
РОССИЙСКАЯ АКАДЕМИЯ НАУК

ВОПРОСЫ ЯЗЫКОЗНАНИЯ



5

2024



НАУКА

— 1727 —

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Отделение историко-филологических наук

ВОПРОСЫ ЯЗЫКОЗНАНИЯ

Журнал основан в январе 1952 года

Выходит 6 раз в год

5

СЕНТЯБРЬ — ОКТЯБРЬ

Москва

2024

Главный редактор:

В. А. Плунгян

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

Зам. главного редактора:

Н. Б. Вахтин

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

В. И. Подлесская

д. ф. н., проф., Институт языкознания РАН

Редколлегия:

В. М. Алпатов

д. ф. н., проф., академик РАН, Институт языкознания РАН

И. М. Богуславский

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

М. Д. Воейкова

д. ф. н., Институт лингвистических исследований РАН

В. З. Демьянков

д. ф. н., проф., Институт языкознания РАН

Д. О. Добровольский

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

А. Ф. Журавлёв

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

П. В. Иосад

Ph.D., Эдинбургский университет, Великобритания

Н. Н. Казанский

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

В. И. Киммельман

Ph.D., Бергенский университет, Норвегия

Г. И. Кустова

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

А. М. Молдован

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

М. С. Полинская

Ph.D., проф., Мэрилендский университет, США

Е. В. Рахилина

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

Я. Г. Тестелец

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

Л. А. Янда

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

Зав. редакцией: Н. В. Ганнус

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

Отдел рецензий: М. И. Сатина

Статьи отбираются редколлегией журнала на основе анонимного независимого рецензирования.

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

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

Телефон: +7 495 637-25-16

E-mail: voprosy@mail.ru

Сайт: <https://vja.ruslang.ru>

© Российская академия наук, 2024

© Составление. Редколлегия журнала «Вопросы языкознания», 2024

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
2024

Editor-in-chief:

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

Assistant editors:

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

Editorial board:

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

Managing editor: Natalia V. GANNUS

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

Review editor: Maria I. SATINA

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

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

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

Telephone: +7 495 637-25-16

E-mail: voprosy@mail.ru

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

Содержание

О. Е. ПЕКЕЛИС. Временной союз <i>как</i> : синхрония, микродиахрония и типология	7
А. Ю. УРМАНЧИЕВА. Контакты родственных языков как фактор формирования энецких идиомов	25

Information structure and syntax of Uralic languages

E. K. SKRIBNIK, D. O. ZHORNIK. Information structure and syntax of Uralic languages: A foreword	60
N. A. MURAVYEV. The influence of internal possessors on argument coding in Northern Khanty	65
D. O. ZHORNIK. Interaction between grammatical phenomena and information structure in Northern Mansi	80
J. HIRVONEN. The information structure of Meadow Mari	97
И. А. ХОМЧЕНКОВА. Условные и уступительные конструкции с русскими союзами <i>если</i> и <i>хотя</i> в горномарийской речи	116

Обзоры

V. A. ZHUKOVA, A. B. LETUCHIY. William Croft's approach to constructions in comparison with the system of constructions in the Russian Constructicon	135
--	-----

Рецензии

N. A. SERDOBOLSKAYA. [Review of:] A. Behnke, B. Wagner-Nagy (eds.). <i>Clause linkage in the languages of the Ob-Yenisei area: Asyndetic constructions</i> . Leiden; Boston: Brill, 2024	154
M. B. ERMOLAEVA, E. A. LYUTIKOVA. [Review of:] D. G. Krivochen. <i>Syntax on the edge: A graph-theoretic analysis of sentence structure</i> . Leiden; Boston: Brill, 2023	161

Contents

Olga E. PEKELIS. Russian <i>kak</i> as a temporal conjunction: Synchronic, diachronic and typological perspective	7
Anna Yu. URMANCHIEVA. The role of contacts within genetically related languages in the formation of Enets idioms	25

Information structure and syntax of Uralic languages

Elena K. SKRIBNIK, Daria O. ZHORNIK. Information structure and syntax of Uralic languages: A foreword	60
Nikita A. MURAVYEV. The influence of internal possessors on argument coding in Northern Khanty	65
Daria O. ZHORNIK. Interaction between grammatical phenomena and information structure in Northern Mansi	80
Johannes HIRVONEN. The information structure of Meadow Mari	97
Irina A. KHOMCHENKOVA. Conditional and concessive constructions with Russian conjunctions <i>esli</i> 'if' and <i>xotja</i> 'although' in Hill Mari speech	116

Overviews

Valentina A. ZHUKOVA, Alexander B. LETUCHIY. William Croft's approach to constructions in comparison with the system of constructions in the Russian Constructicon ...	135
--	-----

Reviews

Natalia V. SERDOBOLSKAYA. [Review of:] A. Behnke, B. Wagner-Nagy (eds.). <i>Clause linkage in the languages of the Ob-Yenisei area: Asyndetic constructions</i> . Leiden; Boston: Brill, 2024	154
Marina B. ERMOLAEVA, Ekaterina A. LYUTIKOVA. [Review of:] D. G. Krivochen. <i>Syntax on the edge: A graph-theoretic analysis of sentence structure</i> . Leiden; Boston: Brill, 2023	161

Временной союз *как*: синхрония, микродиахрония и типология

© 2024

Ольга Евгеньевна Пекелис

Российский государственный гуманитарный университет, Москва, Россия;
Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия; opekelis@gmail.com

Аннотация: В статье исследуется развитие союза *как* во временном значении после XIX в. На материале Национального корпуса русского языка выявляются семантические и формальные ограничения, определяющие употребление *как* в современном языке, и демонстрируется, что еще два века назад многие из этих ограничений отсутствовали. Произошедшие изменения обобщаются на уровне семантики и грамматики. Семантическую причину изменений мы видим в том, что после XIX в. *как* потерял способность вводить временной ориентир для главной ситуации. С точки зрения грамматики изменения предлагается интерпретировать как сдвиг в степени синтаксической интеграции вводимой союзом *как* клаузы: от тесной интеграции к слабой. Последняя интерпретация подвергает сомнению распространенное представление о том, что временным клаузам в разных языках свойственна высокая степень интеграции.

Ключевые слова: диахрония, подчинение, придаточные обстоятельственные, русский язык XIX в., сложное предложение, союзы

Благодарности: Автор благодарит участников конференции «Связь пропозициональных единиц в предложении и в тексте», состоявшейся в Институте языкознания РАН в сентябре 2023 г., за плодотворное обсуждение материалов, легших в основу настоящей статьи. Автор также признателен анонимному рецензенту журнала «Вопросы языкознания» за ценные замечания и комментарии. Исследование выполнено при поддержке гранта РНФ № 22-18-00120.

Для цитирования: Пекелис О. Е. Временной союз *как*: синхрония, микродиахрония и типология. *Вопросы языкознания*, 2024, 5: 7–24.

DOI: 10.31857/0373-658X.2024.5.7-24

Russian *kak* as a temporal conjunction: Synchronic, diachronic and typological perspective

Olga E. Pekelis

Russian State University for the Humanities, Moscow, Russia; HSE University, Moscow, Russia;
opekelis@gmail.com

Abstract: The paper deals with the evolution of the Russian temporal conjunction *kak* after the 19th century. Based on data from the Russian National Corpus, the paper investigates the semantic and formal restrictions that determine the use of *kak* in modern Russian but that were absent from the Russian language of the 19th century. The changes that have occurred are interpreted from both semantic and grammatical points of view. Semantically, they can be attributed to the fact that after the 19th century *kak* lost the ability to introduce a time reference for the main situation. Grammatically, I argue that the changes are due to a shift in the degree of syntactic integration of the clause introduced by *kak*: from tight integration to a weaker one. The latter interpretation challenges the widespread idea that temporal clauses tend to be tightly integrated cross-linguistically.

Keywords: adverbial complements, complex sentence, conjunctions, diachrony, subordination, 19th-century Russian

Acknowledgements: I am profoundly thankful to the participants of the conference on clause linkers (“Svjaz’ propozicional’nyx edinic v predložanii i v tekste”) that took place at the Institute of Linguistics of the Russian Academy of Sciences in September 2023 for the fruitful discussion of my data. I am also thankful to an anonymous reviewer of *Voprosy Jazykoznanija*, who provided valuable comments and corrections. The work has been supported by the Russian Science Foundation (grant No. 22-18-00120).

For citation: Pekelis O. E. Russian *как* as a temporal conjunction: Synchronic, diachronic and typological perspective. *Voprosy Jazykoznanija*, 2024, 5: 7–24.

DOI: 10.31857/0373-658X.2024.5.7-24

Введение

Союз *как* во временном значении широко употреблялся в древнерусском языке [Лавров 1941: 96 и сл.], но после XVIII в. постепенно уступил место союзу *когда* как основному и наиболее нейтральному маркеру адвербиальных временных клауз [Виноградов, Шведова 1964: 152]. В современном языке временной *как* все еще используется, но занимает периферийное положение. Во-первых, *как* тяготеет к разговорным и просторечным текстам, т. е. стилистически маркирован [Кобозева, Инькова 2018: 232]. Во-вторых, *как* более ограничен в своей дистрибуции, чем *когда*. Например, в отличие от *когда*, *как* не употребляется для выражения неполной одновременности однократных ситуаций [Там же: 233], ср. (1) и (2) ¹. Такие примеры с *как* встречаются в текстах XIX в., но устарели с точки зрения современной нормы.

- (1) *Как я входил, мне слышалось, что вы изволите говорить о собаках* [М. Н. Загошкин. Москва и москвичи (1842–1850)].
- (2) *Когда я входил, мне слышалось <...>*.

В настоящей работе исследуется эволюция временного *как* после XIX в. Мы стремимся не просто выявить изменения в свойствах этого союза, но и предложить им интерпретацию, т. е. обнаружить общее содержание произошедшего с *как* сдвига. Мы формулируем это содержание в двух плоскостях: семантики и грамматики. С семантической точки зрения за изменениями, произошедшими с *как*, стоит, на наш взгляд, потеря союзом *как* способности маркировать зависимую ситуацию как временной ориентир для главной. Ту же идею, но в грамматических терминах, мы формулируем с опорой на типологически ориентированную классификацию адвербиальных клауз, противопоставляющую клаузы по степени синтаксической интеграции с главной клаузой. С этой точки зрения клаузы с *как* после XIX в. эволюционировали от тесно интегрированных к слабо интегрированным. Использование при интерпретации *как* типологически ориентированного понятийного аппарата позволяет не только описать в этих терминах *как* в его синхронном состоянии, но и засвидетельствовать диахронический сдвиг внутри классификации ².

¹ Примеры с указанием источника, если не сказано иное, заимствованы в Национальном корпусе русского языка [НКРЯ].

² Союз *как* может сочетаться с коррелятами *так* (i) или *то*. От одноместного *как* двуместные *как... так* и *как... то* отличаются семантическими [Кобозева, Инькова 2018: 220] и грамматическими [Тестелец 2001: 263] свойствами. В настоящей работе двухместные сочетания с *как* не рассматриваются.

(i) *Ширвиндов как пришел, так сразу на бережок сел <...>* [Григорий Горин. Иронические мемуары (1990–1998)].

Работа имеет следующую структуру. В разделе 1 описываются свойства современного *как* и изменения в этих свойствах по сравнению с языком XIX в. В разделе 2 излагается и обосновывается общее семантическое содержание произошедшего сдвига. В разделе 3 вводятся основные принципы классификации адвербиальных клауз по степени синтаксической интеграции и в этих терминах определяется союз *как*.

1. Свойства временного *как* в синхронии и диахронии

В настоящем разделе описываются изменения в свойствах *как* за последние 200 лет. Изучаемые свойства для удобства изложения разделены на два класса: аспектуально-темпоральные (раздел 1.1) и прочие свойства (раздел 1.2).

1.1. Аспектуально-темпоральные свойства

Мы рассмотрим ограничения на употребление *как* с точки зрения двух семантических параметров: тип таксисного отношения между ситуациями, соединяемыми посредством *как*, и аспектуально-темпоральный тип ситуаций. Предлагаемое описание опирается на каталогизацию разных употреблений *как* в этих терминах, осуществленную в [Кобозева, Инькова 2018: 215–235]. Вслед за авторами этой работы, мы выделяем два типа таксисного отношения, доступных для *как*: предшествование и одновременность, а для последнего случая — полную и неполную одновременность³. При полной одновременности периоды существования двух ситуаций совпадают, при неполной — одна из ситуаций задает временной период, рамками которого ограничена вторая ситуация ([Кобозева, Инькова 2018: 232–233], см. также [Храковский 2009: 20]). Релевантное для *как* аспектуально-темпоральное разграничение касается противопоставления эпизодических ситуаций, т. е. локализованных во времени, и итеративных — повторяющихся неопределенное количество раз в неопределенные моменты или периоды⁴. Наше изложение схематично и отвлекается от тонкостей (о которых см. [Кобозева, Инькова 2018]) в употреблении *как* в разных семантических условиях, но, в соответствии с нашими задачами, сосредоточено на произошедших с *как* микродиахронических сдвигах.

Предшествование, по данным [Там же: 232], встречается в контексте временного *как* заметно чаще, чем одновременность. При этом ограничений на аспектуально-темпоральный тип ситуаций в современном языке нет: *как* может соединять и эпизодические (3), и итеративные ситуации (4).

- (3) *Как придете, прочтете две страницы текста* <...> [Маша Трауб. Домик на Юге (2009)].
- (4) *Михалков как приезжает, спешит на охоту на своем джипе «Тигр»* <...> [Комсомольская правда, 27.07.2007].

³ Временной *как* встречается и в полипредикативных конструкциях со значением следования, ср.: ...она только-только привыкла к тому, что она — пани, как неожиданно оказалось, что она еще и мадам! [Татьяна Устинова. Подруга особого назначения (2003)].

Это употребление стоит особняком, поскольку проявляет признаки сочинительной конструкции [Пекелис 2022], и в настоящей работе рассматривается.

⁴ В [Кобозева, Инькова 2018] итеративные ситуации обозначаются как темпорально не локализованные.

