

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Отделение историко-филологических наук

ВОПРОСЫ ЯЗЫКОЗНАНИЯ

Журнал основан в январе 1952 года

Выходит 6 раз в год

5

СЕНТЯБРЬ — ОКТЯБРЬ

Москва

2022

Главный редактор:

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

Зам. главного редактора:

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

В. И. Подлесская д. ф. н., проф., Российский государственный гуманитарный университет

Редколлегия:

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

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

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

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

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

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

В. А. Дыбо д. ф. н., академик РАН, Институт славяноведения РАН

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

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

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

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

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

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

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

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

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

Л. А. Янда Ph.D., проф., Университет Тромсё — Норвежский арктический университет, Норвегия

Зав. редакцией: Н. В. Ганнус

Зав. отделами: Л. С. Козлов, А. С. Кулева

Статьи отбираются редколлегией журнала на основе анонимного независимого рецензирования.

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

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

Телефон: +7 495 637-25-16

E-mail: voprosy@mail.ru

Сайт: <https://vja.ruslang.ru>

© Российская академия наук, 2022

© Составление. Редколлегия журнала «Вопросы языкознания», 2022

ISSN 0373-658X

RUSSIAN ACADEMY OF SCIENCES

Department of History and Philology

VOPROSY
JAZYKOZNANIJA
(TOPICS IN THE STUDY OF LANGUAGE)

Founded in January 1952

6 issues per year

5

SEPTEMBER — OCTOBER

Moscow

2022

Editor-in-chief:

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

Assistant editors:

Vera I. PODLESSKAYA Russian State University for the Humanities, Moscow, Russia
Nikolai B. VAKHTIN European University at St. Petersburg; Institute for Linguistic Studies (RAS), St. Petersburg, Russia

Editorial board:

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

Managing editor: Natalia V. GANNUS

Editorial staff: Lev S. KOZLOV, Anna S. KULEVA

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

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

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

Telephone: +7 495 637-25-16

E-mail: voprosy@mail.ru

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

ISSN 0373-658X

Содержание

А. Ю. УРМАНЧИЕВА. Исторические ареальные связи мансийского языка	7
О. М. РЫНДИНА, Д. О. ЖОРНИК, Е. В. БАРСУКОВ, С. В. ПОКРОВСКАЯ. Архив В. Н. Чернецова: мансийская этнография, фольклор и лингвистическая перспектива	35

Падежи в языках Северной Евразии

S. A. OSKOLSKAYA. Case in the languages of Northern Eurasia: Preface	51
J. JANHUNEN. Issues of comparative Uralic and Altaic studies (10): On the local case markers in Mongolic	55
T. KOIVUNEN, R. ERKKILÄ. The interaction between spatial cases and relational nouns in the Uralic languages of the Volga–Kama area	64
R. ERKKILÄ. How to distinguish between semantically close cases: A case study on Mor-dvin illative and lative	86
P. S. PLESHAK. The status of Dative case in the Moksha case paradigm	108
A. BUGAEVA. On the exclusively borderline case-marking in Ainu	131

Рецензии

И. А. ХОМЧЕНКОВА. [Рец. на:] A. Guillaume, H. Koch (eds.). <i>Associated motion</i> . Berlin; Boston: De Gruyter Mouton, 2021	151
---	-----

Contents

Anna Yu. URMANCHIEVA. Linguistic areas in the history of the Mansi language	7
Olga RYNDINA, Daria ZHORNIK, Evgeny BARSUKOV, Sofie POKROVSKAYA. Valery Chernetsov's archive: Mansi ethnography, folklore and linguistic perspective	35

Case in the languages of Northern Eurasia

Sofia A. OSKOLSKAYA Case in the languages of Northern Eurasia: Preface	51
Juha JANHUNEN. Issues of comparative Uralic and Altaic studies (10): On the local case markers in Mongolic	55
Tomi KOIVUNEN, Riku ERKKILÄ. The interaction between spatial cases and relational nouns in the Uralic languages of the Volga–Kama area	64
Riku ERKKILÄ. How to distinguish between semantically close cases: A case study on Mordvin illative and lative	86
Polina S. PLESHAK. The status of Dative case in the Moksha case paradigm	108
Anna BUGAEVA. On the exclusively borderline case-marking in Ainu	131

Reviews

Irina A. KHOMCHENKOVA. [Review of:] A. Guillaume, H. Koch (eds.). <i>Associated motion</i> . Berlin; Boston: De Gruyter Mouton, 2021	151
--	-----

Исторические ареальные связи мансийского языка

© 2022

Анна Юрьевна Урманчиева

Томский государственный университет, Томск, Россия; urmanna@yandex.ru

Аннотация: Статья посвящена реконструкции истории контактов мансийского языка с представителями самодийской языковой группы. На современной лингвистической карте Западной Сибири эти языки не находятся в непосредственном контакте, однако лингвистические данные позволяют реконструировать несколько контактных ситуаций различной степени временной глубины. Я предполагаю, что в районе верхней Оби (приблизительно на том ее отрезке, который протекает по территории современной Томской области) существовал исторический лингвистический ареал, в который входили мансийский, селькупский и камасинский языки. На восточной периферии этой зоны происходили не слишком интенсивные контакты селькупов с кетами. Хантыйский язык был распространен севернее мансийского и занимал, возможно, северную периферию данного ареала. При этом ближе всего к мансийскому располагался предок современного восточного хантыйского. Манси соседствовали с южными самодийцами, вероятно, в районе Парабели / Чижипки, где Э. Г. Беккер выделяет зону распространения камасинской топонимики. В период этих контактов, на мой взгляд, камасинский и селькупский уже представляли собой два отдельных языка, но при этом диалектная дифференциация селькупского еще не началась: в статье предложен ряд мансийско-селькупских ареальных изоглосс, в равной мере представленных во всех селькупских диалектах. Впоследствии взаимное расположение хантыйского и мансийского языков на лингвистической карте Западной Сибири изменилось: в верхнеобской ареал проникают восточные ханты. Затем северные и южные ханты расселяются широкой полосой, протянувшейся с севера на юг вдоль течения средней Оби и Иртыша. Это прерывает связь манси с восточными хантами и окончательно отрезает манси от верхнеобского ареала. В этот период в верхнеобском ареале начинает формироваться современный «остяцкий» ареал, включающий селькупский, кетский и хантыйский языки. Формирование этого ареала должно быть отнесено к периоду после диалектной дифференциации входящих в него уральских языков: в этот ареал входят только крайневосточные диалекты хантыйского и только северный и тымский диалекты селькупского.

Ключевые слова: ареальная лингвистика, мансийский язык, реконструкция, самодийские языки, уральские языки, хантыйский язык, языковые контакты.

Благодарности: Исследование выполнено за счет гранта Российского научного фонда (проект № 19-18-00329).

Для цитирования: Урманчиева А. Ю. Исторические ареальные связи мансийского языка. *Вопросы языкознания*, 2022, 5: 7–34.

DOI: 10.31857/0373-658X.2022.5.7-34

Linguistic areas in the history of the Mansi language

Anna Yu. Urmanchieva

Tomsk State University, Tomsk, Russia; urmanna@yandex.ru

Abstract: The article is devoted to the reconstruction of historical contacts of the Mansi language with Samoyedic languages. On the modern linguistic map of Western Siberia these languages are not in direct contact, however, linguistic data make it possible to reconstruct several contact situations of varying degrees of temporal depth. I suppose that in the region of the Upper Ob there was a historical linguistic area, which included the Mansi, Selkup, and Kamas languages. In the eastern periphery of this

zone, some limited contacts between the Selkups and the Kets took place. The Khanty language was widespread north of the Mansi language, possibly, in the northern periphery of this area. The ancestor of the modern Eastern Khanty language was located closest to the Mansi language. The Mansi contacted with the speakers of the Southern Samoyedic languages most probably across the rivers of Chizhapka and Parabel. Within the same area E. G. Bekker identified the zone of distribution of the Kamas toponymy. During the period of these contacts, in my opinion, Kamas and Selkup were already two separate languages, but the dialectal differentiation of Selkup had not yet begun: the article proposes a number of Mansi-Selkup areal isoglosses, equally represented in all Selkup dialects. Subsequently, the relative position of the Khanty and Mansi languages on the linguistic map of Western Siberia changed. The Eastern Khanty penetrated the Upper Ob area. Then the Northern and Southern Khanty settled in a wide zone stretching from north to south along the course of the Middle Ob and Irtysh. This interrupts the connection between the Mansi and the Eastern Khanty and cuts off the Mansi from the Upper Ob area. During this period, the modern “Ostyak” area including the Selkup, Ket, and Khanty languages began to form in the Upper Ob region. The formation of this “Ostyak” area should be attributed to the period after the dialectal differentiation of the both mentioned Uralic languages: this area includes only the easternmost dialects of Khanty and only the Northern and Tym dialects of Selkup.

Keywords: areal linguistics, Khanty, language contacts, Mansi, reconstruction, Samoyedic, Uralic.

Acknowledgements: The study was supported by Russian Science Foundation grant No. 19-18-00329.

For citation: Urmanchieva A. Yu. Linguistic areas in the history of the Mansi language. *Voprosy Jazykoznanija*, 2022, 5: 7–34.

DOI: 10.31857/0373-658X.2022.5.7-34

Введение

Статья посвящена реконструкции истории контактов мансийского языка с представителями самодийской языковой группы. На современной лингвистической карте Западной Сибири эти языки не находятся в непосредственном контакте, однако лингвистические данные позволяют реконструировать несколько контактных ситуаций различной степени временной глубины. Это позволяет расширить наши представления об этнической истории манси, реконструировать последовательные изменения лингвистического ландшафта Западной Сибири (собственно исторические, и притом крайне неполные, опосредованные сведения о котором нам доступны только начиная с XVII в.) и роль мансийского языка в формировании этого ландшафта. Очевидно, что изложенные в статье выводы имеют непосредственное отношение не только к лингвистической, но и к этнической истории манси.

В статье будут последовательно рассмотрены две контактные ситуации — от наиболее древней к наиболее современной. В разделе 1 будут рассмотрены контакты между угорским / обско-угорским праязыками и прасамодийским, в разделе 2 — контакты между мансийским и южно-самодийскими. В третьем разделе предлагается реконструкция истории верхнеобского лингвистического ареала.

Вопрос об отношениях между (обско-)угорскими и самодийскими языками в исторической перспективе рассмотрен прежде всего в книге Е. А. Хелимского [1982] «Древнейшие венгеро-самодийские языковые параллели». В этой книге дается следующая относительная периодизация контактных ситуаций:

«а) Древнейшие венгеро-самодийские языковые параллели имеют чисто генетический характер и восходят к общей для всех уральских языков праязыковой системе. {...}

б) В период, когда в ходе постепенного расселения уральцев создавалась ситуация диалектной непрерывности (характеризующаяся сходимостью соседних и значительным различием удаленных друг от друга говоров) и формировались диалектные ареалы, впоследствии легшие в основу отдельных групп уральских языков — в этот

период ⟨...⟩ возникли предпосылки для появления ареально-генетических угорско-самодийских параллелей, обусловленных соседством доугорских и досамодийских говоров (диалектов). ⟨...⟩

в) ... после распада уральского праязыка на территории Западной Сибири существовали контакты между (ранне)самодийским праязыком, с одной стороны, и доугорскими диалектами финно-угорского праязыка, а затем угорским праязыком, — с другой. Контакты такого рода могли послужить предпосылкой для возникновения общих явлений, легших в основу древнейших ареальных угорско-самодийских параллелей. Возможно, эти контакты были маргинальными и в них не принимали непосредственного участия (во всяком случае, в позднеугорскую эпоху) южные или юго-западные диалекты угорской языковой общности, легшие в основу венгерского языка. ⟨...⟩

г) Контакты, охарактеризованные в п. в), могли, в принципе, продолжиться и в период после распада угорской языковой общности и до распада самодийской, т. е. во второй половине I тыс. до н. э. В таком случае они могли привести к появлению ареальных параллелей между обско-угорскими (но не венгерским) и самодийскими языками. Фактически, однако, параллели такого рода — если они и существуют — нелегко отграничить от результатов более поздних и весьма интенсивных взаимодействий. ⟨...⟩

д) в период после образования отдельных самодийских языков, распространенных, в частности, в бассейне Оби (селькупский) и западносибирской тундре (ненецкий), имели место и продолжают до настоящего времени контакты самодийцев с ханты и манси, хорошо документированные исторически, этнографически и лингвистически. Их результатом являются, в первую очередь, многочисленные хантыйско-селькупские, хантыйско-ненецкие и, в несколько меньшей степени, мансийско-ненецкие ареальные параллели (лексические заимствования, конвергентные явления в развитии языковых систем)» [Хелимский 1982: 64–66].

Настоящее исследование предполагает прежде всего возможность уточнить некоторые процессы, характеризующие периоды б)–д): исследование сосредоточено на тех параллелях, которые имеют ареальную природу.

1. Контакты между угорским / обско-угорским праязыками и прасамодийским

Е. А. Хелимский рассматривает в своей работе шесть языковых явлений, которые могут считаться угорско-самодийскими языковыми параллелями:

- к общеуральскому фонду достаточно надежно возводится система трех типов глагольного спряжения [Хелимский 1982: 70–88];
- к общеуральскому фонду, возможно, возводится суффикс аккузатива личных местоимений [Там же: 88–97]. Но, с одной стороны, как указывает Хелимский, рассматриваемый им показатель может иметь когнаты и в других уральских языках, что делает эту параллель не эксклюзивной угорско-самодийской. С другой же стороны, материальное сопоставление угорской и самодийской морфем кажется мне все же не вполне надежным, так как на угорском материале этот показатель восстанавливается в виде **ki* [Там же: 96], на самодийском же — в виде **kit* [Janhunen 1977: 70], см. также этимологию самодийского слова, предложенную в [Aikio 2006: 17–19], которая также противоречит отождествлению самодийского и угорского показателей;

- *n*-овые показатели местных падежей [Хелимский 1982: 97–102] рассматриваются как ареально-генетическая угорско-самодийская параллель (из общеуральского материала в восточном ареале праязыкового диалектного континуума формируются вторичные показатели местных падежей). К вопросу об *n*-овых падежах я еще вернусь ниже в данной работе;
- как ареальная параллель должно рассматриваться активное использование для образования каузативных глаголов составной суффиксации [Там же: 103–114]. В обско-угорских и самодийских языках в этой функции используются комбинации нескольких суффиксов-транзитивизаторов или нескольких суффиксов с каузативным значением. Но хотя наборы морфологически элементарных суффиксов, образующих составные морфемы, по крайней мере частично пересекаются в обско-угорских и самодийских языках (и часть этих суффиксов может быть возведена на прауральский уровень), образуемые комбинации не тождественны;
- наличие у числительного ‘два’ особой неатрибутивной формы, образуемой с использованием показателя двойственного числа [Хелимский 1982: 114–119], с равной долей вероятности может интерпретироваться и как представляющее собой реликт общеуральского состояния (поскольку обе морфемы — и основа числительного ‘два’, и показатель двойственного числа — имеют общеуральское происхождение), и как возникшее «конвергентным путем или вследствие заимствования» [Там же: 118];
- одинаковое развитие прауральских сибилантов (полностью совпадающее только для обско-угорских и самодийских — в венгерском этот процесс шел несколько иначе) [Хелимский 1982: 119–124]. Хелимский интерпретирует этот процесс как исключительно ареальную (не ареально-генетическую!) параллель. Иначе интерпретируется это явление в [Häkkinen 2007; 2012]. Я. Хяккинен — с опорой прежде всего на составленный им список изоглосс в области исторической фонетики — предлагает нетрадиционное дерево уральской языковой семьи: в результате первичного разделения образовались восточная и западная ветви уральских языков, причем восточную ветвь образуют, с его точки зрения, именно угорские и самодийские языки. В этом списке изоглосс важное место занимают фонетические переходы, описывающие параллельное развитие прауральских сибилантов (1. **s* > **š*; 2. **š* > **ʎ*; 3. **š* > **s*). Предложенный Я. Хяккиневым список изоглосс разбирается в [Живлов 2018]; показано, что именно эти три фонетических изменения являются единственными бесспорными. Тем не менее в докладе М. А. Живлова также показано, что эти изменения произошли в обско-угорских языках независимо — после распада угорской, и даже обско-угорской, языковой общности. Заимствования из индоиранских языков (а также более локальные аланизмы) показывают, что заимствования проникали в мансийский язык уже после того, как в нем произошли соответствующие фонетические переходы, но в хантыйский — до этих фонетических переходов. Соответственно, в обско-угорских языках эти фонетические изменения прошли независимо, причем в мансийском раньше, чем в хантыйском. Из этого следует, что параллельные изменения в венгерском и самодийских также произошли независимо. Это подтверждает вывод об ареальной, а не ареально-генетической природе этой параллели между (обско-)угорскими и самодийскими языками.

Ниже с опорой на [Rédei 1988] (при возможности — с уточнениями реконструкции для самодийских и обско-угорских) я привожу неполный список тех изоглосс, которые могут быть отнесены к угорско-самодийским лексическим параллелям. Таких параллелей всего четыре, и ни для одной из них невозможно надежно показать, что они являются общими угорско-самодийскими инновациями, а не независимым сохранением общеуральских основ.

1.1. Угорско-самодийские параллели

1.1.1. U **kuḍa* ‘утро’ [Rédei 1988: 193]: хант. Кам. *χḍtnə* ‘утром’, V *kōltəχ*, DN *χātenχāt*, O *χālewət* // манс. TJ *kol*, KU So. *χol*, P *kol*, *k_ool* // венг. *holnap* // сам. сельк. *karə* [Alatalo 2004: ст. 2217], камас. *karəld'en*, *korod'on* [Ibid.]. Впрочем, значение ‘утро’ выражается разными, неродственными словами в прибалтийско-финском, саамском, мордовском, марийском и пермском, что может говорить скорее об исконности угорско-самодийского **kuḍa* ‘утро’.

1.1.2. U **pʁnz* ‘пердеть’ [Rédei 1988: 413]: хант. DN O *pān* ‘пук’, V *poṅəu-*, DN *pānət-*, O *ponṅə-* ‘пердеть’ // манс. T *pōṅkh-*, LU *rannat-*, N *ponim-* ‘пердеть’, LO *pon* ‘пук’ // венг. *fiŋg* ‘пук’ // сам. камас. *piṅḍōr-* ‘пердеть’.

1.1.3. U **wʁjz* ‘видеть, замечать’ [Rédei 1988: 588]: хант. V *wu-* ‘видеть’, DN *wo-* ‘видеть, знать, быть знакомым с кем-либо, чем-либо, мочь’, Kaz. *wɔ-* ‘знать, быть знакомым с кем-либо, чем-либо’, Tṛj. *oʃəytə-* ‘найти’, Kaz. *wɔjət-* ‘найти’, O *oś-* ‘знать, быть знакомым с кем-либо, чем-либо, найти’ // манс. LO *wē-*, *wāj-*, *waj*, KU *wā-*, *waj-*, *woj-*, *ū-*, P *wa-*, *wē-*, *waj-*, So. *wā-* ‘видеть’ (словарь [Kálmán, Munkácsi 1985] дает и для мансийского значения ‘знать, быть знакомым с кем-либо, чем-либо’) // венг. *óv-* (старовенг. *ó*) ‘стеречь’ // сам.: т. нен. *je-*, л. нен. *we-*, т. эн. *bie-* ‘караулить стадо’, нган. *baa-* ‘бодрствовать’.

1.1.4. U *sud* ‘з’ ‘палец’ [Rédei 1988: 449]. Это слово представлено в трех угорских языках, из самодийских же оно сохранилось только в маторском: хант. Vj. *jɔj*, Tṛj. *loj*, DT *tüj*, Kaz. *lūj*, O *luj* // манс. TJ *tol'ā*, KU *tol'ə*, P *tul'ə*, So. *tul'a* // венг. *ujj* // сам.: **tāja* > матор. *taja* [Helimski 1997a: ст. 941].

1.2. Обско-угорско-самодийские параллели

1.2.1. U **amz-* ‘сидеть’ [Rédei 1988: 8]: хант. V *aməs-*, DN *oməs-*, O *aməs-* // манс. TJ *ōn-*, KU *onl-*, P *wunl-*, So. *ūnl-* // сам. **āmtā* (с.-сам.) ~ **āmtā* (ю.-сам.) ‘сесть’ [Janhunen 1977: 17–18].

1.2.2. U **antz* [Rédei 1988: 12] (**anti* [Reshetnikov, Zhivlov 2011: 99]) ‘рог’: хант. V *āṅət*, DN *oṅət*, O *anət* // манс. TJ *ēnt*, KU *šnt*, KO *ēnt*, P *ōnt*, So *ānt* // сам. **āntā* [Janhunen 1977: 20]. Несмотря на нерегулярный характер рефлексии интервокального кластера, это сопоставление принимается как несомненное в работах по уральской этимологии.

1.2.3. U **kulz(-kz)* ‘ворон’ [Rédei 1988: 200]: хант. V *kələk*, DN *χuləχ*, O *χoləχ* // манс. TJ *kolāk*, KU *χolχ*, P *kulk*, So. *χuləχ* // сам. **kulājā* [Janhunen 1977: 77].

1.2.4. U **kučə* ‘муравей’ [Rédei 1988: 192–193]: хант. V *kočṅj*, DN *χāšṅaj*, O *χāšṅa* // манс. TJ *kaškēj*: *künšk.*, KU *χošχaj*: *χoňšχ.*, P *koššaj*, LO *χōssuj*, So *χōsswoj* // сам. **kecz*, произв. **keczmkā* (ю.-сам.) ~ *kuczmkā* (нен.) [Janhunen 1977: 65].

1.2.5. U **nusz-* (*nišz-*) ‘скоблить шкуру, очищая от мездры’ [Rédei 1988: 309]. К. Редей под вопросом включает в эту статью марийский глагол *niže-* и приводит хантыйский и самодийский материал. // сам. **nət-* [Janhunen 1977: 97]: нган. *nət-*, *nəzү'əту*, т. нен. *нāць*, л. нен. *naš*, т. эн., л. эн. *nos-*, сельк. *nat-* [Alatalo 2004: ст. 1351], камас. *nāʔl'ām*.

1.2.6. Хорошо известно сопоставление самодийского *pät-* (ю.-сам.) ~ *pätz-* (с.-сам.) ‘опустить в котел’ [Janhunen 1977: 118] с прауральской именной основой, обозначающей сам котел: U **pata* [Rédei 1988: 358]. Это сопоставление упоминается в [Reshetnikov, Zhivlov 2011: 99] со ссылкой на [Напольских 1997: 124] и [Aikio 2002: 50]. Кажется, можно предложить параллельное обско-угорско-самодийское сопоставление: манс. N K *pēti*, [= *pēti*], LM *pēti*, T *pītānt* ‘опускать в котел’ [Kálmán, Munkácsi 1985] // сам. **pet(ā)* ‘(глиняный) сосуд’ [Janhunen 1977: 122]: т. нен. *ped*, сельк. *pita* [Alatalo 2004: ст. 465]. Возможность предлагаемого семантического сопоставления названия сосуда для приготовления пищи с глаголом, означающим ‘опустить в сосуд для приготовления пищи’, подкрепляется

аналогичным соотношением слов в первой паре. Как указал мне анонимный рецензент, фонетические соответствия здесь не позволяют считать эти слова исконно родственными, что является аргументом в пользу контактной версии возникновения этой мансийско-самодийской изоглоссы.

1.2.7. U *čančz ~ čačz- ‘идти’ [Rédei 1988: 53]. Из этой статьи следует выделить основу čača- [Reshetnikov, Zhivlov 2001: 99], которая представлена только в обско-угорских и селькупском: хант. Tj. *soč-*, DN *čuš-*, O *sos-* // манс. Tj *šoš-*, KU *šoš-*, P *šuš-*, So. *sūs-* ‘переходить вброд’ // сельк. *čāča-* ‘идти’ [Alatalo 2004: ст. 1047].

1.2.8. U *kurz ‘тело’ [Rédei 1988: 216]: саам.: N *gorod* ~ *gorod* ‘туша (чаще мертвого зверя, но может быть и туловище живого)’ и т. д. // хант. V *kōr* ‘изображение, образ (зверя, птицы, рыбы и т. д.)’, DN *χōr* ‘птица-манок’, Ts. *χor*, Kaz., O *χōr* ‘изображение (духа-покровителя, человека, животного)’ // манс. (Ahlqv.) *qōr*, *xōr* ‘тень (повторяющая очертания), образ, лицо’ [Kálmán, Munkácsi 1985] N *χuri* [=χuri], LM *khur*, LU *kwār*, P *khwār*, K *khwār* ~ *khwōr*, T *khōr* ‘образ, изображение, образец, тень живого существа’ // сам.: т. нен., л. нен. **kārə-q*=Px, **kārye-q*=Px ‘сам’, л. эн, т. эн. *kere-q*=Px (< *kore-q*=Px) ‘сам’, нган. *kəru-tə*=Px ‘сам по себе, просто так’ (*kərutə-nə* ‘я просто так’); основа в самодийских языках имеет вид **kərV-*: нен. -q, эн. -ʔ и нган. -tə являются показателями дательного падежа¹. Даже если, несмотря на различия в семантике, саамское слово относится к этому же этимологическому гнезду, мы видим четко выделяющийся ареал (обско-угорские и самодийские языки), в котором рассматриваемая основа означает приблизительно ‘подобие, двойник, тень живого существа’ и далее грамматикализуется в северносамодийских в слово со значением ‘сам’. О таком пути грамматикализации в уральских языках см. словарную статью U **íce* ~ *ise* ‘тень, душа-тень’ [Rédei 1988: 79–80], на другом этимологическом материале с прибавлением самодийских языков — [Хелимский 1982: 95–96].

1.2.9. U *pīlz ‘овод’ [Rédei 1988: 416]: хант. VK *peləm*, J *pāləm*, DN *petəm* // манс. Tj *päləm*, KU *pāləm*, P *päləm*, So. *paləm* // сам.: т. нен. *pil’ū*, л. нен. *pirñū*, л. эн. *pil’u* (? < т. нен.), нган. *hil’i*.

1.2.10. U *sätz-, šätz- ‘вычерпывать (из котла), вынимать’ [Rédei 1988: 437]: хант. Vj. *jāt-*, Tj *lāt-*, Kam. *tāt-*, Ni. *tat-* ‘вычерпывать в тарелку содержимое из котла (мясо, рыбу, но не суп)’ // т. нен. *t’edā-*, л. нен. *t’šēta-* ‘вычерпывать из котла (мясо, рыбу)’, л. эн., т. эн. *tedi-* ‘вынуть мясо из котла’, нган. *četə-* ‘вычерпать из котла’, сельк. *tit-* ‘положить еду в тарелку’.

1.2.11. U *kälz ‘запекшаяся кровь’ [Rédei 1988: 134]. Удалив из сопоставления ненадежные саамские и хантыйские параллели, мы получаем следующую группу соответствий: манс. KU *kēləp*, VN So. *kēlp* // сам. т. нен., л. нен. *syelw^o*, л. эн., т. эн. *semo*, нган. *сельбу*.

1.2.12. U *perz ‘кора; кожа’ [Rédei 1988: 374]: хант. Vj *pēr* ‘красноватая кожица с внутренней стороны бересты’² // сам.: т. нен. *pīrʔ* ‘красноватая кожица с внутренней стороны бересты, слизистая оболочка рубца, корка на хлебе’ // койб. *pere* ‘кора’.

1.2.13. U *muŋkz ‘стрела с тупым наконечником’ [Rédei 1988: 286]: хант. D *muŋk*, Ni. *muŋk*, O *mōŋk* ‘шарообразный наконечник стрелы’ // манс. (Kann., mitg. Liim.: MSFOu 116: 297) So *māŋk*, ‘шарообразный наконечник стрелы, используемой при охоте на белку’; [Kálmán, Munkácsi 1985] *nāl-***māŋχw* ‘шарообразный наконечник стрелы, используемой при охоте на белку’, N *māŋχw* [*māχw*], LM *mōŋkhw*, LU *mōŋk*, P *mōŋkhw*, K *mōŋkhw* ~ *mōŋkh*,

¹ Т. нен., л. нен. **kəra* ‘манера’, т. эн., л. эн. *kora* ‘манера, мода’, нган. *kəru* ‘манера, привычка’ скорее связаны с общесамодийским **kəra* ‘кривой’ [Janhunen 1977: 55]. При этом не исключено вторичное сближение этих двух основ в результате народной этимологизации, по крайней мере в нганасанском, что объяснило бы гласный второго слога в нган. *kəru* ‘сам по себе’.

² Я благодарю анонимного рецензента за указание на то, что включаемое К. Редеей в эту же словарную статью венг. *bőr* ‘кожа, шкура, пленка на молоке или других жидкостях, скорлупа, (диал.) лыко’ не может быть отнесено к этой этимологии ни из-за анлаутного согласного, ни из-за вокализма корня.

Т *təŋ* (PxSg3 *māŋyā*) ‘обух топора, спинка (противоположная лезвию сторона) ножа’, (N в медвежьих песнях) ‘топор’ // сам.: матор. *tuŋgu*, т. нен. *tūŋk*, т. эн. *tugeʔo*, нган. *munḡkə* ‘тамар, стрела с тупым наконечником’. Это слово можно рассматривать как культурный термин, получивший распространение в Западной Сибири в тот период, когда северно-самодийские уже обособились от южно-самодийских (но маторский еще не оторвался от них, перейдя в ареальную общность с южно-самодийскими, о принадлежности маторского к северносамодийским см. [Helimski 1997b]). При этом распространение этого термина в группе северно-самодийских (как, вероятно, и в целом в Западной Сибири) следует рассматривать именно как ареальное явление: нерегулярный характер гласного в нганасанском (*u* вместо ожидаемого *ü*) говорит в пользу того, что в группе северносамодийских (включая маторский) этот термин представляет не рефлексы единой северносамодийской праформы, а распространявшийся путем межъязыкового заимствования культурный термин.

1.3. Бродячие слова

1.3.1. U **juwz* ‘сосна; ель’ [Rédei 1988: 107]: хант. V *juŋ*, DN O *juχ* ‘дерево’ // манс. TJ *jīw*, KU So. *jīw*, P *jūw* ‘дерево’ // сам. **jeʹ* (? **jeʹw*) ‘сосна’ [Janhunen 1977: 42]. Как указал мне анонимный рецензент, фонетические соответствия между мансийским и хантыйским нерегулярны. Возможно, это субстратное слово, которое, в свою очередь, послужило и источником прасамодийского названия сосны.

1.3.2. U **suŋ(ɜ)lɜ*, *šuŋ(ɜ)lɜ* [Rédei 1988: 451] ‘вид лилии (либо луковица лилии); *Lilium martagon*’: хант. V *loŋəl*, Vj. VK *joŋəl* ‘репа’ (V), ‘лук’ (Vj), ‘любой корнеплод (картофель, репа, лук и т. д.), дикий лук’ (VK) // манс. TJ TČ *taŋkəl* ‘саранка; *Lilium martagon* (желтоватые луковицы съедобны)’ // сам. сельк. *tōkəl* ‘*Lilium martagon*, сарана’ [Alatalo 2004: ст. 1179], камас. *dugul*, матор. *tugul* [Helimski 1997a: ст. 1079]. Я. Алатало [Ibid.: ст. 1179] обращает внимание также на сходство с эвенкийским *тукала* и эвенским *тукэр* [ССТМЯ: 207]. В данном случае мы имеем дело с сибирским культурным термином. В хантыйском это слово представлено только в крайневосточных диалектах, при этом его значение не сохраняется. В селькупском это слово представлено в тымском, северном и кетском диалектах; такой ареал распространения исключает непосредственное заимствование из хантыйского (такие заимствования не достигают кетского диалекта). Исключает гипотезу о заимствовании и фонетика: анлаутный хантыйский *l-* заимствуется в селькупский как *l-*. Все три самодийских языка свидетельствуют об интервокальном **k*, который не может отражать прауральский интервокальный *-ŋ-*. Таким образом, самодийские и обско-угорские слова не могут быть возведены к единой уральской праформе; перед нами случай ареальной инновации — культурного термина, возникшего в уральских языках Сибири уже после разделения ветвей.

1.3.3. U **palz* ‘копье’ [Rédei 1988: 352]: хант. V Vj *pał*, O *pali* ‘копье для охоты на медведя’ // манс. LO So *pāl* ‘хорей’ // нен. *pāli* ‘меч, сабля’ // л. эн. *falli* ‘тунгусский меч’. Редеи предполагает, что это может быть очередным бродячим культурным термином. Этот вывод можно сделать, если посмотреть на широкую географию распространения в русских диалектах Сибири (вплоть до Амура и Колымы) фонетически сходного названия оружия близкой конструкции (рогатины, ножа на длинном древке) — см. статьи *пальма* и *палемка* в [Аникин 2000: 437–438]. Но примечательно, что в Западной Сибири именно обско-угорские и самодийские языки разделяют этот культурный термин.

1.3.4. Предложенную К. Редеи прауральскую этимологию U **šarma* ‘дымовое отверстие чума’ [Rédei 1998: 463] можно с осторожностью рассматривать разве что как очередное бродячее слово: манс. K *šōrəm*, P *surēm*; в словаре [Kálmán, Munkácsi 1985] зафиксирован более широкий ареал распространения этого слова: N LM *surēm* [*surəm*], P *surmē* ~ *surēm*,

К *sgrēm* // сам. т. нен. *sārwa*, л. нен. *hārwa* // т. эн., л. эн. *samaʔa*. Северно-самодийское (собственно, ненецко-энецкое) и мансийское слова не могут восходить к одному и тому же прауральскому источнику, так как продолжением прауральского кластера *-rm-* является прасамодийский *-rm-* и ненецкий *-rm-*, ср. U **śurmz* [Rédei 1988: 490] > сам. **sårma* [Janhunen 1977: 136] (если в прауральском в первом слоге действительно был *u*, по-видимому, мы имеем дело с нерегулярным рефлексом гласного) > т. нен. *sarmə*. Северно-самодийские слова свидетельствуют о прасеверносамодийском **sarwa*, мансийский — о прамансийском **sūrma* (и, вероятно, более раннем *sarma*, об условиях перехода *a* > *ū* в мансийском см., например, [Reshetnikov, Zhivlov 2011]).

1.3.5. Еще один бродячий термин может быть представлен названием куропатки. Редеи предлагает следующую прауральскую этимологию: U **aŋz* ‘куропатка’ [Rédei 1988: 12]. В этой статье названия куропатки объединены с названиями различных видов уток. Если убрать из статьи саамское название длиннохвостой утки (которое сам Редеи приводит под вопросом) и самодийские названия уток (объединение, предложенное в этой статье, критикуется и в [Helimski 1997a: ст. 329]), у нас остается следующий набор: хант. V *ijŋk*, DN *eŋx* // манс. (MSz.) T *āŋghā*, K *əŋkhè*, (Kann.—Liim.: MSFOu 101: 400) P *āŋkə* (Kann.—Liim.: MSFOu 101: 14, 93) KM *ēŋkə*, So. *aŋxa-ŋ* ‘полный куропатками’ // сам.: сельк. *ū* [Alatalo 2004: ст. 2], камас. *uŋŋa* [Donner, Joki 1944: 81], матор. *oŋgo* (*oŋo*) [Helimski 1997a: ст. 329]. При этом обско-угорские формы возводятся к праобскоугорской форме **ijŋki*, а самодийские — к праформе **iŋa* ~ **oŋa*. Объединить фонетически схожие обско-угорскую и самодийскую (распространенную только в южном ареале) праформу в одну этимологическую статью, тем не менее, невозможно, так что и в этом случае речь идет о бродячем термине.

1.3.6. Еще одним бродячим термином, видимо, является название шишки хвойного дерева. Редеи предлагает реконструкцию U **pākz* ‘шишка хвойного дерева’ [Rédei 1988: 362] со следующими рефлексам: марийский J *püül’ü* и т. д. // хант. (OL 15) V *pōki*, DN *pāx*, Kaz. *paw* // манс. KM КО *pōk*, P *pāk*, So *pāk*. // сам.: сельк. *pūkā* [Alatalo 2004: 564]. В действительности вокализм первого слога в селькупском не позволяет считать слово *pūkā* рефлексом праформы *pākz*. Вместе с тем, в ваховском диалекте хантыйского представлено также слово *pik*. Оно не только не восходит к данной праформе, но, по всей видимости, вообще не имеет общехантыйского прототипа, так как содержит неэтимологический гласный: ваховский *i* может восходить к прахантыйскому *ī*, но в стандартном случае после *p-* он переходит в *ū*). Редеи отмечает также нестандартную рефлексацию гласного в кондинском хантыйском *pōw*. Таким образом, в данном узком ареале мы опять обнаруживаем семейство фонетически близких основ, обозначающих единое понятие, которые, однако, несводимы к одному прототипу.

1.3.7. U **ončz* ‘нельма’ [Rédei 1988: 339]. За пределами Западной Сибири это слово представлено только в коми: коми *už* // хант. J *unč-mok* (*mok* ‘малёк’), DN *unč*, O *wus*, *us* // манс. TJ *oškōl* (*kōl* ‘рыба’), KU *ūš*, LU *uš*, So. *ūs* // сам. **ancz* [Janhunen 1977: 18]. Словарная статья Ю. Янхунена составлена буквально следующим образом: «(ср. нган. *jintu* / т. эн. *jiddu*) л. эн *adde* / ср. т. нен. б.-з. *ŋанды* ‘сывка = маленькая нельма’ (? < -ŋt-) / сельк. (Pr.) *wenti*, (D) *Tu ŋāñdž*» [Ibid.]. Ю. Янхунен фактически показывает условность объединения перечисленных самодийских форм в одну этимологическую статью. Приведенные самодийские формы в действительности отражают три праформы:

1. **aŋto* > т. нен. *ŋанду*, л. эн. *adi*;
2. **jintu* > нган. *jintu*, т. эн. *d’idu*;
3. **ančV* > сельк. *uāñčə* — только эта форма является закономерным продолжением предложенной уральской праформы. Кажется, что название нельмы, представленное несколькими фонетически достаточно близкими, но все же не сводимыми друг к другу формами, является в Западной Сибири очередным «бродячим словом»; кажется любопытным отметить, что прауральское слово наряду с обско-угорскими языками сохранил селькупский, тогда как в других самодийских языках представлены

рефлексы иных форм. Либо в северносамодийских прауральское название было замещено (двумя разными) названиями нельмы, либо, возможно, его изначально в северносамодийских и не было, и отнесенное Редее к прауральскому лексическому фонду название нельмы **ončz* в действительности является ареальной инновацией, распространившейся только в обско-угорских, коми и селькупском.

1.3.8. У **kärz* ‘вид рыбы (стерлядь?)’. Исключая единичную, и потому признанную самим Редее ненадежную фиксацию коми слова в материалах Палласа, получаем: хант. V *köri*, DN *kärə*, O *käri* // манс. N *kārāj* ~ *kārēj* ~ *karāj* [*kāraj*], T *kärkēj*, P *kārāj* // сам. т. нен. (ям.) *хыри*, л. нен. *kirī*, сельк. *ķōr(ə)* [Alatalo 2004: ст. 2260] ‘муксун’, камас. *ķuro* ‘ленок’, койб. *churrū* ‘*Salmo coregonoides*’. Для прахантыйского, согласно реконструкции обско-угорского вокализма в [Живлов 2006], восстанавливается форма **keri*, для прамансийского **kārγāj*, для праненецкого **kjri*, для южносамодийских **korə*. Возможно, такие нерегулярные соответствия в общем для данного ареала названий ценной промысловой рыбы объясняются тем, что мы имеем дело с бродячим словом, которое может многократно заимствоваться внутри ареала — ср., например, отчасти сходную историю с названием сига, отраженную в [Аникин 2000: 705].

1.3.9. У **älz*- ‘древесный сок’ [Rédei 1988: 24]. В этой статье приводится саамское слово, а также хантыйское, ненецкое и селькупское. Добавив мансийский, мы получаем следующую картину для интересующего нас западносибирского ареала: хант. V *ōl*, DN *ēlə*, Kaz *äl* // манс. N *il’* ~ *il*, LM *il’* ~ *il’*, LU *il’* ~ *il’*, P *il’*, T *ul’* // сам. [т. нен. *nyl’u*], л. нен. *ñil’i*, сельк. *ülu* [Alatalo 2004: ст. 341]. Для прахантыйского, согласно реконструкции обско-угорского вокализма в [Живлов 2006], восстанавливается форма **ül*, прамансийскую форму восстановить невозможно из-за нерегулярных вокалических соответствий между мансийскими языками / диалектами, в форме т. нен. *n-* не является протезой (перед *i* ожидался бы *η-*) и потому пратундровоненецкая праформа выглядит как *ñil’i*, а вот праформа лесного ненецкого — как **il’i*. Очевидно, что праформы несводимы друг к другу даже в рамках подгрупп, поэтому об общеуральском наследии в данном случае говорить не приходится.

1.3.10. Манс. **ñjra* ‘болото’ > N *ñār* [=ñār], LM *ñār* ~ *ñār*, LU *ñār*, P *ñār*, K *ñōr*, T *ñār* ~ *ñār* ~ *ñēr*; TJ TČ *ñēr*, KU *ñār*, KM *ñēr*, KO *ñēr*, P *ñēr*, VNZ VS *ñēr*, VNK *ñēr*, *ñēr*, LU LM *ñēr*, LO So. *ñār* ‘болото (с неглубокой прозрачной водой, с деревьями и цветами)’ // сам. т. эн., л. эн. *norō* ‘болото’, нган. *n’eryā* ‘земля, покрытая слоем воды, болотистая местность’, сельк. *ñarə* ‘болото’ [Alatalo 2004: ст. 1690]. Я. Алатало указывает на возможность сопоставления с эвенкийским *ñяру* — учитывая это, данное слово можно также отнести к группе западносибирских бродячих слов.

1.4. Термины пространственной ориентации

Ниже в отдельную группу вынесены несколько терминов, связанных с пространственной ориентацией; часть из них вошли в систему послелогов угорских, обско-угорских либо самодийских языков.

1.4.1. Самое широкое распространение имеет показатель, который К. Редее возводит к общеуральскому **nā* [Rédei 1988: 300], определяя его значение как ‘этот; ? тот’. Впрочем, и сам автор этимологического словаря не вполне уверен в правомерности объединения всего приводимого им материала в рамках одной этимологической статьи: эта основа отмечена им как достоверно возводимая на финно-пермский уровень и под вопросом — как возводимая к уральскому праязыку. Действительно, эта этимологическая статья отчетливо разделяется на две части: материал финно-пермских языков (финский *nāta* ‘этот’ // эстонский *nemad* ‘они’, *nee*, *need* (Pl) ‘они, те же самые’ // саамский ? N *nā* ~ *nāvt* (Adv) ‘так, таким образом’ и т. д. // мордовский E *ñe*, *ñeí*, M *ñä*, *ñat* (Pl) ‘этот’ // марийский

KB *ninə*, U *ninə*, B *nine*, *nāne* ‘этот, (Pl) они’ // коми V *na* ‘они’: *na dinə* ‘к ним’, V *naja*, *naje* ‘они’ и т. д.) позволяет восстановить основу с указанным дейктическим значением. Но угорско-самодийский материал дает основание для реконструкции пространственного показателя со значением ‘возле, рядом’. Даже если принять гипотезу К. Редей о том, что финно-пермский, угорский и самодийский показатель восходят к одному источнику и, следовательно, что этот пространственный показатель мог возникнуть из указательного местоимения ‘этот’ (а такое предположение не выглядит очень убедительным), существенно, что этот путь грамматикализации оказался задействован в угорско-самодийском ареале, но остался совершенно невостребованным в финно-пермских языках. Ниже процитирован угорский и самодийский материал, объединенный К. Редей в этой словарной статье:

хант. (Karj. — Vértes: MSFOu. 128: 330) O *-nā* (Lat., Illat.): *χatnā* ‘в дом’, (598) V *-nāt*, *-nat* (Kom.): *öyil-nāt* ‘со своей дочерью’, *pāyäl-nat* ‘со своим сыном’³, (Stein., OstjChr. 149) Šerk. *nataŋ: nāj nataŋen* ‘с тобой’ // манс. (Liim.: MSFOu 127: 78, 80, 82, 85, 88, 89) Tj *-nā*, *-nā*, Ku *-nə*, *-ən*, *-n*, P *-nə*, *-n*, So. *-na*, *-n* (Lat.): Tj *ōš-nā* ‘к городу’, Ku *χāp-nə* ‘в лодку’, P *wor-nə* ‘в лес’, So. *lāyləm-na* ‘на мои ноги’, (ebd. 60, 61) Tj *-nāl*, *-nāl*, Ku So. *-nəl* (Abl., El.): Tj *kül-nā:l* ‘из дома’, *jē-nā:l* ‘от реки’, Ku *ūs-nal* ‘из Тобольска’ (*ūs* ‘город; Тобольск’), So *šēp-nal* ‘из сумки’, (Liim.: MSFOu 127: 120) Tj *-nāt*, *-nāt* (Kom.): *āmp-nāt* ‘с собакой’, *lom-nāt* ‘с моей лошастью’ // венг. (диал.) *-ni*, *-nyi* ‘к’, *-nek*, *-nak* (Dat., Gen.); напротив, на, по направлению к’, (диал.) *-nól*, *-nól*, *-nül*, *-nül* ‘от’, *-nött*, *-nott*, *-nitt*, *-nött* ‘y’ // сам.: **nä*- (послелог) [Janhunen 1977: 99].

Этот угорский материал привлекал для сопоставления с самодийским и Е. А. Хелимский, называвший одной из возможных угорско-самодийских параллелей «п-овые показатели пространственных падежей», о чем уже кратко говорилось в начале статьи. Хелимский привлекает к сопоставлению общесамодийский показатель латива *-ntz* и дополнительно — селькупский показатель датива-аллатива *-nij*. Но это сравнение нельзя признать удачным. Во-первых, нет никаких свидетельств в пользу того, что общесамодийский показатель латива является составным, а не содержит исконный кластер *-nt-*, часто встречающийся и в лексических основах, и в грамматических показателях самодийских языков. Во-вторых, общесамодийский показатель латива и селькупский показатель датива-аллатива занимают разную позицию в словоформах посессивного склонения существительного: общесамодийский показатель латива занимает позицию перед посессивными показателями⁴, селькупский показатель датива-аллатива — после посессивных показателей. Этот факт говорит о более позднем послеложном происхождении селькупского датива-аллатива, и главное — о том, что селькупский датив-аллатив и общесамодийский латив не могут одновременно рассматриваться как варианты самодийских п-овых падежей, так как они представляют явления, имеющие разную датировку. Предложенное Редей сопоставление угорских данных с самодийским послелогом *nä-* выглядит более перспективным. Приведенные Редей угорские данные, кажется, довольно очевидно свидетельствуют о послеложном происхождении перечисленных угорских форм: все они имеют единую основу **nä*, которая, будучи оформленной дополнительно одним из пространственных показателей, получает конкретное значение: латива, элатива, комитатива. Такой же набор значений и такую же морфологическую структуру имеет и послелог **nä-* в самодийских языках, ср. материал тундрового ненецкого ниже. В самодийских языках послелоги могут иметь одну из четырех форм пространственных падежей (пролатива, латива, локатива и элатива). Примеры (1), (2) и (3) иллюстрируют

³ Для ваховского диалекта см. также описание «творительно-совместного падежа» с показателем *-na(mi) / -nä(mi)* в [Терешкин 1961: 51–52].

⁴ В селькупском языке в посессивном склонении три падежа — латив, элатив и локатив — формально не различаются: во всех трех случаях употребляется падежный показатель локатива, восходящий к общесамодийскому локативу, однако позицию в словоформе общесамодийского показателя латива легко установить с опорой на данный других самодийских языков.

употребление послелого т. нен. *ня-* (< **-nā*) в лативной форме (для обозначения направления движения в (1) и для обозначения достижения конечной точки в (2))⁵ и элативной форме (3):

- (1) *уарк-''* *хоё-''* *ня-'* *сауо-да* *хамы-''*
 большой-GENPL хребет-GENPL к-LATADV взгляд-3 упасть.AORRFL-3
 ‘В сторону больших холмов взгляд его упал’ [Терещенко 1990: 89 // Т3, предл. 96].
- (2) *ха-вы* *ня-хаю-та* *ня-'* *тэв-''* *ним''*
 умереть-PTPRAET товарищ-DU-3 к-LATADV прийти-CN NEG.AOR.R3.EMPH
 ‘дошел до убитых товарищей’ [Терещенко 1990: 65 // Т2, предл. 252]
- (3) *уарка* *яв-'* *ня-д* *тир* *тантал-й-''*
 большой море-GEN к-ELADV облако подняться-AORRFL-3
 ‘Со стороны большого моря облако поднялось’ [Терещенко 1990: 226 // Т8, предл. 10].

Локативная форма послелого *ня-* может употребляться в собственно локативном значении (4), но также и в значении латива, причем как для обозначения движения в сторону определенного объекта (5), так и для обозначения достижения конечной точки (6):

- (4) *Тарем'* *сыр-па''-ни* *хано-''* *вар-'* *ня-на* *уб''*
 так посмотреть-VA COND-1 нарта-GENPL край-LAT к-LOCADV один
хан *сяры-Ø*.
 нарта быть.привязанным.AOR-S3
 ‘Если посмотреть, на краю нарт одна нарта привязана’ [Янгасова 2001: 231]
- (5) *хабта-ри* *ёнар-та* *ня-на* *хэ-бён-да*
 бык-LIM тысяча-GENPL.3 к-LOCADV уйти-AUD-3
 ‘В сторону тысячи быков он, слышно, поехал’ [Терещенко 1990: 60 // Т2, предл. 67].
- (6) *седа-ко-'* *ня-на* *то-Ø*
 сопка-DIM-GEN к-LOCADV прийти.AOR-S3
 ‘пришел к сопке’ [Янгасова 2001: 181]

Этот же послелог используется для выражения комитативного значения. С существительными комитатив передается лативной формой послелого (7), в прономинальных формах комитатива (которые представляют собой падежные формы послелого, оформленные соответствующим посессивным показателем) употребляется локативная форма послелого (8):

- (7) *папако-н* *ня-'* *Вэбты''я* *тэта-'* *мят'*
 младший.брат-GEN.1 к-LATADV PN богач-GEN чум.LAT
мядон-зь *мани'* *пяриугу-ми'*
 гостить-VA INF мы.DU отправиться.AOR-O1DU
 ‘Мы с младшим братом отправились в гости в чум богача Вэбты''я’ [Терещенко 1990: 182 // Т6, предл. 121].
- (8) *ня-на-нд*
 к-LOCADV-2
 ‘с тобой’

Такое смешение функций лативной и локативной формы одного и того же послелога (при формальном четком противопоставлении этих форм), которое представлено

⁵ В самодийских языках (особенно в северносамодийских) послелого с лативным значением, как правило, ограничены обозначением направления движения (но не достижения конечной точки: ‘пошел к’, но не ‘подошел к’). С другой стороны, это ограничение несущественно для лативного падежа. Преодоление этого ограничения ненецким послелогом свидетельствует о высокой степени его грамматикализации и фактически — о нарастающей конкуренции с лативом.

примерами (4)–(6) и (7)–(8), не встречается у других серийных послелогов в самодийских языках; возможно, это неразличение — архаичное наследие синкретичной семантики основы послелога: примечательно, что в обдорском диалекте хантыйского падежный показатель *-na* передает не только значение латива, но и значение локатива и инструмента [Николаева 1995: 82–83].

Таким образом, угорские и самодийские языки, очевидно, разделяют общий источник грамматикализации показателя с максимально общим значением ‘у, возле, около’, который употребляется в значении латива, локатива, элатива и комитатива. Наиболее последовательно эта грамматикализация прошла в мансийском и самодийских языках.

1.4.2. Еще одна послеложная основа — общая на этот раз только для обско-угорских и самодийских — основа со значением ‘позади’. К. Редеи под вопросом включает в эту статью также финский материал, но отмечает, что это сопоставление с финским фонетически не вполне достоверно. Обско-угорская и самодийская часть этимологии представлена следующим материалом: U **pujz* ‘задняя часть’ [Rédei 1988: 401]: хант. V O *puj*, DN *pij* ‘задняя часть (DN), задний (V DN), задница (V DN O), комель (O)’ // манс. T P *puj* K *pōj* ‘задняя часть’, T *pujit*, *pojit* ‘позади, следом’, TJ *poj*: *ām pojə mt* ‘позади меня’, So. *puj* ‘задний’ // сам. **puə* (? **puwə*) ‘задняя часть’ [Janhunen 1977: 129]. Во всех самодийских языках эта основа функционирует как послелог (наряду с другими возможностями — например, есть производные от этой основы глаголы ‘быть позади’ и т. д.). Таким образом, здесь мы видим появление в рассматриваемом ареале общей основы для обозначения задней части ориентира. Наблюдается также грамматикализация этой основы: переход ее в сферу послелогов (в южном мансийском и во всех самодийских языках).

1.4.3. Следующая обско-угорско-самодийская параллель при обозначении пространственной части ориентира (верха) представлена следующими словами:

U **numz* ‘верхняя часть, небо; бог’ [Rédei 1988: 308]: хант. V DN O *num* ‘верхний, верхняя часть. небесный’, V *nəmən*, DN *numən*, O *nomən* ‘наверху (V DN O), выше (в том числе по Оби) (O); на; верховный (O)’ (...) // манс. T P N *num*, K *nəm* ‘верхняя часть’, TJ *nomē*, KU *nom*, So. *numi* ‘находящийся наверху’, P *numk* ‘верхний, верховный’, KU *nomən*, P *numən* ‘наверху’, TJ *noŋ*, KU *noχ*, P *nuj*, So. *nəŋχ* ‘наверх’ // сам. **num* ‘небо; погода, бог’ [Janhunen 1977: 104]. Квазиграмматическое (адвербиальное) пространственное значение представлено у этой основы в обско-угорских языках, а также в камасинском. В остальных самодийских языках она обозначает только ‘божество’ или ‘небо’ (что выглядит скорее как заимствование праобскоугорского многозначного слова только в одном из его значений, чем как полная утрата у этого слова во всех самодийских языках каких-либо следов исконного пространственного употребления). При этом в самодийских языках распространение этой основы тяготеет к южному ареалу (самый северный из самодийских языков, в котором представлен рефлекс прасамодийского **num*, — ненецкий, слово представлено также в селькупском, камасинском и маторском). Два последних факта заставляют задуматься о том, что, возможно, эта основа не является общесамодийской, но ее распространение в центральной и южной части самодийского языкового континуума (в эпоху, предшествовавшую полному обособлению отдельных самодийских языков) объясняется контактами с обско-угорским праязыком. Анонимный рецензент обратил мое внимание на то, что селькупская форма восходит скорее к ПС **nom*, которое, как и хант. **nəm-*, может быть регулярным рефлексом **numi* (которое также не выходит за пределы обско-угорской и самодийской групп). Поэтому не исключено, что в самодийском представлен как исконный рефлекс этого корня (в селькупском), так и заимствование (в остальных языках); впрочем, версия независимого заимствования в камасинский, маторский и ненецкий представляется мне все же маловероятной.

1.4.4. Еще одна угорско-самодийская основа (она обозначает заднюю поверхность, оборотную сторону) представлена следующей этимологией: U **kuttz* ‘спина’ [Rédei 1988: 225]: хант. J *kutəl*: *juw kutələnə* ‘под защитой дерева’ // манс. KU *χūtəj*, P *kūtəy* ‘в тени чего-л, под защитой чего-л’, N *χūtäi* ‘за (= скрываясь за какой-либо предмет)’ // венг. *hát*

‘спина; оборотная сторона’ // сам. сельк. *ḡottä* [навзничь] [Alatalo 2004: ст. 1934] // камас. *k'u'd'žë* ‘навзничь’ (на камасинскую параллель указано в [Ibid.]). В этой этимологии представлена довольно древняя, общеугорско-самодийская параллель: название спины, оборотной стороны предмета. Семантическое развитие этой основы пошло разными путями: в венгерском она сохраняет лексическое значение, в обско-угорских развилась в послелог со значением ‘под защитой’ (< *‘за спиной’), в южносамодийских (селькупском и камасинском, которые, на мой взгляд, дольше других сохраняли ареальные связи с обско-угорскими языками, о чем см. ниже, — именно эти ареальные связи могли способствовать сохранению этой основы) — в пространственное наречие ‘навзничь’ (< *‘на спину’). В данном случае мы видим возникновение общей угорско-самодийской основы, но ее сепаратное семантическое развитие в отдельных языках.

1.4.5. Следующая параллель является мансийско-самодийской. К. Редей прибавляет к этой этимологии также финский и мордовский материал и реконструирует на их основе праформу U **jotka*. Но рефлексом U *-*tk-* (и совпавшего с ним в мансийском U *-*sk-*) является *kt / χt / γt*, ср.: U **utka* ‘дорога, след’ [Rédei 1988: 546] > манс. KU *āχt*, P *ekt*, LO *āγt* или U **eskz-* ‘верить’ [Ibid.: 74] > манс. KU *āγt-*, So. *ayt-*. Таким образом, если отказаться от сопоставления с финно-волжскими словами, у нас остается следующий материал [Ibid.: 102–103]: манс. TJ *jät*, KU *jat*, P *jät* (в [Kálmán, Munkácsi 1985] приводятся следующие формы: N *jat* ~ *jot*, LM *jät* ~ *jot*, LU *jät* (~ *ät*), P *jät*, K *jät* ~ *jot* (~ *jöt*), T *jët* ~ *jät*) ‘середина, средний’ // сам. **jotä* ‘середина’ [Janhunen 1977: 47]. Эта этимология представляет собой пример ареально обусловленного возникновения основы, обозначающей среднюю часть ориентира, в обско-угорских и самодийских языках. Дальнейшее развитие этой основы пошло разными путями: в мансийском она осталась в разряде существительных, в самодийских, наряду с сохранением собственно лексического значения, приобрело дополнительные функции послелога ‘среди, посреди’.

1.4.6. Следующая обско-угорско-самодийская параллель — лексема с пространственным значением ‘сторона’: хант. Kг. V *jir*, Kaz. *ḡir*, O *īar* ‘край предмета, боковая поверхность предмета’ [Karjalainen 1948: 181] // манс. LM *jēr* ~ *jer* [*jēr*], T *jēr* ‘направление, область возле предмета, сторона, место’⁶ [Kálmán, Munkácsi 1985]; KU KM Lo So. *jēr* [Honti 1982: 143] // сам. т. нен. *jār*’, л. нен. *jëär* ‘боковая поверхность предмета’ // сельк. *čār* ‘сторона, область возле предмета’ [Alatalo 2004: ст. 1586].

1.4.7. Следующее слово в этой группе представлено названием оконечности, крайней выступающей части ориентира (носок обуви, конец палки и т. п.): U **ukz* ‘голова’ [Rédei 1988: 542]: хант. V *у* ‘голова; носок (обуви)’, DN *иχ* ‘голова (зверя, человека), спичечная головка, оконечность веретена, оконечность любых предметов с утолщенным концом; конец шеста’, Ni. *иχ*, Kaz. *ōχ* ‘голова; носок (обуви); капюшон’ // манс. KU *āwt*, P *ēwt*, So. *āwa*, LM *āw* ‘голова медведя’ // сам. матор. *ugudu* ‘клюв’, т. нен. *ḡiχud*, л. нен. *ḡiχuntī* ‘верхняя губа’, т. эн., л. эн. *ibo* ‘конец’ (аналогичное фонетическое соотношение интервокальных согласных представлено т. нен. *туху*, т. эн. *tubo*), но также т. эн. *иходо* ‘верхняя губа и место под носом’, сельк. *ūkə* ‘оконечность, морда, голова, губа, сосок, нос лодки’.

1.4.8. Еще одной такой параллелью может служить следующая: в мансийском есть послелог / преверб, который в словаре [Kálmán, Munkácsi 1985] переводится как ‘с’, но в действительности, помимо комитативного значения, выражает также значения ‘в отношении чего-либо’ (*atj mænər at woḡrax tæw jətälət* ‘ничего с ним не поделаешь’, *täu am jätəm äk-vujip* ‘он со мной одинаковый’), ‘благодаря чему-либо’ (*taš-kharät jät öli* ‘он живет своими запасами’). В словаре [Kálmán, Munkácsi 1985] приводятся следующие формы этого преверба / послелога: N *jot-* [= *jot-*], LM *jet-* ~ *jät-* ~ *jit-* ~ *jât-*, LU *jätt-*, P *jät-* ~ *jit-*, K *jât-* ~ *jot-*, T *it* ~ *jit-* ‘с’

⁶ По семантике мансийскому *jēr* в большей степени соответствует южнохантыйское DN *ḡōrəχ*, Kг. *ḡūrəχ* [Karjalainen 1948: 183], но фонетически оно несопоставимо с манс. *jēr*. Совпадающее с южнохантыйским *ḡōrəχ*, *ḡūrəχ* слово представлено в кондинском мансийском: K *jürəχ* [Kálmán, Munkácsi 1985], но оно, судя по ареалу его распространения, может быть заимствованием из хантыйского.

(преверб) // N *jot*, P *jät*, K *jot* ~ *jåt*, T *jit* 'с' (послелог). В действительности эти мансийские формы несводимы к одному прототипу. В [Honti 1982: 144] выбраны формы только восточного и северного мансийского, они возводятся к прамансийскому **jõt* и сопоставляются с восточнохантыйскими VVj *jõt* [Steinitz 1966: 416] / *iõi* [Karjalainen 1948: 195], Tj. *jõt*, J. *jõta*. Если считать корректной запись К. Карьялайнена, в соответствии с предложенными в [Живлов 2006] правилами прахантыйская форма будет иметь вид **jut*. Итак, и прахантыйское **jut*, и предложенная Л. Хонти прамансийская форма **jõt* возводятся к праобскоугорскому **jotj*.

Оставшиеся мансийские формы западного и южного диалектов — LM *jet-* ~ *jät-* ~ *jit-* ~ *jåt-*, LU *jätt-*, P *jät-* ~ *jit-*, T *it* ~ *jit-* 'с' (преверб) // P *jät*, T *jit* 'с' (послелог) возводятся к прамансийской форме ПМ **jätV* и далее к обскоугорской праформе ПОУ **jetV*. Праобскоугорское **jetV* является хорошим формальным соответствием самодийского **jet(a)* > т. нен. *ed'*, т. эн. *d'eʔo*, л. эн. *d'ed*, нган. *ɔziɣuʔ*, *ɔɔ*, сельк. *čitə* 'про (что-то говорить), по (чему-то ударить)'. Что касается семантики, можно предполагать, что в мансийском сохраняются только поздние, непространственные употребления (как было показано выше, и в самодийских один из направительных послелогов развивает комитативные употребления).

Таким образом, мы находим в угорских и самодийских языках достаточно систематическое ареально обусловленное развитие общих лексем в сфере пространственной ориентации. Представленность этих лексем в угорских языках заставляет говорить о том, что эта тенденция началась еще до разделения угорской языковой общности (две из восьми рассмотренных этимологий содержат, наряду с мансийским и хантыйским, также венгерский материал) и продолжилась уже после ее распада (оставшиеся шесть этимологий представлены только обско-угорскими параллелями).

2. Контакты между мансийским и южносамодийскими

Этимологии, в которых объединяется материал только обско-угорских и южносамодийских языков, рассмотрены и в предыдущем разделе, хотя их не слишком много (см., например, *čača-* 'идти'). При этом в подобных этимологиях из первого раздела в целом выдерживаются те фонетические соответствия между (обско-)угорскими и самодийскими, которые известны по корпусу общеуральских этимологий. Я считаю возникновение таких параллелей ареально обусловленным, но при этом достаточно давним процессом, в который могли быть в равной или практически в равной степени вовлечены предки и южносамодийских, и северносамодийских языков. В данном разделе будут рассмотрены иные случаи. Во-первых, это сепаратные изоглоссы, объединяющие мансийский с селькупским и мансийский с камасинским. В этом случае фонетические соответствия отличаются от фонетических соответствий «общеуральского» типа. Сепаратные соответствия, в свою очередь, могут быть двух типов. Наиболее очевидный случай представлен такими основами, которые из обско-угорских языков представлены только в мансийском. Более сложными являются случаи, когда слово представлено и в хантыйском, и в мансийском, при этом южносамодийская форма не выводима из праобско-угорской либо из хантыйской, но может быть связана фонетическими соответствиями с мансийской.

2.1. Фонетические соответствия между селькупским и мансийским, представленные в данной основе, говорят в пользу контактного возникновения этого термина в селькупском: фин. *korpi* 'дремучий лес, глушь' [Aikio 2014: 8–10], манс. Tj *kārp*: *kēl'*-k. 'березовый лес' (*kēl'* 'береза') KU *χōrəp* 'березовый лес', P *kōrp* 'лес (пихты, ели и березы, с сухой почвой), еловый лес' So. *χōrp-ur* 'старый лес — кедр, ели, березы и лиственницы' (*ur* — 'гряды') [Rédei 1988: 217]⁷; в слове [Kálmán, Munkácsi 1985] значение 'еловый лес'

⁷ К. Редей не приводит селькупское слово, но приводит мордовское и ненецкое, сопоставление которых с мансийским следует отвергнуть по фонетическим причинам.

отмечено для кондинского, пельмского и нижнелозьвинского диалектов // сельк. *kuārpā* ‘еловый лес; «низкий лес, на бору бывает»’ [Alatalo 2004: ст. 2253]. Поскольку для общеселькупского Я. Алалато восстанавливает веллярный (а не увулярный) анлаутный согласный, данное слово может рассматриваться как заимствование в селькупский из мансийского (заимствование достаточно раннее — такое, что в селькупском гласный первого слова успел подвергнуться дифтонгизации)⁸.

2.2. Селькупское название белки *tapāk* [Alatalo 2004: ст. 972] изолировано на самодийском фоне: предложенное там же Я. Алалато сопоставление с камасинским *t'āžā'β* фонетически не выглядит достаточно убедительным (это сопоставление, по всей видимости, предполагает возможность метатезы второго и третьего согласных в селькупском либо камасинском). К. Редей сопоставляет его с венгерским и хантыйским названием белки и возводит на прауральский уровень: U **sāpз* (*šāpз*) или *tāpз* ‘белка’ [Rédei 1988: 436]: манс. *tap* в N *mā-tapriś* ‘мышь; крот’ (*mā* — ‘земля’, *-riś* — диминутивный суффикс) // венг. *evet* ‘белка’. При этом К. Редей предполагает, что венгерское слово может относиться и к другой, собственно угорской этимологии: Ug. **θāpз* ‘белка’ [Rédei 1988: 873]; в этом случае для мансийского и селькупского он предлагает уральскую праформу **tāpз*. Реконструкция переднего гласного в этом случае, вероятно, основывается на неточной селькупской записи *tāpāk*, которую цитирует Редей. При этом важно, что в правильной селькупской форме *-a-* первого слога может восходить только к прасамодийскому **ə*, что исключает возможность сопоставления с мансийским **tāpз* с опорой на «общееуральские» соответствия в сфере вокализма. Но при этом можно допустить заимствование в селькупский из мансийского формы типа *tāp* с последующим суффиксальным оформлением в селькупском — это предположение, кстати, объясняет колебания в фиксации селькупского гласного первого слога, ср. северноселькупские формы типа *tāpāk* или *tēpeq*: можно предполагать, что нехарактерное для селькупского *ä* первого слога в итоге перешло в *a*. Гипотеза о локальном заимствовании выглядит тем более привлекательной, что есть общесамодийское название белки **t'āpā-* [Janhunen 1977: 149] и маловероятным выглядит предположение о том, что это самодийское название (отсутствующее только в камасинском и селькупском) представляет собой инновацию, а изолированное селькупское название белки представляет собой архаизм, восходящий к прауральской форме.

