

Глагол $L_{(V)}ся2$ имеет только 3-е лицо ед. числа и инфинитив; формы сов – предпочтительно делимитативы с префиксом по-, но допустимы и другие формы сов; см. сноску 9.

Модель управления

—	'X' ⇔ II	'Z' ⇔ III
1. \emptyset ПУСТОЕ (N, СРЕД/ЕД)	1. N _{ДАТ}	1. ADV _(оценочное) 2. ADV _(внутр.сост.) 3. КАК?
	обязательно	обязательно

1. $C_{II.1}$: **Если** 'X' = 'я', 'кто угодно'
или 'X' идентифицирован в ближайшем контексте,
то ГСинт-актант II может быть элидирован:
*Сегодня работается чудесно [мне или кому угодно].
Иван уселся поудобнее. Сегодня работается чудесно [Ивану].*
2. C_{III} : **Если** 'Z' = 'нормально' и $L_{(V)}ся2$ – в НЕСОВ ВИДЕ,
то ГСинт-актант III может быть элидирован:
Сегодня мне работается (нормально) vs. Сегодня мне поработалось ?(нормально).
3. $C_{III.1}$: ADV_(оценочное) – любое оценочное выражение.
4. $C_{III.2}$: ADV_(внутр.сост.) – выражение из списка: *бодро, весело, легко, нормально, с трудом, с удовольствием, спокойно, тяжело, уютно, чудесно...*

Синтаксические свойства

1. Глагол $L_{(V)}\text{ся}2$ получает синтаксический признак «возвр».
2. Из актантов исходного глагола $L_{(V)}$, отличных от Сем-актанта 'X', глагол $L_{(V)}\text{ся}2$ может наследовать только те, которые выражаются в ПСинт-структуре косвенными дополнениями или обстоятельствами.

Примеры

Чтобы вам уютно лежалось и бодро вставалось!

Муратову удобно мыслилось в этих рамках.

Люби, пока (тебе) любится.

Вряд ли тебе у них продуктивно пороботается.

И как (тебе) сегодня прыгнулось? [вопрос к спортсмену-парашютисту]

* * *

Рассмотренные дериватемы выражаются суффиксами $-\text{ся}1$ и $-\text{ся}2$, которые могут присоединяться к возвратному глаголу, уже содержащему $-\text{ся}$; в результате возникают последовательности из двух $-\text{ся}$, недопустимые в русском языке: $*-\text{ся}-\text{ся}$. Поэтому в русской грамматике необходимо правило обязательного морфологического эллипсиса:

$*-\text{ся}-\text{ся} \Rightarrow -\text{ся}$.

Этот эллипсис формально идентичен гаплогогии; см. [Holvoet, 2025].

5. Похожее ≠ то же самое

Описание фраз типов (1) и (2) затрудняется еще и тем, что в русском языке существуют фразы, по существу отличные от фраз типов (1) и (2), но и по смыслу и по форме весьма сходные с ними: в них наблюдается такая же конструкция $N_{\text{ДАТ}} \leftarrow \text{непрям.доп.} - V\text{-ся}$, т.е. безличный возвратный глагол в вершине (с отрицанием или без), имеющий дативное дополнение Экспериенцера/Агенса. Во избежание вредных смешений целесообразно указать два таких известных случая.

Случай 1. $V\text{-ся}$ – диахронически производный, но синхронно «лексикализованный» (простой) глагол, т.е. лексема, обязательно заносимая в словарь. Сюда относятся:

– глаголы зеваться, икаться, чихаться, ... (*Ивану зеваются: 'Иван испытывает позыв к зеванию'*). Эти глаголы должны фиксироваться в словаре в силу непродуктивности данной дериватемы в русском: **Мне кашляется <дрожится, рыдается, храпится, ...>*; заметим, что

в южнославянских языках смысл 'испытывать позыв к' является означаемым продуктивной дериватемы;

– глаголы груститься/взгрустнуться, плакаться/всплакнуться, затосковать, ... (*Ивану взгрустнулось*: 'Иван почувствовал грусть').

Случай 2. V-ся – продуктивно (= синхронно) производный глагол, образованный дериватемой, отличной от двух рассмотренных в настоящей статье дериватем: дериватемой *получиться*. Вот соответствующие примеры:

(19) а. {– *Куда же мы пойдём?*} – *А туда, куда нам пойдется* ['куда получится пойти'].

б. *Ивану вдруг заработалось*.

с. – *А мне само собой совралось*.

д. *Иван очень волновался, что ему не заплачется* ['не получится заплакать'].

е. *Решила отправиться на лекции, но в универ мне не пошло*.

ф. *Иван хотел броситься за ней, но ему как-то не побежалось*.

Заметим, что:

- дериватема *получиться* возможна лишь в разговорном стиле;
- во фразы с производным глаголом *получиться*⊕L_(V), особенно в утвердительные, легко вставляются идиомы 'само собой' и 'как-то' и/или просторечное семантически пустое местоимение оно (*А туда, куда оно пойдется*); здесь наблюдается, по-видимому, контаминация разных конструкций: см. [Поляков, 2025, с. 118];
- производный глагол *получиться*⊕L_(V) употребляется преимущественно в сов виде. Однако он может быть и в несов виде; тогда, если при нем есть отрицание, результирующая фраза неотличима от фразы с производным глаголом *быть_нерасположен*⊕L_(V):

(20) *Мне не встается*.

Выражение (20) соответствует двум разным смыслам: 'Я нахожусь в таком внутреннем состоянии, которое препятствует тому, чтобы я вставал' или 'У меня не получается вставать'. Впрочем, оба смысла очень близки.

Продуктивность этой дериватемы не проверялась.

6. Заключение

Систематическое описание свойств фраз типов (1) и (2) приводит к выводу, что представленная в них «дательнo-бeзличнo-вoзвратнaя

конструкция» есть не что иное, как производные глаголы двух разных типов – типа РАБОТАТЬСЯ₁ и типа РАБОТАТЬСЯ₂ – со своими актантами:

‘X-у не Y-ается’ = ‘[[Пытаясь Y-ать,]] X находится в таком внутреннем состоянии, которое препятствует тому, чтобы X Y-ал’ [Ивану_x не работается.]

‘X-у Y-ается Z-ово’ = ‘Y X-а идет13.3 Z₁-ово или так, что X-у Z₂-ово Y-ать’ [Ивану_x хорошо_{Z₁}/весело_{Z₂} работается.]

Такие глаголы предлагается называть *диспозиционалами* (от англ. *dispositionals*).

Тем самым, проблема описания фраз типов (1) и (2) сводится к проблеме продуктивного синхронного словообразования, или синхронной деривации, в русском языке, более точно, – к проблеме образования глаголов-диспозиционалов с помощью возвратных суффиксов -ся₁ и -ся₂. В этой связи вводится понятие дериватемы – языкового знака, промежуточного между граммемой и лексемой. Приводятся формальные структуры фраз (1) и (2) и полные словарные статьи дериватем БЫТЬ_НЕРАСПОЛОЖЕН и ИДТИ_Z-ОВО.

Диспозиционалы типичны для всех славянских языков – факт, важный для лингвистической типологии. Любопытно, однако, распределение соответствующих дериватем в славянских языках [Mel’čuk, Milićević, to appear]:

- дериватема БЫТЬ_НЕРАСПОЛОЖЕН засвидетельствована только в восточнославянских языках (т.е., кроме русского, в украинском и белорусском);
- дериватема ИДТИ_Z-ОВО имеется, помимо восточнославянских, еще и в западнославянских языках (т.е. в польском, чешском и словацком);
- дериватемы же ХОТЕТЬСЯ [= FEEL_LIKE] и ИСПЫТЫВАТЬ_ПОЗЫВ [= FEEL_URGE] фигурируют только в южнославянских языках (в словенском, сербском, хорватском, болгарском и македонском).

Приложение

Безличный глагол

Безличный глагол – это глагол, который не имеет ГСинт-актанта I; на ПСинт-уровне он получает подлежащее-пустышку (dummy SSynt-subject) в виде семантически пустого местоимения типа англ. *it*, фр. *il*, нем. *es* или рус. \emptyset ^{ПУСТОЕ}_(Н, СРЕДНЕД.).

Безличные глаголы имеются во многих (возможно, практически во всех?) языках. Прототипические безличные глаголы обозначают природные явления и внутренние состояния человека:

- англ. *It dawns* ‘Светает’, нем. *Es schneit* ‘Идет снег’, лат. *Ninguit* ‘Идет снег’, исп. *Nieva* ‘Идет снег’, рус. *Светаает*;
- рус. *Меня тошнит*, *Меня трясет*, *Мне смешно*, *Мне холодно*, лат. *Paenitet* N_{X-ACC} N_{Y-GEN} (*Paenitet me_X senectutis_Y meae* ‘Я_X сожалею о своей старости_Y’), *Pudet* N_{X-ACC} N_{Y-GEN} (*Pudet me_X fratris_Y* ‘Мне_X стыдно за брата_Y’).

NB Термин *безличный глагол* представляется неудачным, т.к. приходится говорить, что безличный глагол употребляется только в третьем **лице** и имеет **личную** форму (которая выступает как синтаксическая вершина – сказуемое – предложения). Но, к сожалению, я не способен предложить никакого лучшего термина. «Трудно жить на этом свете – в нем отсутствует уют!» [по Н. Олейникову].

Путаница с термином *безличный глагол* усугубляется еще и тем, что этот термин часто применяют к глаголам, которые в соответствии со своим смыслом могут употребляться только с подлежащим, обозначающим факт, например, глагол произойти: он употребляется, естественно, лишь в 3-м лице. Такой глагол, однако, имеет ГСинтактант **I** и, в соответствии с приведенным определением, безличным не является. (О безличности в русском языке см. детальный обзор [Летучий, 2011].)

Критерий объединяющей сочетаемости (= критерий Грин–Апресьяна)

Этот лексикографический критерий (см. [Mel’čuk, 2013, с. 330]) позволяет определять целесообразность дизъюнктивного толкования данной лексемы.

Пусть имеется «подозрительная» лексема L*, смысл которой содержит дизъюнкцию семантических компонентов: ‘...σ₁ **или** σ₂...’. Следует ли расщепить ее на лексемы L’ ‘...σ₁...’ и L’’ ‘...σ₂...’ или ее следует сохранить как единую лексему L с дизъюнктивным толкованием?

Если можно построить предложение, в котором наблюдается объединяющая сочетаемость для L* – такая что:

- 1) **либо** L* ‘...σ₁ **или** σ₂...’-synt-[L₁ и/или L₂], причем ‘L₁-sem-σ₁’ и ‘L₂-sem-σ₂’;
- 2) **либо** L* ‘...σ₁ **или** σ₂...’-synt-L₃, причем ‘L₃-sem-σ₁’ и ‘L₃-sem-σ₂’, **то** L* не следует расщеплять на две лексемы.

В результате положительного «ответа» этого критерия получается обычная лексема L с дизъюнкцией в толковании: ‘L’ = ‘...σ₁ **или** σ₂...’.

Примеры

Случай 1, т.е. L^* синтаксически связана с сочиненными лексемами L_1 и/или L_2 так, что L_1 семантически связана с одним из дизъюнктивных компонентов смысла L^* , а L_2 – с другим; возьмем $L^* =$ остыть ‘стать холоднее или холодным’:

Жидкость остыла_{L} сначала немного_{L1}, а затем полностью_{L2}.*

Случай 2, т.е. L^* синтаксически связана с лексемой L' , которая семантически связана со всеми дизъюнктивными компонентами смысла L^* ; $L^* =$ дядя ‘брат отца, или брат матери, или муж тетки’:

Мои_{L'} дяди_{L} Иван – брат отца и Пётр – муж маминой сестры, жили в нашем городе.*

В обоих случаях L^* признается единой лексемой: $L^* = L$ ‘... σ_1 или σ_2 ...’.

Если же нормальное предложение, охарактеризованное в формулировке критерия, построить невозможно (например, получается каламбур), то L^* должна быть расщеплена. Так, каламбурное предложение *Маша приняла Ивана и две таблетки аспирина* указывает на наличие двух лексем принять: вместо L^* мы имеем L' ‘... σ_1 ...’ и L'' ‘... σ_2 ...’.

Библиографический список / References

Апресян, 2005 – Апресян Ю.Д. О Московской семантической школе // Вопросы языкознания. 2005. № 1. С. 1–30. [Apresjan Ju.D. On the Moscow Semantic School. *Voprosy jazykoznanija*. 2005. No. 1. Pp. 1–30. (In Rus.)]

Апресян, 2023 – Апресян Ю.Д. ИДТИ: словарная статья // Активный словарь русского языка. Т. 4. Ч. 1. Ю.Д. Апресян (отв. ред.). М., 2023. С. 43–54. [Apresjan Ju.D. *Idti* ‘go’: dictionary entry. *Aktivnyy slovar russkogo yazyka*. Vol. 4. Part 1. Ju.D. Apresjan (ed.). Moscow, 2023. Pp. 43–54. (In Rus.)]

Булыгина, Шмелёв, 1997 – Булыгина Т.В., Шмелёв А.Д. Языковая концептуализация мира (на материале русской грамматики). М., 1997. [Bulygina T.V., Shmelyov A.D. *Yazykovaya kontseptualizatsiya mira (na materiale russkoy grammatiki)* [Linguistic conceptualization of the world (based on Russian grammar)]. Moscow, 1997.]