Такое же положение дел наблюдалось и в языке XVIII–XIX вв., ср. в (5) *как* в контексте эпизодических ситуаций и в (6) — в контексте итеративных:

- (5) *Как Святослав приехал в Новград, новгородцы <...> Ростислава Юриевича выслали* [В. Н. Татищев. История российская в семи томах. Том второй (1750)].
- (6) *Какъ ночь наступаетъ то они сколько ихъ ни есть въ деревнѣ, сберутся въ одну ватагу, для молитвы назначенную, и тамъ съ пол часа молятся <...>* [Ф. А. Эмин. Непостоянная фортуна или походжение Мирамонда. Ч. 2 (1763)].

Одновременность совместима с *как* прежде всего в том случае, если ситуации итеративные. При этом в контексте неполной одновременности (7)–(8) такое употребление представляется более естественным, чем при полной одновременности (9)⁵.

- (7) *Как из школы идем, всегда встречаем наших мальчишек* [модифицированный корпусный пример].
- (8) *Как зима, обязательно простываю <...>* [Независимая газета (приложение «Кулиса НГ»), 13.11.1999].
- (9) *Как препарат вводится, идет небольшой болевой эффект, меньше минуты, и он уходит* [РИА Новости, 05.12.2020].

В контексте одновременных эпизодических ситуаций употребление *как*, по-видимому, устарело. Для случая неполной одновременности этот факт, иллюстрируемый примерами (1) и (10), неоднократно отмечался (см., наряду с [Кобозева, Инькова 2018: 233], также [РГ 1980: 549]).

- (10) *Какъ я шель домой съ рынку, попался мнѣ навстрѣчу мой учитель <...>* [Ф. А. Эмин. Непостоянная фортуна или походжение Мирамонда. Ч. 2 (1763)].

Полная одновременность эпизодических ситуаций, по мнению [Кобозева, Инькова 2018: 232], выражается с помощью *как* и сегодня. Однако иллюстрирующие эту точку зрения примеры, ср. (11), все относятся к XIX в.:

- (11) *Как последний раз виделись цвела ты, как маков цвет <...>* [П. И. Мельников-Печерский. В лесах. Книга первая (1871–1874)].

Напомним, что в современных текстах временной *как* считается стилистически маркированным независимо от семантики контекста. Это осложняет дифференциацию современных и несовременных употреблений. Наша точка зрения об архаичности использования *как* для выражения полной одновременности эпизодических ситуаций поддерживается контрастами типа (12) и (13). Неудачные примеры (12а) и (13а) выражают (полную) одновременность, а приемлемые (12б) и (13б) выражают предшествование.

- (12) а. ^{??}*Он как входил, громко пел.*
б. ^{OK}*Он как вошел, громко поздоровался.*
- (13) а. ^{??}*Они как шли обратно, ели мороженое.*
б. ^{OK}*Они как вернулись домой, съели по мороженому.*

В языке XVIII–XIX вв. ограничений на употребление *как* при одновременности ситуаций, подобных современным, не было. В НКРЯ встречаются примеры всех четырех логически возможных комбинаций: неполная и полная одновременность эпизодических ситуаций, уже проиллюстрированные в (10) и (11), соответственно, и полная (14) и неполная (15) одновременность вневременных ситуаций.

⁵ Мы благодарим анонимного рецензента за это наблюдение.

- (14) *Как я вспоминаю теперь, то вижу, что многое <...> было неудобно и тяжело <...>* [Л. Н. Толстой. Семейное счастье (1859)].
- (15) *Как улей сажаем, богоявленской водой его кропим...* [В. И. Немирович-Данченко. Святые горы (1880)].

Данные в таблице 1 обобщают сказанное об ограничениях на употребление временного *как* в современном языке.

Таблица 1

Аспектуально-темпоральные ограничения на употребление *как* в современном языке

Такс. отношение	Тип ситуации	
	Эпизодические	Итеративные
предшествование	+	+
одновременность	полная	+/-
	неполная	+

1.2. Другие свойства

Настоящий раздел посвящен более формальным свойствам клаузы с *как*: ограничениям на временную референцию (раздел 1.2.1) и коммуникативный статус (1.2.2), сочетаемости с соотносительными словами (1.2.3) и контрастивными частицами (1.2.4).

1.2.1. Временная референция

В современном языке клауза с временным *как* чаще реферирует к будущему, чем к прошлому. Для препозитивных клауз обнаруживается значимое преобладание референции к будущему над референцией к прошлому по сравнению с клаузами, вводимыми союзом *когда* (см. таблицу 2; двусторонний вариант точного критерия Фишера, $P < 0,05$). Для пост- и интерпозитивных клауз значимого отличия от *когда* не было выявлено (см. таблицу 3). Тем не менее корпусные данные и в этом случае не противоречат предположению о преобладании среди клауз с *как* референции к будущему.

Таблица 2

Частотность форм прошедшего и будущего времени глагола в контексте препозитивных *как* и *когда* (Подкорпус 1980–2020 гг. Основного корпуса НКРЯ)⁶

	Fut	Praet	Всего	Доля praet
<i>Как</i>	17	8	25	0,3
<i>Когда</i>	50	153	203	0,8

⁶ Образец запроса: *как*, first -оказаться & -выясниться & -предполагаться & -уметь & -мочь, v & indic & praet, на расстоянии от 1 до 1 от Слова 1 v & indic & praet, асomma, на расстоянии от 1 до 1 от Слова 2. Примеры отфильтрованы вручную.

Во всех примерах выборки в зависимой клаузе отсутствует выраженное подлежащее, поскольку такое употребление характерно для современного *как*. Ср. контраст в (i)–(iii):

(i) ...он, как подрост, стал сторониться ее [Михаил Шишкин. Всех ожидает одна ночь (1993–2003)].

Таблица 3

**Частотность форм прошедшего и будущего времени глагола
в контексте постпозитивных и интерпозитивных *как* и *когда*
(Подкорпус 1980–2020 гг. Основного корпуса НКРЯ)⁷**

	Fut	Praet	Всего	Доля praet
<i>Как</i>	8	2	10	0,2
<i>Когда</i>	67	58	125	0,5

1.2.2. Коммуникативный статус

Использование временного *как* в современном языке затруднено в случае, если зависимая клауза имеет коммуникативный статус ремы, при этом и зависимая, и главная ситуации локализованы в прошлом. Ср. контраст в (16)–(19). В приемлемых примерах (16), (18) и (19) перечисленные условия не выполнены: в (16) и (19) обе ситуации локализованы в будущем, а в (18), хотя ситуации локализованы в прошлом, зависимая клауза препозитивна и выступает в роли темы⁸. Напротив, в примере (17), мало удачном в контексте союза *как*, условия выполнены: клаузы реферируют к прошлому, зависимая клауза соответствует реме. Для сравнения, союз *когда* уместен во всех четырех примерах, включая (17).

- (16) *Короче, позвонишь/, как (^{OK}когда) прочухаешься* [Екатерина Завершнева. Высотка (2012)].
- (17) <— *Когда же он позвонил?> — [?]Позвонил/, как (^{OK}когда) прочухался\.*
- (18) *Как (^{OK}когда) прочухался/, позвонил\.*
- (19) *Как (^{OK}когда) прочухаешься/, позвонишь\.*

В (17) главная и зависимая ситуации являются эпизодическими и связаны отношением предшествования, т. е. таким отношением, с которым, вообще говоря, современный *как* совместим (см. раздел 2.1). Другая разновидность приемлемого для современного *как* отношения — неполная одновременность итеративных ситуаций. Пример (20) демонстрирует, что и это отношение становится невозможным для *как* (но не для *когда*) при вступлении

(ii)^{??}*Как он попрос, стал сторониться ее.*
(iii)^{OK}*Когда он попрос, стал сторониться ее.*

Кроме того, поскольку в выборку с будущим временем автоматически попали только глаголы совершенного вида, а также глагол *быть*, глаголы несовершенного вида, отличные от *быть*, были также исключены из выборки с прошедшим временем. Как следствие, в выборках отсутствовали примеры с отношением одновременности между ситуациями.

⁷ Образец запроса: Слово 1: *v & indic & praet & first* Слово 2: *когда & acomma* на расстоянии 1 от Слова 1 Слово 3: *-оказаться & -выясниться & -предполагаться & -уметь & -мочь & v & indic & praet* на расстоянии 1 от Слова 2. См. также комментарии в предыдущей сноске.

⁸ Символами «/» и «\» обозначены, соответственно, восходящий акцент, ассоциируемый с темой, и нисходящий акцент, ассоциируемый с ремой, согласно нотации, принятой в [Янко 2001: 36–37].

в силу указанных условий: рематичность зависимой клаузы и референция к прошлому⁹. Модификация этого примера в (21), где референция к прошлому заменена на референцию к будущему, ожидаемо более удачна.

(20) *Он работает с конструктором, и можно было подойти, послушать, самому вмешаться: «Сергей Владимирович, вот так лучше...» — «Да, правильно... Давай так...». Позволял, когда (как) видел, что не просто ротозейничаешь* [Феликс Чуев. Ильяшин (1998)].

(21) ^{OK} *Обязательно позволит, как увидит, что ты не просто ротозейничаешь.*

Отметим также, что в нашей корпусной выборке с *как* и разными типами временной референции (см. раздел 1.2.1) всего два примера, где клауза с *как* отсылает к прошлому и при этом не препозитивна, т. е. позиционно не тяготеет к тематичности (ср. данные в таблице 3). Показательно, что в обоих примерах эта клауза — тема (в этом случае заударная), а не рема:

(22) *В Ленинграде купил, в комиссионке. Обомлел \, как увидел, и не сдержался* [Андрей Дмитриев. Шаги (1987)].

(23) *Сбежал \, как открылась, в лавочку, пива ей принес <...>* [Елена Хаецкая. Синие стрекозы Вавилона (2004)].

Между тем еще в XVIII в. отмеченное ограничение отсутствовало или действовало менее строго: временной *как* мог вводить рему в том числе с референцией к прошлому. Так, оба условия соблюдены в (24) и (25); о рематическом статусе зависимой клаузы свидетельствует не только акцентный паттерн, но и то, что эта клауза постпозитивна и вводит новую информацию:

(24) *Иван-купеческий сын снял с руки перстень и подает старику, но еще больше прежнего удивился \, как увидел, что старика нет * [Сказка пятая о Иване-купеческом сыне (1787)].

(25) *Трудно узнать свою природу, а чужую познать и того труднее. Узнаешь, да поздно. Черепаха ошибку почувствовала \, как начала лететь * [Григорий Сковорода. Разговор, называемый Алфавит или Букварь мира (1760–1775)].

1.2.3. Сочетаемость с соотносительными словами

Современный *как* ведет себя избирательно при сочетании с соотносительными словами в составе главной клаузы. К соотносительным словам, с которыми сегодня *как* сочетается, относятся *всегда, теперь, иногда*. Ср.:

(26) *Короце, Данюшке теперь, как Дуся померла, положена пятая часть* [«Дружба народов», 1998].

(27) *Иногда, как вспомню я, доченька, что ты так далеко, на севере, и одна среди чужих людей, то и плакать мне хочется* [Б. Л. Горбатов. Поединок (1936)].

(28) *Всегда, как приезжала Краля, Мишенька тенью ходил по комнатам и подглядывал, подслушивал их разговоры, сыпал проклятия* [«Волга», 2015].

⁹ В (20) оказываются совмещены референция к прошлому и то, что мы назвали итеративностью: с одной стороны, речь идет о повторении ситуации неопределенное количество раз в неопределенные моменты, с другой стороны — это повторение ограничено прошлым.

Напротив, соотносительные слова *тогда* и *однажды*, а также синонимичное *однажды* выражение *один раз*, в контексте *как* сегодня практически не используются. Однако они встречаются в текстах XIX в. Ср.:

- (29) *Печь нагревается и нагревает не в то время, когда топится, а тогда, как закроют трубу* [Н. А. Бестужев. Письмо И. И. Свиязеву (1851)].
- (30) *Однажды, как батюшка и матушка уехали в гости на целый день к одному старому родственнику, <...> вдруг отворяются двери и входит этот молодой человек в сопровождении няни* [М. П. Погодин. Преступница (1830)].
- (31) *Один раз, как я был в своей комнате, вошла ко мне <...> жена бывшего моего товарища, с ее дочерью* [В. Т. Нарезный. Российский Жилбляз, или Похождения князя Гаврилы Симоновича Чистякова (1814)].

Заметим, что по своим темпорально-аспектуальным характеристикам по крайней мере некоторые из таких примеров отвечают современным требованиям, так что причина того, что они не используются, кроется именно в соотносительном слове. Так, в (29) ситуации связаны отношением предшествования, которое сегодня в контексте *как* допустимо (см. раздел 1.1). Не случайно при элиминации *тогда* пример становится приемлемым:

- (32) *Как закроют трубу, печь нагревается.*

Данные о частотности сочетания *как* с *тогда*, *однажды* и *один раз* начиная с XVIII в. отражены в таблице 4. Можно видеть, что из современного языка все три сочетания ушли.

Таблица 4

Частотность сочетания *как* с *тогда*, *однажды* и *один раз* по периодам (Основной корпус НКРЯ)¹⁰

	1701–1800, объем подкорпуса 6 629 622 слова		1801–1850, объем подкорпуса 17 755 589 слов		1851–1900, объем подкорпуса 62 454 587 слов		1901–1950, объем подкорпуса 99 404 501 слово		1951–2023, объем подкорпуса 291 031 591 слово	
	ipm	N	ipm	N	ipm	N	ipm	N	ipm	N
<i>тогда, как</i>	5	35	4	79	1	79	0,2	15	0,007	2
<i>однажды, как</i>	0	0	0,3	5	0	0	0	0	0	0
<i>один раз, как</i>	0	0	0,06	1	0,05	3	0,01	1	0,003	1

1.2.4. Сочетаемость с контрастивными частицами

Современный *как* не сочетается с контрастивными частицами *только*, *именно*, *как раз* и т. п. Однако такие сочетания обнаруживаются в текстах XVIII — начала XX в. Основного корпуса НКРЯ. Во встретившихся примерах с *только* и *именно* частица модифицирует союз не непосредственно, а через соотносительное слово *тогда*, как в (33) и (34).

- (33) *Поликею легче стало только тогда, как он выехал за город* [Л. Н. Толстой. Поликушка (1863)].