2.3. В обско-угорских языках есть название теленка оленя: манс. **pāčəy* (реконструкция вокализма в соответствии с [Honti 1982: 173]): LO So. *pāsiy*; хант. **pāčiy* (реконструкция вокализма в соответствии с [Живлов 2006]): VVj *pečəy*, Ко. *pečə*, Ni *pešə*, Kaz. *pešī*. В соответствии с изложенными в [Живлов 2006] правилами, праобскоугорская форма реконструируется как **pāčəy*. Этой форме можно сопоставить селькупское название желтого цвета, сохранившееся в кетском и верхнеобском диалектах: *pačəl'* ‘желтый’ [Alatalo 2004: ст. 418b]. Связь названия желтого цвета с названием олененка проявляется в самодийских языках, ср. т. эн. *tasoku* ‘новорожденный олененок; коричневатый цвет’ и т. нен. *tasco* ‘желовато-бурый’ и *tasco* ‘новорожденный олений теленок’. Вероятно, это заимствование является достаточно ранним: оно отражает мансийский интервокальный *č*, впоследствии перешедший в *š*.

2.4. В восточнохантыйских диалектах есть название мертворожденного ребенка: V Vj. Vart. Trj. *pāčak*, Lirk. Mj. Trj. *pāček*, Kam. *pāčex* [Steinitz 1966: 1095]. О мертворожденном ребенке у Штейница сказано, что «он внушает человеку страх, так как он превращается в духа. Чтобы он не бродил, его хоронят в берестяной коробочке и кладут ему на живот камень; тромъеганские ханты считают, что он превращается в птичку, и хоронят его

⁸ Я благодарю анонимного рецензента, предложившего альтернативное объяснение истории этого слова: возможно, мансийскому слову родственно самодийское **kārwlV* ‘лиственница’ > т. нен. *xarw* ‘лиственница’, т. эн. *kami* ‘лиственница’, л. эн. *kamu* ‘лиственница’, нган. *korbi* ‘прямое сухое дерево’, сельк. *kuārkə* ‘сосна’ [Alatalo 2004: 2256], — и в этом случае сельк. *kuārpā* представляет собой вторичное заимствование из мансийского.

с камешком во рту» [Steinitz 1966: 1095]. Это же слово обозначает лесного духа. В селькупском есть слово *pačija* ‘живущий в холмах дух, который воет по ночам и пугает детей’. Общеселькупское распространение этого слова (тымский, северный и кетский диалекты) делают маловероятным заимствование из хантыйского в поздней локальной контактной зоне (контакт между ваховскими хантами и тымскими и северными селькупками). К тому же в случае таких поздних заимствований слова редко получают дополнительное суффиксальное оформление, как правило, являясь точной «копией» хантыйского слова. Общехантыйской формой для этого восточнхантыйского слова было бы, в соответствии с предложенными в [Живлов 2006] правилами, **pičik*. Ему соответствовало бы гипотетическое мансийское **pāčik* и гипотетическое праобскоугорское **pačik*. Как уже говорилось, заимствование из современных хантыйских диалектов маловероятно, праобскоугорская форма фонетически близка к селькупской, но заимствование из праобскоугорского в селькупский — маловероятный сценарий с точки зрения относительной хронологии. Заимствование из прахантыйского **pičik* невероятно по фонетическим основаниям. Остаются две логические возможности: заимствование из правосточнохантыйского (если предполагать, что в нем к этому моменту произошел переход прахантыйского **pičik* в правосточнохантыйское **pāčik*) либо заимствование гипотетической мансийской формы. Второй сценарий, как я надеюсь показать в данной статье, логически более предпочтителен, так как он входит в группу однородных фактов. Это заимствование я также считаю достаточно старым — еще и потому, что в нем отражается мансийский интервокальный *č*, впоследствии перешедший в *š*.

Некоторые заимствования из мансийского в селькупский можно ранжировать по времени в зависимости от того, произошли ли эти заимствования до перехода праобскоугорских *č / č* в мансийский *š / š* или после него. Так, заимствования в пп. 2.2 и 2.3 отражают мансийский **č*. Приводимые ниже в 2.5 и 2.6 заимствования отражают мансийский **č*:

2.5. Манс. КУ *χūs*: *χом-χ*. ‘работник, слуга’, Р *kuš*, Со. *χūs* < праманс. *kōčV* с сохранением праобскоугорского качества гласного (имеет соответствие в венг. *hős* (диал. *hes*) ‘герой, (диал.) неженатый юноша’ [Rédei 1988: 110]) > сельк. *kečə / kočə* ‘слуга’ [Alatalo 2004: ст. 2066] > *keča* ‘младший брат мужа или жены’ [Ibid.: ст. 2067].

2.6. С манс. **pāč* > Т *pōist* ~ *pōst* ~ *pāšt* ~ *pāišt*, ТТ ТČ *pešt* ‘топор’ предположительно можно сопоставить селькупское *piči* ‘топор’ [Ibid.: ст. 514].

В обоих случаях (2.5 и 2.6) мы видим, что мансийский гласный в селькупском отражается как более передний и более узкий — вероятно, это отражает какие-то особенности мансийского произнесения гласных под влиянием последующего мансийского **č*.

2.7. Сельк. *šāk* ‘соль’ [Alatalo 2004: ст. 2579] РТНОК (в конечном итоге слово является иранизмом, что отмечено и в словаре Я. Алатало). Такое же название соли представлено и в мансийском, **čāk* > LM *šāχ* ~ *šex* ~ *šāχ*, LU *šāχ*, P *šāχ*, K *šāχ* (*šāk-*), Т *čix* ~ *čix* ~ *čux* ~ *čōχ* // ТТ ТČ *tšiuχ*, КУ *šāχ*, КМ *šāχ*, КО *šāχ*, Р *šex*, VN VS *šāχ*, LU *šex*. Мансийское слово имеет параллель в венгерском *só*; Я. Алатало (в частной переписке) указал мне также на возможность сопоставления с названием соли в югском и кетском: «The Selkup word *šāk* is identical with Yeniseic **čēy* (only Ket *tē*, Yug *čē*, lacking in the southern languages). The inetymological *š* before back vowel points to a borrowing from Yeniseic, while culturally it is more natural to assume the opposite direction». Тем не менее фонетическая форма мансийского и селькупского слов допускает возможность заимствования из мансийского в селькупский (предположение об этом, как указал мне М. А. Живлов, высказано также в [Korenchy 1972: 78]). В конечном итоге, каков бы ни был путь распространения этого культурного термина, примечательно, что он указывает на то, что селькупский и мансийский в определенный период находились в одном культурном ареале (в который при этом не входил, с одной стороны, хантыйский, с другой — самодийские языки, помимо селькупского).

2.8. Сельк. *širkə* ‘растение’ [Alatalo 2004: ст. 2692] находит параллель в мансийском N *širkə* ~ *širki*, LM *širkə* ~ *širk*, LU *širk*, K *šerk* ‘растение, цветок’. Сюда же может относиться и камасинское *šurgu* ‘вид растения, руссельк. диал. сиб. *михерь* (корни

раньше использовали для приготовления голубой краски для ткани) [Donner, Joki 1944: 96b]. На форму камасинского слова могла повлиять контаминация с заимствованным из тюркского названием вязаля: камас. *šurgend'ek-nò* (*nò* 'трава') < качинский диалект хакасского *šürgeld'ek* (запись К. Доннера). Близкое слово представлено также в хантыйском, ср. казымское *šorəχ* (впрочем, в [Honti 1982: 131] оно сравнивается с манс. К Р LU *šōr*, Со. *šōri*), но в любом случае для хантыйского восстанавливается анлаутный *č-*, что фонетически существенно дальше от селькупского.

2.9. Селькупское *šōšə-kum* / *šūššə-kum* 'человек'; также используется как этноним в нескольких селькупских диалектах [Alatalo 2004: ст. 2656] можно сопоставить с мансийским **šōšəy* ~ **šōšyā* [Honti 1982: 132] > N *sossä* ~ *sässä* [*sossa*], LM *šāššä* ~ *šāši*, LU *šāši*, P *šāšj*, K *sāsē*, коренной, здешний, исконный, ср. в особенности сочетания: N *taw tit ālnē sossä χum* 'он — обитающий здесь, коренной житель'; P *tau tit šāšj khum, pēš-vuul tākw moyät tit' ōli* 'он — здешний человек, он живет на своей земле от рождения'; K *sāsē khqm, tēlēm mōtē tēt ōli* 'он здешний человек, место его рождения здесь'. Подробнее о возможности мансийской этимологии селькупского этнонима см. [Урманчиева 2020]. Соответствующее слово представлено и в хантыйском, ср. DN Ко *čāčə*, Ni. *šōšə*, Kaz *šōšī*, O *sāsi* [Honti 1982: 132], но эти хантыйские диалектные формы восходят, в соответствии с правилами, предложенными в [Живлов 2006], к общехантыйской форме типа **čiči* — и в этом случае очевидно, что мансийские формы фонетически существенно ближе к селькупскому.

В 2.10. Селькупский *s* отражает мансийский *ńs* < *ńc* (при отражении мансийского **ńc* более вероятным кажется селькупский рефлекс *nc*).

2.10. Название селькупского рода с реки Парабель *maš qul* [Прокофьев 1935: 12], вопреки объяснению Г. Н. Прокофьева (повторенному в [Rédei 1988: 263]), не может быть объяснено как 'земляные люди': в селькупском отсутствуют морфологические средства, которые позволили бы получить от гипотетической селькупской основы **ma* производную *maš* со значением 'земной' (к тому же, как показано в [Janhunen 1977: 85], продолжением самодийского **majā* 'земля' является селькупское *mačə* 'гряда, материк (= поросшая лесом возвышенность)' [Alatalo 2004: ст. 799]). С другой стороны, сельк. *maš* является хорошим формальным соответствием манс. **māńcī* > N *mańsi* [*māńsi*], LM *māńš* ~ *mońš*, LU P *mańš*, K *māńš* ~ *mańš*, T *māńši* ~ *mā'ńci* 'манси', что может указывать на мансийское происхождение данного селькупского рода. Селькупский *-s-* регулярно восходит к п-сам. *-ns-*, ср., например, бесспорно этимологически соответствующие друг другу сельк. *kūšə* 'моча' [Alatalo 2004: ст. 2212] < п.-сам. **kunsə* [Janhunen 1977: 77] и манс. N *χuńši*, LM *khuńši* (*kuńši*), LU P *khuńši*, T *khuńšánt* 'мочиться'.

Приведем еще ряд возможных эксклюзивных мансийско-селькупских ареальных изоглосс.

2.11. На эксклюзивную параллель селькупского *kum* 'человек' и мансийского *χum* с тем же значением (это слово не представлено ни в других самодийских языках, ни в хантыйском) обращал внимание еще Г. Н. Прокофьев [1935: 12]. Отдельного комментария в связи с этим заслуживает судьба рефлексов прасамодийского **kāāšā* 'человек' в селькупском, так как она подтверждает то, что селькупское *kum* является именно поздним сепаратным заимствованием из мансийского. Рефлекс общесамодийского **kāāšā* сохранился в двух оборотах [Alatalo 2004: 2202b], вероятно, утративших внутреннюю форму: *ķōňalkusä(j)* 'богач' (от *ķōňəl* 'богатый'; это прилагательное, в свою очередь, производно от *ķōň* 'правитель, князь') и *pūńākusä* 'людоед, чудовище — противник Ичи, героя-трикстера селькупского фольклора', которое в [Ibid.: 495] сопоставляется с названием мифической гигантской птицы *pūńä*, ср. также другое имя этого людоеда, бытующее на Кети: *pūńəyeldžū* [Ibid.], Пюнавельдо в русских записях — оно также доказывает составной характер имени *pūńākusä*. Слово *kāāšā* 'человек', таким образом, безусловно является общесамодийским; даже если бы его рефлексы не сохранились в селькупском, невозможно объяснить, как это слово могло бы распространиться во всех без исключения самодийских языках (см. [Janhunen 1977: 61], где приведен материал всех самодийских языков, за исключением селькупского), не «задев» селькупского. Наличие его рефлексов в селькупском, а также

характер этих рефлексов (реликтовое сохранение во фразеологизированных, застывших оборотах, фактически утративших внутреннюю форму) показывают, что это общесамодийское слово, во-первых, представлено и в селькупском и, во-вторых, что еще важнее, оно в селькупском старше слова *kum*, так как сохранилось только в таких контекстах, откуда новому слову не удалось его вытеснить. Это окончательно подрывает высказанное в [Rédei 1988: 168] предположение о том, что селькупское *kum* ‘человек’, мансийское *kom* ‘человек’, венгерское *him* ‘самец’ и далее, возможно, коми-зырянский этноним *kom* ‘коми’ и коми-пермяцкое слово *vižj-kum(i)* ‘родственник’ исконно родственны друг другу и представляют собой рефлексы прауральского **koj(e)-mz*: к этой прауральской форме, действительно, в конечном итоге (через праугорское **koj-ma*) восходит мансийское слово, но история прасамодийского **kādsā* ‘человек’, в том числе история рефлексов этого слова в селькупском, доказывают, что в селькупском *kum* ‘человек’ появилось позже прасамодийского слова. Соответственно, селькупское *kum* не может представлять собой реликт прауральского слова: оно может быть только поздним сепаратным заимствованием из мансийского.

2.12. В селькупском (в отличие от других самодийских языков) есть два наречия, семантика которых весьма характерна для языков так называемого остяцкого ареала (селькупского, хантыйского и кетского; есть наречия с такой семантикой и в мансийском). Это наречия *konnä* ‘с воды, с огня, от берега в лес’ [Alatalo 2004: ст. 2027] и *karä* ‘на воду, на огонь, из леса на берег’ [Ibid.: ст. 2224]. Эти наречия представлены во всех селькупских диалектах; на общесамодийском фоне обе основы являются изолятами. Кажется, что для пары этих наречий можно предложить мансийскую этимологию. При этом селькупские основы не объясняются из мансийских наречий с соответствующей семантикой, но могут быть объяснены из следующих мансийских форм:

konnä ‘с воды, с огня, от берега в лес’ — ср. манс. N *kwon* [kon], LM LU P K *kwän*, T *kun* преверб и наречие со значением ‘прочь’ [Kálmán, Munkácsi 1985];

karä ‘на воду, на огонь, из леса на берег’ — ср. манс. N *χarä* [χara], LM *kharä* ~ *khärä* ~ *khara* ~ *kharra*, K *khâr* ‘просторный, обширный, широкий, открытое пространство’ [Kálmán, Munkácsi 1985], ср. в особенности N *vitä* *χarü*, *χarü* *äs* ‘полноводная, широкая Обь’ [Ibid.].

Приведенные мансийские формы не имеют когнатов в хантыйском.

2.13. Мансийское N P *por*, LM *por* ~ *poâr*, LU *por* ~ *pâr* ~ *pâr*, K *por* (~ *poré*) ‘раз’ (в значении кратности) можно сопоставить с селькупским *pâr* с тем же значением [Alatalo 2004: ст. 619]. Селькупская долгота, вероятно, объясняется собственно селькупским удлинением в односложной основе.

2.14. Мансийское **kaləŋé* > N *χaliniš* // TJ *kalə-ŋkiš*, TČ *kälə-ŋkiš*, KU *χäləŋš*, KM *kaləŋš*, P *kalŋš*, VNK *kalŋš*, LU *kaləŋš*, LO *χal^hŋš* ‘редко’ (в мансийском это слово, вероятно, связано с **kal* > N *χal* [=χal], LM LU P *khal*, K T *khâl* // TJ *kal*, TČ *kâl*, KU *χâl*, KM *kâl*, KO *kâl*, P VN *kal*, VS *kal*, LU *kal*, LO *χal*, So *χal* ‘1. промежуток 2. середина 3. средний, находящийся в середине) может быть источником сельк. *колык/у* ‘редко’ [Казакевич, Будянская 2010: 41], *колиулоб* /ел./ ‘изредка’ [Быконя и др. 2005: 177].

2.15. Мансийское N *ērīŋ* [ērəŋ], LM *jērīŋ*, P *jērèn* ‘может быть, возможно’ сопоставимо с селькупским *irəŋ* ‘почти, едва’ [Alatalo 2004: ст. 275].

Можно указать также несколько возможных мансийско-камасинских параллелей.

2.16. Мансийское название двери реконструируется в [Honti 1982: 124] как **twz*. Нерегулярный рефлекс представлен в южном мансийском: TJ *ājīw*, TČ *ajəw*. С этой формой можно сопоставить камасинское слово *ājē*, *ājə*, *ājə* ‘дверь’ [Donner, Joki 1944: 3b], не объяснимое ни из самодийского, ни из тюркского языкового материала.

2.17. В камасинском и мансийском представлена уникальная связь между названием мизинца, предплечья и икры ноги (предплечье и икра ноги связаны между собой в самодийских и хантыйском, но без привлечения понятия ‘палец’):

- (9) МАНСИЙСКИЙ:
 а) N *sirëχ-tul'ä* ~ *sërχi-t.* [!], LM *sirχ-kattul'* ~ *kat-tul'-sirχëñ*, LU *sirkën-tul'ä*, P *sèrkëñ-tul'ë*, K *sorkëm-täl'ä* 'мизинец'
 б) LM *kat-sirχ-kattul'* 'предплечье'
 в) LM *ləl-sirχ-kattul'-pōχèl* 'икра ноги'
- (10) КАМАСИНСКИЙ:
 а) *seme* 'мизинец'
 б) *sə^Bmuia*, *sə^Bmiiōñälä* 'предплечье' (*muia* 'палец')
 в) *mui*, *māi*, *māi'ka* 'икра ноги'

2.18. Можно осторожно предположить опосредованное заимствование этнонима в камасинский из мансийского (исходным было мансийское слово, но заимствование произошло из селькупского): если допустить, что родовое название сельк. *maś* (обсуждавшееся выше в 2.10) перешло от селькупов к камасинцам (что могло случиться при переходе определенной группы селькупов на камасинский язык), из этого названия мог бы быть удачно объяснен этноним *kañmāžə* 'камасинец': 'представитель рода Мась с реки Кан' (превращение родового названия в этноним может быть проиллюстрировано названием тундровых энцев *somatu*, происходящим из нганасанского названия одного из родов тундровых энцев).

Есть еще несколько параллелей, объединяющих камасинский одновременно с мансийским и хантыйским (что не говорит обязательно о контакте камасинского с предком обскоугорских — ареальные связи могут точно так же объединять с камасинским именно мансийский, но параллель мансийским данным может обнаруживаться и в близкородственном мансийскому хантыйском).

2.19. Связь названия красного цвета с названием крови представлена в мансийском у слова *kēlp*, в хантыйском у слова *ñol* [Karjalainen 1948: 644b] и в камасинском, ср. *kəm*, *kəm* 'кровь' [Donner, Joki 1944: 29a,b] и *kēmui*, *kūmui*, *kē'm*, *kōmui*, *kūmui* 'красный' [Ibid.: 28b]. Типологически это достаточно распространенный семантический переход (см., например, <https://clics.clld.org/edges/156-946>), но в Западной Сибири языки, в которых произошел такой переход, все же образуют вполне определенный ареал (в частности, аналогичный способ выражения понятия 'красный' сформировался в лесном ненецком под влиянием хантыйского).

2.20. В мансийском, хантыйском и камасинском усматривается связь между названием осы / пчелы и словом 'вонять, пердеть', ср: мансийское N *pāñëχatpā* (редко), LM *pāñnekhatpā* ~ *pāññakhatpā*, LU *pāñnekhatpā*, P *pāñkhatpē* ~ *pāñnekhatne*, K *pāñyep* ~ *khul-pāñyep* 'оса, медуница' [Kálmán, Munkácsi 1985] и N *ponyi*, LM P *pāñni*, K *pāñyi* (inf. *pāñnoχ*), T *pōñnik-* ~ *pōñkh-* 'пердеть' [Ibid.], хантыйское *pāñəñ pus* 'оса, пчела' и *pāñəñ* 'пердун' [Karjalainen 1948: 708a], камасинское *pāñnek^{uu}* 'в', *pāñne:k^{uu}* 'в' 'пчела' [Donner, Joki 1944: 52b] и *pāñōrəl'em*, *pāñōrəl'em* 'пердеть' [Ibid.: 55a].

2.21. Есть параллели и с более широким ареальным распространением, например, представленная в мансийском, хантыйском, камасинском, маторском и селькупском полисемия 'есть' ~ 'гореть'. Прокомментирую подробнее эту изоглоссу. В селькупском в значении 'X сгорел' устойчиво используется выражение 'огонь съел X-а', вероятно, именно на основе этого выражения происходит метафорический переход, в результате которого просто глагол 'есть' используется в значении 'гореть': сельк. *qaj aqū-špa-nda-Ø nū-š sünd'ō-yod?* // что съесть-DUR-LATENT-S3 небо-GEN внутри-LOC // 'Что **горит** на небе?' Аналогично используется глагол 'есть' и в камасинском: *túram am-ō-bi* // дом-1 съесть-STRES-ПРАЕТ-S3 // 'мой дом **сгорел**', *šū am-ō-la?bā* // огонь съесть-STRES-CVB.-S3 // 'огонь **горит**'.

Таким образом, вырисовывается следующая картина. Можно выделить наиболее древний ареал, в котором взаимодействовали праугорский и прасамодийский. В дальнейшем, после обособления венгерского праязыка, продолжился контакт обско-угорского праязыка

с предками современных самодийских языков. С одной стороны, обе эти гипотезы выглядят достаточно тривиально на фоне современных представлений о взаимном расположении промежуточных праязыков уральской языковой семьи; с другой стороны, до сих пор умозрительные представления о возможности контактов между уграми и самодийцами, позже — между обскими уграми и самодийцами не подкреплены собственно лингвистическими свидетельствами, которые я постаралась представить в этой статье. Эти два этапа можно отнести к истории языковых ареалов, в которые входил мансийский язык, только с достаточной степенью условности: очевидно, что эпоха существования и угорско-самодийского, и обско-угорско-самодийского языковых ареалов предшествовала обособлению мансийского языка. Наиболее важным мне представляется вывод о непосредственном контакте мансийского и южносамодийских языков (селькупского и камасинского) в период после распада обско-угорской языковой общности. Этот вывод не следует напрямую ни из взаимного расположения мансийского и селькупского на современной лингвистической карте Западной Сибири, ни из взаимного расположения мансийского, селькупского и камасинского на лингвистической карте Западной Сибири XVII в. [Долгих 1960: карта в приложении]: в историческую эпоху манси были отделены от южных самодийцев хантами. Тем не менее представляется, что на изложенном в статье лингвистическом материале вывод о контактах манси, селькупов и камасинцев можно сделать достаточно надежно. В следующем разделе предлагается еще несколько частных лингвистических аргументов, позволяющих поддержать эту гипотезу.

3. История формирования верхнеобского лингвистического ареала

Как верхнеобской ареал я ниже буду обозначать ту территорию, на которой в настоящее время сложилась так называемая остяцкая этнокультурная общность, в которую входят ханты, селькупы и кеты, и соответствующий ей лингвистический ареал, образуемый хантыйским, селькупским и кетским языками. Сразу необходимо отметить, что ареал этот в его современном виде, судя по всему, образовался достаточно поздно, так как в нем участвуют только отдельные диалекты двух названных уральских языков. По данным селькупского словаря [Alatalo 2004], все заимствования из хантыйского концентрируются в тымском и/или северном диалектах селькупского. Соответственно, в данном случае речь идет о поздней локальной контактной ситуации, в которой со стороны хантыйского также участвуют только крайневосточные диалекты. Заимствования из кетского также концентрируются в основном в двух названных диалектах селькупского. Тем не менее есть несколько заимствований, которые можно признать общеселькупскими (по-видимому, начало контактов с кетским можно датировать периодом до диалектной дифференциации селькупского языка). В противоположность заимствованиям из восточного хантыйского, ограниченным двумя диалектами, рассмотренные выше мансийско-селькупские параллели представлены во всех диалектах селькупского. Это говорит о том, что изначально в верхнеобском ареале присутствовали мансийский, селькупский и кетский — но не хантыйский (возможно, представленный на периферии ареала). Что касается включения в верхнеобской ареал камасинского языка, в пользу этого могут говорить не только приведенные выше факты, которые показывают, что камасинский разделяет некоторые лингвистические особенности языков верхнеобского ареала. Возникает вопрос о том, где географически мог быть локализован камасинский для того, чтобы быть включенным в рассматриваемую контактную ситуацию. Ответ на этот вопрос, по крайней мере отчасти, может быть получен благодаря топонимическим исследованиям Э. Г. Беккер. В работе «Селькупские топонимы Западной Сибири» она выделяет группу топонимов южноселькупского происхождения на территории современного расселения селькупов:

«На территории Томской области, богато представленной гидронимикой селькупского происхождения, встречается ряд названий, имеющих в своем исходе: **ба, ма, ва, ча** (**джа, жа**), **чага** (...) Гидронимы с перечисленными формантами широко распространены на юге Западной Сибири, где по данным исторических актов и записям путешественников еще в XVIII веке проживал ряд самодийских племен (камасинцы, моторы, койбалы, карагасы) (...)

Часть названий на **ба, ва, ма, ча, чага**, встречающиеся на территории Томской области, мы считаем южносамодийскими, поскольку в структурном отношении они однотипны с (...) камасинскими и полностью раскрываются из камасинского языка. Гидронимы на **ча (жа, джа)** составляют замкнутый ареал в юго-западной части Томской области в бассейнах рек Чижалки, п. п. р. Васюган и Парабель, л. п. р. Обь, распространяясь островками к югу и юго-востоку за пределы Томской области, составляя опять ареал плотного залегания в бассейне р. Томь, п. п. р. Обь.

Гидронимы на **ба, ва, ма** разбросаны по всей территории, хотя основная масса сконцентрирована также по левобережью Оби» [Беккер 1965: 13–16].

Далее Беккер анализирует ряд топонимов, показывая, что они могут быть проинтерпретированы исходя из данных камасинского языка. Таким образом, ее исследование позволяет локализовать камасинцев в бассейнах Чижалки, Парабели и далее — в бассейне Томи.

Некоторые представленные в камасинском языке тюркизмы также могли появиться в результате контакта камасинцев с тюрками именно на этих территориях. Так, название соли в камасинском имеет два фонетических варианта: *t'oš, t'uš* [Donner, Joki 1944: 73b]. Вариант *t'oš* не мог быть заимствован из тюркских языков, распространенных в районе Саян, но мог быть заимствован из тюркских языков типа языка томских татар.

Необходимо также вкратце остановиться на некоторых моментах истории хантыйского языка, которые позволяют дополнить картину.

Этнографами обычно отмечаются определенные «черты сходства в культуре манси и восточных хантов (например, охота на лося и почитание его)» [Соколова 2009: 107]. То же самое можно отметить и в отношении некоторых лингвистических черт: например, комитатив и транслатив представлены только в мансийском и восточном хантыйском (причем показатели хантыйских падежей когнатны показателям мансийских). Диалекты северных и южных хантов, сейчас географически разделяющие мансийский и восточный хантыйский, стоят вне этой лингвистической и этнографической общности, что может говорить о том, что именно более поздний приход в эту географическую зону разделил составлявшие единый лингвистический ареал восточный хантыйский и мансийский. При этом мы не обязательно должны предполагать, что восточные ханты занимали современные территории своего расселения: лингвистические предки этой этнографической группы могли прийти сюда позже из более северных зон.

Вообще, если сейчас разделение между манси и хантами проходит по линии запад—восток, есть ощущение, что прежде это разделение должно было проходить по линии север—юг, причем при таком расселении манси занимали более южные территории, а ханты — более северные. О том, что мансийский язык был распространен в том числе и на более южных территориях, чем сейчас, хорошо известно (например, манси жили на Южном Урале в районе реки Чусовой). Но этот факт ничего не говорит о том, какое положение относительно освоенных манси территорий занимали территории, освоенные хантами. Свидетельством в пользу того, что манси изначально занимали более южные территории (как я предполагаю, вплоть до территорий расселения селькупов, чем и объясняются селькупско-мансийские контакты), а ханты — более северные территории, может служить история тюркизмов в обско-угорских языках. Этой теме посвящена статья А. Каннисто [Kannisto 1925]. В этой статье исследователь приводит список мансийских заимствований из тюркских языков и, при наличии параллельных заимствований в хантыйский, — также

хантыйские данные. Для мансийского он приводит 554 слова; по моим подсчетам, только к 130 мансийским словам Каннисто приводит хантыйские параллели. Сам по себе этот факт, однако, не может определенно свидетельствовать о том, что хантыйский язык был меньше вовлечен в контакты с тюркскими языками, чем мансийский: легко можно предположить, что большое количество тюркизмов в мансийском объясняется поздней сильной тюркизацией только одного из мансийских диалектов (а именно, южного). Но Каннисто приводит любопытные подсчеты, в которых он рассматривает распространение тюркизмов в мансийском диалектном континууме. К сожалению, он при этом не стратифицирует заимствования по степени древности: в подсчеты попадают как «старые» тюркизмы, так и русские слова, которые проникли в мансийский через посредство тюркских языков. Тем не менее, если провести аналогичные подсчеты для хантыйского языка по материалам, приведенным Каннисто в данной статье, можно предполагать, что одинаковый подход к материалу мансийского и хантыйского даст примерно одинаковую погрешность, обусловленную отсутствием хронологической стратификации заимствований. Тем самым эти данные все равно могут дать нам некоторое представление о распространении тюркизмов в мансийском и хантыйском диалектных континуумах (мансийский при этих подсчетах Каннисто традиционно разделяет на четыре диалектные зоны — северную, южную, восточную и западную, хантыйский я — также традиционно — в своих подсчетах разделяю на южный, северный и восточный).

Естественно, более всего нас интересуют заимствования, которые дошли до самых дальних (северных) диалектов хантыйского и мансийского (их можно относить к наиболее ранним заимствованиям). В подсчетах, приведенных ниже, в эту группу объединяются заимствования, которые представлены в северном диалекте мансийского либо хантыйского и еще хотя бы в одном диалекте. В хантыйском таких заимствований 18, в мансийском — 54. Следующая группа — заимствования, представленные в нескольких диалектах одновременно (но уже исключая северный). Для мансийского в эту группу попадают 86 заимствований, представленные одновременно в южном, восточном и западном, 34 заимствования, представленные в восточном и западном, 66 заимствований, представленные в южном и восточном и 7 заимствований, представленные в южном и западном диалектах — итого 193 заимствования. В хантыйском таких заимствований, выходящих за пределы одной диалектной группы (то есть представленных одновременно в южном и восточном хантыйском), всего 5. Подсчет тюркских заимствований, представленных только в одном диалекте, дает следующую картину: только в северном мансийском представлено 9 заимствований, только в северном хантыйском — 10 (эти сепаратные заимствования по большей части представляют собой заимствования из русского, которые могли пройти через тюркские языки, прежде чем попасть в обско-угорские), только в восточном мансийском — 73 заимствования, только в восточном хантыйском — 5, только в западном мансийском — 19 заимствований, только в южном мансийском — 206 заимствований, только в южном хантыйском — 83 заимствования. Эти подсчеты позволяют увидеть следующее. Наиболее очевидный факт состоит в том, что основной приток заимствований — это поздние заимствования из тюркских языков в те диалекты мансийского и хантыйского, которые находятся на южной периферии обско-угорского мира, на границе с тюркскими языками. Но за пределами этих двух южных групп представленность тюркизмов в хантыйском и мансийском диалектном континууме существенно различается. В мансийском существенно больше, чем в хантыйском, таких заимствований, которые имеют широкое диалектное распространение и дошли до самого дальнего от тюркских северного диалекта (54 против 18). В мансийском, кроме того, существенно больше заимствований, не дошедших до северного диалекта, но при этом пересекающих границы диалектных групп (193 против 5). Оба этих факта говорят о различной периодизации контактов двух обско-угорских языков с тюркскими: в мансийском существенно большее количество тюркизмов распространилось в период, предшествовавший диалектной дифференциации. Показательно также, что в восточном хантыйском представлено всего 5 сепаратных

тюркизмов (и всего 17 тюркизмов, если считать те, которые представлены не только в восточном, но и в южном хантыйском). Такая сравнительно слабая насыщенность восточного хантыйского тюркизмами (при том, что на современной лингвистической карте восточные хантыйские диалекты имеют протяженную южную границу с тюркскими языками) также должно свидетельствовать о сравнительно позднем проникновении восточных хантов на южные территории. Таким образом, история тюркизмов в обско-угорских языках, на мой взгляд, определенно подкрепляет гипотезу о том, что изначально граница между мансийским и хантыйским языками проходила по линии север–юг и при этом манси занимали более южные территории, хантыйский же язык вступил в контакт с тюркскими языками существенно позже.

Наконец, можно отметить, что лексика, отражающая важные культурные особенности верхнеобского ареала, по-видимому, в хантыйском и мансийском формировалась, по крайней мере отчасти, независимо.

Приведу только один пример: в системе наречий, описывающих ориентацию относительно берега / огня / воды, а также относительно течения реки, нет ни одной совпадающей между мансийским и хантыйским лексемы:

Таблица

Система наречий, описывающих ориентацию относительно берега / огня / воды, а также относительно течения реки, в мансийском и хантыйском

	мансийский	хантыйский
‘с горы вниз, к берегу, к воде, на воду, на огонь, к двери’	N LM P <i>nal</i> [nāl-], N <i>nalu</i> ~ <i>naluw</i> [nāluw], K <i>nalu</i> , T <i>nāلمي</i> [Kálmán, Munkácsi 1985]	Kr. <i>nik</i> , V <i>nik</i> , Kaz. <i>nik</i> , O <i>niχ</i> [Karjalainen 1948: 569ab]
‘на гору вверх, от берега, от воды, с воды, с огня, от двери вглубь жилища’	N <i>pā</i> ~ <i>pāŕ</i> [<i>pā</i> ~ <i>pāŕ</i>], LM LU P K <i>pō</i> , T <i>pā</i> (~ <i>pā</i>) [Kálmán, Munkácsi 1985]	DN <i>uit</i> , V <i>ūt</i> , Kaz. <i>uŕiŕ</i> , O <i>uđdi</i> [Karjalainen 1948: 267ab]
‘вверх по течению; верховья реки’	N <i>alyā</i> ~ <i>alyāl</i> ~ <i>alyäl’è</i> [<i>alyal</i> ~ <i>alya</i>], LM LU <i>ällä</i> , P <i>alle</i> [Kálmán, Munkácsi 1985]	DN <i>təj</i> , V <i>tōj</i> , Kaz. <i>tij</i> , O <i>təj</i> ‘верхнее течение реки; верхушка дерева, вершина оконечность вертикально стоящего предмета’ [Karjalainen 1948: 966ab, 967b]
‘вниз по течению; низовья реки’	N LM P K <i>lui</i> [lūji], LU <i>loi</i> ‘находящийся ниже по течению’ [Kálmán, Munkácsi 1985]; N <i>lāŕxā</i> ~ <i>lāŕxāl</i> ~ <i>lāŕxäl’è</i> [<i>lōŕxā</i> ~ <i>lōŕxal</i>], LM <i>lāŕkxā</i> ~ <i>lāŕxal</i> ~ <i>lāŕkxal</i> ~ <i>lāŕxā</i> ~ <i>lāŕxäl’è</i> ~ <i>lāŕxal</i> ~ <i>lāŕkxal</i> , LU <i>lāŕxā</i> ~ <i>lāŕxā</i> , P <i>lāŕxē</i> ~ <i>lāŕxäl</i> , K <i>lāŕxē</i> ~ <i>lāŕxäl</i> ‘вниз по течению’ [Kálmán, Munkácsi 1985]	DN <i>o uđs</i> , Kaz. <i>ōuđs</i> [Karjalainen 1948: 17b]; Tj <i>ŕim</i> , Kaz <i>ŕim</i> , O <i>ŕim</i> [Karjalainen 1948: 569a]

Итак, можно предполагать, что некоторые черты, характерные для языков верхнеобского ареала, хантыйский и мансийский не унаследовали от праобскоугорского, а приобрели независимо. Здесь можно также упомянуть развитие сибилантов, совпадающее в самодийских, хантыйском и мансийском (в меньшей степени — в венгерском). М. А. Живлов показал, что это явление имеет ареальную природу. С одной стороны, оно должно было начаться в Западной Сибири достаточно рано, так как оно захватило все самодийские языки, с другой стороны, как показал Живлов на материале иранских заимствований,

в хантыйском это явление распространилось позже, чем в мансийском, — таким образом, хантыйский, образно говоря, с опозданием воспринимает ареальные инновации в сравнении с мансийским.

Я предполагаю, что изначально верхнеобской лингвистический ареал составляли мансийский, селькупский и камасинский; на восточной его периферии происходили не слишком интенсивные контакты селькупов с кетами. Хантыйский язык был распространен севернее мансийского и занимал, возможно, северную периферию данного ареала. При этом ближе всего к мансийскому располагался предок современного восточного хантыйского. Манси соседствовали с южными самодийцами, вероятно, в районе Парабели / Чижакки, где Э. Г. Беккер выделяет в том числе и зону распространения камасинской топонимики. В период этих контактов, на мой взгляд, камасинский и селькупский уже представляли собой два отдельных языка, но при этом диалектная дифференциация селькупского еще не началась, так как те явления, которые я рассматриваю как сепаратные мансийско-селькупские ареальные изоглоссы, в равной мере характерны для всех селькупских диалектов. Впоследствии взаимное расположение хантыйского и мансийского языков на лингвистической карте Западной Сибири изменилось: в верхнеобской ареал проникают восточные ханты, затем расселение северных и южных хантов широкой полосой, протянувшейся с севера на юг вдоль течения средней Оби и Иртыша, прерывает связь манси с восточными хантами и окончательно отрезает манси от верхнеобского ареала. В этот период в верхнеобском ареале начинает формироваться современный «остяцкий» ареал, включающий селькупский, кетский и хантыйский языки. Формирование этого ареала должно быть отнесено к периоду после диалектной дифференциации входящих в него уральских языков: в этот ареал входят только крайневосточные диалекты хантыйского и только северный и тымский диалекты селькупского.

СПИСОК СОКРАЩЕНИЙ

Языки и диалекты

Авторские сокращения

венг. — венгерский язык
 диал. — диалектная форма
 камас. — камасинский язык
 койб. — койбальский язык
 л. нен. — лесной ненецкий язык
 л. эн. — лесной энецкий язык
 манс. — мансийский язык
 матор. — маторский язык
 нган. — нганасанский язык
 саам. — саамский язык

сам. — самодийские языки
 сельк. — селькупский язык
 сиб. — сибирское
 с.-сам. — северносамодийские
 т. нен. — тундровый ненецкий язык
 т. эн. — тундровый энецкий язык
 хант. — хантыйский язык
 ю.-сам. — южносамодийские
 ям. — ямальский говор тундрового ненецкого

Сокращения в цитируемых источниках

U — уральский праязык

Коми

V — вычегодский диалект коми языка

Марийский

B — бирский диалект марийского
 J — яранский диалект марийского

KB — козьмодемьянской говор горномарийского
 U — уржумский говор марийского

Мансийский

K — кондинский диалект мансийского
 KM — диалект мансийского на средней Конде
 KO — диалект мансийского на верхней Конде
 KU — диалект мансийского на нижней Конде
 LM — среднелозьвинский диалект мансийского
 LO — верхнелозьвинский диалект мансийского
 LU — нижнелозьвинский диалект мансийского
 N — северный диалект мансийского
 P — пельмский диалект мансийского
 So. — сосьвинский диалект мансийского
 T — тавдинский диалект мансийского

TĚ — тавдинский диалект мансийского (дер. Чандыри)
 TJ — тавдинский диалект мансийского (дер. Янычкова)
 VN — северновагильский диалект мансийского языка
 VNK — северновагильский диалект мансийского языка (дер. Кама)
 VNZ — северновагильский диалект мансийского языка (дер. Заозерная)
 VS — южновагильский диалект мансийского языка

Мордовский

E — эрзянский (эрзя-мордовский) язык

M — мокшанский (мокша-мордовский) язык

Хантыйский

D — демьянковский диалект хантыйского языка
 DN — диалект хантыйского на верхней Демьянке
 DT — диалект хантыйского на нижней Демьянке
 J — юганский диалект хантыйского языка
 Kam. — каминский диалект хантыйского
 Kaz. — казымский диалект хантыйского
 Ko. — кондинский диалект хантыйского языка
 Kr. — красноярский диалект хантыйского языка
 Likr. — хантыйский диалект с. Ликрисовское
 Mj. — хантыйский диалект Малого Югана

Ni. — низямский диалект хантыйского языка
 O — обдорский диалект хантыйского
 Šerk. — шеркальский говор хантыйского языка
 Trj. — тремьюганский диалект хантыйского
 Ts. — хантыйский говор села Цингалы на Иртыше
 V — ваховский диалект хантыйского
 Vart. — вартовский диалект хантыйского
 Vj. — васюганский диалект хантыйского
 VK — хантыйский говор Верхне-Калымска (на Оби между устьями Ваха и Васюгана)

Саамский

N — саамский диалект Норвегии

Селькупский

Ty — тымский диалект селькупского

Глоссы и грамматические обозначения

1, 2, 3 — 1, 2, 3 лицо	Kom. — комитатив
Abl. — аблатив	Lat., LAT — латив
Adv — наречие	LIM — интраклитика с лимитативным значением ('только')
ADV — показатели пространственных падежей (латива, локатива, элатива), используемые в послелогах и наречиях	Loc — локатив
AOR — аорист	NEG — спрягаемый отрицательный глагол
AUD — аудитив	o — объектное спряжение
CN — коннегатив	PL — множественное число
CvB — деепричастие	PN — имя собственное
Dat. — датив	PTPRAET — причастие прошедшего времени
DIM — диминутив	Px — притяжательные показатели
DU — двойственное число	r — рефлексивное спряжение
El., EL — элатив	RFL — особая форма аористного показателя, используемая в рефлексивном спряжении
EMPH — энклитика с эмфатическим значением	s — субъектное спряжение
Gen., GEN — генитив	VACOND — условное деепричастие
Plat. — иллатив	VAINF — инфинитивное деепричастие

Сокращения — литература (в цитируемых источниках)

- Ahlqv. — Ahlqvist A. *Wogulisches Wörterverzeichnis*. Helsinki: Suomalais-Ugrilainen Seura, 1891.
- D — материалы К. Доннера по селькупскому языку из архива Финно-угорского общества (Хельсинки). Kann., mitg. Liim.: MSFOu 116 — Kannisto A., Liimola M. *Wogulische Volksdichtung*. Bd. V. Helsinki: Suomalais-Ugrilainen Seura, 1959.
- Kann.—Liim.: MSFOu 101 — Kannisto A., Liimola M. *Wogulische Volksdichtung*. Bd. I. Helsinki: Suomalais-Ugrilainen Seura, 1951.
- Karj. — Vértés: MSFOu. 128 — Karjalainen K. F., Vértés E. (bearb. und hrsg.). *Grammatikalische Aufzeichnungen aus ostjakischen Mundarten*. Helsinki: Suomalais-Ugrilainen Seura, 1964.
- Liim.: MSFOu 127 — Liimola M. *Zur historischen Formenlehre des Wogulischen. I. Flexion der Nomina*. Helsinki: Suomalais-Ugrilainen Seura, 1963.
- MSz. — Szilasi M. *Vogul szójegyzék* (отдельный оттиск из *Nyelvtudományi Közlemények*, 1896, № 25.)
- OL — Karjalainen K. F. *Zur ostjakischen Lautgeschichte. I. Über den Vokalismus der ersten Silbe*. Helsinki: Suomalais-Ugrilainen Seura, 1905.
- Pr. — материалы Г. Н. и Е. Д. Прокофьевых по Erdélyi I. *Selkupisches Wörterverzeichnis: Tas-Dialekt*. Bloomington: Indiana Univ., 1970.
- Stein., OstjChr. — Steinitz W. *Ostjakische Grammatik und Chrestomatie*. 2., verarb. Aufl. Leipzig: Harrassowitz, 1950.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Аникин 2000 — Аникин А. Е. *Этимологический словарь русских диалектов Сибири. Заимствования из уральских, алтайских и палеоазиатских языков*. М.: Наука, 2000. [Anikin A. E. *Etimologicheskii slovar' russkikh dialektov Sibiri. Zaimstvovaniya iz ural'skikh, altaiskikh i paleoaziatskikh yazykov* [Etymological dictionary of Russian dialects of Siberia. Loanwords from Uralic, Altaic and Paleo-Asiatic]. Moscow: Nauka, 2000.]
- Беккер 1965 — Беккер Э. Г. *Селькупские топонимы Западной Сибири*. Автореф. дис. ... канд. филол. наук. Томск: Изд-во Томского ун-та, 1965. [Bekker E. G. *Sel'kupskie toponimy Zapadnoi Sibiri* [Selkup toponyms of Western Siberia]. Candidate diss. abstract. Tomsk: Tomsk Univ. Press, 1965.]
- Быконя и др. 2005 — Быконя В. В., Кузнецова Н. Г., Максимова Н. П. *Селькупско-русский диалектный словарь*. Томск: Изд-во Томского гос. пед. ун-та, 2005. [Bykonya V. V., Kuznetsova N. G.,

- Maksimova N. P. *Sel'kupsko-russkii dialektnyi slovar'* [Selkup-Russian dialectal dictionary]. Tomsk: State Pedagogical Tomsk Univ. Press, 2005.]
- Долгих 1960 — Долгих Б. О. Родовой и племенной состав народов Сибири в XVII в. М.: Изд-во АН СССР, 1960. [Dolgikh B. O. *Rodovoi i plemennoi sostav narodov Sibiri v XVII v.* [Clans and tribes of Siberian nations in the 16th century]. Moscow: Academy of Sciences of the USSR Press, 1960.]
- Живлов 2006 — Живлов М. А. К вопросу о реконструкции обско-угорского вокализма. *Аспекты компаративистики*. Вып. 2: *Orientalia et Classica. Труды Института восточных культур и античности*. Дыбо А. В. и др. (ред.). М.: РГГУ, 2006, 281–309. [Zhivlov M. A. Towards the reconstruction of Ob-Ugric vocalism. *Aspekty komparativistiki*. No. 2: *Orientalia et Classica. Trudy Instituta vostochnykh kul'tur i antichnosti*. Dybo A. V. et al. (eds.). Moscow: Russian State Univ. for the Humanities, 2006, 281–309.]
- Живлов 2018 — Живлов М. А. Историческая фонетика и внутренняя классификация уральских языков. Доклад на конференции «XIII традиционные чтения памяти С. А. Старостина» (Москва, НИУ ВШЭ, 2018). [Zhivlov M. A. Historical phonetics and the internal classification of Uralic. Talk at 13th Traditional Readings in memoriam S. A. Starostin (Moscow, HSE Univ., 2018).]
- Казакевич, Будянская 2010 — Казакевич О. А., Будянская Е. М. Дialectологический словарь селькупского языка (северное наречие). Екатеринбург: Баско, 2010. [Kazakevich O. A., Budyanskaya E. M. *Dialektologicheskii slovar' sel'kupskogo yazyka (severnoe narechie)* [Dialectological dictionary of Selkup (Northern variety)]. Yekaterinburg: Basko, 2010.]
- Напольских 1997 — Напольских В. В. *Введение в историческую уралистику*. Ижевск: УИИ-ЯЛ УрО РАН, 1997. [Napol'skikh V. V. *Vvedenie v istoricheskuyu uralistiku* [Introduction to the historical study of Uralic]. Izhevsk: Udmurt Institute of History, Language, and Literature, 1997.]
- Николаева 1995 — Николаева И. *Обдорский диалект хантыйского языка*. [Nikolaeva I. *Obdorskii dialekt khantyiskogo yazyka* [Obdorsk dialect of Khanty].] Hamburg: Finnisch-Ugrisches Seminar der Universität Hamburg, 1995.
- Прокофьев 1935 — Прокофьев Г. Н. *Селькупская (остяко-самоедская) грамматика*. Л.: Изд-во Ин-та народов Севера ЦИК СССР, 1935. [Prokof'ev G. N. *Sel'kupskaya (ostyako-samoedskaya) grammatika* [Selkup (Ostyak-Samoyed) grammar]. Leningrad: Institute of the Peoples of the North, 1935.]
- Соколова 2009 — Соколова З. П. *Ханты и манси: взгляд из XXI века*. М.: Наука, 2009. [Sokolova Z. P. *Khanty i mansi: vzglyad iz XXI veka* [Khanty and Mansi: A view from the 21st century]. Moscow: Nauka, 2009.]
- ССТМЯ — Цинциус В. И. (отв. ред.). *Сравнительный словарь тунгусо-маньчжурских языков. Материалы к этимологическому словарю*. Т. 1. Л.: Наука, 1975. [Tsintsius V. I. (ed.). *Sravnitel'nyi slovar' tunguso-man'chzhurskikh yazykov. Materialy k etimologicheskomu slovaryu* [Comparative dictionary of Tungusic. Materials for an etymological dictionary]. Vol. 1. Leningrad: Nauka, 1975.]
- Терешкин 1961 — Терешкин Н. И. *Очерки диалектов хантыйского языка*. Ч. 1: *Ваховский диалект*. М.; Л.: Изд-во АН СССР, 1961. [Tereshkin N. I. *Ocherki dialektov khantyiskogo yazyka* [Sketches of Khanty dialects]. Part 1: *Vakhovskii dialekt* [Vakh dialect]. Moscow; Leningrad: Academy of Sciences of the USSR Press, 1961.]
- Терещенко 1990 — Терещенко Н. М. *Ненецкий эпос. Материалы и исследования по самодийским языкам*. Л.: Наука, 1990. [Tereshchenko N. M. *Nenetskii epos. Materialy i issledovaniya po samodiiskim yazykam* [Nenets epos. Materials and studies in Samoyedic languages]. Leningrad: Nauka, 1990.]
- Урманчиева 2020 — Урманчиева А. Ю. О возможных контактах мансийского и селькупского языков (по данным этнонимии). *Вестник Томского гос. ун-та. Филология*, 2020, 65: 146–157. [Urmanchieva A. Yu. On possible contacts of Mansi and Selkup: Evidence from ethnonymy. *Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya*, 2020, 65: 146–157.]
- Хелимский 1982 — Хелимский Е. А. *Древнейшие венгеро-самодийские языковые параллели*. М.: Наука, 1982. [Helimski E. A. *Drevneishie vengero-samodiiskie yazykovye paralleli* [Earliest Hungarian-Samoyedic linguistic parallels]. Moscow: Nauka, 1982.]
- Янгасова 2001 — Янгасова Н. М. (сост.). *Ненецкие сказки и эпические песни сюдбабц", ярабц"*. Томск: Изд-во Томского ун-та, 2001. [Yngasova N. M. (comp.). *Nenetskie skazki i epicheskie pesni syudbabs", yarabs"* [Nenets tales and syudbabs" and yarabs" epic songs]. Tomsk: Tomsk Univ. Press, 2001.]
- Aikio 2002 — Aikio A. New and Old Samoyed etymologies. *Finnisch-Ugrische Forschungen*, 2002, 57: 9–57.
- Aikio 2006 — Aikio A. New and Old Samoyed Etymologies (Part 2). *Finnisch-Ugrische Forschungen*, 2006, 59: 9–34.

- Aikio 2014 — Aikio A. Studies in Uralic etymology II: Finnic etymologies. *Linguistica Uralica*, 2014, 1: 1–19.
- Alatalo 2004 — Alatalo J. *Sölkupisches Wörterbuch aus Aufzeichnungen von Kai Donner, U. T. Sirelius und Jarmo Alatalo*. Helsinki: Suomalais-Ugrilainen Seura, 2004.
- Donner, Joki 1944 — Joki A. J. *Kai Donners Kamassisches Wörterbuch nebst Sprachproben und Hauptzügen der Grammatik. Bearbeitet und herausgegeben von Aulis J. Joki*. Helsinki: Suomalais-Ugrilainen Seura, 1944.
- Häkkinen 2007 — Häkkinen J. *Kantauralin murteutuminen vokaalivastaavuuksien valossa*. Helsinki: Univ. of Helsinki, 2007.
- Häkkinen 2012 — Häkkinen J. After the protolanguage: Invisible convergence, false divergence and boundary shift. *Finnisch-Ugrische Forschungen*, 2012, 61: 7–28.
- Helimski 1997a — Helimski E. *Die Matorische Sprache. Wörterverzeichnis — Grundzüge der Grammatik — Sprachgeschichte*. Unter Mitarbeit von Beáta Nagy. Szeged: Szeged Univ., 1997.
- Helimski 1997b — Helimski E. *Mator and Northern Samoyed*. Materials for a paper read on May 16, 1997 for the Finno-Ugric Society, Helsinki.
- Honti 1982 — Honti L. *Geschichte des obugrischen Vokalismus der ersten Silbe*. Budapest: Akadémiai kiadó, 1982.
- Janhunen 1977 — Janhunen J. *Samojedischer Wortschatz: Gemeinsamojedische Etymologien*. Helsinki, 1977.
- Kálmán, Munkácsi 1985 — Kálmán B., Munkácsi B. *Wogulisches Wörterbuch*. Gesammelt von Bernát Munkácsi, geordnet, bearbeitet und herausgegeben von Béla Kálmán. Budapest: Akadémiai kiadó, 1985.
- Kannisto 1925 — Kannisto A. Die tatatrischen Lehnwörter im Wogulischen. *Finnisch-ugrische Forschungen*, 1925, XVII, 1–3: 1–264.
- Karjalainen 1948 — Karjalainen K. F. *Ostjakisches Wörterbuch*. Helsinki: Suomalais-Ugrilainen Seura, 1948.
- Korenchy 1972 — Korenchy É. *Iranische Lehnwörter in den obugrischen Sprachen*. Budapest: Akadémiai kiadó, 1972.
- Rédei 1988 — Rédei K. *Uralisches etymologisches Wörterbuch*. Bd. 1. *Uralische und finnisch-ugrische Schicht*. Budapest: Akadémiai kiadó, 1988.
- Reshetnikov, Zhivlov 2011 — Reshetnikov K., Zhivlov M. Studies in Uralic vocalism II: Reflexes of Proto-Uralic *a in Samoyed, Mansi and Permic. *Journal of Language Relationship*, 2011, 5: 96–109.
- Steinitz 1966 — Steinitz W. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin: Akademie-Verlag, 1966.

Получено / received 22.11.2021

Принято / accepted 25.01.2022

Архив В. Н. Чернецова: мансийская этнография, фольклор и лингвистическая перспектива

© 2022

Ольга Михайловна Рындина

Томский государственный университет, Томск, Россия; rynom_97@mail.tomsknet.ru

Дарья Олеговна Жорник

Институт языкознания РАН, Москва, Россия; Томский государственный университет, Томск, Россия; daria.zhornik@yandex.ru

Евгений Владимирович Барсуков

Томский государственный университет, Томск, Россия; barsukovevg@mail.ru

Софья Владимировна Покровская

Институт языкознания РАН, Москва, Россия; Томский государственный университет, Томск, Россия; sofie.v.pokrovskaya@gmail.com

Аннотация: В статье приводятся основные результаты междисциплинарного исследования, посвященного комплексному анализу этнолингвистической идентичности народа манси и объединяющего такие области научного знания, как этнография, фольклористика и лингвистика. Работа опирается на богатый полевой архив известного исследователя обских угров В. Н. Чернецова, хранящийся в Музее археологии и этнографии им. В. М. Флоринского Томского государственного университета. Помимо анализа этнографической методики проведен лингвистический разбор мансийских текстов с учетом системы записи, применявшейся Чернецовым. Отмечены сложности в интерпретации фольклорных мансийских текстов, обусловленные как объективным процессом трансформации языка, так и спецификой его фиксации в полевых условиях. В заключение приведена краткая жанрово-сюжетная характеристика мансийского фольклора, выявленная в ходе анализа архивных материалов.

Ключевые слова: документация языков, коллекции текстов, мансийский язык, персоналии, полевая лингвистика, уральские языки, фольклор, Чернецов В. Н.

Благодарности: Исследование выполнено при поддержке Российского научного фонда (проект № 19-18-00329 «Интерпретация языковой и культурной истории народа манси: этнографические, фольклорные и лингвистические материалы архива В. Н. Чернецова»). Также авторы хотели бы выразить благодарность В. А. Плунгяну за ценные замечания и комментарии к данному тексту.

Для цитирования: Рындина О. М., Жорник Д. О., Барсуков Е. В., Покровская С. В. Архив В. Н. Чернецова: мансийская этнография, фольклор и лингвистическая перспектива. *Вопросы языкознания*, 2022, 5: 35–50.

DOI: 10.31857/0373-658X.2022.5.35-50

Valery Chernetsov's archive: Mansi ethnography, folklore and linguistic perspective

Olga Ryndina

Tomsk State University, Tomsk, Russia; rynom_97@mail.tomsknet.ru

Daria Zhornik

Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia;
Tomsk State University, Tomsk, Russia; daria.zhornik@yandex.ru

Evgeny Barsukov

Tomsk State University, Tomsk, Russia; barsukovevg@mail.ru

Sofie Pokrovskaya

Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia;
Tomsk State University, Tomsk, Russia; sofie.v.pokrovskaya@gmail.com

Abstract: This article presents the main results of a multidisciplinary study aimed at a complex analysis of the ethnolinguistic identity of the Mansi people. The study encompasses ethnography, folklore studies and linguistics. It investigates the rich data gathered by Valery Chernetsov, a renowned researcher of the Ob-Ugric peoples. The archive is currently kept at the Vasily Florinsky Museum of Archaeology and Ethnography (Tomsk State University). Multiple tasks were performed within this study, such as analysis of the researcher's methodology, linguistic analysis of Mansi texts, including the transcription system employed by Valery Chernetsov. We have highlighted the difficulties that arise during the interpretation of Mansi folklore texts, which are explained both by the process of language change and by the specific conditions of gathering data in the field in the early twentieth century. In conclusion, we present a brief characteristic of Mansi folklore genres, which were discerned during the analysis of Valery Chernetsov's data.

Keywords: Chernetsov V. N., field linguistics, folklore, language documentation, Mansi, personalia, texts collections, Uralic

Acknowledgements: This research was supported by the Russian Scientific Foundation grant (project No. 19-18-00329 "Interpretation of linguistic and cultural history of the Mansi people: Ethnographic, folklore and linguistic data from Valery Chernetsov's archive"). The authors would also like to thank Vladimir Plungian for his valuable comments on this paper.

For citation: Ryndina O., Zhornik D., Barsukov E., Pokrovskaya S. Valery Chernetsov's archive: Mansi ethnography, folklore and linguistic perspective. *Voprosy Jazykoznanija*, 2022, 5: 35–50.

DOI: 10.31857/0373-658X.2022.5.35-50

Введение: фигура В. Н. Чернецова

Валерий Николаевич Чернецов (1905–1970) — этнограф, фольклорист и антрополог, ставший одной из ключевых фигур в отечественном финно-угроведении. Он родился в Москве в семье архитектора, закончил Московский радиотехнический институт по специальности радиотехника. В 1923–1925 гг. работал радиотехником в составе геодезической экспедиции на Северном Урале. По долгу службы находился в постоянных контактах с местным населением — манси, начал осваивать мансийский язык и заинтересовался культурой народа [Васильев 1970: 194]. В 1930 г. В. Н. Чернецов окончил этнографическое отделение географического факультета Ленинградского университета по специальности этнография финно-угорских народов и стал работать в Институте народов Севера и Музее антропологии и этнографии АН СССР. Научными исследованиями его деятельность не ограничивалась, в те годы он также активно участвует в разработке мансийской письменности и создании учебной и популярной литературы на мансийском языке. В 1940 г. Чернецов вернулся в Москву и поступил на работу в Институт истории материальной культуры (ныне — Институт археологии РАН). С этого времени сфера его научных изысканий сместилась в сторону археологии Северо-Западной Сибири (и особенно в область изучения наскальных изображений, где ему принадлежат выдающиеся открытия);

вплоть до своей безвременной смерти в 65 лет к этнолингвистическим проблемам манси он уже почти не обращался.

Период этнографических исследований В. Н. Чернецова был отмечен чрезвычайно интенсивной полевой деятельностью. В течение 13 лет ежегодно, с 1925 по 1938 г. (за исключением года окончания университета), Чернецов совершал этнографические экспедиции сроком от 3-х месяцев до 2-х лет. Основным объектом его полевых изысканий стала культура манси, проживавших на территории Тюменской области:

- июль — ноябрь 1925 г. — манси р. Лозьва,
- июнь 1926 г. — апрель 1927 г. — манси р. Северная Сосьва (участие в проведении Приполярной переписи),
- июль — август 1927 г. — манси р. Лозьва,
- июль 1931 — февраль 1932 г. — манси р. Северная Сосьва,
- август 1933 — март 1934 г. — манси рр. Конда, Юконда, Тап, Северная Сосьва,
- июль 1935 — январь 1936 г. — манси и ханты рр. Северная Сосьва, Средняя Обь,
- декабрь 1936 — январь 1937 г. — манси и ханты рр. Северная Сосьва, Средняя Обь,
- август — декабрь 1938 г. — манси р. Тагил.

Одна экспедиция, 1928–1929 гг., проходила на п-ве Ямал и была посвящена изучению культуры ненцев.

Работа В. Н. Чернецова приходится примерно на середину истории изучения мансийского языка. С одной стороны, у него были значимые предшественники, авторы первых научных исследований манси и собиратели текстов, такие как А. Регули, А. Альквист, Б. Мункачи, А. Каннисто. С другой стороны, во второй половине XX в. работа по изучению мансийских диалектов и фольклора тоже продолжалась достаточно интенсивно: большой вклад в развитие этой области внесла Е. И. Ромбандеева, а также А. А. Баландин и М. П. Вахрушева. В XXI в. исследования по мансийскому языку проводятся в Венгрии (К. Шипёц, М. Бакро-Надь, Б. Биро, Г. Секей), Финляндии (У.-М. Кулонен, С. Виртанен), Германии (группа под руководством Е. К. Скрибник) и, конечно же, России (в первую очередь это группа ученых Обско-угорского института прикладных исследований и разработок; двое из авторов настоящей работы также с 2017 г. занимались полевой работой с языком верхнелозьвинских манси).

1. Тематика и методика полевой работы В. Н. Чернецова

Методика полевой работы В. Н. Чернецова основывалась на методе включенного наблюдения и методе эмпатии. В противовес отстраненному процессу документации, затрудненному при работе с этносами, чья ментальность предполагает минимум раскрытия перед представителями других этнолингвистических групп, методы Чернецова базируются на сопереживании происходящим событиям, стремлении разделить ту же шкалу ценностей и оценочных категорий, что и изучаемый этнос. Данный подход заключается в стремлении к максимальному сближению с изучаемой культурой, вживанию в нее настолько, чтобы быть воспринятым народом в качестве «своего» [Рындина, Барсуков 2020: 236]. Такая методика работы позволила исследователю активно практиковать метод углубленного (глубинного) интервью, который широко используется в том числе в социологических исследованиях. В современных условиях его суть состоит в беседе исследователя с респондентом по заранее подготовленному сценарию с записью на диктофон. Во время активной работы Чернецова аудиотехнологии были еще труднодоступны, поэтому записи этнограф вел письменно.

Следуя методу эмпатии, Чернецов довел его практически до совершенства: в 1930-е гг. завершается его интеграция в мансийское сообщество, он приобретает широкую

известность как *Лозум хум* ‘человек с Лозьвы’. Авторитету исследователя способствовало и то, что ученый не раз становился посредником при урегулировании различных конфликтов, принимал участие в решении внутренних вопросов сообщества. «Мое прозвище Лозум хум прильнуло ко мне очень плотно. Сколько раз повторяется одно и то же: подходит вогул¹, улыбаясь, здоровается: “Лозум хум — манси хурип хум², как остяк”. Мой авторитет чувствуют и признают все» [Источники 1987: 148].

Исходя из задач полевой работы, первостепенное внимание Чернецов уделял регулярному изучению мансийского языка. Страницы практически всех полевых блокнотов и тетрадей 1920-х гг. отражают эти усилия, направленные прежде всего на освоение лексики (многочисленные глоссарии, в том числе и тематические, «тезаурусного» типа), но также и грамматики (основные парадигмы, перевод отдельных предложений). В 1930-е гг. Чернецов продолжал целенаправленно углубляться в область лингвистики. Так, с целью дальнейшего изучения мансийского языка в 1931 г. в с. Няксимволь им был заключен договор с Кириллом Николаевичем Сампильталовым о том, что он будет переводчиком не менее двух недель с оплатой в день 4 руб. [МАЭС 869 Д. 11: вкладыш 5]. В процессе этой работы были собраны и систематизированы сведения о морфологии и синтаксисе, фразеологии, внимание было уделено также диалектному членению мансийского языка. В записях 1933–1934 гг. приводится анализ морфем — классификация суффиксов, материалы по страдательному заложу и сослагательному наклонению, зафиксированы диалектные различия, в частности сосьвинских и сыгвигнских (ляпинских) идиомов, записаны отдельные фразы и целые сюжеты с переводом [МАЭС 869 Д. 14: 75–109]. В целом в экспедиции 1933–1934 гг. Чернецов уже демонстрирует очень высокий уровень владения мансийским языком, позволяющий ему записывать большие тексты. По мере работы значительно возрастает и скорость фиксации: «Теперь мне осталось писать уже немного, страниц 100. При наличии хорошего рассказчика это легко можно сделать в 2 дня, но сейчас почти все в лесу (ноябрь — период активной охоты. — *Авторы*), и удастся хватать лишь урывками» [Источники 1987: 183].

Со временем тематика полевых исследований претерпевает изменения. В экспедиции 1933–1934 гг. на рр. Конда, Юконда, Тап и Северная Сосьва на первый план выходит фольклор. Наиболее информативный материал о Медвежьем празднике, включающий и обрядовый фольклор, зафиксирован в экспедиции 1936–1937 гг. на р. Северная Сосьва и Обь. В. Н. Чернецов непрерывно вел записи, переезжая из одного селения в другое. В результате такой напряженной полевой работы оказались зафиксированы Медвежий праздники в четырех селениях — Ильпи-пауль, Сури-пауль, Ялп-ус и Так-як. В ходе непосредственного наблюдения велись беглые записи, в том числе и нотные, и зарисовки, которые затем расшифровывались в стационарных условиях. По возможности, прежде всего при наличии должного освещения, велась и фотосъемка. Фиксации подлежали обрядовые действия, культовая атрибутика, в том числе костюмы активных участников праздника, обрядовый фольклор. В результате был создан полевой вариант энциклопедии Медвежьего праздника — важнейшего атрибута традиционной культуры манси и обских угров в целом.