Зализняк, 1992 – Зализняк Анна А. Исследования по семантике предикатов внутреннего состояния. München etc., 1992. [Zaliznyak Anna A. *Issledovaniya po semantike predikatov vnutrennego sostoyaniya* [Research on the semantics of internal state predicates]. München etc., 1992.]

Иванова, 2007 – Иванова Е.Ю. Уникальна ли русская конструкция *мне не работает?* или О чрезмерном усердии в поиске национальноязыковой специфики // Филология и человек. 2007. № 3. С. 1–10. [Ivanova E.Yu. Is the Russian construction *mne ne rabotaetsja* ‘I can’t work’ unique? or About excessive zeal in searching for national-language specifics. *Filologiya i chelovek*. 2007. No. 3. Pp. 1–10. (In Rus.)]

Иванова, 2022 – Иванова Е.Ю. Предикаты желания в болгарском языке в сопоставлении с русским // *Онтология на ситуациите за състояние – лингвистично моделиране. Съпоставително изследване за български и руски*. С. Коева, Е. Иванова, Й. Тишева, А. Циммерлинг (ред.). София, 2022. С. 427–457. DOI: 10.7546/STONTB-GRU2022.15 [Ivanova E.Yu. Predicates of desire in the Bulgarian language in comparison with Russian. *Ontologija na situaciite za състояние – лингвистично моделиране. Съпоставително изследване за български и руски*. S. Koeva, E. Ivanova, J. Tiševa, A. Zimmerling (eds.). Sofia, 2022. Pp. 427–457. (In Rus.). DOI: 10.7546/STONTB-GRU2022]

Летучий, 2011 – Летучий А.Б. Безличность. Материалы для проекта корпусного описания русской грамматики. М., 2011. URL: <http://rusgram.ru> (accessed: 09.01.2025). [Letuchiy A.B. Bezlichnost. Materialy dlya proekta korpusnogo opisaniya russkoi grammatiki [Impersonality. Materials for the project of corpus description of Russian grammar (rusgram.ru)]. Ms. Moscow, 2011. URL: <http://rusgram.ru> (accessed: 09.01.2025).]

Летучий, 2014 – Летучий А.Б. Между пассивом и декаузативом: русские модальные пассивы // *Acta Linguistica Petropolitana*. 2014. № 3. С. 364–394. [Letuchiy A.B. Between the passive and the decausative: Russian modal passives. *Acta Linguistica Petropolitana*. 2014. No. 3. Pp. 364–394. (In Rus.)]

Летучий, 2016 – Летучий А.Б. Возвратность // *Материалы к Корпусной грамматике русского языка. Глагол*. Ч. I. В.А. Плунгян (отв. ред.). СПб., 2016. С. 268–337. [Letuchiy A.B. Reflexivity. *Materialy k Korpusnoy grammatike russkogo yazyka*. Part I. Verb. V.A. Plungjan (ed.). St. Petersburg, 2016. Pp. 268–337. (In Rus.)]

Мельчук, 1997 – Мельчук И.А. Курс общей морфологии. Т. I. М.; Вена, 1997. [Mel'čuk I.A. Kurs obshchey morfologii [General Morphology]. T. I. Moscow; Vienna, 1997.]

Поляков, 2025 – Поляков Д.К. «Ну, как вам съездилось, Лидия Андреевна?»: безличный модальный пассив глаголов совершенного вида в русском художественном тексте // *Rhema. Рема*. 2025. № 1. С. 98–121. DOI: 10.31862/2500-2953-2025-1-98-121 [Polyakov D. “Nu, kak vam s'ezdilos', Lidiya Andreevna?»: Impersonal modal passive of perfective verbs in Russian literary text. *Rhema*. 2025. No. 1. Pp. 98–121. (In Rus.). DOI: 10.31862/2500-2953-2025-1-98-121]

Фичи, 2011 – Фичи Ф. Об одной модальной функции рефлексивных конструкций // *Слово и язык. Сборник статей к восьмидесятилетию академика Ю.Д. Апресяна / под ред. И.М. Богуславского, Л.Л. Иомдина, Л.П. Крысына*. М., 2011. С. 645–651. [Fici F. On a modal function of reflexive constructions. *Slovo i yazyk. Sbornik statey k vosmidesyatiletiju akademika Yu.D. Apresyana*. I.M. Boguslavskiy, L.L. Iomdin, L.P. Krysin (eds.). Moscow, 2011. Pp. 645–651. (In Rus.)]

Фичи, Жукова, 2012 – Фичи Ф., Жукова Н.Н. О грамматико-семантических свойствах конструкций типа *Вчера мне легко работалось* // *Русский язык в научном освещении*. 2012. № 1 (=23). С. 18–38. [Fici F., Zhukova N.N. On the grammatical and semantic properties of constructions of the type “Vchera мне legko rabotalos”. *Russkiy yazyk v nauchnom osveshchenii*. 2012. No. 1 (= 23). Pp. 18–38. (In Rus.)]

Goldberg, Suttle, 2010 – Goldberg A.E., Suttle L. Construction grammar. *Wires Cognitive Science*. 2010. Vol. 1. No. 4. Pp. 468–477. DOI: 10.1002/wcs.22

Holvoet, 2025 – Holvoet A. Haplogy and diplogy of grammatical markers in Balto-Slavic. URL: https://www.researchgate.net/publication/395472753_Haplogy_and_diplogy_of_grammatical_markers_in_Balto-Slavic (date accessed: 15.10.2025).

Mel'čuk, 1993 – Mel'čuk I. Cours de morphologie générale. Vol. 1 : Introduction + Le mot. Montréal ; Paris, 1993.

Mel'čuk, 2013 – Mel'čuk I. Semantics: From Meaning to Text. Vol. 2. Amsterdam; Philadelphia, 2013.

Mel'čuk, 2023 – Mel'čuk I. General Phraseology: Theory and Practice. Amsterdam; Philadelphia, 2023.

Mel'čuk, Milićević, 2023 – Mel'čuk I., Milićević J. Dispositionals in Russian and Serbian. To appear.

Wierzbicka, 1992 – Wierzbicka A. The Russian language. *Wierzbicka A. Semantics, culture, and cognition*. Oxford University Press, 1992. Pp. 395–441.

Статья поступила в редакцию 04.09.2025

The article was received 04.09.2025

Сведения об авторе / About the author

Мельчук Игорь Александрович – кандидат филологических наук; заслуженный профессор и сотрудник лингвистической лаборатории «Смысл – Текст» кафедры лингвистики, Монреальский университет, Канада

Igor Mel'čuk – PhD in Linguistics; Distinguished Professor and Fellow at the Meaning – Text Laboratory of the Department of Linguistics, University of Montreal, Montreal, Canada

ORCID: <https://orcid.org/0000-0002-4520-0554>

E-mail: igor.melcuk@umontreal.ca

DOI: 10.31862/2500-2953-2025-4-106-139

УДК 81-114.4

Е.А. Миронова¹, Ш.Ш. Унарокова², Ю.А. Ландер¹

¹ Национальный исследовательский университет
«Высшая школа экономики»,
105066 г. Москва, Российская Федерация

² Адыгейский государственный университет,
385000 г. Майкоп, Российская Федерация

Абазинские и адыгейские демонстративы: моноцентричность под видом полицентричности

Статья посвящена анализу систем демонстративов в ашхарском диалекте абазинского и в темиргоевском диалекте адыгейского языков. Исследование основано на эксперименте по анкете Д.П. Вилкинса (Wilkins, 2018), выявляющей параметры, которые влияют на выбор демонстративов. Несмотря на некоторые различия, рассматриваемые системы сходны в том, что для них значимы расстояние от говорящего, ассоциированность с адресатом и видимость. При этом ни один из демонстративов не используется последовательно в ситуациях, когда референт находится рядом с адресатом: выбор также зависит от дистанции. Мы предполагаем, что данные системы не полицентричны, а, напротив, ориентированы на говорящего, и ассоциированность с адресатом – лишь один из параметров, ограничивающих включенность референта в зону говорящего.

Ключевые слова: абазинский язык, адыгейский язык, абхазо-адыгские языки, демонстративы, дейксис, моноцентричность, полицентричность, экспериментальные данные

Благодарности. Данные по абазинскому языку собирались в экспедиции, организованной при поддержке Фонда образовательных инициатив Национального исследовательского университета «Высшая школа экономики» в рамках конкурса «Открываем Россию заново». Исследование Ю.А. Ландера выполнено при поддержке гранта РФФИ № 24-18-00879 «Принципы формирования полисемии и колексификации: границы возможного».

© Миронова Е.А., Унарокова Ш.Ш., Ландер Ю.А., 2025

Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License



Авторы благодарны участникам абазинского лингвистического проекта Национального исследовательского университета «Высшая школа экономики» и слушателям семинара Международной лаборатории языковой конвергенции Национального исследовательского университета «Высшая школа экономики» и Третьей конференции “Anatolia – The Caucasus – Iran” (Ереван, май 2025 г.) за обсуждение данных и основных положений исследования, а также анонимным рецензентам статьи за замечания и исправления. Мы выражаем особую признательность участникам экспериментов, результаты которых описаны в статье. Ответственность за все ошибки лежит на авторах работы.

ДЛЯ ЦИТИРОВАНИЯ: Миронова Е.А., Унарокова Ш.Ш., Ландер Ю.А. Абазинские и адыгейские демонстративы: моноцентричность под видом полицентричности // Рема. Rhema. 2025. № 4. С. 106–139. DOI: 10.31862/2500-2953-2025-4-106-139

DOI: 10.31862/2500-2953-2025-4-106-139

E. Mironova¹, Sh. Unarokova², Yu. Lander¹

¹ HSE University,
Moscow, 105066, Russian Federation

² Adyge State University,
Maykop, 385000, Russian Federation

Abaza and West Circassian demonstratives: Monocentricity under the guise of polycentricity

The paper examines demonstrative systems in Ashqarawa Abaza and Temirgoy West Circassian. The study is based on an experiment using the questionnaire (Wilkins, 2018), which identifies parameters affecting the choice of demonstratives. Despite certain differences, both systems are similar in that they show the relevance of distance from the speaker, association with the addressee, and visibility. However, no demonstrative is used consistently when the referent is near the addressee, since distance from the speaker is also significant. We argue that these systems are not polycentric but rather speaker-oriented, where association with the addressee prevents inclusion in the speaker's domain.

Key words: Abaza, West Circassian, Northwest Caucasian languages, demonstratives, deixis, monocentricity, polycentricity, experimental data

Acknowledgments. The Abaza data were collected in a field trip supported by the Foundation for Educational Initiatives of HSE University (“Rediscovering Russia” program). Yury Lander’s research was supported by the Russian Science Foundation, grant No. 24-18-00879 “Principles of polysemy and colexification: The boundaries of what is possible”.

We thank the participation of the Abaza field project (HSE University) and the audience of the seminar of the Linguistic Convergence Laboratory at HSE University and of the Third International Conference “Anatolia – The Caucasus – Iran” (Yerevan, May 2025) for discussion, and two anonymous reviewers for their notes and corrections. We are particularly grateful to the participants of the experiments whose results are presented in the paper. All errors are ours.

FOR CITATION: Mironova E., Unarokova Sh., Lander Yu. Abaza and West Circassian demonstratives: Monocentricity under the guise of polycentricity. *Rhema*. 2025. No. 4. Pp. 106–139. DOI: 10.31862/2500-2953-2025-4-106-139

1. Введение

Традиционно считается, что основная функция указательных местоимений (демонстративов) связана с дейктическим пространственным употреблением – указанием на объект (референт) через его отношение к месту речевого акта. Эта функция – экзофорическая – противостоит разнообразным эндофорическим функциям, предполагающим в первую очередь обращение к информации внутри дискурса (в частности, анафору). Однако экзофорическая функция является первичной по отношению к эндофорическим, даже если последние ее начинают вытеснять [Diessel, 1999; Levinson, 2018].

Тем не менее и в отношении экзофорических употреблений языки различаются по тому, каким образом они соотносят место речевого акта и референт. Как утверждается в литературе, основные противопоставления здесь определяются ассоциацией референта с говорящим или с адресатом (лично-ориентированная система) или с разным расстоянием от говорящего (дистантно-ориентированная система); см. [Anderson, Keenan, 1985, p. 282; Bhat, 2007, p. 177; Diessel, 1999, p. 39; Diessel, 2013]¹. Например, система латинских указательных местоимений описывается как лично-ориентированная: *hic* предполагает, что референт ассоциируется с говорящим, *iste* – что референт ассоциируется с адресатом, *ille* – что референт находится вне сферы коммуникации говорящего

¹ Ср., впрочем, основанную на специфических допущениях попытку свести разные системы к одной (лично-ориентированной) в [Lander, Haegeman, 2018].

и адресата [Pinkster, 2015, p. 1137]. Напротив, противопоставление демонстративов вроде *этот* и *тот* в русском языке связывается в первую очередь с (соответственно) близостью и дальностью от говорящего [Падучева и др., 2011].

Дистантно-ориентированные системы также называются моноцентричными, т.к. они содержат один дейктический центр (или «ориго» [Bühler, 1982], «indexical ground» [Hanks, 2011]) – говорящего, относительно которого определяется расположение референта [Diessel, 1999, p. 41; Schulze, 2008, p. 247–248; Diessel, 2013]. Лично-ориентированные системы, наоборот, отличаются тем, что они включают, в частности, демонстративы, ориентированные не на говорящего, а на адресата или на общий дейктический центр двух локуторов; соответственно, эти системы полицентричны.