¹⁰ Образцы запросов доступны по ссылкам: <https://ruscorpora.ru/s/eVZAo>, <https://ruscorpora.ru/s/dN102>, <https://ruscorpora.ru/s/aO98R>. В целях репрезентативности в выборке со словом *тогда* от каждого текста был отобран один пример. В выборках с *однажды* и *один раз*, в силу их очень маленького объема, такой фильтрации не проводилось.

- (34) ...ведь надобно ж, чтоб это случилось именно тогда, как я в отсутствии [М. Ю. Лермонтов. Герой нашего времени (1839–1841)].

В сочетаниях с частицей *как раз* частица модифицирует союз без участия соотносительного слова:

- (35) *Как раз, как тебе проснуться, поспеет* [Н. Н. Алексеев. Заморский выходец (1900)].
(36) *Тогда я взял себя в руки и решил только с прицелу стрелять и как раз как решил вылетел дупель* [М. М. Пришвин. Дневники (1927)].

Все такие примеры производят впечатление архаичных.

1.2.5. Свойства *как*: обобщение

Мы продемонстрировали одну сочетаемостную тенденцию и три запрета на сочетаемость современного временного *как*:

- клаузы, соединяемые с помощью *как*, чаще реферируют к будущему, чем к прошлому;
- клауза, вводимая с помощью *как*, не может иметь одновременно статус ремы и референцию к прошлому;
- клауза, вводимая с помощью *как*, не сочетается с соотносительными словами *тогда*, *однажды* и *один раз*;
- клауза, вводимая с помощью *как*, не сочетается с контрастивными частицами.

В языке XVIII–XIX вв. перечисленные запреты не действовали¹¹.

2. Временной *как*: интерпретация произошедшего сдвига

В настоящем разделе предпринята попытка сформулировать содержание сдвига, произошедшего с временным *как* после XIX в. (раздел 2.1). Основным доводом в пользу предлагаемой интерпретации служит ее объяснительная сила: объяснение получают, как мы стремимся показать, изменения и в семантических, и в формальных свойствах *как* (раздел 2.2).

2.1. Общее содержание сдвига

Мы предполагаем, что сдвиг, произошедший с *как* после XIX в., имеет следующее общее содержание: союз *как* во временном значении потерял способность вводить временной ориентир для главной ситуации. Как следствие, *как* утратил сочетаемость с контекстами, которые подчеркивают у зависимой клаузы статус временного ориентира.

Поясним, что мы понимаем под временным ориентиром. В наиболее прозрачном случае временной ориентир выражается с помощью обстоятельства времени в составе простого предложения, ср.: *Он приехал в 7 часов*. В составе полипредикативной конструкции временной ориентир — это прежде всего такая клауза, которая отвечает на вопрос *когда?*, ср. *Он приехал, когда было 7 часов*. Клауза-ориентир может и не содержать указания на конкретное время, но она все равно должна отвечать на (явный или подразумеваемый) вопрос *когда?*. Ср.:

¹¹ Вопрос о том, действовала ли в языке XVIII–XIX вв. упомянутая тенденция, остается за рамками работы из-за технических трудностей, связанных с его исследованием.

(37) *Возможно, он приехал, когда я уже уехал* (Google).

Такого рода контексты, согласно нашей интерпретации, после XIX в. ушли из сферы употребления *как*.

Закономерен, однако, вопрос, в чем же, если не в выражении временного ориентира, состоит функция современного временного *как*? Мы думаем, что в конструкции с *как* выражается упорядоченность ситуаций на временной оси, т. е. то, как соотносятся ситуации по своей временной локализации. Например, в (38) говорится о том, что ситуация ‘он придет’ наступит после ситуации ‘будет потеплее’.

(38) *Придет, как будет потеплее* (Google).

Правда, такое прочтение в (38) не является единственно возможным: ничто как будто не мешает считать ситуацию ‘будет потеплее’ ответом на вопрос *Когда он придет?*, т. е. считать временным ориентиром для ситуации ‘придет’. То, что союз *как*, тем не менее, используется в (38), объясняется, на наш взгляд, нейтрализацией различий между двумя прочтениями: контекст в (38) допускает оба прочтения, и в этих условиях современный *как* допустим. Но есть и такие контексты, где интерпретация зависимой клаузы как временного ориентира предпочтительна, и тогда *как* оказывается неуместен с точки зрения современной нормы. (37) как раз относится к таким контекстам, ср. невозможность замены *когда* на *как*:

(39) *“Возможно, он приехал, как я уже уехал.*

Оговоримся, что нейтрализация двух прочтений может возникать и в контексте прямого вопроса со словом *когда*. В этом случае ответить на такой вопрос оказывается возможным с помощью союза *как*:

(40) *Когда вы выходите? — Как темнеет* [пример анонимного рецензента].

Вместо клаузы с *как* в качестве ответа в (40) можно использовать сочинительную конструкцию, ср. (41). В контексте сочинения прочтение в значении упорядоченности не просто уместно — оно вытесняет прочтение в значении ориентира.

(41) *Когда вы выходите? — Темнеет, и сразу выйдем.*

Почему, в отличие от вопроса типа (40), в контексте типа (39) значение временного ориентира оказывается единственно возможным? Это связано с тем, что в (39) не просто подразумевается вопрос с *когда*, ответом на который служит клауза с *как*, — этот вопрос к тому же является контрастивным, т. е. предполагает сравнение двух потенциальных временных ориентиров (‘он приехал, когда я уже уехал’ vs. ‘он приехал в какой-то другой момент’). Контрастивный вопрос с *когда?*, таким образом, однозначно предопределяет прочтение ответа в значении ориентира и исключает возможность использовать союз *как* с позиций современной нормы.

Заметим, наконец, что противопоставление двух прочтений — временная ориентация vs. упорядочивание на временной оси — не является новым и для описания временных клауз уже использовалось. В этих терминах предлагалось различать, например, русские временные союзы *прежде чем* и *перед тем как* [Барентсен, Пупынин 1999: 113; Храковский 2005: 37].

2.2. Временной *как*: объяснение свойств

Продолжая разграничение, принятое в разделе 1, в настоящем разделе мы предлагаем объяснение сначала для аспектуально-темпоральных свойств *как* (раздел 2.2.1), а затем для остальных (раздел 2.2.2).

2.2.1. Аспектуально-темпоральные свойства

За последние 200 лет *как* потерял способность соединять эпизодические ситуации в конструкции со значением одновременности (см. подробнее раздел 1.1). Напротив, с эпизодическими ситуациями в конструкции предшествования, а также с итеративными ситуациями независимо от типа таксисного отношения между ними, *как* используется по сей день (при одновременности, предположительно, реже, чем при предшествовании, а при полной одновременности — реже, чем при неполной). Таким образом, современный *как* предпочел предшествование одновременности, а итеративные ситуации эпизодическим.

Нам кажется, что это направление эволюции получает объяснение, если развитие *как* интерпретировать как потерю способности обозначать временной ориентир. В самом деле, как раз те комбинации аспектуально-темпоральных признаков, которые ушли из сферы употребления *как* — одновременность и эпизодичность — в наибольшей степени могут ассоциироваться с понятием временного ориентира. Ср. прототипическую конструкцию с временным ориентиром *Он пришел в 7 часов*: ситуация ‘наступило 7 часов’, во-первых, имеет конкретную временную локализацию, т. е. является эпизодической в наших терминах, а во-вторых, эта ситуация произошла одновременно с ситуацией ‘он пришел’, для которой она служит временным ориентиром. И предшествование, и итеративность представляют собой отступления от прототипического временного ориентира и поэтому дольше сохраняются в сфере употребления *как*.

В контексте итеративных ситуаций, *как* мы предположили, неполная одновременность сохранилась в современном языке лучше, чем полная. Эту тенденцию тоже можно связывать с потерей союзом *как* способности вводить ориентир. Полная одновременность предполагает совпадение двух ситуаций во времени. Это сближает зависимую ситуацию с временным ориентиром, несмотря на итеративность. Напротив, при неполной одновременности (ср. *Как зима, обязательно простужаюсь*) зависимая ситуация служит своего рода фоном для главной, но не временным ориентиром. В качестве ориентира она не может квалифицироваться потому, что в некоторые моменты зависимая ситуация имеет место, а главная — нет (*Как зима, обязательно простужаюсь* не означает ‘я простужен всю зиму’).

2.2.2. Другие свойства

В современном языке временной *как* чаще вводит ситуации с референцией к будущему, чем с референцией к прошлому, а при рематическом статусе зависимой клаузы — вовсе не употребляется с ситуациями, локализованными в прошлом (см. подробнее разделы 1.2.1 и 1.2.2). Последний запрет до XX в. отсутствовал. Как кажется, все это — следствие того, что *как* потерял способность вводить временной ориентир. В самом деле, временной ориентир по определению отвечает на вопрос *когда?*, поэтому должен получить статус ремы. Референция к прошлому тоже предрасполагает к прочтению клаузы, вводимой *как*, в значении временного ориентира: ситуации, реферирующие к прошлому, уже произошли и имеют конкретную временную локализацию. Можно предположить, что сочетание этих двух признаков — референция к прошлому и статус ремы — вместе предопределяют прочтение зависимой клаузы как временного ориентира для главной. Это несовместимо со свойствами современного *как*, но не запрещалось в языке XIX в.

Избирательность *как* при сочетании с соотносительными словами получает такое же объяснение: *как* не сочетается с теми соотносительными словами, которые предопределяют интерпретацию вводимой *как* клаузы как временного ориентира для главной. Закономерным образом, в языке XIX в. аналогичных запретов на сочетаемость не было. Напомним, что к совместимым с *как* относятся слова *теперь*, *иногда* и *всегда*, к запрещенным — слова *тогда*, *однажды* и выражение *один раз* (см. подробнее раздел 1.2.3).

В контексте *тогда* временная клауза с необходимостью вводит временной ориентир для главной. Так, в (42) клауза с *когда* описывает момент, в который *наступит равновесие системы*. Показательно, что *тогда, когда* может вводить ответ на вопрос о времени — с вопросительным словом *когда* (43) или без него (44).

- (42) *Равновесие системы наступит тогда, когда давление воды в цилиндрах на одинаковых уровнях будет одинаковым* [Владимир Лукашик, Елена Иванова. Сборник задач по физике. 7–9 кл. (2003)].
- (43) *Так когда Вы мне выдадите бумагу? — обратился я вновь к Рычкину. — Тогда, когда пойду туда, где она лежит, — ответил грубо он* [В. И. Максимов. Дневник научного сотрудника (2004)].
- (44) — *Я должен делать это каждый день? — Тогда, когда будет получаться само. Но весьма желательно, чтобы это получалось само каждый день* [Виктор Пелевин. S.N.U.F.F (2011)].

Соотносительное слово *однажды* и выражение *один раз* задают временную локализацию главной клаузы, конкретизацией которой служит зависимая клауза. Например, в (45) говорится, что *Бабка-Ежка подпрыгнула от злости однажды*, а именно тогда, когда *червяк слишком громко захрапел*. Таким образом, *однажды* и его синонимы тоже подчеркивают роль зависимой клаузы как временного ориентира для главной.

- (45) *Но вот однажды, когда червяк уж слишком громко захрапел во сне, Бабка-Ежка подпрыгнула на месте от злости и принялась колотить кулаками в стенку* [«Мурзилка», 2001].

Этот эффект связан в том числе с тем, что *однажды* и *один раз* сами по себе обозначают неопределенную временную локализацию, поэтому задача уточнить эту локализацию целиком ложится на временную клаузу. Если вместо *однажды* используется конкретное обозначение времени, ограничения на употребление *как* ослабевают. Ср.:

- (46) *Ты начнешь спуск через полчаса, как я уйду к перевалу* [Сергей Наумов. На расстоянии крика (1978)].

Совместимые с временным *как* соотносительные слова *теперь*, *иногда* и *всегда* отличаются от *однажды* и *один раз* как раз тем, что не навязывают клаузе с *как* (или навязывают в меньшей степени) роли временного ориентира. *Иногда* и *всегда*, в отличие от *однажды*, вообще не указывают на однократную локализацию главной ситуации. *Теперь* указывает на совпадение ситуации с моментом речи или моментом повествования, т. е. с таким моментом, локализация которого известна и поэтому не требует уточнения в том смысле, в каком ее требуют *однажды* или *один раз*.

Наконец, похожее объяснение получает запрет на сочетание *как* с контрастивными частицами, действующий в современном языке и отсутствовавший в языке XVIII–XIX в. (см. раздел 1.2.4). Как известно, при контрасте происходит сопоставление выделенного элемента с множеством других, аналогичных элементов (см., например, [Янко 2001: 47]). В случае с временным *как* этим достигается сопоставление момента, в который произошла зависимая ситуация, с некоторыми другими моментами. Например, в (47) (=36) говорится, что *дупель вылетел, когда я решил стрелять*, а не в какой-то другой момент. В результате временная ориентация главной ситуации относительно зависимой получает подкрепление (о контрастивности *как* способе подчеркнуть значение ориентира см. также раздел 2.1).

- (47) *решил <...> стрелять <...> и как раз как решил вылетел дупель* [М. М. Пришвин. Дневники (1927)].

3. Временной *как* и типология адвербиальных клауз

Наша идея о том, что эволюция временного *как* за последние 200 лет состояла в потере способности вводить временной ориентир, пока что оторвана от общих — теоретических и типологических — представлений об адвербиальных, в частности временных, клаузах. В настоящем разделе мы стремимся восполнить этот пробел, предложив интерпретацию произошедшего сдвига в терминах общепринятой, типологически ориентированной классификации адвербиальных клауз. В разделе 3.1 описаны основные параметры такой классификации, раздел 3.2 посвящен описанию в этих терминах временного *как*.