Одним из факторов, способствующих формированию лингвистической составляющей архива, стал высокий приоритет национального образования и преподавания мансийского языка в глазах исследователя, осознающего базовую роль языка в процессе функционирования и трансляции этнической культуры. Во время полевых изысканий ученый регулярно собирал материал для разработки и дальнейшего усовершенствования букваря, в составлении которого он принимал самое активное участие. При первой же возможности Чернецов посещал местные школы. Так, в школе пос. Шаим, расположенного возле р. Конда, он провел урок мансийского языка. При этом ученый констатировал

¹ Вогул (устар.) — то же, что манси.

² Букв.: ‘Лозьвинский человек, такой, как манси человек’.

минимальное представительство вогульских детей в группе, изучающей вогульский язык, но отметил поголовное желание всех школьников его изучать.

2. Состав архива В. Н. Чернецова

В ходе многолетних экспедиций был собран уникальный материал по различным областям культуры коренных народов Северо-Западной Сибири. Часть полевых материалов В. Н. Чернецова хранится в Музее антропологии и этнографии им. Петра Великого (Кунсткамере), однако большая их часть была безвозмездно передана вдовой и соратником ученого В. И. Мошинской в Музей археологии и этнографии Сибири Томского государственного университета при посредничестве Н. В. Лукиной — фонд № 869 [Лукина 1985: 71–72].

В состав этого фонда включены 157 индивидуально пронумерованных единиц хранения (дел). Их формируют разнообразные материалы, характеризующие исследовательскую деятельность В. Н. Чернецова: полевые дневники, фиксация конкретных этнографических сюжетов — «реальные записи», записи текстов на мансийском языке, рукописные и машинописные статьи, рисунки, фотографии, негативы и позитивы на стекле и пленках, киноматериалы [Там же: 72–74]. Архивные изобразительные материалы фонда включают 481 рисунок, 761 фотоснимок, более 600 негативов и позитивов на стекле и фотопленке, киноматериалы в четырех фильмоштатах.

Согласно сведениям, указанным самим Чернецовым, сборы материала по мансийскому фольклору имеют следующую территориально-временную характеристику:

1933 г. — рр. Конда и Северная Сосьва (550 с. тетрадного формата);

1934–1936 гг. — р. Северная Сосьва (230 с.);

1937 г. — рр. Северная Сосьва и Обь (140 с.);

1938 г. — р. Лозьва (100 с.).

Информантами являлись большей частью мужчины со слабым знанием или без знания русского языка, а также женщины, не владевшие русским языком. Запись сказок происходила, как правило, в бытовой обстановке, по вечерам или в лесу, а обрядовый фольклор фиксировался во время исполнения соответствующих ритуалов. Тексты на кондинском диалекте сопровождался переводом на русский язык и примечаниями, а записи, сделанные на других диалектах, не имели подобного авторского перевода [МАЭС 869 Д. 38: 230]. Характеристика фольклорных записей Чернецова на мансийском языке, а также история их изучения неоднократно становились центральной темой научных публикаций [Скрибник 1999; Лукина 2017].

В составе фонда № 869 тексты на мансийском языке сформированы в 16 дел, диапазон номеров 39–54. Определение объема архивных мансийских текстов является непростой задачей, это обусловлено прежде всего тем, что рукописные записи выполнены на листах разного формата. Источниковедческий анализ и систематизация всех дел фонда № 869, содержащих тексты на мансийском языке, позволяют сделать следующие выводы: для фиксации фольклорных данных исследователь использовал 11 общих тетрадей, 5 тонких ученических тетрадей, 3 блокнота, отдельные листы формата, близкого к А4. Общее количество страниц, содержащих непосредственно записи на мансийском языке, равняется 1546 (размер тетрадного листа, формат близкий А5). Кроме этого, в архиве имеются записи, сделанные на отдельных листах формата А4 или близкого ему, зафиксирована 41 страница такого размера. В некоторых делах содержатся листы-вкладыши, как правило, это небольшие кусочки бумаги либо фрагменты тетрадных листов с записями. В 16 делах архивного собрания, содержащих тексты на мансийском языке, отмечено наличие 63 таких листов-вкладышей.

Если дневники и полевые «реальные записи» ученого, хранящиеся в архиве Музея археологии и этнографии ТГУ, в подавляющей части опубликованы [Источники 1987],

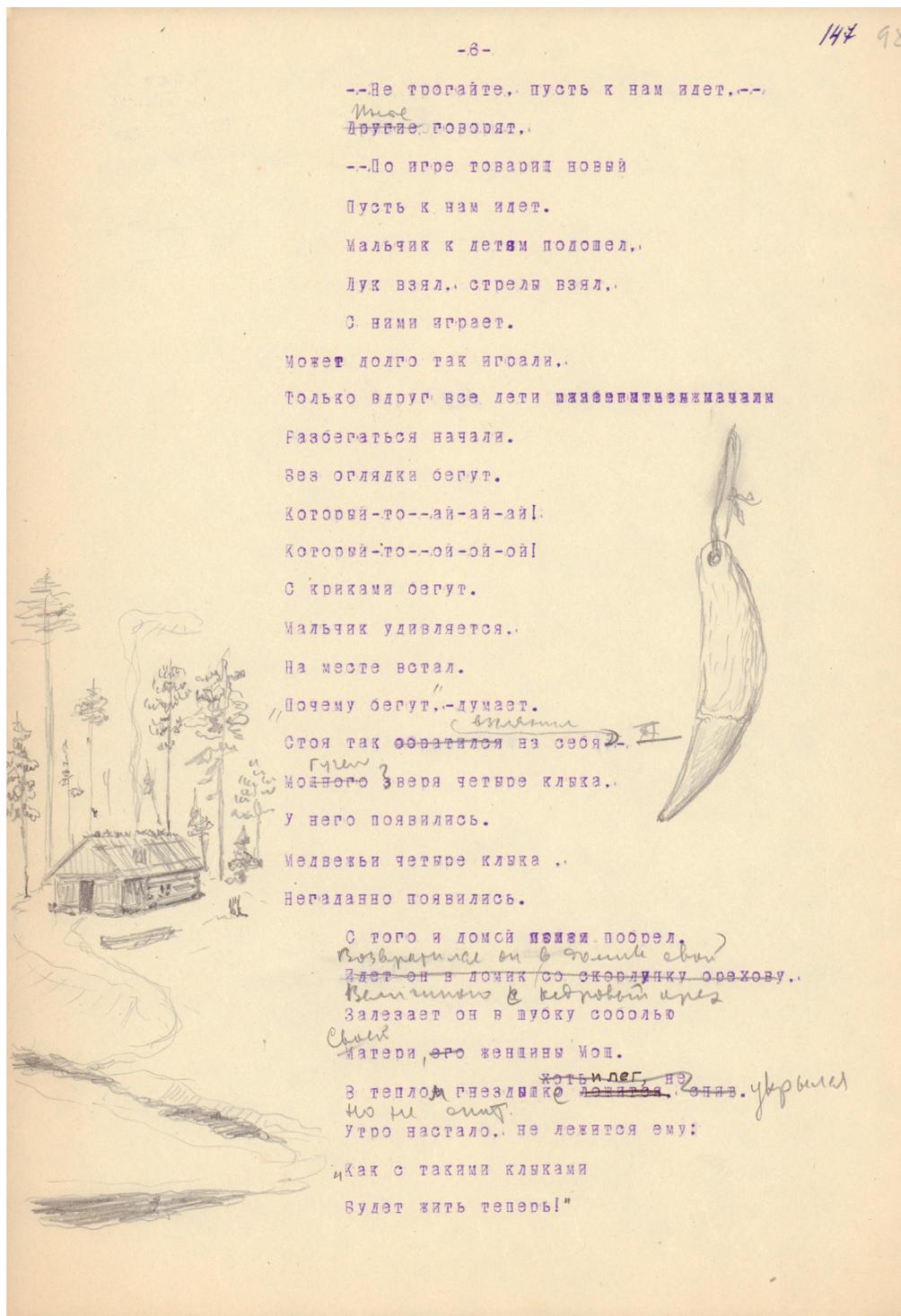


Рис. 1. Машинописный вариант фольклорного текста с рисунками В. Н. Чернецова (Д. 38. Л. 147)

то с фольклорными материалами дело обстоит иначе. Самим Чернецовым были подготовлены и изданы две публикации мансийского фольклора в переводе на русский язык (без оригинальных текстов): «Вогульские сказки» и «Про мышонка». В первом издании [Чернецов 1935] представлены разнообразные фольклорные сюжеты и жанры — мифы, сказки (волшебные, бытовые и про животных), легенды, бывальщины. Второе [Чернецов, Чернецова 1936] рассчитано на детей, соответственно, отобраны и фольклорные материалы преимущественно эвристического и дидактического характера. Второй формат публикации ученым мансийского фольклора (также в переводе на русский) состоял в привлечении фольклорных сюжетов в качестве доказательной базы в научных исследованиях.

Практически все опубликованные русскоязычные версии фольклорных мансийских текстов собраны в архиве В. Н. Чернецова в одном деле [МАЭС 869 Д. 38]. Они представлены здесь машинописными и рукописными вариантами текстов, часто с авторскими пометками, сделанными от руки чернилами или карандашом, иногда с авторскими рисунками. Неопубликованные тексты из указанного дела впервые увидели свет в отдельном научном издании из серии «Сказки и мифы народов Востока», посвященном широкой публикации повествовательного фольклора обских угров [Мифы 1990]. Здесь представлен обрядовый фольклор Медвежьего праздника [Там же: № 116–117], героический эпос [Там же: № 149], легенды [Там же: № 110–111, 153, 155, 158–159], но наиболее репрезентативны сказки — волшебные, главным образом об Эква-пырище [Там же: № 123–124; 133–134, 138–139], бытовые, прежде всего о сверхъестественных существах-менквах [Там же: 171–173, 176; 197, 200], о животных [Там же: № 192, 194].

3. Характеристика мансийских текстов из архива В. Н. Чернецова

История систематизации и подготовки к публикации текстов, записанных В. Н. Чернецовым по-мансийски и не имеющих перевода, наиболее долгая и драматичная; она уходит в прошлое столетие и насчитывает несколько десятилетий. Узнав об указанных архивных материалах, венгерская фольклористка Ева Шмидт высказала идею о возможности их публикации в Венгрии. В июне 1984 г. копии фольклорных текстов были переданы в Институт этнографии Венгерской академии наук. В 1985 г. Б. Кальман опубликовал несколько текстов из архива Чернецова, см. [Kálmán 1985]. Е. Шмидт подготовила аннотации к текстам, содержащимся в делах № 39–47, 51, 53, 54. В аннотациях были указаны следующие позиции: название текста, жанр, сюжет, информант, место и год записи, номер дела и страницы (под которыми текст зафиксирован в архиве МАЭС ТГУ), а также имелись ценные примечания о внешних признаках записи, например о расположении текста на странице, об использованной В. Н. Чернецовым транскрипции, ссылки на имеющиеся публикации. В 1999 г. в недавно открывшийся Березовский научный фольклорный фонд народа манси был передан машинописный вариант аннотаций, подготовленный этнографом Н. В. Лукиной, а в 2010 г. — и аутентичный рукописный вариант. Предполагалась публикация аннотаций, но в 2018 г. Березовский фольклорный фонд был закрыт, а в 2019 г. рукописный и машинописный варианты аннотаций Е. Шмидт были переданы в архив Обско-угорского института прикладных исследований и разработок (г. Ханты-Мансийск), где они и хранятся в настоящее время [Лукина 2019: 327–328]. Публикация аннотаций в виде каталога содержится в указанной статье и осуществлена в рамках нашего научного проекта. Одним из авторов данной статьи в 2020 г. была опубликована «Сказка о железном волке», один из текстов тетради 50 [Жорник 2020]. Некоторые тексты были опубликованы Е. К. Скрибник, см., в первую очередь, [Скрибник 1999], а также материалы сайта <https://www.babel.gwi.uni-muenchen.de/>.

В целом данный проект (начатый в 2019 г. на базе Национального исследовательского Томского государственного университета и поддержанный фондом РФФ) ставил целью

обработку и интерпретацию наименее доступной и сложной части научного наследия Чернецова, прежде всего текстов, записанных на кондинском, верхнелозьвинском, сосьвинском и обском диалектах мансийского языка в различных районах расселения этноса. Чернецов на продвинутом уровне овладел мансийским языком и самостоятельно записывал за рассказчиками длинные повествования. Записи исследователь вел в режиме реального времени, в связи с чем периодически в текстах встречаются пропуски, которые усложняют их разбор и понимание. Записи выполнены простым карандашом и, за исключением тетради 43, не сопровождаются переводом.

Тексты на мансийском языке в материалах Чернецова представлены сначала в системе записи на основе Единого северного алфавита (тетради 39–43), а затем — в смешанной записи на основе ЕСА и финно-угорской транскрипции (иначе называемой уральским фонетическим алфавитом, см. [Setälä 1901]) с определенными модификациями для каждого диалекта. Укажем некоторые примеры записи мансийских слов в такой системе в сопоставлении с современной мансийской орфографией (сосьвинский диалект):

Таблица

**Сопоставление системы записи В. Н. Чернецова
с современной мансийской орфографией и записью в IPA**

Запись В. Н. Чернецова	Современная орфография	IPA	Перевод
ri ^y	пйг	ri ^ɨ	сын
mènm ^y tas	минмыгтас	minm ^ə təs	пошел
ròtuw	рутув	rutəv	наш род
ðle't	блэгыт	ɔləjt	(они) есть, находятся

В текстах Чернецов использует общепринятые знаки препинания: точки, запятые, восклицательные и вопросительные знаки, а также тире для различения реплик персонажей в диалогах. Расстановка знаков препинания, вероятно, в целом соответствует интонационному членению мансийского текста.

Тетради, в которых производились записи, аккуратно оформлены. Каждая тетрадь озаглавлена годом и местом записи, а также номером тетради в данной поездке, например на титульном листе дела № 48 сделана надпись «Поездки на Сосьву и Обь в 1937 г. № 6». За заголовком следует содержание тетради: перечисляются тексты, включенные в данную тетрадь, с их номерами, названиями, номерами начальных страниц текста, и, в некоторых случаях, местом записи или указанием автора рассказа. Приведем в качестве примера надпись из той же тетради № 48: «tèrəl mōjt. ɔl'ein Степ. 3.», где «tèrəl mōjt» — название сказки («Сказка о желудке»), «ɔl'ein Степ.» — имя рассказчика (Степан Ольцин), а «3» — страница начала текста в тетради.

Сами тексты Чернецов записывал на правых страницах тетрадей, а левые оставлялись для комментариев к тексту (преимущественно таких, как переводы отдельных слов или выражений, или другие стилистические варианты той же самой фразы). Соответственно, нумерация осуществлялась только для правых страниц: номер располагается, как правило, в правом верхнем углу страницы тетради. В начале текста на отдельной строчке всегда дано заглавие текста, и лишь затем начинается сам рассказ. Новый текст начинается с новой страницы, а не продолжает старую.

Зачастую в текстах, по-видимому, для скорости записи, помечались многоточием повторяющиеся места, например³:

³ Все примеры с поморфемной аннотацией отгlossированы авторами данной статьи.

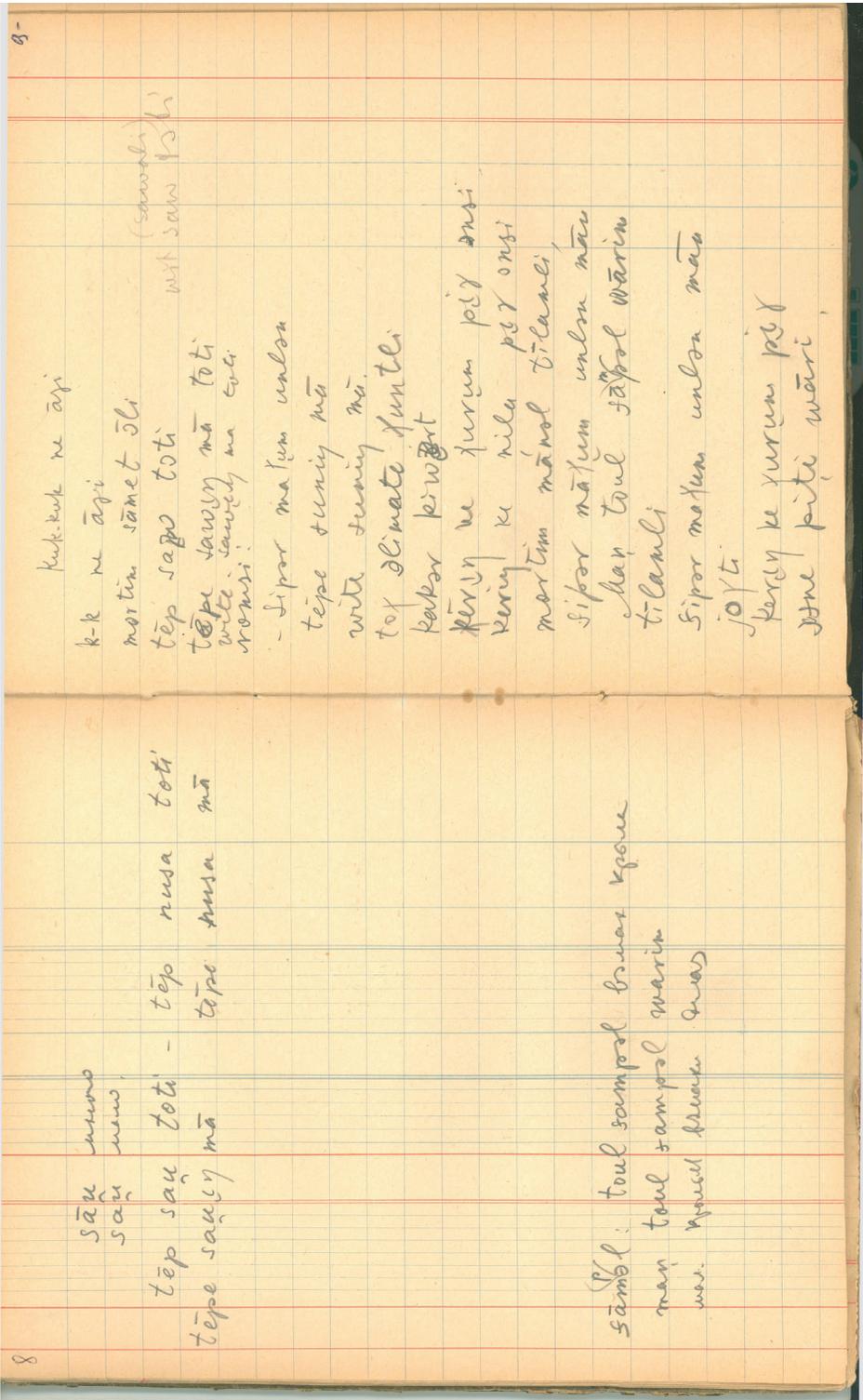


Рис. 2. Рукописный вариант фольклорного текста, записанного В. Н. Чернецовым

- (1) $\chi\delta\tau\lambda$ min-ne sat eəs jəŋχ-i-ten,
солнце идти-PTCP.NPST семь раз кружиться-NPST-3DU
 $\chi\delta\tau\lambda$... χət-ten.
солнце идти-PTCP.NPST шесть раз кружиться-NPST-3DU

‘По солнечному кругу семь раз кружатся, по солнечному (кругу) шесть (раз круж)атся’.

Однако в некоторых случаях, когда, судя по всему, носитель рассказывал слишком быстро, возникало крайне много пропусков, из-за чего трудно восстановить цельную картину фольклорного сюжета, представленного в тексте.

В некоторых случаях можно наблюдать зачеркивания фрагментов текста, что, вероятно, происходило в случае оговорок рассказчика или же описок самого исследователя. Помимо зачеркиваний наблюдаются также переносы слов с помощью стрелочек или линий в другие места предложения. Иногда можно увидеть слова или отрывки фраз в скобках, что, предположительно, обозначает допустимость опущения находящихся в скобках слов. В других случаях скобки несут уточняющую функцию.

Для повторяющихся либо же общеизвестных персонажей фольклорных сюжетов Чернецов использует сокращения по первым буквам имени персонажа, которые записываются в верхнем регистре, например: MX = mis χum ‘Мисхум’.

Иногда почерк Чернецова не вполне разборчив из-за высокой скорости записи текстов, и его становится сложно интерпретировать, особенно с учетом некоторой непоследовательности записи одних и тех же слов, а также, в некоторых случаях, наличия омонимичных слов или форм. Непоследовательность записи в большинстве ситуаций касается символов, представляющих гласные звуки. Например, часто можно увидеть в одном слове наличие долготы в части вхождений и ее отсутствие в другой части вхождений, см. окончание поессивного показателя двойственного числа в следующем примере:

- (2) $\chi\delta sa$ t̄rm-ēn χδsa-^{ly} ɔl-i-tēn,
долгий жизнь-POSS.3DU долгий-TRANS жить-NPST-3DU
wāti t̄rm-en waty-^{ly} ɔl-i-tēn.
короткий жизнь-POSS.3DU короткий-TRANS жить-NPST-3DU

‘Долгую жизнь долго живут, короткую жизнь коротко живут’.

Также в записи одного и того же слова могут встречаться различные символы, отражающие качество гласного, будь то символы для обозначения самих гласных звуков или же диакритики для обозначения дополнительных характеристик звука. К примеру, слово ‘город’ встречается в записи ūs в одних случаях, и в записи ū̄s — в других. Следующим примером является основа глагола ‘пойти’, где вариативно встречается символ i или же e: min-as ‘пойти-PT[3SG]’, mēn-as ‘пойти-PT[3SG]’. Зачастую подобную вариативность можно наблюдать в грамматических показателях, особенно в показателях словоизменяемых категорий, к примеру, у показателя лица субъекта глагола: ealt-s-ət ‘зайти-PT-3PL’ vs. ealt-s-ot ‘зайти-PT-3PL’.

Важно отметить, что о непоследовательности речь идет в том случае, если различные записи одного слова встречаются в пределах одной тетради или группы тетрадей (по географической и диалектной принадлежности). В случае разнодиалектных тетрадей речь идет уже о диалектной вариативности.

4. Работа с мансийскими текстами из архива Чернецова в рамках проекта

В рамках нашего проекта в ходе активной работы с носителями верхнелозьвинского диалекта мансийского языка родом из д. Тресколье (в первую очередь, Верой Владимировной Анямовой, а также Татьяной Прокопьевной Бахтияровой) были составлены подстрочники

для 10 тетрадей с мансийскими текстами. Первоначально осуществлялся компьютерный набор всех текстов из тетрадей Чернецова в текстовые компьютерные файлы, а уже затем на основе файлов велась работа по переводу. Для каждой сессии отбирался один или несколько текстов из переводимой тетради, и по предложению в течение нескольких часов осуществлялся перевод на русский язык. Обсуждению подлежало развитие сюжета в данном фрагменте текста, а также неизвестные слова, малопонятные конструкции и сложные для анализа элементы. Подобные фрагменты, как правило, встречались в текстах на обском диалекте (тетради 47, 48 и 53), наиболее далеком от родного диалекта консультантов. Тексты на кондинском диалекте (тетрадь 39), относящиеся к ныне исчезнувшей восточной диалектной группе мансийского языка (или же, как считают некоторые исследователи, к самостоятельному восточномансийскому языку), были крайне сложны для восприятия нашими консультантами. По этой причине работа над кондинскими текстами велась отдельно, с помощью восточномансийских словарей [Munkácsi, Kálmán 1986; Кузакова 1994; Баландин 1958] и грамматики [Kulonen 2007].

Сложности, с которыми пришлось столкнуться в процессе перевода фольклорных мансийских текстов, обусловлены как объективным процессом трансформации языка, так и спецификой его фиксации в полевых условиях. Так, на страницах тетрадей зафиксировано внушительное количество слов, не встречающихся в синхронном срезе лексики в наши дни и не обнаруживаемых в словарях. Порой даже при смысловой прозрачности слов синтаксис оказывался настолько далек от современного северномансийского, что общий смысл предложений был крайне тяжело восстановить. Отдельной трудностью стала расшифровка всех встречающихся в текстах аббревиатур, обозначающих имена мифологических персонажей. В некоторых случаях это общеизвестные персонажи, либо же персонажи, чьи имена сначала фигурируют в тексте полно и лишь затем превращаются в аббревиатуры. Иногда, однако, аббревиатура вводится сразу, и произвести идентификацию не удается даже с помощью носителя. Серьезную сложность представляли фрагменты текста со множеством исправлений: зачеркиваний, переносов или попыток написать один символ поверх другого. Поскольку все записи велись простым карандашом, то не всегда представлялось возможным установить приоритет одной стрелки над другой, последовательность зачеркиваний и переносов, что могло обусловить неточность предложенных интерпретаций. То же касается случаев, когда почерк становится неразборчивым. Неоднозначность толкований порождали и ситуации наличия текста в скобках: то, что рассказчик произнес, а затем попросил опустить; то, что рассказчик не произносил, но затем попросил дополнить; то, что Чернецов считал нужным дополнить за рассказчиком или другие варианты.

Следующим этапом обработки переведенных мансийских текстов становилось глоссирование с использованием общепринятых типологических глосс. На начальном этапе осуществлялось ручное глоссирование, затем была реализована специальная программа по автоматическому глоссированию мансийских текстов с ручной коррекцией. В исследовательской перспективе — сопоставление фонетических, лексических и грамматических особенностей различных диалектов.

На данном этапе обработки мансийских текстов из архива В. Н. Чернецова наиболее ярко заметными оказываются лексические особенности этих материалов. Напомним, что целиком записанные тексты исследователь не переводил (кроме тетради 43). Тем не менее на полях тетрадей иногда встречаются пометы, прямо относящиеся к лексике, например: *tēpəl* — *sory* (второе слово представляет собой общеизвестное обозначение желудка, первое слово, вероятно, с тем же значением, не встречается в словарях), *an* — *akw* (второе слово обозначает собой более распространенный вариант слова «один», чем первое). Обычно Чернецов выносил на поля редкие слова либо диалектные варианты более известных слов⁴. Некоторые из них незнакомы современным носителям мансийского языка и/

⁴ Большая часть записей на полях в тетрадях Чернецова не относится к лексическим пометам и представляет собой разъяснения к предложениям текста: это может быть полный вариант фразы,

или не встречаются словарях. Многие слова расшифровать так и не удалось ни с помощью словарей, ни с помощью носителей, например: *wes*, *eṗt*- (основа глагола), *tāry*- (основа глагола), *siwr*. Многие из этих лексических единиц встретились в текстах на обском диалекте, поэтому требуется консультация носителей этого диалекта. Также непрозрачны имена некоторых персонажей, например *ont wöltan*.

5. Особенности фонетики мансийского языка через призму материалов В. Н. Чернецова

Отличительной особенностью текстов архива является то, что они приводятся в авторской транскрипции. Таким образом, рукописи содержат материалы, подходящие не только для изучения грамматических и лексических аспектов языка, но и для анализа мансийской фонетики. Наибольшая динамика феноменов современного северного мансийского языка наблюдается именно в данной области, поэтому рассмотрим содержательную сторону архива с лингвистической стороны именно на ее примере.

Транскрипционная система, используемая В. Н. Чернецовым, претерпевает изменения в ходе полевой работы, что объяснимо процессом погружения в язык. Так, например, для обозначения фонемы <β> в более ранних тетрадах наблюдается вариативность между губно-губным [w] и губно-зубным [v] (действительно, в наши дни из-за влияния языка экспансии в северном мансийском фигурируют оба этих аллофона, но довольно сложно предположить их возникновение на том этапе), в то время как более поздние записи содержат обозначения, соответствующие губно-губному аппроксиманту [w] / [u]. Данный переход свидетельствует об адаптации фонетического слуха русскоязычного исследователя к мансийской речи и развитию способности идентифицировать звуки, не ассоциируя их с фонемами русского языка. Соответственно, спорные моменты между фонетическими данными разрешаются в пользу более поздних тетрадей в те моменты, когда это не объяснимо диалектными различиями. Общим для всех записей является использование диакритических знаков, надстрочных буквенных символов и расстановка пунктуации, состав используемых знаков и графем изменчив.

Диакритические знаки используются Чернецовым для обозначения мягкости согласных, в качестве комплексного обозначения аппроксимантов, для различения количества и качества гласных. В силу широкого диапазона и меняющейся стратегии использования графических средств можно не только наблюдать эволюцию фонетического слуха исследователя, но и косвенно верифицировать акустические феномены. Так, непоследовательным на первый взгляд выглядит использование диакритик для обозначения долготы / краткости гласных: для долгих фонем <a: >, <o: >, <e: >, <i: > чаще всего используется знак макрона, в то время как долгая фонема <u: > регулярно обозначается соответствующей графемой без диакритики. Для обозначения же краткой фонемы <u > в первом слоге используется кратка (y-образный надстрочный знак), что совпадает с результатами инструментальных исследований наших дней, обнаруживающих не краткий гласный [u], а непосредственно редуцированный гласный [ə] на месте ожидаемой краткой фонемы <u > в первых слогах.

При ближайшем рассмотрении знаки надстрочного индекса в системе Чернецова передают те звукотипы, которые, очевидно, подвергаются ассимилятивным процессам. Безусловно, позиционная вариативность и преобразование фонем с течением времени — связанные явления, поэтому столь подробно транскрибированные тексты прошлого века являются важным источником данных. Наибольшей иллюстративностью обладают следующие примеры:

записанной в самом тексте кратко, либо же другой способ сформулировать смысл, переданный в данном фрагменте текста.

архива можно выделить мифы, героический эпос, сказки, легенды и предания и демонологические рассказы с доминированием фабулат. Мифы включают в себя классические сюжеты о первичном мироустройстве и обрядовый фольклор, воспроизводимый главным образом на Медвежьем празднике в виде песен о деянии духов высшего ранга, обитающих в Верхнем мире, о спуске на землю медведя, сына верховного бога Торума. Героический эпос воспевают деяния богатырей, которые точно локализованы благодаря привязке к конкретным географическим реалиям — рекам или населенным пунктам: герои из селений Сатыги, Леуши, богатыри с р. Воря, верховий р. Казым, устья р. Кос-Я. Сказки представлены во всем своем многообразии жанров — волшебные, бытовые и о животных. Среди волшебных доминируют сюжеты, связанные с путешествиями Мир-сусне-хума, еще одного сына Торума, выступающего в роли культурного героя, по разным мирам, и с Эква-пырисем (эманацией Мир-сусне-хума). Среди многообразия последних популярны «Эква-пырись и шаман», «Эква-пырись и русский богатырь», «Эква-пырись обманщик и вор». Еще одна сюжетная линия волшебных сказок задана деяниями прародителей фратрий Пор и Мось. Бытовые сказки отличает социальная направленность сюжета — о глупом зяте, умном старике, а также переложение дидактических функций сказочных персонажей на материальные предметы, например зоб, стельку. Животный мир в сказках выполняет познавательно-назидательную функцию в межпоколенческой коммуникации и представлен мышатами, лисятами, широким орнитоморфным набором. Зафиксированные легенды имеют характер этнологических и повествуют о происхождении природных (светила, скалы, реки) и культурных реалий (серебряные вещи). Сюжеты фабулат связаны преимущественно со встречами со злыми лесными духами — менквами и с попаданием в загробный мир.

При этом следует иметь в виду условность жанровой классификации. Так, граница между мифами и героическим эпосом в обско-угорском фольклоре порой весьма подвижна в силу сильной тенденции к «героизации мифологических персонажей», а условность деления на прозаические и поэтические жанры объясняется тем, что один и тот же сюжет может быть исполнен как в поэтической, так и прозаической форме [Мифы 1990: 32–34]. Следует учитывать и специфику мифопоэтического мышления, в лоне которого зародились и функционируют многие фольклорные сюжеты: потусторонний, «заглазный» мир мыслится в нем как реально существующий, а не фантастический, поэтому и деление реальности на естественную и сверхъестественную, что играет важную роль при характеристике жанровых признаков, лишается своей принципиальности.

Заключение

Архив В. Н. Чернецова представляет собой объемное собрание материалов по культуре и языку различных групп манси. Эти данные могут быть использованы для исследований в различных областях науки: фольклористике, этнографии, лингвистике.

Уникальность собранных исследователем фольклорных материалов обусловлена системными трансформациями самой мансийской культуры. Фольклор как единая система, построенная на традиционном мировоззрении и диалектах мансийского языка, начал разрушаться, и самое начало этого процесса зафиксировали сборы В. Н. Чернецова. С одной стороны, ученый продолжил линию на фиксацию обрядового фольклора, заложенную его зарубежными предшественниками — финскими и венгерскими фольклористами. В научном плане этот пласт фольклорного материала трудно переоценить: без него невозможно решение важнейших вопросов этнокультурной истории манси. С другой стороны, Чернецов был одним из первых исследователей, кто стал активно собирать малые прозаические жанры фольклора — демонологические рассказы, легенды и предания, а также бытовые сказки и сказки о животных.

Большую ценность представляет архив и для лингвистов — в частности, за счет полноты охвата диалектов мансийского языка: в архиве представлены бесценные записи многих вымерших на сегодняшний день идиомов, в том числе и полностью исчезнувший восточномансийский.

В данной статье в качестве предварительных результатов исследований мы привели наиболее характерные примеры из области фонетики и лексики. После включения текстов из архива Чернецова в электронный корпус появится возможность подробного анализа и других уровней мансийского языка. Методичная фиксация вариантов мансийского языка дает достаточно полное представление о наблюдавшихся в указанные годы языковых процессах в сопоставлении с современными данными. Высокая степень подробности и точность транскрипционной записи исследователя позволяет не только сопоставлять диалекты мансийского языка, но и наблюдать динамику развития: используемые транскрипционные приемы очевидно фиксируют как субъективное слуховое восприятие, так и объективные феномены, частично отражая их анализ. Записи, сделанные в режиме реального времени, также представляют уникальный материал для потенциального анализа дискурсивных явлений — на письме отображаются паузы, оговорки, повторения. Все это зачастую опускается в письменной передаче устной речи, вследствие чего многие дискурсивные особенности стираются. Теперь, когда трудоемкий этап технической работы по компьютерному набору, переводу и глоссированию текстов близится к завершению, у лингвистов появится возможность изучать различные явления мансийского языка в 20-е и 30-е гг. XX в. в сопоставлении как с более ранними записями (в первую очередь, А. Каннисто и А. Регули), так и с современными полевыми данными.

СПИСОК СОКРАЩЕНИЙ

3 — третье лицо	POSS — посессивный показатель
DU — двойственное число	PST — прошедшее время
EMPH — эмфатическая частица	PTCP — причастие
NPST — непрошедшее время	S — субъектное спряжение
O — объектное спряжение	SG — единственное число
PL — множественное число	TRANS — транслатив

СПИСОК ИСТОЧНИКОВ

- Баландин 1958 — Баландин А. Н., Вахрушева М. П. *Мансийско-русский словарь с лексическими параллелями из южно-мансийского (кондинского) диалекта*. Л.: Учпедгиз, 1958.
- Жорник 2020 — Жорник Д. О. Мансийская сказка о железном волке. *Родной язык. Лингвистический журнал*, 2020, 2: 151–201.
- Кузакова 1994 — Кузакова Е. А. *Словарь манси: восточный диалект*. М.: Российская академия наук, Институт этнологии и антропологии имени Миклухо-Маклая, 1994.
- Мифы 1990 — *Мифы, предания, сказки хантов и манси* / Сост., предисл. и примеч. Н. В. Лукиной. М.: Наука, 1990.
- Чернецов 1935 — Чернецов В. Н. *Возульские сказки. Сборник фольклора народа манси (вогулов)*. Л.: Гослитиздат, 1935.
- Чернецов, Чернецова 1936 — Чернецов В. Н., Чернецова И. Я. *Про мышонка. Сказки и песни народа манси (вогулов)*. Л., 1936.
- Чернецов 1937 — Чернецов В. Н., Чернецова И. Я. *Краткий мансийско-русский словарь*. М.; Л.: Учпедгиз, 1937.
- МАЭС 869 Д. 11 — *Архив В. Н. Чернецова*. МАЭС № 869. Д. 11. 35 л., 7 вкладышей.
- МАЭС 869 Д. 14 — *Архив В. Н. Чернецова*. МАЭС № 869. Д. 14. 109 л., 4 вкладыша.
- МАЭС 869 Д. 38 — *Архив В. Н. Чернецова*. МАЭС № 869. Д. 38. 230 л.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Васильев 1970 — Васильев В. И., Симченко Ю. Б., Соколова З. П. Валерий Николаевич Чернецов. *Советская этнография*, 1970, 4: 194–197. [Vasil'ev V. I., Simchenko Yu. B., Sokolova Z. P. Valery Nikolaevich Chernetsov. *Sovetskaya etnografiya*, 1970, 4: 194–197.]
- Источники 1987 — Лукина Н. В., Рындина О. М., Марков Г. Е. (ред.). *Источники по этнографии Западной Сибири*. Томск: Изд-во Томского ун-та, 1987. [Lukina N. V., Ryndina O. M., Markov G. E. (eds.). *Istochniki po etnografii Zapadnoi Sibiri* [Sources for the ethnography of West Siberia]. Tomsk: Tomsk Univ. Press, 1987.]
- Лукина 1985 — Лукина Н. В., Рындина О. М. Этнографический архив В. Н. Чернецова (К 80-летию со дня рождения ученого). *Советская этнография*, 1985, 5: 70–74. [Lukina N. V., Ryndina O. M. Valery Chernetsov's ethnographical archive (On 80 years since the scholar's birthday). *Sovetskaya etnografiya*, 1985, 5: 70–74.]
- Лукина 2017 — Лукина Н. В. Мансийские фольклорные тексты в архивном фонде В. Н. Чернецова. *Вестник угроведения*, 2017, 2: 101–113. [Lukina N. V. Mansi folklore texts in Valery Chernetsov's archive collection. *Vestnik ugrovedeniya*, 2017, 2: 101–113.]
- Лукина 2019 — Лукина Н. В., Попова С. А. Каталог фольклорных текстов В. Н. Чернецова (по аннотациям Е. Шмидт). *Вестник Томского гос. ун-та. Культурология и искусствоведение*, 2019, 35: 327–340. [Lukina N. V., Popova S. A. Catalogue of Valery Chernetsov's folklore texts, based on E. Schmidt's annotations. *Vestnik Tomskogo gosudarstvennogo universiteta. Kul'turologiya i iskusstvovedenie*, 2019, 35: 327–340.]
- Рындина, Барсуков 2020 — Рындина О. М., Барсуков Е. В. Архив В. Н. Чернецова как творческая лаборатория исследователя (1920-е гг.). *Вестник Томского гос. ун-та. Культурология и искусствоведение*, 2020, 37: 228–238. [Ryndina O. M., Barsukov E. V. Valery Chernetsov's archive as the scholar's creative laboratory (the 1920s). *Vestnik Tomskogo gosudarstvennogo universiteta. Kul'turologiya i iskusstvovedenie*, 2020, 37: 228–238.]
- Скрибник 1999 — Скрибник Е. К. О мансийских языковых материалах из архива В. Н. Чернецова. *Языки и фольклор коренных народов Сибири*, 1999, 5: 293–302. [Skribnik E. K. On Mansi linguistic materials from Valery Chernetsov's archive. *Yazyki i fol'klor korennykh narodov Sibiri*, 1999, 5: 293–302.]
- Шмидт 1984 — Шмидт Е. Проблемы современного фольклора северных обских угров. *Искусство и фольклор народов Западной Сибири*. Лукина Н. В. (отв. ред.). Томск: Изд-во Томского ун-та, 1984, 95–108. [Shmidt E. Problems of contemporary folklore of northern Ob-Ugrians. *Iskusstvo i fol'klor narodov Zapadnoi Sibiri*. Lukina N. V. (ed.). Tomsk: Tomsk Univ. Press, 1984, 95–108.]
- Kálmán 1985 — Kálmán B. V. N. Csereyecov vogul szövegei. *Ethnographia*, 1985, 96(4): 572–577.
- Kulonen 2007 — Kulonen U.-M. *Itämansin kielioppi ja tekstejä*. Helsinki: Suomalais-Ugrilainen Seura, 2007.
- Munkácsi, Kálmán 1986 — Munkácsi B., Kálmán B. *Wogulisches Wörterbuch*. Budapest: Akadémiai kiadó, 1986.
- Setälä 1901 — Setälä E. N. Über Transskription der Finnisch-ugrischen Sprachen. *Finnisch-ugrische Forschungen*, 1901, 1: 15–52.

Получено / received 22.11.2021

Принято / accepted 24.05.2022

Case in the languages of Northern Eurasia: Preface

© 2022

Sofia A. Oskolskaya

Institute for Linguistic Studies, Russian Academy of Sciences, St. Petersburg, Russia; sonypolik@mail.ru

Падеж в языках Северной Евразии: предисловие

Софья Алексеевна Оскольская

Институт лингвистических исследований РАН, Санкт-Петербург, Россия; sonypolik@mail.ru

DOI: 10.31857/0373-658X.2022.5.51-54

The special issue “Case in the languages of Northern Eurasia” was inspired by the conference “Cases in Uralic, Altaic and Paleo-Asiatic languages”. It was held at the Institute for Linguistic Studies in St. Petersburg on 21–23 October, 2021. The aim of the conference was to survey various aspects of the case systems in the languages of Northern Eurasia.

Case is a grammatical category typical for nominal parts of speech such as nouns, pronouns, verbal forms with nominal morphology (participles and others). Cases mark syntactic and semantic relationship between constituents. Besides marking common roles (e.g., definite direct object or recipient), some cases can have functions that go beyond the default ones: case can express TAME semantics like in Estonian or it can reflect the information structure of an utterance like in Japanese. The talks presented at the conference had different goals. Some of them introduced new data from descriptive studies of case markers or case systems in the languages of Northern Eurasia: e.g., the talk “Dative case in Udihe: Semantic map” by Elena Perekhval'skaya. Some of the presentations focused on peripheral or rare functions of case markers: “Case forms functioning as temporal adverbials in Votic and Ingrian” by Fedor Rozhanskiy and Elena Markus, “Case marking in (con)verbal morphology in Nivkh” by Ekaterina Gruzdeva and others. A number of talks were dedicated to the evolution of case systems. They comprise both historical comparative studies (“On the origin of new allative in Ulcha” by Alexander Pevnov and other talks) and observations of dialectal variation (“Terminative, abessive and comitative cases in Finnic languages and dialects of Ingria” by Mehmet Muslimov and other talks).

All the talks were united by the area of Northern Eurasia. Although the languages under discussion have different genetic affiliation (most of them belong to the Uralic, Tungusic, Turkic, Mongolic, Chukotko-Kamchatkan families), they are likely to share some linguistic features due to intensive language contacts, see [Nichols 1992]. Therefore, the conference became a workshop where scholars could exchange data and interpretations on case markers and systems in these languages.

This issue includes five papers proceeding conference talks that provide new data interesting for general linguistic theory.

The paper by Juha Janhunen “On the local case markers in Mongolic” opens the issue. It is the tenth historical comparative work in the series “Issues of comparative Uralic and Altaic

Studies” published by the author. Janhunen provides a robust historical analysis of the Mongolic dative-locative markers **-DUR* and **-D-A-*, comparing them with the Turkic locative marker **-DA* and Tungusic dative-locative marker **-dOO*. He argues that all these markers are not genetically connected despite their superficial similarity. The Mongolic marker **-DUR* was derived by suffixalizing the spatial **dota-r(-a)* ‘inside’. The Mongolic suffix **-D-A-* is a secondary form combined from a preceding spatial formative **-d-* and the monomorphemic locative marker **-A*. Tungusic **-dOO* also originates from the spatial **dOO* ‘inside’. Finally, Turkic **-DA* is a monomorphemic locative suffix that can be reconstructed as **-tA*. Thus, Janhunen proves that these dative-locative suffixes have different origin and their similarity is accidental, although it is not excluded that their similarity has been supported due to areal contacts. This research provides new information significant for the Altaic hypothesis. Some scholars consider the dative-locative markers in Mongolic, Tungusic, and Turkic to be of the same origin, which is used as one of the arguments for the hypothesis that the three language groups belong to a single family, known as Altaic. The analysis by Janhunen shows that the similarity of the dative-locative markers in these languages cannot be used as an argument in favor of the Altaic hypothesis. This paper can be also helpful for general linguistics, in particular areal linguistics, providing a sample where similar features in neighboring languages could have emerged by coincidence.

The next three papers are dedicated to the cases in the Uralic languages of the Volga–Kama area. Tomi Koivunen and Riku Erkkilä tested the hypothesis proposed in [Levinson 2003] that more prototypical relations between Trajector and Landmark are usually expressed by structurally simpler forms. They investigated the variation between plain spatial cases and relational nouns expressing two spatial relations, CONTAINMENT and SUPPORT¹, in the Uralic languages of the Volga–Kama area (Erzya, Moksha, Meadow Mari, Hill Mari, Udmurt, Komi-Zyrian, and Komi-Permyak). The corpus data partially support the hypothesis only in expression of SUPPORT: non-prototypical relations tend to be expressed by relational nouns, i.e. by more complex forms. However, prototypical relations seem to be expressed equally often by plain spatial cases and relational nouns. Also, the hypothesis is not supported in expression of CONTAINMENT: there seem to be no correlation between prototypicality of the situation and simplicity of the form. Koivunen and Erkkilä suggest another parameter that can influence the choice between a plain spatial case marker and a relational noun: the latter is preferred in situation where the relation between the Trajector and the Landmark should be specially focused. Therefore, the contribution of Koivunen and Erkkilä brings new data significant for a general cognitive hypothesis. It can lead to revision of the hypothesis or to exploration of the interaction between prototypicality of situation and other factors influencing the variation of simpler and more complex forms. The latter case would not cancel the general hypothesis, but would explain how other features push the prototypicality factor out.

The next study, presented in Riku Erkkilä’s paper “How to distinguish between semantically close cases: A case study on Mordvin illative and lative”, has been developed in the same cognitive framework. The author applied a classification of Landmark types to describe the variation of two GOAL case markers, lative and illative, in Erzya and Moksha. The corpus data revealed that the illative is preferred when the Landmark entity is a “zero-dimensional” object (like ‘mushroom’), a 2D bounded entity (like ‘field’), a 3D bounded entity (like ‘barrel’) except for a dwelling or ‘home’, as well as an abstract or temporal entity, while the lative usually occurs with a Landmark which is a 3D unbounded entity (like ‘sky’), a dwelling, ‘home’, or an institution (like ‘school’).

These two studies demonstrate different applications of the same theoretical framework. In the first case (paper by Koivunen and Erkkilä), new language data can influence a general hypothesis, while in the second case (paper by Erkkilä) general linguistic theory helps to describe language data more thoroughly.

In the next paper of the issue, Polina Pleshak discusses a special status of the dative case in the Moksha language. The dative has an ambiguous position in the Moksha case system:

¹ These spatial relations are also known as IN and SUPER localizations.

morphologically, it patterns with the nominative and genitive (structural cases) and is opposed to the other (inherent) cases, while its functions are characteristic of an inherent case. Pleshak suggests a syntactic explanation that dative is a P head (in terms of some formal approaches listed in the article), like other inherent cases, but unlike them, dative has a non-relational nature and assigns genitive to its complement, which makes its morphological behavior similar to that of structural cases. This syntactic analysis corresponds also to the history of the dative case: it is a grammaticalized postposition, i.e. it developed from a non-relational P head. This is a case important for linguistic theory, which demonstrates that the grammaticalization processes may be reflected in the synchronic state of the language not so straightforward as it is usually expected to be.

The final paper of the special issue provides an exciting example of so-called “exclusively borderline case-marking”² which is attested in Ainu. Anna Bugaeva shows that Ainu displays a rather complex system of expression of case relations. There are no structural cases, and arguments of the verb remain unmarked. Some peripheral semantic relations are expressed by postpositional case particles. A large number of relations can be expressed by so-called adverbial postpositions that have developed from transitive verbs. Bugaeva argues that exclusively borderline case-marking in Ainu correlates with an extensive use of applicatives, which serve as a functional alternative to structural case markers. Moreover, she suggests a hypothesis that borderline case-marking may generally correlate with the head-marking language type.

To sum up, the five contributions of the special issue “Case in the languages of Northern Eurasia” represent studies that have been carried out in different approaches. All of them are significant for various issues of general linguistics: in areal linguistics, cognitive linguistics, grammaticalization theory and linguistic typology in particular.

I am grateful to the editors of *Voprosy Jazykoznanija* for giving an opportunity to compile this issue and for preparing it.

REFERENCES

- Iggesen 2013 — Iggesen O. A. Number of cases. *The world atlas of language structures online*. Dryer M. S., Haspelmath M. (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. <http://wals.info/chapter/49>.
- Levinson 2003 — Levinson S. C. *Space in language and cognition. Explorations in cognitive diversity*. Cambridge: Cambridge Univ. Press, 2003.
- Nichols 1992 — Nichols J. *Linguistic diversity in space and time*. Chicago: The Univ. of Chicago Press, 1992.

Получено / received 24.07.2022

Принято / accepted 01.08.2022

² The term is from the chapter of WALS on the feature “Number of cases” by Oliver A. Iggesen [2013].

Issues of comparative Uralic and Altaic studies (10): On the local case markers in Mongolic

© 2022

Juha Janhunen

University of Helsinki, Helsinki, Finland; asiemajeure@yahoo.com

Abstract: This paper discusses the diachronic status of the local case markers in Mongolic, as compared with the other “Micro-Altaic” languages. In spite of some striking superficial similarities, it is shown that the local case markers in Mongolic, Turkic, and Tungusic have separate origins, though it cannot be ruled out that they have secondarily influenced each other in the context of areal contacts. Mongolic has two secondary “dative” (dative-locative) markers, **-DUr* and **-D-A-*, of which the former is likely to be based on the spatial **dota-r(-)* ‘inside’, while the latter involves a combination of the spatial formative **-d-* with the primary “locative” (locative-dative) case marker **-A*. The suffixalization of the spatial **dota-r(-)* in Mongolic is paralleled by the analogous development of the spatial **doo* ‘inside’ in Tungusic. The formal similarity of the spatial roots **dota-* and **doo* can be accidental, but it could also reflect a material connection of either the areal or the genetic type.

Keywords: Altaic, case, dative, grammaticalization, locative, Mongolic, reconstruction, relational nouns, Tungusic, Turkic

For citation: Janhunen J. Issues of comparative Uralic and Altaic studies (10): On the local case markers in Mongolic. *Voprosy Jazykoznanija*, 2022, 5: 55–63.

DOI: 10.31857/0373-658X.2022.5.55-63

Вопросы сравнительной урало-алтаистики (10): О показателях местных падежей в монгольских языках

Юха Янхунен

Хельсинкский университет, Хельсинки, Финляндия; asiemajeure@yahoo.com

Аннотация: В статье обсуждается диахронический статус показателей местных падежей в монгольских языках в сопоставлении с прочими «микроалтайскими» языками. Мы показываем, что, несмотря на некоторые явные поверхностные сходства, показатели местных падежей в монгольских, тюркских и тунгусо-маньчжурских языках имеют различное происхождение, хотя нельзя исключать возможность вторичного взаимовлияния в ходе ареальных контактов. В монгольских языках существует два вторичных «дативных» (дативно-локативных) показателя — **-DUr* и **-D-A-*, — из которых первый, по-видимому, восходит к пространственному **dota-r(-)* ‘внутренность’, а второй — к сочетанию пространственного форматива **-d-* с первичным «локативным» (локативно-дативным) показателем **-A*. Суффиксализация пространственного **dota-r(-)* в монгольских языках аналогична развитию пространственного **doo* ‘внутренность’ в тунгусо-маньчжурских языках. Формальное сходство корней **dota-* и **doo* может быть случайным, но возможно и наличие между ними связи ареального или генетического характера.

Ключевые слова: алтайские языки, грамматикализация, датив, локатив, монгольские языки, падеж, реконструкция, реляционные имена, тунгусо-маньчжурские языки, тюркские языки

Для цитирования: Janhunen J. Issues of comparative Uralic and Altaic studies (10): On the local case markers in Mongolic. *Voprosy Jazykoznanija*, 2022, 5: 55–63.

DOI: 10.31857/0373-658X.2022.5.55-63

Introduction

The Altaic Hypothesis was originally based on the assumption of the existence of a genetic relationship between three language families: Turkic, Mongolic, and Tungusic, collectively also known as “Micro-Altaic” or “Core Altaic”. Later, in the context known as “Macro-Altaic”, Koreanic and Japonic were added to this framework. However, in spite of claims made to the contrary, the assumption of any kind of (pre)historical bond of Koreanic and Japonic with “Micro-Altaic”, or with each other, can only be supported by structural similarities, which, moreover, also encompass the Uralic languages. By contrast, the mutual connections between the three “Micro-Altaic” entities involve a significant number of actual material parallels, many of which even exhibit regular sound correspondences. While the majority of these material parallels can be shown to be due to borrowing in the context of a complex network of areal contacts, some are based on accidental convergence. The present paper takes up the question concerning the local case markers in Mongolic, as compared with Turkic and Tungusic. Like in the earlier papers (1–9) in this same series, the aim is to demonstrate the invalidity of some of the arguments conventionally quoted in favour of the Altaic Hypothesis. A critical analysis of the material parallels shared by the “Micro-Altaic” entities will ultimately show what, if anything, remains in support of a genetic relationship between any two, or all three, of these entities.

The local case markers offer one of the seemingly most obvious parallels between the “Micro-Altaic” languages. If we take the comparison to the extreme, we may pick the element transcribed as <-de>, which functions as a locative marker in, for instance, Turkish, Buryat, and Manchu, as in Turkish *ev* ‘house’: *ev-de* ‘in the house’, Buryat *müren* ‘river’: *müren-de* ‘at the river’, Manchu *ba* ‘place’: *ba-de* ‘at the place’. It is no wonder that early Altaists took such evidence at face value and assumed that the locative markers in the three “Micro-Altaic” language families descend from a common “Proto-Altaic” prototype [Ramstedt 1952: 35–38]. It was also proposed [Ibid.] that the suffix has a “cognate” in the multifunctional Korean element *-dae* [-de], as in *yeo-dae* ‘here’, which goes back to the “quasi-free noun” + *toy* ‘place (where)’ [Martin 1992: 131, 807–808, 824]. Continuing with such superficial similarities, we should not forget Japanese *-de*, which marks the locative case for dynamic events, e.g. *niwa* ‘yard’: *niwa-de* ‘in the yard’ [Kiyose 1995: 29–32]. However, Japanese *-de* (< **n-i-te*>), which also bears the instrumental function, is, like the other Japanese locative marker *-ni* (< **n-i*>), of a copular origin and very recent as a case suffix [Vovin 2003: 172–180]. What is more important from the point of view of Altaic comparisons is that the seeming similarity of the locative markers in the “Micro-Altaic” languages also turns out to be an illusion, for they all have separate origins.

Indeed, Manchu *-de* represents the neutralized reflex of the Common Tungusic suffix **-dOO* (> **-dO*), which is conventionally identified as the “dative” case marker, though it has a wide range of uses, some of which are better characterized as locative [Benzing 1956: 83]. Although reconstructible as a true suffixal case marker already for Proto-Tungusic, this element actually represents a secondarily cliticized and suffixalized variant of the spatial noun **+doo* ‘inside’, as specialists on Tungusic have noted long ago [Avrorin 1959: 172; Novikova 1960: 226]. This is also evident from the fact that the dative marker serves as a base for two complex case markers, the locative in **-dOO-lAA* (also with dative functions, and used especially after consonant stems) and the ablative in **-dOO-kl*, which contain the primary locative and ablative markers **-lAA* and **-kl*, respectively [Benzing 1956: 83–84]. The locative form **doo-laa* ‘inside’ is still synchronically attested in the modern Tungusic languages as a free-standing spatial item.

The Turkic locative suffix may for most Turkic languages, including modern Turkish, be abstracted as *-DA*, in which the initial consonant varies morphophonologically depending on the quality of the preceding segment (*D = t d*), while the vowel follows the rules of vowel harmony (*A = a e*). For Proto-Turkic, the suffix may be reconstructed in the shape **-tA*, with an invariant initial **t*, as still attested, alongside with †*-dA*, in Old Turkic, e.g. †*köl* ‘lake’: LOC †*köl-te* [Erdal 2004: 173–174]. Moreover, it originally marked not only the locative (‘where?’) but also

the ablative (‘from where?’) function (locative-ablative), which explains how it came to mark partitivity in Yakut [Räsänen 1957: 51–62]. This constitutes an important typological difference with regard to both Mongolic and Tungusic, in which the locative function is primarily associated with that of the dative (dative-locative). It happens that the Turkic suffix **-tA* has a formal analogy in the Uralic ablative suffix, conventionally reconstructed identically as **-tA*, but, again, Uralic has originally a three-way distinction between the locative, ablative, and dative functions. The Uralo-Turkic comparison is also invalidated by chronological and formal problems, in that the two protolanguages existed at very different time levels, and the Proto-Uralic form of the suffix should correctly be reconstructed as **-tI* [Janhunen 2014: 326–327]. We are, then, dealing with another superficial similarity with no consequences for linguistic genealogy.

As far as Mongolic is concerned, the situation is also more complicated than it superficially seems. The Mongolic local case markers are discussed in more detail below.

1. The Mongolic dative in **-DUr*

In the modern Mongolic languages, the case that is conventionally labelled the “dative” (dative-locative) is marked by a suffix that may be reconstructed as Proto-Mongolic **-DU*, in which the initial consonant had the variants *d* (after sonorants) and *t* (after obstruents, including **r*), while the vowel was represented by the harmonic values **u* **ü*. This form of the suffix, with essentially unchanged variation patterns, is today preserved in a few in this respect archaic Mongolic languages, notably Khammigan Mongol and Ordos. In most languages outside of the Central Mongolic group the alternation of the initial consonant has been levelled to an invariant *d* [t], while the suffix vowel has been generalized as an invariant *u* [u], which can secondarily also be represented as *i* (as in Shira Yughur and Halchighul Mongghul), *a* (as in Bonan), or *e* [ə] (as in Santa) (for the details, see the relevant chapters in [Janhunen (ed.) 2003] passim). In the innovative members of the Central Mongolic group, as well as in Daghur, all short (single) vowels in non-initial syllables have undergone reduction, with or without the loss of vowel harmony, resulting in *a* *e* in Buryat vs. *e* [ə] : Ø (zero) in Mongolian proper, Oirat, and Daghur, as in **mören* ‘river’ : DAT **mören-dü* > Buryat *müren-de* [murəndə] = Khalkha *mören-d* [mөрənt].

The dative suffix **-DU* is attested also in Late Middle Mongol sources, especially in those that reflect the actual spoken language of Western Middle Mongol [Rykin 2018: 232], but in Written Mongol, as well as in the Middle Mongol sources reflecting the Written Mongol orthographical representation, the suffix is attested as †-*DUr*, that is, †-*dUr* (Written Mongol **dur**) after sonorants and †-*tUr* (Written Mongol **tur**) after obstruents, including **r*. In Written Mongol, the suffix shape †-*DUr* has survived until the present day, though in the twentieth century it has largely been replaced by †-*DU* (Written Mongol **du** vs. **tu**) in the practical orthography of Modern Written Mongol. It may be noted that the shape †-*DUr* has not been attested in any early source reflecting the actual spoken language, which means that it had become archaic already in Middle Mongol times, and its continuing use in Written Mongol (as well as in vPhags.pa Mongol following Written Mongol conventions) was based on a tradition that traced back to the period when the orthographical principles were first created. This is, in fact, also true of the other case markers, which in the Written Mongol tradition are until the present day written with conventions reflecting a spoken language clearly more archaic than Proto-Mongolic, cf. e.g. GEN (after consonant stems) **uv** = †-*Un* vs. Proto-Mongolic **-Vin*, ACC **i** = †-*I* vs. Proto-Mongolic **-ii* (later also > *-i-g*), ABL **vca** = †-*AcA* vs. Proto-Mongolic **-A.AsA*, INSTR **bar** (after vowel stems) = †-*bAr* ~ **ijar** (after consonant stems) = †-*IyAr* vs. Proto-Mongolic **-A.Ar*.

As far as the origin of the suffix **-DUr* is concerned, it has been proposed in the past that this form was a secondary extension of the “primary” dative suffix **-DU*. The extension would have taken place by the addition of the “directive suffix” **-r*, also attested in the composition of the prolativ (prosecutive) suffix **-x-U.r* > *-g.UUr*, as in Khalkha *dee-* ‘above’ : PROL *dee-g*.

üür ‘along the top’ [Poppe 1955: 197–198, 205]. Another point of comparison could be offered by the instrumental suffix, originally **-pAr* > *-A.Ar*, which also ends in **r*. However, such explanations miss the point that **-DŪr* is clearly older than **-DU*, for which reason the latter must also formally be secondary to the former. Therefore, the only reasonable explanation is that **-DU* is an irregular truncation of **-DŪr*, conditioned by the frequent use of this case form. The dropping of the final **r* must have happened immediately before the Middle Mongol period, since the older form of the suffix entered Written Mongol, whose orthographical principles were fixed no more than, at most, a couple of hundred years before the rise of the historical Mongols.

Searching for the true origin of the suffix **-DŪr*, the possibility comes close at hand that it represents a trace of the suffixalized spatial **dotor* ‘inside’ or its locative form **dotor-a*. The process would have been the same as in the case of Tungusic, where the dative suffix can be traced back to the spatial **doo* ‘inside’. The original form of the spatial **dotor(-)* may with some confidence be reconstructed as **dota-r(-)*, based on the primary root **dota-* and the spatial formative **-r-* [Nugteren 2011: 318]. The spatial root **dota-* is also attested with other spatial formatives, including **-n-* and **-g-* > **-x-*, as in LOC **dota-n-a* : LAT **dota-g-si* : PROL **dota-x-U.r*. A parallel root variant **doca-* : LOC **doca-x-a* is present in Buryat and Khamnigan. The reduction of **+dota(-a)* to **-DŪr* in suffixal use would have been a rather natural development. In this process the suffix vowel was raised to **U*, since **O* would have been incompatible with most stem vowels in the phonotactic system of Pre-Proto-Mongolic. Also, the representation of the sequence **+dot-* as the alternating pair **d* (after sonorants) vs. **t* (after obstruents, including **r*) would have involved an automatic adaptation to the patterns already prevalent in the language (though the status of **r* as an “obstruent” remains an interesting topic that calls for an explanation).

2. The Mongolic locative in **-A*

Since the dative in **-DŪr* turns out to be a relatively recent secondary member of the Mongolic case paradigm, it is not surprising that Mongolic also has another, apparently older, case in the same function. To signal the formal difference, this case is here termed the “locative”, but it had exactly the same functional range as the “dative” (locative-dative), with both cases marking both location at (‘where?’) and movement to (‘whither?’). The locative marker may be reconstructed as **-A*. In view of its simple form, this looks like an old grammatical element which was probably once used on all nominal stems. However, by the Proto-Mongolic period its occurrences had been restricted to consonant stems. It is possible that the form of this marker became incompatible with vowel stems, or that it tended to be lost after stem-final vowels. This was probably one of the reasons why the language needed the introduction of the new secondary dative marker in **-DŪr*.

In Written Mongol and Middle Mongol, the locative in **-A* is attested after all kinds of nominal words ending in a consonant [Weiers 1969: 87–91]. Importantly, nouns ending in a diphthongoid sequence with **i* as the latter component could also take this marker, as in †*taulai* ‘hare’ : LOC †*taulai-a*, suggesting that the postvocalic final **i* could function as a consonantal glide (cf. paper No. 5 in the present series [Janhunen 2021]). Gradually, however, the use of **-A* became restricted to a small number of fixed expressions, such as †*gajar* ‘land, earth’ : LOC †*gajar-a* and †*köser* ‘earth, ground’ : LOC †*köser-e*. Traces of this usage are still preserved in Modern Mongolian, but they are obscured by the reduction and loss of final vowels, which has removed the distinction with regard to the plain basic form, as in Khalkha *xöser xebtex* ‘to lie on(to) the ground’ < **köser-e kebte-kü* (Written Mongol **guisar e gabdagu**). Occasionally, a modern reading pronunciation with a long vowel is used, as in 2 PL *taa* : LOC *ta.n-aa* (Written Mongol **tav e**) ‘to you’ (in addresses, as in letters). Many spatial forms, containing the spatial formatives **r* **n* **x* are originally locatives in **-A*. A case in point is *doter* ‘inside’ < **dota-r-a*, but

there are also several other examples, including *xoin* < **koi-n-a* ‘behind’, *gaden* < **gada-n-a* ~ *gadaa* < **gada-x-a* ‘outside’, and others.

More productively, the locative in †-*A* was widely used in Classical Written Mongol as an alternative to the dative in †-*DUr* in several quasiconverbial constructions, e.g. PTCP PRF LOC †*ire-gsen-e* (Written Mongol **viragsav e**) ~ DAT †*ire-gsen-diir* (Written Mongol **viragsav dur**) ‘after (one) had come’. The marker †-*A* is also preserved in the petrified composition of the final converb in *-*rA* = *-*r-A* ‘in order to’ and its synonym in *-*kU.i-A* > modern -*xAA* ~ -*xAyAA*, as well as in the terminative converb in *-*tAl-A* [Poppe 1955: 270–271, 278–279]. Very possibly, it is also present in the always adverbially used comitative case marker *-*IUXA* = *-*IUX-A* < *-*IUG-A* [Ramstedt 1952: 41–42], in which the initial part is identical with, and borrowed from, the Turkic proprietive suffix in -*IXg* < *-*IUG* [Erdal 1991: 139–155]. It cannot, however, be ruled out that the final vowel in Mongolic COM *-*IUXA* is a direct reflex of the former final vowel in the Turkic (or, more exactly, Pre-Proto-Bulgharic) original, later regularly lost on the Turkic side.

Finally, the locative in †-*A* has diachronically served as the base for the ablative in *-*A-cA* > Proto-Mongolic *-*A.AsA*. Although the development **c* > *s* in the suffix consonant is irregular, it may be taken for granted that the ablative element had originally the shape *-*cA*, as also reflected in the Middle Mongol sources following the Written Mongol orthographical conventions. A formal analogy to the element *-*cA* is offered by the Turkic equative in *-*cA*, as in DEM PROX *bu* : EQU *bu.n-ca* ‘like this’ [Räsänen 1957: 70–71], which also has a locative function in Khalaj [Doerfer 1971: 165], but its marginal status in the Turkic case system makes a connection with the Mongolic ablative rather unlikely (cf. also [Tenishev (ed.) 1988: 101–103]). In Mongolic, *-*cA* seems to have been the primary ablative marker, as still attested in Middle Mongol, e.g. ABL †*aka-ca* ‘elder brother’, †*morin-ca* ‘horse’ ([Weiers 1969: 101–107]; cf. also [Azılı 2021: 46] and passim), and also later in lexicalized relicts like DEM PROX †*e.xü.n-ce* (Written Mongol **vguvca**) ‘from this’ : DIST †*te.xü.n-ce* (Written Mongol **taguvca**) ‘from that’. The locative element *-*A*- may have been generalized to the ablative marker from spatialials of the type LOC *koi-n-a* : LOC ABL *koi-n-a-ca* > ABL *koi-n-aca* ‘from behind’.