При выборе указательного местоимения могут играть роль и другие факторы, кроме того, языки разнятся в деталях функционирования лично-ориентированной и/или дистантно-ориентированной систем. Этому способствует то, что выбор указательного местоимения зачастую не является результатом очевидного отнесения референта к той или иной области, но строится вокруг прототипических употреблений демонстративов и языки могут в разной степени позволять отход от таких прототипических употреблений [Wilkins, 2018, p. 55]. В этой связи интересно сопоставить функционирование демонстративов, входящих в сходные системы в родственных языках, поскольку внутривидовая типология особенно продуктивна для изучения континуальных параметров [Кибрик, 2005, с. 194; Шлуинский, 2014].

В этой статье мы рассмотрим неконтрастивные пространственные дейктические употребления указательных местоимений в ашхарском диалекте абазинского языка² и в темиргоевском диалекте адыгейского языка (оба языка принадлежат абхазо-адыгской семье). В основе нашего описания лежат экспериментальные данные, собранные по анкете [Wilkins, 2018], которая широко используется при анализе систем указательных местоимений для разных языков; см., например, многочисленные работы в сборнике [Levinson et al., 2018]. Альтернативный способ анализа трехчастных систем демонстративов предложен в [Ростовцев-Попель, 2009]. Как и в анкете [Wilkins, 2018], метод основан на инсценированных ситуациях, однако здесь используются контрастивные контексты и рассматриваются лишь четыре диагностических теста. Они

² С лингвистической точки зрения ашхарский диалект, по-видимому, ближе к абхазскому языку, чем к основному абазинскому диалекту – тапантскому [Chirikba, 1996, p. 14]. Впрочем, у нас нет оснований считать, что абхазская и абазинская системы демонстративов принципиально различаются.

сфокусированы прежде всего на значении среднего дейктика, которое считается ключевым различием трехчастных систем. Мы используем анкету [Wilkins, 2018], поскольку она позволяет систематически варьировать больше параметров и меньше зависит от предварительных теоретических допущений.

Из ареально близких обсуждаемым нами языков описание, основанное на этой анкете, есть для агульского [Мерданова, Федорова, 2002] и рутульского [Mukhin, Nasledskova, 2025] языков, принадлежащих нахско-дагестанской семье. Типология демонстративов в дагестанских языках также обсуждается в [Федорова, 2001]. Для абхазо-адыгских языков, насколько нам известно, специальных исследований семантики демонстративов не предпринималось.

Как мы увидим, системы функционирования демонстративов в рассмотренных идиомах следуют сходным принципам, но различаются в примечательных деталях. При этом абазинская и адыгейская системы показывают, что границы между лично-ориентированным и дистантно-ориентированным типами не всегда очевидны.

В разделе 2 рассматриваются подходы к описанию демонстративов в абхазо-адыгских языках, противопоставляющие лично-ориентированный и дистантно-ориентированный анализ. В разделе 3 описывается методология исследования, в частности, условия эксперимента и процедуры сбора и анализа данных. В разделе 4 представлены результаты эксперимента: отдельно для ашхарского и для адыгейского анализируются значения демонстративов и сопоставляются факторы, влияющие на выбор указательного местоимения. В разделе 5 сопоставляются обе системы, выявляются их сходства и различия. В разделе 6 предлагаются интерпретации полученных данных в более широком типологическом и теоретическом контексте. Наконец, в разделе 7 формулируются основные выводы.

2. Подходы к изучению демонстративов в абхазо-адыгских языках

Абхазо-адыгская семья включает пять языков:

- абхазо-абазинская группа – абхазский, абазинский;
- адыгская группа – адыгейский, кабардино-черкесский;
- убыхский язык (последний носитель умер в 1992 г.).

В таблице 1 приведены основные приименные формы демонстративов в каждом из языков (в абхазском и абазинском они совпадают с независимыми демонстративами, в других языках неприименное указание требует дополнительной морфологии). Разряды демонстративов условно обозначены римскими цифрами, которые в разных описаниях могут соответствовать ярлыкам «ближний», «средний», «дальний» или

«проксимальный», «медиальный», «дистальный». Мы не используем эти термины, чтобы избежать соответствующих ожиданий от семантики демонстративов, которая обсуждается в этой статье.

Таблица 1

Инвентарь демонстративов в абхазо-адыгских языках
[Demonstrative inventory in the West Caucasian languages]

	Абхазский [Abkhaz]	Абазинский [Abaza]	Убыхский [Ubykh]	Кабардино-черкесский [Kabardian]	Адыгейский [West Circassian]
I	<i>ару/абри</i>	<i>ару/абару</i>	<i>и</i>	<i>мы</i>	<i>мы</i>
II	<i>ани/абни</i>	<i>ани/абани</i>	–	<i>мо</i>	<i>мо/о/у</i> ¹
III	<i>уу/убри</i>	<i>ауу/абауу</i>	<i>уэ</i>	<i>а</i>	<i>а</i>
(IV)	–	–	–	–	<i>джа</i>

¹ Варианты *о* и *у* представлены в шапсугском и бжедугском диалектах адыгейского языка. В литературе встречается утверждение, что к сочетанию этой местоименной основы и местоимения *мы* восходит и форма *мо* [Шагиров, 1977, с. 269; Кумахов, 1989, с. 93–94].

Во всех языках, за исключением убыхского, есть по меньшей мере три указательных местоимения. В адыгейском языке можно выделить и четвертое указательное местоимение (*джа*), в литературе почти не описанное (см., однако, [Барыльникова, 2013; Ландер, Унарокова, 2024]), но у него собственно пространственные употребления, по-видимому, почти отсутствуют – в частности, оно не появляется при ответе на используемую нами анкету. Для убыхского языка выделяется только два местоимения.

Противопоставление между демонстративами в убыхском языке описывается исключительно через расстояние (см., например, [Fenwick, 2011, с. 79]) и далее не обсуждается: эта система отличается от того, что наблюдается в других языках семьи. Указательные местоимения остальных абхазо-адыгских языков исследователи описывают на основании разных принципов – как лично-ориентированного, так и дистантно-ориентированного, хотя при этом указываются и дополнительные противопоставления вроде «видимый vs невидимый».

Лично-ориентированный подход. Дж. Хьюит в обзоре абхазо-адыгской семьи выделяет три уровня дейксиса, которым должны соответствовать три демонстратива: (1) около говорящего, (2) около адресата / видимый, (3) далекий или невидимый [Hewitt, 2005, p. 104].

Похожим образом А.Н. Генко описывает абазинскую систему указательных местоимений через тройное противопоставление: «предмет, о котором говорится, находится в ближайшем соседстве с самим говорящим» (соответствует абазинскому *ари*), «предмет, о котором говорится, находится в некотором отдалении от говорящего, в соседстве с тем, к кому речь обращена» (соответствует *ани*), «предмет, о котором говорится, находится в соседстве с третьими лицами, близкими или далекими, присутствующими или отсутствующими» (соответствует *ауи*) [Генко, 1955, с. 102]. Аналогично (впрочем, наряду с дистантно-ориентированным способом описания) представляет контраст между абазинскими демонстративами и К. Ломтатидзе [Ломтатидзе, 2006, с. 135].

Наконец, Г.В. Рогава и З.И. Керашева считают, что к аналогичной системе восходит тройное противопоставление в адыгейском языке, предполагая, что «указательные местоимения исторически подразделялись на *мы-р*, *мо-р* и *а-р* (здесь *-р* – показатель именительного/абсолютного падежа. – Е.М., Ш.У., Ю.Л.) не по признакам близости и дальности, а по принципам нахождения предмета около первого, второго и третьего лиц: первого лица – *мы-р*, второго лица – *мо-р*, третьего лица – *а-р*» [Рогава, Керашева, 1966, с. 85].

Дистантно-ориентированный подход встречается в описании всех абхазо-адыгских языков.

Общим для дистантно-ориентированных описаний абхазского и абазинского языков оказывается представление местоимений ряда *ари* как указывающих на объект, располагающийся близко к говорящему, и местоимений ряда *ани* как указывающих на видимый объект, располагающийся на расстоянии от говорящего. Дистантные местоимения ряда *ауи* трактуются по-разному:

- как относящиеся к объекту, который говорящий не видит (см. для абхазского языка [Chirikba, 2003, р. 33; Yanagisawa, 2013, р. 61], для абазинского [Клычев, 2000, с. 131]);
- как относящиеся к далекому объекту – в отличие от объекта, располагающегося неподалеку и в поле зрения (см. [Шадури, 2006, с. 43] для абхазского языка);
- как относящиеся к объекту, который находится рядом с третьим лицом (для абазинского см. [Ломтатидзе, 2006, с. 135], в том числе и в дистантно-ориентированном описании).

В абазинской грамматике Н.Т. Табуловой, впрочем, сообщается, что демонстративы ряда *ауи* указывают «на видимый или невидимый предмет (лицо), находящийся от говорящего не на отдаленном расстоянии»

[Табулов, 1976, с. 94]. Так или иначе, во всех этих описаниях выявляется дополнительный параметр видимости.

Те же сложности возникают и для адыгских языков. Здесь чаще всего противопоставление, лежащее в основе системы демонстративов, тоже трактуется как противопоставление, основанное на дистанции и видимости. Почти все описания сходятся в том, что указательные местоимения *мы* и *мо* различаются по тому, что первое указывает на объект около говорящего, а второе – на объект на некотором расстоянии. Особняком стоит работа Т.Н. Конук, в которой *мо* представляется как указывающее на объект между говорящим и адресатом [Конук, 2022, с. 147]. Иногда одновременно подчеркивается, что референт при использовании *мо* должен быть видим; ср. [Яковлев, 1948, с. 68; Colarusso, 1992, с. 53–54] для кабардино-черкесского языка и [Яковлев, Ашхамаф, 1941, с. 289; Керашева, 1957, с. 60; Paris, 1989, с. 177] для адыгейского языка (последняя работа при этом заявляет, что объект должен видеть адресат, что сближает ее с лично-ориентированным подходом). Впрочем, в неопубликованной работе Д.В. Барыльникова допускается использование *мо* при невидимом объекте, если говорящий специфицирует жестом направление, в котором он располагается (жест в этой работе вообще понимается как почти необходимое условие для появления *мо*) [Барыльникова, 2013]. В описании кабардино-черкесского языка видимость требуется для местоимения *мы* [Кумахов, 2006, с. 120]; ср. формулировку «на некотором расстоянии от говорящего, но видимо».

Проблема возникает с третьим указательным местоимением – *a*. Как указывалось выше, грамматика адыгейского языка определяет его исключительно через дистанцию («указывает на вещь (лицо), которая находится дальше от первого и второго лиц» [Рогава, Керашева, 1966, с. 85]). Другие источники настаивают, что значение *a* должно определяться через невидимость; ср. [Яковлев, Ашхамаф, 1941, с. 120; Керашева, 1957, с. 60; Paris, 1989, с. 177] для адыгейского языка и [Colarusso, 1992, с. 63–64] для кабардино-черкесского языка. В книге Н.Ф. Яковлева, впрочем, дается более аккуратная формулировка: «предмет невидимый или далекий, чуть видимый» [Яковлев, 1948, с. 68]. Д.В. Барыльникова (2013) также говорит лишь о том, что невидимость часто сопровождает употребление *a*.

Ниже мы увидим, что и в ашхарском диалекте абазинского языка, и в адыгейском языке выбор демонстратива не определяется исключительно ни связью с говорящим или адресатом или третьим лицом, ни дистанцией, ни видимостью – но все эти факторы оказываются важными.

3. Методология исследования

3.1. Описание эксперимента

Эксперимент, основанный на анкете [Wilkins, 2018], был организован следующим образом. Инсценировалась ситуация, содержащая по меньшей мере говорящего (консультанта), адресата и объект, который выступал в качестве референта указательного местоимения. Консультант должен был перевести предложение с демонстративом в соответствии с этой ситуацией с русского языка на абазинский или адыгейский язык. При этом в предложении на русском языке демонстратив обычно заменялся на пропуск или повторялись два варианта предложения (с демонстративами *этот, тот*), чтобы предотвратить эффект прайминга. Консультанта просили не использовать жест рукой, если не сказано иное, т.к. наличие жеста может влиять на выбор демонстратива [Levinson, 2018, p. 32–33].

Анкета состоит из 25 ситуаций, которые в схематическом виде приведены на рисунке 1. Говорящий отмечен на них как S, адресат – как A. Референтом в ситуациях 1–2 является часть тела, в ситуациях 3–5 – насекомое, в остальных ситуациях – любой предмет подходящего размера.

Пример ситуации (рис. 2) и соответствующего предложения представлены далее.

Ситуация (см. рис. 2): референт (красная книга) находится прямо перед говорящим, адресат сидит сбоку, т.е. он видит референт, но не может достать рукой.

Задание: обратившись к исследователю, переведите на целевой язык предложение «Эта/та книга красная» в соответствии с ситуацией.

Результат (1)–(2):

(1) абазинский язык

Ари апхъага къапцунI.

arəj á-pχ'aga qapš'ə-w-ṗ

dem(I) sp-книга красный-ipfv-ṗpst

(2) адыгейский язык

Мы тхылгыр плъыжсы.

mə tχələ-g pləž'ə

dem(I) книга-abs красный

В ситуациях варьируются следующие параметры: расстояние от референта до говорящего (в [Wilkins, 2018] выделяется семь степеней удаленности, таблица 2), расстояние от референта до адресата (те же семь степеней удаленности), видимость референта для говорящего и для адресата, наличие пространственных границ (оконная рама, дверной

проем), наличие третьего человека. Анкета также содержит минимальную пару по наличию/отсутствию жеста (ситуации 4 и 5). Предполагается, что для каждого демонстратива одни ситуации будут более типичными, а другие – более периферийными или невозможными.