3.1. О типологически ориентированной классификации адвербиальных клауз

В генеративной литературе принято выделять три типа адвербиальных клауз, различающихся степенью интеграции с главной клаузой: **центральные** клаузы (central adverbial clauses) — интегрированные наиболее тесно, **периферийные** клаузы (peripheral adverbial clauses), промежуточные по степени интеграции, и **неинтегрированные** клаузы (non-integrated adverbial clauses), демонстрирующие самую слабую интеграцию ([Frey 2012; 2016; Endo 2012; Badan, Naegeman 2022] и др.). Интеграция клаузы определяется позицией, занимаемой ею в синтаксической структуре: чем ниже расположена позиция присоединения клаузы в составе главной клаузы, тем выше степень интеграции. Центральные клаузы, как принято считать, присоединяются (attach) к главной клаузе на уровне проекции T(ense) P(hrase), периферийные клаузы — выше, на уровне проекции иллокутивной силы (Force), тогда как неинтегрированные клаузы не образуют с главной клаузой единой синтаксической составляющей, а соединяются с ней только семантически [Frey 2012; 2016]. Эта картина основывается на представлениях о расширенной проекции глагола (48) [Rizzi 1997: 297; Cinque 1999: 106, 124; Krifka 2013: 5].

(48) Force > Modality_{epistemic} > (Negation) > TP > (Negation) > VP

Базовое синтаксическое различие между клаузами разных типов проявляется в различии их семантических и формальных свойств. Так, только центральные клаузы сочетаются с контрастивными частицами и способны служить ответом на частный вопрос. Например, причинные союзы *оттого что* (50), (52) и *потому что* (49), (51) обладают обоими свойствами. Напротив, причинный союз *ибо* (49), (50), (53) не обладает ни одним из свойств, а значит, вводимая им клауза не может относиться к классу центральных.

- (49) «Ну почему же нет-то?» — «Потому что (^{??}ибо) унизительно, и вдруг кто-то знакомый в магазин придет» [Игорь Сахновский. Девочки с ушками (2015)].
- (50) — *Отчего погибло царство У?* — *Оттого, что (^{??}ибо) много воевало и много побеждало* [«Знание — сила», 2005].
- (51) *Я в том городке оказался именно потому, что был инженером дивизии* [А. А. Бушков. Дверь в чужую осень (сборник) (2015)].
- (52) *Мы проснулись только оттого, что они топали как слоны* [«Бельские просторы», 2013].
- (53) *А мне все-таки очень нравятся все наши преподы, (^{??}только, ^{??}именно) ибо я патриот нашей маленькой, но гордой кафедры* [Переписка в icq между agd-ardin и Колючий друг (16.01.2008)].

Заметим, что в [Belyaev 2015] оба свойства обсуждаются в других терминах — сочинения и подчинения: сочетаемость с частицами и способность быть ответом на вопрос признаются признаками семантического (противопоставленного синтаксическому) подчинения. С одной стороны, это описание по своей сути похоже на описание в терминах центральных, периферийных и неинтегрированных клауз: отличие сочинения от подчинения тоже касается степени интеграции. С другой стороны, однако, осмысление указанных свойств как различий внутри класса адвербиальных — т. е. подчиненных — клауз, как кажется, лучше вписывается в более общую картину. Так, согласно [Belyaev 2015: 41], такое известное свойство подчинительного союза, как способность занимать позицию в начале сложного предложения (см., например, [Тестелец 2001: 262]), сигнализирует о синтаксическом подчинении. Но при системном воплощении этого подхода оказывается, что некоторые базовые адвербиальные союзы в русском языке устроены противоположным образом на уровне семантики и синтаксиса. Например, союзы *хотя* и *раз*: семантически они сочинительные (ср. ^{??}*Приду, только раз приглашаешь*; ^{??}*Приду, именно хотя ты и не приглашаешь*), а синтаксически подчинительные (^{OK}*Раз приглашаешь, приду*; ^{OK}*Хотя ты и не приглашаешь, я приду*). Этот вывод противоречит интуиции: расхождение в семантических и синтаксических признаках ощущается как исключительное свойство, а не свойство частотных базовых единиц¹². Описание в терминах центральных, периферийных и неинтегрированных клауз позволяет избежать этого противоречия (см. более подробную аргументацию в [Pekelis 2022]).

К свойствам центральных клауз относится также способность союза помещаться в сфере действия оператора эпистемической модальности, входящего в состав главной клаузы [Haegeman, Schönenberg 2023: 136]: соответствующая проекция расположена выше в структуре (48), чем проекция TP, считающаяся местом присоединения центральных клауз. Так, союзы *потому что* и *оттого что* снова ведут себя одинаково в этом отношении: в (54) и (55) оба находятся в сфере действия оператора *вероятно*. В (54) предположение, выражаемое *вероятно*, касается причины того, что они неловко себя чувствовали (при расположении союза вне сферы действия *вероятно* (54) выражало бы, наоборот, причину предположения). Так же и в (55) предположение относится к причине того, что друг редко пишет.

- (54) *Они, вероятно, неловко себя чувствовали, оттого что впервые ехали вместе, да еще не одни* [Владимир Высоцкий. Где Центр? (1975)].
- (55) *Какой-то гвардии майор, осетин, пытался ухаживать за мной и старался доказать, что, вероятно, мой друг редко пишет, потому что нашел другую* [«Известия», 21.06.2002].

Обратное соотношение — расположение союза вне сферы действия оператора эпистемической модальности — свидетельствует, наоборот, о более слабой интеграции, чем та, которая свойственна центральным клаузам. Такое употребление союза, называемое эпистемическим [Sweetser 1990], считается признаком периферийных адвербиальных клауз [Freu 2016; 2023]. Ср. эпистемическое употребление союза *потому что* в (56). Заметим, что *оттого что* к такому употреблению не способен (56), т. е. в этом отношении отличается от *потому что*. Это говорит о том, что *оттого что* всегда вводит центральные клаузы, а *потому что* может вводить и периферийные. Последнее обстоятельство иллюстрирует отсутствие одно-однозначного соответствия между типами клауз и адвербиальными союзами: некоторые союзы (например, *потому что*)

¹² Например, к семантически подчинительным и синтаксически сочинительным, по данным [Culicover, Jackendoff 1997], относится конструкция с английским союзом *and* в условном значении, ср. *You drink another can of beer and I'm leaving*.

могут вводить клаузы разных типов; другие (например, *оттого что*) специализируются на каком-то одном типе клауз.

- (56) ...они, наверное, еще оба курили, потому что (^{??}*оттого что*), когда он вошел, комната была синяя от дыма [Ю. О. Домбровский. Хранитель древностей / Приложение (1964)].

К свойствам союзов, вводящих неинтегрированные клаузы, относится способность союза употребляться иллокутивно [Frey 2016; 2023], т. е. соединять пропозицию одной клаузы с иллокутивной модальностью, входящей в смысл другой [Sweetser 1990]. Так, союз *потому что* допускает иллокутивное употребление, в отличие от *оттого что* (57). Это еще раз подтверждает тесную интеграцию клауз с *оттого что* и обнаруживает способность *потому что* вводить (не только центральные и периферийные, но и) неинтегрированные клаузы.

- (57) *Запаситесь чаем с печеньками, потому что (^{??}оттого что) этот репортаж будет такой же длинный, как и небоскреб* [Ирина Плыткевич. Как строится небоскреб Лахта Центр (24.08.2017)].

Важным доводом в пользу описанной классификации адвербиальных клауз служит тот факт, что свойства, ассоциируемые с разными типами клауз, коррелируют друг с другом. Так, контрастивное выделение не характерно для союзов, употребленных эпистемически, потому что способность к контрасту служит признаком центральных клауз, а способность к эпистемическому употреблению — признаком периферийных. Например, в английском эпистемическое употребление союза *because* не совместимо с клефтовой конструкцией [Sweetser 1990: 76–77], ср. (58в)–(58г), хотя вне клефтовой конструкции встречается, ср. (58а)–(58б).

- (58) а. *John came back because he loved her.*
 б. *John loved her, because he came back.*
 в. *It was because he loved her that John came back.*
 г. [?]*It was because he came back that John loved her.*

Аналогично союз *потому что* не сочетается с контрастивными частицами, в том числе с отрицанием, при условии эпистемического употребления, ср. (59а)–(59б).

- (59) а. *Асфальт мокрый, не потому что дождь прошел, а потому что с крыши капает* (пример из [Belyaev 2015: 54]).
 б. [?]*Дождь прошел, не потому что асфальт мокрый, а потому что с крыши капает* (пример из [Ibid.: 55]).

Похожим образом коррелируют между собой способность употребляться эпистемически (признак периферийных клауз) и способность служить ответом на частный вопрос (признак центральных клауз). Так, клауза с *потому что*, употребленная эпистемически, как в (60а), менее удачна в качестве ответа на частный вопрос (60б), чем неэпистемически употребленная клауза (61). В [Заика 2019: 24] эта закономерность отмечается как типологически распространенная.

- (60) а. *Иван дома, потому что в его окне горит свет.*
 б. — *Почему Иван дома?* — [?]*Потому что в его окне горит свет.*
 (61) — *Почему Иван дома?* — *Потому что он приболел.*

3.2. Эволюция временного *как*: типологический ракурс

Напомним, что произошедший с *как* сдвиг мы предложили интерпретировать как потерю способности вводить временной ориентир. В терминах классификации адвербиальных клауз, общие принципы которой изложены в предыдущем разделе, эта интерпретация, как мы думаем, равносильна следующей: в языке XIX в. клаузы с *как* относились к классу центральных, а в современном языке степень их интеграции снизилась — сегодня *как* вводит периферийные или даже неинтегрированные клаузы¹³.

Приведем аргументы в пользу такого предположения. Во-первых, с ролью ориентира мы связываем способность временной клаузы отвечать на (явный или подразумеваемый) вопрос *когда?*. Такая способность ограничена у современного *как* (поэтому *как* не вводит клаузу-рему с локализацией в прошлом), но присутствовала у *как* в языке XIX в. (см. подробнее раздел 1.2.2). Между тем служить ответом на частный вопрос могут именно центральные клаузы (см. раздел 3.1).

Во-вторых, после XIX в. клауза с *как* утратила способность к контрастивному выделению (см. раздел 1.2.4) — еще одно свойство центральных клауз (см. раздел 3.1).

В-третьих, современный *как* не может помещаться в сфере действия оператора эпистемической модальности, что свойственно слабо интегрированным (не центральным) клаузам. Поэтому *как* неуместен вместо *когда* в (62а), где выражаемое *когда* временное отношение помещено в сферу действия оператора *наверное*. Обратим внимание, что ситуации в (62а) связаны отношением предшествования, т. е. таксисным отношением такого типа, который вообще говоря с современным *как* совместим (см. раздел 1.1). Ср. пример (62б) — с *как* и без эпистемического оператора¹⁴.

(62) а. *Но ничего, залезу, наверное, в воду, когда ("как") она уже зацветет (шутка)*
[Письмо девушки из Перми сестре в Москву (2001)].

б. ^{OK}*Залезу в воду, как зацветет.*

Наконец, в-четвертых, о слабой интеграции клауз с современным *как* свидетельствует тот факт, что *как* допускает употребление, близкое иллюкутивному [Кобозева, Инькова 2018: 234], ср. (63). Такое употребление, напомним, ассоциируется с неинтегрированными клаузами (см. подробнее раздел 3.1).

(63) *Как войдете, ейная квартира слева, прямо тут* [Дарья Донцова. Уха из золотой рыбки (2004)].

Заключение

В работе предпринята попытка проследить эволюцию временного союза *как* после XIX в. и обобщить эту эволюцию в семантических и грамматических терминах. Мы

¹³ С идеей о том, что в более ранние исторические периоды *как* мог вводить тесно интегрированные клаузы, согласуется способность *как* употребляться в клаузах с причастиями в древнерусском языке. Ср.:

...а како *доспѣво* (действ. прич.) *буду* (Берестяная грамота 68 (1240–1260)).

‘Я, как только управлюсь, приеду’ (перевод с сайта www.gramoty.ru).

¹⁴ Если наше предположение о произошедшем с *как* сдвиге справедливо, в языке XIX в. *как* должен был встречаться в контекстах типа (62а). Утверждать это уверенно мы не можем, поскольку обнаружить подтверждающие примеры в корпусе нам не удалось.

воспользовались семантическим понятием временного ориентира, предположив, что после XIX в. *как* потерял способность вводить такой ориентир. Грамматическое обобщение опиралось на классификацию адвербиальных клауз, известную из литературы по теории грамматики и типологии: в этих терминах, клаузы с *как* за последние два века трансформировались из тесно интегрированных в слабо интегрированные. Предлагаемое описание поддерживается тем, что позволяет объяснить многие свойства и современного *как*, и *как* в языке XIX в., т. е. проливает свет на эволюцию этих свойств.

Вывод о развитии *как* в сторону ослабления интеграции ставит и некоторые более общие вопросы. Временные клаузы принято относить к тесно интегрированным, в отличие, например, от причинных, уступительных или сопоставительных клауз [Haegeman 2003; Badan, Haegeman 2022]. Представленные данные о союзе *как* противоречат этой точке зрения и, возможно, требуют ее пересмотра.