In the context of Altaic comparisons, the Mongolic locative marker *-*A* has been compared with the Turkic dative marker, which also appears as *-*A* in several Turkic languages, especially in Oghuzic, as well as in the possessive declension [Tenishev (ed.) 1988: 86–91, 96–98]. The suffix form *-*A* is, however, secondary to the primary form *-*ka* (> -*ga* > -*A*), which is still synchronically dominant in the majority of the Common Turkic languages [Erdal 2004: 184]. Moreover, in view of the velar vocalism of the personal pronoun datives SG 1 *manja* (< **man-ka*) ‘to me’ : 2 *sanja* (< **san-ka*) ‘to thee’, it is reasonable to assume that the dative suffix originally had the invariably velar form *-*ka* ([Doerfer 1977], first suggested in [Ramstedt 1952: 31–32]). This may mean that it represents a relatively recently grammaticalized and suffixalized independent spatial word. The similarity of the Turkic and Mongolic elements must, therefore, be illusory. In Chuvash, the dative function is expressed by a syncretic case which also functions as the accusative, marked -(*C*)*A* [Räsänen 1957: 58–60]. Obviously, in view of the diffuse nature of the Chuvash data any comparison with Mongolic would be highly speculative.

3. The Mongolic dative in *-*D-A*

There is, however, also a third suffix expressing the dative (dative-locative) function in Mongolic. This suffix is conventionally reconstructed as *-*DA*, and because of its form it has occasionally been confused with the modern reflexes of the suffix *-*DUr* > -*DU* > -*DA* > -*D* (cf. e.g. [Poppe 1955: 199, Table]). Even so, the two suffixes are distinct and have different origins. Historically, †-*DA* is attested as a relatively rare alternative to †-*DUr* on all stem types throughout the Middle Mongol period, e.g. †*eke* ‘mother’ : DAT †*eke-de*, PL †*noya-d* ‘noblemen’ :

DAT †*noya-d-ta* [Weiers 1969: 72–76]. Later, it survived in Written Mongol in several lexicalized items, such as †*exüri* (**vguri**) ‘long time (ago)’: DAT †*exüri-de* (**vgurida**) ‘all the time’, †*nasu/n* (**nasuv**) ‘(year of) age’: DAT *nasu-da* (**nasuda**) ‘throughout one’s lifetime, always’. It is also present in the pronominal spatial PROX DAT †*e.n-de* ‘here’: DIST DAT †*e.n-de* ‘there’, from which the corresponding ablatives can also be formed, i.e. DAT ABL †*e.n-de-ce* ‘from here’: †*te.n-de-ce* ‘from there’. These pronominal ablatives, in turn, have provided the model for the complex ablative marker **-DA-cA*, as attested on regular nouns in Middle Mongol and later, e.g. †*beve-de-ce* ‘from the body’ [Rykin 2020: 66], and preserved also in modern relict forms, as in **ger-te-ce* > *ger-t.ees* ‘from home’ (cf. also [Azılı 2021: 46–47, 50–51]).

In spite of its conspicuous similarity with the Turkic locative (locative-ablative) suffix **-tA* (> *-DA*), the Mongolic suffix **-DA* is an entirely separate formation, for it is actually a complex sequence of two elements: the spatial formative **-d-* and the “locative” case marker **-A* (as discussed above). The division of the suffixal complex **-D-A* into its two components becomes obvious from the example of the spatial root **uri-* ‘front part’, which yields the spatial nominal **uri-d* ‘former, previous’ and further the spatial locative **uri-d-a* ‘before (of time)’. If the diminutive suffix **-kAn* is added, we get the basic form DIM **uri-d-kan* (Written Mongol **vuritqav**) ‘somewhat earlier’ and the locative form DIM LOC **uri-d-kan-a* (Written Mongol **vuritqav e**) ‘a little before’, in which the locative marker follows the diminutive suffix. This means that the spatial formative **-d-* is not a case marker in its own right, but a denominal derivational suffix which can be separated from the following case marker by the diminutive suffix. Two other spatial forms with the sequence **-d-A* are **derge-d-e* ‘(by the) side (of)’ and **dum-d-a* ‘(in the) middle (of)’, but the pronominal spatial forms **e.n-d-e* ‘here’ and **te.n-d-e* ‘there’ also originally belong to this type.

It remains an enigma why the complex marker **-D-A* > **-DA* was not systematically generalized to the dative (dative-locative) function, but was replaced by the innovative suffix **-DUr* > **-DU* in the regular nominal declension. The marker **-DA* remained, however, in use in two special positions: before the reflexive marker **-pAn* > **-xAn* > *-A.A/n*, e.g. **aka* ‘elder brother’: DAT RX **aka-da-xan* (Written Mongol **vaq e daqhav**) > Modern Mongolian *ax-daa/n* ‘to one’s own elder brother’, and before the marked nominative (“relativizing”) suffix **-ki*, e.g. **ger* ‘house, home’: DAT NOM **ger-te-ki* > Modern Mongolian *ger-t.e-x* ‘(the one) located in the house ~ at home’. In Written Mongol, the reflexive dative can also be formed with the suffix **-DUr*, e.g. †*aka-dur-ıyan* (**vaq e dur ijav**), with no functional difference with regard to the forms based on **-DA-*. In addition, the suffix †*-DUr* — but not †*-DA* — is in Middle Mongol occasionally attested in combination with a following third-person possessive marker **-i*, e.g. DAT PX SG 3 †*jüg-tür-i* ‘in that direction’, PTCP PRF DAT PX SG 3 †*kür.ü-gsen-dür-i* ‘when (he) had arrived’ (cf. [Rykin 2020: 65–66]).

It may be noted that the combination of the complex dative with the marked nominative suffix **-ki*, which in Mongolic yields **-D-A-ki* > DAT NOM **-DA-ki*, is formally confusingly similar to the Turkic suffixal sequence LOC NOM **-tA-ki* with an identical function. In view of its shape and function, the suffix **-ki* [Ramstedt 1952: 234] must involve a borrowing from Turkic (Turko-Bulgharic) to Mongolic, though the chronology and typology of this element still presents a number of unanswered questions (cf. [Luutonen 2011]). In this situation, one comes to wonder whether the entire suffix complex **-DA-ki* could also have been borrowed from Turkic to Mongolic. While this possibility cannot be ruled out, it is more likely that Turkic and Mongolic arrived in this case, as in many others, at a similar result out of different presuppositions. We are, therefore, most probably dealing with a case of “shared drift” between two interacting languages [Janhunen 2012].

Another instance of “shared drift” may have been present between Khitan and Jurchen. In Khitan, the dative (dative-locative) function was expressed by a suffix for which the technical readings †*-de* †*-do* †*-dü* have been established, with the vowel quality depending on the preceding stem vowel [Kane 2009: 136–138]. For the moment there is no way to know whether this suffix was a cognate of Mongolic **-D-A* or **-DUr*, though the apparently late origin of **-DUr* would

perhaps speak in favour of **-D-A*. At the same time, forms very similar to those established for Khitan are attested in Jurchen, perhaps conditioned by vowel harmony [Kiyose 1984/1991: 348–349, 354–355] and predating the later uniform Manchu suffix *-de*. As there is no reason to assume that this would be a case of suffixal borrowing between Khitan and Jurchen, it is more likely that the two languages were on the way towards developing an identical form for the same function out of their own resources, which on the Jurchen side reflected the Proto-Tungusic secondary dative suffix **-dOO*.

Conclusion

It has been shown above that the dative-locative markers in the three “Micro-Altaiic” languages have several different sources. The only original monomorphemic case markers seem to have been the Turkic locative-ablative **-tA* (later > **-DA*) and the Mongolic “locative” **-A*. Mongolic built a secondary dative-locative marker of the shape **-D-A* by combining the primary locative marker with a preceding spatial formative **-d-*. Also, Mongolic introduced a new dative-locative form in **-DUr* by suffixalizing the spatial **dota-r(-a)* ‘inside’. In a similar way, Tungusic innovated a dative-locative of the shape **-dOO* and an accompanying set of complex case forms by suffixalizing the spatial **dOO* ‘inside’ and some of its primary case forms. Secondary reductive developments led to similar forms in sets of interacting languages, including Turkic and Mongolic, on the one hand, and Khitan and Jurchen, on the other. However, the only certain case of suffixal borrowing is involved in the history of the marked nominative suffix **-ki*.

In this connection we should not ignore the fact that the two spatial roots, Mongolic **dota-* (~ **doca-*) and Tungusic **dOO(-)*, show a certain material similarity, which extends also to the Mongolic spatial formative **-d-*. This similarity can be accidental, but it is not impossible that these elements are interconnected in one or more ways. There are other spatials that show similarities between Mongolic and Tungusic, notably Mongolic **dexe-* < **dege-* ‘upper part, top’ ~ Tungusic **deg-* ‘to rise, to fly’ and perhaps Mongolic **dum-d-* ‘middle’ ~ **duli* ~ **düli* = Tungusic **duli-* ‘middle’ [Doerfer 1985: 20 Nos. 18, 20]. This is not enough to confirm a genetic connection, but it is also not typical for spatials to be borrowed. Altogether, it is exactly this type of similarities that should be considered if traces of a true genetic relationship were to be searched for between the “Micro-Altaiic” languages (cf. [Janhunen 1996]). For the time being, the material is insufficient and inconclusive.

NOTES ON THE NOTATION

While the asterisk (*) is used to denote linguistic reconstructions, based on the comparative evidence of attested spoken languages, the dagger (†) is used to denote “readings” deduced from the orthographical representation of historical languages recorded only in written form, and in non-Roman scripts. Written Mongol is transliterated according to principles of the Balk–Janhunen system (as first introduced in [Balk, Janhunen 1999]).

ABBREVIATIONS

1, 2, 3 — first, second, third person
 ABL — ablative
 COM — comitative
 DAT — dative
 DEM — demonstrative
 DIM — diminutive

DIST — distal
 EQU — equative
 GEN — genitive
 INSTR — instrumental
 LAT — lative
 LOC — locative

NOM — nominative
 PL — plural
 PRF — perfective
 PROL — prolative
 PROX — proximal

PTCP — participle
 PX — possessive suffix
 RX — reflexive suffix
 SG — singular

REFERENCES

- Avrorin 1959 — Avrorin V. A. *Grammatika nanaiskogo yazyka* [Nanai grammar]. Vol. 1: *Foneticheskoe vvedenie i morfologiya imennykh chastei rechi* [Phonetic introduction and nominal morphology]. Moscow; Leningrad: Academy of Sciences of the USSR Press, 1959.
- Azılı 2021 — Azılı K. On ablative forms in Altaic languages: An archaic compound morpheme in Mongolian +*daca*. *Festschrift in honor of Marcel Erdal*. Nevskaya I., Şirin H., Ağa F. (eds.). = *Journal of Turkish Studies*, 2021, special edition, vol. 1: 45–52.
- Balk, Janhunen 1999 — Balk M., Janhunen J. A new approach to the Romanization of Written Mongol. *Writing in the Altaic World*. Janhunen J., Rybatzki V. (eds.). Helsinki: Finnish Oriental Society, 1999, 17–27.
- Benzing 1956 — Benzing J. *Die tungusischen Sprachen: Versuch einer vergleichenden Grammatik*. Mainz: Franz Steiner, 1956.
- Doerfer 1971 — Doerfer G. *Khalaj materials*. Bloomington (IN): Indiana Univ., 1971.
- Doerfer 1977 — Doerfer G. Zu türkisch *bana* 'mir', *sana* 'dir'. *Central Asiatic Journal*, 1977, 21: 208–214.
- Doerfer 1985 — Doerfer G. *Mongolo-Tungusica*. Wiesbaden: Otto Harrassowitz, 1985.
- Erdal 1991 — Erdal M. *Old Turkic word formation: A functional approach to the lexicon*. Vols. 1–2. Wiesbaden: Otto Harrassowitz, 1991.
- Erdal 2004 — Erdal M. *A grammar of Old Turkic*. Leiden: Brill, 2004.
- Janhunen 1996 — Janhunen J. Prolegomena to a comparative analysis of Mongolic and Tungusic. *Proc. of the 38th Permanent International Altaistic Conf.* Stary G. (ed.). Wiesbaden: Harrassowitz, 1996, 209–218.
- Janhunen (ed.) 2003 — Janhunen J. (ed.). *The Mongolic languages*. London: Routledge, 2003.
- Janhunen 2012 — Janhunen J. Non-borrowed non-cognate parallels in bound morphology: Aspects of the phenomenon of shared drift with Eurasian examples. *Copies versus cognates in bound morphology*. Johanson L., Robbeets M. (eds.). Leiden: Brill, 2012, 23–46.
- Janhunen 2014 — Janhunen J. Ural-Altaic: The polygenetic origins of nominal morphology in the Trans-eurasian zone. Robbeets M., Bisang W. (eds.). *Paradigm change: In the Transeurasian languages and beyond*. Amsterdam: John Benjamins, 2014, 311–335.
- Janhunen 2021 — Janhunen J. Issues of comparative Uralic and Altaic studies (5): The status of glides in Mongolic. *Studies in Asian historical linguistics, philology and beyond. Festschrift for Alexander Vovin*. Kupchik J., Alonso de la Fuente J. S., Miyake M. (eds.). Leiden: Brill, 2021, 261–274.
- Kane 2009 — Kane D. *The Khitan language and script*. Leiden: Brill, 2009.
- Kiyose 1984/1991 — Kiyose G. N. Vowel harmony of the Tyr dialect of Jurchen in the XV century. *Japanese linguistics and Altaic linguistics*. Kiyose G. N. Tokyo: Meiji Shoin, 1991, 347–358. (First published 1984.)
- Kiyose 1995 — Kiyose G. N. *Japanese grammar: A new approach*. Kyoto: Kyoto Univ. Press, 1995.
- Luutonen 2011 — Luutonen J. *Chuvash syntactic nominalizers: On *-ki and its counterparts in Ural-Altaic languages*. Wiesbaden: Harrassowitz, 2011.
- Martin 1992 — Martin S. E. *A reference grammar of Korean: A complete guide to the grammar and history of the Korean language*. Rutland (VT); Tokyo: Charles E. Tuttle Co., 1992.
- Novikova 1960 — Novikova K. A. *Očerki dialektov evenskogo yazyka: Ol'skii govor* [Sketches of Ewen dialects: The Ola variety]. Part 1. Moscow; Leningrad: Academy of Sciences of the USSR Press, 1960.
- Nugteren 2011 — Nugteren H. *Mongolic phonology and the Qinghai-Gansu languages*. Proefschrift. Leiden: Netherlands Graduate School of Linguistics, 2011.
- Poppe 1955 — Poppe N. *Introduction to Mongolian comparative studies*. Helsinki: Suomalais-Ugrilainen Seura, 1955.
- Ramstedt 1952 — Ramstedt G. J. *Einführung in die altaische Sprachwissenschaft*. Vol. 2: *Formenlehre*. Bearbeitet und herausgegeben von Pentti Aalto. Helsinki: Suomalais-Ugrilainen Seura, 1952.
- Räsänen 1957 — Räsänen M. *Materialien zur Morphologie der türkischen Sprachen*. Helsinki: Societas Orientalis Fennica, 1957.

- Rykin 2018 — Rykin P. Noun cases in the language of the Sino-Mongol glossary *Dada Yu / Beilu Yiyu* from the late Ming period. *Journal Asiatique*, 2018, 306(2): 229–234.
- Rykin 2020 — Rykin P. O. Non-standard case forms in the “History of Altan Khan”. *Tomsk Journal of Linguistics and Anthropology*, 2020, 4: 60–72.
- Tenishev (ed.) 1988 — Tenishev E. R. (ed.). *Sravnitel'no-istoricheskaya grammatika tyurkskikh yazykov: Morfologiya* [Comparative-historical grammar of Turkic: Morphology]. Moscow: Nauka, 1988.
- Vovin 2003 — Vovin A. *A reference grammar of classical Japanese prose*. London: RoutledgeCurzon, 2003.
- Weiers 1969 — Weiers M. *Untersuchungen zu einer historischen Grammatik des Präklassischen Schriftmongolisch*. Wiesbaden: Otto Harrassowitz, 1969.

PREVIOUS PAPERS IN THE SERIES

1. Issues of comparative Uralic and Altaic studies (1): The case of Proto-Mongolic *x. *Journal of the Center for Northern Humanities*, 2017, 10 (On the occasion of the retirement of Professor Toshiro Tsumagari): 97–104.
2. Issues of comparative Uralic and Altaic studies (2): Medial *p in Pre-Proto-Mongolic. *Essays in the history of languages and linguistics. Dedicated to Marek Stachowski on the occasion of his 60th birthday*. Németh M., Podolak B., Urban M. (eds.). Kraków, 2017, 297–305.
3. Issues of comparative Uralic and Altaic studies (3): The Turkic plural in *s. *Altai Hakpo*, 2017, 27: 155–168.
4. Issues of comparative Uralic and Altaic studies (4): On the origin of the Uralic comparative marker. *The state of the art of Uralic studies: Tradition vs. innovation*. Marcantonio A. (ed.). Roma: Sapienza, Università editrice, 2018, 49–57.
5. Issues of comparative Uralic and Altaic studies (5): The status of glides in Mongolic. *Studies in Asian historical linguistics, philology and beyond. Festschrift for Alexander Vovin*. Kupchik J., Alonso de la Fuente J. S., Miyake M. (eds.). Leiden: Brill, 2021, 261–274.
6. Issues of comparative Uralic and Altaic studies (6): Uralic copulas and their analogues in other Eurasian languages. Holopainen S., Kuokkala J., Saarikivi J., Virtanen S. (eds.). *Jomas siməŋ nəkwe wōtur ēpost samən patum — Scripta miscellanea in honorem Ulla-Maija Forsberg*. Helsinki: Suomalais-Ugrilainen Seura, 2020, 129–140.
7. Issues of comparative Uralic and Altaic studies (7): On privative verbs in Mongolic and beyond. *Festschrift in honor of Marcel Erdal*. Nevskaya I., Şirin H., Ağca F. (eds.). = *Journal of Turkish Studies*, 2021, special edition, vol. 1: 153–161.
8. Issues of comparative Uralic and Altaic studies (8): On the velar nasal in Mongolic. *Festschrift in Honor of Toshiro Tsumagari's 70th Birthday*. Abashiri: Hokkaido Museum of Northern Peoples, 2021, 62–73.
9. Issues of comparative Uralic and Altaic studies (9): Medial intervocalic *k and *g in Mongolic. *Historical linguistics and philology of Central Asia: Essays in Turkic and Mongolic studies*. Khabtagaeva B. (ed.), with the assistance of Zs. Olach. Leiden: Brill, 2021, 280–297.

Получено / received 02.02.2022

Принято / accepted 29.03.2022

The interaction between spatial cases and relational nouns in the Uralic languages of the Volga–Kama area

© 2022

Tomi Koivunen

University of Turku, Turku, Finland; Udmurt State University, Izhevsk, Russia; tokrkoj@utu.fi

Riku Erkkilä

Ludwig-Maximilians-Universität München, Munich, Germany; University of Helsinki, Helsinki, Finland; riku.erkkila@helsinki.fi

Abstract: In this paper we take a look at the variation between plain spatial cases and relational nouns in expressions of two spatial relations, namely CONTAINMENT and SUPPORT, in the Uralic languages of the Volga–Kama area (Erzya, Moksha, Meadow Mari, Hill Mari, Udmurt, Komi-Zyrian, and Komi-Permyak). By studying the variation in the coding of these relations, we tested the proposition put forward in literature that structurally more simple forms are usually used to code semantically more simple relations. In the most prototypical case of CONTAINMENT there is a three-dimensional landmark in which the trajector is situated, and in the most prototypical case of SUPPORT there is a two-dimensional landmark which shields the trajector from the effects of gravity. To test the proposition, we gave the landmarks in our data prototypicality scores based on how well they conform to the spatial prototypes of CONTAINMENT and SUPPORT. Next, we looked at what means (plain case or relational noun) was used to code spatial relations with different prototypicality values. The results show that in the studied languages the proposed generalization of simpler meaning coded by simpler form does not explain the variation of means of coding very well. Especially, the CONTAINMENT relation has almost equal numbers of prototypical and non-prototypical relations coded by plain case and relational nouns. Finally, we propose an alternative explanation based on the focus of attention as the motivation of the variation.

Keywords: case, cognitive linguistics, relational nouns, spatial semantics, Uralic

Acknowledgements: The data on Mordvin languages has been collected within the project “Descriptive Grammar of Mordvin” (University of Helsinki) funded by Kone Foundation. We would like to thank the anonymous reviewer for their comments, which helped to shape the paper.

For citation: Koivunen T., Erkkilä R. The interaction between spatial cases and relational nouns in the Uralic languages of the Volga–Kama area. *Voprosy Jazykoznanija*, 2022, 5: 64–85.

DOI: 10.31857/0373-658X.2022.5.64-85

Взаимодействие между пространственными падежами и реляционными именами в уральских языках волго-камского ареала

Томи Койвунен

Университет Турку, Турку, Финляндия; Удмуртский государственный университет, Ижевск, Россия; tokrkoj@utu.fi

Рику Эрркиля

Мюнхенский университет имени Людвиг и Максимилиана, Мюнхен, Германия;
Хельсинкский университет, Хельсинки, Финляндия; riku.erkkila@helsinki.fi

Аннотация: В статье рассматривается вариативность между формами пространственных падежей и реляционными именами в выражении двух пространственных отношений — нахождения ВНУТРИ и СВЕРХУ — в уральских языках волго-камского ареала (мордовских, марийских и пермских). Исследуя вариативность в кодировании указанных отношений, мы проверяли существующую в литературе гипотезу о том, что структурно более простыми формами обычно выражаются семантически более простые отношения. В наиболее прототипическом случае нахождения ВНУТРИ траектор располагается в границах трехмерного ориентира, а в наиболее прототипическом случае нахождения СВЕРХУ траектор сохраняет положение в пространстве, не падая вниз, благодаря опоре на двумерный ориентир. Для проверки гипотезы мы ранжировали встречающиеся в нашем материале ориентиры по степени прототипичности, то есть близости к описанным выше прототипам. Затем мы рассмотрели средства, которыми выражаются пространственные отношения в более и менее прототипических случаях, — падежная форма или реляционное имя. Результаты показывают, что в исследуемых языках описанная выше закономерность не может удовлетворительным образом объяснить выбор средств кодирования. В частности, для нахождения СВЕРХУ наблюдается почти равное число случаев кодирования прототипических и непрототипических отношений падежами и реляционными именами. В заключение мы предлагаем альтернативное объяснение, при котором наблюдаемая вариативность мотивируется фокусом внимания.

Ключевые слова: когнитивная лингвистика, падеж, пространственная семантика, реляционные имена, уральские языки

Для цитирования: Koivunen T., Erkkilä R. The interaction between spatial cases and relational nouns in the Uralic languages of the Volga–Kama area. *Voprosy Jazykoznanija*, 2022, 5: 64–85.

DOI: 10.31857/0373-658X.2022.5.64-85

1. Background

In this paper, we investigate one part of the system of spatial expressions in the Uralic languages of Volga–Kama area, namely the interaction between plain case inflection in nouns, and the use of relational nouns inflected in cases (traditionally called postpositions or adverbs according to their syntactic status). We are specifically interested in what conditions the choice between plain cases and relational nouns expressing CONTAINMENT¹ and SUPPORT in utterances.

It has been claimed that prototypical relations are expressed by general spatial expressions cross-linguistically [Levinson 2006: 171–173] and basic spatial terms are used to express basic spatial relations, e.g., CONTAINMENT and SUPPORT [Landau 2020: 93]. On the other hand, the variation between plain cases and relational nouns has been linked to focus of attention in Finnish [Ojutkangas 2008], a distant relative language of those studied here. The aim of this study is to analyze the variation evident in the coding of spatial relations, pinpoint the conditioning factors, and compare the results with previous research.

We restrict the comparison to two relations, namely CONTAINMENT and SUPPORT-FROM-BELOW (SUPPORT). There are two reasons for this. Firstly, these relations are considered (among some others) core spatial concepts or spatial primitives [Landau 2019; 2020], and they both are expressed with relational nouns in the languages studied here. Secondly, plain spatial cases exhibit variation with relational nouns even when the relation between the trajector (TR) and the landmark (LM) is prototypical. This is unexpected, as spatial cases are the general spatial expressions and basic spatial terms in the studied languages (cf., e.g., [Levinson 2006: 171–173; Levinson, Wilkins 2006a: 15–17]). This variation makes these two relations special among the spatial relations in the studied languages, which is the other reason we focus only on them.

The languages covered in this study are the Mordvin languages Erzya (MdE) and Moksha (MdM), the Mari languages Meadow Mari (MaM) and Hill Mari (MaH), and the Permian

¹ In cognitive linguistics it is customary to mark schematic relations with small capitals. We adopt this practice here.

languages Udmurt (Udm.), Komi Zyrian (KoZ), and Komi Permyak (KoP). They all belong to the Uralic language family and exhibit variation in spatial expressions typical for the family, i.e. some spatial relations are expressed by spatial case inflection on content nouns, and others with NPs consisting of a content noun and a relational noun inflected in a spatial case.

The structure of the paper is as follows: In this section we discuss the theoretical background, the structure of the inventory of spatial expressions in the studied languages, and previous research on the subject. In Section 2 we consider the data used in the study, as well as the methodology used to analyze the data. In Section 3 we present the analysis of the data. In Section 4 we compare the results with previous research, and in Section 5 we draw some conclusions on the matter.

1.1. Theoretical prerequisites

In this study we make use of the concepts and findings presented in the fields of cognitive linguistics and cognitive psychology. From cognitive linguistics the most important concepts for our study are conceptualization, relation, prototypicality, and active zone. From cognitive psychology we adopt the idea of prelinguistic spatial concepts.

Conceptualization refers to the activity of converting language-external (real-world) information to linguistic form. This is done by a conceptualizer, i.e. the person producing speech. In conceptualization, the conceptualizer chooses, among other things, what s/he wants to communicate about a situation, what viewpoint s/he adopts, what s/he focuses on in the situation, and what are the best linguistic elements to express these choices. The aim is to choose such elements that the conceptualizer gets her/his ideas across to her/his interlocutor as clearly as possible [Jiang, Yang 2021].

Relations are one of the two basic categories of entities in cognitive linguistics, the other being things. Things are usually expressed by content nouns, and relations by other linguistic elements, including other word classes and inflection. In the cognitive linguistics framework, verbs are seen to express relations evolving in time, i.e. processes [Langacker 1987: 244–275],² and other kinds of relational expressions depict atemporal relations. Atemporal relations are, in contrast to processes, stable in time, i. e. they lack a temporal profile.³ Spatial relations are a type of atemporal relation. In an atemporal relation, one thing, the TR, is related to another thing, the LM, in some domain. For spatial relations, this domain is space [Langacker 1987: 214–243].

Prototypicality is the phenomenon where some entities are considered representing a category better than some others. The idea is based on Eleanor Rosch's studies on categorization of things (see, e.g., [Rosch 1975]), but in modern cognitive linguistics any level of language is seen to be capable of exhibiting prototypicality [Lewandowska-Tomaszczyk 2007; Janda 2007]. Prototypicality of linguistic elements can be viewed from two perspectives: that of the coded situation, and that of the use of a linguistic element. From the first perspective, some situation is prototypically coded with some linguistic element, but it can still be coded with other elements, if the conceptualizer so wants. From the second perspective, some linguistic element is prototypically used to code some situation but can still be used to code a range of other situations as well. The first case is more relevant to the study at hand, as the variation of coding spatial relations is basically variation in the choice of one spatial element over the other by a conceptualizer.

One choice of linguistic element being preferred in a certain context is called salience of the prototypical member of a category [Taylor 2019: 132–133]. Salience of a member of a category essentially means that that member is seen as the best example of the category. Salience often (but not always) correlates with the frequency of the element, and this, in turn, correlates with the entrenchment of the element into the cognition of the language user, i.e. with how easily

² This includes states expressed by verbal predicates, as they have a temporal profile.

³ In cognitive linguistics framework the verbal predicate is seen to add temporality into clauses.

accessible it is when the conceptualizer needs to express something [Taylor 2019: 132–133; Divjak, Caldwell-Harris 2019]. For our study this means that the most suitable way to express a certain spatial relation is the most salient expression. The expression, then, is judged most suitable by the conceptualizer. In addition, the expression is entrenched in the cognition, and thus it is easily accessed by the conceptualizer and is probably the most frequent expression.

Active zone in a spatial relation is the part of the LM that is in focus in the relation. There are various reasons for a part of the LM to become the active zone in an utterance. For example, the action of the TR can affect or be confined just to a certain part of the LM, or the relation itself can have such configuration that some part of the LM automatically becomes the active zone. In a prototypical relation, the active zone is not highlighted, but rather inferred from the context. If, however, the relation is somehow unexpected, it is usually necessary to mark the active zone with additional linguistic material [Langacker 1987: 271–274].

Spatial primitives⁴ are prelinguistic concepts, which means that they are common to all humans regardless of their language. The two most typical, and most relevant for our study are CONTAINMENT and SUPPORT [Landau 2019; 2020]. The evidence for such primitives comes from studies of children’s conceptualization of spatial situations (see [Landau et al. 2017: 758–774]; for SUPPORT also [Lakusta et al. 2021]), and from experiments testing the spatial memory of adults (see, e.g., [Munnich et al. 2001]). Actually, CONTAINMENT and SUPPORT has been shown to cover different subrelations, but they do have prototypes, namely FULL CONTAINMENT (one object (TR) is contained inside another (LM)) and SUPPORT-FROM-BELOW (an object (TR) is shielded against the force of gravity by a plane (LM)) [Landau et al. 2017]. Hereafter, we mean exactly those prototypical relations by the terms CONTAINMENT and SUPPORT.

1.2. Spatial expression systems of the studied languages

Spatial expressions prototypically denote a relation between a TR and a LM in space. They can also be used to express other kinds of relations (e.g., temporal or causal), but these are not considered here. When spatial relations are concerned, the LM is always the referent of the phrase marked with the spatial expression.

All languages under study have a spatial case system with a three-way distinction between LOCATION, SOURCE (the starting point of movement), and GOAL (the endpoint of movement). In addition, Mordvin and Permic languages have dedicated cases to mark PATH, i.e. the route of movement [Zlatev 2007: 331–332]. The basic case system is thus unidimensional and does not express any configural information [Creissels 2011: 614–615, 618].⁵ This means that the plain cases express only the relation between the TR and the LM, and the configural information, if not specified with other means, is inferred with the help of encyclopedic knowledge about the prototypical configuration of the given TR and LM.⁶ Mordvin and Permic languages have more than four cases in their spatial case system, meaning that they make additional distinctions in their spatial expressions. These include DIRECTION and LIMIT of the action, among other distinctions. However, these are not considered further in this study, as additional semantics makes

⁴ Also called “innate spatial concepts” [Landau 2020].

⁵ Mordvin languages do have different declensions called basic and definite, but this does not affect the cases studied in this paper. In Erzya, all studied cases can be formed in both declensions, and in Moksha a neutral postpositional construction with the postposition *e- ~ ez-*, which is inflected in the cases of the basic declension [Hamari, Ajanki 2022: 400–401]. This postpositional construction, then, is in opposition with the inflected relational nouns.

⁶ Permyak dialects do have a series of spatial cases that define the configuration between the TR and the LM, which are a rather recent grammaticalization of the inflected relational noun *vil-* ‘top’ [Baker 1985: 175–191]. As these forms are not part of the literary standard, they are not considered in this paper.

it more difficult to determine if the case is used to express a prototypical relation. In addition, due to historical reasons, the original SOURCE case in Mari languages has been substituted by an element that is traditionally considered a postposition. This element, however, functions exactly like a case, i.e. it can be added after a relational noun to express SOURCE [Saarinen 2022: 437]. Therefore, we count it as the source marker of Mari languages. The relevant cases of the studied languages are shown in Table 1.

Table 1

The plain cases expressing the four basic spatial relations in the studied languages⁷

	Erzya	Moksha	Meadow Mari	Hill Mari	Udmurt	Komi-Zyrian	Komi-Permyak
LOCATION	-so / -se (inessive)	-sa / - <i>tsa</i> (inessive)	- <i>fte</i> / - <i>fto</i> / - <i>ftø</i> (inessive)	- <i>ftə</i> / - <i>ftë</i> (inessive)	-in (inessive)	-in (inessive)	-in (inessive)
SOURCE	-sto / -ste (elative)	-sta / - <i>tsa</i> (elative)	<i>gätē</i> (SOURCE postposition) ⁸	<i>gētš</i> (SOURCE postposition)	-ic (elative)	-ic (elative)	-ic (elative)
GOAL	-s (illative) ⁹	-s / - <i>ts</i> (illative)	- <i>fke</i> / - <i>fko</i> / - <i>fko</i> / - <i>f</i> (illative)	- <i>fka</i> / - <i>fke</i> / - <i>f</i> (illative)	-e / -i (illative)	-ə (illative)	-ə (illative)
PATH	-ka / -ga / -va (prolative)	-ka / -ga / -va / - <i>gæ</i> (prolative)	—	—	-ti / -eti / -jeti / -iti (prolative)	-əd (prolative); -ti (transitive)	-ət (prolative); -ti (transitive)

⁷ We use the terminology traditionally used about the cases in Uralistics. It is worth noting, however, that the cases called inessive, elative, and illative do not express any notion of CONTAINMENT like the labels seem to suggest, but are essentially locative, ablative, and lative, i.e. neutral spatial cases [Haspelmath 2011: 515–516].

⁸ An anonymous reviewer noted that we have not included the Mari unproductive morphological elative into the table. However, the unproductive elative occurs only in a few stems, and, e.g., Alhoniemi [2010: 63] considers all these kinds of forms as adverbs or particles. Also, there are unproductive inessive and illative endings; they occur, in particular, with the Mari relational noun stems *ymba(l)*- (MaM), *βē(l)*- (MaH) ‘on’, which are within our scope: *ymbalne* (MaM), *βēlnē* (MaH) (iness.); *ymbate* (MaM), *βēlets* (MaH) (elat.); *ymbak* (MaM), *βēk(ë)* ~ *βēlkë* (MaH) (illat.). In our study, these unproductive endings are included only with relational nouns. There is also a longer form of the source postposition *gätēan* (MaM), *gētšēn* (MaH), but it is far less frequently used than *gätē*, *gētš* (cf. *гыч* 37581 occurrences vs. *гычын* 303 occurrences in Meadow Mari web corpus [Arkhangelskiy 2019b]) and does not seem to have any additional meaning, see, e.g., [Alhoniemi 2010: 62]. This is the reason why we left the longer form out of this study. The question about whether the shorter and longer SOURCE postposition forms actually have any functional differences would deserve its own study.

⁹ In addition to illative, the Mordvin languages have two other GOAL-cases, namely lative and (definite) dative. These are not included in the sample for the following reasons: In a pilot dataset it was found out that there are practically no lative forms of the containment relational nouns in the corpus used (only 6 instances in Moksha and 4 in Erzya). Moreover, especially in Erzya but to some extent also in Moksha the lative seems to be used to mark movement to the vicinity of an object, rather than movement onto the top of an object. Therefore, the lative cannot be compared with the other cases studied here. The definite dative is not included in the data because it is not attested in relational nouns. As we are interested in the variation between a plain case and a relational noun inflected in a case, it is not useful to study definite dative here. A study where the whole GOAL-case system of Mordvin languages is analyzed must be left for future research.

In addition to spatial cases, the languages in question use a rather large number of relational nouns to express spatial relations. All the relational nouns in the languages we study are inflected in spatial cases just like all other nouns. Syntactically they can function either as a head of a noun phrase (traditionally the relational nouns are regarded as postpositions in this position), or as adverbs. We focus on the relational noun phrases here, as in this position the semantic difference between plain case inflection and the relational noun is most evident.

Unlike cases, relational nouns express configurational information pertaining to the relation between the TR and the LM. Most of the relational nouns express axial information (cf. [Levinson 2003: 24–56]), e.g., *az-* (Udm.) ‘front’, *ber-* (Udm.) ‘back’, *al-* (MdE) ‘under’, and *ver’-* (MdE) ‘above’. In other words, they specify a frame of reference that in turn specifies a search domain in relation to the LM where the TR is to be found. This additional information entails that such relational nouns and plain cases are not interchangeable, as plain cases never express axial information. However, some relational nouns, like Udm. *pu/k-* ‘in’ and *vil-* ‘top’ express configuration without axial information. These relational nouns should, in principle, be interchangeable with plain cases in situations where the prototypical relation between the TR and the LM coincide with the semantics of the relational noun. This means that, for example, with LMs having a prominent inside, the plain case should theoretically express containment by itself and using a relational noun like *pot-* ‘inside’ (MdE) does not add any information that is not deductible from the situation. This is exemplified by (1) and (2).

- (1) ERZYA (MokshEr-V.3/V.2 additions/Erzya/Syatko/Syatko-2006_4/Syatko-2006_4_108-116.txt)
 oʽuʒ-t-bezʽguta-t varʽsta-sʽtʽi vejke ombo-tʽsʽe lang-s, pejdʽezʽev-sʽtʽi
 [name]-PL-[name]-PL glance-PST1.3PL one other-ORD top-ILL smirk-PST1.3PL
 di sova-sʽtʽi kudo-s.
 and enter-PST1.3PL house-ILL
 ‘Otyuzh and Bezʽguta glanced at each other, smirked, and entered the house.’
- (2) ERZYA (MokshEr-V.3/V.2 additions/Erzya/Syatko/Syatko-2004_2/Syatko-2004_2_50-55.txt)
 “Adʽa, Sanʽa, nʽej-at erʽzʽa-nʽi pisatʽelʽ-tʽi!” di tʽjopavʽtʽ-inʽek
 PCL [name] see-PRS.2SG Erzya-GEN writer-PL and dive-PST1.1PL
 kudo-nʽtʽi pot-s.
 house-GEN.DEF inside-ILL
 ‘“Let’s go, Sanya, you will see Erzyan writers!” — and we entered [dived] (inside) the house.’

1.3. Previous research

The study of spatial relations in cognitively oriented research is abundant. Both monographs (e.g., [Levinson 2003; Tyler, Evans 2003]) and edited volumes (e.g., [Hickmann, Robert 2006; Levinson, Wilkins 2006b; Evans, Chilton 2010]), have been published, not to mention countless papers on the subject. These publications have treated a number of issues in spatial language and cognition, but the study of variation between different types of spatial expressions has not been so extensive.

Stephen Levinson and his collaborators studied the expression of static LOCATION in various languages and proposed that every language has a basic locative construction (BLC), that is, the construction used in basic locative function (as an answer to a *where*-question), when the TR and the LM are in a prototypical relation [Levinson 2006: 171–173; Levinson, Wilkins 2006a: 15–17]. Parameters for the probability of the BLC are the following (adapted from [Levinson, Wilkins 2006c: 515]):

1. close contact ↔ separation;
2. independent TR ↔ attached TR ↔ part-whole configuration;
3. contained TR ↔ contained LM;
4. inanimate TR/LM ↔ animate TR/LM
(this parameter refers to situations where at least one of the elements is inanimate/animate);
5. relatively small TR (compared to LM) ↔ relatively large TR;
6. prototypical relation between the TR and the LM ↔ non-prototypical relation;
7. prototypical TR (three-dimensional physical object) ↔ two- or one-dimensional ↔ negative space (e.g., a crack or a hole in an object).

These parameters predict the use of the BLC in a language. If the spatial relation expressed has properties shown on the left of the arrows, it is more likely to be expressed by the BLC. It is also shown that different parts of a motion event (SOURCE, GOAL, and PATH) can be expressed by the same means as in the BLC [Levinson, Wilkins 2006c: 527–537], e.g., with cases in a language that uses case-marking in the BLC.

In Uralic linguistics, some noteworthy research has been done on the matter at hand. Krista Ojutkangas has studied the variation between the inner spatial cases of Finnish (inessive, elative, and illative)¹⁰ and the relational noun *sisä-* ‘in’. Her conclusion is that expressions with the relational noun are used to focus the attention at the CONTAINMENT-relation between the LM and the TR, if the relation is somehow meaningful to the discussion, unexpected etc. [Ojutkangas 2008]. Jane Klavan and her associates have studied extensively the variation between the outer spatial cases of Estonian (adessive, ablative, and allative), and the relational noun *pea-* ‘on’. All studies point to the direction that the semantics of the LM and/or the relation between the TR and the LM affect the choice between a plain case and a relational noun inflected in case, but in addition morphosyntactic factors (e.g., the length of the phrase), lexical determination, and the used register play a role in the assignment (see, e.g., [Klavan et al. 2011; Klavan, Veismann 2017; Klavan 2021]).

Maria Usacheva [2011] has shown in her research on Beserman Udmurt that the variation between a plain case and a relational noun in expressions of CONTAINMENT and SUPPORT depends on information structure. What is meant under information structure by Usacheva can be subsumed, from the perspective of cognitive linguistics, under the notion of conceptualization, more specifically as a part of focal adjustments [Langacker 1987: 116–137]. Other manifestations of such process, where the language user chooses what linguistic elements s/he wants to use in describing a situation are, e.g., focus of attention (see [Ojutkangas 2008]) and register (see, e.g., [Klavan, Veismann 2017]).

2. Data and methodology

Our data was collected from various corpora of Uralic languages. Because so many different sources are used, the size and composition of the corpora vary. The primary genres presented, however, are newspaper texts, other non-fiction (e.g., historical texts, memoirs), and fictional prose (both novels and short stories). In addition, some corpora also contain poetry. Almost all the texts are written in the standard orthography of each language, except the Hill Mari corpus, where an obsolete version of orthography is mostly used, and rare instances of non-standard orthography in other corpora.¹¹ The sources are described in Table 2.

¹⁰ Note that in Finnish these cases actually express CONTAINMENT and are contrasted in this respect to another set of spatial cases expressing SUPPORT [Karlsson 2018: 228–229].

¹¹ Antconc [Anthony 2014] was used to extract data from some of the corpora, e.g., MokshEr.

Table 2

Corpora used in the study

Language	Corpora
Erzya	“Syatko” subcorpus of the MokshEr corpus [MokshEr 2010]
Moksha	Subcorpora in Moksha of the MokshEr corpus [MokshEr 2010]
Meadow Mari	Marko East [Marko (no year)], Oncyko [Oncyko 2000], Meadow Mari corpus [Arkhangelskiy 2019b], Wanca (Meadow Mari) [Helsingin yliopisto et al. 2019]
Hill Mari	Marko West [Marko (no year)], Wanca (Hill Mari) [Helsingin yliopisto et al. 2019]
Udmurt	Pilot version of Udmurt corpus (relational nouns; presently included into [Arkhangelskiy 2018]); Udmurt corpus (content nouns) [Arkhangelskiy 2018]
Komi-Zyrian	Komi Zyrian Web Corpus [Arkhangelskiy 2019a], Komi Corpus [Fu-Lab team 2021]
Komi-Permyak	Komi Permyak text collection from the University of Turku [Permyak 2009]

The data consists of three subsets. There are spatial expressions with relational nouns *pot(mo)-* (MdE), *potmə-* (MdM), *kørgø*, *kørgə-* (MaM), *kørgë-* (MaH), *puŋk-* (Udm.), *pitŋk-* (KoZ, KoP) ‘in’, spatial expressions with relational noun *lang-* (MdE, MdM), *ymba-* (MaM), *βë(l)-* (MaH), *vil-* (Udm), *vil-/viv-* (KoZ, KoP) ‘on’, and spatial expressions with plain cases. For each subset, a random sample of 20 examples of each spatial case under study was collected. As the number of cases present in each language and subset differs somewhat (see Section 1.2), the samples are of different sizes.

The data includes some instances where the combination of a content noun and a relational noun forms a compound. An example of such a compound is *vož vil* (Udm.) ‘meadow [lit. top of meadow]’, which can vary rather freely with *vož* ‘meadow’, and can even rarely take grammatical cases, e.g., *vož vil-len* ‘of the meadow [meadow top-GEN]’. Such compounds are included in our dataset because from cognitive linguistics point of view frequent collocations, i.e. *vož vil* ‘top of meadow’, where *vil* ‘top’ is a relational noun that frequently is attested with content nouns possessing a salient surface, such as *vož* ‘meadow’, on the one hand, and fixed compounds, such as *vož vil* ‘meadow’, on the other, form a continuum of construals that are entrenched to different degrees [Schmid 2014: 254–259]. Besides, they are exceedingly rare in the data. For example, *vož vil* is attested only once in the sample.¹²

Langacker [1987: 290–291] notes that a compound is a composite form, where one of the linguistic elements functions as a profile determinant, i.e. as an element that defines what kind of entity the compound denotes. In the case of compounds formed from a content noun and a relational noun it can be argued that the relational noun is the profile determinant, i.e. in *vož vil* ‘meadow’ the compound defines a part–whole relationship where the most salient part of the entity (top) is profiled. In such compounds it is possible that the element expressing the whole (e.g., *vož* ‘meadow’) can semantically override the profile determinant [Heyvaert 2009: 244–248], presumably leading to the loss of its semantic contribution altogether. This does not, however, change the fact that the relational noun is the profile determinant. Similarly, Langacker points out [1987: 288–290] that in a prepositional construction (collocation) in English the preposition

¹² An anonymous reviewer expressed concerns that including lexical compounds to the data might skew the results. Even though we think that comparing the behavior of compounds and collocations, it seems unlikely that including compounds would drastically alter our findings. A thorough comparison of the differences in, e.g., entrenchment of compounds compared to collocations must be left for further research.

is the profile determinant. This is because it expresses the relevant relation in the construction. As relational nouns in Uralic languages also express relevant relations in similar constructions (collocation), it can be concluded that relational nouns function as profile determinants. This means that functionally relational nouns used in compound and in relational noun constructions are similar enough to be compared. In addition, in expressions where the compound is inflected in a spatial case the difference between compound and collocation can be expected to be minimal, as in both situations the construal conveys a relation in physical space. The number of examples in each language and subset in total is shown in Table 3. The complete dataset is publicly available on Zenodo [Koivunen, Erkkilä 2022].

Table 3

The number of examples in the data

	Moksha	Erzya	Meadow Mari	Hill Mari	Udmurt	Komi-Zyrian	Komi-Permyak
Plain case	80	80	60	60	80	100	80
Relational noun ‘in’	80	80	60	60	80	100	80
Relational noun ‘on’	80	80	60	60	80	80	100
Total	240	240	180	180	240	280	260

The methodology used in the analysis of the data is based on the fact that humans have language-independent spatial primitives, concepts of certain spatial relations which are common to people regardless of their language (discussed in Section 1.1). We used the spatial primitives of CONTAINMENT and SUPPORT to establish hypothetical prototypes against which we compared the examples found in the data. The prototypical situation of CONTAINMENT includes a LM with distinct inside and outside, and a solid boundary between them. In addition, the LM wholly encompasses the TR and restricts its movements. The prototypical situation of SUPPORT includes a vertical, solid LM that holds the TR so that the TR won’t fall down. In both cases, the TR is prototypically smaller than the LM.

For LOCATION, the prototypes can be used as such, but for the three other relations, unfolding of the action in time must be considered. This means that an instance of SOURCE conforming to the prototype has the TR and the LM in the prototypical configuration when the action begins, and conversely, the prototypical configuration applies to GOALS in the end of the action. For PATH, the prototypical configuration must apply at least in the middle part of the action.¹³ For the cases with a moving TR, the entrenchment of the portrayed action was assessed. This means that we estimated if the action portrayed is typically associated with the spatial relation in question. For example, actions of ‘putting’ and ‘taking’ are strongly associated with INCLUSION, and in these cases the relation is highly entrenched, as things are habitually put in and taken out of CONTAINERS. On the other hand, e.g., ‘singing’ and ‘flying’ would not have such association.

Based on these parameters, we gave every example in our data a prototypicality score ranging from 4 (prototypical) to 1 (non-prototypical). Prototypicality score 4 entails that the example conforms totally to the hypothetical prototype and, in case of non-static relations, displays action associated with the relation. Score 3 is given when some part of the hypothetical prototype is not present, or the action is not typical for the relation. With score 2, more than one trait of the prototype missing, and the action is not typical for the given spatial relation, and finally, score 1 is given when the example exhibits almost no commonalities with the hypothetical prototype. The scale is left coarse on purpose, because with our data and methodology a more fine-grained

¹³ The “at least” constraint is necessary because PATH cases habitually also express a trajectory that is entirely inside the LM.

analysis would most probably lead to fruitless speculations. Such endeavors should involve native speakers to assure the accuracy of the results.

Because we are interested in the expressions of these spatial primitives in the languages under study, we have delimited the dataset to relational nouns expressing the said spatial primitives. It is possible that the languages have relational nouns that express superficially similar or perceptually close relations with different relational nouns, but these are excluded as they do not express only the spatial primitive, but rather spatial primitive and, in addition, some other configuration relevant in the relation between the TR and the LM. A case in point is the relational noun *pɔl-* (Udm.), *pi-* (KoZ) ‘inside some medium; among’. This relational noun is excluded because being inside a medium or among something is not prototypical CONTAINMENT (see, e.g., [Landau et al. 2017]). These relations are of course frequently conceptualized as CONTAINMENT, and being inside a medium arguably is a type of CONTAINMENT, but it cannot be considered prototypical according to the criteria we posited above.

An anonymous reviewer points out, that leaving out relational nouns like *pɔl-* (Udm), *pi-* (KoZ) ‘inside some medium; among’ skews the results of our study, as relational nouns of, e.g., Mari languages do not make similar distinctions, and thus our data for Mari languages includes different relations than our data for Permic languages. This is a relevant point when one is comparing the semantics of relational nouns across different languages, but our aim is to study how certain non-linguistic spatial relations, namely prototypical CONTAINMENT and prototypical SUPPORT are coded in languages (see Section 1). More precisely we are interested in the variation between plain cases and relational nouns in these situations. If a language (like Mari languages) uses the same relational noun to code prototypical and non-prototypical CONTAINMENT with the same relational noun, the non-prototypical instances are coded as such and analyzed accordingly, as explained above.

It is typical for languages to make different kinds of distinctions in the spatial domain. This means that different languages can use the same marker to express prototypical CONTAINMENT and additional CONTAINMENT-like relations, whereas another language can use different markers for these relations. An illustrative example is given in [Talmy 2000: 192–195], where the English preposition *in* is compared to markers of enclosure of Atsugewi (a Californian language). The meanings that are expressed in English by one preposition need in total 13 markers. Similar results of a typological study are shown in [Levinson, Wilkins 2006c: 514–526]. They analyzed the marking of location in 10 different languages and concluded that every language shows some differences in coding of location, but there is a core set of spatial relations that are frequently coded similarly. Thus, it can be concluded that different languages have spatial markers that have different extensions, but the conceptual core (e.g., prototypical CONTAINMENT and SUPPORT) is fairly constant across languages, and this is important for us because we are comparing prototypical instances, not the semantic extent of different markers.

3. Analysis

In this section, we will discuss the coding of CONTAINMENT and SUPPORT in our data. First, we show the frequencies of different types of relations and their prototypicality in our data. Then, in subsection 3.1, we discuss the CONTAINMENT relation, and in subsection 3.2 the SUPPORT relation. In 3.3, we take a quick look at those situations in our data where a plain spatial case does not express either of the relations under study, and in 3.4, we give a short summary of the analysis.

The frequencies of different types of relations in our data are given in Tables 4–6 below. The number of instances of expressing relations with different prototypicality values are shown for each language; in addition, in Table 6 the number of spatial relations that do not qualify either as CONTAINMENT or SUPPORT are given (the row “neither”).

Table 4

The prototypicality of CONTAINMENT-relations expressed by relational nouns

Language	Prototypicality score				Total
	4	3	2	1	
Moksha	19	40	12	9	80
Erzya	22	36	13	9	80
Meadow Mari	39	7	1	13	60
Hill Mari	19	8	11	22	60
Udmurt	30	31	13	6	80
Komi-Zyrian	69	16	6	9	100
Komi-Permyak	68	9	2	1	80
Total	266	147	58	69	540
Total %	49.26 %	27.22 %	10.74 %	12.78 %	100 %

Table 5

The prototypicality of SUPPORT-relations expressed by relational nouns

Language	Prototypicality score				Total
	4	3	2	1	
Moksha	42	15	21	2	80
Erzya	31	18	22	9	80
Meadow Mari	34	7	3	16	60
Hill Mari	35	5	0	20	60
Udmurt	46	11	15	8	80
Komi-Zyrian	34	2	1	43	80
Komi-Permyak	51	2	6	41	100
Total	273	60	68	139	540
Total %	50.56 %	11.11 %	12.59 %	25.74 %	100 %

Table 6

The prototypicality of spatial relations expressed by plain cases

Relation	Prototypicality score				Total	Total %
	4	3	2	1		
CONTAINMENT	81	135	35	34	285	52.78 %
SUPPORT	71	45	5	15	136	25.19 %
CONTAINMENT + SUPPORT	152	180	40	49	421	77.96 %
neither	—	—	—	—	119	22.04 %
Total	—	—	—	—	540	100 %
Total %	28.15 %	33.33 %	7.41 %	9.07 %	100 %	

The frequencies shown in the tables will be discussed in the following subsections.

3.1. Expressions of CONTAINMENT

As explained above, the Uralic languages of the Volga–Kama area use both relational nouns and plain cases to express CONTAINMENT. According to the corpus material we used, all the studied languages use relational nouns mostly for the prototypical CONTAINMENT relation. There are slight differences among the languages. The prototypicality of the usage of CONTAINMENT relational nouns seems to grow geographically from west to east, that is from Mordvin languages towards Mari and Permic, so that Komi varieties (both Zyrian and Permyak) exhibit the most prototypical usage of CONTAINMENT relational nouns. Taking to account all the languages studied here, about 49 % of all usage of relational nouns expressing CONTAINMENT represent score 4 (most prototypical); 27 % represent score 3; 11 % represent score 2; and 13 % represent score 1 (most non-prototypical). 76 % of all instances are used to express somewhat prototypical and 24 % somewhat non-prototypical relations.

According to our data, the usage of cases in the CONTAINMENT function is not as clear. Previous research points to the direction that the morphologically simplest form should cover the most prototypical instances. That means the plain cases should cover more prototypical CONTAINMENT relations than the relational nouns do. There is much more variation in plain cases in different languages. For example, in Moksha, there are more non-prototypical than prototypical relations expressed by plain cases. The most common score in our scale is 3 which means that, in CONTAINMENT relation, the use of case in its entirety is somewhat prototypical but not as prototypical as it should be according to earlier studies. Taking into account all the languages studied here, 28 % of plain cases in CONTAINMENT function represent score 4; 47 % represent score 3; 12 % represent both score 2 and score 1. About 76 % of all instances are used to express somewhat prototypical and 24 % somewhat non-prototypical relations.

The situation described above is surprising to some extent. The simpler form (the case) seems not to express the most prototypical CONTAINMENT relations, although according to the earlier studies they should be used in these situations. However, the somewhat prototypical and non-prototypical relations are expressed almost equally by both relational nouns and plain cases. Interestingly, it seems that the most significant differences in the usages of relational nouns and plain cases can be found between expressing relations of the scores 4 and 3. The state of affairs that relations of score 4 are expressed more often by relational nouns than by plain cases is very unexpected, as this means that relational nouns are more prototypical in expressing the CONTAINMENT-relation than plain cases are.

Looking at all the languages under study, Moksha shows the lowest share of examples with the score 4 with relational nouns, that is, 24 %. Erzya shows the second lowest share of such examples, that is, 28 %. Then go Hill Mari with 32 % and Udmurt with 38 %. The next one is Meadow Mari with 67 %. Finally, Komi languages show the highest number of score 4 relations: the number in Komi-Permyak is 67 % and in Komi-Zyrian 69 %.

How should we explain the differences in the usage of relational nouns expressing CONTAINMENT between the different languages? The influence of Russian could be one explaining factor. In practice, all the speakers of the languages studied here are (at least) bilingual, Russian being clearly the language that has the biggest influence on all Volga–Kama languages. In Russian, one cannot use only case to express CONTAINMENT or SUPPORT; one has to add one of the spatial prepositions, which distinguish (roughly) between CONTAINMENT and SUPPORT. These kinds of foreign patterns could easily have affected minority languages. Nevertheless, the Russian influence is considered to be the strongest, firstly, in Mordvin languages (Erzya and Moksha), and secondly, in Komi languages (Komi-Zyrian and Komi-Permyak). However, these are the languages whose CONTAINMENT relational nouns show almost the extreme shares of prototypicality

in our data, that is, the Mordvin languages have a low share of score 4 in relational nouns expressing CONTAINMENT, while Komi languages, in contrast, have the highest score 4. Based on this, it seems that the Russian influence is not the crucial factor regarding the prototypicality of CONTAINMENT expressions.¹⁴

When the possibility of Russian influence is probably excluded, there does not seem to be any clear reason for the differences. However, the trend is similar in every language we studied: the prototypical use of relational nouns for expressing prototypical CONTAINMENT relation seems to be a very common feature of the Uralic languages of the Volga–Kama area.

The prototypicality of the plain cases expressing CONTAINMENT varies among the languages. Considering all languages together, the share of the score 4 is 28 %; that of the score 3 is 47 %; that of the score 2 is 12 %; and that of the score 1 is also 12 %. Again, according to the previous studies, the use of plain cases should be the most prototypical way to express CONTAINMENT. This means that the score 4 should have the largest score, but that is not the case.

There are differences among separate languages: Meadow Mari has an especially high share of the score 4 (50 %). The next ones are in order: Moksha (33 %), Erzya (31 %), Komi-Permyak (31 %), Hill Mari (22 %), Komi-Zyrian (21 %), and finally Udmurt (18 %) with a rather low score. The Mordvin languages have a similar share of the score 4, but the Mari and Permic groups do not show such conformity. The most striking is the difference between Komi-Permyak and Udmurt.

However, if we count scores 4 and 3 together (that is, “at least somewhat prototypical” examples), we get the following numbers (from the highest share to the lowest): Meadow Mari — 93 %, Erzya — 86 %, Udmurt — 82 %, Hill Mari — 78 %, Komi-Zyrian — 69 %, Moksha — 69 %, Komi-Permyak — 65 %. Here, Udmurt does not show a low share anymore, but Moksha and the Komi languages do, the others being no further than 15 percentage points from each other. This looks more like the situation we would expect: relatively prototypical CONTAINMENT relation is shown by plain case very often. Nevertheless, it is hard to explain why relational nouns have a higher number of score 4 than plain cases.

There may be some lexically motivated use of the relational nouns. In Meadow Mari, for instance, 65 % of all illative forms of the relational noun expressing CONTAINMENT are used in a context where the predicate includes the verb *puraf* ‘to enter’ or some form derived from it. The same happens in Hill Mari, where the share of the etymologically and semantically same predicate *pəraf* ‘to enter’ or its derivation is 55 %. The word meaning ‘enter’ is obviously very plausible in utterances with a CONTAINMENT relational noun in illative form. An example is shown in (3).

(3) MEADOW MARI (MarkoEast 35:15)

oʃ	yp-an	saməək	sotnik-lan	kətəkəraλ-e,	omsa-m
white	hair-DER.AD	young	sotnik-DAT	shout-PST1.3SG	door-ACC
kot̩ɛərɪk	ʃupʃəl-o,	mont̩ɛa	gaje	ʃy̆te-aŋ-ʃe	pört körgə-ʃ
ONOM	pull-PST1	bathroom	like.PO	SOOT-DER.V-PCT.PACT	house inside-ILL
pur-en	ʃogaλ-e				
enter-CNV	stand.up-PST1.3SG				

‘(S/he) shouted to a white-haired young sotnik, pulled a creaking door, went into a house sooty like a sauna.’

3.2. Expressions of SUPPORT

As is the case with the CONTAINMENT relation, also the SUPPORT relation has similarly twofold expression in the languages studied. According to the literature, the plain cases should be used primarily for prototypical SUPPORT relation in every language studied here. In addition, there is

¹⁴ There are other historical and also present contact languages as well, most importantly Turkic languages Chuvash, Tatar, and Bashkir, but research about their possible influence has to be left for future studies.

also a relational noun ‘on’ in all the languages. This is a symmetric situation compared to the expression of CONTAINMENT relation. In every language we study, the same plain cases are used to express both relations, but the relational noun is different.

As it is said in Section 3.1, the relational noun is used mostly in a prototypical CONTAINMENT relation. The situation seems to be similar — and even clearer — in the SUPPORT relation. Again, this is against the assumption that more prototypical relation is expressed by a simpler form, a plain case. When a relational noun is used, the score 4 (most prototypical) is the most common score in the SUPPORT relation in almost every language studied in this paper. Score 4 of all the languages studied by us represents a share of 51 %; score 3 represents a share of 11 %; score 2 represents a share of 13 %; and score 1 represents a share of 26 %. In the different languages the shares of score 4 are as follows: Moksha — 53 %; Erzya — 39 %; Meadow Mari — 57 %; Hill Mari — 58 %; Udmurt — 77 %; Komi-Zyrian — 43 %; Komi-Permyak — 51 %.

However, there are differences between the languages. The most striking difference is that, unlike any other of the languages, the Komi varieties also have a very high share of score 1 (most non-prototypical) with relational nouns expressing SUPPORT. In Komi-Zyrian it is even more common than score 4, that is 54 % (score 1) compared to 43 % (score 4). In Komi-Permyak the share is close to score 4 as well, specifically 41 % (score 1) compared to 51 % (score 4). Here, the Russian influence is very probable; the use of prepositions of Russian (the functional equivalent of relational nouns) is briefly described in Section 3.1. It seems likely that the Russian preposition *na* very often corresponds to a relational noun expressing SUPPORT, namely *vil-/viv-* in Komi. Considering the fact that a Russian speaker has to use one of the prepositions *v* or *na* to express CONTAINMENT or SUPPORT, it is clear that they have a variety of more abstract usages as well. It seems, according to the data, that Russian *v* ‘in’ mostly corresponds to a plain case and Russian *na* ‘on’ to a relational noun expressing SUPPORT. Considering the fact that Komi has undergone vast Russian influence and interference, this correspondence explains much about the non-prototypical usage of the relational nouns expressing SUPPORT in both Komi varieties.

Score 1 is prominent also in Mari, although not as clearly as in Komi. In Meadow Mari it is 27 % and in Hill Mari 33 %. The Mordvin languages Erzya and Moksha and another Permic language Udmurt show clearly smaller values. In Erzya, the most non-prototypical value (score 1) is given for 11 % of all our sentences, while in Moksha the same value is only 3 %, and in Udmurt it is 10 %. It is somewhat surprising that the number of score 1 is so different in the geographical extremities of the area under research, since both Mordvin and Komi varieties are considered to have the largest Russian influence. If the exceptional non-prototypical usage of the support relational noun in Komi could be explained by Russian influence, then why does Mordvin show the complete opposite?

If we look at the prototypicality by combining the “somewhat prototypical” scores 4 and 3, the result is rather clear. Only the Komi varieties are exceptional as they get shares only around 50 %. In Moksha, the score “4 + 3” is 71 %; in Erzya 61 %; in Meadow Mari 68 %; in Hill Mari 67 %; in Udmurt 71 %; in Komi-Zyrian 45 %; in Komi-Permyak 53 %. However, this should definitely not be the case according to the literature. All the languages show a high share of prototypicality in relational nouns expressing SUPPORT.

The SUPPORT relation is also expressed by the same plain cases as the CONTAINMENT relation is, and these cases should be used in more prototypical relations. Altogether the prototypicality percentages of plain case used in SUPPORT relation are as follows: 52 % for score 4; 33 % for score 3; 4 % for score 2; and 11 % for score 1. The plain cases are used to express SUPPORT in 25 % of all its occurrences in our data. Of this share, the languages get score 4 as follows: Moksha — 73 %; Erzya and Meadow Mari — 53 %; Hill Mari — 41 %; Udmurt — 36 %; Komi-Zyrian — 67 %; and Komi-Permyak — 50 %. Here, the shares of the score 4 differ drastically. Udmurt and Hill Mari have shares under 50 % and they are rather far below all other languages in this case. In contrast, Moksha is the only language to get a share more than 70 %, although Komi-Zyrian gets close. There does not seem to be any correlation between the closest related languages.

Again, when we add together scores 4 and 3, we get different results: Moksha — 100 %; Erzya — 79 %; Meadow Mari — 94 %; Hill Mari — 71 %; Udmurt — 96 %; Komi-Zyrian — 75 %; and Komi-Permyak — 81 %. Here we can see that if a plain case is used to code a SUPPORT relation, it is usually a more or less prototypical relation. All the languages have a share of 70 % or more, Moksha even 100 %. Here, one has to remember that plain cases are used much more commonly to code CONTAINMENT than SUPPORT. It means, *inter alia*, that prototypicality rates for SUPPORT in this case are more susceptible to accidental distortion.

3.3. Expressions of neither CONTAINMENT nor SUPPORT

Along with the hypothetically most basic relations of CONTAINMENT and SUPPORT, the plain cases are used to express a range of other relations as well. There is a 22 % share of the other relations coded in all the plain case usage. Interestingly, that is almost as much as SUPPORT, the corresponding share being 25 %. These other functions are, e.g., ceremonies, languages, government bodies, and usage like ‘at work’, ‘to the family’. Clear temporal expressions are totally excluded in this study.

The exact share of “neither CONTAINMENT nor SUPPORT” examples differs among the languages: the share in Moksha is 33 %; in Erzya 24 %; in Meadow Mari 25 %; in Hill Mari 42 %; in Udmurt 9 %; in Komi-Zyrian 18 %; and in Komi-Permyak 11 %. However, these expressions are not the main focus of this study. Therefore, we will not consider them further.

3.4. A brief summary of the analysis

The main observation that can be made based on our data is that relational nouns are widely used to express the CONTAINMENT and SUPPORT relations. Especially the prototypical SUPPORT relations are expressed extremely often by relational nouns. There is variation between the languages, but the trend is clear. Some of the details can probably be explained by language contacts and the effect of lexical semantics. Some spatial relations represent neither CONTAINMENT nor SUPPORT. In these situations, plain cases are a clear choice instead of relational nouns.¹⁵ However, Komi varieties widely use the relational noun expressing SUPPORT for this purpose as well, which is probably Russian influence.

Plain cases are also used to code prototypical relations as well to some extent, since the score 3 is the most common in utterances with plain cases expressing CONTAINMENT. With plain cases expressing SUPPORT, score 4 is the most common, but it has to be considered that the usage of plain cases in that function is relatively small. However, when the plain cases are used to code SUPPORT, its prototypicality seems to be high.

4. Comparison of our results and previous explanations

The studies of Levinson and associates presented in Section 1.3 imply that in unidimensional spatial case systems BLC should be expressed by a plain spatial case, as it expresses the unmarked relation between TR and LM [Levinson, Wilkins 2006a: 15–17]. This derives

¹⁵ The relational nouns can of course be used in non-spatial contexts to express relations other than CONTAINMENT or SUPPORT. However, as our study is restricted to how to express spatial relations, non-spatial uses of relational nouns are not considered.

from the fact that in a unidimensional case system the cases do not express additional configurational information about the relation between the TR and the LM, as was discussed in Section 1.2. In addition, plain cases should express SOURCE, GOAL, and PATH in basic motion events, as it is possible for languages to use the same spatial elements to express both BLC and motion [Levinson, Wilkins 2006a: 17–19; 2006c: 527–537]. On the other hand, CONTAINMENT and SUPPORT are more marked relations between the TR and the LM [Levinson 2003: 100], and therefore should be expressed by other means than plain cases when language has a unidimensional spatial system.