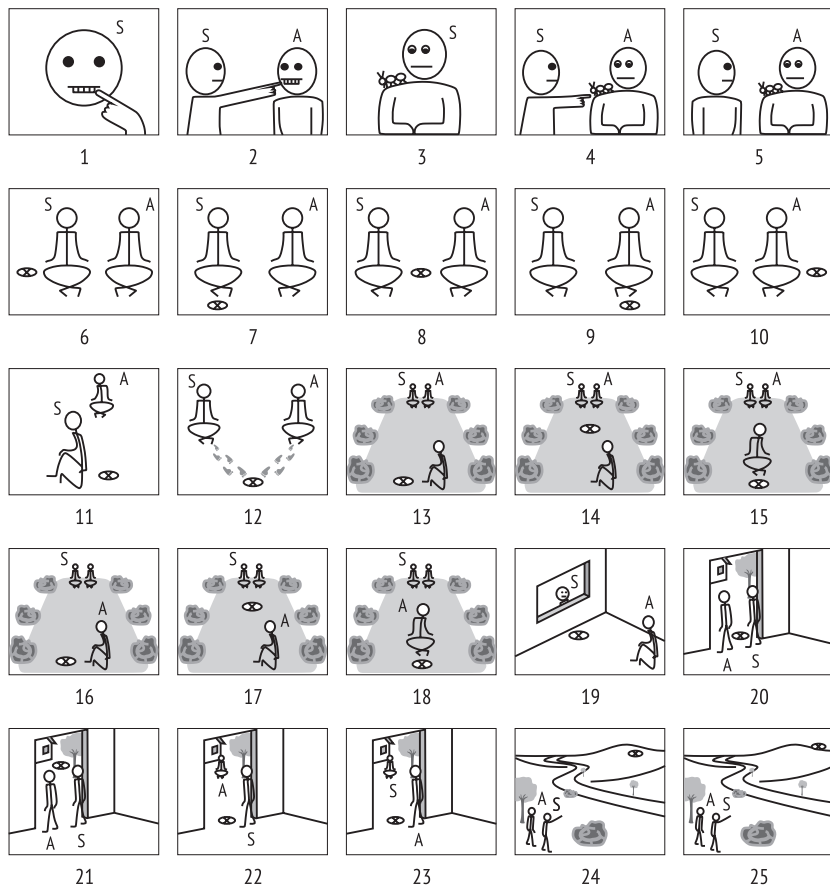


Рис. 1. Анкета по [Wilkins, 2018]

Говорящий отмечен как S, адресат – как A. Референтом в ситуациях 1–2 является часть тела, в ситуациях 3–5 – насекомое, в остальных ситуациях – любой предмет подходящего размера

Fig. 1. Questionnaire based on [Wilkins, 2018]

The speaker is marked as S, the addressee as A. The referent in situations 1–2 is a body part, in situations 3–5 it is an insect, in the remaining situations it is any object of a suitable size

Таблица 2

**Шкала расстояния от говорящего до референта по [Wilkins, 2018]
[Speaker-referent distance scale according to [Wilkins, 2018]]**

1	Референт – часть тела [The referent is a body part]
2	Референт касается тела [The referent is in contact with the body]
3	До референта можно дотянуться рукой [The referent is within arm's reach]
4	Референт в легкой доступности (в нескольких шагах) [The referent is within easy access (a few steps from the speaker)]
5	Референт в нескольких десятках метров [The referent is a few tens of meters away]
6	Референт на расстоянии более 100 м [The referent is more than 100 meters away]
7	Референт на расстоянии нескольких километров [The referent is several kilometers away]

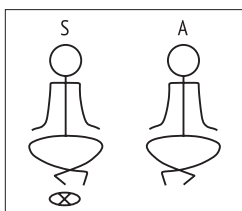


Рис. 2. Пример ситуации

Fig. 2. Example of a situation

за другими исследователями [Cutfield, 2018, р. 106; Bohnemeyer, 2018, р. 201; Mukhin, Nasledskova, 2025, р. 461], было решено не контролировать этот параметр.

В анкете Д. Уилкинса к каждой ситуации предлагались дополнительные вопросы для учета фактора первого/непервого упоминания. Однако в большинстве случаев в разных ситуациях использовался один и тот же объект, и просьба к консультанту вообразить, что он упоминается впервые, была бы по меньшей мере неестественна. Поэтому вслед

3.2. Сбор и анализ данных

Данные для абазинского языка были собраны в ходе Западнокавказской лингвистической экспедиции Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) в аул Старо-Кувинск (Карачаево-Черкесия) в июле 2024 г. В эксперименте участвовали носители кувинского говора ашхарского диалекта абазинского языка. Всего приняли участие 11 человек: 6 женщин и 5 мужчин от 45 до 85 лет.

Данные для темиргоевского диалекта адыгейского языка были собраны у носителей в Адыгее (город Майкоп и аул Хачемзий). В опросе приняли участие 10 человек: 9 женщин и 1 мужчина от 19 до 73 лет.

Результаты экспериментов представлены на рисунках 3–8 в виде таблиц, отражающих вероятность употребления каждого демонстратива в каждой из ситуаций. Результаты округлены до одного знака после запятой. Изображения ситуаций распределены по убыванию вероятности слева направо и сверху вниз. Предложенные в нескольких ответах производные формы учитывались как соответствующие им простые формы демонстративов, но в силу небольшого числа этих форм это не должно повлиять на наши выводы.

4. Принципы устройства ашхарской и адыгейской системы демонстративов

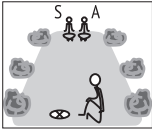
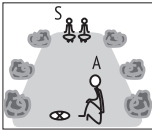
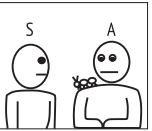
4.1. Предварительные наблюдения

В первую очередь нужно отметить, что ни ашхарская, ни адыгейская системы не могут быть описаны как исключительно дистантно-ориентированные или лично-ориентированные.

Расстояние не может являться единственным фактором, т.к. этому противоречат, например, ситуации 13 и 16, в которых расстояние от говорящего одинаковое, но различается расположение адресата (таблица 3).

Таблица 3

Ситуации, иллюстрирующие неединственность факторов [Situations illustrating that multiple factors are involved]

Ситуация [Situation]	Ашхарский [Ashqarawa]	Адыгейский [West Circassian]
13 	<i>ани</i> – 0,45 <i>ауи</i> – 0,45	<i>мо</i> – 0,8 <i>а</i> – 0,2
16 	<i>ани</i> – 0,27 <i>ауи</i> – 0,73	<i>мо</i> – 0,35 <i>а</i> – 0,7
5 	<i>апи</i> – 0,64 <i>ани</i> – 0,18 <i>ауи</i> – 0,27	<i>мы</i> – 0,25 <i>мо</i> – 0,7 <i>а</i> – 0,15

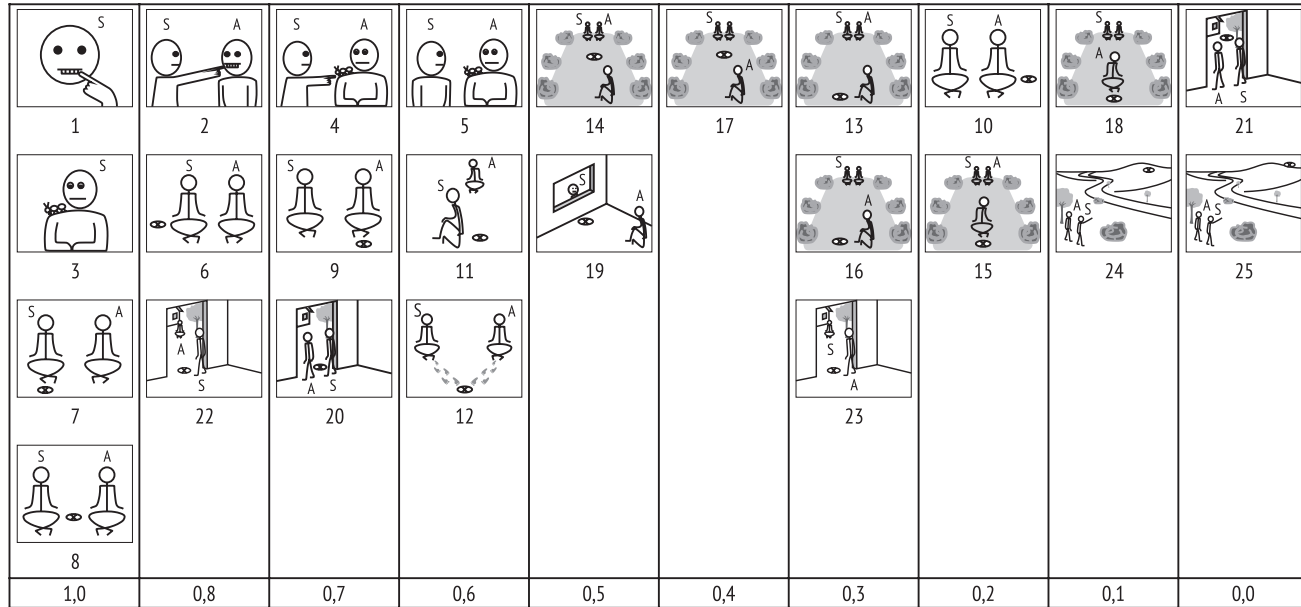


Рис. 3. Абазский язык: вероятность использования *aru*

Fig. 3. Abaza: Probability of using *aru*

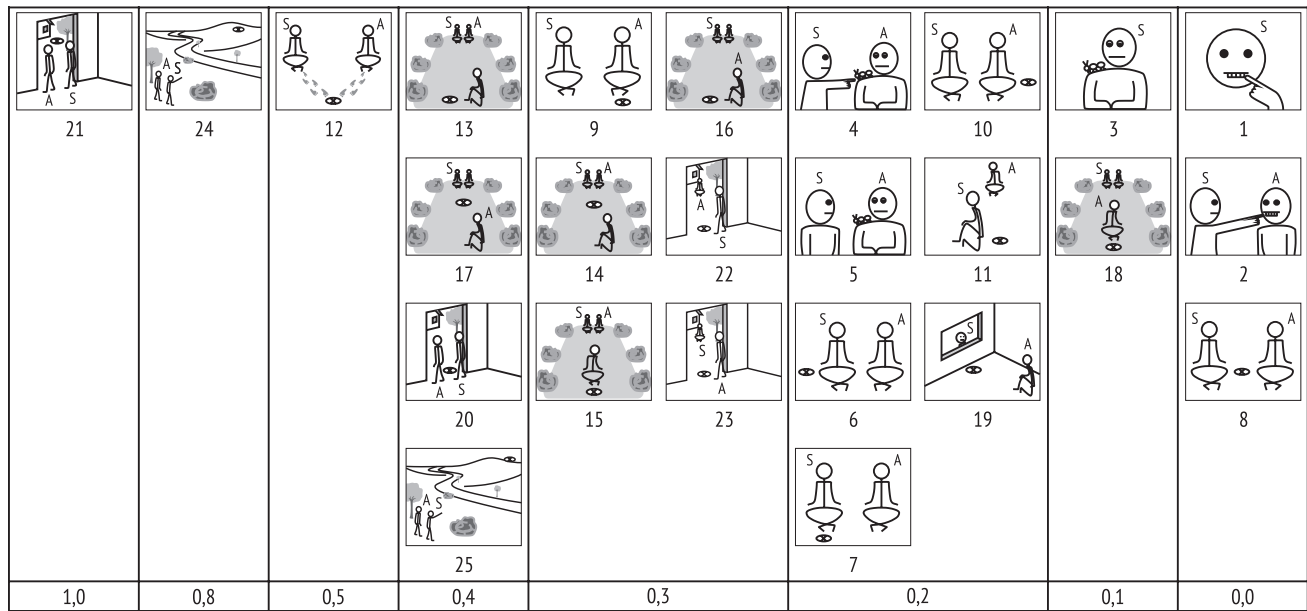
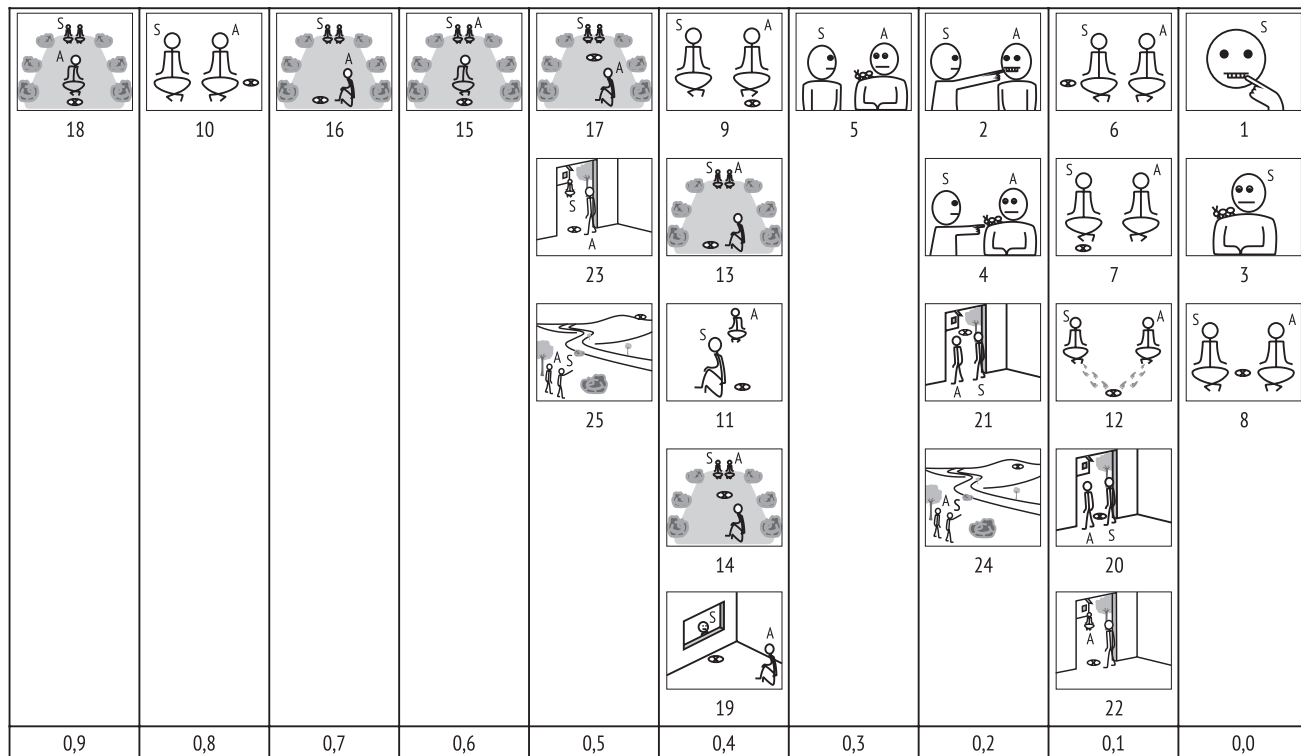