Отдельный вопрос касается причин произошедшего сдвига. Материал статьи не позволяет продвинуться в поиске ответа, поскольку раскрывает не причины, а только содержание сдвига. Вопрос о причинах, в свою очередь, является частью более общего и, насколько нам известно, малоисследованного вопроса о закономерностях семантической эволюции коннекторов.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Барентсен, Пупынин 1999 — Барентсен А., Пупынин Ю. О союзах предшествования в русском и английском языках. [Barentsen A., Pupynin Yu. On conjunctions of precedence in Russian and English.] *Semantyka a konfrontacja językowa* 2. Greń Zb., Koseska-Toszeva V. (eds.). Warszawa: Slawistyczny Ośrodek Wydawniczy, 1999, 105–122.
- Виноградов, Шведова 1964 — Виноградов В. В., Шведова Н. Ю. *Очерки по исторической грамматике русского литературного языка XIX века. Изменения в строе сложноподчиненного предложения в русском литературном языке XIX века*. М.: Наука, 1964. [Vinogradov V. V., Shvedova N. Yu. *Ocherki po istoricheskoi grammatike russkogo yazyka XIX veka. Izmeneniya v stroe slozhnopodchinnenogo predlozheniya v russkom literaturnom yazyke XIX veka* [Essays on the historical grammar of the Russian literary language of the 19th century. Changes in the structure of complex sentences in the Russian literary language of the 19th century]. Moscow: Nauka, 1964.]
- Заика 2019 — Заика Н. М. Полипредикативные причинные конструкции в языках мира: пространство типологических возможностей. *Вопросы языкознания*, 2019, 4: 7–32. [Zaika N. M. Polypredicative reason constructions in the world's languages: Typological parameters. *Voprosy Jazykoznanija*, 2019, 4: 7–32.]
- Кобозева, Инькова 2018 — Кобозева И. М., Инькова О. Ю. *Как и его двухместные варианты. Семантика коннекторов: контрастивное исследование*. Инькова О. Ю. (ред.). М.: ТОРУС ПРЕСС, 2018, 168–239. [Kobozeva I. M., Inkova O. Yu. *Kak and its two-place variants. Semantika konnektorov: kontrastivnoe issledovanie*. Inkova O. Yu. (ed.). Moscow: TORUS PRESS, 2018, 168–239.]
- Лавров 1941 — Лавров Б. В. *Условные и уступительные предложения в древнерусском языке*. М.; Л.: Изд-во Академии наук СССР, 1941. [Lavrov B. V. *Uslovnye i ustupitelnye predlozheniya v drevnerusskom yazyke* [Conditionals and concessive clauses in Old Russian]. Moscow; Leningrad: Academy of Sciences of the USSR Press, 1941.]
- НКРЯ — Национальный корпус русского языка [Russian National Corpus]. <http://www.ruscorpora.ru>.
- Пекелис 2022 — Пекелис О. Е. Русские временные клаузы на шкале семантико-синтаксической интеграции (на примере сочинительного союза *когда*). *Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной конференции «Диалог»*, 2022, 21: 420–435. [Pekelis O. E. Russian temporal clauses in terms of semantic-syntactic integration (a case study of a coordinating conjunction *kogda* ‘when’). *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conf. “Dialogue”*, 2022, 21: 420–435.]
- РГ 1980 — Шведова Н. Ю. (гл. ред.). *Русская грамматика*. Т. 2: *Синтаксис*. В 2 т. М.: Наука, 1980. [Shvedova N. Yu. (ed.). *Russkaya grammatika* [Russian grammar]. Vol. 2: *Sintaksis* [Syntax]. In 2 vols. Moscow: Nauka, 1980.]

- Тестелец 2001 — Тестелец Я. Г. *Введение в общий синтаксис*. М.: РГГУ, 2001. [Testelecs Ya. G. *Vvedenie v obshchii sintaksis* [Introduction to general syntax]. Moscow: Russian State Univ. for the Humanities, 2001.]
- Храковский 2005 — Храковский В. С. Таксис следования в современном русском языке. *Проблемы функциональной грамматики. Полевые структуры*. Бондарко А. В. (ред.). СПб.: Наука, 2005, 29–85. [Xrakovskij V. S. Taxis of following in modern Russian. *Problemy funktsionalnoi grammatiki. Polevye struktury*. Bondarko A. V. (ed.). St. Petersburg: Nauka, 2005, 29–85.]
- Храковский 2009 — Храковский В. С. Таксис: семантика, синтаксис, типология. *Типология таксисных конструкций*. Храковский В. С. (ред.). М.: Знак, 2009, 11–113. [Xrakovskij V. S. Taxis: Semantics, syntax, and typology. *Tipologiya taksisnykh konstrukcii*. Xrakovskij V. S. (ed.). Moscow: Znak, 2009, 11–113.]
- Янко 2001 — Янко Т. Е. *Коммуникативные стратегии русской речи*. М.: Языки славянской культуры, 2001. [Yanko T. E. *Kommunikativnye strategii russkoi rechi* [Communicative strategies of Russian speech]. Moscow: Yazyki slavyanskoi kul'tury, 2001.]
- Badan, Haegeman 2022 — Badan L., Haegeman L. The syntax of peripheral adverbial clauses. *Journal of Linguistics*, 2022, 58(4): 697–738. doi:10.1017/S002226721000463.
- Belyaev 2015 — Belyaev O. Cause in Russian and the formal typology of coordination and subordination. *Donum semanticum: Opera linguistica et logica in honorem Barbarae Partee a discipulis amicisque Rossicis oblata*. Arkadiev P., Kapitonov I., Lander Yu., Rakhilina E., Tatevosov S. (eds.). Moscow: Yazyki slavyanskoi kul'tury, 2015, 35–66.
- Cinque 1999 — Cinque G. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford Univ. Press, 1999.
- Culicover, Jackendoff 1997 — Culicover P. W., Jackendoff R. Semantic subordination despite syntactic coordination. *Linguistic Inquiry*, 1997, 28: 195–218.
- Endo 2012 — Endo Y. The syntax discourse interface in adverbial clauses. *Main clause phenomena: New horizons*. Lobke A., Haegeman L., Nye R. (eds.). Amsterdam: John Benjamins, 2012, 365–384.
- Frey 2012 — Frey W. On two types of adverbial clauses allowing root-phenomena. *Main clause phenomena: New horizons*. Lobke A., Haegeman L., Nye R. (eds.). Amsterdam: John Benjamins, 2012, 405–429.
- Frey 2016 — Frey W. On some correlations between formal and interpretative properties of causal clauses. *Co- and subordination in German and other languages: Special issue of Linguistische Berichte 21*. Reich I., Speyer A. (eds.). Hamburg: Buske, 2016, 153–179.
- Frey 2023 — Frey W. On the categorial status of different dependent clauses. *Propositionale Argumente im Sprachvergleich / Propositional Arguments in Cross-Linguistic Research: Theorie und Empirie / Theoretical and Empirical Issues*. Hartmann J. M., Wöllstein A. (eds.). Tübingen: Narr, 2023, 364–410.
- Haegeman 2003 — Haegeman L. The syntax of adverbial clauses and its consequences for topicalization. *Current studies in comparative Romance linguistics*. Coene M., De Cuyper G., D'Hulst Y. (eds.). Antwerp: Univ. of Antwerp, 2003, 61–90.
- Haegeman, Schönenberger 2023 — Haegeman L., Schönenberger M. English rationale *since* and a reassessment of the typology of adverbial clauses. *Micro- and macro-variation of causal clauses: Synchronic and diachronic insights*. Jędrzejowski Ł., Fleczonek C. (eds.). Amsterdam: John Benjamins, 2023, 129–166.
- Krifka 2013 — Krifka M. Response particles as propositional anaphors. *Proceedings of SALT*, 2013, 23: 1–18. <https://doi.org/10.3765/salt.v23i0.2676>.
- Pekelis 2022 — Pekelis O. On the oppositions that underlie the distinctions displayed by Russian causal clauses. *Linguistics*, 2022, 5(60): 1399–1449. <https://doi.org/10.1515/ling-2020-0151>.
- Rizzi 1997 — Rizzi L. The fine structure of the left periphery. *Elements of grammar: A handbook of generative syntax*. Haegeman L. (ed.). Dordrecht: Kluwer, 1997, 281–337.
- Sweetser 1990 — Sweetser E. *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge Univ. Press, 1990. <https://doi.org/10.1017/CBO9780511620904>.

Контакты родственных языков как фактор формирования энецких идиомов

© 2024

Анна Юрьевна Урманчиева

Институт лингвистических исследований РАН, Санкт-Петербург, Россия;
urmanna@yandex.ru

Аннотация: В статье реконструируется история формирования двух близкородственных идиомов: лесного и тундрового энецкого, относящихся вместе с распространенными в том же регионе нганасанским и лесным и тундровым ненецким к северной подгруппе самодийских языков. Приводятся аргументы в пользу того, что общененецкий язык мог сформироваться в процессе перехода с языка, близкого к нганасанскому, на язык, близкий к (тундровому) ненецкому. При этом энецкий — вероятно, в силу географически центрального положения в группе северно-самодийских языков Приполярья и малочисленности носителей — на протяжении всей своей истории испытывал влияние нганасанского и ненецкого. Поэтому первичные энецко-нганасанские изоглоссы субстратного характера важно отделять от вторичных контактных изоглосс, объединяющих с нганасанским общененецкий и в более позднюю эпоху — тундровый энецкий. В статье будут рассмотрены нганасанско-энецкие изоглоссы (фонологические, морфонологические, морфологические) и предложена их стратификация на первичные субстратные и вторичные контактные.

Ключевые слова: нганасанский язык, ненецкий язык, самодийские языки, субстрат, энецкий язык, этническая история, языковой сдвиг, языковые контакты

Для цитирования: Урманчиева А. Ю. Контакты родственных языков как фактор формирования энецких идиомов. *Вопросы языкознания*, 2024, 5: 25–59.

DOI: 10.31857/0373-658X.2024.5.25-59

The role of contacts within genetically related languages in the formation of Enets idioms

Anna Yu. Urmanchieva

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

Abstract: The article examines the history of two closely related idioms: Forest and Tundra Enets, which, together with Nganasan and Forest and Tundra Nenets, spoken in the same region, belong to the Northern subgroup of Samoyedic languages. Arguments are given in favor of the hypothesis that the Protoenets language could have been formed in the process of transition from a language close to Nganasan to a language close to (Tundra) Nenets. At the same time, Enets — due to its geographically central position in the group of Northern Samoyedic languages of the Circumpolar region and the small number of speakers — has been influenced by Nganasan and Nenets throughout its history. Therefore, it is important to differentiate between the primary Enets-Nganasan isoglosses of a substrate nature and the secondary contact isoglosses, which unite Protoenets with Nganasan and, in a later period, Tundra Enets. The article examines Nganasan-Enets isoglosses (phonological, morphophonological, morphological) and proposes their stratification.

Keywords: Enets, ethnic history, language contacts, language shift, Nenets, Nganasan, Samoyedic, substrate

For citation: Urmanchieva A. Yu. The role of contacts within genetically related languages in the formation of Enets idioms. *Voprosy Jazykoznanija*, 2024, 5: 25–59.

DOI: 10.31857/0373-658X.2024.5.25-59

1. Введение

В работе рассматривается история формирования двух близкородственных северносамодийских идиомов: лесного и тундрового энецкого (в том случае, если надо обозначить оба идиома, также используется термин «энецкий»). Северная подгруппа самодийских языков состоит из занимающего обособленное положение нганасанского и объединяющихся в промежуточную подгруппу лесного и тундрового ненецкого, лесного и тундрового энецкого и маторского, см. [Helinski 1997/2022]. Не углубляясь в дискуссии о соотношении языков внутри ненецко-энецко-маторской подгруппы, подчеркнем наиболее важный для данной работы факт (который отражается и в других классификациях, например, предложенных в [Katz 1987; Janhunen 1998]): нганасанский в лингвистическом отношении находится дальше от энецкого, чем ненецкий. В данной работе приводятся аргументы в пользу того, что история формирования энецкого могла быть сложнее, чем дивергенция общего ненецко-энецко(-маторского) праязыка. Энецкий язык мог сформироваться в процессе перехода с языка, близкого к нганасанскому, на язык, близкий к (тундровому) ненецкому: в ряде случаев удается показать, что более раннее реконструируемое состояние энецкого демонстрирует большее сходство с нганасанским.

Вместе с тем в географическом отношении энецкий занимает (и занимал в исторически обозримое время) промежуточное положение между ненецкими идиомами и нганасанским и развивался в постоянном контакте с ними. Поэтому на картину первичных энецко-нганасанских изоглосс субстратного характера (а также на массив ненецко-энецких адстратных изоглосс, которые и обуславливают отнесение энецкого к одной подгруппе с ненецким) накладываются вторичные контактные изоглоссы, среди которых — общеэнецко-нганасанские, а также более поздние, объединяющие тундровый энецкий с нганасанским и лесной энецкий — с тундровым ненецким. Мы сосредоточимся на нганасанско-энецких изоглоссах (фонологических, морфонологических, морфологических) и их стратификации на первичные субстратные и вторичные контактные.

2. Историческая фонетика: влияние нганасанского субстрата

2.1. Сохранение дифтонгов

Влиянием нганасанского субстрата можно объяснить рефлексацию в энецком прасамодийских вокалических последовательностей вида $*V\bar{a} \sim *V\bar{\bar{a}}$ в первом слоге. Из всех самодийских языков (не только северносамодийских) такие вокалические последовательности в виде дифтонгов сохраняются в двух языках — в энецком и нганасанском. В ненецком эти вокалические последовательности упростились, другие самодийские языки, в том числе селькупский и камасинский, также указывают на их историческую неустойчивость, ср.:

$*k\bar{a}\bar{a}$ – ‘умереть’ > эн. *kaa-*, нг. *ku̯a-*, но т. нен. *xa-*, ср. также ск. *k̄i-*, км. *k'u-¹*;

¹ Прасамодийские реконструкции приводятся по [Janhunen 1977] с учетом изменений в реконструкции прасамодийского вокализма, предложенных в [Helinski 2005]. Материалы по отдельным самодийским языкам приводятся по следующим источникам: для нганасанского и энецкого

- **roâj* ‘год’ > эн. *ria*, нг. *xuo*, но т. нен. *po*, ср. также ск. *po*, км. *p’je*;
 **woâj* ‘остров’ > эн. *huo*, нг. *nyai*, но т. нен. *jo*, ср. также ск. *ko*;
 **piâ* ‘позади’ > эн. *pio*, нг. *xuo*, но т. нен. *pî*, ср. также ск. *pî*, км. *p’u-*, *pî-*;
 **tîâ* ‘облако’ > эн. *t’io-ri*, нг. *chi-py*, *chi-zu*, но т. нен. *tyi-r°*, ср. также ск. *tî*, км. *tî*;
 **teâ* ‘домашний олень’ > эн. *tia*, нг. *maa*, но т. нен. *tî*, ср. также км. *t’o*;
 **tîâj-* ‘рукав’ > эн. *tio-jo*, нг. *chi-de*, но т. нен. *tyi*, ср. также ск. *tî-naŋ*, км. *ti*.

Модель дивергенции промежуточных праязыков предполагала бы необходимость реконструировать вокалические последовательности для ненецко-энецкого языка с последующим устранением их в ненецком и сохранением — в энецком². С нашей точки зрения (которую мы надеемся подтвердить совокупностью рассмотренных в статье фактов), логичнее трактовать сохранение таких последовательностей в энецком как черту, унаследованную этим языком от субстрата нганасанского типа.