First, we take a look at how the parameters adapted from [Levinson, Wilkins 2006c: 515] for situations where BLC is probably used fit our data. The parameters are following (shown also in Section 1.3):

1. close contact ↔ separation
2. independent TR ↔ attached TR ↔ part-whole configuration
3. contained TR ↔ contained LM
4. inanimate TR/LM ↔ animate TR/LM (this parameter refers to situations where at least one of the elements is inanimate/animate)
5. relatively small TR compared to LM ↔ relatively large TR
6. prototypical relation between TR and LM ↔ non-prototypical relation
7. prototypical TR (three-dimensional physical object) ↔ two-or one-dimensional ↔ negative space (or hole).

The BLC is more likely, according to [Levinson, Wilkins 2006c: 515–519], when the spatial relation includes values on the left in the parameters. This applies to the use of plain cases in our data, as seen in (4) and (5).

(4) HILL MARI (MarkoWest 301:39)

pørt-əftē tsela-ok əl-ən, no kězēt-æt fukə ma+ʃon-əm
house-INE all-FOC be-PST2.3SG but now-FOC much something-ACC
 etʃe mo-mə agəl.
 still find-PTCP.PASS NEG.3SG

‘Everyone was in the house but now many things are still missing.’

In example (4), the parameters implying the use of BLC show following values:

1. TR and LM are in close contact. Being inside a house makes TR close to LM.
2. TR is independent from LM. The person functioning as TR is not attached to LM and is not a part of it.
3. TR is contained, as is typical to situations where TR is inside LM.
4. LM is inanimate. LM (house) is an inanimate thing.
5. TR is relatively small compared to LM. People tend to be smaller than houses.
6. The relation between TR and LM is prototypical. People habitually are inside of houses.
7. TR is prototypical. People are three-dimensional and physical, and they are objects, if we conflate the notion of object here with the notion of thing as opposed to relation [Langacker 1987: 183–274].

(5) MOKSHA (MokshEr-V.3/V.2 additions/Moksha/Moksha/Varia-2/Pinjasov.txt)

Son af smenda-zʲ as'kølda-sʲ **kijaksə-va.**
 3SG NEG dare-CNV step-PST1.3SG **floor-PROL**

‘It (a chick) timidly walked along the floor.’

In (5), the same parameters manifest as follows:

1. TR and LM are in close contact. Walking on a floor includes contact between TR and LM.
2. TR is independent from LM. The entity (in this case a chick) functioning as TR is not attached to LM and is not a part of it.

3. TR is not contained. The relation portrayed in the example is one of SUPPORT, where TR is prototypically not contained.
4. LM is inanimate. LM (floor) is an inanimate thing.
5. TR is relatively small compared to LM. Chicks tend to be smaller than floors.
6. The relation between TR and LM is prototypical. Floors are meant to be walked on.
7. TR is prototypical. Chicks are three-dimensional, physical and objects (cf. point seven in the analysis of example 4).

In both languages six or seven out of seven parameters show values typical for the use of BLC.¹⁶ This shows that in Hill Mari and Moksha the spatial cases are used as BLC with both CONTAINMENT and SUPPORT. However, if we look at examples where a relational noun is used to express the relation, we find such examples as (6) and (7).

(6) KOMI-ZYRIAN (Parma Gor 17.05.2018)

sci-jas ju-əni, mij udʒal-igkosti-nis mu pitʃk-ic
 one-PL say-3PL that.CONJ work-CNV-PX3PL ground inside-ELA
pet-i-s radon gaz.
 exit-PST1-3SG radon gas

‘Some people say that while they worked there came radon gas out of the ground.’

Example (6) shows the following appearances of the BLC-parameters. The relational noun is inflected in relative, so the parameters refer to the situation before the action:

1. TR and LM are in close contact. TR exits from inside of LM.
2. TR is independent from LM. TR (radon gas) is not attached to LM, nor is it an inalienable part of it.
3. TR is contained in the beginning of the depicted situation.
4. Both TR and LM are inanimate. Radon gas and ground are inanimate things.
5. TR is relatively smaller than LM. It is of course hard to assess the size of gas, but usually there is less gas than ground when the former escapes from the other.
6. The relation between TR and LM is more or less prototypical. If there is gas in the ground, it can quite easily start to move because of work done to the ground.
7. TR is not prototypical. Gas is not, strictly speaking, three-dimensional or physical.

The anonymous reviewer commented on this analysis that they do not think that the relation in (6) is prototypical CONTAINMENT, as the LM is not an actual CONTAINER. This is a valid point, as the LM (in (6) specifically the inside of *mu* ‘ground’) in fact possess only some of the properties of a prototypical CONTAINER, e.g., it is three-dimensional and restricts the actions of the TR. However, in this specific case two facts point to the direction that the relation should be considered more prototypical CONTAINMENT than not. Firstly, in a relation between the TR and the LM, the TR is the profile determinant [Langacker 1987: 288–290].¹⁷ This means that the relation tells something about the location of the TR. Secondly, based on encyclopedic knowledge, language users know that it is highly unlikely that the relation between *gaz* ‘gas’ and *mu* ‘ground’ would be one of SUPPORT. Therefore, the prototypical spatial relation in (6) is CONTAINMENT.

(7) UDMURT (Ošmes 19.03.2015)

Budos-jos-ti iʃka, voz vil-ic turin bəm-jos-ti
 plant-PL-ACC pull.PRS.3SG meadow top-ELA hay small.haystack-PL-ACC

¹⁶ In the Moksha example parameter three is irrelevant, as the example portrays a SUPPORT-relation.

¹⁷ Note that in Section 2 it is said that relational noun is the profile determinant in a construction. This refers to a lower level of organization. In a construction with relational noun, relational noun is the profile determinant, but when it functions as a LM in a higher-level construction, the TR of that construction is the profile determinant [Langacker 1987: 288–290].

in-oz ik dʒut-e.
 sky-TERM PCL raise-PRS.3SG

‘She pulls the plants, she will make the small haystacks from top of the meadow rise up to the sky.’

Example (7) shows the following manifestations of the parameters. As the example has a relational noun inflected in relative, the parameter refers to the situation in the beginning of action, as in (6) above:

1. TR and LM are in close contact. TR is on the LM, as is typical for haystacks on fields.
2. TR is independent from LM. TR (haystack) is not attached to LM, nor is it an inalienable part of it.
3. TR is not contained. The relation portrayed in the example is one of SUPPORT, where TR is prototypically not contained.
4. Both TR and LM are inanimate. Haystacks and fields are inanimate things.
5. TR is relatively smaller than LM. A haystack is usually significantly smaller than a field. In addition, the exact meaning of the Udmurt compound is ‘(small) haystack’, so it is evident that the size difference between TR and LM is clear.
6. The relation between TR and LM is prototypical. Haystacks tend to be on fields.
7. TR is prototypical. Haystack is a three-dimensional physical object.

From these examples we see that, in addition to plain cases, also relational nouns can be used in the BLC, and they tend not to be any worse fit than plain cases. Our analysis shows, then, that the proposition expressed by Levinson and Wilkins does not hold for all languages, at least in prototypical CONTAINMENT and SUPPORT relations. The authors themselves underline that the proposed generalizations are tendencies, but the percentage of contradicting examples discussed in Section 3 seems to require some refinement of the original proposition. A recent study on the constructions expressing location in Secoya (a Tucanoan language of Amazonia) points to the same direction, i.e. BLC is not a sustainable explanation [Vallejos, Brown 2021]. It was shown in the study that Secoya has as many as four different constructions that can be used to express basic location, and thus the language lacks a single BLC, in contrast to the proposition of Levinson and Wilkins. What their theory predicts correctly, however, is that more complicated expressions are used overwhelmingly in non-prototypical situations, as exemplified by (8).

- (8) ERZYA (MokshEr-V.3/V.2 additions/Erzya/Syatko/Syatko-2007_2/Syatko-2007_2_119-139.txt)
- | | | | | | |
|--------------------------|-------------------------|-------------------|---------------------|------------------|--------------------------|
| targa-sinze | stolⁱ | pot-sto, | st ⁱ -i, | nolda-sinze | ser ⁱ ga-nzo, |
| pull-PRS.3SG>3PL | table | inside-ELA | stand.up-PRS.3SG | free-PRS.3SG>3PL | body-POSS.3SG |
| must _{nom} must | mer ⁱ -i | kurgo-nzo. | | | |
| ONOM | say-PRS.3SG | mouth-POSS.GEN | | | |

‘She pulled them (pants) from inside the table, stood up, freed them along her body, silent laughter sounds from her mouth.’

The relation is non-prototypical, as usually things are in a SUPPORT relation to table, and not in CONTAINMENT relation, as in (8). As predicted, the relation is then expressed by a relational noun, not with a plain case. This is the situation across the studied languages, as almost 24 % of the CONTAINMENT relations and over 38 % of the SUPPORT relations expressed by a relational noun are non-prototypical (scores 2 and 1), whereas only a little over 21 % of CONTAINMENT and SUPPORT relations expressed with a plain case are non-prototypical. The picture gets a bit more complicated if we look at the relations expressed by a plain case separately. With SUPPORT relation, a plain case is used with non-prototypical relations in little less than 15 % of cases, which is in line with the prediction put forward. However, with CONTAINMENT relation the plain case is used for non-prototypical relations in little over 24 % of cases, which is almost the same amount as the relational noun is used (little less than 24 %). Clearly, the relational noun expressing

CONTAINMENT is used to express something other than non-prototypicality of the relation between the TR and the LM.

In addition, the parameters put forward by Levinson and Wilkins do not capture the non-prototypicality of the relations we are studying. This can be seen by looking at the values the proposed parameters give for example (8). For five or six of the seven parameters, these values are characteristic of prototypical relations: the TR and the LM are in close contact, the TR is independent from the LM, the TR is contained, the TR and the LM are inanimate, the TR is prototypical, and the TR is somewhat smaller than the LM. One of the values, namely that the TR is contained, actually constitutes the non-prototypicality of the situation in (8) (contrary to what it does in plenty of other cases). The assessment of the prototypicality of a relation should thus be somehow incorporated into the model. One possibility is to use spatial prototypes, as we have done in this study, but this method does not cover all cases, e.g., if the TR is VICINAL to the LM.

CONTAINMENT does not behave like expected compared to the proposition discussed above. According to the proposition, the BLC is the simplest type of spatial expression in a language, and it is used to express the prototypical relation between the TR and the LM. Our study shows, however, that in the Uralic languages of the Volga–Kama area a plain case, the simplest means in these languages to express spatial relations, and a relational noun inflected in spatial case, which is a (structurally) far more complex expression, are used almost interchangeably to express CONTAINMENT relation. This applies both to prototypical and non-prototypical cases.

Of CONTAINMENT relations expressed by a plain case in our data over 75 % are prototypical (scores 4 and 3), and over 24 % are non-prototypical (scores 2 and 1). The same numbers for relational nouns are over 76 % and almost 24 %, respectively. Both means for expressing CONTAINMENT are, from a point of view of expressing basic spatial relations, practically interchangeable in our data. Based on this result the concept of BLC cannot be implemented on the languages under study, as it is impossible to say which of the alternatives is more basic. This seems to point towards another explanation for the variation between plain cases and relational nouns put forward in the literature, namely that relational nouns are used to focus the attention to the CONTAINMENT relation if the relation is somehow relevant [Ojutkangas 2008]. This line of research should be continued with native speakers, as the corpus data we used is not very suitable for assessing focus of attention. In addition, the findings of Jane Klavan and her associates, namely that the properties of the referent of the LM and the relation between the TR and the LM are meaningful in the variation between a plain case and relational noun (see, e.g., [Klavan 2021]), and the research on information structure by Usacheva [2011] could possibly help to explain the variation. This, however, must also be left for further studies.

SUPPORT conforms a little better to Levinson's comment about the markedness of the relation [Levinson 2003: 100], but there are still some unexpected results. In our data, prototypical SUPPORT (scores 4 and 3) coded with a relational noun covers a little under 62 % of all instances where SUPPORT is coded by a relational noun. On the other hand, the more prototypical half of SUPPORT relations coded by a case (scores 4 and 3) make up over 85 % of all instances of SUPPORT coded by a case. The high percentage of prototypical relations marked by plain cases is expected, but it is unexpected that about two thirds of SUPPORT relations expressed by relational nouns are also prototypical. Non-prototypical cases, on the other hand, are usually coded with relational nouns, as over 38 % of examples with relational nouns represent non-prototypical relations (score 2 or 1), whereas only about 15 % of SUPPORT relations coded by plain cases are non-prototypical. It seems, then, that when the SUPPORT relation cannot be inferred with the help of encyclopedic knowledge pertaining to the properties of the LM referent, the studied languages prefer to explicitly code the SUPPORT relation with a relational noun.

These results seem to point to the direction that the prototypicality of the spatial relation, on the one hand, and the focus of attention, on the other hand, affect the conceptualization of SUPPORT in the languages studied. First, as noted above, the prototypicality of the relation affects the choice between a plain case and a relational noun. If the relation is non-prototypical, a relational noun is preferred. However, if it is prototypical, some other factor must affect the decision.

Here, too, the findings in [Ojutkangas 2008] about the variation in the expressions of CONTAINMENT in Finnish, namely that conceptualizers use the relational noun as a focus device when the relation between the TR and the LM seems to explain the variation between plain cases and relational nouns in prototypical cases. As with CONTAINMENT relation, the research should be continued with native speakers, though.

5. Conclusions

In this paper, we have focused on variation between plain spatial cases and relational nouns inflected in spatial cases as expressions of two primitive spatial relations, namely CONTAINMENT and SUPPORT. We studied the phenomenon using corpus data collected from seven Uralic languages, Erzya, Moksha, Meadow Mari, Hill Mari, Udmurt, Komi Zyrian, and Komi Permyak, all spoken in the Volga–Kama area in central European Russia. All these languages have a uni-dimensional spatial case system, i.e. they do not refer to additional dimensions of spatiality (e.g., ‘inside’ or ‘on top of something’) besides LOCATION, SOURCE, GOAL, and PATH. To complement their case system, all the languages have systems of relational nouns, of which those expressing CONTAINMENT and SUPPORT were under scrutiny.

It was discussed that previous studies have proposed that languages use structurally simpler forms to express more basic relations between the TR and the LM [Levinson 2003; Levinson, Wilkins 2006b]. We tested this hypothesis by assessing the prototypicality of spatial relations coded by plain cases, on the one hand, and by relational nouns, on the other hand, in the Uralic languages of the Volga–Kama area. The expected result was that plain cases would be used in prototypical situations, and relational nouns in non-prototypical situations.

We gave all examples in our data a prototypicality value on a four-point scale from 4 (prototypical) to 1 (non-prototypical). The prototypicality of a relation was based on the spatial primitives introduced in cognitive psychology (e.g., [Landau 2019; 2020]). By comparing the properties of the relation between the TR and the LM in our examples to the properties of the spatial primitives, we could evaluate the examples reliably. What we found out was that the plain cases are used to express prototypical relations, as expected, but most of the situations coded by relational nouns were also prototypical. However, the non-prototypical relations in our data were coded almost exclusively by relational nouns, which is in line with the idea that less prototypical relations are coded by structurally more complex forms. As an explanation for the variation in the coding of prototypical relations, we propose the idea put forward by Krista Ojutkangas [2008], namely that relational nouns are used to conceptualize situations where special attention should be paid to the relation between the TR and the LM. However, a more precise study of this matter should be conducted with the help of native speakers, as it is usually hard to decide from corpus data if the relation between the TR and the LM is in focus or not.

ABBREVIATIONS

1, 2, 3 — 1st, 2nd, 3rd person
 ACC — accusative
 CNV — converb
 CONJ — conjunction
 DAT — dative
 DEF — definite declension
 DER.AD — adjectival derivation
 DER.V — verbal derivation
 ELA — elative

FOC — focus particle
 FUT — future tense
 GEN — genitive
 ILL — illative
 INE — inessive
 NEG — negational element
 ONOM — onomatopoeic element
 ORD — ordinal derivation
 PCL — particle

PL — plural
 PO — postposition
 POSS — possessive declension
 PROL — prolativ
 PRS — present tense
 PST1 — first past tense
 PST2 — second past tense

PTCP.ACT — active participle
 PTCP.PASS — passive participle
 PX — possessive suffix
 SG — singular
 TERM — terminative
 > — marking object conjugation persons

REFERENCES

- Alhoniemi 2010 — Alhoniemi A. *Marin kieliooppi* [Grammar of Mari]. Helsinki: Société Finno-ougrienne, 2010.
- Anthony 2014 — Anthony L. *AntConc (version 3.4.4w)*. Tokyo: Waseda Univ., 2014. <https://www.laurenceanthony.net/software>
- Arkhangelskiy 2018 — Arkhangelskiy T. *Udmurt corpus*. 2018. <http://udmurt.web-corpora.net/index.html>.
- Arkhangelskiy 2019a — Arkhangelskiy T. *Komi-Zyrian corpus*. 2019. <http://komi-zyrian.web-corpora.net/index.html>.
- Arkhangelskiy 2019b — Arkhangelskiy T. *Meadow Mari corpus*. 2019. http://meadow-mari.web-corpora.net/index_en.html.
- Baker 1985 — Baker R. *The development of the Komi case system. A dialectological investigation*. Helsinki: Société Finno-Ougrienne, 1985.
- Creissels 2011 — Creissels D. Spatial cases. *The Oxford handbook of case*. Malchukov A., Spencer A. (eds.). Oxford: Oxford Univ. Press, 2011, 609–625.
- Divjak, Caldwell-Harris 2019 — Divjak D., Caldwell-Harris C. L. Frequency and entrenchment. *Cognitive linguistics. Foundations in language*. Dąbrowska E., Divjak D. (eds.). Berlin: De Gruyter Mouton, 2019, 61–86.
- Evans, Chilton 2010 — Evans V., Chilton P. (eds.). *Language, cognition and space. The state of the art and new directions*. London: Equinox, 2010.
- Fu-Lab team 2021 — *Korpus komi yazyka* [Corpus of the Komi language]. 2021. <http://komicorpora.ru/>.
- Hamari, Ajanki 2022 — Hamari A., Ajanki R. Mordvin (Erzya and Moksha). *The Oxford guide to Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 392–431.
- Haspelmath 2011 — Haspelmath M. Terminology of case. *The Oxford handbook of case*. Malchukov A., Spencer A. (eds.). Oxford: Oxford Univ. Press, 2011, 505–517.
- Helsingin yliopisto et al. — *Wanca 2016, Korp Version*. Kielipankki, 2019. <http://urn.fi/urn:nbn:fi:lb-2019052401>.
- Heyvaert 2009 — Heyvaert L. Compounding in cognitive linguistics. *The Oxford handbook of compounding*. Lieber R., Štekauer P. (eds.). Oxford: Oxford Univ. Press, 2009, 233–254.
- Hickmann, Robert 2006 — Hickmann M., Robert S. (eds.). *Space in languages. Linguistic systems and cognitive categories*. Amsterdam: John Benjamins, 2006.
- Janda 2007 — Janda L. Inflectional morphology. *The Oxford handbook of cognitive linguistics*. Geeraerts D., Cuyckens H. (eds.). Oxford: Oxford Univ. Press, 2007, 632–649.
- Jiang, Yang 2021 — Jiang C., Yang K. Concepts and conceptualization. *The Routledge handbook of cognitive linguistics*. Wen X., Taylor J. R. (eds.). New York: Routledge, 2021, 255–267.
- Karlsson 2018 — Karlsson F. *Finnish. A comprehensive grammar*. London: Routledge, 2018.
- Klavan 2021 — Klavan J. The alternation between exterior locative cases and postpositions in Estonian web texts. *ESUKA — JEFUL*, 2021, 12: 153–188.
- Klavan et al. 2011 — Klavan J., Kesküla K., Ojava L. The division of labour between synonymous locative cases and adpositions. The Estonian adessive and the adposition *peal* ‘on’. *Case, animacy and semantic roles*. Kittilä S., Västi K., Ylikoski J. (eds.). Amsterdam: John Benjamins, 2011, 113–134.
- Klavan, Veismann 2017 — Klavan J., Veismann A. Are corpus-based predictions mirrored in the preferential choices and ratings on native speakers. Predicting the alternation between the Estonian adessive case and the adposition *peal* ‘on’. *ESUKA — JEFUL*, 2017, 8: 59–91.
- Koivunen, Erkkilä 2022 — Koivunen T., Erkkilä R. Dataset of CONTAINMENT and SUPPORT in the Uralic languages of the Volga–Kama area. <https://doi.org/10.5281/zenodo.7062965>.
- Lakusta et al. 2021 — Lakusta L., Hussein Y., Wodzinski A., Landau B. The privileging of ‘support-from-below’ in early spatial language acquisition. *Infant Behavior and Development*, 2021, 65: 1–10.
- Landau 2019 — Landau B. Space in semantics and cognition. *Semantics. Typology, diachrony and processing*. Von Heusinger K., Maienborn C., Portner P. (eds.). Berlin: De Gruyter Mouton, 2019, 341–365.

- Landau 2020 — Landau B. Learning simple spatial terms. Core and more. *Topics in Cognitive Science*, 2020, 12: 91–114.
- Landau et al. 2017 — Landau B., Johannes K., Skordos D., Papafragou A. Containment and support. Core and complexity in spatial language learning. *Cognitive Science*, 2017, 41: 748–779.
- Langacker 1987 — Langacker R. W. *Foundations of Cognitive Grammar*. Vol. I: *Theoretical prerequisites*. Stanford: Stanford Univ. Press, 1987.
- Levinson 2003 — Levinson S. C. *Space in language and cognition. Explorations in cognitive diversity*. Cambridge: Cambridge Univ. Press, 2003.
- Levinson 2006 — Levinson S. C. The language of space in Yéli Dnye. *Grammars of space. Explorations in cognitive diversity*. Levinson S. C., Wilkins D. P. (eds.). Cambridge: Cambridge Univ. Press, 2006, 157–205.
- Levinson, Wilkins 2006a — Levinson S. C., Wilkins D. P. The background to the study of the language of space. *Grammars of space. Explorations in cognitive diversity*. Levinson S. C., Wilkins D. P. (eds.). Cambridge: Cambridge Univ. Press, 2006, 1–23.
- Levinson, Wilkins 2006b — Levinson S. C., Wilkins D. P. (eds.). *Grammars of space. Explorations in cognitive diversity*. Cambridge: Cambridge Univ. Press, 2006.
- Levinson, Wilkins 2006c — Levinson S. C., Wilkins D. P. Patterns in the data. Towards a semantic typology of spatial description. *Grammars of space. Explorations in cognitive diversity*. Levinson S. C., Wilkins D. P. (eds.). Cambridge: Cambridge Univ. Press, 2006, 512–552.
- Lewandowska-Tomaszczyk 2007 — Lewandowska-Tomaszczyk B. Polysemy, prototypes, and radial categories. *The Oxford handbook of cognitive linguistics*. Geeraerts D., Cuyckens H. (eds.). Oxford: Oxford Univ. Press, 2007, 139–169.
- Marko (no year) — *MARKO — Corpus of Mari language*. Univ. of Turku, (no year). <https://finno-ugric-corpora.utu.fi/cqpweb/>.
- MokshEr 2010 — MokshEr V. 3. *Mokšan ja ersän sähköinen korpus*. Turun yliopisto, 2010. <https://finno-ugric-corpora.utu.fi/cqpweb/>.
- Munnich et al. 2001 — Munnich E., Landau B., Doshier B. A. Spatial language and spatial representation. A cross-linguistic comparison. *Cognition*, 2001, 81: 171–207.
- Ojutkangas 2008 — Ojutkangas K. Mihin suomessa tarvitaan *sisä*-grammeja [What are the *sisä* ‘in’ grams are needed for in Finnish]. *Virittäjä*, 2008, 112: 382–400.
- Oncyko 2000 — *Oncyko corpus*. Univ. of Turku, 2000.
- Permyak 2009 — *Turku Komi-Permyak Corpus*. Univ. of Turku, 2009. <https://finno-ugric-corpora.utu.fi/cqpweb/>.
- Rosch 1975 — Rosch E. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 1975, 104: 192–233.
- Saarinén 2022 — Saarinén S. Mari. *The Oxford guide to Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 432–470.
- Schmid 2014 — Schmid H.-J. Lexico-grammatical patterns, pragmatic associations and discourse frequency. *Constructions collocations patterns*. Herbst T., Schmid H.-J., Faulhaber S. (eds.). Berlin: De Gruyter Mouton, 2014, 239–293.
- Talmy 2000 — Talmy L. *Towards a cognitive semantics*. Vol. 1: *Concept structuring systems*. Cambridge (MA): MIT Press, 2000.
- Taylor 2019 — Taylor J. R. Prototype effects in grammar. *Cognitive linguistics. Key topics*. Dąbrowska E., Divjak D. (eds.). Berlin: De Gruyter Mouton, 2019, 127–147.
- Tyler, Evans 2003 — Tyler A., Evans V. *The semantics of English prepositions. Spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge Univ. Press, 2003.
- Usacheva 2011 — Usacheva M. *Lokativnye padezhi v sostave grupp s prostranstvennym znacheniem v permskikhazykakh*. Candidate diss., Moscow State Univ., 2011. http://beserman.ru/publications/usacheva_diss_thesis.pdf.
- Vallejos, Brown 2021 — Vallejos R., Brown H. L. Locative construals. Topology, posture, disposition, and perspective in Secoya and beyond. *Cognitive Linguistics*, 2021, 32: 251–286.
- Zlatev 2007 — Zlatev J. Spatial semantics. *The Oxford handbook of cognitive linguistics*. Geeraerts D., Cuyckens H. (eds.). Oxford: Oxford Univ. Press, 2007, 318–350.

How to distinguish between semantically close cases: A case study of Mordvin illative and lative

© 2022

Riku Erkkilä

Ludwig-Maximilians-Universität München, Munich, Germany; University of Helsinki,
Helsinki, Finland; riku.erkkila@helsinki.fi

Abstract: In this paper I discuss the distribution of GOAL-cases in Erzya and Moksha (Mordvin, Uralic). Both languages have two productive GOAL-cases, namely illative and lative, but only one productive case for every other spatial relation (LOCATION, SOURCE, PATH). Furthermore, the two GOAL-cases do not seem to exhibit any clear division of labor (e.g., marking inner and outer GOALS). Previous studies have proposed different solutions, but they do not cover all the occurring variation, and thus are not general enough. In my study I propose that the best way to account for this phenomenon is to investigate the similarities in the language-external situations that are coded with each case. To accomplish this, I adopt the Cognitive Linguistics framework, and look at the conceptualizations of situations where each case is used. I assess the types of nouns used as Landmarks and divide them into eight categories: 0D objects, 2D bounded and unbounded entities, 3D bounded and unbounded entities, institutions, abstract entities, and temporal entities. The bulk of this paper treats the first five categories, i.e. spatial relations. I show that there are clear preferences of coding GOAL with one or the other case when the Landmark represents a certain category. Then I show that similar preferences are at work when coding GOALS with other types of Landmarks as well. After establishing these preferences, I go on to propose a solution to the variation of GOAL-marking in Mordvin languages based on the specificity of the relation between the Trajector and the Landmark.

Keywords: case, cognitive linguistics, illative, lative, Mordvin, spatial semantics, Uralic

Acknowledgements: I would like to thank Professor Elena Skribnik and an anonymous reviewer for their valuable comments on the manuscript of this paper.

For citation: Erkkilä R. How to distinguish between semantically close cases: A case study of Mordvin illative and lative. *Voprosy Jazykoznanija*, 2022, 5: 86–107.

DOI: 10.31857/0373-658X.2022.5.86-107

О разграничении семантически близких падежей: латив и иллатив в мордовских языках

Рику Эрркиля

Мюнхенский университет имени Людвиг и Максимилиана, Мюнхен, Германия;
Хельсинкский университет, Хельсинки, Финляндия; riku.erkkila@helsinki.fi

Аннотация: В статье обсуждается дистрибуция направительных падежей в эрзянском и мокшанском языках (мордовская ветвь уральской семьи). В обоих языках имеется два продуктивных направительных падежа — латив и иллатив — при том, что для прочих пространственных ролей (Места, Источника, Маршрута) существует лишь по одному падежу. Между двумя направительными падежами в мордовских языках не наблюдается четкого распределения функций (такого, как маркирование внешнеместных и внутреннеместных отношений). В предшествующих исследованиях предлагались различные подходы к лативу и иллативу, но они не покрывают всей наблюдаемой вариативности, в связи с чем их можно признать недостаточно обобщенными.

В настоящем исследовании я подхожу к проблеме, опираясь на анализ сходств между внеязыковыми ситуациями, которые кодируются тем или иным падежом. Для этого я использую аппарат когнитивной лингвистики и рассматриваю различные варианты концептуализации ситуаций, для которых используется тот или иной направительный падеж. Сущности, играющие роль Ориентира в пространственной ситуации, я подразделяю на восемь классов: «нульмерные» (точные) объекты, ограниченные и неограниченные двумерные объекты, ограниченные и неограниченные трехмерные объекты, учреждения, абстрактные сущности и временные сущности. Основная часть статьи посвящена рассмотрению первых пяти классов, то есть чисто пространственных Ориентиров. Я показываю, что разные типы пространственных Ориентиров явным образом обнаруживают разные предпочтения относительно маркирования тем или иным направительным падежом. Далее я демонстрирую, что аналогичные предпочтения имеют место и для остальных типов Ориентиров. Наконец, затем я предлагаю подход к вариативности маркирования Конечной точки в мордовских языках, основанный на специфичности отношения между Траектором и Ориентиром.

Ключевые слова: иллатив, когнитивная лингвистика, латив, мордовские языки, падеж, пространственная семантика, уральские языки

Для цитирования: Erkkilä R. How to distinguish between semantically close cases: A case study of Mordvin illative and lative. *Voprosy Jazykoznanija*, 2022, 5: 86–107.

DOI: 10.31857/0373-658X.2022.5.86-107

1. Background

In this paper I will address an interesting feature of Mordvin languages (a branch of the Uralic family), namely the fact that both languages of that group, Erzya and Moksha, exhibit two spatial cases expressing GOAL in otherwise symmetrical spatial case systems. One of these, lative, belongs to an older case system, whereas the other, illative, is a newer development. There seem not to be an evident distinction made with these cases¹. Previous studies have addressed this phenomenon only partially, as they have failed to consider all the parameters at work in the variation the GOAL-cases exhibit.

I will propose that the conceptualization of the relation between the Trajector (TR) and the Landmark (LM)² is the key to understanding the variation between the two GOAL-cases. In the analysis, I focus on the concrete spatial LMs, as these are the core area of spatial expressions. However, I will briefly comment on other types of LMs as well, as these make up a significant portion of my data. Usage of abstract LMs and abstract meanings expressed by spatial cases are semantic extensions from the concrete spatial usage. In a wider perspective, this paper argues that aspects of human cognition, especially conceptualization, can affect the case system, especially semantic cases, of languages, and these aspects should be considered when describing case systems of case-rich languages.

The data for this study is collected from a corpus, but the study is not intended to be a quantitative one. Rather, it is based on qualitative analysis of the data, identification of recurring patterns, and explaining them with the help of different theories of the Cognitive Linguistics framework. In addition, findings of cognitive psychology pertaining to the perception and processing of spatial scenes are used in the identification and analysis of patterns. The corpus is used as a convenient way to get access to large amount of natural language data. The study can be considered as a first step in a multimethod investigation, where natural language data is first analyzed qualitatively, then quantitatively, and finally the analysis is verified by forced choice or acceptability rating tests with native speakers [Arppe, Järvikivi 2007].

¹ Compare e. g. Finnish, a distant relative of Mordvin languages, where there are two series of spatial cases (so-called inner and outer local cases) that express a distinction between inner and outer/proximal LOCATION (see e. g. Karlsson 2018: 228–229).

² The concepts of Trajector and Landmark are discussed in Section 2.1.

The structure of the paper is as follows: In Section 1, I will discuss the structure of the system of spatial expressions in Mordvin languages, and the previous studies that have dealt with the problem at hand. In Section 2, I outline the theoretical framework and methodology used in this study, as well as discuss the data and its restrictions. In Section 3, I analyze the spatial LM types found in my data, and in Section 4, I discuss the non-spatial LM types. In Section 5, I briefly touch upon some additional evidence for my viewpoint based on the grammatical system of Mordvin languages, and finally in Section 6, I draw some conclusions.

1.1. The system of spatial expressions in Mordvin languages

Mordvin languages have three declensions, traditionally called basic, possessive, and definite declension. The basic declension has a spatial case system which makes a four-way distinction of spatial relations. Each of the relations is expressed by one case except for the GOAL-relation, which has two distinct cases expressing it. In the other declensions goal-marking is a bit different. In possessive declension the lative is lacking, and in the definite declension the only case to express the GOAL-relation is the dative.³ The system is shown in Table 1.

Table 1

Productive Mordvin spatial case system in the basic declension

	[-TO]	[+TO]
[-FROM]	inessive	illative / lative
[+FROM]	elative	prolative

The classification is based on [Ylikoski 2015].

The situation apparently derives from an older system of spatial cases which is reflected in the inflection of some relational nouns. The older system lacked inessive, elative, and illative. Instead, the LOCATION was expressed by “locative” (morphologically just the vowel following the stem [Bartens 1999: 163–164])⁴, and SOURCE by ablative, which survives in the modern Mordvin languages as a partitive case. Original GOAL-case was the lative, and PATH-case the prolative, which both survive in the productive case inflection of Mordvin languages (see, e.g., [Bartens 1999: 75–77; Serebrennikov 1967: 19–26]). However, as there are two GOAL-cases, they must differ in their meaning somehow. This is consistent with the Cognitive Linguistics’ view of language (e.g., [Hamawand 2021: 243–245]). In addition, this view is supported by the fact that the older SOURCE-case has gotten a new meaning, and the older LOCATION-case is confined only to the inflection of some relational nouns. If the cases of the old and new paradigm were synonymous one should also expect variation in the marking of SOURCE and LOCATION.

The current system probably evolved due to the restructuring of the Mordvin spatial case paradigm. Of the old spatial cases at least the lative, ablative, and prolative go back to Proto-Uralic [Bartens 1999: 75–76]. In some stage of the Uralic languages the Western group developed a new set of spatial cases, the so-called *s*-cases [Ylikoski 2016], which have reflexes in all the present-day Western Uralic languages (Finnic, Saami, and Mordvin). This seems to have led to the marginalization of the old case paradigm in Mordvin languages. The “locative” stopped to be

³ In Erzya plural definite plural illative is sometimes also possible. However, this phenomenon is not discussed further here. In Moksha, all spatial cases in the definite declension are expressed by periphrastic constructions.

⁴ Serebrennikov’s [1967: 32–33] view seems to differ from the analysis of Bartens. The crucial point is that previously there was no inessive, but some other locative case that is no longer productive.

productive anywhere but in the paradigm of certain relational nouns, and ablative turned into a partitive in content noun inflection. The meaning or distribution of lative did not suffer such drastic losses, but it nonetheless became more marginal in content noun inflection⁵. However, it stayed productive with relational nouns. This, in turn, led to a situation where the most frequent model for the use of lative was to express movement to or towards an unbounded region relative to a LM. This vagueness has thereafter affected the meaning of lative. This development would have made possible the contrast between the two GOAL-cases without a secondary distinction expressed by two spatial case series.

Even though both GOAL-cases are part of the productive case inflection, they do have somewhat different contexts of occurrence. The most notable difference is that relational nouns in Mordvin can be divided in two groups. One group uses exclusively lative as its GOAL-case. Such relational noun is, e.g., *alo-* (E)⁶, *al-* (M) ‘underside’. The other category usually takes illative as its GOAL-case, but lative is also possible. Relational noun *lango* (E), *langa* (M) ‘top’ is an example of this group [Bartens 1999: 163–164]. Secondly, there seems to be a division of the preferred GOAL-case according to the semantics of the referent of the inflected content noun. According to the established view (cf. Section 1.2), content nouns taking typically, or even exclusively, lative as their GOAL-case are mostly geographical names (e.g., names of cities), institutions, and to a smaller extent, two-dimensional entities (e.g., *paks’a* (E), *paks’æ* (M) ‘field’; on two-dimensional entities, cf. Section 3), whereas illative is preferred with other kinds of entities (e.g., abstract things). This division is not absolute, however, but represents a strong tendency of preferences.

1.2. Previous studies

The difference of illative and lative has been described differently in the literature. The oldest grammars that discuss the meanings of cases [Ahlquist 1861 (M); Wiedemann 1865 (E)] have the opposite stance towards the description of the GOAL-cases: Ahlquist [1861: 18–20] gives both cases separate meanings, for lative that of GENERAL GOAL, and for illative INNER GOAL. On the other hand, Wiedemann [1865: 45–46] denies lative the status of case altogether. In his view, lative ending is just a suffix deriving relational nouns and adverbs with the meaning of direction towards the LM.

In the grammars of the Soviet era, the stance towards lative again differs a lot. In the oldest grammar of this period [Evseviev 1928], the lative is not mentioned at all. In the two following grammars [Kolyadenkov 1954; 1959], the lative is discussed only in connection with relational nouns, i.e. it is not regarded as a part of the productive inflectional paradigm. The grammar [Kolyadenkov, Zavodova 1962] views the lative as a variant (allomorph) of the illative that highlights the meaning of OUTER GOAL in contrast to the illative proper, which emphasizes the meaning of INNER GOAL [Ibid.: 145–146]. However, they also say that lative can express INNER GOAL as well, and that the predicate has the biggest effect on the choice between the GOAL-cases. Tsygankin [1980: 170–172] includes both illative and lative as cases with separate meanings in the nominal paradigm.

Alhoniemi [1985] finds three varying parameters in the use of the GOAL-cases in Moksha. They differ syntactically, as nouns in the lative can take a demonstrative attribute, and nouns in the illative can take genitive and possessive attributes. The usual referents of the inflected nouns for each case also differ: nouns in the lative usually denote buildings, institutions, nature places, objects, or directions; nouns in illative usually have other kinds of referents. The meanings

⁵ The survival of the older spatial case paradigm in relational nouns can probably be attributed to the routinization of constructions where relational nouns were frequently inflected in spatial cases. Routinized constructions tend to resist change in languages. [cf. Schmid 2014: 242–247.]

⁶ The abbreviations E (Erzya) and M (Moksha) are used to refer to the languages.

of predicates with which the GOAL-cases appear also differ. Bartens [1996] shows that illative and lative are used to express different semantic roles in Mordvin languages.

The grammars [Tsygankin et al. 2000: 85–86] (E) and [Alyamkin 2000: 67–68] (M) show that the lative has fewer meanings, which are more concrete than those of the illative. A recent description of Moksha local cases [Kozlov 2018: 167–175] attests three differences between the GOAL-cases: firstly, they appear partly in different functions (compared to lative, illative has additional non-spatial meanings); secondly, the cases prefer different types of LM entities in spatial meanings; and thirdly, the frequency of the use of LM noun as a location affects the choice between the GOAL-cases (lative is preferred with LMs that denote typical locations while illative is preferred with other kinds of nouns).

Thus, the previous literature argues that there are two major differences between the GOAL-cases in Mordvin: the number of meanings attested, and the preference of different types of LM entities. However, these claims are not entirely true, as many of the meanings attested only with illative in the previous literature are possible, albeit rare, also with lative. The claim that the two cases favor different kinds of LM entities is, on the other hand, an undeniable truth. However, it has not yet been explained satisfactorily.

2. Theory, methodology, and data

In this section, the most important theoretical concepts as well as the methodology used to analyze the data are explained. In addition, I will discuss the data used in this study.

2.1. Theoretical prerequisites

The theoretical background I use in this paper is that of Cognitive Linguistics in general. In Cognitive Linguistics, the spatial elements (e.g., spatial cases) are seen to express atemporal relations between a Trajector (TR) and a Landmark (LM) [Leino 1993; Svorou 2007]. This means that the meaning or semantic content of the case is a certain relation that holds between two entities. TR is usually the more active participant of the relation that is located in some place with the help of a more stable entity, the LM⁷ [Langacker 1987: 231–243]. This means that what is at issue when studying semantics of spatial cases is the type of relation between the TR and the LM. This relation can be of different types: it can describe different configurations in space, or in other cognitive domains, like time, possession, or causation.

One important notion for the current study is conceptualization. Conceptualization is the process that combines the mind-external world, the mental capabilities of human, and language to yield concepts [Jiang, Yang 2021: 258–263]. Concept is not a mirror image of the external world, but rather a mental image filtered through the perception, previous experiences, and communicative needs of the conceptualizer, i.e. they are embodied [Ibid.: 257–258]. Conceptualization can highlight and background different elements of the real-world situation according to the needs of the conceptualizer, e.g., when trying to communicate a meaning to someone else [Ibid. 2021: 260]. In such case, the conceptualizer chooses the aspects of the situation at hand most important from her / his point of view and selects the most appropriate linguistic material to convey the information. What is most important depends on what part of a situation, perspective etc. the conceptualizer wants to express, and what is most appropriate depends on the conventions of the language used.

⁷ Morphosyntactically LM is the word inflected with the case, the content noun in a relational noun construction, or unmarked when a relational noun is used as a plain adverbial.

Conventionalized conceptualizations are entrenched in the mind of the language user, i.e. it is stored in her/his long-term memory. Entrenched linguistic units are routinized to express a certain linguistic meaning in a language community. This routinization is partly due to the relative frequency of use of a linguistic structure to conceptualize similar extralinguistic situations in comparison to other possibilities [Schmid 2014: 274–279]. A concept closely tied to entrenchment in Cognitive Linguistics is salience. Salience is the phenomenon where an entity attracts the attention of a conceptualizer. Different kinds of salience can be distinguished, but the most important for current discussion is pragmatic salience. Pragmatic salience is salience of some linguistic element in a certain context, i.e. a certain linguistic element is preferred over other options as an expression of a certain extralinguistic situation. Repeated use of a pragmatically salient expression in similar contexts can lead to pragmatic entrenchment of the expression as the primary way of expressing the situation. This can override general frequency effects that affect entrenchment [Ibid.].

An important notion closely tied with that of conceptualization is categorization, i.e. the process of comparing entities⁸ and mentally assembling them as groups based on their similarity, which is one of the most basic human cognitive processes. The entities can be real-world objects, speech sounds, linguistic meanings etc. Categorization is a part of the conceptualization process, where the conceptualizer includes the entity conceptualized into the category fitting best in her/his opinion. There are many theories of how categorization works in the human cognition [Wen, Fu 2021: 174–181], but in this paper categorization is viewed to be based on prototypicality of entities. This view is based on the studies by Eleanor Rosch (e.g., [1978]), where she proved that categories have a central member (a prototype) and other members that have something in common with the prototype but differ from it in other respects. Later, prototype theory has been shown to be applicable to all linguistic categories [Taylor 2003: 170]. Categorization according to prototypes is based, in turn, on the human ability to recognize common patterns in entities. When a person encounters a new situation, s/he compares it to previous encounters, and assigns it to a category whose properties most closely resemble those of the newly encountered entity.

2.2. Methodology of analysis

The focus of this paper is the relation between the TR and the LM in spatial expressions, so the analysis is mostly confined to them. In this study, I am analyzing what kind of LMs take each GOAL-case in Mordvin languages. This, in turn, reveals the type of relation between the TR and the LM. This matter is discussed in Section 3.3. What is at stake in this study is not the behavior of certain parts of speech, but how semantic content is expressed by GOAL-cases in Mordvin. Therefore, the analysis pertains to all linguistic elements that show productive inflection in GOAL-cases.

The first step in the analysis was to determine what forms to analyze. The choice was made to confine the analysis to productive inflection, as there the choices of the conceptualizer can be seen most easily. In productive inflection, the conceptualizer chooses a lexical form and a grammatical morpheme from the inventory of the language and combines them to produce an expression that, in her/his mind, describes best the situation s/he wants to communicate. Here, we assume that a given form is an instance of productive inflection in two alternative cases:

1. The inflected word is a content noun. Content nouns are considered to productively inflect in all cases of the paradigm.
2. The inflected word is a relational noun and can be inflected in one of the spatial case series discussed in Section 1.1, i.e. inessive, elative, and illative, or “locative” (a vowel following the stem [Bartens 1999: 164–165]), ablative, and lative.

⁸ Entity is a cover term used in Cognitive Linguistics for all things and relations [Langacker 1987: 198].

According to the principle of entrenchment (see, e.g., [Schmid 2010; 2014]), linguistic elements that are typically used in the same construction are probably stored as chunks (cf. [Schmid 2014: 249–250]) in the mind of the language user. At the same time such repeated constructions are conventionalized to express a certain meaning in the speech community [Ibid.: 250–251]. This applies especially to the relational nouns, which take frequently same spatial cases to express same meanings. However, from a Cognitive Linguistics point of view it is perfectly acceptable to study these kinds of elements alongside other inflection. There are two reasons to this: First, it is probable that other typical collocations of word and case, including content noun inflection, are also stored as chunks (cf. [Cappelle et al. 2010: 198–200]), and as such there is no clear limit of what should and what should not be considered “productive” inflection in the traditional sense. Secondly, the element in the end of a relational noun that varies paradigmatically (the case morpheme) has the same semantics as its identical counterpart in content noun inflection. Therefore, these elements can be compared regardless of the part-of-speech status of the inflected form, i.e. whether it is a content noun, relational noun, postposition, or adverb.

Secondly, LMs were grouped according to the following principles. The referents of the spatial LMs are divided into five categories according to their properties. These categories are marked by two parameters: boundedness and number of dimensions. There are three degrees of dimension, namely zero, two, and three. Dimension here means the functional dimensions of a LM entity, i.e. a field is two-dimensional even though the growth on it essentially adds a third dimension to the referent. This is especially important to keep in mind with zero-dimensional referents, e.g., *pango* (E), *panga* (M)⁹. These are of course not strictly zero-dimensional, but they are conceptualized as a whole, so they have only one functional dimension. Boundedness, on the other hand, tells if the LM noun has an exact border between its inside and outside or not. Here again it is a case about functional boundedness, so, e.g., a forest is bounded at least to some degree, even though it would slowly fade to the countryside. One prominent category of unbounded entities in my data are LMs coded with relational nouns, as relational nouns only define an area in relation to the referent of LM [Carlson 2010]. Because of the functionality, all zero-dimensional LM entities are automatically considered bounded.

In addition, three further categories, namely institution, abstract entity, and temporal entity are posited. These differ from the five other categories in that they do not exhibit dimensions, and their boundedness is debatable. This is because they are not strictly speaking spatial entities, but rather entities that are conceptualized with metaphorical spatial properties. In addition, institutions are differentiated from abstract entities and temporal entities by the fact that they designate societal, political, legal etc. forces. These can have a discrete place in space, but this place, contrary to actual spatial entities, does not define them.

The division is based on common human cognitive properties. The most important is categorizing entities according to prototypes, as discussed in Section 2.1. The division of spatial entities into two- and three-dimensional is based on the spatial prototypes of CONTAINMENT and SUPPORT-FROM-BELOW [Landau 2020]. These prototypes are attested even in children that cannot yet produce speech [Lakusta et al. 2021; Landau 2019: 350–353]. This means that they are pre-linguistic and thus should be present in the cognitive faculty of humans regardless of their language.

CONTAINMENT involves a LM that is (conceptualized as) a CONTAINER, which prototypically is three-dimensional, and consists of an inside, an outside, and a solid boundary separating them. Other common properties of a CONTAINER are the capacity of holding TR and restricting its movements, and the occlusion of TR when the configuration is viewed from outside. SUPPORT-FROM-BELOW, on the other hand, involves a LM that is (conceptualized as) a two-dimensional PLANE capable of shielding TR from the effect of gravity, i.e. stopping TR from falling (cf. [Landau et al. 2017]).

⁹ This type of Landmarks could also be called e.g., “point-like”. The term “zero-dimensional” is used here because it is in line with the terms used of other spatial Landmarks.

Boundedness is based on the human cognitive ability to conceptualize entities as bounded or unbounded. Unbounded entities are conceptualized as continuing indefinitely, whereas bounded entities are demarcated as discrete entities separate from their surroundings. Bounded entities usually also include a boundary that clearly separate them from other entities surrounding them [Talmy 2000: 50–55].

The spatial prototypes of CONTAINMENT and SUPPORT-FROM-BELOW together with the concept of boundedness yield the categories of 2D and 3D bounded and unbounded LMs. 0D objects are a residual class of spatial entities that are not conceptualized as CONTAINERS or PLANES because of the lack of required functional dimensions. Non-spatial entities are categorized as abstract or temporal entities depending on their semantics. Finally, institutions are sort of an a posteriori category, which is posited because certain LM entities that are partly spatial and partly abstract exhibit uniform behavior in GOAL-marking. This analysis yields the categories shown in Table 2.

Table 2

Landmark types used in the analysis of data

	Prototypical LM	Non-prototypical LM
0D object	<i>pango</i> ‘mushroom’ (E)	<i>s’el’ved’i</i> ‘tear’ (E)
2D bounded	<i>paks’æ</i> ‘field’ (E)	<i>kart’fa langa</i> ‘top of stick’ (M)
2D unbounded	<i>vi’f’an’æ langa</i> ‘top of grass’ (M)	<i>front</i> ‘front’ (E)
3D bounded	<i>kud</i> ‘house’ (M)	<i>vel’an’i pe</i> ‘end of village’ (M)
3D unbounded	<i>men’al’i langa</i> ‘top of sky’ (M)	<i>ver’ie</i> ‘above’ (E)
institution	<i>fkola</i> ‘school’ (E)	—
abstract entity	<i>vij</i> ‘strength’ (E)	—
temporal	<i>i’l’æd’i</i> ‘evening’ (M)	—

For spatial LMs (first five rows) an example of all the prototypicality categories is shown. The prototypicality scores of this study do not apply to non-spatial LMs (three rows in the bottom).¹⁰ For further examples see [Erkkilä 2022].

The categories posited here are not absolute but represent prototypes. This naturally leads to the fact that there are better and worse examples of the prototypes in the data, i.e. every LM referent conforms to the prototype to a degree. However, this depicts reality well, as conceptualization of situations include highlighting some properties and finding similarities between known and novel described scenes. To underline this, every example in the data [Erkkilä 2022] is given a score from 1 to 4 describing how prototypical they are for their LM type. 4 marks the most prototypical cases, whereas 1 means that the LM entity is only marginally representative of its category. However, this division might be too fine-grained for the type of data.¹¹ Therefore, in this paper the categories are collapsed to prototypical (prototypicality scores 3 and 4) and non-prototypical (prototypicality scores 1 and 2).

The prototypicality categories are based on general capacities and properties of human cognition. The first is the capacity of perceive similarities in objects and spatial scenes and group these entities together according to them [Evans 2010: 29–36]. The second is the structuring of categories according to prototypicality. This means that categories of entities have more representative members that are more central to the category, and less representative members that

¹⁰ These types of LMs do, however, exhibit a prototype structure like all categories. To give the reader an idea of what kind of entities the categories consist of, I have included examples of prototypical representatives to the table.

¹¹ This was brought to my attention by an anonymous reviewer, for which I am thankful.

are more peripheral. However, all members of one category have some common properties that makes it possible for the conceptualizer to recognize that central and peripheral members belong to the same category [Rosch 1975; 1978]. These capacities are pre-linguistics and therefore they work in the same way in all humans. Categorization is affected by physical and social interactions of humans [Lakoff 1987: 5–154], and therefore they may be language or culture specific. In this study all posited categories except institutions are based on spatiotemporal (physical) properties of entities, so the effect of language or culture is minimal.

It is worth noting, that conceptualization of LM entities is not fixed, but done in each situation separately. As mentioned in Section 2.1, the conceptualizer construes the conceptualization of the situation to fit the message s/he is trying to convey. This means that, for example, in a situation where approaching a mushroom is conceptualized from the point of view of a fly (e.g., in a fairytale), the LM type of the mushroom could very well be 2D bounded. However, as the analysis in Section 3 shows, there are strong tendencies to conceptualize certain entities consistently in the same way. This effect is based on the entrenchment of linguistic units in the minds of language users and the conventionalization of certain units in the language community to express certain meanings. These conventionalizations are called pragmatic associations by Schmid [2014: 251–254], and they are used to tie syntagmatic associations [Ibid.: 248–251] to extralinguistic situations. Pragmatic associations are conventionalized ways of referring to similar extralinguistic situations with one and the same linguistic elements. This allows us to group different extralinguistic situations into bigger categories.

2.3. The data

The data is collected from the MokshEr V. 3 corpus [MokshEr 2010] created at the Research unit of Volgaic languages at University of Turku.¹² The corpus consists mainly of newspaper texts in both languages, but there is also prose and poetry, and as such represents mainly the literary variant of each language. For the Moksha data, the whole Moksha subcorpus was used, but for Erzya only the subcorpus containing the periodical “Sjatko” is used for the collection of data. The complete data set is freely published in Zenodo [Erkkilä 2022].

For each language, 200 examples of each case were collected, giving a grand total of 800 examples. The data was collected with the help of Antconc.¹³ Because the corpus is not annotated, only word-final allomorphs of the cases were considered, as the word-internal allomorphs of illative would have resulted in a huge number of false positives, and the raw data included a huge number of false positives as it was. In addition, some lexicalized adverbs were not considered, as they do not represent productive inflection. These adverbs include lexicalized forms such as *ods* (E, M); *odov* (E), *odu* (M) ‘again’ (originally illative and lative forms of *od* ‘new’), *kor'ias* (E) ‘corresponding, in comparison to’ (a lexicalized illative form of *kor'ia* ‘id.’). All forms that do represent productive inflection as defined in Section 2.2 are included in the data.

Corpus data is frequently used as the source of quantitative analysis, but in this research, it is used only as a convenient source of real-life language data. As is pointed out in different sources, e.g., [Arppe et al. 2010: 3–6; Luodonpää-Manni et al. 2017: 3–12; Newman 2010: 94], corpus data is a good source for qualitative analysis, as it represents real language use. In Cognitive Linguistics there has in recent times been a demand for actual language data (cf. [Divjak et al. 2016]), but using this kind of data does not necessarily entail using statistical methods.

¹² Presently the corpus is available online at <http://digilang.utu.fi/>, but a permit is necessary for its use.

¹³ Antconc [Anthony 2014] is a concordance tool which enables searches on corpus consisting of text files. This tool was used to collect the data, as when data collection for this study was made, the MokshEr corpus consisted only of text files without any search interface.

Qualitative analysis of corpus data can serve as a first step in theory forming, and such analysis can be later operationalized in quantitative research (cf. [Langacker 2016]). The role of corpus data in the present study is exactly this.

3. The distribution of GOAL-cases between spatial LMs

In this section I will look at the distribution of different kinds of spatial LM entities between the two cases under study. First, I will consider the two cases separately (Sections 3.1 and 3.2), and then I will draw conclusions of the distribution of GOAL-cases. The data is shown in Table 3.

Table 3

Occurrences of GOAL-cases with different types of LM entities

Language and case	Proto-typicality	0D object	2D bounded	2D unbounded	3D bounded	3D unbounded	institution	abstract entity	temporal	total
Erzya illative	prototypical	16	29	8	20	8	0	0	0	81
	non-prototypical	12	10	3	10	8	0	0	0	43
	N/A	0	0	0	0	0	5	62	9	76
	total	28	39	11	30	16	5	62	9	200
Erzya lative	prototypical	0	4	12	53	18	0	0	0	87
	non-prototypical	2	0	1	4	88	0	0	0	95
	N/A	0	0	0	0	0	16	2	0	18
	total	2	4	13	57	106	16	2	0	200
Moksha illative	prototypical	32	24	7	13	14	0	0	0	90
	non-prototypical	12	15	0	9	9	0	0	0	45
	N/A	0	0	0	0	0	1	54	10	65
	total	44	39	7	22	23	1	54	10	200
Moksha lative	prototypical	0	6	4	81	9	0	0	0	100
	non-prototypical	0	1	2	7	49	0	0	0	59
	N/A	0	0	0	0	0	32	7	2	41
	total	0	7	6	88	58	32	7	2	200

N/A = prototypicality scale is not applied in this study to the entity in question.

As the focus of this paper is on concrete spatial LMs, I will not discuss institutions, abstract entities, or temporal LMs in detail in this section. Some remarks on the basis of Table 3 are still in order. As can be seen, in Mordvin languages illative is overwhelmingly used to mark GOALS when the LM is abstract. There are 62 tokens of GOAL with abstract entities in Erzya, and 54 tokens in Moksha against two and seven tokens of lative, respectively. The same tendency applies

to temporal LMs, which have illative nine tokens in Erzya, and 10 tokens in Moksha against zero and two tokens with lative, respectively. The third non-spatial category of LM entities, institutions, show the opposite tendency of preference of GOAL-case. Lative as the GOAL-case of institutions has 16 tokens in Erzya, and 32 tokens in Moksha against five and one tokens, respectively. The connection of these tendencies to the use of GOAL-cases with spatial LMs will be discussed in Section 4.

The explanation put forward in [Kozlov 2018: 170] should be mentioned here. According to that study, the illative is the preferred GOAL-case in contexts where the function of a landmark in a spatial expression is not typical for the noun (“В случае если для существительного не характерна функция ориентира в пространственных выражениях [...]”). This is said to be based on the frequency of use of the LM noun in spatial contexts, i.e. illative is preferred for nouns that do not frequently express places. As non-spatial entities are not concrete places, this explanation seems plausible. However, from the point of view of Cognitive Linguistics frequency of a (meaning of a) lexeme does not alone predict the entrenchment of a construction (i.e. the occurrence of a certain lexeme with a certain GOAL-case). The frequency effects evident in language use result from entrenchment of linguistic chunks and the routinization of them as means of communicating a certain meaning in a language community [Schmid 2014: 242–247]. Especially important is to notice that with frequent lexemes and frequent constructions both the lexeme (i.e. its meaning) and the construction it frequently appears in can be separately entrenched. If this is the case, the frequency of the use of a GOAL-case with certain types of nouns does not predict the use of that GOAL-case with them, but rather is a symptom of entrenched conceptualizations of different extralinguistic situations.¹⁴

A similar explanation is given for differential place marking in Haspelmath [2019] (among others). Haspelmath argues that typologically languages can present asymmetrical marking in nouns expressing place. He presents a spatial reference scale of the following kind: human noun > common inanimate noun > place name / topo-noun [Ibid.: 323]. In this scale, nouns expressing meanings on the left are more prone to take a place marker (case or adposition) than those on the right, and the marker tends to be longer for the nouns on the right.

The usefulness of this typology is somewhat questionable in the present study, as we are dealing with symmetrical place marking in Mordvin languages. The GOAL-cases do not differ in their existence in the studied contexts, nor is one of them longer than the other. What is of interest for this study is the distribution of the cases among contexts, and especially the underlying conceptualizations affecting the distribution. In addition, the typological approach evokes frequency asymmetries and expectations derived from them [Haspelmath 2019: 329–330]. For typological research, form-frequency correspondence can yield fruitful generalizations, but in a study in the Cognitive Linguistics framework alternative explanations need to be put forward (see, e.g., [Schmid 2014: 274–279]). However, the observation that nouns denoting place entities and relational spaces [Haspelmath 2019: 322–324], on the one hand, and nouns denoting entities that are not typically conceived as places [Ibid.: 317–322], on the other, tend to behave in a certain way across languages is interesting. It seems to point to a typical way of conceptualization of such entities cross-linguistically.

3.1. Illative

In addition to the abstract LM entities discussed above, the illative is the dominant GOAL-case in combination with two types of LM entities. These are zero-dimensional and

¹⁴ In addition, the definition of a noun that does “not characteristically have the function of a landmark in a spatial expression” is vaguely defined and cannot reliably be studied as a factor explaining variation in the use of GOAL-cases.

bounded two-dimensional entities. 0D objects, as discussed in Section 2.2, are not literally zero-dimensional, but rather do not have specific functional dimensions. This means that the LM entity as a whole is meaningful for the relation between the TR and the LM at the end of the action expressed by the predicate. There are two main types of 0D objects in the data: small objects and humans. Human referents are usually LMs in clauses with a perceptual predicate and are used with the relational noun *lango* (E), *langa* (M) ‘top’. As targets of perceptive action such LMs do not have functional dimensions and are meaningful for the relation between the TR and the LM as a whole, i.e. are wholly perceived by the TR, as shown in (1).

(1) MOKSHA

(MokshEr-V.3/V.2 additions/Moksha/Jakshter_tyashtenya/JT-2005_5/JT-2005_5_14-15.txt)

lʲædi-ks-nʲæ izʲə-zʲ soda ruzə-nʲ kælʲ-tʲ,
 stay-NMLZ-NOM.PL.DEF NEG.PST1.3PL>3SG know.CNG Russian-GEN language-GEN.DEF
 a sʲas panʒa-də-nʲ kurk-t van-tʲsʲitʲ **general-snən**
 and therefore open-NMLZ-GEN mouth-NOM.PL look-PST1.3PL **general-GEN.POSS.3SG**

lang-s.**top-ILL**

‘The others didn’t know the Russian language, and therefore looked at their general with open mouths.’

2D bounded entities, on the other hand, possess two functional dimensions, i.e. they are planes. In prototypical case they are capable of supporting other object, cf. *fa* (M) ‘table’. In addition, as bounded entities they have a border separating them from their surrounding. Prototypically, the border is absolute, as in *eziem* (E) ‘bench’, where it is quite obvious what constitutes the entity. In less prototypical cases, such as *kuzʲtʲembe* (E) ‘yard’, the LM sort of fades to the surroundings. Unlike with 0D objects, however, it is not necessarily the case that the whole LM is important for the relation between the TR and the LM in the end of some action, but with prototypical LM entities of this category it is. Even in less prototypical cases the boundedness of the LM makes it easy to locate the TR in the LM, which makes the LM entity more salient in the relation. A prototypical instance of 2D bounded entity is shown in (2).

(2) ERZYA (MokshEr-V.3/V.2 additions/Erzya/Syatko/Syatko-2006_7/Syatko-2006_7_76-101.txt)

ketʲkin a pomnʲa-sʲi, koda lʲisʲ-sʲ uʲfo-v, molʲʲ-sʲ
 [name] NEG remember-PRS.3SG>3SG REL exit-PST1.3SG outside-LAT go-PST1.3SG
 oʲfo-nʲ sado-v, oza-sʲ **eziem lang-s.**
 city-GEN garden-LAT sit-PST1.3SG **bench top-ILL**

‘Kechkin does not remember how he went outside, went to the city park, [and] sat down on the bench.’

What is common to these two types of LM entities is that they can easily be distinguished from their surroundings in the language-external world. In most cases it is obvious where the LM entity ends and something else begins, and even in less prototypical cases only the border between LM entity and surroundings is fuzzy, the core area of LM entity is still clear in these cases. This also makes the relation between the TR and the LM unambiguous, as either the whole LM is affected by the action of the TR, or the conceptualizer can locate the TR inside the borders of the LM easily. This makes the LM, and the relation between the TR and the LM specific¹⁵ as discussed in Section 3.3.

It is also noteworthy that in both languages the entities belonging to these LM types are usually prototypical. The percentage of prototypical instances in the sample is as follows: 0D object

¹⁵ The terms ‘specific’ and ‘specificity’ are used a bit differently but comparably, i.e. to refer to the specificity of the relation between an expression and real-world situation, elsewhere [e.g. Hamawand 2021: 245–246].

57 % (E), 73 % (M); 2D bounded LMs 74 % (E), 62 % (M). This, too, can be seen as a manifestation of the tendency of favoring illative with specific LMs. Prototypical instances of 0D objects and 2D bounded LMs are more specific than non-prototypical because they manifest better the properties which define these categories, i.e. clear boundaries and distinctness from their surroundings.

The opposite situation is present with the 3D unbounded LMs, which do not readily take illative as their GOAL-case. Prototypically, a 3D unbounded entity is a volume without clear borders to distinguish it from its surrounding. An example of a 3D unbounded LM is *men'əlv'i s'er'i* (M) 'height of sky'. This kind of LM cannot be clearly separated from its surroundings, and thus do not allow the conceptualizer to locate the TR in it straightforwardly. In addition, in such cases the TR is always interacting only with a relatively small part of the LM. The most numerous group of 3D unbounded LMs is comprised of relational nouns, as they define an unbounded area in relation to the LM entity. From the point of view of conceptualization, both types of LMs (content nouns and relational nouns) are used to express similar entities, i.e. volumes that do not have clear boundaries and do not readily allow the conceptualizer to locate the TR [cf. Carlson 2010]. This, of course, is relative to some extent, but it is justified to study both types of nouns together instead of artificially carve up a coherent category based on human categorization (cf. Section 2.2 on the principles of categorization). A typical case of relational noun *ekf-* (E) 'backside' is shown in (3).

- (3) ERZYA (MokshEr-V.3/V.2 additions/Erzya/Syatko/Syatko-2004_7/Syatko-2004_7_32-47.txt)
 s'este min'ek jutk-sto kajav-s'i ikel'e-v ojam j'egor, joma-s'i
 then IPL.GEN middle-ELA rush-PST1.3SG front-LAT [name] vanish-PST1.3SG

katjamo ekf-s.

smoke backside-ILL

'Then from between us Oyam Yegor rushed forward, [and] vanished behind the smoke.'

It would be logical for 2D unbounded LMs to behave like 3D unbounded LMs, as they, too, denote an area hardly distinguishable from its surrounding. However, the data is too scarce with this kind of LMs to say anything about them. This is probably due to the fact that 2D unbounded entities are naturally quite rare.

It is also somewhat counterintuitive that 2D and 3D unbounded LMs also tend to be prototypical instances of their categories, as shown by the following percentages: 2D unbounded LMs 73 % (E), 100 % (M); 3D unbounded LMs 50 % (E), 61 % (M). However, the number of examples of both categories, and especially of 2D unbounded LMs is so small in the data that this cannot really be considered a significant result. Therefore, more research on the behavior of 2D unbounded entities would be in order. The results with the use of lative with 3D unbounded LMs discussed in Section 3.2 support this claim.

The previous discussion has pointed to the direction that illative is preferred when the LM is easily distinguishable, and the action of the TR affects the whole LM or a significant portion of it. In addition, LM should enable conceptualizer to locate the TR within it with ease. Considering these properties, it is a bit surprising that both in Erzya and Moksha lative is used more frequently with 3D bounded entities in my data. This discrepancy has a logical explanation though. The lative is used in both languages as the almost only GOAL-case with two groups of words, which both are categorized as 3D bounded in this study. These are words denoting dwellings, i.e. villages and cities, including their names, and the word *kudo* (E), *kud* (M) 'home'¹⁶. When these two groups are subtracted from the uses of lative, the distribution of GOAL-cases seems more as expected. The number of examples after the subtraction is shown in Table 4.

¹⁶ Note that the same noun *kudo* (E), *kud* (M) in the meaning 'house' takes the illative. The reason why 'home' takes the lative is discussed in Section 4.

Table 4

The distribution of GOAL-cases among different kinds of 3D bounded LMs

Landmark	Erzya		Moksha	
	illative	lative	illative	lative
'home'	0	22	0	19
cities and villages	1	19	117	26
other	29	16	21	43
total	30	57	22	88

The category "other" comprises all the LMs that are not cities, villages, or the word 'home'.

From Table 4 it can be seen that illative in Erzya behaves as expected after the most numerous exceptions are cleaned away. However, illative in Moksha is still used less than lative as the GOAL-case with 3D bounded LMs. It is possible that the sampling process has omitted 3D bounded LMs with illative in the Moksha data, in which case the situation in the data could be seen as a sort of statistical artefact. However, if lative is actually used more than illative as the GOAL-case of 3D bounded LMs in Moksha, a special explanation would be needed for that. The uniformity of the use of the GOAL-cases in other contexts, both within Moksha and compared to Erzya, seems to point to the direction of a statistical artifact. Still, further study on the matter seems to be necessary.