Рис. 4. Абазинский язык: вероятность использования *ани*

Fig. 4. Abaza: Probability of using *anəj*

Рис. 5. Абазинский язык: вероятность использования *ayu*Fig. 5. Abaza: Probability of using *ayu*

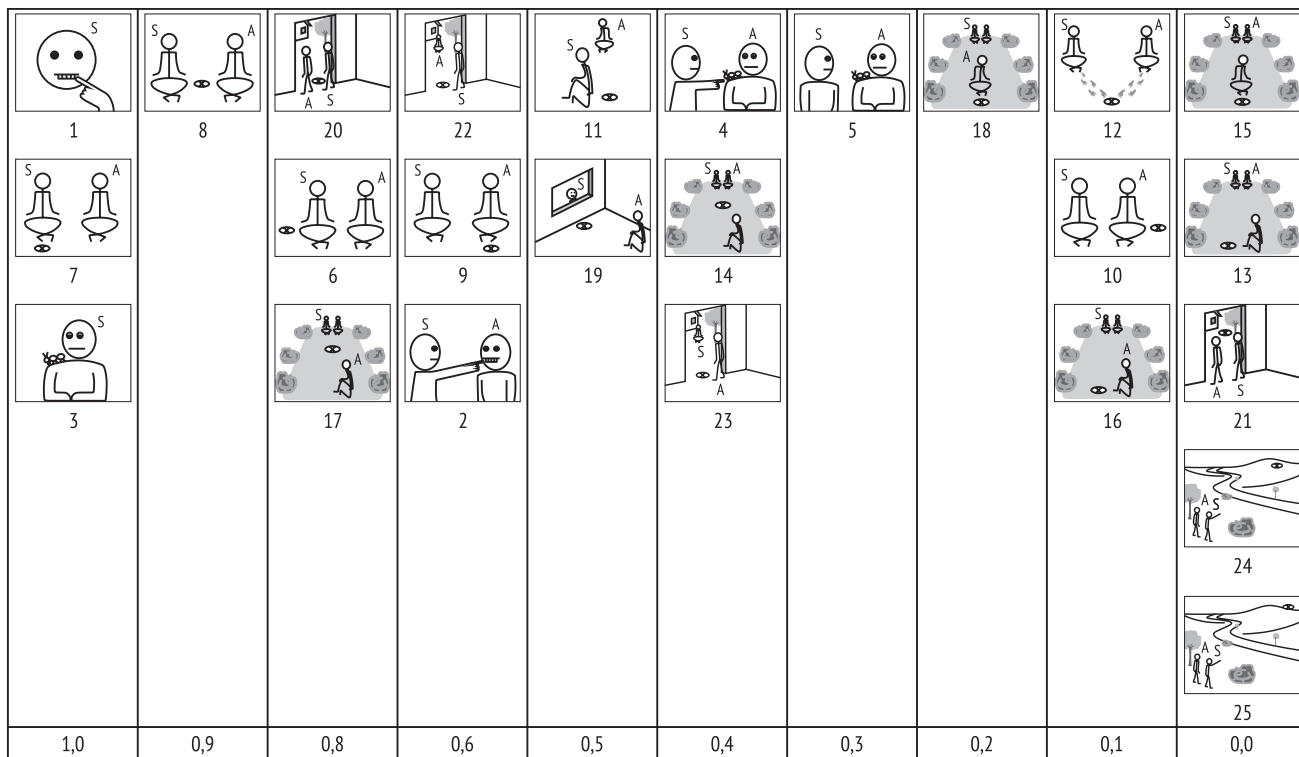


Рис. 6. Адыгейский язык: вероятность использования *мы*

Fig. 6. West Circassian: Probability of using *та*

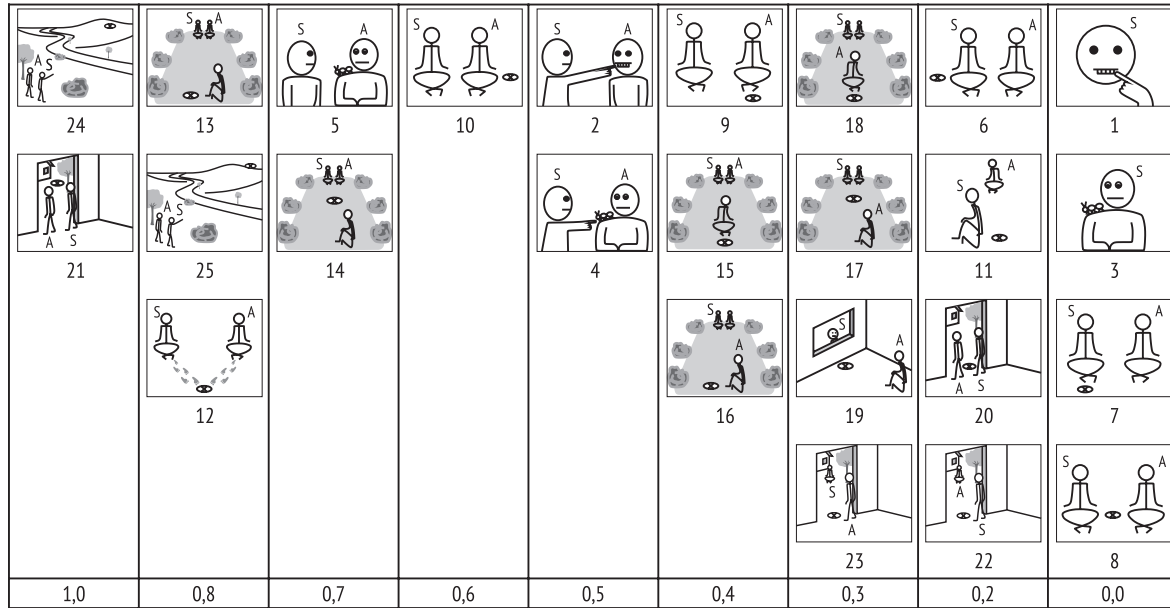


Рис. 7. Адыгейский язык: вероятность использования *mo*

Fig. 7. West Circassian: Probability of using *mwe*

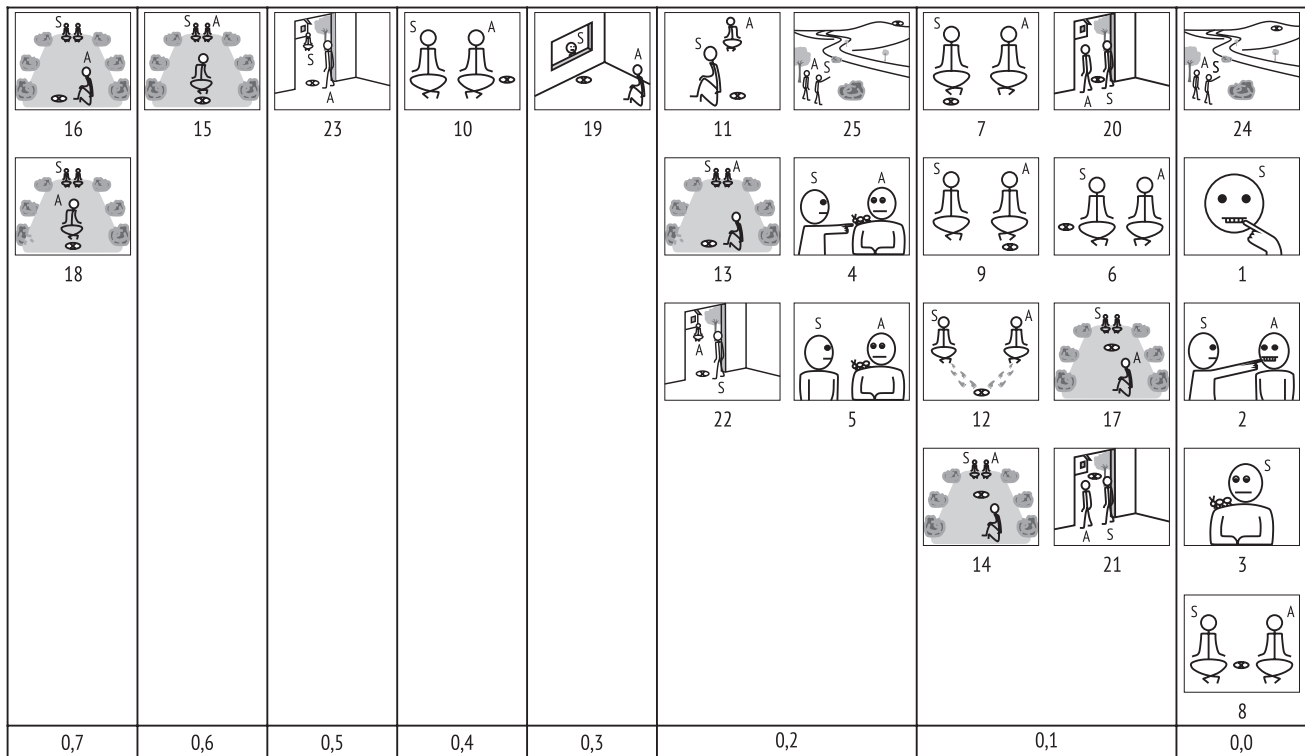


Рис. 8. Адыгейский язык: вероятность использования *a*

Fig. 8. West Circassian: Probability of using *a*

Описать употребление демонстративов исключительно через ассоциацию с первым или вторым лицом тоже не получится. Это показывает сопоставление ситуаций 16 и 5, в которых референт находится ближе к адресату, чем к говорящему, но предпочитают разные демонстративы (см. таблицу 3). Обозначение референтов в сравниваемых ситуациях различается для обоих языков, однако это различие не идентично.

Далее для экспериментов сначала проводится анализ значения каждого из демонстративов, а потом сопоставляются разные факторы. При этом релевантное различие между вероятностью появления демонстративов условно принимается в 0,2: другими словами, если разница составляет меньше 0,2, мы считаем расхождение с большой вероятностью случайным (хотя иногда все же отмечаем).

4.2. Значения ашхарских демонстративов

Вероятность употребления демонстратива первого разряда *ari* (см. рис. 3) в первую очередь зависит от расстояния от говорящего до референта. В ситуациях, в которых вероятность 0,5 и ниже, референт находится на большом расстоянии от говорящего (не в легкой доступности, т.е. >4 по шкале расстояния; ср. ситуации 17, 13, 16, 23, 15, 18, 24, 21, 25). В ситуациях, в которых вероятность выше 0,5, расстояние, наоборот, небольшое (≤ 4 по шкале).

Кроме того, невидимость референта говорящему сильно понижает вероятность *ari* (ситуации 10, 11). Также понижают вероятность использования *ari*, несмотря на другие благоприятные условия, такие факторы, как наличие пространственных границ (ситуации 22, 20, 19) и нахождение референта в явной зоне адресата, т.е. когда речь идет о части тела адресата или о насекомом на нем или когда объект располагается непосредственно перед адресатом (ситуации 2, 4, 9, 5). Указательный жест увеличивает вероятность *ari*, если референт находится близко к говорящему, но в зоне адресата (ситуации 4 vs 5).

Наиболее типичная ситуация для употребления демонстратива второго разряда *ani* (см. рис. 4) – если локуторы находятся рядом друг с другом, а референт им виден и достаточно удален от них обоих (>4 по шкале: ситуации 21, 24). Нахождение референта между локуторами (внутри коммуникативной диады [Jungbluth, 2003]; ситуация 17) и наличие третьего человека (ситуации 13, 14) сильно снижает вероятность употребления *ani*.