2.2. Различение рефлексов **a* в зависимости от сингармонического ряда слова

Существенное уточнение в правила рефлексии двух прасамодийских гласных — **ä* и **a* (**e* и **ä* в реконструкции [Janhunen 1977]) были внесены в работе [Гусев 2022], где показано, что каждый из этих гласных может иметь в ряде самодийских языков (нганасанском, селькупском, камасинском и, возможно, маторском) по два рефлекса — в зависимости от сингармонического ряда слова, который можно определить по сингармонической характеристике корня в нганасанском³. Поскольку раздвоение рефлексов **a* и **ä* представлено в языках двух ветвей — в нганасанском и южносамодийских — это явление в нганасанском следует рассматривать как архаизм, а его отсутствие в ненецком и энецком — как общую инновацию⁴.

Однако в лесном энецком в трех позициях противопоставление рефлексов **a* (U) и **a* (I) фонологизировано.

Во-первых, это анлаутная позиция:

- (1) а. переднерядные рефлексы
 **ar-* ‘ходить за водой’ > л. эн. *erir-*, ср. нг. *нялы*²- (I);
 **atз-* ‘верх’ > л. эн. *edo*, ср. нг. *нячи-* (I) ‘выше по реке’;
 **al-tä-* ‘радоваться’ > л. эн. *ede-*, ср. нг. *нялты-* (I);

использовались неопубликованные словари, собранные Е. А. Хелимским (для нганасанского также дополнительно словарь [Костеркина и др. 2001]), для тундрового ненецкого — [Salminen 1998], для маторского — [Helinski 1997], для селькупского — [Alatalo 2004], для камасинского — [Donner, Joki 1944] (иногда с дополнениями по [Castrén 1855]).

² В ненецком, по-видимому, сохранялись в определенный период только вокалические последовательности с первым гласным верхнего подъема, о чем свидетельствует долгота гласных в случаях типа т. нен. *tyi-r°* ‘облако’, *pî* ‘позади’, *tyi* ‘рукав’ (отметим, что в ненецком в данном случае имеет место монофтонгизация, нехарактерная для нганасанского и энецкого, сохраняющих разное качество первого и второго гласного в вокалических последовательностях), следов сохранения иных вокалических последовательностей в ненецком нет.

³ Поэтому для обозначения сингармонического ряда слова мы употребляем обозначения, принятые для нганасанского: (I) — переднерядные основы, (U) — заднерядные основы.

⁴ Вероятно, это явление существовало в статусе аллофонического варьирования, и в языках, утративших сингармонизм, произошло выравнивание, а в языках, сингармонизм сохранивших, — фонологизация этого противопоставления.

b. заднерядные рефлексы

**amtā* ‘рог’ > л. эн. *nado* (при т. эн. *edo*), ср. нг. *уамтэ* (U);

**aṇtā* ‘лезвие’ > л. эн. *nado* (при т. эн. *edo*), ср. нг. *уаче* (U);

**aṇ* ‘рог’ > л. эн. *naʔ* (при т. эн. *eʔ*), ср. нг. *уау* (U);

**antā-mā* ‘взмолиться’ > л. эн. *nadoo-* (при т. эн. *edoo-*), ср. нг. *уантэмэ-* (U).

Во-вторых, это позиция после анлаутного *j:

(2) a. переднерядные рефлексы

**jalä* ‘день’ > л. эн. *d'ere*, ср. нг. *дялы* (I);

b. заднерядные рефлексы

**jaʔke* ‘дым’ > л. эн. *d'aki* (так же т. эн.), ср. нг. *дякагэ* (U);

**jaṇkājw-* ‘не иметься’ > л. эн. *d'agi-* ‘не иметься’ (т. эн. *d'igu-*), ср. нг. *дяугуй-* (U);

**jaʔ(ā)* ‘выковать’ > л. эн. *d'ado-* (т. эн. *d'edo-*), ср. нг. *дяʔ-* (U).

В-третьих, это позиция после анлаутного *ñ. Для этой позиции не найдено примеров слов, относящихся к переднерядному гармоническому классу⁵, но это и не так важно: для нас наиболее существенно показать, что в трех рассматриваемых позициях рефлексация *a в заднерядных словах в лесном энецком отличается от стандартной (которой является переход *a > эн. e):

(3) заднерядные рефлексы:

**ñaj-* ‘толстый’ > л. эн. *naus-* ‘быть толстым’ (в т. эн. когнатного слова нет), ср. нг. *нячегээ* (U);

**ñajä-* ‘голый; лысый’ > л. эн. *nae-xa* (т. эн. *eo-do*), нг. *няадэ* (U);

**ñajə-rä* ‘мездра’ > л. эн. *naara* (т. эн. *eara*), ср. нг. *няэру* (U).

Если оставаться в рамках стандартной модели дивергенции языков северносамодийской подгруппы, факты лесного энецкого должны были бы заставить нас реконструировать гармонию гласных, влиявшую на рефлексацию *a, даже для периода после диалектной дифференциации протоэнецкого.

Однако одновременно мы наблюдаем в обоих энецких идиомах переход *a в передний гласный (вне зависимости от сингармонической характеристики корня) — переход, характерный для ненецкого, но не для нганасанского. В частности, в энецком, как и в ненецком, перед *a в словах **любого сингармонического ряда** происходит переход *k > s, см. о нем в следующем подразделе.

В итоге мы должны признать, что в лесном энецком сосуществуют, с одной стороны, противопоставление рефлексов *a (U) и *a (I) как заднерядного и переднерядного гласных (в трех позициях), с другой стороны — переход *a в переднерядный гласный вне зависимости от сингармонической характеристики корня. Первый переход при этом предполагает сохранение гармонии гласных вплоть до этапа независимого существования лесного и тундрового энецкого, второй — устранение гармонии гласных и продвижение **любого** *a вперед еще на этапе гипотетического общененецкоэнецкого. Иными словами, мы наблюдаем в энецком два исторических несовместимых друг с другом явления.

Эту ситуацию можно объяснить, если исходить из нашего предположения о том, что энецкий сформировался в результате перехода с языка типа нганасанского на язык типа ненецкого: в энецком на периферии сохраняются некоторые субстратные явления, генетически общие с наблюдаемыми в нганасанском, и одновременно представлены продуктивные явления, объединяющие его с ненецким.

⁵ Впрочем, в отсутствие южносамодийских соответствий мы не можем определить л. эн. *edo* ‘верх’, ср. нг. *нячи-* (I) ‘выше по реке’, и л. эн. *ede-* ‘радоваться’, ср. нг. *няты-* (I) ‘радоваться’, как слова, представляющие вокалический анлаут или анлаутный *ñ-.

2.3. Трактовка **ü* как заднерядного гласного и развитие **k*

В нганасанском — единственном из всех самодийских — гласный **ü* ведет себя как заднерядный (при сохранении переднерядной гармонической характеристики корня). В частности, анлаутный **k* перед **ü* ведет себя так же, как в заднерядных словах, то есть не переходит в сибилант, ср. **küðŋ* (**küŋ* в [Janhunen 1977: 79]) ‘пуп’ > т. нен. *syúh*, т. эн. *s’uʔ*, ск. *śōñ*, км. *šāŋ*, *šāñ*, но нг. *киш*.

Условия перехода анлаутного **k* в сибилант в энецком и ненецком совпадают, но отличаются от представленных в нганасанском, где этот переход не только не происходит перед **ü*, но и чувствителен к сингармоническому ряду слов с **ä* и **a*, см. таблицу 1.

Таблица 1

Переход **k* > *s/š* перед передними гласными в самодийских языках

Гласный корня	Ненецкий	Энецкий	Нганасанский	Селькупский	Камасинский	Маторский
* <i>a</i> (U)	<i>sya-</i> <i>syal°h</i> ‘подмышка’ [Janhunen 1977: 63]	<i>se-</i> <i>seriʔ</i> ‘подмышка’ [Janhunen 1977: 63]	<i>ka-</i> <i>кали</i> ^C ‘подмышка’ [Janhunen 1977: 63]	<i>kā-</i> <i>ḱāliñ</i> ‘подмышка’ [Janhunen 1977: 63]	<i>ka-</i> <i>ḱ’ālêŋ</i> ‘подмышка’ [Janhunen 1977: 63]	<i>kū-</i> (ср. <i>kädər-</i> ‘чихать’ [Janhunen 1977: 63])
* <i>a</i> (I)	<i>sya-</i> ⁶	<i>se-</i>	<i>ša-</i> <i>сярымсы</i> ⁷ ‘устыдиться’ [Janhunen 1977: 68]	<i>šo-</i> <i>šoräši-</i> ‘стыдиться’ [Janhunen 1977: 68]	<i>še-</i> <i>šēräl’em</i> ‘стыдиться’? ср. кой- бальск. <i>сарлянь</i> [Janhunen 1977: 68]	нет данных
* <i>ä</i> (U)	<i>syē-</i> <i>syēsər/cy°</i> ‘шуршать’ [Гусев 2022: 47]	<i>se-</i> <i>sesana-</i> ‘шуршать’ [Гусев 2022: 47]	<i>ka-</i> <i>касырса</i> ‘шуршать’ [Гусев 2022: 47]	<i>kā</i> <i>ḱātaŋ</i> ‘иней’ [Alatalo 2004: 273]	<i>ka-</i>	нет данных
	<i>syedyah</i> ‘иней’ [Alatalo 2004: 273]				<i>ḱadaŋ</i> ‘первый снег’ [Alatalo 2004: 273]	
	<i>syempye(sy°)</i> ‘держаться на поверхно- сти, не про- валиваясь’, <i>syempyad°</i> ‘поплавок’ [Alatalo 2004: 1836]	т. эн. <i>sebetuʔ</i> ~ <i>sebotuʔ</i> , л. эн. <i>seboruʔ</i> ‘поплавок’		<i>ḱāmpə-</i> ‘пла- вать на по- верхно- сти’ [Alatalo 2004: 1836]		

⁶ Для энецкого и ненецкого нет примеров, позволяющих по нганасанскому соответствию определить слово с **a* как относящееся к переднему ряду, но нет оснований реконструировать иные рефлексy.

⁷ В [Janhunen 1977: 68] восстанавливается **ker-*, но нганасанский (и койбальский, см. в ячейке камасинского материала в данной строке таблицы) указывают на **kär-*.

Гласный корня	Ненецкий	Энецкий	Нганасанский	Селькупский	Камасинский	Маторский
*ä (I)	<i>syē-</i> <i>syerə</i> (sy°) ‘надеть’ [Janhunen 1977: 68]	<i>se-</i> <i>sero-</i> ‘надеть’ [Janhunen 1977: 68]	<i>še-</i> <i>sepaḍi</i> ‘надеть’ [Janhunen 1977: 68]	<i>še-</i> <i>šēr-</i> ‘надеть’ [Janhunen 1977: 68]	<i>še-</i> <i>šērəl’em</i> ‘надеть’ [Janhunen 1977: 68]	<i>ke-</i> <i>ker-</i> ‘надеть’ [Janhunen 1977: 68]
*e	<i>syi-</i> <i>syiq</i> ‘подобие’ [Janhunen 1977: 62; Helimski 2005: 33]	<i>s’i-</i> <i>s’idoḡo</i> ‘тень живого существа’ [Janhunen 1977: 62; Helimski 2005: 33]	<i>si-</i> <i>syḷ, syḷaḡkə</i> ‘тень живого существа’ [Janhunen 1977: 62; Helimski 2005: 33]	<i>ši</i> <i>ši-</i> основа аккузатива личных местоимений [Janhunen 1977: 62; Helimski 2005: 33]	<i>ši-</i> <i>šimḍa</i> ‘кто?’ < *ke- [Helimski 2005: 30]	<i>ki-</i> <i>kim</i> (? kim) ‘кто?’ < *ke- [Helimski 2005: 30]
*i	<i>syi-</i> <i>syidya</i> ‘два’ [Janhunen 1977: 71]	<i>si-</i> <i>s’ide</i> ‘два’ [Janhunen 1977: 71]	<i>ši-</i> <i>cumi</i> ‘два’ [Janhunen 1977: 71]	<i>ši-</i> <i>šitə</i> ‘два’ [Janhunen 1977: 71]	<i>ši-</i> <i>šide</i> ‘два’ [Janhunen 1977: 71]	<i>ki-</i> <i>kide</i> ‘два’ [Janhunen 1977: 71]
*ü	<i>syu-</i> <i>syüh</i> ‘пуп’ [Janhunen 1977: 79] ⁸	<i>šu-</i> <i>s’u?</i> ‘пуп’ [Janhunen 1977: 79]	<i>ki-</i> <i>kuḷu</i> ‘пуп’ [Janhunen 1977: 79]	<i>šü-, šö-</i> <i>šöñ</i> ‘пуп’ [Janhunen 1977: 79]	<i>šü-, šV-</i> <i>šāñ, šāñ</i> ‘пуп’ [Janhunen 1977: 79]	<i>kü-</i> <i>küj</i> ‘пуп’ [Janhunen 1977: 79]

Однако непродуктивные диминутивные суффиксы показывают в энецком и ненецком другие условия перехода *k в сибилант. В энецком представлен непродуктивный диминутивный суффикс, имеющий два гармонических варианта:

эн. -*igu* < ПСС **-jeṅko* (U) ~ эн. -*id’u* < **-jensö* < ПСС **-jeṅkō* (I)

Для ненецкого также можно рассмотреть непродуктивный диминутивный суффикс:

т. нен. -*yeko* < ПСС **-jeṭ’ko* (U) ~ -*yeco* < **-jeṭ’sö* < ПСС **-jeṭ’kō* (I).⁹

В таблицах 2 и 3 приводится распределение двух алломорфов диминутивных суффиксов в энецком и ненецком в зависимости от корневого гласного (в том случае если когнатное слово имеется в нганасанском и известна его сингармоническая характеристика, при слове указывается сингармонический ряд).

⁸ В [Janhunen 1977: 79] дана праформа **küj*, современные записи позволяют уточнить праформу в виде **küñ*.