3D bounded entities are prototypically CONTAINERS that can hold a TR (e.g., *par'n/æ* (M) 'barrel'). Other typical properties are concealment and restriction of movement of the TR. They enable the conceptualizer to locate the TR within the LM with high precision, as the LM has distinct boundaries and thus is separated from its surrounding. This makes the LM, and as a corollary the relation between the TR and the LM in the end of the action, specific. As with 0D objects and 2D bounded LMs, 3D bounded LMs also tend to exhibit the prototypical characteristics of their category. The relevant percentages are 67% (E), 59% (M).

3.2. Lative

The use of the lative shows a somewhat complementary picture compared to the use of the illative. The lative is used as the preferred GOAL-case with 3D unbounded LMs, and it is practically non-existent with 0D objects and 2D bounded LMs. In addition, the lative is used as the GOAL-case with 3D bounded LMs when the entity is a dwelling or has the meaning 'home'.

Most of the examples with 3D unbounded entities are explained with the fact that a relational noun that obligatorily takes the lative is used (see Section 1.1), but content nouns are present in the data as well. The use of the lative with this type of LMs seems to be the core area of the use of this case. This probably derives from the fact that many relational nouns cannot choose their GOAL-case freely but are always inflected in lative. Such relational nouns are, e.g., *al-* (E, M) 'underside'; *ikel-* (E), *ingäl-* (M) 'front'; *jon-* (E) 'direction'; *pel-* (E), *pæl-* (M) 'side, direction'; *firi-* (M) 'side'; *udal-* (E), *ftal-* (M) 'backside'; *veri-* (E), *væri-* (M) 'above'. This, then, is not a direct reflex of the effect of LM type to the choice of the GOAL-case. The use of a relational noun is a direct reflex of the conceptualization of a certain situation, though. Therefore, it is justified to compare relational nouns with content nouns in a study where the aim is to analyze the effects of conceptualization and categorization of extra-linguistic entities to variation in case marking. After all, relational nouns express relational areas, which are naturally unbounded [Carlson 2010], and therefore it is natural that they are categorized together with content nouns expressing unbounded entities.

In addition, relational nouns taking exclusively lative as their GOAL-case are a plausible origin of the variation in the present-day Mordvin languages. As mentioned in Section 1.1, the spatial case system of Mordvin languages was probably restructured at some point of development of the Western Uralic language group, before Proto-Mordvin separated from it. In this case, the relative frequency of relational nouns taking lative kept the case productive through the restructuring, as the relational nouns are probably stored as chunks consisting of stem and case marker in the minds of the speakers. The LM type relational nouns represent still function as the prototype of LMs preferring lative as their GOAL-case, but this prototype has spread to new contexts, e.g., to recent Russian loans [Alhoniemi 1985].

The prototypical 3D unbounded LM, as mentioned in Section 3.1, is an entity that has three functional dimensions and does not possess a clear boundary against its surroundings. In addition, with 3D unbounded LMs the relation between the TR and the LM is not exact, as either TR affects only a minor part of the LM or the TR is not readily located by the conceptualizer because of the fuzziness of the LM entity. Relational nouns conform to this prototype very well, as they define an area in relation to some entity. This area has usually three functional dimensions, and the acceptability of the relation slowly fades as the relation between the TR and the LM changes. This phenomenon has been studied by Laura Carlson [2010] with English prepositions, and as both prepositions and relational nouns express a relative area, it is safe to assume that the use of Mordvin relational nouns conform to similar restrictions. The only property differing from the prototype with relational nouns is that they do not refer to an entity, but to a relation. An example of prototypical 3D unbound LM with relational noun is *kardaz udal/-* (E) ‘backside of the pen’, and with a content noun *men’al/-* (M) ‘sky, heaven’.

This tendency also explains the lack of 0D objects and 2D bounded LMs inflected in lative. As the LM entity is highly distinct and TR either affects LM significantly or occupies a significant portion of it, 0D objects and 2D bounded LMs do not conform to the prototypical situations where lative is used as the GOAL-case. On the other hand, 2D unbounded LMs should take lative as their GOAL-case, but as mentioned in Section 3.1, the data is too scarce to verify this. The putative percentages of prototypical 2D unbounded entities taking lative as their GOAL-case in the data are 92 % (E), and 67 % (M).

The case of dwellings, briefly commented in Section 3.1, seems to pose a problem for the analysis presented here. Dwellings overwhelmingly marked with lative in my data, 95 % (E) and 96 % (M) of cases, even though they refer to quite prototypical 3D bounded entities. This problem is, however, only superficial. There are two main reasons for the use of lative with dwellings. First one is the fact also pointed out in [Kozlov 2018: 168–171], that the relation between the TR and the LM with dwelling as referent is highly generic, i.e. there is normally no other possible relation for a TR and a LM in these cases than a pure spatial relationship. This, then, allows the relation to be conceptualized as unspecific, even though the referent of the LM is specific. Second reason is the physical nature of the relation between the TR and the LM. When the endpoint of, e.g., movement of TR is a dwelling, TR does not interact with the whole LM when it arrives (except in a marginal case where the whole dwelling is only a few buildings), but rather occupies one place there and interacts with the LM in one place at a time. This is exemplified in 4, where the LM *Omskæj* ‘(city of) Omsk’ gives only a rather rough estimate of the whereabouts of TR.

(4) MOKSHA

(MokshEr-V.3/V.2 additions/Moksha/Moksha/Moksha-2006_4/Moksha-2006_4_49-57.txt)

tʲætʃəm-s af soda-fə-ks lʲadə-nd-i, koda kondurʃkin
this.day-ILL NEG know-PASS.PTCP-TRA stay-CONT-PRS.3SG REL [name]

POV-sʲ **omskæ-i**
end.up-PST1.3SG **Omsk-LAT**

‘Until this day it remains unknown how Kondurshkin ended up in Omsk.’

In (4), the relation between the TR and the LM is unspecific in the sense that we can only say that TR is inside the borders of LM, but not where it exactly is. This is parallel to the situation

where LM is 3D unbounded. The reason, then, why lative is used frequently with 3D bounded LMs is twofold: Firstly, the relation between the TR and the LM when LM is a dwelling is rather unspecific, and secondly, dwellings are categorized in this study as less prototypical 3D bounded entities (score 3), because they possess three functional dimensions and more or less sharp boundary between their inside and outside. This explanation does not cover the cases where *kudo* (E), *kud* (M) ‘home’ takes lative as its GOAL-case, though, but I will discuss this matter in Section 4 in greater detail.

In addition, Bernhardt [2019: 59–63] has shown that names of countries and natural formations (e.g., rivers) take almost exclusively lative as their GOAL-case. This also applies to all six examples (one in Erzya, five in Moksha) of names of countries found in my data. Countries are categorized as less prototypical 3D bounded entities in this study because they usually are not flat, and one can be in a country, e.g., flying over it. However, they are less prototypical because they do not possess a tangible, solid boundary that separates them from their surrounding. Countries, like dwellings, do not facilitate let the conceptualizer to easily locate the TR, nor does the TR interact with a significant part of the LM at one time. Therefore, the use of lative as the preferred GOAL-case of countries parallels that of cities.

The distribution of the lative, just like that of the illative shows sensibility to the specificity of either the LM entity, the relation between the TR and the LM, or both. In the case of the lative, however, the unspecific instances attract it as their GOAL-case.

3.3. The distribution of GOAL-cases across different types of spatial LMs

The distribution of GOAL-cases in Mordvin languages emerging from the discussion in previous sections is as follows. Illative is virtually the only GOAL-case used with 0D objects and 2D bound LMs. In addition, illative in Erzya is the primary GOAL-case with 3D bound entities when the special cases of LMs denoting dwellings and ‘home’ are discarded. In Moksha, lative dominates even after this operation with 3D bound LMs, but this state of affairs probably results from the overall small amount of 3D bounded LMs with illative in the data. The lative, on the other hand, is the GOAL-case of choice with 3D unbounded LMs and LMs denoting dwellings (and with ‘home’).

The analysis of data revealed that specificity is one major parameter affecting the choice between GOAL-cases. Specificity is tied to the capacity of the conceptualizer to locate the TR in LM and/or estimate the effect of TR’s action on LM. For recap, specificity of LM means that the LM entity is easily distinguishable from its surroundings, and the specificity of the relation between the TR and the LM means that either the action of TR affects the whole LM, after the action TR occupies a significant portion of LM, or both. Specificity of LM presupposes specificity of the relation between the TR and the LM, as specific LM facilitates the locating of the TR in the LM, and thus makes it easier to evaluate the effect of TR to LM. Non-specificity is then the opposite of this situation. Unspecific LMs make it hard to locate TR in them, and unspecific relations leave the effect of TR to LM unclear. This conceptualization is based on the embodied aspect of conceptualization in general. The conceptualization of a relation between the TR and the LM as specific or unspecific relies on the previous experience of the conceptualizer used as a basis of categorizing situations and applying the proper linguistic elements (in the view of the conceptualizer) to code the conceptualized relation. There are also other parameters affecting the choice of GOAL-case, like what sense the inflected word expresses (see Table 3 in Section 2.2). However, it is possible that specificity also underlies these parameters as it seems to be the most basic one. This is a case of conceptualization affecting the structure of language (cf. [Jiang, Yang 2021: 260]).

0D objects and bounded 2D LMs show high specificity and therefore take the illative as their GOAL-case. Same goes for 3D bounded entities with the restrictions given above. 3D bounded

LMs denoting dwellings show high non-specificity in the location of TR after the action, i.e. the TR does not occupy a significant portion of the LM. 3D unbounded LMs show high non-specificity and therefore their GOAL-case is usually lative. For 3D unbounded LMs that are expressed by relational nouns obligatorily taking lative the relation is not this straightforward. The relation they express are undeniably unspecific, but this is not marked by the tendency to take lative as their GOAL-case. Rather the situation where frequently occurring 3D unbounded LMs always take lative as their GOAL-case has probably affected the rise of a situation where specificity affects the choice of GOAL-case. 2D unspecific LMs should, in principle, behave like 3D unbounded LMs, but the data allows no reliable analysis of behavior of this kind of entities.

Finally, it is worth noting that the analysis is based on prototypicality. This means that prototypical cases conform to the parameters discussed here. However, there are always less prototypical cases in language, and the discussion here describes them less well. These less prototypical cases tend to have some connection to the prototypical cases. For example, the use of lative with prototypical 3D bounded LMs in Erzya is non-prototypical, as these entities tend to be highly specific. Still, we find examples such as (5) in the data.

- (5) ERZYA (MokshEr-V.3/V.2 additions/Erzya/Syatko/Syatko-2008_8/Syatko-2008_8_40-47.txt)
 tʃugunov ard-sʲ vokzalo-v.
 [name] drive-PST1.3SG station-LAT
 ‘Chugunov drove to the railway station.’

In cases like this, there probably is a connection in the conceptualization to either unspecific LM or unspecific relation between the TR and the LM. The data has unfortunately too few examples of non-prototypical use of the GOAL-cases that anything certain could be said about their connections to the prototypical cases.

It is worth noting that Haspelmath [2019] has noted cross-linguistical tendencies of GOAL-marking that can be applied to the distribution of goal-cases of Mordvin languages. He notes that nouns denoting entities not typically conceived to be places take usually different marking than nouns typically conceived to be places [Ibid.: 317–322]. In addition, he notes that place names and relational nouns are cross-linguistically conceived as nouns expressing typical places [Ibid.: 322–324]. He then proceeds to explain the differences in place marking by a principle of expected behavior of nouns. If a noun is expected to act as a place it does not need overt marking, but if its functioning as a place is unexpected, marking is needed [Ibid.: 329–330]. He derives the expected behavior from frequencies, but if that is disregarded, we are dealing with a conceptualization pattern where the type of the noun is salient, and the marking associated with each type of noun is entrenched in the mind of the language user. In the case of variation of goal-cases in Mordvin languages the type of the noun corresponds to the seven categories of entities presented in Section 2.2.

4. The non-spatial LM types

As was shown in Table 3 in Section 3, there is a large amount of non-spatial LMs in the data. These form three categories, namely institutions, abstract entities, and temporals. Institutions are entities like *ferma* (M) ‘farm’ that do occupy a place in space, but consist of more than just the spatial parts, as discussed in Sections 2 and 3. Institutions are almost exclusively marked with lative in my data. Abstract entities are entities that do not occupy space, like *viĵ* (E) ‘strength’. Temporals, like *iľædʲ* (M) ‘evening’, could be seen as a subcategory of abstract LMs, but they are categorized here separately, as time constitutes its own cognitive domain. Both of these take almost exclusively illative as their GOAL-case.

The key to the use of lative with institutions seems to be the structure of an institution. As discussed, institutions habitually consist of one or more buildings, certain persons working in it

and activities that are performed there or have some connection to the activity done in the institution. In addition, TR usually has a reason to interact with the institution, which does not compass all places, persons, and possible activities the institution consists of. Furthermore, some of the activities associated with an institution are not located in a place associated with the institution. E.g., homework is part of school and asking help from a police officer on the street is part of police institution, but they happen outside the prototypical site of the institution. In other words, the relation between the TR and the LM is unspecific even though the LM itself is separable from its surroundings, much like with dwelling LMs.

A note on *kudo* (E) / *kud* (M) ‘home’ is in order here. From synchronic point of view, home can be considered on par with institutions, as it consists of a lot more than just a building, and as such the relation between the TR and the LM could be considered unspecific. Another factor contributing to the non-specificity of home is no doubt that home is an entity habitually related with people, and when a TR, e.g., moves home it does not matter where the home is situated physically. Historically, however, the frequency of the word has probably helped to retain the lative as the GOAL-case of this specific entity, a lot like with relational nouns. It can be hypothesized that these factors have interacted to yield the situation visible today, but further research is needed to settle this matter. It is interesting to note that the synonymous words with the meaning ‘house’ take illative. This is in line with the supposition of specificity, as houses are 3D bounded entities, and TR interacts with a noticeable part of LM in the end of some action.

The use of illative with abstract nouns seem to be an anomaly with respect to the specificity-based explanation. Entities like ‘work’ could quite naturally be categorized together with institutions, but the tendency in Mordvin languages is that institutions as GOALS are coded with lative, whereas abstract things are coded with illative. As a working hypothesis it can be mentioned that this distribution seems to imply that abstract entities are conceptualized as 0D objects. As shown in Section 3.1, 0D objects as GOALS are overwhelmingly coded with illative because TR affects the whole LM in these cases, and thus makes the relation specific. The same explanation would then appear also to temporal entities, probably because they are abstract as well. This hypothesis seems to explain the seemingly anomalous coding of abstract GOALS with illative, and institutional GOALS with lative, but the distribution and its causes should be studied further.

5. Reflexes of specificity in the grammar of Mordvin languages: The distribution of declensions

In addition to the arguments presented above, one can find additional evidence of the feature of specificity directing the choice between the GOAL-cases in the Mordvin languages. Lative is lacking in the nominal declensions inherently specific, i.e. the possessive and definite declension. These declensions express more specific referents than the basic declension as possession is a relation where possessor (TR) possesses (affects) the whole thing possessed (LM). Definite declension, by definition, marks definite entities. Definiteness can be seen as a grammatical reflex of specificity, as definite entities are more salient in discourse and more accessible in cognition, and thus more clearly delimited than non-definite entities. In Mordvin languages the definite declension is used when the hearer can identify the referent of the inflected word [Bernhardt 2021: 37–40]. Identifiable referent is mentally easily accessible for conceptualization, as it is salient in the discourse. If LM is marked with definite declension the easy access facilitates the assessment of the relation between the TR and the LM. This parallels the specific/unspecific distinction outlined in Section 3.3.

The relational nouns of Mordvin languages inflected in spatial cases can take possessive suffixes [Muravieva, Kholodilova 2018: 221–223]. This applies also to lative, which is incapable of taking possessive inflection in the content noun paradigm. The possessive marking on relational nouns does not, however, mark possession in a strict sense. Instead, they cross-refer the

entity in relation to which the TR is located [Ibid.: 231–234]. Because of this possessive inflection of relational nouns does not express the same level of specificity as in content noun inflection. However, it is possible that on a scale of specificity relational nouns with possessive inflection are more specific than relational nouns without it. The effect of possessive inflection on specificity of relations expressed by relational nouns should be studied further.

In contrast, illative is well capable to appear freely with possessive inflection (6), and even in definite declension in Erzya, as long as it is in plural (7). In Moksha, the definite declension is confined to nominative, genitive, and dative, and all other definite cases are formed with the help of the postpositional stem *e-/ez-*. This type of forming definite illative is marginally productive in Moksha [Bernhardt 2021: 30–34].

(6) MOKSHA (MokshEr-V.3/V.1/Moksha/Moksha/2002/3.txt)

məziarda korotʃkov praɛ:a-sʲ “kelʲgoma jalga-nts” marʲta i
 when [name] say.goodbye-PST1.3SG dear friend-POSS.GEN.3SG with and
 valɡ-sʲ poʒəzd-sta, zʲepə-zə-nza ilʲad-sʲiʲ kafta
 descend-PST1.3SG train-ELA pocket-ILL-POSS.3SG stay-PST1.3PL¹⁷ two
 tʲoʒætʲ-tʲ tsalkovaj.
 thousand-PL rouble

‘When Korochkin said goodbye to his “dear friend” and got off the train, 2000 roubles were left in his pocket.’

(7) ERZYA (MokshEr-V.3/V.2 additions/Erzya/Syatko/Syatko-2007_6/Syatko-2007_6_102-111.txt)

viʃka velle-tʲnʲe-s gazo-sʲ sval patʲʲi-nʲe-v-i
 small village-PL.DEF-ILL gas-NOM.DEF always deliver-FRQ-PASS-PRS.3SG
 ʃka-sto-nzo.
 time-ELA-POSS.3SG

‘To small villages the gas will always be delivered on [its] time.’

6. Conclusion

In this paper, I have discussed the underlying reasons for the variation exhibited by the GOAL-cases in Mordvin languages. This variation is shown to involve the type of LM in the following way. Illative is used typically when the LM entity is:

- 1) a 0D object, a 2D bound entity, or a 3D bounded entity that is not a dwelling or does not denote ‘home’;
- 2) an abstract or temporal entity.

In contrast, the lative is typically used when the LM entity is:

- 1) a 3D unbounded entity, or a 3D bounded entity that is a dwelling or denotes ‘home’;
- 2) an institution.

The distribution of GOAL-cases between the different types of LM entities was shown to involve specificity of the relation between the TR and the LM. Specificity in spatial contexts (with 0D objects, and 2D and 3D bounded and unbounded entities) involves two partly overlapping conceptualizations: the extent of the action of the TR to the LM, and the size of the area of interaction between the TR and the LM. When the action of the TR affects most or all of the LM, or if the TR interacts with a significant part of the LM, the relation is specific, and illative is used to mark GOAL. In opposite case the relation is unspecific, and lative is the GOAL-case of choice.

¹⁷ The use of GOAL-case with verbs ‘stay’, ‘leave’ etc. is a tendency of some Uralic languages which does not affect the analysis here [Huumo 2006].

This analysis of spatial relations was then extended to the non-spatial LM entities present in the data. Based on the use of illative as the primary GOAL-case with abstract and temporal entities it was hypothesized that these types of LMs are conceptualized as 0D objects in Mordvin languages. Institutions as LMs, on the other hand, behave more like 3D unbounded entities, as they take almost exclusively lative as their GOAL-case. Finally, the differences between the declensions (basic, possessive, and definite) in Mordvin languages are used to get additional support for the result. Lative is excluded from possessive and definite declensions altogether, whereas illative can be used in possessive declension freely, and in definite declension in plural. This distribution parallels the specificity each declension conveys.

These results also point to the direction that the previous studies have failed to exhaustively describe the mechanism underlying the distribution of the GOAL-cases. The difference is not in the senses each case can express, nor is it in the type of the LM referent, in the sense that one can posit a list of entities taking one or the other GOAL-case. Instead, the more general tendency of conceptualizing the relation between the TR and the LM either as specific or unspecific explains the variation. The explanations put forward in previous studies touch only parts of the more universal explanation posited here. However, as this study is based on qualitative analysis of the data, further quantitative research is needed to check and corroborate the claims put forward here. In addition, a parallel study on acceptability among native speakers could shed more light to the question.

From a more general point of view, this study has tried to show that there is no such thing as synonymous cases. By employing the theory and methodology of Cognitive Linguistics it is possible to explain superficially unexplainable cases satisfactorily.

ABBREVIATIONS

1, 3 — 1 st , 3 rd person	NOM — nominative
ADJZ — adjectivizer	PASS — passive
CNG — connegative	PL — plural
CONT — continuative	POSS — possessive
DEF — definite declension	PRS — present tense
FRQ — frequentative	PRT — partitive (ablativ)
GEN — genitive	PST — past (in participles and nominal conjugation)
ELA — elative	PST1 — 1 st past tense
ILL — illative	PTCP — participle
INE — inessive	SG — singular
LAT — lative	TRA — translative
LOC — locative	> — separates subject and object in object conjugation
NEG — negative element	
NMLZ — nominalizer	

REFERENCES

- Ahlquist 1861 — Ahlquist A. *Versuch einer Mokscha-Mordwinischen Grammatik nebst Texten und Wörterverzeichnis*. St. Petersburg: Kaiserliche Akademie der Wissenschaften, 1861.
- Alhoniemi 1985 — Alhoniemi A. Über die Wohin-Kasus im Mordwinischen. *Ural-Altäische Jahrbücher*, 1985, 5: 45–53.
- Alyamkina 2000 — Alyamkina N. S. (ed.). *Mokshen' kyal'. Morfologiya* [The Moksha language. Morphology]. Saransk: Krasnyi Oktyabr', 2000.
- Anthony 2014 — Anthony L. *AntConc (version 3.4.4w)*. Tokyo: Waseda Univ., 2014. <https://www.lauranceanthony.net/software>.
- Arppe et al. 2010 — Arppe A., Gilquin G., Glynn D., Hilpert M., Zeschel A. Cognitive Corpus Linguistics. Five points of debate on current theory and methodology. *Corpora*, 2010, 5: 1–27.

- Arppe, Järvikivi 2007 — Arppe A., Järvikivi J. Every method counts. Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory*, 2007, 3: 131–159.
- Bartens 1996 — Bartens R. Über die Deklinationen im Mordwinischen. *Finnisch-Ugrische Forschungen*, 1996, 53: 1–113.
- Bartens 1999 — Bartens R. *Mordvalaiskielten rakenne ja kehitys* [The structure and development of the Mordvin languages]. Helsinki: Suomalais-Ugrilainen Seura, 1999.
- Bernhardt 2019 — Bernhardt M. Definiteness marking on proper nouns in Mordvin. *Ural-Altische Jahrbücher*, 2019, 27: 5–74.
- Bernhardt 2021 — Bernhardt M. *Määräisyys ja objektin merkitseminen mordvalaiskielissä* [Definiteness and the marking of the object in the Mordvin languages]. Ph.D. diss., Univ. of Turku, 2021.
- Cappelle et al. 2010 — Cappelle B., Shtyrov Y., Pulvermüller F. Heating up or cooling up the brain? MEG evidence that phrasal verbs are lexical units. *Brain & Language*, 2010, 115: 189–201.
- Carlson 2010 — Carlson L. Parsing space around objects. *Language, cognition and space. State of the art and new directions*. Evans V., Chilton P. (eds.). London: Equinox, 2010, 115–137.
- Divjak et al. 2016 — Divjak D., Levishina N., Klavan J. Cognitive Linguistics. Looking back, looking forward. *Cognitive Linguistics*, 2016, 27: 447–463.
- Erkkilä 2022 — Erkkilä R. *Dataset of Mordvin GOAL-cases*. <https://doi.org/10.5281/zenodo.5940175>
- Evans 2010 — Evans V. The perceptual basis of spatial representation. *Language, cognition and space. State of the art and new directions*. Evans V., Chilton P. (eds.). London: Equinox, 2010, 21–48.
- Evseviev 1928 — Evseviev M. E. *Osnovy mordovskoi grammatiki. Erzyan'-grammatika s prilozheniem obratsov mokshanskikh sklonenii i spryazhenii* [Basics of Mordvin grammar. Erzya grammar with examples of Moksha declension and conjugation]. Moscow: Central Press of the Peoples of the USSR, 1928.
- Hamawand 2021 — Hamawand Z. Construal. *The Routledge handbook of cognitive linguistics*. Wen X., Taylor J. R. (eds.). New York: Routledge, 2021, 242–254.
- Haspelmath 2019 — Haspelmath M. Differential place marking and differential object marking. *STUF — Language Typology and Universals*, 2019, 72: 313–334.
- Huumo 2006 — Huumo T. “I woke up from the sofa”. Subjective directionality in Finnish expression of a spatio-cognitive transfer. *Grammar from the human perspective. Case, space and person in Finnish*. Helasvuo M.-L., Campbell L. (eds.). Amsterdam: John Benjamins, 2006, 41–65.
- Jiang, Yang 2021 — Jiang C., Yang K. Concepts and conceptualization. *The Routledge handbook of cognitive linguistics*. Wen X., Taylor J. R. (eds.). New York: Routledge, 2021, 255–267.
- Karlsson 2018 — Karlsson F. *Finnish. A comprehensive grammar*. London: Routledge, 2018.
- Kolyadenkov 1954 — Kolyadenkov M. N. *Grammatika mordovskikh (erzyanskogo i mokshanskogo) yazykov* [Grammar of Mordvinic (Erzya and Moksha) languages]. Part 2: *Sintaksis* [Syntax]. Saransk: Mordvin Book Publishing House, 1954.
- Kolyadenkov 1959 — Kolyadenkov M. N. *Struktura prostogo predlozheniya v mordovskikh yazykakh. Predlozhenie i ego chleny* [Clause structure in Mordvinic. Clause and its members]. Saransk: Mordvin Book Publishing House, 1959.
- Kolyadenkov, Zavodova 1962 — Kolyadenkov M. N., Zavodova R. A. (eds.). *Grammatika mordovskikh (mokshanskogo i erzyanskogo) yazykov* [Grammar of Mordvinic (Moksha and Erzya) languages]. Part 1. *Fonetika i morfologiya* [Phonetics and morphology]. Saransk: Mordvin Book Publishing House, 1962.
- Kozlov 2018 — Kozlov L. S. Locative cases. *Elementy mokshanskogo yazyka v tipologicheskom osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 154–182.
- Lakoff 1987 — Lakoff G. *Women, fire and dangerous things. What categories reveal about the mind*. Chicago: Univ. of Chicago Press, 1987.
- Lakusta et al. 2021 — Lakusta L., Hussein Y., Wodzinski A., Landau B. The privileging of ‘support-from-below’ in early spatial language acquisition. *Infant Behavior and Development*, 2021, 65: 1–10.
- Landau 2019 — Landau B. Space in semantics and cognition. *Semantics. Typology, diachrony and processing*. Von Heusinger K., Maienborn C., Portner P. (eds.). Berlin: De Gruyter Mouton, 2019, 341–365.
- Landau 2020 — Landau B. Learning simple spatial terms. Core and more. *Topics in Cognitive Science*, 2020, 12: 91–114.
- Landau et al. 2017 — Landau B., Johannes K., Skordos D., Papafragou A. Containment and support. Core and complexity in spatial language learning. *Cognitive Science*, 2017, 41: 748–779.
- Langacker 1987 — Langacker R. W. *Foundations of Cognitive Grammar*. Vol. I. *Theoretical prerequisites*. Stanford: Stanford Univ. Press, 1987.
- Langacker 2016 — Langacker R. W. Working towards a synthesis. *Cognitive Linguistics*, 2016, 27: 465–477.

- Leino 1993 — Leino P. *Suomen kielen kognitiivista kieliooppia 1* [Cognitive grammar of Finnish language 1]. *Polysemia. Kielen moniselitteisyys* [Polysemy. The ambiguity of language]. Helsinki: Helsingin yliopiston suomen kielen laitos, 1993.
- Luodonpää-Manni et al. 2017 — Luodonpää-Manni M., Penttilä E., Viimaranta J. Introduction. *Empirical approaches to Cognitive Linguistics. Analyzing real-life data*. Luodonpää-Manni M., Penttilä E., Viimaranta J. (eds.). Cambridge: Cambridge Scholars Publishing, 2017, 1–21.
- MokshEr 2010 — MokshEr V. 3. *Mokšan ja ersän sähköinen korpus* [Electronic corpus of Erzya and Moksha]. Turun yliopisto, 2010.
- Muravieva, Kholodilova 2018 — Muravieva A. M., Kholodilova M. A. Postpositions and relational nouns. *Elementy mokshanskogo yazyka v tipologicheskoy osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 212–248.
- Newman 2010 — Newman J. Balancing acts. Empirical pursuits in Cognitive Linguistics. *Quantitative methods in cognitive semantics. Corpus-driven approaches*. Glynn D., Fischer K. (eds.). Berlin: De Gruyter Mouton, 2010, 79–99.
- Rosch 1975 — Rosch E. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 1975, 104: 192–233.
- Rosch 1978 — Rosch E. Principles of categorization. *Cognition and categorization*. Rosch E., Lloyd B. B. (eds.). Hillsdale (NJ): Lawrence Erlbaum, 1978, 27–48.
- Serebrennikov 1968 — Serebrennikov B. A. *Istoricheskaya morfologiya mordovskikh yazykov* [Historical morphology of the Mordvin languages]. Moscow: Nauka, 1968.
- Schmid 2010 — Schmid H.-J. Does frequency in text instantiate entrenchment in the cognitive system? *Quantitative methods in cognitive semantics. Corpus-driven approaches*. Glynn D., Fischer K. (eds.). Berlin: De Gruyter Mouton, 2010, 101–133.
- Schmid 2014 — Schmid H.-J. Lexico-grammatical patterns, pragmatic associations and discourse frequency. *Constructions Collocations Patterns*. Herbst T., Schmid H.-J., Faulhaber S. (eds.). Berlin: De Gruyter Mouton, 2014, 239–293.
- Svorou 2007 — Svorou S. Relational constructions in cognitive linguistics. *The Oxford handbook of cognitive linguistics*. Geeraerts D., Cuyckens H. (eds.). Oxford: Oxford Univ. Press, 2007, 726–752.
- Talmy 2000 — Talmy L. *Towards a cognitive semantics*. Vol. 1. *Concept structuring systems*. Cambridge: MIT Press, 2000.
- Taylor 2003 — Taylor J. R. *Cognitive grammar*. 3rd edn. Oxford: Oxford Univ. Press, 2003.
- Tsygankin 1980 — Tsygankin D. V. (ed.). *Grammatika mordovskikh yazykov. Fonetika, grafika, orfografiya, morfologiya* [Grammar of the Mordvinic languages. Phonetics, graphics, orthography, morphology]. Saransk: Ogarev Mordvin State Univ., 1980.
- Tsygankin et al. 2000 — Tsygankin D. V., Agafonova N. A., Imaikina M. D., Mosin M. V., Tsyпкаikina V. P., Abramova E. A. (eds.). *Erzyan' kel' . Morfemika, valon' teema dy morfologiya* [The Erzya language. Morphemics, word-formation, and morphology]. Saransk: Krasnyi Oktyabr', 2000.
- Wen, Fu 2021 — Wen X., Fu Z. Categorization. *The Routledge handbook of cognitive linguistics*. Wen X., Taylor J. R. (eds.). New York: Routledge, 2021, 173–190.
- Wiedemann 1865 — Wiedemann F. J. *Grammatik der Ersa-Mordwinische Sprache nebst einem kleinen Mordwinisch-Deutschen und Deutsch-Mordwinischen Wörterbuch*. St. Petersburg: Kaiserliche Akademie der Wissenschaften, 1865.
- Ylikoski 2015 — Ylikoski J. From compound nouns to case marking. Prolatives in South Saami and Lule Saami. *Finnisch-Ugrische Mitteilungen*, 2015, 39: 101–155.
- Ylikoski 2016 — Ylikoski J. The origins of the western Uralic s-cases revisited. Historiographical, functional-typological and Samoyedic perspectives. *Finnisch-Ugrische Forschungen*, 2016, 63: 6–78.

The status of Dative case in the Moksha case paradigm

© 2022

Polina S. Pleshak

University of Maryland, College Park, USA; ppleshak@umd.edu

Abstract: In this paper, I discuss the structure of the Moksha case system and its implications for linguistic theory. Based on their morphological properties, Moksha (Uralic) cases are divisible into two groups, which seem to correspond to structural and inherent cases. Dative, however, presents a puzzle: distributionally, it behaves as an inherent case, but morphologically it patterns with the structural ones. To resolve this discrepancy, I argue that inherent cases are always headed by a P, providing a finer classification of P heads (relational vs. non-relational) based on their morphosyntactic properties. Synchronically, inherent “case” markers are bound counterparts of free-standing postpositions and relational nouns. Diachronically, the latter grammaticalize to the former. The nature of a free-standing element gives rise to two different patterns once it is grammaticalized into a case marker, and thus the case marker continues to echo either the relational nature (< relational noun) or non-relational nature (< postposition) of the element it came from. Dative is a non-relational P head that assigns structural genitive to its complement, explaining why the morphological behavior is the one found with structural cases. On the other hand, other inherent cases are relational P heads, which take bare complements.

Keywords: case, dative, grammaticalization, Moksha, Mordvin, noun phrase, postpositions, relational nouns, Uralic

Acknowledgements: I am grateful to my advisors Masha Polinsky and Omer Preminger for their discussions and advice, as well as to my friends and fieldwork colleagues from Lomonosov Moscow State University. This research could never have happened without my language consultants. All errors are my responsibility.

For citation: Pleshak P. S. The status of Dative case in the Moksha case paradigm. *Voprosy Jazykoznanija*, 2022, 5: 108–130.

DOI: 10.31857/0373-658X.2022.5.108-130

Статус дательного падежа в мокшанской падежной парадигме

Полина Сергеевна Пleshак

Университет Мэриленда, Колледж-Парк, США; ppleshak@umd.edu

Аннотация: В данной статье обсуждается структура падежной системы мокшанского языка, а также следствия наличия такой системы для лингвистической теории. По морфосинтаксическим свойствам мокшанские падежи могут быть разделены на две группы, которые в целом соответствуют структурным и ингерентным падежам. Однако датив представляет собой проблему, поскольку, имея дистрибуцию ингерентного падежа, он группируется со структурными падежами с морфосинтаксической точки зрения. В статье предлагается решение данной проблемы: все ингерентные падежи в синтаксической структуре являются вершинами категории P. Далее предлагается детальная классификация вершин P (реляционные vs. нереляционные), основанная на их морфосинтаксических свойствах. Синхронно показатели ингерентных падежей представляют собой связанные с морфологической точки зрения соответствия свободностоящим послелогам и реляционным именам. Диахронически же первые возникают из последних в процессе грамматикализации. Таким образом, различие в наблюдаемых паттернах обусловлено природой свободностоящих

элементов, грамматикализующихся в «падежные» аффиксы. Эти аффиксы сохраняют реляционную или нереляционную природу своего источника грамматикализации. В частности, датив является нереляционной вершиной P, приписывающей структурный генитив своему комплементу. Это объясняет, почему датив имеет морфосинтаксические свойства структурного падежа. Прочие же ингерентные падежи являются реляционными вершинами P и могут присоединять немаркированные комплементы.

Ключевые слова: грамматикализация, датив, именная группа, мокшанский язык, мордовские языки, падеж, послелогои, реляционные имена, уральские языки

Для цитирования: Pleshak P. S. The status of Dative case in the Moksha case paradigm. *Voprosy Jazykoznanija*, 2022, 5: 108–130.

DOI: 10.31857/0373-658X.2022.5.108-130

1. Introduction

The focus of this paper is the case system of Moksha (Mordvin < Finno-Ugric < Uralic). I will first present and analyze the Moksha case data and then use these data to discuss their implications for linguistic theory.

Cases in Moksha can be morphologically divided into two main groups (see [Simonenko, Leontjev 2012] for such a division within several Finno-Ugric languages). In one group of forms, the case marker follows the possessive marker, and number has to be expressed (1a–b). This group includes nominative (NOM), genitive (GEN), and dative (DAT). In the rest of the Moksha cases, the case marker precedes the possessive marker, and the form is not specified for number (1c).

- (1) a. *vel'ə-c'ə-n'd'i* b. *vel'ə-t'n'ə-n'd'i* c. *vel'ə-sə-t*
 village-2SG.POSS.SG-DAT village-2SG.POSS.PL-DAT village-IN-2SG.POSS
 'to your village' 'to your villages' 'in your village(s)'

It may be tempting to link the contrast between the two groups to the opposition between “structural” and “inherent” cases [Chomsky 1981], which differ in terms of their licensing conditions. Structural cases are assumed to be assigned by a particular head; for example, nominative is assigned by T. In the configurational theory of Case [Marantz 1991], “lexical” Case (= “inherent” in Chomsky’s terminology) is licensed by a head, in contrast to all other Cases of the disjunctive hierarchy that are assigned based on a specific syntactic configuration of DPs. In general, inherent case is assigned to forms associated with particular thematic roles, and it is common to characterize DAT as inherent. If so, the grouping of DAT with NOM and GEN, two unambiguously structural cases, is unexpected. Using this dilemma as my starting point, I argue that NOM and GEN are genuine case forms, that is, case features on a DP head, while DAT groups with other case forms which all include an underlying P head. Further still, I show that Moksha has two types of P heads with different morphosyntactic properties, and demonstrate that the difference between two groups of P heads is in their relational vs. non-relational nature, as shown in Figure 1.

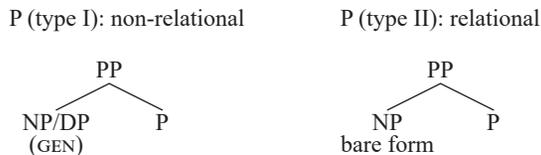


Figure 1. Two types of P heads in Moksha

Based on my proposal, the Moksha DAT patterns with structural cases NOM and GEN, because it is actually a PP with a non-relational head (Type I), which in turn embeds GEN.

The data for my analysis are taken from a recent grammatical description of one of the Moksha dialects [Toldova et al. 2018], created as a result of a fieldwork project organized by the Department of Linguistics of Lomonosov Moscow State University in 2013–2018 as well as from my own fieldwork notes on the same dialect taken in 2014–2019.

The structure of the paper is as follows. In Section 2, I provide some theoretical background. In Section 3, I familiarize the reader with the basics of Moksha morphosyntax necessary to understand the problem and the analysis. In Section 4, I go into details describing the properties of different types of P heads in Moksha. Finally, in Section 5, I discuss my analysis and its predictions.

2. Theoretical background

I adopt the **hierarchical model of the nominal phrase**. The noun phrase is represented as a sequence of functional layers above the projection of the lexical noun (NP), such as NumP (number features), PossP (person and number of possessor) and DP (referentiality, definiteness). Schematically, this can be represented as in Figure 2 [Abney 1987; Szabolcsi 1984; Alexiadou et al. 2007; Lyutikova 2018 and references therein].

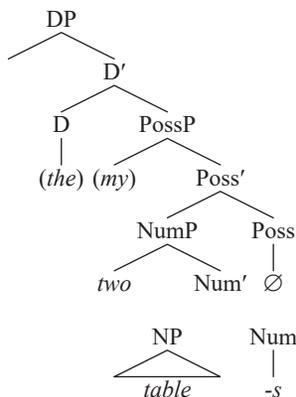


Figure 2. The structure of the noun phrase *the two tables / my two tables*

Not all noun phrases have to contain the same amount of structure [Cinque (ed.) 2002; Déchaine, Wiltschko 2002], and noun phrases that lack some higher projections can be referred to as “small nominals” [Pereltsvaig 2006]. Evidence for small nominal status comes from the particular syntactic properties of these nominals, such as restriction on modification by possessors or demonstratives, or the impossibility of nominal morphology marking number or possession.

As mentioned above, it is customary to recognize the distinction between structural and inherent cases based on their licensing properties. Empirical evidence for this distinction comes from the behavior of structural cases in passive and raising constructions (see [McFadden 2004; Woolford 2006], among others). Arguments that bear structural cases change their case marking (ACC direct objects are assigned NOM after they move to the subject position, see (2)), while arguments in inherent cases do not (for instance, dative arguments preserve their marking, see (3)).

(2) PASSIVE: ACC OBJECTS → NOM SUBJECTS

- a. *Hans schieb-t mein-en Bruder.*
 Hans[NOM] push-PRS.3SG my-ACC brother[ACC]
 ‘Hans pushes my brother.’

- b. **Mein Bruder** / **mein-en Bruder ist geschoben worden.*
 my[NOM] brother[NOM] my-ACC brother[ACC] be.PRS.3SG pushed become
 ‘My brother was pushed.’ (adapted from [McFadden 2004: 6])

(3) PASSIVE: DAT OBJECTS PRESERVED

- a. **Hans hilf-t mein-em Bruder.**
 Hans help-PRS.3SG my-DAT brother[DAT]
 ‘Hans helps my brother.’

- b. **Mein-em Bruder** / **mein Bruder ist geholfen worden.*
 my-DAT brother[DAT] my[NOM] brother[NOM] be.PRS.3SG helped become
 ‘My brother was helped.’ (adapted from [McFadden 2004: 6])

There are also certain expectations regarding the functional distribution of these cases [Woolford 2006]. Structural cases are assigned to a structural position, and are not related directly to a specific semantic role. Inherent cases, on the other hand, are assigned based on semantic roles. For instance, DAT is strongly associated with the role of recipient or goal. Since mainstream syntactic tests, such as passivization, raising, agreement, control of PRO, and reflexive binding, can reveal the difference in syntactic positions rather than the nature of the case itself, I rely on the diagnostics of the restriction on case mismatch in free relatives¹ (see Section 3.1). I will also use functional distribution of the cases discussed above to determine the inherent / structural status of DAT.

Along with the mainstream approach, which recognizes the division of cases into structural and non-structural (inherent or lexical), a group of approaches has been proposed, which treat inherent cases as instances of P (adpositional head)² [Bresnan, Grimshaw 1978; Emonds 1985; 1987; Nikanne 1993; McFadden 2004; Asbury 2008; Caha 2009; Dékány 2011; Polinsky 2016]. The basic idea behind these approaches is that, when a DP gets its case for semantic reasons, there is in fact a P in the structure that connects the DP to the predicate.³ Thus, one can represent the taxonomy of P heads as shown in Figure 3.

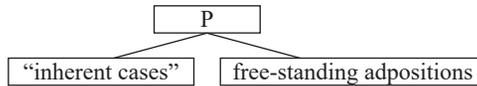


Figure 3. Taxonomy of P heads

In what follows, I will develop my analysis of the Moksha data based on the outlined assumptions.

3. Basics of the Moksha nominal morphosyntax

In this section, I first outline the basic functions of Moksha cases and show that, based on their functional distribution, they can be divided into two classes: structural (NOM and GEN) and

¹ I am grateful to an anonymous reviewer for suggesting this diagnostic.

² What is represented in my trees as a single P head can be further split into several functional layers, such as AxPart, PlaceP, and PathP (see [Riemsdijk 1990; Svenonius 2010], among others). Hence, the combination of free-standing adpositions (P elements) with “case” morphemes (also P elements) is not a problem for my account, e.g. *lank-sə* [top-IN] ‘on’.

³ There is still debate regarding whether the morphological inherent case is the exponence of the P head itself or else is assigned by a null P head. This distinction does not impact my analysis; consequently, I do not discuss the pros and cons of the two different mappings. For the purposes of this paper, I follow [Asbury 2008] in assuming that inherent cases are exponents of P heads proper.

inherent (all the rest). Then I discuss the Moksha nominal categories. In addition to case, three categories are expressed with special morphological affixes within the nominal domain: number, possession, and definiteness. Based on the interaction of case with the other nominal categories, Moksha cases seem to be divided into two slightly different categories: one which consists of NOM, GEN, DAT; and one which consists of the remaining cases. After that, I provide some more background, showing that case affixes are not the only way to encode syntactic relations between nominals and the predicate: for certain relations, constructions with postpositions and relational nouns can also be used. The interaction of nominal categories within the case paradigm, as well as competition between bound morphological affixes (case markers) and free-standing elements (postpositions and relational nouns), will play an important role in my analysis.

3.1. Types of cases and their distribution

Descriptions of the Moksha case paradigm include up to 16 cases [Evseiev 1963; Kolyadenkov, Zavodova 1962; Feoktistov 1975; Plaksina 2002; Kholodilova 2018]. It has already been established [Kholodilova 2018: 81] that some of the nominal markers lack properties typical of case markers, such as compatibility with other nominal categories and the absence of lexical or phonological restrictions on the nouns with which they combine (see [Corbett 2012]). In this work, I focus on the first nine markers from the case paradigm in Table 1. Only the first three of them are compatible with all other nominal categories, as only these three have a definite declension, which consists of markers that express number and definiteness in addition to case. None of the cases I consider have lexical or phonological restrictions on their distribution.

Table 1

Moksha nominal declension, a fragment [Kholodilova 2018: 66]

Case	Indefinite declension		Definite declension	
	SG	PL	SG	PL
Nominative	-∅	-t / -t'	-s'	-(t')n'ə
Genitive	-ən' / -ən'n'ə		-t'	-(t')n'ə-n'
Dative	-ən'd'i		-t'i	-(t')n'ə-n'd'i
Ablative	-də / -tə / -d'ə / -t'ə			
Inessive	-sə			
Elative	-stə			
Illative	-s			
Lative	-u / -v / -i			
Prolative	-əva / -ga / -ka			

NOM, whose exponent is phonologically null in Moksha, is the case of subjects, both intransitive (4a) and transitive (4b). Indefinite direct objects also appear in this form (4c). Definite direct objects, as shown in (4b), are marked with the definite GEN.

(4) a. INTRANSITIVE SUBJECT

maša-n' s'ora-c juma-s' paks'a-v
 Mary-GEN boy-3SG.POSS.SG[NOM] get.lost-PST.3[SG] field-LAT
 'Mary's son got lost in the field.'

b. TRANSITIVE SUBJECT + DEFINITE DO

maša-n' s'ora-c juksta-z'ə kaza-t' paks'a-v
 Mary-GEN boy-3SG.POSS.SG[NOM] forget-PST.3SG.O.3SG.S goat-DEF.SG.GEN field-LAT
 'Mary's son forgot the goat in the field.'

c. INDEFINITE DO

maša-n' s'ora-c exer' juksta-j kodama bəd'ə kaza paks'a-v
 Mary-GEN boy-3SG.POSS.SG always forget-NPST.3[SG] which INDEF goat field-LAT
 'Mary's son always forgets a goat in the field.'

Adapted from [Toldova 2018: 552]

In addition to marking definite direct objects, GEN encodes adnominal dependents (5) and complements of all but one postposition (6) (see below on the postposition *baška*).

(5) ADNOMINAL DEPENDENT

ava-t' sumka-c pra-s'
 woman-DEF.SG.GEN bag-3SG.POSS.SG fall-PST.3[SG]
 'The woman's bag fell.'

(6) COMPLEMENT OF A POSTPOSITION

vel'a-n'akə-n' esə er'ε-s' koz'ε, koz'ε al'ε
 village-1PL.POSS-GEN in.IN live-PST.3[SG] rich rich man
 'There lived a rich man in our village.' [Toldova, Pleshak 2018]

Because they are assigned in particular syntactic configurations,⁴ NOM and GEN function as prototypical structural cases. DAT, on the other hand, marks recipients (7), goals (8), external arguments in passives (9), and external arguments in debitive constructions (10). A restricted set of predicates take the experiencer in DAT; in (11), 'shame' is one such predicate. Based on its functional distribution, the Moksha DAT can be classified as an inherent case (see [Woolford 2006] on the overview of functional distribution of inherent cases).

(7) RECIPIENT

pet'ε maks-s' pan'čf-t mon' s'tər'-əz'ə-n'd'i
 Peter give-PST.3[SG] flower-PL 1SG.GEN girl-1SG.POSS.SG-DAT
 'Peter gave flowers to my daughter.' [Kholodilova 2018: 87]

(8) GOAL

pet'ε suva-s' mon' kud-əz'ə-n'd'i
 Peter enter-PST.3[SG] 1SG.GEN house-1SG.POSS.SG-DAT
 'Peter entered my house (lit. to my house).' [Ibid.]

(9) EXTERNAL ARGUMENT OF A PASSIVE CONSTRUCTION

t'ε karaf-s' šuv-əv-s' vas'ε-n'd'i
 this pothole-DEF.SG dig-PASS-PST.3[SG] Vasya-DAT
 'This pothole was dug by Vasya.' [A. Kozlov et al. 2016: 541]

(10) EXTERNAL ARGUMENT OF A DEBITIVE CONSTRUCTION

s'ora-n'ε-t'i luv-əma t'ε kn'iga-s' / kn'iga-t'
 boy-DIM-DEF.SG-DAT read-NZR this book-DEF.SG book-DEF.SG.GEN

'The boy has to read this book.' (lit. 'The reading of this book occurs to the boy.')

[Zakirova 2018: 768]

⁴ I assume that the complement of a postposition or relational noun is also a particular syntactic position (in particular syntactic positions, structural case is assigned rather than inherent case).

(11) EXPERIENCER

i konečno mon'-d'eja-n ul'-s' viz'ks
 and of.course (Rus.) 1SG.OBL-PRON.DAT-1SG.POSS be-PST.3[SG] shame

‘And, of course, I was ashamed.’ [Moksha Corpus; from Toldova 2018: 555]

Ablative (ABL) encodes the theme of ingestive verbs [Toldova 2018: 556]; in addition, ABL is selected by certain other verbs (e.g. *pel'əms* ‘be afraid’), and by the postposition *baška* ‘except’, the only postposition that does not require a GEN complement. This postposition will not be considered further. The rest of the cases that I consider here are locative cases, which encode primarily spatial arguments and adjuncts [L. Kozlov 2018].

(12) SPATIAL ARGUMENT

mon povfta-jn'ə kart'ina-t' s't'ena-s
 1SG hang-PST.3.O.1SG.S painting-DEF.SG.GEN wall-ILL

‘I hung the painting on the wall.’ [L. Kozlov 2018: 168]

(13) SPATIAL ADJUNCT

mon kel'k-sa komət'-n'ə-m-s žuvar-s
 1SG like-NPST.3SG.O.1SG.S jump-FREQ-INF-ILL sand-ILL

‘I like jumping into the sand.’ [Ibid.: 169]

In addition to spatial relations proper, locative cases may encode other meanings. For instance, inessive has also instrumental (14) and comitative uses.

(14) INSTRUMENTAL PARTICIPANT

al'ε-z'ə šuft-t' ker'-si uz'ər'-sa
 man-1SG.POSS.SG tree-DEF.SG.GEN cut-NPST.3.O.3SG.S axe-IN

‘My father is cutting the tree with an axe.’ [Ibid.: 157]

What matters for the present study is that, except for NOM and GEN, which are assigned based on purely structural reasons, all other cases are linked to particular semantic roles. Setting DAT aside, lexical-case markers and postpositional phrases (discussed in Section 3.3) can equally be used to encode the same notions.

In addition to the functional distribution, morphosyntactic evidence supports DAT’s status as inherent case. This evidence comes from the restrictions on case assignment in free relatives. Free relatives are relative clauses with no overt internal or external head [de Vries 2002: 15]. They include a relative pronoun which receives case inside the relative clause (see *mejsə* ‘in what’ in (15)). At the same time, the constituent formed by the free relative (see the bracketed clause in (15)) receives case within the matrix clause. Free relatives are cross-linguistically subject to the matching condition [Grimshaw 1977]. Under this condition, the relative pronoun in a free relative clause must receive the same case as the one assigned to the entire constituent, i.e. free relative, within the matrix clause. If the relative pronoun in a free relative clause receives a case different from the case assigned to the entire constituent by the matrix licensor, case mismatch occurs. As reported by [Privizentseva 2019: 51–52], case mismatch in Moksha is only possible if NOM or GEN is assigned by the matrix clause (15). If any other case, including DAT, is assigned by the matrix clause licensor, case mismatch becomes ungrammatical (16). This diagnostic again separates DAT from the Moksha structural cases.

(15) NOM in the main clause, IN in the relative clause

tosə ašč-i [mej-sə mon mol'-an oš-u]
 there be.situated-NPST.3SG what-IN 1SG go-NPST.1SG city-LAT

‘There lies the thing in which I will go to the city.’ (adapted from [Privizentseva 2019: 51])

(16) DAT in the main clause, IN in the relative clause

**urdaz-s' pec', [mej-sə mon jaka-n oš-u]*
 mud-DEF.SG adhere.PST.3SG what-IN 1SG go-NPST.1SG city-LAT

‘The mud adheres to the dress in which I am going to the city.’ [Ibid.: 52]

In sum, the two structural cases, NOM and GEN, form a natural class, and contrast with all the other cases. In the next section, I will show that DAT morphosyntactically patterns with structural cases.

3.2. Affix ordering and the expression of number and definiteness

To begin, DAT patterns with the structural cases in the expression of number. In Moksha, number is encoded morphologically only in NOM, GEN, and DAT. Consider (17)–(18), in which (17a) and (18a) are unambiguously singular, while (17b) and (18b) are unambiguously plural. The same marker that encodes person and number features of the possessor encodes the number of the possessum (the head noun). In contrast to these examples, the noun phrase in (19) is not specified for number; depending on the context, it can be either singular or plural.

- (17) NOMINATIVE (STRUCTURAL) POSSESSED NP → NUMBER DISTINCTION
 a. *vel'ə-z'ə* [village-1SG.POSS.SG] ‘my village’
 b. *vel'ə-n'ə* [village-1SG.POSS.PL] ‘my villages’
- (18) DATIVE (INHERENT) POSSESSED NP → NUMBER DISTINCTION
 a. *vel'ə-z'ə-n'd'i* [village-1SG.POSS.SG-DAT] ‘to my village’
 b. *vel'ə-n'ə-n'd'i* [village-1SG.POSS.PL-DAT] ‘to my villages’
- (19) INESSIVE (INHERENT) POSSESSED NP → NO NUMBER DISTINCTION
vel'ə-sə-n [village-IN.POSS.1SG] ‘in my village / in my villages’

In addition to the expression of number, DAT patterns with structural cases in the expression of definiteness. As shown in Table 1, only NOM, GEN, and DAT have definite forms. In order to express definiteness in cases other than NOM, GEN and DAT, one has to use constructions with free-standing postpositions:

- (20) *vel'ə-t'esə* [village-DEF.SG.GEN in.IN] ‘in the village’

Yet another property that DAT shares with structural cases concerns the order of affixes; the possessive affix precedes the case affix, as shown in (21) for GEN and in (22) for DAT. However, in (23), where the inessive is used, the possessive affix follows the case affix.

- (21) GENITIVE → POSS+CASE
vel'ə-z'ə-n' [village-1SG.POSS.SG-GEN] ‘of my village’
- (22) DATIVE → POSS+CASE
vel'ə-z'ə-n'd'i [village-1SG.POSS.SG-DAT] ‘to my village’
- (23) INESSIVE (INHERENT) → CASE+POSS
vel'ə-sə-n [village-IN.POSS.1SG] ‘in my village / in my villages’

Upon considering how (21) and (22) differ from (23), one might posit phonologically driven metathesis⁵. Consider (24). One could assume that the form in (24a) starts as *vel'ə-n-zə* [village-POSS.1SG-ILL], with the case marker following the possessive marker. Following this line

⁵ A post-syntactic metathesis does not have to be phonologically driven; for instance, while considering Mari case patterns that are similar to the Moksha ones presented here, [Guseva, Weisser 2018] propose an account based on morphological metathesis. However, such a morphological analysis is stipulative and uses ad-hoc, language-specific features to capture the data. Meanwhile, the syntactic analysis proposed here is based on the independently motivated division of categories into relational and non-relational.

of reasoning, one would stipulate that the resulting string /nzə/ is phonotactically implausible in Moksha, which would trigger metathesis. However, in (24b), the possessive affix itself is /nzə/, so an explanation in terms of syllabic restrictions is untenable.

- (24) a. CVN-STRUCTURE
vel'ə-zə-n [village-ILL-POSS.1SG] 'to my village / to my villages'
- b. NCV-STRUCTURE
vel'ə-nzə [village-3SG.POSS.PL] 'his/her villages'

3.3. Obligatory possessive agreement

Finally, DAT patterns with structural cases in terms of obligatory possessive agreement. Possessed nominals in Moksha show ϕ -feature agreement between the nominal head and its GEN dependent (' ϕ -feature' is a cover term for person and number, and, in other languages, gender as well). Not all GEN dependents trigger agreement on the head [Pleshak 2021]; noun phrases not denoting proper names of humans, if marked with GEN of the indefinite declension, do not trigger agreement. All other GEN dependents trigger obligatory possessive agreement on the head noun if the head noun is marked with NOM, GEN, or DAT, (25)–(26). In other case forms, possessive agreement with possessors that are not expressed by personal pronouns is optional, as illustrated for illative in (27).

- (25) GENITIVE (STRUCTURAL) → OBLIGATORY AGREEMENT
t'ɛ-n' panar-ənc / **panar-t'* *mon ton'-d'ej-t*
 this-GEN dress-3SG.POSS.SG.GEN dress-DEF.SG.GEN 1SG 2SG.GEN-PRON.DAT-POSS.2SG
n'ɛ-f-n'-in'ə
 see-CAUS-IPFV-PST.3.O.1SG.S
 'I showed you that one's dress.'
- (26) DATIVE (INHERENT) → OBLIGATORY AGREEMENT
t'ɛ-n' ruc'ɛ-n'a-nc̣ti / **ruc'ɛ-n'ɛ-t'i*
 this-GEN shawl-DIM-3SG.POSS.SG.DAT shawl-DIM-DEF.SG.DAT
mon aškəd-in'ə kuc'uf-n'ə-n'
 1SG wrap-PST.3.O.1SG.S spoon-DEF.PL-GEN
 'I wrapped the spoons in her shawl.' [Pleshak, Kholodilova 2018: 291]
- (27) ILLATIVE (INHERENT) → OPTIONAL AGREEMENT
t'ɛ-n' ruc'ɛ-n'a-zə-nzə / ^{OK}*ruc'ɛ-n'a-s* *mon*
 this-GEN shawl-DIM-ILL-3SG.POSS shawl-DIM-ILL 1SG
aškəd-in'ə kuc'uf-n'ə-n'
 wrap-PST.3.O.1SG.S spoon-DEF.PL-GEN
 'I wrapped the spoons in her shawl.' [Ibid.]

Possessors expressed by personal pronouns form a special class; as opposed to other possessors, they **always** trigger agreement on the head noun, irrespective of its case (28).⁶

⁶ Agreement with 1st and 2nd person possessors is strictly obligatory (28), whereas the absence of agreement with 3rd person pronominal possessors is perceived by many speakers as less natural but is not ruled out completely:

- (i) AGREEMENT WITH 3rd PERSON PRONOMINAL POSSESSOR
ton n'ej-əv-at pek mazi-stə son' panar-sə-nzə / ?*panar-sə*
 2SG see-PASS-NPST.2SG very beautiful-EL 3SG.GEN dress-IN-3SG.POSS dress-IN
 'You look very nice in her dress.' [Pleshak, Kholodilova 2018: 290]

(28) AGREEMENT WITH 1ST PERSON PRONOMINAL POSSESSOR

*ton n'ej-əv-at pek mazi-stə mon' panar-sə-n / *panar-sə*
 2SG see-PASS-NPST.2SG very beautiful-EL 1SG.GEN dress-IN-1SG.POSS dress-IN

'You look very nice in my dress.' [Ibid.: 289]

The properties of different cases in Moksha are summarized in Table 2. We see that DAT shares morphosyntactic properties with the structural cases and differs in that regard from other cases.

Table 2

Morphosyntactic properties of structural cases, DAT, and other inherent cases

Case	Number expression	Definiteness expression	Affix order	Agreement with non-pronominal possessors
Structural: NOM, GEN	yes	yes	POSS+CASE	obligatory
Inherent: DAT	yes	yes	POSS+CASE	obligatory
Inherent: other	no	no	CASE+POSS	optional

As I already mentioned, the meanings expressed by inherent cases can alternatively be expressed by postpositional phrases or phrases with relational nouns, both of which belong to the PP category. For instance, to encode an instrument, a postposition or the inessive case can be used, as shown in (29) and (30) respectively.

(29) INSTRUMENTAL PARTICIPANT: FREE-STANDING POSTPOSITION

son laz-ənd-i pen'ga-t es' uz'ər'-ənc mar̥tə
 3SG split-FREQ-NPST.3[SG] firewood-PL REFL axe-3SG.POSS.SG.GEN with

'He splits logs with his axe.' [L. Kozlov 2018: 159]

(30) INSTRUMENTAL PARTICIPANT: BOUND MORPHEME (INESSIVE)

son laz-ənd-i pen'ga-t es' uz'ər'-sə-nzə
 3SG split-FREQ-NPST.3[SG] firewood-PL REFL axe-IN-3SG.POSS

'He splits logs with his axe.' [Ibid.]

In the next section, I describe the properties of postpositions and relational nouns, comparing them to the properties of inherent cases.

4. Postpositions, relational nouns and two types of P heads

Scholars of Finno-Ugric languages have paid a great deal of attention to the proposal that postpositions are grammaticalized from nouns and share some properties with them. It is common to assume either a continuum or cline between the two classes [Maitinskaya 1982; Grünthal 2003; Biryuk 2005; Arkhangelskiy, Usacheva 2015]. According to grammaticalization studies, each element on such a continuum is characterized by a cluster of properties; in the cline considered here, only some of the characteristic properties are nominal (for example, case assignment to the complement or possessor-like ϕ -feature agreement). The closer an element is to the nouns on the cline, the more nominal properties it exhibits. Postpositions are seen as a category that lost nominal properties in the process of grammaticalization. Researchers also recognize a separate category between regular nouns and postpositions, relational nouns, which share some properties with regular nouns and some with postpositions, see (31). Crucial for the discussion here is the divide between relational nouns and postpositions.

(31) nouns > **relational nouns** > **postpositions** > case markers

Relational nouns mostly denote spatial relations. Consider (32) from English, where *top* is a relational noun, which obligatorily takes an argument (*of the tree* in this case).

- (32) *top* *of the tree*
 relational noun internal argument

Even though some evidence justifies each of the grammaticalization steps in (31) (see discussion below), the continuum in (31) fails to account for the synchronic grammar of individual Moksha speakers. This is for two reasons. First, postpositions do not form a homogenous class based on their properties. Second, some case markers share properties with relational nouns, but not with all postpositions. Thus, the continuum is empirically insufficient in terms of synchronic grammar. Further, it does not fully account for possible diachronic change in the language because the transition from nouns to case markers does not always go according to the pathway in (31).

In order to capture differences and similarities between relational nouns, particular postpositions, and certain case morphemes, I propose a synchronic classification of P heads (Section 4.1) and argue that some case morphemes are bound counterparts of relational nouns while others are bound counterparts of genuine postpositions. I use possessive agreement and complement marking as diagnostics for a particular type of postpositions, as these two properties are always bundled together in this domain of Moksha grammar. Based on these diagnostics, we find two types of structurally-ambiguous items among Moksha relational nouns and postpositions. First, relational nouns grammaticalize to a “more functional” type, which I call *grammaticalized relational nouns*. Second, grammaticalized relational nouns can shift to genuine postpositions, in which case further semantic bleaching is involved. Alternatively, they can shift to bound morphemes via form-reduction. The recognition of categories such as *relational noun*, *grammaticalized relational noun*, and *genuine postposition* indicates that grammaticalization processes can subdivide into (i) processes that turn a lexical category to a functional category, and (ii) processes that are responsible for a shift between functional categories.

This requires replacing the continuum in (31) with a more complex structure, as shown in Figure 4.

Lexical categories	Common nouns	Relational nouns	Adverbs/other
Pheads	free-standing	Relational postpositions (grammaticalized relational nouns)	Non-relational postpositions (genuine postpositions)
	bound	Case markers (type I)	Case markers (type II)

Figure 4. Synchronic classes and possible grammaticalization paths

In contrast to an approach in which relational nouns share the same cline with common nouns, merely representing a different stage on that cline, my proposal is that relational nouns and common nouns belong to different lexical classes, each with its own properties, already prior to grammaticalization. In other words, it is crucial that relational nouns proper are not derived from common nouns via the grammaticalization process and always exhibit distinct properties. Common nouns do not grammaticalize into postpositions. Relational nouns, on the other hand, serve as a source of grammaticalization for some postpositions (I call such postpositions *grammaticalized relational nouns*).^{7,8} However, relational nouns are not the only grammaticalization

⁷ It may be the case that the difference between non-grammaticalized and grammaticalized relational nouns is more terminological than conceptual. In previous research, the term ‘relational noun’ refers to already grammaticalized elements. In this paper, I call those elements *grammaticalized relational nouns*, distinguishing them from their source, namely *relational nouns proper*.

⁸ Outside the Finno-Ugric tradition, relational nouns have also been seen as a source of grammaticalization for adpositions, which can later develop into inherent cases, cf. the scheme in [Lehmann 1985: 304]:

source for postpositions (hence the column “Adverbs / other” in Figure 4). Because postpositions arise from different sources, they may differ in their properties (as they do in Moksha).

Looking deeper into properties of different types of postpositions, we can observe that some elements are structurally ambiguous between grammaticalized relational nouns and genuine postpositions (see the right arrow from “relational postpositions” to “non-relational postpositions” in Figure 4). This is due to further grammaticalization and the loss of “relational” properties by postpositions. We can compare the shift from relational to non-relational postpositions to the grammaticalization of relational nouns into case markers, where the relational nature of a grammaticalized element is clearly preserved (see Section 5.1).

In summary, the proposal advanced here differs from the hierarchy in (31) in two respects: it does not consider common nouns as a diachronic source of adpositions, and it recognizes two different paths in the grammaticalization of relational nouns, which can turn into genuine postpositions or “case” markers.

4.1. Different types of P elements in Moksha

In this section, I discuss the way in which the class of P elements in Moksha is not uniform independently of their grammaticalization stage. Rather, it includes elements with different morphosyntactic properties. In arguing for the heterogeneity of Moksha Ps, I will focus on the case marking of the complement and possessive agreement on the head. Based on these properties, we can distinguish between genuine postpositions and grammaticalized relational nouns. This approach is different from that of [Muravieva, Kholodilova 2018], who divide the class of P elements in Moksha into postpositions and relational nouns based on their morphosyntactic properties but attribute difference between the two to the current grammaticalization stage.

4.1.1. Different types of complements of Ps

Before I proceed with the diagnostics that involve different types of complements, I need to provide some necessary background regarding the types of complements that can be taken by the elements I am considering. First, Ps can take personal pronouns as their complements. Personal pronouns are assigned GEN in this position and, as it is with possessors in noun phrases, trigger obligatory agreement on the head. This fact is relevant for my analysis of DAT in Section 5.

- (33) *pin'ə-s' tu-s' mon' mel'ə-n / *mel'ə*
 dog-DEF.SG go-PST.3[SG] 1SG.GEN after.LOC-1SG.POSS after.LOC
 ‘The dog went after me.’ [Muravieva, Kholodilova 2018: 231]

Second, complements of Ps include nominals marked for definiteness (34) or possession (35). I consider such nominals to be full DPs. Like personal pronouns, full DPs are assigned GEN. Such DPs can optionally trigger agreement, depending on the type of the P head (see below).