Демонстратив третьего разряда *au* (см. рис. 5) противопоставляется демонстративу I по расстоянию от говорящего до референта (ситуации 18, 16, 17, 23), демонстративам I и II – по видимости (ситуации 18, 10, 15, 25, 27), и отдельно демонстративу II – по ориентированности

на адресата (ситуация 16). Таким образом, наиболее типичным употреблением для *ауи* оказывается ситуация, когда референт находится на большом расстоянии (>4 по шкале) от говорящего, но близко к адресату (т.е. в зоне адресата) и при этом не виден говорящему, чему полностью соответствует ситуация 18. Доказательством влияния расположения адресата в контекстах, где демонстратив I не должен употребляться, выступает ситуация 16. При этом *ауи* может выбираться в ситуации невидимости, даже если референт находится недалеко от говорящего (ситуация 10). У двух носителей употребление *ауи* ограничивается контекстами, где референт не виден говорящему.

По результатам эксперимента мы можем обобщить основные значения демонстративов в ашхарском диалекте абазинского языка:

- *ари* (I) обозначает объект, находящийся близко к говорящему;
- *ани* (II) обозначает объект, находящийся одинаково далеко от говорящего и от адресата;
- *ауи* (III) обозначает объект, находящийся далеко от говорящего и близко к адресату, или невидимый объект.

4.3. Сопоставление факторов в ашхарской системе

Факторы, влияющие на выбор демонстративов, могут конкурировать друг с другом. Как было замечено, для указательного местоимения первого разряда основным фактором является «близость» к говорящему. Близким оказывается референт, до которого можно дотянуться рукой (ситуации 1–9). Если референт находится в нескольких шагах/метрах от говорящего, возникает вариативность (ситуация 12: I – 0,64, II – 0,55, III – 0,09; ситуация 14: I – 0,55, II – 0,27, III – 0,36), но все еще более вероятен I.

Интересно посмотреть, как соотносятся фактор расстояния и фактор ассоциации со вторым лицом. Объект, находящийся ближе к адресату, чем к говорящему, может обозначаться демонстративом III, но выбор демонстратива I все же преобладает, если референт на небольшом расстоянии от говорящего (ситуация 5: I – 0,64, II – 0,18; III – 0,27; ситуация 9: I – 0,73, II – 0,27; III – 0,45). В ситуации 10 демонстратив I уже почти не употребляется, но эта ситуация не очень иллюстративна, потому что референт мог быть не виден говорящему (I – 0,18, III – 0,82). Можно сделать вывод, что при выборе демонстратива I расстояние важнее, чем включение в зону адресата. Если расстояние «маленькое» (и референт виден), предпочитается демонстратив I, а если «большое», приходится выбирать между демонстративами II или III. На «большом» расстоянии от говорящего включение референта в зону адресата уже оказывается более значимым. Демонстратив III выбирается, если референт находится в зоне адресата (ситуация 16: II – 0,27, III – 0,73),

а II – если референт равноудален от обоих локуторов (ситуация 21: II – 1, III – 0,18).

Видимость тоже является определяющим фактором. На «большом» расстоянии для обозначения референта, который не виден говорящему, используется демонстратив III (ситуация 15: II – 0,27, III – 0,64; ситуация 18: III – 0,91; ситуация 25: II – 0,36, III – 0,55). В ситуации 10, когда референт находится на некотором расстоянии от говорящего, близко к адресату и при этом может быть не виден, все еще предпочитается III (0,82). Однако ситуация 11 показывает, что на «маленьком» расстоянии наблюдается вариативность (I – 0,64, III – 0,36) с перевесом в сторону демонстратива I. Таким образом, непосредственная близость к говорящему важнее невидимости.

Наличие пространственных границ «отдаляет» референт как от говорящего (ситуация 20: I – 0,73, II – 0,36; ситуация 19: I – 0,55, II – 0,36), так и от адресата (ситуация 23: I – 0,27, II – 0,27, III – 0,55). Однако в обоих случаях преобладает демонстратив, который бы ожидался без учета фактора пространственной границы. Дверной проем можно считать менее жесткой границей, чем оконная рама (т.к. он легче преодолевается), что отражается в результатах эксперимента.

В ситуации 17 (I – 0,45, II – 0,45, III – 0,55), когда референт находится между локуторами на большом расстоянии, наблюдается сильная вариативность, т.к. эта ситуация не является типичной ни для одного из демонстративов. Референт расположен далеко от обоих локуторов, но в их «общей зоне». Вариативность в ситуациях 19 и 23 также можно связать не только с наличием пространственных границ, но и с тем, что, если пространственная граница «отдаляет» референт, система не предлагает однозначного варианта его обозначения, т.к. он находится между локуторами.

Сопоставление ситуаций 13 и 24 показывает, что наличие третьего человека тоже может иметь значение. В ситуации 13 (I – 0,27, II – 0,45, III – 0,45) ожидалось бы *ани*, т.к. референт удален от обоих локуторов, однако мы видим вариативность *ани/ауи*. Видимо, это связано с тем, что референт находится рядом с третьим человеком.

Анкета не фокусировалась на том, чтобы определить, влияет ли на выбор указательного местоимения наличие жеста. Минимальная пара (ситуации 4 и 5) – плохой диагностический контекст для ашхарского, т.к. референт находится на расстоянии руки от говорящего и, соответственно, здесь предполагается использование ближнего дейктика. Только у одного носителя было зафиксировано различие в выборе демонстратива в этих ситуациях (I и III соответственно). Кроме того, один носитель в ситуации 12 употребил демонстратив I при наличии жеста, и когда его

попросили убрать жест, ответ изменился на II. Тем не менее этот фактор, безусловно, требует отдельного изучения.

4.4. Значения темиргоевских адыгейских демонстративов

Результаты эксперимента показали, что адыгейский также имеет трехчастную систему пространственных дейктиков, состоящую из местоимений *мы*, *мо* и *а*; четвертый демонстратив *джа* при ответах не появился ни разу.

Демонстратив первого разряда *мы* (см. рис. 6) выбирается, если референт включен в зону говорящего: если он расположен на расстоянии ≤ 3 по шкале от говорящего и не находится ближе к адресату, чем к говорящему. Если референт находится между или почти между локуторами на одинаковом расстоянии от них, вне зависимости от величины расстояния (ситуации 8, 17, 20) тоже выбирается *мы*. Указательный жест увеличивает вероятность *мы*, если референт находится близко к говорящему, но в зоне адресата (ситуации 4 vs 5). Если референт не виден говорящему, вероятность *мы* понижается (ситуации 6, 7 vs 11). Наличие пространственной границы между говорящим и референтом также понижает вероятность *мы* (ситуации 19, 22).

Демонстратив второго разряда *мо* (см. рис. 7) появляется в двух случаях, если референт не включен в зону говорящего:

1) референт одинаково удален от обоих локуторов (≥ 4 по шкале от говорящего и адресата: ситуации 12, 13, 14, 21, 24, 25);

2) референт находится на небольшом расстоянии от обоих локуторов, но ближе к адресату, чем к говорящему (≤ 3 по шкале от адресата: ситуации 2, 4, 5, 10).

Вопреки выводам Д.В. Барыльниковой (2013), в условиях эксперимента *мо* оказывался легко допустим и без жеста рукой.

Наконец, демонстратив третьего разряда *а* (см. рис. 8) появляется, если референт находится далеко от говорящего, но близко к адресату (ситуации 16, 18, 23) или если референт не виден говорящему (ситуации 15, 18). Однако даже в типичных для III контекстах *а* конкурирует с *мо*, и вероятность его появления не превышает значения 0,7.

Обобщим значения демонстративов:

- *мы* (I) обозначает объект, находящийся в зоне говорящего;
- *мо* (II) обозначает объект, находящийся одинаково далеко от обоих локуторов или находящийся на небольшом расстоянии от говорящего, но ближе к адресату;
- *а* (III) используется, когда объект расположен далеко от говорящего, но близко к адресату, или когда референтом выступает невидимый объект.

4.5. Сопоставление факторов в адыгейской системе

Сперва проанализируем взаимодействие факторов расстояния и ассоциации с первым и вторым лицом. На небольшом расстоянии от говорящего выбирается демонстратив первого разряда, если референт не находится ближе к адресату. В этом случае факт соотносённости с зоной локутора важнее, чем расстояние. Нахождение референта в зоне адресата сильно понижает вероятность демонстратива I, который в таком контексте конкурирует с демонстративом II (ситуация 2: I – 0,55, II – 0,45; ситуация 9: I – 0,57, II – 0,43; ситуация 4: I – 0,4, II – 0,45; ситуация 5: I – 0,25, II – 0,7).

На большом расстоянии от говорящего референт, расположенный ближе к адресату, обозначается демонстративом III (ситуация 16: II – 0,35, III – 0,7). Если же референт удален от обоих локуторов, выбирается демонстратив II. Таким образом, демонстратив II можно считать немаркированным указательным местоимением, которое выбирается, когда не могут быть употреблены демонстративы I или III. Иногда демонстратив II – как немаркированный компонент – возникает и в контекстах, типичных для демонстратива III (ситуация 16: 0,35, ситуация 18: 0,30), но выбор демонстратива II в этих ситуациях не является последовательной стратегией.

Если референт находится в общей зоне, т.е. между или почти между локуторами, предпочитается демонстратив I, вне зависимости от расстояния (ситуация 8: I – 1; ситуация 17: I – 0,75, II – 0,3; ситуация 20: I – 0,8, II – 0,2).

Невидимые объекты могут обозначаться демонстративом III, но фактор видимости является значимым лишь для некоторых участников эксперимента. Минимальные пары по видимости, когда референт находится рядом с третьим лицом, показывают, что этот фактор повышает вероятность демонстратива III (ситуация 13: III – 0,2, ситуация 15: III – 0,57). Однако в минимальных парах на большом расстоянии и без наличия третьего лица фактор видимости оказывается не столь существенным (ситуация 24: II – 0,96, III – 0,03; ситуация 25: II – 0,8, III – 0,2). Кроме того, вариативность, связанную с фактором видимости, показывают ситуации 10 (II – 0,63, III – 0,43) и 11 (I – 0,53, III – 0,23). Таким образом, хотя III и может обозначать невидимые объекты, в некоторых ситуациях и для некоторых носителей параметр видимости незначим.

Пространственные границы, как и в ашхарском, «отделяют» референт от локутора, отделенного границей (ситуация 19: I – 0,45, III – 0,3; ситуация 20: I – 0,8, II – 0,2; ситуация 22: I – 0,6, II – 0,2; ситуация 23: I – 0,35, III – 0,5). Если этот эффект имеет место, то референт может быть

отнесен к зоне локутора, с которым он находится в едином пространстве, или может восприниматься как удаленный от обоих участников.

Минимальная пара по наличию жеста показывает, что, если референт находится на небольшом расстоянии от говорящего, но ближе к адресату, жест может способствовать выбору I (ситуация 4 с жестом: I – 0,4, II – 0,45; ситуация 5 без жеста: I – 0,25, II – 0,7).

5. Сопоставление ашхарской и адыгейской систем демонстративов

Очевидно, дистрибуция ситуаций для абазинского и адыгейского языков похожа. Демонстратив I обозначает референт, находящийся близко к говорящему. Демонстратив II обозначает референт, находящийся далеко от обоих локуторов. Демонстратив III используются, когда референт расположен далеко от говорящего, но близко к адресату или когда он не виден говорящему. Итак, системы двух родственных языков схожи несмотря на то, что они не содержат очевидных родственных элементов. Однако эти системы неидентичны.

Основное различие состоит в выборе демонстративов, когда референт находится на небольшом расстоянии от говорящего, но ближе к адресату. Ни в одном из языков преобладающий в этом случае демонстратив не соответствует тому, который характерен для референта на большом расстоянии от говорящего и близко к адресату. В ашхарском материале в таких ситуациях предпочитается I, который выбирается на основании расстояния (≤ 4 по шкале от говорящего). В адыгейском используется не I, который обычно выбирается на основании включения референта в зону говорящего и не III, ассоциированный с адресатом на большом расстоянии, а II – немаркированный компонент системы.

Кроме того, между языками обнаруживается отличие в дистанции, на которой может находиться объект, обозначающийся демонстративом I. Если до референта можно дотянуться рукой (≤ 3 по шкале), он может маркироваться I как в ашхарском, так и в адыгейском. В случае, когда референт находится на расстоянии нескольких шагов или метров от говорящего, в ашхарском предпочитается I, а в адыгейском – немаркированный II. Как видно, для ашхарского граница пространства говорящего немного шире.

Адыгейская система, в отличие от ашхарской, имеет правило трактовки ситуаций, когда референт находится между локуторами (на одинаковом расстоянии). Такой объект относится к зоне говорящего и, соответственно, преимущественно обозначается I. В ашхарском же, если такие ситуации не поддаются какому-либо другому правилу, возникает

сильная вариативность, показывающая, что эти случаи не вписываются в типичные значения ни одного из демонстративов. Аналогичная проблема описывается для бразильского португальского [Meira, Guirardello-Damian, 2018, p. 127] и для языка гоэмаи [Hellwig, 2018, p. 142].

Как в ашхарской системе, так и в адыгейской демонстратив III ассоциируется также с невидимостью говорящему. Фактор видимости является определяющим для ашхарского абазинского, т.е. выбор демонстратива явно зависит от этого фактора. В адыгейском же он влияет на выбор демонстратива только у некоторых носителей.

Также и в ашхарском абазинском, и в адыгейском на выбор демонстративов может влиять наличие пространственных границ и наличие жеста, но влияние этих факторов невелико.

Итак, мы можем обобщить основные различия систем.

1. В ситуации, когда референт находится недалеко от говорящего, но ближе к адресату, в ашхарском предпочитается демонстратив I, а в адыгейском – демонстратив II.