⁹ В нганасанском перехода *k > s в суффиксах нет. Соответствующие диминутивные суффиксы всегда выступают в нганасанском в форме -*AṬKU*₀ либо -*AḌKU*₀. При этом *k сохраняется в нганасанском также и в тех переднерядных алломорфах суффиксов, гласный которых в ином случае — в анлауте слова — вызывает переход *k > нг. s, ср. *сярү-кяли* (I) [стыд-сарг] ‘бесстыжий’ < **karz-kali* (I). Поскольку в одном и том же фонетическом контексте в анлауте корня переход *ka (I) > нг. s’a происходит, но в суффиксе перехода нет, можно предположить, что в нганасанском произошло вторичное выравнивание в суффиксе для устранения слишком значительного различия между алломорфами (заглавными буквами здесь и далее обозначается морфонологическая запись нганасанских морфем).

Таблица 2

**Корневой гласный и гармонический ряд слова с диминутивными суффиксами
-igu ~ -id'u в тундровом энецком**

Гласный корня	Суффикс т. эн. -igu	Суффикс т. эн. -id'u
*ā	<i>patuδ-igu</i> (U) 'рыхлый снежок'	
*ǣ	т. эн. <i>kol-igu</i> , л. эн. <i>kor-igu</i> (U) 'ящичек'	
*ǧ		<i>et'-id'u</i> 'молодой' (соотв. т. нен. <i>ηасу</i> ¹⁰)
*a (U)	<i>meδ-igu</i> (U) 'чумик'	
*a (I)		нет примеров
*ä (U)	<i>enet'e-igu</i> (U) 'человечек'	
*ä (I)		<i>d'eddos'i-id'u</i> (I) 'Енисейчик'
*e (U)	<i>usu-igu</i> (U) 'суставчик'	
*e (I)		<i>sir-id'u</i> 'солица' ¹¹
*o	<i>poδer-igu</i> (U) 'лямочка'	
*u	<i>tuδ-igu</i> (U) 'сальце'	
*ü	<i>s'u-igu</i> (I) 'пупочек'; <i>d'ur-igu</i> (I) 'сотенка'	
*j		нет примеров
*i		<i>ñi-id'u</i> (I) 'имечко'
*e		ср. <i>bid-it'u</i> 'водичка' ¹²

Для ненецкого также рассмотрим два варианта диминутивного суффикса *-yeke* ~ *-ye-су*, дополнительно внеся в таблицу 3 также информацию о других деривационных суффиксах с подобным алломорфическим варьированием.

Таблица 3

**Корневой гласный и гармонический ряд слова с суффиксами с *k или *Ck,
имеющими вариант с т. нен. k либо c**

Гласный корня	Суффикс с т. нен. k	Суффикс с т. нен. c
*ā	<i>xad-yeke</i> (U) 'ель'	
*ǣ		нет примеров
*ǧ		<i>xæbyi-су</i> ¹⁰ (I) 'короткий'
*a (U)	<i>myad-yeke</i> (U) 'чумик'	
*a (I)	<i>yal-yeke</i> (I) 'серый'	
*ä (U)	<i>nyenesyi-yeke</i> (U) 'человечек'	

¹⁰ На этимологический передний ряд указывает чередование *yø* // *ye* в форме АСС.РЛ в ненецком.

¹¹ Значения 'лед', 'соль' и 'белый' передаются в северносамодийских языках одним корнем. В переднерядном варианте — наряду с более распространенным заднерядным — оно встретилось в нганасанских текстах и словарных материалах в формах *сырты* 'его соль', *сырты* 'его лед', *сыриалыдыа* 'белесый', *сыриайси* 'белый', *сырлиа* 'ледяной', букв. 'только лед'.

¹² В этом слове представлен переднерядный вариант *-it'u* другого диминутивного суффикса, эн. *-iku* (U) ~ *-it'u* (I), но в энецком слов с ним почти нет.

Гласный корня	Суффикс с т. нен. <i>k</i>	Суффикс с т. нен. <i>c</i>
<i>*ä</i> (I)		<i>nyer^o-суи-na</i> (I) ‘раньше’
<i>*e</i> (U)	<i>wen-yeko</i> (U) ‘собака’	
<i>*e</i> (I)		<i>ñeda-ncyo-</i> (I) ‘защищаться’
<i>*o</i>	<i>xor-yeko</i> (U) ‘прирученный молодой олень’	
<i>*u</i>	нет примеров	
<i>*ü</i>		<i>tyur^o-суиq-</i> ‘иметь что-л. опорой, использовать что-л. в качестве опоры’; <i>nyüdye-суиq-</i> ‘тащить, волочить за собой’
<i>*j</i>	<i>min-yeko</i> ‘герой, от лица которого ведется ярабц’	
<i>*i</i>		<i>yüid-yecyo</i> (I) ‘десяток’
<i>*e</i>		<i>yüid-yecyo</i> (I) ‘водичка’

Можно видеть, что выбор одного из вариантов суффикса (с велярным или с сибилантом) и в ненецком, и в энецком скоррелирован с сингармоническим классом слова в прасамодийском: слова с **e*, **a*, **ä* и **ə* (**ə* (U) ~ **ä* (I)) сохраняют в этом случае различие двух сингармонических классов. В энецком, однако, особым образом ведет себя гласный **ü*: хотя это, несомненно, переднерядный гласный, при основах с корневым **ü* выбирается заднерядный вариант суффикса: эн. *s'u-igu* (I) ‘пупочек’; *d'ur-igu* (I) ‘сотенка’. В ненецком в словах с **ü* употребляется переднерядный вариант суффикса, ср. имперфективные глаголы *tyur^o-суиq-* ‘иметь что-л. опорой, использовать что-л. в качестве опоры’; *nyüdye-суиq-* ‘тащить, волочить за собой’.

Диминутивный энецкий суффикс *-igu* ~ *-id'u* не имеет параллелей в ненецком, но находит этимологическое соответствие в нганасанском *-AḶKU_o*. Логично предположить, что энецкий суффикс в данном случае сохраняет архаичное распределение алломорфов субстратного нганасанского суффикса, впоследствии утратившего алломорфическое варьирование этого типа в нганасанском.

Итак, история развития **k* также позволяет выделить в энецком два хронологических пласта. К более инновационному относится изоглосса, объединяющая ненецкий с энецким: переход анлаутного **k* в сибилант, зависящий от качества гласного, но не от сингармонической характеристики корня. Исследование поведения **k* в суффиксах в энецком и ненецком позволяет выявить более архаичное состояние (периода сохранения гармонии гласных), в котором по одному из параметров энецкий демонстрирует сходство уже с нганасанским: это сохранение велярного в словах с **ü* (I) в диминутивных суффиксах.

3. Историческая фонетика: энецкие инновации, возникшие под вторичным влиянием нганасанского

Мы рассмотрим ряд собственно энецких инноваций в сфере исторической фонетики и постараемся показать, что причиной их зарождения могли стать вторичные контакты с нганасанским языком. Энецкие системные инновации касаются прежде всего консонантизма.

3.1. Судьба интервокальных кластеров

Интервокальная позиция обнаруживает наибольшее своеобразие энецкого в сравнении с ненецким и нганасанским: в энецком устранены все интервокальные кластеры, ср., например: **kâmpa* > нг. *ко́нху*, т. нен. *хатра*, но т. эн. *kaba* ‘волна’.

Развитие кластеров в трех северносамодийских языках представлено в таблице 4 (для нганасанского также приводятся результаты чередования по ритмической и слоговой градации, о которых подробнее см. после таблицы):

Таблица 4

Интервокальные кластеры в северносамодийских языках

ПСС	Тундровый ненецкий	Энецкий		Нганасанский		
		т. эн.	л. эн.	РГ, сильная ступень	РГ, слабая ступень	СГ, слабая ступень
кластеры, принимающие участие в чередовании ступеней						
<i>*mp</i>	<i>mb</i>	<i>b</i>		<i>ɲh</i> (< <i>mp</i>)	<i>^ch</i> (< <i>p</i>)	<i>mb</i>
<i>*nt</i>	<i>n</i>	<i>d</i> / <i>dd</i> ¹³		<i>nt</i>	<i>^ct</i>	<i>nd</i>
<i>*ɲk</i>	<i>ɲg</i>	<i>g</i> / <i>gg</i>		<i>ɲk</i>	<i>^ck</i>	<i>ɲg</i>
<i>*ns</i>	<i>nc</i>	<i>d'</i>	<i>s</i>	<i>ns</i>	<i>^cs</i>	<i>nd'</i>
кластеры, не принимающие участия в чередовании ступеней						
<i>*lt</i>	<i>l</i>	<i>d</i> / <i>dd</i>		<i>lt</i>		
<i>*mt</i>	<i>md</i>	<i>d</i> / <i>dd</i>		<i>mt</i>		
<i>*ɲt</i>	<i>nd</i>	<i>d</i> / <i>dd</i>		<i>jt'</i>		
<i>*rk</i>	<i>rg</i>	<i>g</i> / <i>gg</i>		<i>rk</i>		
<i>*rs</i>	<i>rc</i>	<i>d'</i>	<i>s</i>	<i>rs</i>		
<i>*rm</i>	<i>rm</i>	<i>m</i>		<i>rm</i>		
<i>*rw</i>	<i>rw</i>	<i>m</i>		<i>rb</i>		
<i>*ps</i>	<i>bs</i>	<i>t'</i>		<i>s</i>		
<i>*pt</i>	<i>bt</i>	<i>t</i>		<i>pt</i>		

Вероятно, упрощение кластеров произошло в энецком относительно недавно: и Г. Ф. Миллер, и М. А. Кастрен последовательно записывают удвоенные согласные на месте ряда кластеров, ср. **ântaj* ‘лодка’ > *oddu* (Г. Ф. Миллер), *óddu* (М. А. Кастрен); **ârka* ‘большой’ > *agga* (Г. Ф. Миллер, М. А. Кастрен). Я благодарю рецензента за указание на то, что двойные согласные на месте кластеров в тундровом энецком фиксировал еще К. И. Лабанаускас во второй половине XX века (см. материалы, опубликованные в [Лабанаускас 2002]).

Такое упрощение фонотактики в энецком языке само по себе странно: географически энецкий занимает центральную позицию в гипотетическом континууме диалектов прасеверносамодийского, так что скорее мы ожидали бы более или менее консервативное развитие. Как кажется, такая радикальная перестройка могла быть спровоцирована «столкновением» в ситуации языкового контакта двух различных морфонологических

¹³ После знака / приводится вариант записи этимологического кластера у Г. Ф. Миллера (материалы XVII в., см., например, в публикации [Donner 1932: 36–50]) и М. А. Кастрена (материалы XIX в. [Castrén 1855]) в том случае, если эти варианты отличаются от современных.

систем — системы языка, уже утратившего чередование (типа ненецкого) и системы языка с чередованием ступеней согласных (нганасанского типа).

Суть нганасанского чередования ступеней (см. [Хелимский 1994: 195–198]) сводится к тому, что ряд одиночных согласных и кластеры вида «гоморганные назальный + шумный» реализуются по-разному в зависимости, во-первых, от четности / нечетности слога (ритмическая градация, ср. реализацию кластера **nt* после нечетного и после четного количества слогов: *ня-нтэ* [друг-GEN.2] ‘твоего друга’ vs. *нюо-мэ* [ребенок-GEN.2] ‘твоего ребенка’), во-вторых, в зависимости от открытости / закрытости слога (слоговая градация, ср. реализацию кластера **nt* перед открытым и перед закрытым слогом в формах вспомогательного отрицательного глагола: *ни-нты ту-?* [NEG-PRAES.S3 прийти-CN] ‘не пришел’ vs. *ни-нды? ту-?* [NEG-PRAES.S3PL прийти-CN] ‘не пришли’).

При утрате чередования ступеней система упрощается за счет обобщения одного из алломорфов корня либо суффикса (как правило, сильноступенного): чередование ступеней имеет уральское происхождение [Хелимский 1995/2000], однако из всех уральских языков оно сохранилось только в прибалтийско-финских (прежде всего в части слоговой градации), саамском и самодийских, причем из самодийских — только в нганасанском и кетском диалекте селькупского.

На наш взгляд, сложнее дело может обстоять именно в том случае, когда язык без чередования ступеней испытывает влияние со стороны родственного языка с чередованием ступеней. Можно предположить, что в ситуации энецко-нганасанского билингвизма носители энецкого (L1) испытывали значительное когнитивное давление, когда в языке L1 не чередуются те же самые кластеры, которые чередуются в «модельном» нганасанском языке (L2). Фонотактическое упрощение интервокальной позиции в энецком могло быть способом разрешить этот конфликт: если из системы устраняется кластер, он перестает отождествляться с таким же кластером нганасанского языка. Этот способ устранения кластеров затем, вероятно, распространился и на нечередующиеся кластеры.

Отметим, что сходный процесс «деназализации» кластеров ск. *mp*, *nt*, *nč*, *ŋt*, *mt*, *ŋk* прошел — достаточно поздно, в XIX–XX вв. — в нижнекетском и верхнеобском селькупском. Это также контактная зона, в которой встречаются два идиома, различающиеся наличием (кетский селькупский) vs. отсутствием (верхнеобский селькупский) чередования ступеней согласных. Для кетского диалекта М. А. Кастрен фиксирует сильноступенные формы следующих кластеров: **nč* (*sündde* ‘нутро’), **nč* (*kondze* ‘корень’), **nt* (*anddu* ‘лодка’), **ŋt* (*aŋdde* ‘лезвие’), **mt* (*amdde* ‘рог’). В северной части верхнеобской диалектной зоны и в нижнекетском имеем деназализованные формы типа верхнеобского *tan ad-dokaŋt* ‘к твоему обласку’ < *tan andokaŋt*, см. [Филиппова 2011: 83].