- (34) *oš-t' tərva-sə*
 city-DEF.SG.GEN coast-IN
 ‘near the city’ [Ibid.: 215]
- (35) *al'e-z'ə-n' vasc*
 man-1SG.POSS.SG-GEN place.ILL
 ‘instead of my father’ [Ibid.: 215]

relational noun > secondary adposition > primary adposition > agglutinative case affix > fusional case affix. See also the mention of a grammaticalization path from relational nouns to *relational* grammatical markers in [Heine, Kuteva 2002: 81].

Finally, Ps can take nominals that are not marked for definiteness as their complements.⁹ These can be either bare or marked with the indefinite GEN (36). If the phrase in the complement has its own modifiers, as in (36), GEN marking is always allowed. However, if it is unmodified, some Ps only allow for a bare form. Therefore, I use unmodified complements as my diagnostic.

- (36) *kav'or-s' kodamə bəd'ə morkš / morkš-an alə*
 rug-DEF.SG what INDEF table table-GEN under.LOC
 'The rug is under a table.' [Ibid.: 228]

I consider nominals without overt definiteness or possessive marking to be small nominals, namely NPs, because they do not allow plural marking (37), and therefore have no number features [Pleshak 2021]. Such small nominals trigger no agreement on either of the P head types.

- (37) **ber'ə-n' s'orə-t'n'ə ašč-ij-t' morkš-t lank-sə*
 garden-GEN seed-DEF.PL be.situated-NPST.3-PL table-PL on-IN
 Intended: 'The vegetables lay on tables.'

4.1.2. Genuine postpositions

Elements of this class select the complement in GEN, and non-pronominal possessors do not show agreement. For instance, the postposition *kolga* 'about' in (38) requires GEN marking from its complement *zabər* 'fence'. Even though this complement is not marked for number or definiteness, the bare form of the complement is ungrammatical. Example (39) shows that possessive agreement between the postposition and its non-pronominal argument is impossible, even though this complement is marked with a definite GEN, which can trigger agreement on nouns (see Section 3.3 about agreement with definite possessors).

- (38) INDEFINITE COMPLEMENT → GENITIVE MARKING
*pet'ε s'id'ə-stə koɾn'-i zabər-an' / *zabər kolga*
 Peter often-EL say.FREQ-NPST.3[SG] fence-GEN fence about
 {Peter repairs his house.} 'Peter frequently talks about fences.' [Muravieva, Kholodilova 2018: 229]
- (39) DEFINITE COMPLEMENT → NO AGREEMENT
*t'ε šuft-t' kolga / *kolga-nzə*
 this tree-DEF.SG.GEN about about-3SG.POSS
 'about this tree' [Koryakov, Kholodilova 2018: 16]

4.1.3. Grammaticalized relational nouns

The term "grammaticalized" implies some level of semantic bleaching (loss of specified lexical content) and/or reduction in the phonological form of a given item. Elements of this class combine with a non-modified indefinite complement and allow for possessive agreement. Consider (40). An unmodified complement of the element *lankšə* 'on' cannot appear in GEN unless it is additionally marked for definiteness. This is in contrast to *kolga* 'about', which requires GEN (38). In (41), we see that *lankšə* can bear a possessive marker showing agreement with its complement. Again, this is in contrast to *kolga*, which disallows possessive agreement with a non-pronominal complement (39).

⁹ I do not refer to complements as definite or indefinite based on the presence or absence of a definite marker respectively; this is because the mapping between semantics and morphosyntactic marking is not one-to-one, see [Pleshak 2021] for discussion.

(40) INDEFINITE COMPLEMENT: UNMARKED

morkš / **morkš-ən*’ *lank-sə* *put-f* *kn’iga*
 table table-GEN top-IN put-PTCP.RES book

‘There is a book on a table.’

(41) DEFINITE COMPLEMENT: AGREEMENT

morkš-n’ə-n’ *lank-sə-st* *ašč-ij-t’* *čejn’ək-t*
 table-DEF.PL-GEN top-IN-3PL.POSS be.situated-NPST.3-PL teapot-PL

‘There are teapots on the tables.’ [Muravieva, Kholodilova 2018: 218]

Based on these two critical diagnostics, we find that two groups of P elements exhibit opposite morphosyntactic properties, summarized in Table 3.

Table 3

Properties of two types of P heads in Moksha

P	Indefinite unmodified complement	Possessive agreement with a non-pronominal DP
Type I	GEN / *bare	*
Type II	bare / *GEN	OK

4.2. Grammaticalized and non-grammaticalized relational nouns

Grammaticalization still plays an important role in the distribution of properties of synchronically-distinguishable classes. In this section, I elaborate on the parts of the grammaticalization approach that are required to account for Moksha grammatical forms.

First of all, we see elements such as *lank-sə* ‘on’, which are relational nouns (with locative meanings) that have functional properties. Crucially, the same item can have one set of properties while in subject and direct object positions, and another set of properties when used in locative adjuncts or arguments. In the former case, such a relational noun belongs to the same lexical category as common nouns (N). In the latter case, a relational noun is grammaticalized into a functional category. It is worth noting that relational nouns in subject and object positions, although members of the lexical category N, have properties distinct from common nouns. In particular, unlike common nouns, they can take unmarked NPs as complements (42) [Muravieva, Kholodilova 2018: 229; Pleshak, Kholodilova 2018: 285], see (43).

(42) *kədə* [*kodamə-vək* *morkš*] *lank-s’* *urdaz-u*, *šta-k* *son’*
 if what-ADD table top-DEF.SG dirt-ATTR wash-IMP.3SG.O.SG.S 3SG.GEN
 ‘If the surface of any table is dirty, wash it.’ [Muravieva, Kholodilova: 228]

(43) ^{??}*mon* *af* *jařc-s’-an* *senger’e* *mar’* *ket’-tə*
 1SG NEG eat-FREQ-NPST.1SG green apple skin-ABL
 ‘I don’t eat the skin from green apples.’ [Pleshak, Kholodilova 2018: 285]

Example (42) contains a non-grammaticalized relational noun, while (43) contains a common noun. I assume that the adnominal dependent in (42) is the complement of the relational noun, whereas in (43), the dependent occupies the specifier position of that noun. The complement of the relational noun is in the bare form (42), and the specifier of the common noun cannot be bare (43). The contrast in case marking between (42) and (43) shows that even non-grammaticalized relational nouns are distinct from common nouns, which is important when we consider hypothetical grammaticalization “from nouns to postpositions”. As I indicated in the introduction to Section 4, the term “relational nouns” refers to not yet grammaticalized nouns like the

one in (42); further grammaticalized elements are referred to as “grammaticalized relational nouns” (this term corresponds to “relational nouns” in previous research). I will provide justification for this distinction in Section 4.2.1.

These grammaticalized relational nouns (under the name of “relational nouns”) were claimed to be still less grammaticalized than genuine postpositions. Indeed, based on a different set of ambiguous lexemes, it is possible to show that one and the same item can have properties of a grammaticalized relational noun in contexts with locative meaning and properties of a genuine postposition in contexts without locative meaning. Therefore, grammaticalized relational nouns can be further grammaticalized to postpositions. My main contribution to the discussion of grammaticalization processes is that the step “grammaticalized relational nouns > genuine postpositions” is not obligatory, because in some instances, a bound morpheme assumed to encode case shares properties with grammaticalized relational nouns, rather than genuine postpositions. Crucially, not every bound morpheme is similar to grammaticalized relational nouns; some morphemes share their properties with genuine postpositions. I will discuss the latter type in Section 5.

4.2.1. Relational nouns: from lexical to functional category

Following [Muravieva, Kholodilova 2018], I adopt the view that Moksha relational nouns have indeed undergone the process of grammaticalization, and we can still see different stages of this process. Items that have undergone further grammaticalization, e.g., postpositions, if they start with a sonorant, allow for a phonological process (a final /n/ can be omitted) on the boundary between the P head and its complement [Ibid.: 216–218]. Muravieva and Kholodilova provide the following argument. As the minimal pair in (44) shows, *n*-omission is sensitive to the first consonant on the P element. In (44a), the postposition begins with a sonorant /v/, and *n*-omission is allowed. In (44b), on the other hand, the preposition begins with an obstruent /t/, and *n*-omission is impossible. Example (45), on the other hand, illustrates that the absence of *-n* cannot be explained by case alternation between GEN and NOM. The two sentences in (45) are parallel, but in (45b), GEN marker does not contain *-n*. Thus, the difference between the phonological forms of NOM and GEN is not related to the presence of *-n* in the latter. Therefore, the process of *n*-omission is phonological.

(44) a. *vas'ε komæc' pet'ε-n' / ^{OK}pet'ε vel'f*
 Vasya jump.PST.3[SG] Petya-GEN Petya[GEN] through.top

b. *vas'ε komæc' pet'ε-n' / *pet'ε turks*
 Vasja jump.PST.3[SG] Petja-GEN Petja[GEN] through

‘[Vasya and Petja were playing Leapfrog.] Vasya jumped over Petya.’

[Muravieva, Kholodilova 2018: 216–217]

(45) a. *mon saj-an jalga-z'ə-n' / jalga-z'ə maŋtə*
 1SG come-NPST.1SG friend-1SG.POSS.SG-GEN friend-1SG.POSS.SG[GEN] with

‘I’ll come with my friend.’

b. *son sa-j jalga-nc / jalga-c maŋtə*
 3SG come-NPST.3SG friend-3SG.POSS.SG-GEN friend-3SG.POSS.SG[NOM] with

‘He’ll come with his friend.’

[Ibid.: 217]

Based on this criterion, [Muravieva, Kholodilova 2018] claim that some P-like elements in Moksha can function as either relational nouns or postpositions, depending on the context. For instance, in (46), *langə* ‘on’ functions as a relational noun when it assumes the form *lanksnə*; it is used in an argument (subject) position and does not allow for *n*-drop. On the other hand, in (47), the same element *langə* when expressed as *lanksə* is a postposition. While empirically

correct, this generalization is still confusing terminologically. To phrase it differently, items such as *langə* can sometimes be relational nouns, and sometimes grammaticalized relational nouns.

- (46) RELATIONAL NOUN \rightarrow *n*'-OMISSION IS IMPOSSIBLE
morkš-n'ə-n' / **morkš-n'ə* *lank-snə* *vad'a-ft*
 table-DEF.PL-GEN table-DEF.PL on-3PL.POSS grease-PTCP.RES-PL
 'The surfaces of the tables are dirty.' [Muravieva, Kholodilova 2018: 218]

- (47) GRAMMATICALIZED RELATIONAL NOUN \rightarrow *n*'-OMISSION IS POSSIBLE
čejn'ək-t *ašč-ij-t'* *morkš-n'ə-n'* / *morkš-n'ə* *lank-sə*
 teapot-PL be.situated-NPST.3-PL table-DEF.PL-GEN table-DEF.PL[GEN] top-IN
 'The teapots are on the tables.' [Ibid.: 217]

The contrast between a non-grammaticalized noun and a grammaticalized relational noun is further confirmed by agreement data. Compare (47), without possessive agreement, and possible *-n*' omission, and (48), where possessive agreement is indexed on the head, and *-n*' cannot be omitted.

- (48) RELATIONAL NOUN \rightarrow *n*'-OMISSION IS IMPOSSIBLE
morkš-n'ə-n' / **morkš-n'ə* *lank-sə-st* *ašč-ij-t'* *čejn'ək-t*
 table-DEF.PL-GEN table-DEF.PL top-IN-3PL.POSS be.situated-NPST.3-PL teapot-PL
 'There are teapots on the tables.' [Muravieva, Kholodilova 2018: 218]

It is worth noting that elements like *langə* take unmarked complements both in their non-grammaticalized and grammaticalized instantiations, as was shown by (42) and (40) respectively. The two indicators of grammaticalization are morphophonological processes on the boundary and restriction on possessive agreement. Note that neither nouns nor postpositions from the cline in (31) can take an unmarked complement. This means that they do not obtain the ability to take unmarked complements due to grammaticalization, but rather retain said ability despite grammaticalization.

4.2.2. Shifting between two functional categories

Items whose grammaticalization process involved further semantic bleaching in Moksha lose their relational-noun nature and acquire properties of genuine postpositions. The semantic bleaching may include the complete loss of spatial meaning. Consider (49)–(50) with the P element *kuvalməva* in its locative function. In this function, *kuvalməva* is a grammaticalized relational noun; this is evidenced by the unmarked form of the complement (49), and possessive agreement with a non-pronominal complement (50).

- (49) GRAMMATICALIZED RELATIONAL NOUN, LOCATIVE MEANING \rightarrow UNMARKED COMPLEMENT
pet'ε *s'id'ə-stə* *jaka-j* *zabər* / *ʔzabər-ən'* *kuvalmə-va*
 Peter often-EL walk-NPST.3[SG] fence fence-GEN length-PROL
 'Peter often walks along fences.' [Muravieva, Kholodilova 2018: 229]

- (50) RELATIONAL NOUN, LOCATIVE MEANING \rightarrow AGREEMENT
ki-t' *kuvalmə-va(-nzə)* *jota-s'* *mašina*
 road-DEF.SG.GEN length-PROL-3SG.POSS pass-PST.3[SG] car
 'A car passed by on the road.' [Ibid.: 233]

In (51)–(52), where the same element serves a non-locative function, it does not manifest its relational nature. As a morphosyntactic reflex of that, the indefinite complement must bear GEN marking (51), and the head must not agree with non-pronominal arguments (52).

- (51) POSTPOSITION, NON-LOCATIVE MEANING → GENITIVE COMPLEMENT
*pet'ε s'id'ə-stə koɾn'-i zabər-ən' / *zabər kuvalmə-va*
 Peter often-EL say.FREQ-NPST.3[SG] fence-GEN fence length-PROL
 'Peter frequently talks about fences.' [Muravieva, Kholodilova 2018: 229]
- (52) POSTPOSITION, NON-LOCATIVE MEANING → NO AGREEMENT
*es' er'a-f-ənc kuvalmə-va / *kuvalmə-va-nzə*
 REFL live-PTCP.RES-3SG.POSS.SG.GEN length-PROL-3SG.POSS length-PROL-3SG.POSS
at'ε-z'ə pek kel'k-si az-ən-kšn'-əma-nc
 old.man-1SG.POSS.SG very like-NPST.3SG.S.3SG.O say-FREQ-FREQ-NZR-3SG.POSS.SG.GEN
 'My grandfather loves talking about his life.' [Ibid.: 233]

In this section, I motivated the distinction between two stages of grammaticalization from the lexical category of nouns to the functional category of postpositions: relational nouns > grammaticalized relational nouns (lexical to functional), and grammaticalized relational nouns > genuine postpositions (between two functional categories). Additionally, I showed that the grammaticalization process starts with a special class of nominals, not with common nouns. In the next section, I discuss the situation in which grammaticalized relational nouns turn into bound "case" morphemes without going through the stage of genuine postpositions.

5. Inherent-case markers as postpositional heads

5.1. Inherent locative cases as relational P heads

The morphosyntactic properties of locative case markers are inherited from relational nouns (grammaticalized and non-grammaticalized). In particular, they inherit the selection of unmarked indefinite complements, as opposed to the selection of an *n'*-marked indefinite complement. I propose that relational P heads expressed by bound morphemes retain this requirement on their complement from relational nouns proper.

The main evidence for morphosyntactic similarities between locative bound morphemes and relational nouns is provided by possessive constructions. If a complement has a possessor (expressed explicitly as a GEN DP or a null element triggering agreement), there are two strategies available for relational nouns (grammaticalized and non-grammaticalized).

In the first strategy, the complement is a full DP; possessive agreement is indexed on that DP, and the relational noun cannot agree with the possessor of its complement (53).

- (53) POSSESSOR AGREEMENT WITHIN THE DP → NO AGREEMENT WITH THE POSSESSOR ON P
*put-it' maɾ'-n'ə-n' mon' morkš-əz'ə-n' lank-s / *langə-zə-n*
 put-IMP apple-DEF.PL-GEN I.GEN table-1SG.POSS.SG-GEN on-ILL on-ILL-1SG.POSS
 'Put the apples on my table.'

In the second strategy of possessor marking, the complement of the relational noun is an unmarked small nominal.

When the complement of a relational noun is a small nominal, there is no syntactic room for the possessor in the complement phrase itself, and possessor agreement occurs within the projection of the relational noun.

- (54) *put-it' maɾ'-n'ə-n' mon' morkš langə-zə-n / *lank-s*
 put-IMP apple-DEF.PL-GEN I.GEN table on-ILL-1SG.POSS on-ILL
 'Put the apples on my table.'

As I argued earlier, P heads encoded by bound morphemes inherit the selection of an unmarked complement from relational nouns. Therefore, when considering possessive phrases,

we expect possessive agreement on the P element. Possessive agreement markers follow locative case markers (see Section 3.2), so this prediction is borne out. Consider (55), repeated from earlier example (19):

- (55) INESSIVE (INHERENT) → CASE+POSS
vel'ə-sə-n [village-IN-POSS.1SG] 'in my village / in my villages'

The properties of free-standing and bound grammaticalized relational P elements are represented schematically in Table 4. Regardless of its phonological realization, a relational P does not allow GEN marking on the complement, and, if the complement is not a full DP, possessive agreement is on the P head itself. In other words, a P head can be realized as a bound morpheme only if its complement is a small nominal, morphologically unmarked.

Table 4

Summary: Free-standing and bound relational Ps in Moksha

	Free-standing grammaticalized relational noun	Inherent case
no possessor of the complement	<p>● Bare NP complement</p> <pre> graph TD PP1[PP] --- NP1[NP] PP1 --- P1[P] NP1 --- morkš[morkš] P1 --- lanksə[lanksə] </pre> <p>'on a table'</p>	<pre> graph TD PP2[PP] --- NP2[NP] PP2 --- P2[P] NP2 --- morkš[morkš] P2 --- sə[sə] </pre> <p>'in a table'</p>
+ possessor of the complement	<p>● Bare NP complement</p> <p>● Possessive agreement with the possessor of the complement (on the P-head!)</p> <pre> graph TD PossP1[PossP] --- PossDP1[PossessorDP] PossP1 --- Poss1[Poss'] PossDP1 --- mon'[mon'] Poss1 --- PP1[PP] Poss1 --- Possn1[Poss n] PP1 --- NP1[NP] PP1 --- P1[P] NP1 --- morkš[morkš] P1 --- langəzə[langəzə] </pre> <p>'on(to) my table'</p>	<pre> graph TD PossP2[PossP] --- PossDP2[PossessorDP] PossP2 --- Poss2[Poss'] PossDP2 --- mon'[mon'] Poss2 --- PP2[PP] Poss2 --- Possn2[Poss n] PP2 --- NP2[NP] PP2 --- P2[P] NP2 --- kud[kud] P2 --- sə[sə] </pre> <p>'in my house'</p>

This analysis explains why the affix order in locative cases is CASE+POSS. It also captures the fact that no number or definiteness distinctions can be encoded together with bound locative morphemes. Note as well that locative morphemes are attached to bare stems directly, with no other case markers (e.g. GEN) preceding them. This is in contrast to DAT, which always follows the GEN morpheme. I discuss this ordering in more detail in the next section.

5.2. Inherent dative case as a non-relational P head

We can now turn to the main question: why is Moksha DAT different from the other inherent cases? The answer is this: even though morphosyntactically DAT seems to pattern with the structural cases, which are analyzed as features on D heads, it is a P head.

First, DAT affixes (-*ti*; -*n'd'i*) historically develop from postpositions [Serebrennikov 1967: 18]. Moreover, there is synchronic evidence from the DAT pronominal paradigm that it is postpositional in nature. Consider (56), which represents 2nd person singular DAT form. (56a) contains a GEN form of the personal pronoun followed by the DAT marker, which is then followed by the possessive agreement marker. (56b) demonstrates that the pronoun itself can be dropped. The same structure is exhibited by postpositions taking pronominal complements.

In (57), a genuine postposition *kolga* requires a GEN form of its pronominal complement and agrees with it. The pronoun in the complement position can be omitted. In contrast to DAT, GEN has no separate identifiable marker (*ton* ' [2SG.GEN] 'your'), and NOM is the unmarked form (*ton* [2SG] 'you').

- (56) a. *ton'-d'ejə-t* b. *t'ej(ə)-t*
 2SG.GEN-PRON.DAT-POSS.2SG PRON.DAT-POSS.2SG
 'to you' 'to you'

- (57) (*ton*') *kolga-t*
 2SG.GEN about-POSS.2SG
 'about you'

DAT differs from other inherent cases with respect to its pronominal paradigm. I argue that this difference follows from the explanation for the morpheme ordering and restrictions on expression of number and definiteness: DAT is a non-relational P head, contrasted to relational P heads. Pronominal paradigm has gaps in places of other inherent cases, and free-standing postpositions are used instead (58). This is expected in my analysis, because pronouns must be full DPs, and they are known to be ungrammatical in syntactic positions where small nominals occur [Pereltsvaig 2006: 446–447, 457; Lyutikova 2019: 335]. Therefore, DPs are allowed as complements of DAT P, but not as complements of bound relational Ps (locative “cases”).

- (58) ³*ton'-cə-t*
 2SG.GEN-IN-POSS.2SG
 'in you' [Kholodilova 2018: 98]

The data shown so far suggest that DAT is expounded by a postposition. Crucially, DAT is a P head of the non-relational type. In contrast to relational P heads, DAT requires GEN marking on its complement. This in turn means that DAT markers always properly contain GEN markers, as shown by the paradigm in Table 5. The element *-t'i/-ti/-d'i* is always preceded by a GEN form of the corresponding noun.¹⁰ Such ordering is predicted if datives are underlyingly PPs whose P takes a GEN complement.

Table 5

A fragment of Moksha case paradigm: examples of GEN and DAT markers

	GEN	DAT
INDEF	<i>-n'</i>	<i>-n'd'i</i>
DEF	<i>-t'</i>	<i>-t'i</i>
POSS 1SG	<i>-z'ə-n'</i>	<i>-z'ə-n'd'i</i>
POSS 3SG	<i>-nc</i>	<i>-ncti</i>

Crucially, GEN does not get overwritten by the dative P because it is the case of the complement required by this P head. As the GEN complement occurs in its base (non-derived) position, the structural case on this complement does not get erased because it does not undergo any syntactic displacement. The corresponding structures with possessed and unpossessed complements of P heads are represented in Table 6, both for free-standing postpositions and DAT.

¹⁰ The definite form looks like an exception here; however, it is plausible to assume that one of the *t'*-s in the onset of **-t'i* has undergone deletion.

Table 6

Free-standing and bound non-relational Ps in Moksha

	Free-standing postposition	Dative case
no possessor of the complement ● GEN complement	<pre> PP / \ NP_{GEN} P morkš-ən' kolga </pre> <p>‘about a table’</p>	<pre> PP / \ NP_{GEN} P morkš-ən' -d'i </pre> <p>‘to a table’</p>
+ possessor of the complement ● GEN complement ● possessive agreement within the GEN phrase	<pre> PP / \ DP P morkš-əz'ə-n' kolga </pre> <p>‘about my table’</p>	<pre> PP / \ DP P morkš-əz'ə-n' -d'i </pre> <p>‘to my table’</p>

Once we recognize that GEN is present in DAT structures, the parallelism in the syntactic properties of DAT and GEN is explained. This analysis also predicts that number and definiteness should be expressed in the nominals that are assigned DAT, since it is independently known that both categories are expressed in GEN. Next, this analysis of DAT as a non-relational P head correctly predicts affix ordering (like in GEN), since all possessive agreement must happen inside the DP complement of P. Finally, this proposal also predicts the difference in morphosyntactic properties between DAT and inherent cases, because the latter represent a different type of relational P heads independently attested in Moksha.

6. Conclusions

In this paper, I analyzed the unexpected morphosyntactic behavior of the Moksha DAT case. Based on their morphosyntactic properties, Moksha cases can be divided into two classes. Setting DAT aside, the two classes map onto the theoretical distinction between structural and inherent cases. DAT, however, poses problems for this initial classification, since it shows behavior characteristic of prototypical inherent cases, but its morphosyntactic properties align with those of the structural cases. The proposed analysis of Moksha inherent cases as instantiations of postpositional phrases and relational noun phrases resolves this problem. Based on the marking of the complement and possibility of ϕ -feature agreement, I introduced the distinction between relational and non-relational P heads. The distinction is needed in Moksha independent of any analysis of the DAT, because free-standing elements in PPs exhibit different properties. I argued that only relational P heads take bare complements, while non-relational P heads take GEN complements. DAT is analyzed as a PP headed by a non-relational P, whereas other PPs in Moksha are headed by relational Ps. The “dative” P selects a GEN complement, and the presence of that complement, in one of the structural cases, accounts for the similarity between the dative “case” on the one hand, and the structural cases (nominative and genitive), on the other.

These findings support the broader idea that the mapping between syntax and form does not have to be direct. For example, separate P heads can be encoded either by bound morphemes or by free-standing forms. This case study also suggests that divergence between form and meaning is important for grammatical theory not only in synchronic syntactic derivation (see [Beard 1986; McCawley 1987], among others), but in diachrony as well, particularly in grammaticalization. I have shown that the grammaticalization process behind the synchronically

distinguishable classes of P heads is less straightforward than previously assumed. For instance, the nature of a free-standing postposition (relational vs. non-relational) is not rendered obsolete in the course of grammaticalization to a bound morpheme in Moksha, but can be manifested in the morphosyntactic properties of emergent case markers.

ABBREVIATIONS

1, 2, 3 — 1 st , 2 nd , 3 rd person	NOM — nominative
ABL — ablative	NPST — nonpast tense
ACC — accusative	NZR — nominalizer
ADD — additive	O — object
ATTR — attributivizer	OBL — oblique stem
CAUS — causative	PASS — passive voice
DAT — dative	PL — plural
DEF — definite	POSS — possession
DIM — diminutive	PRS — present tense
EL — elative	PROL — prolative
FREQ — frequentative	PRON — pronominal
GEN — genitive	PST — past tense
ILL — illative	PTCP — participle
IMP — imperative	REFL — reflexive pronoun
IN — inessive	RES — resultative
INDEF — indefinite	Rus. — Russian
IPFV — imperfective	S — subject
LAT — lative	SG — singular
NEG — negation	

REFERENCES

- Abney 1987 — Abney S. *The English noun phrase in its sentential aspect*. Ph.D. diss., Massachusetts Institute of Technology, 1987.
- Alexiadou et al. 2007 — *Noun phrase in the generative perspective*. Alexiadou A., Haegeman L., Stavrou M. (eds.). Berlin: Mouton de Gruyter, 2007.
- Arkhangel'skiy, Usacheva 2015 — Arkhangel'skiy T., Usacheva M. Syntactic and morphosyntactic properties of postpositional phases in Beserman Udmurt as part-of-speech criteria. *SKY Journal of Linguistics*, 2015, 28: 103–137.
- Asbury 2008 — Asbury A. *The morphosyntax of case and adpositions*. Ph.D. diss., Utrecht: LOT, 2008. http://www.lotpublications.nl/Documents/180_fulltext.pdf.
- Beard 1986 — Beard R. *On the separation of derivation from morphology: Toward a lexeme*. Bloomington (IN): Indiana Univ. Linguistics Club, 1986.
- Biryuk 2005 — Biryuk O. L. *Morfosintaksicheskie kriterii vydeleniya poslelogov v ural'skikh yazykakh* [Morphosyntactic criteria for distinction of the category of postpositions in Uralic]. BA thesis, Moscow State Univ., 2005.
- Bresnan, Grimshaw 1978 — Bresnan J., Grimshaw J. The syntax of free relatives in English. *Linguistic Inquiry*, 1978, 9: 331–391.
- Caha 2009 — Caha P. *The Nanosyntax of Case*. Ph.D. diss., Univ. of Tromsø, 2009.
- Chomsky 1981 — Chomsky N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.
- Cinque (ed.) 2002 — Cinque G. (ed.). *The cartography of syntactic structures*. Vol. 1: *Functional structure in DP and IP*. New York: Oxford Univ. Press, 2002.
- Corbett 2012 — Corbett G. G. *Features*. Cambridge: Cambridge Univ. Press, 2012.
- Déchaine, Wiltschko 2002 — Déchaine R. M., Wiltschko M. Decomposing pronouns. *Linguistic Inquiry*, 2002, 33(3): 409–442.
- Dékány 2011 — Dékány É. K. *A profile of the Hungarian DP: The interaction of lexicalization, agreement and linearization with the functional sequence*. Ph.D. diss., Univ. of Tromsø, 2011.

- Emonds 1985 — Emonds J. *A unified theory of syntactic categories*. Dordrecht: Foris, 1985.
- Emonds 1987 — Emonds J. The invisible category principle. *Linguistic Inquiry*, 1987, 18: 613–632.
- Evseviev 1963 — Evseviev M. E. *Izbrannye trudy* [Selected works]. Vol. 4: *Osnovy mordovskoi grammatiki* [The basics of the Mordvin grammar]. Saransk: Mordvin Publishing House, 1963. (First published 1929.)
- Feoktistov 1975 — Feoktistov A. P. Mordovskie yazyki [Mordvinic languages]. *Osnovy finno-ugorskogo yazykoznaniya 2. Pribaltiisko-finnskie, saamskii i mordovskie yazyki* [The basics of the Finno-Ugristics 2. Finnic, Saami, and Mordvinic languages]. Lytkin V. I. et al. (eds.). Moscow: Nauka, 1975, 248–343.
- Grimshaw 1977 — Grimshaw J. B. *English wh-constructions and the theory of grammar*. Ph.D. diss., Univ. of Massachusetts Amherst, 1977.
- Grünthal 2003 — Grünthal R. *Finnic adpositions and cases in change*. Helsinki: Société Finno-Ougrienne, 2003.
- Guseva, Weisser 2018 — Guseva E., Weisser Ph. Postsyntactic reordering in the Mari nominal domain. *Natural Language & Linguistic Theory*, 2018, 1–39.
- Heine, Kuteva 2002 — Heine B., Kuteva T. *World lexicon of grammaticalization*. Cambridge: Cambridge Univ. Press, 2002.
- Kholodilova 2018 — Kholodilova M. A. Imennaya morfologiya [Nominal morphology]. *Elementy mokshanskogo yazyka v tipologicheskom osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 63–121.
- Kolyadenkov, Zavodova 1962 — *Grammatika mordovskikh (mokshanskogo i erzyanskogo) yazykov* [Grammar of the Mordvinic (Moksha and Erzya) languages]. Part 1. *Fonetika i morfologiya* [Phonetics and Morphology]. Kolyadenkov M. N., Zavodova R. A. (eds.). Saransk: Mordvin Publishing House, 1962.
- Koryakov, Kholodilova 2018 — Koryakov Yu. B., Kholodilova M. A. Obshchie svedeniya o mokshanskom yazyke i issleduemom govore [Basic information about the Moksha language and the variety under study]. *Elementy mokshanskogo yazyka v tipologicheskom osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 6–18.
- A. Kozlov et al. 2016 — Kozlov A. A., Muraviev N. A., Nikiforova S. O. Semantika i morfosintaksis odnoi mokshanskoi aspektual'noi derivatsii [Semantics and morphosyntax of one Moksha aspectual derivation]. *Acta linguistica Petropolitana*, 2016, XXII(1): 534–547.
- L. Kozlov 2018 — Kozlov L. S. Lokativnye padezhi [Locative cases]. *Elementy mokshanskogo yazyka v tipologicheskom osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 154–182.
- Lehmann 1985 — Lehmann Ch. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile*, 1985, 20: 303–318.
- Lyutikova 2018 — Lyutikova E. A. *Struktura imennoi gruppy v bezartiklevom yazyke* [Noun phrase structure in an articleless language]. Moscow: “YaSK” Publishing House, 2018.
- Lyutikova 2019 — Lyutikova E. A. Ierarkhicheskaya struktura imennoi gruppy [Hierarchical structure of the noun phrase]. *Elementy tatarskogo yazyka v tipologicheskom osveshchenii*. Tatevosov S. G., Pazel'skaya A. G., Suleimanov D. Sh. (eds.). Moscow: Buki Vedi, 2019, 330–338.
- Maitinskaya 1982 — Maitinskaya K. E. *Sluzhebnye slova v finno-ugorskikh yazykakh* [Functional words in Finno-Ugric languages]. Moscow: Nauka, 1982.
- Marantz 1991 — Marantz A. Case and licensing. *Eastern States Conf. on Linguistics*. Germán F., Westphal B. A., Hee-Rahk Ch. (eds.). Ithaca (NY): CLC Publications, 1991, 234–253.
- McCawley 1987 — McCawley J. D. A case of syntactic mimicry. *Functionalism in syntax*. Dirven R., Frid V. (eds.). Amsterdam: John Benjamins, 1987, 459–470.
- McFadden 2004 — McFadden Th. *The position of morphological case in the derivation: A study on the syntax-morphology interface*. Ph.D. diss., Univ. of Pennsylvania, 2004.
- Muravieva, Kholodilova 2018 — Muravieva A. M., Kholodilova M. A. Poslelogi i relyatsionnye imena [Postpositions and relational nouns]. *Elementy mokshanskogo yazyka v tipologicheskom osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 212–248.
- Nikanne 1993 — Nikanne U. On assigning semantic cases in Finnish. *Case and other functional categories in Finnish syntax*. Holmberg A., Nikanne U. (eds.). New York: Mouton de Gruyter, 1993, 75–88.
- Pereltsvaig 2006 — Pereltsvaig A. Small nominals. *Natural Language & Linguistic Theory*, 2006, 24(2): 433–500.
- Plaksina 2002 — Plaksina T. A. *Areal'noe issledovanie severo-zapadnykh govorov moksha-mordovskogo yazyka* [Areal study of North-Western varieties of Moksha]. Candidate diss., Evseviev Mordvin State Pedagogical Univ., 2002.

- Pleshak 2021 — Pleshak P. The exponence of caseless NPs in Moksha. *Proc Ling Soc Amer*, 2021, 6(1): 628–640. <https://doi.org/10.3765/plsa.v6i1.4999>.
- Pleshak, Kholodilova 2018 — Pleshak P. S., Kholodilova M. A. Imennaya gruppa [Noun phrase]. *Elementy mokshanskogo yazyka v tipologicheskoy osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 272–310.
- Polinsky 2016 — Polinsky M. *Deconstructing ergativity: Two types of ergative languages and their features*. Oxford: Oxford Univ. Press, 2016.
- Privizentseva 2019 — Privizentseva M. Free relative clauses and timing of case assignment in Moksha Mordvin. *Proc. of the 49th Meeting of the North East Linguistic Society*. Baird M., Göksu D., Pesetsky J. (eds.). Amherst (MA): GLSA Publications, 2019, 49–59.
- Riemsdijk 1990 — van Riemsdijk H. Functional prepositions. *Unity in diversity: Papers presented to Simon C. Dik on his 50th birthday*. Pinkster H., Gene I. (eds.). Boston (MA): De Gruyter Mouton, 1990, 229–241.
- Serebrennikov 1967 — Serebrennikov B. A. *Istoricheskaya morfologiya mordovskikh yazykov* [Historical morphology of Mordvinic languages]. Moscow: Nauka, 1967.
- Simonenko, Leontiev 2012 — Simonenko A., Leontiev A. Morfosintaksis imennogo kompleksa [Morpho-syntax of the noun complex]. *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisaniya. Formal'nyi i funktsional'nyi podkhody*. Kuznetsova A. A. (ed.). Moscow: Russkie Slovarei, 2012, 259–337.
- Svenonius 2010 — Svenonius P. Spatial P in English. *Mapping spatial PPs. The cartography of syntactic structures 6*. Cinque G., Rizzi L. (eds.). Oxford: Oxford Univ. Press, 2010, 127–160.
- Szabolcsi 1984 — Szabolcsi A. The possessor that ran away home. *The Linguistic Review*, 1984, 3(1): 69–102.
- Toldova 2018 — Toldova S. Yu. Predikatsii s glagol'nym skazuemym [Verbal predication]. *Elementy mokshanskogo yazyka v tipologicheskoy osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 546–615.
- Toldova et al. 2018 — *Elementy mokshanskogo yazyka v tipologicheskoy osveshchenii* [Elements of the Moksha language in typological perspective]. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018.
- Toldova, Pleshak 2018 — Toldova S., Pleshak P. Grammatical and lexical case distinction in Moksha and Hill Mari. Talk at the 4th Workshop on Languages of the Volga–Kama Sprachbund, Moscow, HSE Univ. / Institute of Linguistics, 16–18 November 2018.
- de Vries 2002 — de Vries M. *The syntax of relativization*. Academisch Proefschrift. Amsterdam: Univ. of Amsterdam, 2002.
- Woolford 2006 — Woolford E. Lexical case, inherent case, and argument structure. *Linguistic Inquiry*, 2006, 37: 111–130.
- Zakirova 2018 — Zakirova A. Nominalizatsii [Nominalizations]. *Elementy mokshanskogo yazyka v tipologicheskoy osveshchenii*. Toldova S. Yu., Kholodilova M. A. et al. (eds.). Moscow: Buki Vedi, 2018, 753–778.

Получено / received 01.02.2022

Принято / accepted 29.03.2022

On the exclusively borderline case-marking in Ainu

© 2022

Anna Bugaeva

Tokyo University of Science, Tokyo, Japan; bugaeva@rs.tus.ac.jp

Abstract: This paper argues that the system of marking case relations in Ainu represents a typical example of borderline case-marking since arguments are unmarked and it is difficult to distinguish case markers from adverbs. For instance, dative, instrumental, comitative, and other cases are marked by postpositional adverbs, which can occur without respective NPs and some can even take indexing. It is shown that such an unusual zero anaphora-like behavior of adverbial “case” postpositions can be explained by their relatedness to respective transitive verbs out of which they have not yet fully grammaticalized. I adduce a tentative hypothesis that case-markers in Ainu are relatively young and innovative, while applicatives, which are a functional alternative to case markers, are old and genuine, and that the borderline case-marking may generally correlate with the head-marking language type, which is most broadly attested in the Pacific Rim.

Keywords: Ainu, applicative, case, head marking, Pacific

Acknowledgements: The present study was supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology Grant-in-Aid Research Program, project “Towards understanding dynamics of language change in Ainu” (No. 17K02743; 2017–2022) (Kiban C, Principal Investigator: Anna Bugaeva).

For citation: Bugaeva A. On the exclusively borderline case-marking in Ainu. *Voprosy Jazykoznanija*, 2022, 5: 131–150.

DOI: 10.31857/0373-658X.2022.5.131-150

Об исключительно пограничном падежном маркировании в айнском языке

Анна Бугаева

Токийский университет естественных наук, Токио, Япония; bugaeva@rs.tus.ac.jp

Аннотация: В настоящей статье утверждается, что система маркирования падежных отношений в айнском языке представляет собой типичный случай исключительно пограничного падежного маркирования, поскольку актанты не маркируются, а также представляется затруднительным разграничение падежных показателей и наречий. Например, дательный, творительный, совместный и другие падежи маркируются постпозиционными наречиями, которые могут употребляться отдельно без соответствующих именных групп, а некоторые даже могут принимать согласовательные аффиксы. Как показано в статье, подобное необычное, напоминающее анафору, поведение наречных «падежных» постпозитивов может быть объяснено их происхождением и все еще неполной грамматикализацией от переходных глаголов. Автор выдвигает предварительную гипотезу о том, что падежи в айнском языке относительно молодые и представляют собой новый тип, в то время как аппликативы, которые являются функциональной альтернативой падежных показателей, архаичны и исконны, и что пограничное падежное маркирование, возможно, коррелирует с явлением вершинного маркирования, которое особенно характерно для языков «тихоокеанского кольца».

Ключевые слова: айнский язык, аппликатив, вершинное маркирование, падеж, тихоокеанский ареал

Благодарности: Настоящее исследование выполнено при поддержке программы научных грантов Министерства образования, культуры, спорта, науки и технологий Японии, проект «Изучение динамики языковых изменений в айнском языке» (№ 17K02743; 2017–2022) (тип гранта С, руководитель — Анна Бугаева).

Для цитирования: Bugaeva A. On the exclusively borderline case-marking in Ainu. *Voprosy Jazykoznanija*, 2022, 5: 131–150.

DOI: 10.31857/0373-658X.2022.5.131-150

1. Introduction

Ainu (isolate), a head-marking language formerly spoken in Hokkaido, Sakhalin, and Kurils, has exclusively borderline case-marking. This means that nominative, accusative, and genitive cases are absent and that there is overt marking only for concrete or “peripheral” case relations and delimiting case markers from other morphological-functional entities is difficult [Iggesen 2013]. Moreover, adnominal phrases clearly employ not case markers, but existential transitive verbs, which form a relative-clause-like structure.

This paper argues that exclusively borderline case-marking in Ainu correlates with an extensive use of applicatives, which are regarded here as a head-marking alternative to case markers. Overall, I hypothesize that borderline case-marking may generally correlate with the head-marking language type, but a typological survey based on a broad sample of languages is a matter of future research.

The data in this paper are mainly from South Hokkaido dialects of Chitose and Saru, and include the author’s own fieldwork, as well as previous documentation on Ainu, particularly *A Glossed Audio Corpus of Ainu Folklore* [Nakagawa et al. 2016–2021], which is available online. South Hokkaido Ainu dialects belong to the Southwestern Hokkaido group (SW), which shows contrast with the Northeastern Hokkaido group (NE).

After a few introductory comments on the typological profile of Ainu in Section 2, it will be shown in Sections 3 and 4 that Ainu demonstrates a typical example of exclusively borderline case-marking as defined above. Next, Section 5 focuses on marking syntactic relations in adnominal phrases, while Section 6 discusses applicatives as head-marking case markers. Section 7 summarizes the paper.

2. Typological features of Ainu

Ainu is strikingly different from other languages of Northeast Asia such as Altaic languages, Korean, and Japanese. Like these languages, Ainu is SOV and agglutinating, but unlike them it is polysynthetic, incorporating, and head-marking. According to Bugaeva [2022], most key features in Ainu are related to head-marking, which is widely attested in the languages of the Pacific Rim, i.e., coastal areas around the Pacific including northern Australia, the north coast of New Guinea, Oceania, coastal eastern Asia, western coastal North America, Mexico and Central America, and western coastal South America [Nichols 2017].

Though unrelated to head-marking, Ainu shows mixed alignment¹ depending on person and number (see the Table): nominative–accusative in the first person singular, neutral in the second person singular and plural, zero-marked third person, tripartite alignment in the first person plural (exclusive) and “fourth person” (see below), and hierarchical alignment when it comes to the interaction of the first person subject and second person object (2>1).

¹ An excellent account of verb indexing in Ainu can be found in Comrie [to appear].

Table

Person-number affixes in South Hokkaido Ainu [Bugaeva 2012: 471]

		A	S	O
1	SG	<i>ku- / k-</i>		<i>en-</i>
	PL	<i>ci- / c-</i>	<i>-as</i>	<i>un-</i>
2	SG		<i>e-</i>	
	PL		<i>eci-</i>	
4		<i>a-</i>	<i>-an</i>	<i>i-</i>

S — intransitive subject, A — transitive subject (also used for possessors), O — object.

“Fourth person” is a common label for the indefinite (= impersonal passive), first person plural inclusive, second person (singular and plural) honorific, and logophoric (= person of the protagonist in folktales), which is conventionally translated as ‘I’.

Other typological relevant features of Ainu include verbal plurality partially marked by suppletion, extensive voice system encoded by five causative, three applicative, two reflexive, reciprocal, and antipassive affixes, ability of intransitive verbs to function as nouns, a four-term evidential system, possessibility oppositions in nouns, and the absence of pure tense markers and non-finite verbal forms.

3. Arguments are unmarked for case

As mentioned in Section 1, exclusively borderline case-marking, which is a key feature of Ainu, has the following characteristics:

- (a) arguments are unmarked for case;
- (b) only “peripheral” participants are case-marked;
- (c) case markers are almost indistinguishable from other word classes.

This section will focus on feature (a), see examples below.

Example (1) is an intransitive clause with a single subject argument being indexed on the predicate. Though *cóka* ‘we (EXCL)’ appears in (1), it is most natural not to use pronominal NPs and to omit nominal NPs unless you need them for emphasis or introducing new participants in discourse.

- (1) *cóka mina-as*
 1PL.EXCL laugh-1PL.EXCL.S
 ‘We laughed.’ [Tamura 1979: 3]

Example (2) is a transitive clause with two arguments being indexed on the predicate. However, such cases are rare because all third person arguments are zero-marked, as the object in (3) and (4) and subject in (5), and also because of above-mentioned hierarchical alignment (2>1), which never results in a straightforward addition of two markers [Bugaeva 2012: 474].

- (2) *hi-oy-oy. nei ta pak-no eci-un-nunuke wa iyayiraykere.*
 thanks where LOC till-ADV 2PL.A-1PL.EXCL.O-be.kind.to and thank.you
 ‘Thanks. You have always been so kind to us. Thank you.’ [Tamura 1979: 19]
- (3) *cep e-koyki ya?*
 fish 2SG.A-catch Q
 ‘Did you catch fish?’ [Tamura 1979: 4]

- (4) *cep ku-koyki ruwe un*
 fish 1SG.A-catch INFR.EV SFP
 ‘Yes, I caught fish.’ [Tamura 1979: 4]
- (5) *ku-sa-ha en-oputuye*
 1SG.A-elder.sister-PSD 1SG.O-push
 ‘My elder sister pushed me.’ [Tamura 1979: 15]

Moreover, it is not only the subject and object NPs that are unmarked for case, but also the genitive is lacking. Instead, in the possessive phrase *ku-sa-ha* (1SG.A-elder.sister-PSD) ‘my elder sister’, the “possessed” suffix (PSD) *-hV* on the head (= possessee) registers a possessor, while person-marking prefix (A) indexes a possessor, for details see Bugaeva et al. [2022].

4. “Peripheral” cases: How to distinguish them from other word classes?

Only “peripheral” participants are case-marked in Ainu, but, in fact, it is quite difficult to come up with a really neat closed list of cases in Ainu because they do not enter into paradigmatic relations and because it is not easy to distinguish them from other word classes. Also, peripheral case markers vary considerably across the dialects.

It is generally accepted in modern Ainu studies that there are two classes of words expressing case relations: postpositional particles and postpositional adverbs [Tamura 1988/2000: 129–135, 95–102; Kirikae 1989; 2003: 40–42; Okuda 1998; 1999; Satō 2008: 33–42], but that was not the case in the earlier studies of Ainu [Kindaichi 1931/1993: 179–188; Kindaichi, Chiri 1936/1974: 35–38; Chiri 1942/1973: 575–582; 1953/1976: 1–20; Fukuda (Tamura) 1960/2001: 57; Asai 1969/2001: 36–37; Refsing 1986: 164–170], which lump all case markers together and some even mix in a few other “non-case” particles.

According to Tamura [1988/2000: 129–135, 95–102], Okuda [1999], and Satō [2008: 33–42], the following three criteria can be used to distinguish genuine postpositional case particles from postpositional adverbs:

- (i) ability to attach directly only to locative nouns of category 2 (most place names, *kim* ‘mountains’, *pis* ‘beach’, *rep* ‘ocean’, *ya* ‘shore’, *wor* ‘water’, *uni* ‘home’, *rik* ‘high place’, *ra* ‘low place’, *pena* ‘upstream’, *pana* ‘downstream’ etc.) as in (6), while with ordinary nouns an additional use of locative nouns of category 1 is required (e.g., *or* ‘place’, *sam* ‘near’, *ka* ‘top’, *corpok* ‘below’ etc.)², as in (7a), cf. (7b);
- (ii) inability to be used independently;
- (iii) inability to take verbal pronominal affixes of the subject or object.

However, a close look reveals that criteria (i) and (ii) partially intersect³. Thus Tamura [1988/2000: 129–135], who prioritizes (i) over (ii), comes up with a list of five postpositional particles for the Saru dialect of Ainu: locative *ta*, allative / location of action *un*, ablative *wa*, traversal *peka* ‘over’, and mutative *ne* ‘as’. As we can see, all but one case markers are locative / spatial and indeed they may be preceded by locative nouns only, as in (6) and (7a), but the last case marker *ne* ‘as’, though meeting criterion (ii), does not meet (i) and allows for direct use after ordinary nouns as in (8), so it should better be excluded from the list⁴.

² For the details of categorization of nouns in Ainu see Nakagawa [2022].

³ If a case marker does not meet (iii), it also does not meet (i) and (ii), in other words, the ability to be indexed correlates with the ability to occur with all types of nouns and implies independent use, which automatically makes it fall into the category of postpositional adverbs.

⁴ This is also hinted at in Satō [2008: 36].

- (6) *sine ar suy ne suy pis ta sa-n-an*
 one one/half times COP again beach LOC front.place-INTR.SG-4.S
 ‘One day, I was again down on the beach (fishing).’
 [Nakagawa et al. 2016–2021: K7708241UP.005]
- (7) a. *a-kor ekasi or-o ta MATA sa-p-an*⁵
 1SG.A-have grandfather place-PSD LOC again front.place-INTR.PL-4.S
 ‘I went down to my grandfather’s place again.’ [Onseishiryō 5: 28]
- b. **a-kor ekasi ta MATA sa-p-an*
 1SG.A-have grandfather LOC again front.place-INTR.PL-4.S
 Intended meaning: ‘I went down to my grandfather again.’
- (8) *onne ekasi ne a-e-kar*
 be.old grandfather/elder as 4.A-2SG.O-make
 ‘I will transform you into an old man.’
 [Nakagawa et al. 2016–2021: K7807151UP.045]

Satō [2008: 33–38] lists the above postpositional case particles plus the destination/degree *pak(no)* ‘till, until’,⁶ which also meets criterion (i), cf. (9), but does not meet criterion (ii), cf. its independent use in (10); Satō does not discuss the latter fact.

- (9) *TOKIYO pak-no k-arpa rusuy korka henpak to*
 Tokyo till-ADV 1SG.S-go.SG DESI but how.many day
ne kor sirepa-an pe an?
 COP and arrive-4SG.S NMLZ exist.SG
 ‘I want to go to Tokyo, but how long would it take to arrive (there)?’ [Satō 2008: 36]
- (10) *pak-no ka ku-ye*
 till-ADV even/also 1SG.A-say/tell
 ‘I will say till [here].’ (= ‘I finished speaking.’) [Tamura 1993: 76]

And finally, Okuda [1999], who uses data of the Shizunai dialect of Ainu (NE), prioritizes (ii) over (i) and as a result comes up with a much longer list of bound case postpositions, further subcategorizing them by criterion (i) into those that can follow only locative nouns (namely, locative *ta*, allative/location of action *un/unno/en/enno*, ablative *wa/wano*, and traversal *peka* ‘over’) and those that can follow any noun (*ekari* ‘around, along’, *ekopas* ‘against’, *epitta* ‘all over’, *epeka* ‘for, with a purpose of’, *hekota* ‘facing, towards’, *ne* ‘as’, *okari* ‘around’, and *tomotuye* ‘across’). Syntactically, this is probably a less self-contradicting classification, but semantically it poses many questions, since the latter group of words is semantically closer to a much broader open class of postpositional adverbs such as *esoro* ‘along a river/mountain downstream/downwards’, *kari* ‘through’, *okannatki* ‘going full circle around’, *pekanō* ‘facing’, *turasi* ‘along a river/mountain upstream/upwards’, *eirpak* ‘simultaneously’, *os* ‘after’ etc. (27 items altogether), which generally allow for independent use and most of them for the subject/object indexing [Okuda 1998: 133–146]. Surprisingly, when classified on purely syntactic grounds, postpositional “case” adverbs also appear to include such canonical peripheral cases as *eun* ‘dative’, *tura(no)* ‘comitative’, and *ani/ari*⁷ ‘instrumental’ (see Section 4.2.2).

Unfortunately, I am not able to offer an essentially new classification of Ainu case markers simply because we are dealing with “the exclusively borderline case-marking”, where “delimiting

⁵ According to Tamura’s Japanese translation, the subject referent is singular, and it is unclear why *sap* ‘come down (PL)’ is used.

⁶ In fact, Satō [2008: 36, 40] subclassifies *pak(no)* into a destination postpositional case particle ‘to/till’ and degree postpositional adverb ‘till/until, as’.

⁷ *Ani* is used in SW dialects of Ainu, while *ari* is used in NE.

case markers from other morphological-functional entities is difficult” by definition [Iggesen 2013]. For the sake of convenience, I suggest applying strictly both criteria (i) and (ii) as diagnostic, which will leave us with only four postpositional case particles: locative *ta*, allative / location of action *un*, ablative *wa*, and *peka* ‘over’ (see Section 4.1); all the rest should go into postpositional adverbs (Section 4.2) based on their independent use capacity and, in most cases, also very peripheral adverbial semantics⁸. We can further classify postpositional adverbs into those not allowing for independent use (Section 4.2.1), those allowing for independent use but not allowing for indexing (Section 4.2.2), and those allowing for both independent use and indexing (Section 4.2.3).

Also, compared to case markers in European languages, Ainu case markers show little poly-functionality; this and other issues will be discussed in the ensuing sections. Note that some markers have variants with an adverbial morpheme *-no* (ADV) without a clear semantic difference.

4.1. Postpositional case particles

4.1.1. Locative *ta*

This particle indicates a fixed temporal or spatial location: ‘at, in’ [Tamura 1988/2000: 130], cf. the first and second occurrences of *ta* respectively in (11). Moreover, unlike a fixed spatial location in (12), the second occurrence of “location” in (11) is more of a motion, yet the one in which an achievement (destination) is emphasized. When the motion itself is emphasized, an allative particle *un* is used instead (cf. Section 4.1.2).

- (11) *utur-u ta atuy teksam ta*
 space/time.interval-PSD LOC sea/ocean next.to LOC
sa-n-an wa peray-an
 front.place-INTR.SG-4.S and fish-4.s

‘Sometimes I go down to the seashore to fish.’

[Nakagawa et al. 2016–2021: K7708241UP.003]

- (12) *ora ape sam ta an a-aca ram-no*
 then fire near LOC exist.SG 4.A-uncle.PSD be.low-ADV
kane i-e-rankarap
 somewhat 4.O-to.APPL-greet

‘My uncle who was sitting near the hearth politely greeted me.’

[Nakagawa et al. 2016–2021: K8010291UP.068]

Note also that the first *ta* in (11) occurs after a locative noun *utur-u* lit. ‘the interval of (it)’, while the second *ta* in (11) and *ta* in (12) after ordinary nouns *atuy* ‘sea / ocean’ and *ape* ‘fire’ respectively, which are necessarily followed by locative nouns *teksam* ‘next to’ and *sam* ‘near’.

The origin of *ta* ‘at, in’ is not clearly known but it may be related to the transitive verb *ta* ‘be / exist at’ appearing in adnominal phrases (see Section 5).

4.1.2. Allative / location of action *un*

This particle indicates the place towards which a motion or action goes: ‘to, towards’ [Tamura 1988/2000: 131] and therefore combines allative (13) and locative of action (14) functions.

⁸ This approach is most like Satō [2008], but there are some differences, cf. discussion above.

- (13) *a-aca-ha ror-ke un a-sike-he a-an-u*
 4.A-uncle-PSD seat.of.honour.in.a.house-place ALL 4.A-luggage-PSD 4.A-exist-TR.SG
 ‘I placed my luggage at my uncle’s place of honor.’
 [Nakagawa et al. 2016–2021: K8010291UP.067]
- (14) *hoka un tuy-pa wa hoka o ne ya nen-nen ki*
 flame ALL cut-TR.PL and flame put.sth.on.sth COP Q any-any do
 ‘He cut (some meat) in the flames and placed it on the fire (=sacrificed it to the Fire god-
 dess), and so on.’ [Nakagawa et al. 2016–2021: K8010291UP.108]
- As mentioned in Section 4.1.1, not only allative *un* but also locative *ta* can occur with motion verbs such as *arpa/paye* ‘go (SG / PL)’ in (15) and (16). As previously pointed out in Tamura [1988/2000: 133], there is usually a semantic difference between those allative uses: *un* typically focuses on motion (15), while *ta* focuses on achievement (16). Yet, there are a few cases where *un* too seems to focus on achievement and the distinction between the two is less straightforward [Satō 2008: 38], cf. also example of *ta*-marked caused motion in (17).
- (15) *a-uni un paye-an pe ne na*
 4.A-home.PSD ALL go.PL-4.S NMLZ COP SFP
 ‘So let’s go to my house.’ [Nakagawa et al. 2016–2021: K8010291UP.160]
- (16) *orowa KONTA oyak ta arpa-an*
 then then different.place LOC go.SG-4.S
 ‘Then I went to live in another place.’ [Nakagawa et al. 2016–2021: K7708241UP.090]
- (17) *ora su oha hi kusu ora*
 then pan be.open/vacant thing/place/time because then
soy ta rur-i a-kuta
 outside LOC soup-PSD 4.A-throw.out
 ‘When the pot was empty, I threw the soup outside.’
 [Nakagawa et al. 2016–2021: K7908032UP.158]

4.1.3. Ablative *wa*

This particle “indicates the starting point of an action” [Tamura 1988/2000: 132].

- (18) *Kusur wa topattumi ek hine*
 Kusur ABL night.raid(er) come.SG and
 ‘The night raiders from Kusur came (and destroyed the village).’
 [Nakagawa et al. 2016–2021: K7803231UP.088]

Wa is often followed by the adverb-forming suffix *-no*, especially when indicating time, which results in *wa-no* as in (19).

- (19) *KONTA a-yup-ihī utar nani ne ancikari*
 then 4.A-older.brother-PSD PL immediately this/that night
wa-no sake kar pa hine
 from-ADV liquor make PL and
 ‘So, my older brothers started making sake immediately from that night.’
 [Nakagawa et al. 2016–2021: K7807151UP.174]

According to Tamura [1988/2000: 133], “after a word indicating a place to sit, *wa*, not *ta*, is frequently used.” She claims that the usage in example (20) has the nuance of “Go to your mother’s side.”

- (20) *hapo harki-sam-a wa rok yan*
 mother left-near-PSD ABL sit.PL IMP.POL
 ‘Go sit on your mother’s left side.’ [Tamura 1988/2000: 133]

And finally, *wa(no)* is used to encode a demoted animate Agent in the impersonal passive construction in *a-*, cf. instrumental *ani*, which is used to encode a demoted inanimate Agent in Section 4.2.2.1.1; for details see Bugaeva [2011].

- (21) *cicikew-nitney anakne mintuci TONO or-o wa-no*
 Chichikew-nitnei TOP Mintuchi lord place-PSD ABL-ADV
kamuy a-ona or-o wa-no
 god 4.A-father.PSD place-PSD from-ADV
ar-teyne mosir a-ko-o-terke ruwe ne
 completely-be.wet land 4.A-to.APPL-to.APPL-jump INFR.EV COP
 ‘Chichikew-nitnei was kicked down to the nether world underground by Mintuchi lord and my god father.’ lit. ‘Someone kicked Chichikew-nitnei... by Mintuchi lord and my god father.’ [Nakagawa et al. 2016–2021: K8303243UP.203–204]

4.1.4. Traversal *peka* ‘through, around, over’

According to Tamura [1988/2000: 133], this particle indicates the extension of the range of an action, i.e., shows that an action occurs over a wide area, otherwise if an action occurs at one point, the locative *ta* is used. However, it should be added that if “a wide area” semantic component applies, *peka* can replace not only the locative *ta* as in (22) and (23), but also the allative / location of action *un* as in (24) and (25), and even ablative *wa* as in (26).

- (22) *o-harki-so-un*
 bottom.POSS.PFX-left.(from.the.east.window)-seat.near.fireplace-belong.to
ape-kes-utur kopak peka mo-no a
 fire-end-space side.of TRV quiet-ADV sit.SG
 ‘She sat by the far side of fire at the left upwind seat.’
 [Nakagawa et al. 2016–2021: K8010291UP.119]
- (23) *tononoski paki peka e-soy-ne aynu*
 noon time TRV head.POSS.PFX-outside-COP human
hum as kusu soy-ne-an
 sound stand.SG because outside-COP-4.s
 ‘Around noon, I could hear someone approaching the house so I went outside.’
 [Nakagawa et al. 2016–2021: K7708242UP.115]
- (24) *pis peka arki pa hine oka-an a*
 beach TRV come.PL PL and exist.PL-4.S PRE.SG
 ‘They came to the beach to where we live.’
 [Nakagawa et al. 2016–2021: K7908051UP.191]
- (25) *okkayo hekaci tek-e an-i kane wa kotan*
 man child hand-PSD hold-TR.SG doing.so and village
or peka sinewe kor oka-an
 place TRV visit and exist.PL-4.s
 ‘He held my son’s hand as they walked around the village paying visits.’
 [Nakagawa et al. 2016–2021: K8010301UP.128]
- (26) *a-ona a-unu-hu ror peka utur*
 4.A-father.PSD 4.A-mother-PSD seat.of.honor TRV the.lower.seat

peka *harki-so* *turasi* *ek*
 TRV left.(from.the.east.window)-seat.near.fireplace along.upstream come.SG
 ‘Father and mother came down from the high and right seats (respectively) to (my) left seat.’
 [Nakagawa et al. 2016–2021: K8303243UP.212]

Interestingly, the traversal *peka* can not only take inanimate nouns with the meaning of location and time but also animate nouns as in (27), which should, of course, be followed by locative nouns; this usage is not reported in previous studies.

(27) *oraun a-mat-ne-po utar ka nispa or peka a-kor-ar*
 then 4.A-woman-COP-child PL even rich.man place TRV 4.A-have-INDF.CAUS
 ‘I give away my daughters in marriage to wealthy men.’
 [Nakagawa et al. 2016–2021: K7803232UP.143]

The case particle *peka* originates in the transitive verb *péka* ‘catch (something moving through the air)’ [Tamura 1988/2000: 134].

4.2. Postpositional “case” adverbs

This section does not intend to provide an exhaustive list of case-markers in question. As mentioned, it is hardly possible because of the exclusively borderline character of case-marking in Ainu. The biggest problem is where to draw a line between postpositional case adverbs and all other adverbs, given the ability of most of them to occur independently; recall examples from Okuda [1998] in Section 4. Thus, I am going to focus on a few representative examples mainly from our corpus [Nakagawa et al. 2016–2021] trying to sub-categorize them.

4.2.1. Not allowing for independent use

Contrary to Okuda [1999: 162–164], who provides a list of eight bound adverbs expressing case relations⁹, I could attest only two such words, i.e., mutative *ne* ‘as/into’ and *pes* ‘along a river/mountain downstream/downwards’. I suggest that the ability of independent use closely correlates with the presence of and yet strong connection to respective transitive verbs, which normally allow for an anaphoric omission of the object. It may not be accidental that no respective transitive verb has been attested for *pes* ‘along a river/mountain downstream/downwards’ (Section 4.2.1.2) and that the mutative *ne* ‘as/into’ (Section 4.2.1.1) originates in a very special transitive copula verb *ne* ‘be/become sb/sth’ (for details see Bugaeva [2012: 473–474]), which does not allow for the omission of its object (= copula complement).

4.2.1.1. Mutative *ne* ‘as/into’

This particle means ‘as’ with action verbs (28), but is better translated as ‘into’ with action verbs of transformation as in (29) and (8).

(28) *yaykata po-ho ne i-e-ynnonnitak*
 by.oneself child-PSD as 4.O-about.APPL-pray
 ‘He prayed for me as for his own son.’
 [Nakagawa et al. 2016–2021: K7803232UP.091]

⁹ In fact, most of Okuda’s apparent case particles do allow for independent use.

- (29) *aynu ne yay-kar-an*
 human as REFL-make-4.s
 ‘I transformed myself into a human.’ [Nakagawa et al. 2016–2021: K7807151UP.027]

4.2.1.2. *Pes* ‘along a river/mountain downstream/downwards’

Rivers and mountains are major orientation points in the worldview of the Ainu, which is naturally reflected in their language. In fact, every motion verb of ‘going’ or ‘coming’ is specified for a particular direction of motion regarding the location of rivers / mountains. There are also a few adverbs additionally specifying the direction of movement.

- (30) *pet pes sa-p-an akusu*
 river downstream front.place-INTR.PL-4.s then
 ‘We went down the river (and came to a large village).’
 [Nakagawa et al. 2016–2021: K8108011UP.035]
- (31) *orano situ pes a-nin-pa-nin-pa*
 then ridge downstream 4.A-pull.along-TR.PL-pull.along-TR.PL
 ‘(I tied up four of them, two Japanese and two Ainu,) and pulled them down the mountain ridge.’ [Nakagawa et al. 2016–2021: K8108011UP.087]

Though the respective transitive verb is not registered in dictionaries, its ditransitive causative version *pes-te* ‘make sb / sth follow sth down’ is well attested [Tamura 1996: 524].

4.2.2. Allowing for independent use

Postpositional “case” adverbs can further be subdivided into those that take indexing and those that do not. The reasons why they do not take indexing may vary: they can either be grammaticalized or simply cooccur only with third person participants, recall that the third person is zero-marked (Section 3).

4.2.2.1. Without indexing

Ani ‘by’, *tomotuye* ‘across’, and *turasi* ‘along a river / mountain upstream / upwards’ are mentioned in Tamura [1988/2000: 100–102] as object indexing taking postpositional adverbs, but I could not attest such a use either in Tamura’s materials or in mine.

4.2.2.1.1. Instrumental *ani* ‘by’

This adverb can express instrumental proper (‘using, by means of’) (32) and a number of other closely related semantic functions (33).

- (32) *i-ruska haw ani ene haw-e-an hi*
 ANTIP-be.angry.with.sth voice INS like.this voice-PSD-exist.SG NMLZ
 ‘He said this in an angry voice.’ [Nakagawa et al. 2016–2021: K8106233UP.111]
- (33) *ruyanpe ani sorekusu nen ka arpa*
 storm INS exactly who even go.SG
 ‘No one would want to walk in the storm.’
 [Nakagawa et al. 2016–2021: K8106233UP.120]

Ani is also used for encoding a demoted inanimate Agent in the impersonal passive construction in *a-*, cf. ablative *wa(no)*, which is used to encode a demoted animate Agent in Section 4.1.3; for details see Bugaeva [2011].

- (34) *ci-kosinninu-p ani a-i-maketa ruwe ne*
 1PL.EXCL.A-have.sth.as.an.amulet-thing INS 4.A-4.O-beat INFR.EV COP
 ‘I lost because of the power of the amulet.’ lit. ‘Someone beat me by the thing that we use as an amulet.’ [Nakagawa et al. 2016–2021: K8010291UP.311]

Postpositional adverb *ani* originates in the transitive verb *ani* ‘hold sth/sb’ and allows for independent use without a preceding NP, cf. (35a) with an NP and (35b) without it. This is possible because, just as many other postpositional adverbs, *ani* is still in an early stage of grammaticalization from the verb, though it does not any longer allow for indexing and plural marking on the stem (cf. *an-pa* ‘hold sth (PL)’).