2. С точки зрения расстояния для ашхарского граница пространства говорящего немного шире, чем для адыгейского.

3. Если референт находится между локуторами на большом расстоянии от обоих, в адыгейском предпочитается демонстратив I, а в ашхарском наблюдается вариативность.

4. Фактор видимости в ашхарском более значим, чем в адыгейском.

6. Интерпретация результатов

Ашхарская и адыгейская системы не являются лично-ориентированными или дистантно-ориентированными в строгом смысле, поскольку для них одновременно существенны несколько разных факторов. Далее мы рассмотрим возможные подходы к анализу этих двух систем в более широком типологическом и теоретическом контексте.

Ассоциация с адресатом – важный параметр при выборе демонстратива в обоих языках, что свойственно лично-ориентированным системам. Тогда, используя терминологию С. Левинсона [Levinson, 2018], можно описать набор элементов ашхарской системы демонстративов как S-proximal (близко к говорящему, I), A-proximal (близко к адресату, III) и S&A-distal (далеко от говорящего и адресата, II). В адыгейском же выделяется S-proximal (близко к говорящему, I), A-proximal (близко к адресату, III) и немаркированный элемент системы (II). В обоих языках A-proximal ограничен контекстами, в которых референт находится далеко от говорящего. Правда, в ашхарском A-proximal возникает и в контекстах, когда референт недалеко от говорящего (ситуация 5: 0,27, ситуация 9: 0,45), но ни один из опрошенных носителей не придерживается

последовательно такой стратегии. Это можно объяснить тем, что для ассоциации с адресатом наиболее типичны контексты с референтом, находящимся далеко от говорящего, в иных же контекстах возникает конкуренция между ассоциацией с говорящим и ассоциацией с адресатом.

Похожим образом описывается система кининского рутульского [Mukhin, Nasledskova, 2025], содержащая S-proximal, A-proximal и S(&A)-distal демонстративы. A-proximal тоже в основном используется, если референт находится далеко от говорящего и близко к адресату. В контекстах, где референт находится близко к адресату, могут возникать S-proximal и (у некоторых носителей) S(&A)-distal, в зависимости от расстояния от говорящего. Однако это отличается от ситуации в адыгейском: немаркированный компонент не может быть назван S(&A)-distal, т.к. он встречается и в ситуациях, в которых референт находится близко, на расстоянии руки от говорящего (ситуации 2, 4, 5).

При таком описании ашхарских и адыгейских данных существенным оказывается не только фактор (не)ассоциированности с одним из локуторов, но и параметр расстояния от говорящего. В системе нет элемента, который всегда употребляется, если референт находится в зоне адресата, т.к. выбор демонстратива зависит и от дистанции. Таким образом, параметр ассоциированности с одним из локуторов нужно рассматривать наравне с параметром расстояния от говорящего и другими параметрами, которые влияют на включенность референта в зону говорящего.

Мы полагаем, что рассматриваемые системы не являются лично-ориентированными в строгом смысле и не полицентричны, т.к. ассоциированность с адресатом – лишь один из параметров, «отдаляющих» референт от говорящего. Основываясь на этом, мы предлагаем два возможных описания обсуждаемых систем.

Первое описание фактически подсчитывает количество факторов, препятствующих включению референта в зону говорящего. Соответствующее представление систем ашхарских и адыгейских демонстративов предлагается в таблице 4. Основные параметры – расстояние от говорящего и ассоциированность с адресатом – удобно анализировать как ограничения на включенность в зону говорящего. Объект в меньшей степени ассоциирован с говорящим, если он от него отдален и/или если он находится в зоне адресата. Именно поэтому признаки в таблице 4 сформулированы отрицательно³.

³ Для описания наших данных эти признаки могут быть сведены к бинарному противопоставлению, т.к. для «среднего» расстояния нет специального демонстратива, в таком случае возникают колебания между ближним и средним/дальним демонстративом. Однако нужно оговориться, что граница между «+» и «-» для расстояния различается: в ашхарском I может использоваться при большем расстоянии от говорящего, чем в адыгейском (см. раздел 5).

Таблица 4

**Ашхарская и адыгейская системы демонстративов:
два параметра**
[Ashqarawa and West Circassian demonstrative systems:
Two parameters]

Ашхарский [Ashqarawa]	He [далеко от говорящего] [Not [far from the speaker]]	He [в зоне адресата] [Not [in the addressee's domain]]	Адыгейский [West Circassian]
I	+	+	I
	+	–	II
II	–	+	
III	–	–	III

Итак, комбинации значений параметров выстраиваются в шкалу включенности в зону говорящего, отображаемой количеством «+»: референт включен в зону говорящего (+ +), один из параметров мешает референту быть включенным в зону говорящего (+ – или – +) и, наконец, референт никак не может ассоциироваться с говорящим (– –). С этими контекстами можно сопоставить демонстративы, которые в них обычно употребляются в ашхарском и в адыгейском. Во втором контексте (+ –) два языка обнаруживают различие: если референт недалеко от говорящего и в зоне адресата, в ашхарском выбирается демонстратив I, а в адыгейском – демонстратив II.

Как было установлено в разделе 4, значимым также может быть параметр видимости. Если невидимость влияет, выбирается демонстратив III, однако этот параметр не всегда оказывается решающим. Предложенная схема описания позволяет отразить и это; см. таблицу 5, дополнительно включающую параметр видимости. Мы не можем однозначно упорядочить все контексты, соответствующие одному демонстративу, поэтому их порядок отчасти условный. Ашхарский и адыгейский ведут себя по-разному в ситуации 10, когда говорящий и адресат сидят рядом, а референт находится сбоку от адресата, так, что он не виден говорящему (+ – –). В ашхарском в таком случае выбирается III, а в адыгейском предпочитается II (т.е. параметр видимости не значим). Также в ситуациях, когда референт находится далеко от говорящего и адресата и не виден, в адыгейском наблюдается вариативность между II и III. Такое широкое использование адыгейского демонстратива II отражает его «немаркированность», но его значения могут быть закономерно описаны с помощью предлагаемого подхода.

**Ашхарская и адыгейская системы демонстративов:
три параметра**
**[Ashqarawa and West Circassian systems of demonstratives:
Three parameters]**

Ашхарский [Ashqarawa]	Не [далеко от говорящего] [Not [far from the speaker]]	Не [в зоне адресата] [Not [in the addressee's domain]]	Видимый [Visible]	Адыгейский [West Circassian]
I	+	+	+	I
	+	+	–	
	+	–	+	II
II	–	+	+	
III	+	–	–	II / III
	–	+	–	
	–	–	+	III
	–	–	–	

Второе возможное описание по-прежнему учитывает то, что ассоциация с адресатом мешает включению референта в зону говорящего, но при этом исходит из того, что основное противопоставление в этих системах возникает между демонстративом первого разряда (референт рядом с говорящим или ассоциирован с говорящим) и демонстративом второго разряда (референт не рядом с говорящим и не ассоциирован с говорящим). Демонстратив III, в свою очередь, появляется, когда говорящий вообще не выступает как точка отсчета, с которой соотносится объект. Это может касаться трех типов случаев:

(i) референт не виден говорящему;

(ii) референт находится настолько близко к адресату, но при этом настолько далеко от говорящего, что адресат перенимает роль точки отсчета – с этим связан комментарий некоторых носителей адыгейского к ситуации 16, согласно которому демонстратив II не должен в ней употребляться, поскольку он невозможен для адресата;

(iii) отсылка к референту не носит собственно дейктический характер – именно поэтому местоимения разряда III обычно употребляются в качестве анафорических (см. [Генко, 1955, с. 103–104] для абазинского языка и [Arkadiev, Lander, 2020, p. 27; Унарокова, 2025] для адыгейского языка).

Так или иначе, рассматриваемые системы демонстративов могут быть охарактеризованы как ориентированные на включенность в зону говорящего, т.е. они в широком смысле моноцентричны, несмотря на то, что отношения между адресатом и референтом являются релевантными.

7. Выводы

По итогам экспериментального исследования на основе анкеты [Wilkins, 2018] мы можем описать основные значения демонстративов в двух абхазо-адыгских идиомах: ашхарском абазинском и темиргоевском адыгейском. В таблице 6 приведены условия, при которых референт в рассматриваемых языках обозначается тем или иным демонстративом.

Таблица 6

Значения демонстративов в ашхарском и адыгейском [The meanings of demonstratives in Ashqarawa and West Circassian]

	Ашхарский [Ashqarawa]		Адыгейский [West Circassian]	
I	<i>ари</i>	Близко (≤ 4 по шкале) к S [Close to S (≤ 4 on the scale)]	<i>мы</i>	Близко (≤ 3 по шкале) к S [Close to S (≤ 3 on the scale)]
II	<i>ани</i>	Далеко от S и A [Far from both S and A]	<i>мо</i>	Далеко от S и A или недалеко от S, но ближе к A [Far from both S and A or close to S but closer to A]
III	<i>ауи</i>	Далеко от S и близко к A или не виден S [Far from S and close to A or not visible to S]	<i>а</i>	Далеко от S и близко к A или (иногда) не виден S [Far from S and close to A or (sometimes) not visible to S]

Примечание. S – говорящий, A – адресат. [Note. S is the speaker, A is the addressee]

Таким образом, основные факторы, влияющие на выбор демонстратива, – ассоциация с одним из локуторов, расстояние от говорящего и (не-)видимость. В меньшей степени влияет наличие пространственных границ, третье лицо и наличие жеста. Выбор ни одного из демонстративов не задается исключительно параметром дистанции (как в моноцентричных дистантно-ориентированных системах) или параметром ассоциированности с говорящим или адресатом (как в полицентричных лично-ориентированных системах).

Эти данные указывают на неединственность факторов, влияющих на выбор демонстратива. В то же время, как мы показали в разделе 6, несмотря на релевантность адресата, рассматриваемые системы можно

интерпретировать как ориентированные на говорящего и, соответственно, моноцентричные (при том что на включенность референта в зону говорящего влияют разные параметры). Не исключено, что и некоторые другие системы, которые описываются как лично-ориентированные и полицентричные, могут быть проанализированы аналогично – как ориентированные на говорящего и моноцентричные, где ассоциированность с адресатом является лишь одним из параметров, влияющих на включенность в зону говорящего. Этот вопрос остается открытым для будущих исследований.

Библиографический список / References

Барыльникова, 2013 – Барыльникова Д.В. Пространственный дейксис в темиргоевском диалекте адыгейского языка. Рукопись. М., 2013. [Barylnikova D.V. Prostranstvennyi deiksis v temirgoevskom dialekte adygeiskogo yazyka [Spatial deixis in the Temirgoy dialect of the West Circassian language.] Ms. Moscow, 2013.]

Генко, 1955 – Генко А.Н. Абазинский язык. Грамматический очерк наречия тапанта. М., 1955. [Genko A.N. Abazinskii yazyk. Grammaticheskii ocherk narechiya tapanta [Abaza language: A grammatical sketch of the Tapanta dialect]. Moscow, 1955.]

Керашева, 1957 – Керашева З.И. Особенности шапсугского диалекта адыгейского языка. Майкоп, 1957. [Kerasheva Z.I. Osobennosti shapsugskogo dialekta adygeiskogo yazyka. [Specifics of the Shapsug dialect of the West Circassian language]. Майкоп, 1957.]

Кибрик, 2005 – Кибрик А.Е. Константы и переменные языка. СПб., 2005. [Kibrik A.E. Konstanty i peremennyye yazyka [Constants and variables of language]. St. Petersburg, 2005.]

Клычев, 2000 – Клычев Р.Н. Абазинский литературный язык. Поиски и находки. Карачаевск, 2000. [Klychev R.N. Abazinskii literaturnyi yazyk. Poiski i nakhodki [The Abaza literary language: Searches and findings]. Karachaevsk, 2000.]

Кумахов, 1989 – Кумахов М.А. Сравнительно-историческая грамматика адыгских (черкесских) языков. М., 1989. [Kumakhov M.A. Sravnitelno-istoricheskaya grammatika adygsikh (cherkesskikh) yazykov [Comparative-historical grammar of the Adyghe (Circassian) languages]. Moscow, 1989.]

Кумахов, 2006 – Кабардино-черкесский язык. Т. I. Создание письменности, фонетика и фонология, морфология, синтаксис / под ред. М.А. Кумахова. Нальчик, 2006. [Kabardino-cherkesskii yazyk [Kabardian-Circassian language]. Vol. I. Creation of writing system, phonetics and phonology, morphology, syntax. Nalchik, 2006.]

Ландер, Унарокова, 2024 – Ландер Ю.А., Унарокова Ш.Ш. Принципы выделения местоименных серий (на материале адыгейских демонстративов) // Вопросы языкознания. 2024. № 2. С. 93–121. [Lander Yu.A., Unarokova Sh.Sh. Defining pronominal series: Demonstratives in West Circassian. *Voprosy jazykoznanija*. 2024. No. 2. Pp. 93–121. (In Rus.)]

Ломтатидзе, 2006 – Ломтатидзе К. Абазинский язык (краткое обозрение). Тбилиси, 2006. [Lomtaticidze K. Abazinskii yazyk (kratkoe obozrenie) [The Abaza language (a brief survey)]. Tbilisi, 2006.]