- (35) a. *a-kor ekasi inaw ani sake ani a-nomi.*
 4.A-have grandfather inaw INS rice.wine INS 4.A-pray
 ‘I worshipped my grandfather with *inaw*-willow prayer sticks and wine.’
 [Nakagawa et al. 2016–2021: K7807151UP.265]
- b. *ani e-aynu-wen-te pa pe ne*
 INS by.APPL-human-be.bad-CAUS PL NMLZ COP
 ‘They were suffering because of (it).’ [Nakagawa et al. 2016–2021: K7908051UP.233]

4.2.2.1.2. Dative *eun* ‘to/towards’

Two-place transitive (ditransitive) verbs like *kore* ‘give sth/sb to sb’ do not require this dative adverb, but intransitives as in (36) and one-place transitives as in (37)–(40), especially transfer and speech verbs, do when it is necessary to encode a Recipient/ Addressee, as in (37), (36), and (39), or Beneficiary (38), or Goal (40).

- (36) *sekor ona-utar-i eun haw-e-an pe ne*
 QUOT father-PL-PSD DAT voice-PSD-exist.SG NMLZ COP
 ‘As she told her father and relatives.’
 [Nakagawa et al. 2016–2021: K8106233UP.218]
- (37) *makan a-ray-ke ki pe a-unu eun a-se*
 how 4.A-die-CAUS do NMLZ 4.A-mother.PSD DAT 4.A-shoulder
 ‘Once, I took to my mother some game I had caught.’
 [Nakagawa et al. 2016–2021: K8010291UP.338]
- (38) *cise eun ki o-tuy-e*
 house DAT reeds bottom.POSS.PFX-cut-TR.SG
 ‘Go cut some grass for a house.’ [Nakagawa et al. 2016–2021: K8010291UP.227]

Just like the instrumental marker *ani* in the previous section, the dative marker *eun* perfectly allows for independent use as if it still were a verb *eun* ‘be at/on, stick to’ with an anaphorically omitted object (in parenthesis). Yet, there is no conjunction *wa* ‘and’ between it and the next verb, which shows that it has already undergone a certain degree of grammaticalization and is no longer a verb in this context.

- (39) *iteki eun itak ki, nen ka haw-e-an yak pirka.*
 PROH DAT speech do what even voice-PSD-exist.SG if be.good
 ‘It’s better not to say anything to (mother).’
 [Nakagawa et al. 2016–2021: K8010291UP.465]

- (40) *sem ka a-kar wa eun ni a-rura*
 porch even 4.A-make and DAT firewood 4.A-carry
 ‘I built a porch and piled firewood in (there) (so that you can build a fire throughout the winter).’ [Nakagawa et al. 2016–2021: K7908032UP.224]

4.2.2.1.3. *Kari* ‘through’

This adverb too can attach to an NP (41) and be used independently (42).

- (41) *kuca or ta saramanto kari*
 hunting.hut place LOC ceiling.flue through
tur-se-an hi ne an hine
 fall-INTR-4.S thing/place/time COP exist.SG and
 ‘I had fallen through the skylight of the hunting cabin.’
 [Nakagawa et al. 2016–2021: K8106233UP.026]

- (42) *atuy sinpuy an hine kari ahu-p*
 sea well exist.SG and through inside.house-INTR.PL
 ‘There was a sea bubbling [spot] and through [it] they entered’. [Onseishiryō 3: 28]

According to Tamura [1996: 281], the corresponding transitive verb *kari* ‘turn, spin sth / sb’ is attested only in compounds.

4.2.2.1.4. *Tomotuye* ‘across’

Although this adverb can occur either with an NP (43) or independently (44), the latter use is rare; cf. the transitive verb *tomotuyya* ‘cross, cut cross sth (PL)’ in (59).

- (43) *atuy tomotuye ikesuy-an wa arpa-an*
 sea across go.away.in.anger-3.s and go.SG-4.s
 ‘I went away in anger crossing the sea.’ [Onseishiryō 1: 32]
- (44) *tomotuye tem ne aynu hinak wa ek*
 across fathom COP human where ABL come.SG
 ‘A [huge] man, [who] was a fathom (= 1.829 m) across [his width], came in.’ [Kubodera 1977: 450]

4.2.2.1.5. *Turasi* ‘along a river / mountain upstream / upwards’

This adverb originates in a transitive verb, namely *turasi* ‘follow a river upstream’ [Tamura 1996: 738].

- (45) *pet turasi paye-an pa ruwe ne*
 river upstream go.PL-4.S PL INFR.EV COP
 ‘We went up the river.’ [Nakagawa et al. 2016–2021: K7803231UP.159]
- (46) *a-e-kay kane hine nay turasi*
 4.A-2SG.O-carry.on.one’s.back doing.so and stream upstream
sinot-an kor cep-po-koyki-an kor arpa-an
 play-4.s and fish-DIM-catch-4.s and go.SG-4.s
 ‘I was playing upstream at the mountain stream with you on my back. I was catching small fish and walking up the stream.’ [Nakagawa et al. 2016–2021: K7908051UP.026]

Here is an example of the independent use of this adverb.

- (47) *ane ru ne wa an nankor kusu,*
 thin footprint COP and exist.SG probably because
turasi e-arpa, e-arpa ayne an-an uske
 upstream 2SG.S-go.SG 2SG.S-go.SG finally exist.SG-4.S place
un-no e-arpa nankor
 ALL-ADV 2SG.S-go.SG probably

‘There are probably thin footprints, if you go and go following them upstream you will go to the place where I am.’ [Onseishiryō 3: 8]

4.2.2.2. With indexing

Seven adverbs were chosen to be included here. Six of them take object indexing and only one, namely *renkayne* ‘accordingly’ (Section 4.2.2.2.6), subject indexing.

4.2.2.2.1. *Akkari* ‘than’

This adverb is used to encode the standard of comparison, as in (48) and (49); the respective transitive verb *akkari* means ‘pass sth, surpass sth / sb’. Cf. independent use with indexing in (49).

- (48) *teeta kamuy TONO akkari*
 long.ago god lord than
asur-as TONO a-ne
 reputation-stand.SG lord 4.A-COP

‘I became a more famous lord than previous big lords.’

[Nakagawa et al. 2016–2021: K8108011UP.157]

- (49) *a-yup-utar-i ka i-akkari pa rupne pa pe ne*
 4.A-older.brother-PL-PSD even 4.O-than year big PL NMLZ COP

‘Brothers were older than me.’ [Nakagawa et al. 2016–2021: K7908051UP.365]

4.2.2.2.2. *Ekari* ‘aiming, facing’

This adverb can be used with temporal and spatial NPs or independently, being indexed if necessary (51). The respective transitive verb *ekari* means ‘wander about’.

- (50) *onuman-ipe ekari ek hani!*
 evening-food aiming come.SG EXCLM
 ‘Come for dinner!’ [Tamura 1996: 86]

- (51) *poro hure seta i-ekari arpa*
 big red/pink dog 4.O-facing go.SG

‘A large red dog came at me.’ [Nakagawa et al. 2016–2021: K7708242UP.140]

Ekari-(no) can also be used as an adverbial conjunction ‘just at that time’ after a clause.

- (52) *sa-n-an ekari-no a-kosmac-i nuwap ruwe ne*
 front.place-INTR.SG-4.S aiming-ADV 4.A-daughter-in-law-PSD give.birth INFR.EV COP

‘Just at that time I came down (from the mountains), the older son’s wife was giving birth.’ [Nakagawa et al. 2016–2021: K7803233UP.105]

4.2.2.2.3. *Okari* ‘around’

Okari, which originates in the transitive verb ‘go round sth’, can be used either after an NP (53) or independently with indexing (54).

- (53) *sine to mosir okari apkas nispa a-ne*
 one day land around walk rich.man 4.A-COP
 ‘I was such a wise man that I could walk throughout the country in just one day.’
 [Nakagawa et al. 2016–2021: K8010281UP.003]
- (54) *i-okari mik kor*
 4.O-around bark and
 ‘The dog barked around me.’ [Nakagawa et al. 2016–2021: K7708242UP.093]

4.2.2.2.4. *Os* ‘after’

Just like all previous adverbs in this section, *os* ‘after’ can be used after an NP or without it independently being indexed if needed, but unlike them, it has no corresponding transitive verb. It is a genuine adverb, i.e., not the one converted from a verb.

- (55) *e-pis-ne e-yup-utar-i os*
 head.POSS.PFX-beach-COP 2SG.A-older.brother-PL-PSD after
sa-p-an pe ne na
 front.place-INTR.PL-4.S NMLZ COP SGST
 ‘Let’s go down the mountain after your brothers who went to the beach.’
 [Nakagawa et al. 2016–2021: K8109193UP.044]
- (56) *i-os ahun hine i-e-rankarap ruwe ne*
 4.O-after enter.SG and 4.O-to.APPL-greet INFR.EV COP
 ‘He came in after me in and greeted me.’ [Nakagawa et al. 2016–2021: K7708242UP.119]

4.2.2.2.5. *Pak(no)* ‘till, until; as’

As mentioned in Section 4, *pak(no)* has either a spatial or temporal destination function ‘to/ till’ or a degree function ‘as’:

- (57) *kuca soy pak-no arpa-an sekor yaynu-an*
 hunting.hut outside till-ADV go.SG-4.S QUOT think-4.s
 ‘I decided to go to the hunting hut.’ [Nakagawa et al. 2016–2021: OI990205UP.062]
- (58) *an-noski pak-no u-w-e-newsar-an kor an-an*
 night-middle till-ADV RECIP-EP-about.APPL-talk-4.s and exist.SG-4.s
 ‘We were talking till midnight.’ [Nakagawa et al. 2016–2021: OI000118UP.177]
- (59) *re-n pak-no pet tom-o-tuy-pa hine*
 three-person.CLF till-ADV river middle-PSD-cut-TR.PL and
 ‘As many as three people, crossed the river.’
 [Nakagawa et al. 2016–2021: K8007292UP.013]

Pak(no) can be used independently as in (10) and can also take object indexing, as in (60).

- (60) *i-pak-no rametok e-kor wa he e-i-ki ya?*
 4.O-till-ADV bravery 2SG.A-have and FOC 2SG.S-ANTIP-do Q
 ‘Did you do it being as brave as me?’ [Nakagawa 1995: 319]

Both *pak* and *pak-no* (the latter more frequently) can also attach to a clause and function as an adverbial conjunction, as in (61).

- (61) *mata epitta e-e pak-no kim ta ka a-e-p a-hop-pa*
 winter all.over 2SG.A-eat till-ADV mountains LOC even 4.A-eat-thing 4.A-leave-TR.PL
 ‘I put up as much food in the mountains as [it is enough] for you to eat for the whole winter.’ [Nakagawa et al. 2016–2021: K7908032UP.073]

Synchronically, there is no corresponding *pak* verb, but it probably existed in the past, most likely with a meaning ‘reach sth / sb’. There is additional evidence for this in the adnominal use of *pak*, cf. Section 5, example (81).

4.2.2.2.6. *Renkayne* ‘accordingly’

This adverb is special in that it allows not for the object indexing like all other adverbs in this section, but for the subject indexing, as in (63).

- (62) *kamuy renkayne e-pa wa e-siknu-re hi ne aan kus*
 god accordingly 2SG.A-find and 2SG.A-be.alive-CAUS NMLZ COP ADM because
 ‘You must have found [that girl] with the help of gods and revived her.’ [Nakagawa et al. 2016–2021: K8106233UP.133]

- (63) *ku-renkayne k-ek*
 1SG.A-accordingly 1SG.S-come.SG
 ‘I came because I wanted to (at my own intention or convenience).’
 [Tamura 1988/2000: 96]

4.2.2.2.7. *Comitative tura(no)* ‘with’

This adverb is used to encode a comitative participant. Syntactically it can be part of a comitative group of the intransitive subject (64), transitive subject (65), object (66) or even be an adjunct of the intransitive predicate, which can possibly be interpreted as its adverbial modifier, as in (67) and (68).

- (64) *a-ak-ih i tura-no ekimne arpa/paye-an*¹⁰
 4.A-younger.brother-PSD COM-ADV to.the.mountains go.SG/PL-4.S
 ‘I went to the mountains to hunt together with my younger brother.’
 [Nakagawa et al. 2016–2021: OI980905UP.058]

- (65) *KONTO isimne nea nispa tura-no a-ri*
 then tomorrow that rich.man COM-ADV 4.A-skin
wa a-sup a wa a-e
 and 4.A-cook.PL and 4.A-eat
 ‘The following day the elder and I butchered [the bear], cooked and ate its meat.’ [Nakagawa et al. 2016–2021: K8106233UP.154]

- (66) *a-ur-ih i anakne a-ninke-he tura-no eci-e-i-hok yakne*
 4.A-fur-PSD TOP 4.A-gallbladder-PSD COM-ADV 2PL.A-by.APPL-ANTIP-buy when
 ‘If you sell my hide with my gallbladder (= bear’s gallbladder), [that money will make you wealthy].’ [Nakagawa et al. 2016–2021: K8106233UP.151]

¹⁰ Interestingly, the intransitive predicate with the comitative group can take either singular form or plural.

- (67) *tane-po rur tura-no ipe-an pe ne*
 now-DIM soup COM-ADV eat-4.S NMLZ COP
 ‘It was the first time I ate (= had food) with a soup.’
 [Nakagawa et al. 2016–2021: K7803231UP.079]
- (68) *mina tura... sekor haw-e-an-an akusu*
 laugh/laughter COM QUOT voice-PSD-exist.SG-4.S then
 ‘With a laughter I said the following...’
 [Nakagawa et al. 2016–2021: K7708241UP.064, 068]

It is noteworthy that examples of the comitative group of the object, as in (66), are very rare. They are most likely to trigger a different instrumental interpretation, as in (69) and (70), which was overlooked in previous studies. Recall that comitative and instrumental polysemy is common cross-linguistically [Stolz et al. 2013].

- (69) *kam pirka hi cep pirka hi tura-no*
 meat be.good NMLZ fish be.good NMLZ COM-ADV
ho-pun-pa-re pa ruwe ne
 bottom.PFX-lift-TR.PL-CAUS PL INFR.EV COP
 ‘They sent [their souls] by using (= sacrificing) delicious meat and delicious fish.’
 [Nakagawa et al. 2016–2021: K7803231UP.154]
- (70) *kanna kamuy sorekus sake tura inaw tura a-nomi kor oka-an*
 upper god exactly liquor COM inaw COM 4.A-worship and exist.PL-4.S
 ‘We worshipped the god of Lightning with sake and prayer sticks.’
 [Nakagawa et al. 2016–2021: K8010291UP.520]

And finally, let us consider an example of *tura* used independently, which is hardly surprising, because this adverb originates in and is still closely related to the transitive verb *tura* ‘follow sb’.

- (71) *a-kor Pananpe i-tura wen kur ne wa*
 4.A-have Pananpe 4.O-COM poor person COP and
 ‘My dear Pananpe was as poor as I’, lit. ‘... was poor together with me.’ (How did you become so wealthy?) [Nakagawa et al. 2016–2021: K8106231UP.006]

5. Adnominal phrases

An adnominal phrase like **Satporo wa* (ABL) *kur* ‘a man from Sapporo’ is not possible in Ainu. Various semantic relations between two nouns (put aside possessive relations as in (5)) can be encoded with a few existential transitive verbs (72), which appear between two nouns and form a relative-clause-like structure, as in examples (73)–(81) below.

- (72) Locative (‘at / from’) relations:
ta ‘be / exist at’ (cf. locative *ta* in Section 4.1.1),
un ‘live at’ (cf. allative / location of action *un* in Section 4.1.2);
- Membership relations:
ne ‘be(come) sb’ (cf. mutative *ne* in Section 4.2.1.1);
- Proprietary (‘with’) relations:
us ‘attach to’,
o ‘be located inside of’,
ot ‘be in abundance at’,
kor ‘have / rule’;

6. Applicatives as head-marking case markers

As suggested in Bugaeva [2017], Ainu case markers have a marginal status. In fact, there is a viable alternative to them, i.e., three highly polyfunctional applicative prefixes, namely instrumental *e-*, as in (12), (28), (35b), (56), (58), and (66), dative *ko*, as in (21), and locative *o-*, as in (21) and (83b); for details see Bugaeva [2010]. It has been hypothesized that applicatives in Ainu have not always been voice markers serving to promote a semantically peripheral participant to direct object, cf. (83a) and (83b).¹¹ Originally, they may have been used just for encoding participants on the verb, i.e., as a polysynthetic equivalent of adpositional marking (83b).

- (83) a. *a-kasuy wa a-se wa sem or un a-rura*
 4.A-help and 4.A-shoulder and porch place ALL 4.A-carry

‘I helped by carrying it on my back **to** the shed.’

[Nakagawa et al. 2016–2021: K8010291UP.179]

- b. *nea mat-ih i uni opitta a-o-rura ruwe ne*
 that wife-PSD home.PSD all 4.A-to.APPL-carry INFR.EV COP

‘I carried everything **to** his wife’s house.’

[Nakagawa et al. 2016–2021: K8010281UP.213]

As mentioned above, for a few exceptions such as the ablative case particle *wa* ‘from’ (Section 4.1.3) and traversal *peka* ‘through, around, over’ (Section 4.1.4), almost all case particles and adverbial postpositions not only originate in transitive verbs but are still at a rather early stage of grammaticalization from them; recall that the latter allow for independent use and some of them even for indexing. On the other hand, it is not so easy to produce a good etymology for applicative markers, cf. an attempt of that in Bugaeva [2010]. Let me speculate that case markers in Ainu are young and innovative, while applicative markers are old and genuine. Of course, more solid evidence is necessary to prove this hypothesis. Overall, I hypothesize that borderline case-marking may generally correlate with the head-marking language type.

7. Concluding remarks

This paper aims to describe how case relations are expressed in Ainu. It is shown that arguments are unmarked for case and that there is no genitive marker either (Section 3). However, a few peripheral semantic relations such as locative, allative, ablative, and traversal can be expressed by postpositional case particles (Section 4.1). In addition, there is an open class of more than ten so-called adverbial postpositions (Section 4.2), which, unlike case markers in other languages, can be used independently, i.e., without respective NPs, and moreover some can even take indexing. The semantic range of those adverbial postpositions varies significantly: from such canonical case relations as dative, instrumental, and comitative to really peripheral and language-specific ones like ‘along a river / mountain upstream / upwards’ (*turasi*) or ‘along a river / mountain downstream / downwards’ (*pes*). Naturally, there is a huge problem of setting a limit between postpositional case adverbs and all other adverbs. This system of marking case relations is claimed to represent a typical case of the exclusively borderline case-marking as defined in Iggesen [Iggesen 2013].

¹¹ Here a two-place transitive verb *rura* ‘carry sth / sb’ turns into a three-place transitive verb *o-rura* ‘carry sth / sb **to**’, but there are also many cases when an intransitive verb turns into a two-place transitive as in (12) and (28).

It is argued that the unusual behavior of adverbial “case” postpositions can be explained by their relatedness to respective transitive verbs out of which they have not yet fully grammaticalized. It may not be accidental that adnominal phrases too employ transitive verbs, not case markers. Furthermore, Ainu possesses a good deal of applicatives on the verb, which are so highly polyfunctional and ubiquitous that they can be regarded as a functional alternative to case-markers described in this paper. Based on this and other evidence, it is hypothesized that case-markers in Ainu are likely to be young and innovative, while applicatives are old and genuine, and that the borderline case-marking may generally correlate with the head-marking language type, which is most broadly attested in the Pacific Rim.

ABBREVIATIONS

1, 2, 3, 4 — 1 st , 2 nd , 3 rd , “4 th ” person	DESI — desiderative	PFX — prefix
— — relative clause gap	DIM — diminutive	PL — plural
A — transitive subject or possessor	HON — honorific	POSS — possessive
ABL — ablative	EP — epenthetic consonant	PROH — prohibitive
ADM — admirative	EXCL — exclusive	PSD — possessed
ADV — adverbial	EXCLM — exclamation	RECIP — reciprocal
ALL — allative	FOC — focus	REFL — reflexive
ANTIP — antipassive	IMP.POL — imperative polite	Q — question particle
APPL — applicative	INCL — inclusive	QUOT — quotative
CAUS — causative	INDF — indefinite	S — intransitive subject
CLF — classifier	INFR.EV — inferential evidential	SG — singular
COM — comitative	INS — instrumental	SFP — sentence final particle
COP — copula	INTR — intransitive	TOP — topic
DAT — dative	LOC — locative	TR — transitive
	NMLZ — nominalizer	TRV — traversal
	O — object	

REFERENCES

- Asai 1969/2001 — Asai T. Ainugo no bunpō: Ainugo Ishikari hōgen no gairyaku [Ainu grammar: Outline of the grammar of the Ishikari dialect of Ainu]. *Ainu minzokushi* [Ainu ethnic records]. Ainu Bunka Honzon Taisaku Kyōgikai (ed.). Tokyo: Daiichi Hōki Shuppan, 1969, 771–800. Reprinted in *Ainugo kō* 4. Tokyo: Yumani Shobō, 2001, 21–50.
- Bugaeva 2010 — Bugaeva A. Ainu applicatives in typological perspective. *Studies in Language*, 2010, 34(4): 749–801.
- Bugaeva 2011 — Bugaeva A. A diachronic study of the impersonal passive in Ainu. *Impersonal constructions cross-linguistically*. Malchukov A., Siewierska A. (eds.). Amsterdam: John Benjamins, 2011, 517–546.
- Bugaeva 2012 — Bugaeva A. Southern Hokkaido Ainu. *The languages of Japan and Korea*. Tranter N. (ed.). London: Routledge, 2012, 461–509.
- Bugaeva 2017 — Bugaeva A. Polysynthesis in Ainu. *Handbook of Polysynthesis*. Fortescue M., Mithun M., Evans N. (eds.). Oxford: Oxford Univ. Press, 2017, 882–905.
- Bugaeva et al. 2022 — Bugaeva A., Nichols J., Bickel B. Appositive possession in Ainu and around the Pacific. *Linguistic Typology*, 2022, 26(1): 43–88. <https://doi.org/10.1515/lingty-2021-2079>.
- Bugaeva, 2022 — Bugaeva A. Ainu: A head-marking language of the Pacific Rim. *Handbook of the Ainu language*. Bugaeva A. (ed.). Berlin: Mouton de Gruyter, 2022, 23–55.
- Chiri 1942/1973 — Chiri M. Ainu gohō kenkyū [A study of Ainu grammar], 1942. Reprinted in: *Chiri Mashihō chosakushū* [Selected writings of Chiri Mashihō]. Vol. 3. Tokyo: Heibonsha, 1973, 455–586.
- Chiri 1953/1976 — Chiri M. *Bunrui Ainugo jiten: Shokubutu-hen, dōbutsu-hen* [Categorical Ainu dictionary: Plants and animals], 1953. Reprinted in: *Chiri Mashihō chosakushū* [Selected writings of Chiri Mashihō]. Extra vol. 1. Tokyo: Heibonsha, 1976.

- Comrie, to appear — Comrie B. Verb indexing (“agreement”) in languages of northern Asia. *The languages and linguistics of Northern Asia: a comprehensive guide*. Vajda E. (ed.). Berlin: De Gruyter Mouton, to appear.
- Fukuda (Tamura) 1960/2001 — Fukuda (Tamura) S. Ainugo Saru hōgen no jodōshi: Ainugo no joshi ni tsuiteno hōkoku sono 1 [Auxiliary verbs in the Saru dialect of Ainu: Report on particles in Ainu 1]. *Ainugo kō*, No. 4. Tokyo: Yumani Shobō, 2001, 57–68. (First published 1960.)
- Iggesen 2013 — Iggesen O. A. Chapter 49. Number of cases. *The world atlas of language structures online*. Dryer M. S., Haspelmath M. (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. <http://wals.info/chapter/49> (accessed on 27 March 2022).
- Kirikae 1989 — Kirikae H. “*Ainu shinyōshū jiten*” *tekisuto, bunpō kaisetsu-tsuki* [Dictionary of collections of Ainu divine epics, with texts and grammatical notes]. Sapporo: Hokkaido daigaku bungakubu gengogaku kenkyūshitsu, 1989.
- Kirikae 2003 — Kirikae H. *Ainu shin'yōshū jiten* [Dictionary of collections of Ainu divine epics]. Tokyo: Daigaku Shorin, 2003.
- Kindaichi 1931/1993 — Kindaichi K. Ainu yūkara gohō tekiyō [An outline grammar of the Ainu epic poetry]. *Ainu jojishi yūkara no kenkyū* 2, 1931, 3–233. Tokyo: Tōyō Bunko. Reprinted in *Ainugogaku kōgi* 2 [Lectures on Ainu studies 2] (Kindaichi Kyōsuke zenshū: Ainugo I), vol. 5. Tokyo: Sanseidō, 1993, 145–366.
- Kindaichi, Chiri 1936/1974 — Kindaichi K., Chiri M. *Ainu gohō gaisetsu* [An outline of Ainu grammar]. Tokyo: Iwanami Shoten, 1936. Reprinted in *Chiri Mashiho chosakushū* 4 [Selected writings of Chiri Mashiho 4]. Tokyo: Heibonsha, 1974, 3–197.
- Kubodera 1977 — Kubodera I. *Ainu jojishi: Shin'yō seiden no kenkyū* [Studies of the Ainu epic poem *Shin'yō seiden*]. Tokyo: Iwanami Shoten, 1977.
- Nakagawa 1995 — Nakagawa H. *Ainugo Chitose hōgen jiten* [An Ainu dictionary: Chitose dialect]. Tokyo: Sōfūkan, 1995.
- Nakagawa, 2022 — Nakagawa H. Parts of speech — with a focus on the classification of nouns. *Handbook of the Ainu language*. Bugaeva A. (ed.). Berlin: Mouton de Gruyter, 2022, 473–513.
- Nakagawa et al. 2016–2021 — Nakagawa H., Bugaeva A., Kobayashi M., Yoshikawa Y. (eds.). *A glossed audio corpus of Ainu folklore*. Tokyo: National Institute for Japanese Language and Linguistics, 2016–2021. <https://ainu.ninjal.ac.jp/folklore/en/> (accessed on 27 March 2022).
- Nichols 2017 — Nichols J. Polysynthesis and head marking. *Handbook of polysynthesis*. Fortescue M., Mithun M., Evans N. (eds.). Oxford: Oxford Univ. Press, 2017, 59–69.
- Okuda 1998 — Okuda O. Ainugo Shizunai hōgen no kōchi-fukushi [Postpositional adverbs in the Shizunai dialect of Ainu]. *Bulletin of the Hokkaido Ainu Culture Research Center*, 1998, 4: 127–149.
- Okuda 1999 — Okuda O. Ainugo Shizunai hōgen no kaku-joshi [Case particles in the Shizunai dialect of Ainu]. *Bulletin of the Hokkaido Ainu Culture Research Center* 1999, 5: 153–169.
- Onseishiryō 1 — Tamura S. *Ainugo onseishiryō* 1 [Ainu audio materials 1]. Tokyo: Waseda Daigaku Gogaku Kyōiku Kenkyūjo, 1984.
- Onseishiryō 3 — Tamura S. *Ainugo onseishiryō* 3 [Ainu audio materials 3]. Tokyo: Waseda Daigaku Gogaku Kyōiku Kenkyūjo, 1986.
- Onseishiryō 5 — Tamura S. *Ainugo onseishiryō* 5 [Ainu audio materials 5]. Tokyo: Waseda Daigaku Gogaku Kyōiku Kenkyūjo, 1988.
- Refsing 1986 — Refsing K. *The Ainu language*. Aarhus: Aarhus Univ. Press, 1986.
- Satō 2008 — Satō T. *Ainugo bunpō no kiso* [Basic Ainu grammar]. Tokyo: Daigaku Shorin, 2008.
- Stolz et al. 2013 — Stolz T., Stroch C., Urdze A. Chapter 52. Comitatives and instrumentals. *The world atlas of language structures online*. Dryer M. S., Haspelmath M. (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. <https://wals.info/chapter/52> (accessed on 27 March 2022).
- Tamura 1979 — Tamura S. *Aynu Itak: Aynugo nyūmon* [The Ainu language: An introduction into Ainu]. Tokyo: Waseda Daigaku Gogaku Kyōiku Kenkyūjo, 1979.
- Tamura 1988/2000 — Tamura S. *The Ainu language*. Tokyo: Sanseidō, 2000. (Transl. from Japanese, first published 1988.)
- Tamura 1993 — Tamura S. *Ainugo onsei shiryō (1–6) Goi* [Ainu audio materials 1–6. Lexicon]. Vol. 3. Tokyo: Waseda Daigaku Gogaku Kyōiku Kenkyūjo, 1993.
- Tamura 1996 — Tamura S. *Ainugo Saru hōgen jiten* [A dictionary of the Saru dialect of Ainu]. Tokyo: Sōfūkan, 1996.

[Рец. на: / Review of:] **A. Guillaume, H. Koch (eds.). *Associated motion***. (Empirical approaches to language typology, 64.) Berlin; Boston: De Gruyter Mouton, 2021. vii + 923 p. ISBN 978-3-11-069200-6.

Ирина Андреевна Хомченкова

Институт русского языка им. В. В. Виноградова РАН, Москва, Россия; Научно-образовательная школа Московского государственного университета имени М. В. Ломоносова «Сохранение мирового культурно-исторического наследия», Москва, Россия; irina.khomchenkova@yandex.ru

Irina A. Khomchenkova

Vinogradov Russian Language Institute, Russian Academy of Sciences, Moscow, Russia; Scientific and Educational School of Lomonosov Moscow State University “Preservation of the World Cultural and Historical Heritage”, Moscow, Russia; irina.khomchenkova@yandex.ru

DOI: 10.31857/0373-658X.2022.5.151-160

В декабре 2017 г. в Канберре состоялся воркшоп “Associated motion”¹ («Сопутствующее движение») в рамках международной конференции ALT (Association for Linguistic Typology), посвященный этой малоизученной категории². Рецензируемый сборник был издан через три с половиной года под редакцией организаторов воркшопа, Антуана Гийома и Гарольда Коха, дополненный статьями от авторов, которые не принимали участия в конференции.

Сопутствующее движение определяется редакторами сборника как «глагольная грамматическая категория, отличная от категорий времени, аспекта, модальности и глагольной ориентации, которая добавляет к ситуации, выражаемой глаголом, компонент перемещения»³ (с. 3), см. иллюстрацию в (1). Как отмечают редакторы, перемещение, выражаемое с помощью этих показателей, выходит за рамки типологии глаголов движения Л. Талми (см. в первую очередь [Talmy 1985]), поскольку он учитывал лишь случаи, в которых факт перемещения выражается с помощью глагольных корней.

- (1) Кавиненья (< таканские < пано-таканские)
- | | |
|-----------------|--|
| <i>ba-</i> | ‘видеть O’ |
| <i>ba-ti-</i> | ‘пойти и увидеть O’ |
| <i>ba-na-</i> | ‘прийти и увидеть O’ |
| <i>ba-aje-</i> | ‘видеть O, идя’ |
| <i>ba-be-</i> | ‘видеть O, приходя’ |
| <i>ba-kena-</i> | ‘увидеть O и пойти’ |
| <i>ba-dadi-</i> | ‘видеть O, пока O отдаляется’ |
| <i>ba-tsa-</i> | ‘видеть O, пока O приближается’ (с. 3) |

¹ Программа воркшопа доступна по ссылке: http://www.ddl.cnrs.fr/fulltext/Guillaume/Workshop_programs/Associated%20motion_ALT2017_Epire.pdf.

² Похожая тематика также обсуждалась в Канберре в июне 2018 г. на воркшопе “Directionality or associated motion? Evidence from diverse languages” («Глагольная ориентация или сопутствующее движение? Факты из разных языков»), который был организован одним из авторов рецензируемого сборника Л. Рид (гл. 9). Из других авторов сборника в этом воркшопе также участвовали Г. Кох (гл. 1, гл. 7) и Д. Шоккин (гл. 10). Программа воркшопа доступна по ссылке: <https://www.dynamicsoflanguage.edu.au/news-and-media/past-events/past-events-in-2018/article/?id=workshop-directionality-or-associated-motion-evidence-from-diverse-languages-22-june>.

³ “Associated Motion is (...) a verbal grammatical category, separate from tense, aspect, mood and direction, whose function is to associate, in different ways, different kinds of translational motion (spatial displacement / change of location) to a (generally non-motion) verb event”.

Показатели сопутствующего движения могут различаться по ряду параметров:

- синтаксический статус перемещающегося участника ситуации (субъект vs. прямой объект, ср. *ba-be-* ‘видеть O, приходя’ и *ba-tsa-* ‘видеть O, пока O приближается’);
- таксисные отношения между сопутствующим движением и ситуацией, выражаемой исходным глаголом (предшествование, одновременность, следование, ср. *ba-ti-* ‘пойти и увидеть O’, *ba-aje-* ‘видеть O, идя’, *ba-kena-* ‘увидеть O и пойти’);
- направление движения (‘к’/‘от’, ‘вверх’/‘вниз’, ‘туда’/‘обратно’ и др.);
- глагольный аспект (перфектив vs. имперфектив).

На основе первых двух параметров А. Гийом предложил две имплицитивные иерархии (‘если в языке есть показатель следования, то в нем должны быть показатели одновременности и/или предшествования, если в языке есть показатель одновременности, то в нем должен быть показатель предшествования’, а также ‘если в языке есть показатель, обозначающий перемещение референта прямого объекта, то в нем обычно есть и показатели, обозначающие перемещение референта субъекта’) [Guillaume 2016].

Книга включает пять частей и три указателя (предметный, языковой и именной). **Первая часть** (“Perspectives and general issues”) содержит шесть теоретических и типологических глав о категории сопутствующего движения и прочих категориях, так или иначе связанных с ней. В этих главах рассматриваются история изучения сопутствующего движения, терминология, основные типологические наблюдения (А. Гийом и Г. Кох), морфологическое кодирование сопутствующего движения и глагольной ориентации в языках мира (Д. Росс), выражение сопутствующего движения и глагольной ориентации с помощью сериальных глагольных конструкций (Дж. Лавстранд и Д. Росс), связь категорий сопутствующего движения и глагольной ориентации (М. Драйер; А. Белькади), экспериментальные методики исследования показателей сопутствующего движения (М. Вюйерме).

В **первой (вводной) главе** (“Introduction: associated motion as a grammatical category in linguistic typology”) А. Гийом и Г. Кох обсуждают общие вопросы, связанные с категорией сопутствующего движения (о которых было частично сказано выше), а также приводят краткий обзор вошедших в книгу глав.

Феномен сопутствующего движения был обнаружен в 1980-х гг. на материале языков центральной Австралии (самая ранняя работа принадлежит редактору рецензируемого сборника Г. Коху [Koch 1984], который и предложил использовать этот термин). В двухтысячные годы термин был применен для описания показателей с похожей семантикой в языках Южной Америки (одна из ранних работ принадлежит другому редактору сборника — А. Гийому [Guillaume 2000]). Первыми были исследованы системы с большим количеством противопоставлений, но затем были описаны и более простые системы с одним или двумя показателями категории сопутствующего движения.

Следующим этапом в изучении этой категории стало включение в рассмотрение систем, в которых сопутствующее движение выражается с помощью показателей глагольной ориентации (или «ориентивов» в терминологии [Плунгян 2002]) дейктического типа (deictic directionals). Эти показатели выражают глагольную ориентацию (в общем случае присоединяясь к глаголам движения) и могут добавлять компонент перемещения, присоединяясь к прочим глаголам. А. Белькади, описавшая такие системы в языках Африки, назвала их «дейктическим сопутствующим движением» (“deictic associated motion”, D-AM) [Belkadi 2015: 50]. В примере (2) на с. 6 рецензируемого издания приводится иллюстрация этого типа системы в кабийском языке: в (2a) показатель =*d* с вентивным значением (‘к дейктическому центру’) присоединяется к глаголу движения и функционирует как показатель глагольной ориентации, обозначая направление прыжка, а в (2b) он присоединяется к глаголу, не обозначающему движение, и функционирует как показатель сопутствующего движения, добавляя к событию ‘читать’ компонент перемещения.

(2) КАБИЛЬСКИЙ (< БЕРБЕРСКИЕ)

a. *t-jjlb=d* *γr tabla*.
 3SG.F-прыгать.PRF=VEN к стол

‘Она прыгнула на стол (в направлении говорящего)’.

b. *t-γra=d* *taktaf*
 3SG.F-читать.PRF=VEN книга

‘Она прочитала книгу (где-то еще) и вернулась (к местоположению говорящего)’.

(с. 6)

Системы D-AM отличаются от специализированных систем сопутствующего движения. В частности, значения этих показателей в «контекстах сопутствующего движения» не фиксированы и могут зависеть от семантики глагола, времени, аспекта и модальности, прагматических характеристик. Так, в сомалийском языке единица *soo* с вентивным значением (‘к дейктическому центру’) может выражать предшествующее движение (3i), одновременное движение (3ii), последующее движение (3iii), а также сочетание предшествующего и сопутствующего движения (3iv).

(3) СОМАЛИЙСКИЙ (< КУШИТСКИЕ)

soo seexo
 VEN спать.IMP.2SG

i. ‘Приходи и спи здесь’. {непосредственная коммуникация}

ii. ‘Отдохни (пока едешь сюда)’. {телефонный разговор}

iii. ‘Отдохни (а потом мы встретимся)’. {телефонный разговор}

iv. ‘Иди отдохни (а потом возвращайся сюда)’. / ‘Иди отдохни, (а потом мы встретимся)’. {непосредственная коммуникация} (с. 8)

На материале 20 языков Африки А. Белькади делает обобщения относительно зависимости семантического класса глагола и значения показателей в системах D-AM: глаголы на левой части шкалы (4) обычно вызывают «ориентационные» прочтения, а глаголы на правой части — прочтения «сопутствующего движения» [Belkadi 2016: 64]. В скобках приводятся некоторые глаголы, рассмотренные в [Ibid.]. Эта классификация также обсуждается в гл. 20.

(4) Движение по траектории (‘входить’, ‘спускаться’, ‘приезжать’) > способ движения (‘бегать’, ‘прыгать’, ‘танцевать’) > каузативное перемещение (‘брать’, ‘давать’) > глаголы восприятия (‘смотреть’, ‘находить’) > явления природы и телесные выделения (‘мочиться’, ‘выкорчевывать’) > деятельность, не подразумевающая движения (‘пить’, ‘спать’) > состояние (‘оставаться’).

Вторая глава (“A cross-linguistic survey of Associated Motion and Directionals”) авторства Д. Росса представляет типологический обзор категорий сопутствующего движения и глагольной ориентации на материале сбалансированной выборки из 325 языков. Рассматривая лишь морфологическое маркирование, Д. Росс демонстрирует, что показатели, выражающие сопутствующее движение, зафиксированы в 83 языках (26%), а показатели глагольной ориентации — в 114 языках (35%). В 57 языках (18%) есть показатели обоих типов, причем в 47 из них (14%) для маркирования этих двух категорий могут использоваться одинаковые показатели. В целом иерархии, сформулированные в [Guillaume 2016] на материале языков Южной Америки, подтверждаются (хотя Д. Росс отмечает, что количество показателей одновременности ненамного превышает количество показателей следования). К недостаткам изложения можно отнести то, что автор, приводя список языков и данные по ним в приложении, не указывает источники, предпочитая ссылаться на свою диссертацию [Ross 2021], которая на момент написания рецензии не была доступна. Было бы удобно, если бы автор оформил языковые данные в виде вспомогательных материалов, размещенных на сайте издания (как это сделано для гл. 4 и гл. 13 М. Драйера).

Третья глава (“Serial verb constructions and motion semantics”) Дж. Лавстранда и Д. Росса посвящена обсуждению выражения сопутствующего движения и глагольной ориентации с помощью сериальных глагольных конструкций (конструкций, в которых как минимум два глагола встречаются в одной клаузе без каких-либо показателей сочинения или подчинения). Семантика перемещения или ориентации была зафиксирована в 101 из 125 рассмотренных языков с сериальными глагольными конструкциями. Самыми частотными оказались «ориентационные» сериальные глагольные конструкции, в которых с помощью двух глаголов выражается одно событие перемещения (70 языков), и сериальные глагольные конструкции «предшествования» и «движения с целью» (“motion-sum-purpose”; 68 языков). В конструкции предшествования указывается на совершение обоих событий, в то время как в конструкции «движения с целью» выражается лишь намерение совершить второе событие, ср. русск. *Я пошел на рынок и купил арбуз* (“но арбузов не было) и *Я пошел на рынок, чтобы купить арбуз, но арбузов не было*. Поскольку по переводам примеров в грамматиках не всегда возможно различить эти два типа конструкций, Д. Росс и Дж. Лавстранд рассматривали их вместе. Сериальные глагольные конструкции «одновременности» и «следования» были зафиксированы лишь в пяти и шести языках выборки соответственно. К недостаткам главы также можно отнести отсутствие ссылок на источники в приложении.

Авторы справедливо отмечают, что исключение сериальных глагольных конструкций (и других конструкций с несколькими глаголами) при изучении категории сопутствующего движения нецелесообразно и целостное понимание этой категории не может быть достигнуто, если принимать во внимание лишь морфологические способы ее кодирования. Так, утверждение А. Гийома и Г. Коха (гл. 1) о том, что сопутствующее движение не зафиксировано в языках Европы, верно лишь отчасти — эта категория действительно не находит морфологического выражения в этих языках, однако кодируется с помощью глагольных конструкций (ср., например, англ. *go and do*).

В **четвертой главе** (“Associated motion and directionals: where they overlap”) М. Драйера обсуждается связь между показателями сопутствующего движения и глагольной ориентации на материале «удобной» выборки (convenience sample) из 56 языков. Автор приводит многочисленные примеры, в которых одна и та же морфема может либо вводить компонент перемещения, либо обозначать ориентацию. Он утверждает, что категорию глагольной ориентации нельзя считать частным случаем категории сопутствующего движения, а случаи, когда показатели сопутствующего движения используются как ориентивы, объясняются тем, что это «естественное расширение» значений этих показателей, поскольку кодирование дополнительного глагольного события с глаголами движения оказывается избыточным.

В **пятой главе** (“Deictic directionality as associated motion: motion, complex events and event integration in African languages”), также обсуждая связь категорий глагольной ориентации и сопутствующего движения, А. Белькади приходит к противоположным выводам: «дейктическая глагольная ориентация (как минимум в некоторых языках) является видом сопутствующего движения» (с. 166). Она утверждает, что ориентивы, как и показатели сопутствующего движения, вводят ситуацию, дополнительную к ситуации, заданной исходным глаголом, однако в случае ориентивов эта дополнительная ситуация оказывается полностью интегрированной в ситуацию движения, выраженного исходным глаголом. Одним из аргументов является трактовка примеров (5) из [Rappaport Hovav 2008]: в обоих случаях группы с предлогом *to* вводят ситуацию перемещения (бочки и мяча соответственно), однако в (5a) она происходит одновременно с ситуацией ‘катить’, в отличие от примера (5b), в котором ситуация бросания имела место, а перемещение мяча к Мэри — нет.

(5) Английский

а. *John rolled the barrel to the store.*

‘Джон прикатил бочку к магазину’.

- b. *I threw the ball to Mary (but aimed badly and she didn't catch it).*
 ‘Я бросил мяч Мэри (но плохо целился, и она его не поймала)’. (с. 186)

Шестая глава (“A visual stimulus for eliciting associated motion”) М. Вюйерме описывает набор иллюстраций («История об охоте»), которые были специально разработаны для записи рассказов, содержащих показатели сопутствующего движения. В результате записи текстов от 14 носителей языка эсе-эхха (< ТАКАНСКИЕ < ПАНО-ТАКАНСКИЕ) были получены примеры на большую часть показателей сопутствующего движения в этом языке. Среди интересных наблюдений автора можно отметить то, что количество показателей сопутствующего движения в текстах варьировало по носителям неочевидным образом. Можно было бы ожидать, что молодые носители, использующие язык менее активно, и носители из региона, где язык больше не передается детям, будут употреблять меньше показателей сопутствующего движения, чем взрослые носители и носители из региона, где язык передается детям, однако такого эффекта не наблюдалось. Автор отмечает, что на результаты могли повлиять, например, степень комфорта в процессе выполнения задания (чем более человек расслаблен, тем больше показателей сопутствующего движения встречается в его речи).

Остальные четыре части книги посвящены различным ареалам: Австралия и Океания (ч. II), Америка (ч. III), Африка (ч. IV) и Азия (ч. V).

Вторая часть, посвященная языкам Австралии и Океании, содержит четыре главы. В **седьмой** (почти стостраничной) главе (“Associated motion in the Pama-Nyungan languages of Australia”) Г. Кох рассматривает сопутствующее движение в языках пама-ньюнга. Автор приводит анализ сложной системы сопутствующего движения в арандийском языке кайте-те (дополненный по сравнению с работой [Koch 1984]), сравнивает системы сопутствующего движения всех арандийских языков, а затем обсуждает и данные ряда других языков пама-ньюнга. Автор демонстрирует, что показатели этой категории присутствуют в 41 языке (количество сократится до 35, если сгруппировать некоторые диалекты) из 100 рассмотренных. В статье приводятся обобщения относительно систем сопутствующего движения в этих языках: они варьируют по уровню сложности и параметрам. Также обсуждается географическая дистрибуция (сложные системы группируются в центре Австралии; простые системы распространены в остальных частях континента, где говорят на языках пама-ньюнга) и диахронические источники показателей сопутствующего движения (самый частотный — конструкция, состоящая из нефинитного глагола и глагола движения).

Оставшиеся три главы в этой части посвящены отдельным языкам. Д. Осгарби (гл. 8, “Mudburra associated motion in an areal perspective”) обсуждает данные языка мутпурра (< ПАМА-НЬЮНГА) с двумя показателями сопутствующего движения (‘от дейктического центра’; ‘к дейктическому центру’), которые могут добавлять предшествующее или одновременное событие. Присоединяясь к глаголам движения, эти показатели могут обозначать как ориентацию (например, ‘упасть по направлению к говорящему’), так и сопутствующее движение (например, ‘прийти и упасть’), а их сочетание с прочими глаголами может обозначать сопутствующее движение, но не ориентацию. Эти показатели не сочетаются со всеми значениями времени и модальности и присоединяются после показателей ТАМ, а не до, что нетипично для других пама-ньюнгских языков. Предполагается, что эти свойства могут быть обусловлены контактами с соседними миндийскими языками. Интересно, что в одном из миндийских языков (вампая) также существуют два показателя сопутствующего движения. При этом показатели сопутствующего движения в языке мутпурра произошли от показателей глагольной ориентации, а показатели сопутствующего движения в языке вампая — от глаголов.

Автор также приводит таблицу (с. 341 рецензируемого издания), обобщающую различия между показателями глагольной ориентации и сопутствующего движения, см. таблицу. Он считает, что показатели сопутствующего движения могут обозначать как ориентацию, так и сопутствующее движение, присоединяясь к глаголам движения. Показатели

«ориентационного» сопутствующего движения («дейктического») сопутствующего движения в терминах А. Белькади) в свою очередь не могут вводить дополнительное событие перемещения с глаголами движения, но могут обозначать сопутствующее движение с глаголами других лексических классов. Показатели глагольной ориентации не выражают сопутствующее движение ни с глаголами движения, ни с прочими глаголами.

Таблица

Значение показателей глагольной ориентации и сопутствующего движения в зависимости от типа глагола

	Глагольная ориентация	«Ориентационное» сопутствующее движение	Сопутствующее движение
Глаголы движения	Ориентация	Ориентация	Ориентация (/Сопутствующее движение)
Прочие глаголы	Ориентация	Ориентация / Сопутствующее движение	Сопутствующее движение

Л. Рид и К. Линдси (гл. 9, “Now the story’s turning around: Associated motion and directionality in Ende, a language of Papua New Guinea”) описывают показатель глагольной ориентации (‘к дейктическому центру’) в языке энде (Папуа — Новая Гвинея), который в сочетании с глаголами, не обозначающими движение, может кодировать сопутствующее движение (выражая предшествование, одновременность или следование). Авторы предлагают новый термин «инферентивное сопутствующее движение» (inferential associated motion) в противовес «специализированному сопутствующему движению». Термин «дейктическое сопутствующее движение» (deictic associated motion), введенный А. Белькади, они считают не вполне удачным, поскольку специализированные показатели сопутствующего движения также имеют дейктический компонент в своем значении.

Д. Шоккин (гл. 10, “Preverbal directionals as markers of associated motion in Paluai (Austronesian; Oceanic)”) рассматривает 10 показателей глагольной ориентации в австронезийском языке палуай (балуан-пам), которые употребляются как в предглагольной, так и в постглагольной позиции. Автор демонстрирует, что эти показатели, присоединяясь к глаголам, не обозначающим движение, могут функционировать как показатели сопутствующего (предшествующего) движения, но лишь в предглагольной позиции. Он заключает, что различные свойства рассмотренных им показателей, которые зависят от позиции, могут служить примером полиграмматикализации [Craig 1991] — процесса, при котором один элемент может эволюционировать в несколько разных грамматических маркеров.

Третья часть, посвященная языкам Америки, также содержит четыре главы. Две главы посвящены конкретным языкам. А. Таллман (гл. 11, “Associated motion in Chácobo (Pano) in typological perspective”) описывает данные паноанского языка чакобо (< ПАНОАНСКИЕ < ПАНО-ТАКАНСКИЕ), в котором он выделил семь показателей сопутствующего движения (в отличие от предыдущих исследований, в которых были описаны лишь три показателя). Эти показатели могут также выражать ориентационные и аспектуальные значения. Автор рассматривает, приносят ли эти маркеры компонент перемещения при их употреблении с глаголами, не выражающими движение, и приходит к выводу, что эта диагностика не дает однозначного ответа — корпусные данные показывают, что частотность выражения сопутствующего движения варьирует в зависимости от типа движения: одновременное (в 70 % случаев показатель кодирует сопутствующее движение) > последующее (39 %) > предшествующее (27 %).

А. Видаль и Д. Пэйн (гл. 12, “Pilagá directionals and the typology of associated motion”) исследуют язык пилага (< гуйкуру). В этом языке богатая система глагольной ориентации, насчитывающая 14 показателей. Авторы описывают три показателя, которые также могут функционировать как показатели сопутствующего движения. Они отмечают, что

в конструкциях с сопутствующим движением перемещающимся объектом оказывается либо прямой объект переходного глагола, либо субъект непереходного глагола, но не субъект переходного глагола. Авторы предполагают, что такая абсолютивная стратегия может объясняться объединением выражения сопутствующего движения с аппликативной функцией.

Тринадцатая глава (“Associated motion in North America (including Mexico and Central America)”) М. Драйера имеет континентальный охват: автор исследует сопутствующее движение в Северной и Центральной Америке. Эта категория отсутствует в пяти семьях, в то время как в 29 семьях и в некоторых языках-изолятах она засвидетельствована. Большинство языков имеет простые системы (с одним или двумя показателями). Наибольшее количество показателей зафиксировано в юто-ацтекском языке тимбиша (панаминт) с семью единицами и в алгонкинском языке меномини с шестью единицами. К недостаткам главы можно отнести отсутствие раздела с общими выводами.

Четырнадцатая глава (“Associated motion in the Otomi family”) Н. Эрнандес-Грина и Э. Паланкара описывает сопутствующее движение в трех языках отоми (< ОТО-МАНГСКИЕ). Рассматриваемые языки различаются по степени сложности системы сопутствующего движения: от одного до четырех семантических противопоставлений. Во всех трех языках зафиксирован вентив (‘к дейктическому центру’), который может выражать предшествующее, одновременное и последующее движение. Особенностью категории сопутствующего движения в этих языках является то, что ее значения выражаются кумулятивно с видо-временными и модальными показателями и лицом субъекта и с помощью чередований в корне.

Четвертая часть, посвященная языкам Африки, содержит пять глав. Четыре из них посвящены отдельным языковым семьям. **В пятнадцатой главе** (“Associated motion in Bantu languages”) Р. Геруа, Х. Гибсон и Б. Персон описывают сопутствующее движение в 55 языках банту, большинство из которых имеет простые системы, состоящие из одного показателя (обычно итива ‘от дейктического центра’) или двух (итив ‘от дейктического центра’, вентив ‘к дейктическому центру’). Также обсуждается синтаксический статус перемещающегося участника ситуации (в общем случае — субъект), такисные отношения между сопутствующим движением и ситуацией, выражаемой исходным глаголом (в общем случае — предшествующее движение; только один язык противопоставляет предшествующее и одновременное движение) и дальнейшую эволюцию значений соответствующих показателей в некоторых языках.

В шестнадцатой главе (“Associated motion and deictic directional in Atlantic languages”) С. Вуазен рассматривает данные 28 атлантических языков (< КОНГО-АТЛАНТИЧЕСКИЕ). Она показывает, что в части этих языков есть специализированные системы сопутствующего движения (в терминологии автора “exclusive associated motion”), которые содержат от одного до трех показателей (самый частотный обозначает предшествующее движение от дейктического центра). Она также рассматривает D-AM системы (где показатели глагольной ориентации могут обозначать сопутствующее движение с некоторыми глаголами) и AM-D системы (где показатели сопутствующего движения начинают употребляться в качестве показателей глагольной ориентации).

В семнадцатой главе (“Ventive, associated motion and aspect in Jóola Fóoñi (Atlantic)”) Д. Кресселя и А. Бассене обсуждается показатель глагольной ориентации (вентив) в атлантическом языке дьола-фони, который может функционировать как показатель сопутствующего движения (с глаголами, не обозначающими движение). Этот показатель также может обозначать, что человек, присутствующий в дейктическом центре, делал или будет делать что-то, находясь где-то в другом месте, а также выражать аспектуальное значение (‘развитие процесса в направлении некоторого результата’).

Восемнадцатая глава (“The extension of associated motion to direction, aspect and argument structure in Nilotic languages”) Д. Пэйн посвящена показателям глагольной ориентации (в частности, итиву и вентиву) в нилотских языках, которые в сочетании с глаголами, не обозначающими движение, могут функционировать как показатели

сопутствующего движения (в южнонилотских языках — только в сочетании со специальным показателем, который также обсуждается в гл. 19). Утверждается, что они могут менять модель управления глагола. (О моделях управления и показателях сопутствующего движения см. также гл. 22.) Они имеют и аспектуальные функции: итив маркирует имперфективность, а вентив — инхоативность / перфективность. Интересно, что в (восточнонилотском) масайском языке эти показатели почти всегда выражают одновременное движение, что противоречит иерархии ‘предшествующее движение > одновременное движение’ из [Guillaume 2016]. Кроме того, эти показатели употребляются чаще всего с пациенсом переходных глаголов и реже всего с агенсом переходных глаголов. Тем самым, мы наблюдаем абсолютивную стратегию, что не вполне укладывается в иерархию ‘субъект движения’ > ‘объект движения’.

Девятнадцатая глава (“The ‘along’-deictic-directional verb suffix complex in Kupsapiny”) К. Кавати описывает данные южнонилотского языка купсабини. Показатели итива и вентива могут маркировать глагольную ориентацию с глаголами движения. С глаголами, не обозначающими движение, эти показатели могут обозначать (последующее) сопутствующее движение субъекта, но лишь в сочетании со специальной морфемой (ср. также обсуждение в гл. 18) — и это сочетание показателей всегда привносит аспектуальный компонент (продолжительности, ‘for some time’).

В двадцатой главе (“At the intersection of associated motion, direction and exchoative aspect in the Koman languages”) М. Отеро исследует показатели глагольной ориентации в пяти команских языках⁴. Он рассматривает три класса глаголов: глаголы движения, глаголы действительности, не подразумевающие движение, и стивитивы. Присоединяясь к динамическим глаголам, не обозначающим движение, эти показатели могут обозначать последующее движение субъекта (предшествующее и одновременное движение выражается с помощью сериальных глагольных конструкций). Однако в сочетании со стивитивными глаголами значение одного из показателей варьирует по языкам: он может как обозначать сопутствующее движение, так и маркировать терминативный (“exchoative” в терминологии автора) аспект (6). Такое поведение показателей интересно в связи с иерархией А. Белькади (3), согласно которой стивитивные глаголы наиболее склонны к прочтениям, содержащим сопутствующее движение.

(6) Кома (< команские)

mányà p'èl-kí-n (i=jángú)
манго быть_красным-DD2-3PL LOC=Я.

‘Манго были красные (в Янгу)’. (сейчас уже не красные) (с. 796)

Пятая часть, содержащая две главы, посвящена языкам Азии. В **двадцать первой главе** (“Associated motion in Sino-Tibetan, with a focus on Gyalrongic and Kiranti”) Г. Жак, А. Лаосуа и Ш. Чжан описывают сопутствующее движение в сино-тибетских языках, фокусируясь на гьялронгских языках (в которых зафиксированы два показателя) и на языках киранти (в которых количество показателей варьирует от нуля до семи). В частности, они показывают, что в гьялронгских языках (в отличие от языков киранти) высока степень интеграции событий (например, в сферу действия отрицания попадают как лексический глагол, так и перемещение, кодируемое показателем сопутствующего движения).

Двадцать вторая глава (“Associated motion in Tungusic languages: A case of mixed argument structure”) Б. Пакендорф и Н. М. Стойновой посвящена описанию показателя (предшествующего) сопутствующего движения в пяти тунгусских языках. Одно из наблюдений авторов заключается в том, что глаголы с этим показателем могут вводить пространственные аргументы (цель, источник, траектория) в дополнение к аргументам основного глагола или вместо них, что нетипично для показателей этой категории. Впрочем, как отмечают авторы, «смешанная» аргументная структура может быть обусловлена нечеткой

⁴ Иногда эти языки включают в нило-сахарскую макросемью, но такая трактовка дискуссионна.

границей между категориями сопутствующего движения ('V и затем X') и движения с целью ('V, чтобы X'): в «целевых» прочтениях событие движения оказывается центральным (движение V произошло, но событие X еще не наступило).

Следует отметить некоторую непоследовательность терминологического аппарата книги. Редакторы решили не унифицировать используемые термины (см. таблицу всех использованных терминов в главах на с. 10). Для обозначения одного феномена в разных главах используются различные термины, например, итив и андатив, ср. в гл. 13 на с. 138 «андативная частица *āzà* (отгlossированная как 'it' от 'ITIVE')». Одни и те же термины могут использоваться для обозначения различных феноменов (например, термин «андатив» используется в разных главах для обозначения как ориентации от дейктического центра, так и отсутствия ориентации). Глоссирование в разных главах (а иногда и в пределах одной главы) также не унифицировано. Также хотелось бы отметить, что в рецензируемом издании не хватает отрицательного языкового материала, который мог бы сделать некоторые утверждения более убедительными.

Кажется не вполне удачным и обобщение данных в первой главе. Редакторы отмечают, что категория сопутствующего движения пересекается с категорией движения с целью (особенно в будущем времени или в императиве), с глагольной множественностью (ср. показатель «амбулатива» или «дистрибутива», имеющий значение 'делать в различных местах'), с аспектом и др. Они лишь упоминают эти смежные категории, ссылаясь на главы, в которых эта проблематика обсуждается. Тем не менее расширение этой главы более подробным обсуждением таких точек пересечения и включением релевантных примеров позволило бы читателям с первых страниц более глубоко погрузиться в исследуемую проблематику.

Интересно, что разные авторы приходят к различным выводам. Так, М. Драйер (гл. 4) считает, что категорию глагольной ориентации нельзя рассматривать как частный случай категории сопутствующего движения, в отличие от А. Белькади (гл. 5). Некоторые авторы не относят показатели движения с целью к категории сопутствующего движения (например, Дж. Лавстранд и Д. Росс, гл. 3), в то время как некоторые относят (М. Драйер, гл. 4). Таким образом, категория сопутствующего движения оказывается либо очень узкой (если учитывать лишь морфологические показатели, при этом исключая показатели движения с целью), либо весьма широкой (если учитывать все способы выражения сопутствующего движения, включая глагольные конструкции, а также конструкции движения с целью). Такие нечеткие границы этой категории могут затруднять ее холистическое понимание.

Подведем итог: рецензируемый сборник статей, несмотря на предварительный характер сделанных выводов и обобщений, несомненно, является важнейшим шагом в исследовании категории сопутствующего движения. Эта девятисотстраничная книга содержит много ценных сведений о языках из разных ареалов, типологических и теоретических наблюдений, а также очерчивает возможные направления дальнейших исследований.

СПИСОК СОКРАЩЕНИЙ

2, 3 — 2, 3 лицо	PL — множественное число
DD2 — показатель глагольной ориентации	PRF — перфект
F — женский род	SG — единственное число
IMP — императив	VEN — вентив
LOC — локатив	

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

Плунгян 2002 — Плунгян В. А. О специфике выражения именных пространственных характеристик в глаголе: категория глагольной ориентации. *Исследования по теории грамматики*. Вып. 2: *Грамматикализация пространственных значений в языках мира*. Плунгян В. А. (ред.). М.: Русские

- словари, 2002, 57–98. [Plungian V. A. On the specifics of expressing nominal spatial characteristics in the verb: The category of verbal orientation. *Issledovaniya po teorii grammatiki*. No. 2: *Grammatikalizatsiya prostranstvennykh znachenii v yazykakh mira*. Plungian V. A. (ed.). Moscow: Russkie Slovarei, 2002, 57–98.]
- Belkadi 2015 — Belkadi A. Associated motion with deictic directionals: A comparative overview. *SOAS Working Papers in Linguistics*, 2015, 17: 49–76.
- Belkadi 2016 — Belkadi A. Associated motion constructions in African languages. *Africana Linguistica*, 2016, 22: 43–70.
- Craig 1991 — Craig C. Ways to go in Rama: A case study in polygrammaticalization. *Approaches to grammaticalization*. Traugott E. C., Heine B. (eds.). Amsterdam: John Benjamins, 1991, 455–492.
- Guillaume 2000 — Guillaume A. Directionals versus associated motions in Cavineña. *LACUS Forum XXVI: The lexicon*. Melby A. K., Lommel A. L. (eds.). Fullerton (CA): The Linguistic Association of Canada and the United States, 2000, 395–401.
- Guillaume 2016 — Guillaume A. Associated motion in South America: Typological and areal perspectives. *Linguistic Typology*, 2016, 20(1): 81–177.
- Koch 1984 — Koch H. The category of “associated motion” in Kaytej. *Languages in Central Australia*, 1984, 1: 23–34.
- Rappaport Hovav 2008 — Rappaport Hovav M. Lexicalized meaning and the internal temporal structure of events. *Theoretical and crosslinguistic approaches to the semantics of aspect*. Rothstein S. D. (ed.). Amsterdam: John Benjamins, 2008, 13–42.
- Ross 2021 — Ross D. *Pseudocoordination, serial verb constructions and multi-verb predicates: The relationship between form and structure*. Ph.D. diss. Urbana: Univ. of Illinois at Urbana-Champaign, 2021.
- Talmy 1985 — Talmy L. Lexicalization patterns: semantic structure in lexical form. *Language typology and syntactic description*. Vol. 3. Shopen T. (ed.). Cambridge: Cambridge Univ. Press, 1985, 57–148.

Получено / received 22.11.2021

Принято / accepted 23.11.2021

ВОПРОСЫ ЯЗЫКОЗНАНИЯ

научный журнал Российской академии наук
(свидетельство о СМИ ПИ № ФС77-77284 от 10.12.2019 г.)

Оригинал-макет подготовлен С. С. Белоусовым

Адрес редакции: 119019, Москва, ул. Волхонка, 18/2,
Институт русского языка им. В. В. Виноградова РАН, редакция журнала «Вопросы языкознания»,
тел.: +7 495 637-25-16, e-mail: voprosy@mail.ru

Подписано к печати 10.10.2022 Дата выхода в свет 24.10.2022
Формат 70×100¹/₁₆ Уч.-изд. л. 15,5 Тираж 280 экз. Зак. 5/5а Цена свободная

Учредители: Российская академия наук, Институт русского языка им. В. В. Виноградова РАН

16+

Издатель: Российская академия наук
20 экземпляров распространяется бесплатно

Исполнитель по контракту № 4У-ЭА-134-21 ООО «Интеграция: Образование и Наука»
105082, г. Москва, Рубцовская наб., д. 3, стр. 1, пом. 13–14

Отпечатано в ООО «Институт информационных технологий»
Адрес типографии: 170021, Тверская область, г. Тверь, Дачная улица, дом 33, офис 1



ИСТОРИЧЕСКИЙ ФАКУЛЬТЕТ

НА БАЗЕ ИНСТИТУТА ВСЕОБЩЕЙ ИСТОРИИ РАН



БАКАЛАВРИАТ

История
Культурология
Археология



МАГИСТРАТУРА

Модели всемирной истории
Культура массовых коммуникаций
Теория и практика археологических исследований



АСПИРАНТУРА

Исторические науки
и археология
Теория и история культуры

Комплексные образовательные программы разработаны специалистами исторического факультета с учетом последних научных достижений и современных общемировых тенденций.

В основе образовательного процесса — передовые технологии обучения, направленные на развитие мышления и творческого потенциала личности, достижение успеха в профессиональной среде. Студенты факультета с первого курса погружаются в мир академической науки, слушают лекции ведущих российских ученых с мировыми именами и сами участвуют в научных мероприятиях. В образовательные программы, помимо обязательных дисциплин, предусмотрены федеральными стандартами, включены уникальные авторские учебные курсы.

5 ПРИЧИН ПОСТУПИТЬ В ГАУГН



ВЫДАЮЩИЕСЯ ПРЕПОДАВАТЕЛИ

Научные сотрудники ведущих институтов РАН, включая академиков, членов-корреспондентов, докторов и кандидатов наук.



ИНДИВИДУАЛЬНЫЙ ПОДХОД

Преподаватель общается с каждым студентом индивидуально, помогает в выборе вектора профессионального развития.



МЕЖДУНАРОДНОЕ СОТРУДНИЧЕСТВО

Отдельные лекции читают приглашенные специалисты из других стран. Большое внимание уделяется языковой подготовке.



УДОБСТВО

Факультеты находятся в Москве в непосредственной близости от метро. Обучение в магистратуре и аспирантуре в основном проходит в вечернее время. Подать документы можно онлайн.



СТУДЕНЧЕСКАЯ ЖИЗНЬ

Студенты ГАУГН могут участвовать в многочисленных студенческих клубах («Что? Где? Когда?», Клуб политического анализа, Китайский разговорный клуб и др.).



ГОСУДАРСТВЕННЫЙ
АКАДЕМИЧЕСКИЙ
УНИВЕРСИТЕТ
ГУМАНИТАРНЫХ НАУК

НА БАЗЕ РОССИЙСКОЙ АКАДЕМИИ НАУК
с 1994 года



Преподаватели – ведущие российские ученые

- более 30% – доктора наук
- более 45% – кандидаты наук



Стажировки в:

- ведущих научно-исследовательских организациях
- органах государственной власти
- крупнейших общественных организациях
- бизнес-структурах



Интеграция науки
и образования



Бюджетные
места



Насыщенная
студенческая жизнь



Отсрочка от армии

БАКАЛАВРИАТ

МАГИСТРАТУРА

АСПИРАНТУРА

НАПРАВЛЕНИЯ ПОДГОТОВКИ:

• История

• Философия

• Политология

• Социология

• Международные отношения

• Зарубежное регионоведение

• Востоковедение и африканистика

• Психология

• Культурология

• Археология

• Менеджмент

• Юриспруденция

• Экономика

КОНТАКТНАЯ ИНФОРМАЦИЯ



Горячая линия: +7 (499) 238-04-12



facebook.com/gaugn



instagram.com/gaugn_/



gaugn.ru



E-mail: info@gaugn.ru



vk.com/gaugn