Мерданова, Федорова, 2002 – Мерданова С.Р., Федорова О.В. Дейктическая система хпокского диалекта агульского языка // Исследования по теории грамматики. Вып. 2. Грамматикализация пространственных значений / под ред. В.А. Плуңгяна. М., 2002. С. 220–231. [Merdanova S.R., Fedorova O.V. The deictic system of the Huppuq' dialect of Agul. *Issledovaniya po teorii grammatiki*. Vol. 2: Grammaticalization of spatial meanings. V.A. Plungian (ed.). Moscow, 2002. Pp. 220–231. (In Rus.)]

Падучева et al., 2011–2020 – Падучева Е.В., Пекелис О.Е., Стойнова Н.М. Местоимение. Материалы для проекта корпусного описания русской грамматики (rusgram.ru). Рукопись. Москва, 2011–2020. URL: <http://rusgram.ru/new/chapter/pos/pronoun/> (дата обращения: 09.01.2025). [Paducheva E.V., Pekelis O.E., Stoinova N.M. Mestoimenie. Materialy dlya proekta korpusnogo opisaniya russkoi grammatiki (rusgram.ru) [Pronoun. Materials for the project of corpus description of Russian grammar (rusgram.ru)]. Ms. Moscow, 2011–2020. URL: <http://rusgram.ru/new/chapter/pos/pronoun/> (accessed: 09.01.2025).]

Рогава, Керашева, 1966 – Рогава Г.В., Керашева З.И. Грамматика адыгейского языка. Краснодар; Майкоп, 1966. [Rogava G.V., Kerasheva Z.I. Grammatika adygeiskogo yazyka [Grammar of the West Circassian language]. Krasnodar; Maikop, 1966.]

Ростовцев-Попель, 2009 – Ростовцев-Попель А.А. Типология демонстративов: средние дейктики // Вопросы языкознания. 2009. № 2. С. 22–34. [Rostovtsev-Popiel A.A. Typology of demonstratives: Medial deictics. *Voprosy jazykoznanija*. 2009. No. 2. Pp. 22–34. (In Rus.)]

Табулова, 1976 – Табулова Н.Т. Грамматика абазинского языка. Фонетика и морфология. Черкесск, 1976. [Tabulova N.T. Grammatika abazinskogo yazyka. Fonetika i morfologiya [Grammar of the Abaza language: Phonetics and morphology]. Cherkessk, 1976.]

Унарокова, 2025 – Унарокова Ш.Ш. Конкуренция указательных местоимений «а» и «джа» // Сохранение и развитие родного языка в условиях глобализации: современные методы и технологии: материалы VII Международной научно-методологической конференции, посвященной Международному десятилетию языков коренных народов (2022–2032) (Майкоп, 20–21 февраля 2025 г.) / под ред. С.Х. Анчек. Майкоп, 2025. С. 88–91. [Unarokova Sh.Sh. Competition between demonstrative pronouns “a” and “dzha”. *Sokhranenie i razvitie rodnogo yazyka v usloviyakh globalizatsii: sovremennye metody i tekhnologii: materialy VII Mezhdunarodnoi nauchno-metodologicheskoi konferentsii, posvyashchennoi Mezhdunarodnomu desyatiletiyu yazykov korennykh narodov (2022–2032) (Maikop, 20–21 fevralia 2025 g.)*. S.Kh. Anchek (ed.). Maikop, 2025. Pp. 88–91. (In Rus.)]

Федорова, 2001 – Федорова О.В. Пространственная типология указательных местоимений дагестанских языков // Вопросы языкознания. 2001. № 6. С. 28–56. [Fedorova O.V. Spatial typology of demonstratives in Dagestania languages. *Voprosy jazykoznanija*. 2001. No. 6. Pp. 28–56. (In Rus.)]

Шаги́ров, 1977 – Шаги́ров А.К. Этимологический словарь адыгских (черкесских) языков. Т. 1 (А–Н). М., 1977. [Shagirov A.K. Etimologicheskii slovar' adygsikh (cherkessikh) yazykov [Etymological dictionary of Adyghe (Circassian) languages]. Vol. 1 (A–N). Moscow, 1977.]

Шадури, 2006 – Шадури И. Морфология абхазского языка. Тбилиси, 2006. [Shaduri I. Morfologiya abkhazskogo yazyka [Morphology of the Abkhaz language]. Tbilisi, 2006.]

Шлуинский, 2014 – Шлуинский А.Б. Внутригенетическая типология: методологические заметки // Язык. Константы. Переменные. Памяти Александра Евгеньевича Кибрика / под ред. М.А. Даниэля, Е.А. Лютиковой, В.А. Плунгяна и др. СПб., 2014. С. 127–139. [Shluinsky A.B. Intra-genetic typology: Methodological notes. *Yazyk. Konstanty. Peremennye. Pamyati Aleksandra Evgenevicha Kibrika*. M.A. Daniel, E.A. Lyutikova, V.A. Plungian et al. (eds.). St. Petersburg, 2014. Pp. 127–139. (In Rus.)]

Яковлев, Ашхамаф, 1941 – Яковлев Н., Ашхамаф Д. Грамматика адыгейского литературного языка. М.; Л., 1941. [Yakovlev N., Ashkhamaf D. Grammatika adygeiskogo literaturnogo yazyka [Grammar of the West Circassian literary language]. Moscow; Leningrad, 1941.]

Яковлев, 1948 – Яковлев Н.Ф. Грамматика литературного кабардино-черкесского языка. М.; Л., 1948. [Yakovlev N.F. Grammatika literaturnogo kabardino-cherkesskogo yazyka [Grammar of the Kabardian-Circassian literary language]. Moscow; Leningrad, 1948.]

Anderson, Keenan, 1985 – Anderson S.R., Keenan E.L. Deixis. *Language Typology and Syntactic Description*. T. Shopen (ed.). Vol. 3. Cambridge University Press, 1985. Pp. 259–308.

Arkadiev, Lander, 2020 – Arkadiev P., Lander Y. The Northwest Caucasian Languages. *The Oxford Handbook of the Languages of the Caucasus*. Oxford University Press, 2020.

Bhat, 2007 – Bhat D.N.S. Pronouns. Oxford University Press, 2007.

Bohnemeyer, 2018 – Bohnemeyer J. Yucatec Demonstratives in Interaction: Spontaneous versus Elicited Data. *Demonstratives in Cross-Linguistic Perspective*. S. Levinson, S. Cutfield, M. Dunn et al. (eds.). Cambridge University Press, 2018. Pp. 176–205.

Bühler, 1982 – Bühler K. The Deictic Field of Language and Deictic Words. *Speech, Place and Action*. R. Jarvella, W. Klein (eds.). Amsterdam, 1982 [1934]. Pp. 9–30.

Chirikba, 1996 – Chirikba V.A. Common West Caucasian: The Reconstruction of Its Phonological System and Parts of Its Lexicon and Morphology (CNWS Publications 48). Leiden, 1996.

Chirikba, 2003 – Chirikba V.A. Abkhaz. München, 2003.

Colarusso, 1992 – Colarusso J. A Grammar of the Kabardian Language. Calgary, 1992.

Cutfield, 2018 – Cutfield S. Dalabon exophoric uses of demonstratives. *Demonstratives in Cross-Linguistic Perspective*. S. Levinson, S. Cutfield, M. Dunn et al. (eds.). Cambridge University Press, 2018. Pp. 90–115.

Diessel, 1999 – Diessel H. Demonstratives: Form, Function and Grammaticalization (Typological Studies in Language, Vol. 42). Amsterdam, 1999.

Diessel, 2013 – Diessel H. Distance contrasts in demonstratives. *WALS Online* (v2020.4). M.S. Dryer, M. Haspelmath (eds.). 2013. URL: <http://wals.info/chapter/41> (дата обращения: 21.08.2025).

Fenwick, 2011 – Fenwick R.S.H. *A Grammar of Ubykh*. München, 2011.

Hanks, 2011 – Hanks W.F. Deixis and Indexicality. *Handbook of Pragmatics*. W. Bublitz, N.R. Norrick (eds.). Berlin, 2011. Pp. 315–346.

Hellwig, 2018 – Hellwig B. “See This Sitting One”: Demonstratives and deictic classifiers in Goemai. *Demonstratives in Cross-Linguistic Perspective*. S. Levinson, S. Cutfield, M. Dunn et al. (eds.). Cambridge University Press, 2018. Pp. 134–149.

Hewitt, 2005 – Hewitt G. North West Caucasian. *Lingua*. 2005. Vol. 115. No. 1–2. Pp. 91–145.

Jungbluth, 2003 – Jungbluth K. Deictics in the conversational dyad: Findings in Spanish and some cross-linguistic outlines. *Deictic Conceptualization of Space, Time and Person*. Pragmatics and Beyond New Series. Vol. 122. F. Lenz (ed.). Amsterdam, 2003.

Konuk, 2022 – Konuk N.M. Description grammaticale de l’abzakh – tcherkesse occidental. PhD dis. Université de la Sorbonne Nouvelle – Paris III. 2022.

Lander, Haegeman, 2018 – Lander E., Haegeman L. The Nanosyntax of Spatial Deixis. *Studia Linguistica*. 2018. Vol. 72. No. 2. Pp. 362–427.

Levinson, 2018 – Levinson S. C. Introduction: Demonstratives: Patterns in diversity. *Demonstratives in Cross-Linguistic Perspective*. S. Levinson, S. Cutfield, M. Dunn et al. (eds.). Cambridge University Press, 2018. Pp. 1–42.

Levinson et al., 2018 – Demonstratives in Cross-Linguistic Perspective. S. Levinson, S. Cutfield, M. Dunn et al. (eds.). Cambridge University Press, 2018.

Meira, Guirardello-Damian, 2018 – Meira S., Guirardello-Damian R. Brazilian Portuguese: Non-contrastive exophoric use of demonstratives in the Spoken Language. *Demonstratives in Cross-Linguistic Perspective*. S. Levinson, S. Cutfield, M. Dunn et al. (eds.). Cambridge University Press, 2018. Pp. 116–133.

Mukhin, Nasledskova, 2025 – Mukhin T., Nasledskova P. Demonstrative pronouns in Kina Rutul. *Studies in Language*. 2025. Vol. 49. No. 2. Pp. 456–477.

Paris, 1989 – Paris C. West Circassian (Adyghe: Abzakh Dialect). *The Indigenous Languages of the Caucasus*. Vol. 2: North West Caucasus. B.G. Hewitt (ed.). Delmar, 1989. Pp. 157–260.

Pinkster, 2015 – Pinkster H. *The Oxford Latin Syntax*. Vol. 1. Oxford University Press, 2015.

Schulze, 2008 – Schulze W. Deictic strategies in Udi // Удинский сборник: грамматика, лексика, история языка / под ред. М.Е. Алексеева, Е.А. Майсака, Д.С. Ганенкова, Ю.А. Ландера. М., 2008. С. 241–308. [Schulze W. Deictic strategies in Udi. *Udinskiy sbornik: grammatika, leksika, istoriya yazyka*. М.Е. Alekseev, Т.А. Maisak, Д.С. Ganenkov, Ю.А. Lander (eds.). Moscow, 2008. Pp. 241–308.]

Wilkins, 2018 – Wilkins D.P. The demonstrative questionnaire: “THIS” and “THAT” in comparative perspective. *Demonstratives in Cross-Linguistic Perspective*. S. Levinson, S. Cutfield, M. Dunn et al. (eds.). Cambridge, 2018. Pp. 43–71.

Yanagisawa, 2013 – Yanagisawa T. *A Grammar of Abkhaz*. Tokyo, 2013.

Статья поступила в редакцию 06.09.2025

The article was received 06.09.2025

Сведения об авторах / About the authors

Миронова Елена Александровна – стажер-исследователь Международной лаборатории языковой конвергенции, Национальный исследовательский университет «Высшая школа экономики», г. Москва

Elena A. Mironova – research intern at the Linguistic Convergence Laboratory, HSE University, Moscow, Russian Federation

ORCID: <https://orcid.org/0009-0000-1174-5938>

E-mail: elalmironova@edu.hse.ru

Унарокова Шамсет Шабановна – ассистент кафедры адыгейской филологии, научный сотрудник лаборатории экспериментальной лингвистики, Адыгейский государственный университет, г. Майкоп

Shamset Sh. Unarokova – assistant at the Adyghe Philology Department, research fellow at the Laboratory of Experimental Linguistics, Adyghe State University, Maykop, Russian Federation

ORCID: <https://orcid.org/0009-0001-4343-7267>

E-mail: Unarokova.shamset@yandex.ru

Ландер Юрий Александрович – кандидат филологических наук; ведущий научный сотрудник Международной лаборатории языковой конвергенции, доцент Школы лингвистики, Национальный исследовательский университет «Высшая школа экономики», г. Москва

Yury A. Lander – PhD in Linguistics; Leading research fellow at the Linguistic Convergence Laboratory, associate professor at the School of Linguistics, HSE University, Moscow, Russian Federation

ORCID: <https://orcid.org/0000-0003-1798-7174>

E-mail: yulander@yandex.ru

Издание
подготовили
к печати:
редактор
А. А. Козаренко,
корректор
А. А. Алексеева,
обложка, макет,
компьютерная
верстка
Н. А. Попова

Rhema. Рема

2025.4

Сайт журнала: rhema-journal.com
E-mail: rhema.pema@gmail.com

Авторы статей несут полную ответственность за точность приводимой информации, цитат, ссылок и списка литературы.

Перепечатка материалов, опубликованных в журнале, невозможна без письменного разрешения редакции.

Подписано в печать 29.12.2025.
Формат 60×90 1/16. Гарнитура «Liberation Serif».
Объем 8,75 п. л. Тираж 1000 экз.