То, что упрощению кластеров может способствовать также сам факт интенсивных языковых контактов, можно подтвердить более далекой аналогией — историей прауральских кластеров в языках различных ветвей уральской языковой семьи. Кластеры вида «носовой + смычный» подверглись деназализации в трех языках: прапермском, венгерском и (спорадически) в марийском, ср.:

Таблица 5

Интервокальные кластеры в венгерском, прапермском и прамарийском

ПУ	Венгерский	Прапермский	Прамарийский
<i>*mp</i>	<i>*kompa</i> ‘волна’ > <i>hab</i>	<i>*kumpi</i> ‘кочка’ > <i>*gib äd</i>	<i>*lämpi</i> ‘теплый’ > <i>*liwə</i>
<i>*mt</i>	<i>*tumti</i> ‘знать’ > <i>tud</i>	<i>*kamti</i> ‘крышка’ > <i>*kud</i>	(но <i>*kamti</i> ‘крышка’ > <i>*komdəs</i>)
<i>*nt</i>	<i>*jänti</i> ‘тетива’ > <i>ideg</i>	<i>*künti</i> ‘туман, дым’ > <i>*küd</i>	<i>*jänti</i> ‘тетива’ > <i>*jið aŋ</i> (но <i>*kanta</i> ‘нести’ > <i>*kändə-</i>)
<i>*nč</i>	<i>*ponči</i> ‘хвост; ручка’ > <i>far</i>	<i>*ponči</i> ‘хвост; ручка’ > <i>*būž</i>	<i>*künči</i> ‘коготь’ > <i>*kič</i> (но <i>*künčä-</i> ‘царапать’ > <i>*künčə-</i>)

ПУ	Венгерский	Прапермский	Прамарийский
<i>*h̥c</i>	<i>*kuñci</i> ‘моча’ > <i>húgy</i>	<i>*kuñci</i> ‘моча’ > <i>*küž</i>	<i>*kuñci</i> ‘моча’ > <i>*küž</i>
<i>*h̥c</i>	<i>*lāñcā</i> ‘овод’ > <i>légy</i>	<i>*lāñcā</i> ‘овод’ > коми <i>lāž</i>	?
<i>*h̥s</i>	<i>*joŋsi</i> ‘поклон’ > <i>ij</i>	?	?
<i>*h̥k</i>	<i>*jāŋi</i> ‘лед’ > <i>jég</i> (праугорский <i>*h̥k</i> < прауральского <i>*h̥</i>)	<i>*vēŋka</i> ‘ручка’ > <i>*vug</i>	<i>*jāŋkā</i> ‘болото’ > <i>*jāgīr</i> ‘болотистый лес’

Венгерский относится к угорской ветви, прапермский и прамарийский каждый образуют отдельные ветви прафиннопермского языка, то есть в генетическом отношении все эти языки далеки друг от друга, и их не объединяет никакой промежуточный праязык, к которому мы могли бы относить денализацию кластеров как специфическую инновацию. Эти языки объединяет между собой другое — вовлечение в сферу влияния болгарского государства и контакты с носителями языка или языков болгарского типа (см., например, [Напольских 2006/2018] с анализом релевантной литературы). При этом само по себе наличие кластеров тюркской фонотактики не противоречит; таким образом, если и в этом случае связывать упрощение фонотактики в перечисленных уральских языках с языковыми контактами, то в качестве причины такого упрощения мы должны рассматривать не копирование какой-либо конкретной (фонотактической) модели языка L2, а саму контактную ситуацию как таковую. Интересно также обратить внимание на то, что в венгерском упростился целый ряд других прауральских кластеров, ср.: **močki-* ‘мыть’ > *mos-*, **níčkā-* ‘тянуть’ > *nyes-*, **męksa* ‘печень’ > *máj* (также праперм. **mus*), **ćilmä* ‘глаз’ > *szēm*, **wajŋi* ‘дышать’ > *vágy*, **mälki* ‘грудь’ > *mell* (также прамарийск. **mīl*, праперм. **mel*), **purki* ‘пурга’ > *forr* (также праперм. **pīr-* ‘пуржить’) и т. д. По числу упростившихся кластеров венгерский является «чемпионом» среди финноугорских языков, что составляет резкий контраст даже с генетически наиболее близкими ему мансийским и хантыйским: обско-угорские языки едва ли не самые консервативные среди уральских с точки зрения сохранения консонантных сочетаний. Возможно, это радикальное упрощение кластеров в венгерском было результатом того, что венгерский является «чемпионом» среди финно-угорских языков также и по сложности и интенсивности языковых контактов.

3.2. Утрата назальной протезы

В ненецком и нганасанском в словах, этимологически имевших вокалический анлаут, представлена назальная протеза: *ń* либо *ŋ*. Условия распределения протетических *ń* и *ŋ* в обоих языках полностью повторяют условия развития **k*: перед теми гласными, перед которыми в данном языке произошел переход **k* > *s*, употребляется протетический *ń*, а перед теми гласными, перед которыми в данном языке **k* сохраняется, употребляется протетический *ŋ*¹⁴.

В энецком назальной протезы нет: она, несомненно, была представлена (см. аргументацию ниже), но впоследствии устранена. Протетический *ŋ-* может (не обязательно) сохраняться в энецком в словах, требующих «усиления» фонетической структуры — только в корнях, которые иначе имели бы вид V(VV), ср. т. эн. *ŋa*, л. эн. *ŋo* ‘нога’ < **āj* (? **aāj*

¹⁴ Возможно, протетическим назальным исходно во всех случаях был **ŋ*, который впоследствии был палатализован в **ń* в тех же контекстах, в которых произошел переход **k* > *s* (напомним, что условия этого перехода различны в северносамодийских языках — как различны и контексты распределения протетических *ŋ* и *ń*). Для нас не является существенным выбор между двумя

[Janhunen 1977: 17]); т. эн. *ḡaa*, л. эн. *ḡa* ‘погода, бог’, ср. нг. *ḡuo* с тем же значением; т. эн., л. эн. *ḡu* ‘шест’ < **ḡj* [Ibid.: 29]; т. эн. *ḡiio*, л. эн. *ḡii* ‘трава’, ср. т. нен. *ḡit* ‘трава’ и т. д. Что касается протетического *ḡ*, он также был устранен в энецком — за одним исключением: он сохраняется в лесном энецком в словах с анлаутным **a* (U):

- л. эн. *nado* ‘рог’ (при т. эн. *edo*), ср. нг. *ḡamtə* (U), т. нен. *nyamt*;
- л. эн. *nado* ‘лезвие’ (при т. эн. *edo*), ср. нг. *ḡače* (U), т. нен. *nyand*;
- л. эн. *naʔ* ‘рот’ (при т. эн. *eʔ*), ср. нг. *ḡau* (U), т. нен. *nyah*;
- л. эн. *nadoo-* ‘взмолиться’ (при т. эн. *edoo-*), ср. нг. *ḡantəmə-* (U), т. нен. *nyantə-*;
- л. эн. *nee* ‘тот, другой’ (также т. эн. *ee-* в *eesaa* ‘девять’, ср. л. эн. *neesaa*), ср. нг. *ḡamiy* (U), т. нен. *nyaby*.

Как можно видеть, протетический *ḡ* в лесном энецком в данном случае депалатализован (что и привело к его сохранению, см. ниже о падении анлаутного *ḡ*) — однако это именно протетический *ḡ*, как в ненецком, а не протетический *ḡ*, как в нганасанском. Таким образом, система протетических назальных в энецком совпадала с ненецкой, а не с нганасанской, даже несмотря на то, что рефлексация гласных в этих словах в лесном энецком зависит от гармонии, как в нганасанском (то есть мы имеем наложение адстратных характеристик на субстратные).

Депалатализация протетического *ḡ* в энецком скоррелирована с депалатализацией **ḡ*. В энецком **ḡ* депалатализуется перед заднерядными гласными, например: т. эн. *naba* ‘заяц’ (в л. эн. *naba* — очевидно, под вторичным влиянием ненецкого) < **ḡāḡā*; эн. *niga* ‘тальник’ < **ḡerka* и т. д. Перед этимологическими передними гласными палатальность *ḡ* (и протетического, и этимологического) сохраняется. Хронологически после депалатализации *ḡ* в энецком произошло устранение назальной протезы (*ḡ* и *ḡ*); вместе с протетическим *ḡ* в энецком отпал также **ḡ* перед теми гласными, перед которыми он сохранял палатализацию (сохранение депалатализованного протетического *ḡ* в лесном энецком в словах типа л. эн. *nado* ‘рог’ < *nado* < **āntā* свидетельствует о том, что исконно назальная протеза была характерна и для энецкого):

Таблица 6

Устранение *ḡ* (протетического и этимологического) в энецком

	ПСС	Энецкий	Ненецкий	Нганасанский
Протетический <i>ḡ</i>	<i>*āḡā</i> (I) ‘мать’	<i>ee</i>	<i>nyebuḡa</i>	<i>немы</i>
Этимологический <i>*ḡ</i>	<i>*ḡārə-</i> (I) ‘передняя часть’	т. эн. <i>ere-</i> , т. эн., л. эн. <i>oro-</i>	<i>nyer^o-</i>	<i>нерə-</i>
	<i>*ḡensə-</i> ‘скользить’	т. эн. <i>ed’u-ir-</i> , л. эн. <i>esu-ir-</i>	<i>nyensə-r-</i>	<i>нянсю-</i>

Одним из возможных объяснений устранения назальной протезы в энецком, возможно, являются вторичные контакты с нганасанским: система назальных протез, совпадающая с адстратной ненецкой, отличалась от системы, представленной в нганасанском (см. таблицу 7), так что восстановление вокалического анлаута может рассматриваться как способ устранить конфликт двух систем, убрав те согласные в анлауте, которые не совпадают в энецкой и нганасанской системах¹⁵.

моделями возникновения назальной протезы: с двумя назальными **ḡ* и **ḡ* или с одним назальным **ḡ*, палатализованным в *ḡ* в ряде контекстов.

¹⁵ Этому способствует то, что в ненецких диалектах протетический назальный также может в редких случаях устраняться (см., например, слова с вокалическим анлаутом в словаре [Lehtisalo 1956]), так что в определенном смысле эти фонемы, вероятно, отличались от полноценного

Таблица 7

Протетические назальные ненецкого и нганасанского

	Нганасанский	Тундровый ненецкий
* <i>â</i>		<i>ŋ</i>
* <i>a</i> (U)	<i>ŋ</i>	<i>ń</i>
* <i>a</i> (I)		<i>ń</i>
* <i>ä</i> (U)	<i>ŋ</i>	<i>ń</i>
* <i>ä</i> (I)		<i>ń</i>
* <i>ɛ</i>		<i>ŋ</i>
* <i>e</i>		<i>ń</i>
* <i>i</i>		<i>ń</i>
* <i>ĩ</i>		<i>ŋ</i>
* <i>o</i>		<i>ŋ</i>
* <i>u</i>		<i>ŋ</i>
* <i>ũ</i>	<i>ŋ</i>	<i>ń</i>
* <i>ǣ</i> // <i>ǥ</i>		<i>ŋ</i>

Возможно, однако, что устранение назальной протезы вызвано и внутренними причинами: нельзя игнорировать тот факт, что конфликт между системами есть только в трех историко-фонетических контекстах из тринадцати рассмотренных в таблице 7, и, напротив, в одном из случаев, где такой конфликт есть, — ср. т. нен. *пуо*, нг. *ɲua* ‘дверь’ — энецкий сохранил протезу, ср. *n* в лесном (*no* из более раннего *ńo*) и *ŋ* в тундровом (*ɲia*). Поскольку форма в лесном энецком при этом прошла через депалатализацию начального *ń*, она, вероятно, не может быть отнесена к недавним заимствованиям (хотя депалатализацию в данном случае можно было бы объяснить и вторичным выравниванием при заимствовании: в энецком слов на *ńo* нет). Я благодарю рецензента, указавшего мне на возможные трудности интерпретации устранения назальной протезы как контактно обусловленного явления.

3.3. Палатализация согласных в энецком

Вторичными контактами с нганасанским, вероятно, можно объяснить и условия еще более позднего процесса — палатализации согласных в энецком. Для датировки этого процесса существенно, что мы можем показать, что в период устранения назальной протезы в энецком согласные перед передними гласными не палатализировались: **n* (непалатальный) перед передними гласными в этот период в энецком не совпал с **ń*, то есть в энецком **n* не отпадает перед передними гласными (в отличие от **ĩ* и протетического *ń*), ср.: **nim* ~ **nim* ‘имя’ > совр. т. эн. *niĩ*? (не ***iĩ*?).

Влияние нганасанского на процесс палатализации согласных в энецком можно видеть в двух моментах. Во-первых, в энецком, как и в нганасанском, палатализуются только дентальные согласные *t*, *n*, *s*, *l* (в ненецком палатализуются все). Таким образом, инвентарь согласных фонем в энецком ближе к нганасанскому, чем к ненецкому. Во-вторых, важно проинтерпретировать следующую ситуацию: как уже говорилось выше, **a* и **ä*

консонантного анлаута менее тесной связью с корнем (скорее с точки зрения внутрисистемной, нежели собственно фонетической).

исторически ведут себя в энецком как передние гласные — как в ненецком: перед ними происходит переход **k > s* и употребляется протетический *ń* (правда, депалатализирующийся в л. эн. перед **a* (U)). Однако в современном энецком дентальные перед этими гласными не палатализируются.

Таблица 8

Палатализация согласных в северносамодийских языках: примеры¹⁶

П(С)С	Нганасанский	Энецкий	Тундровый ненецкий
<i>*a</i> (U)	<i>*tanājw</i> ‘мало’	<i>tan^ua</i>	<i>tɥanyo</i>
<i>*a</i> (I)	<i>*tajbā</i> ‘брюхо’	<i>t^ʰajbā</i>	<i>tyɥw^o</i>
<i>*ä</i> (U)	<i>*sänkä</i> ‘колокольчик’	<i>saŋku</i>	<i>syɥŋka</i>
<i>*ä</i> (I)	<i>*tänä-</i> ‘знать’	<i>t^ʰenj-</i>	<i>tyɥnye-</i>
<i>*i</i>	<i>*tirāmä</i> ‘рыбья икра’	<i>t^ʰirimi</i>	<i>tyɥryɥbya</i>
<i>*e</i>	<i>*tetä</i> ‘дядя, младший брат матери’	<i>tɥtɥd^ʰa</i>	<i>tyɥdɥa</i>
<i>*ü</i>	<i>*türo</i> ‘посох’	<i>t^ʰirä</i>	<i>tyur^o</i>

Таблица 9

Палатализация согласных в северносамодийских языках: обобщение¹⁷

П(С)С	Нганасанский	Энецкий	Тундровый ненецкий
<i>*a</i> (U)	НП		
<i>*a</i> (I)	П	НП	П
<i>*ä</i> (U)	НП		
<i>*ä</i> (I)	П	НП	П
<i>*i</i>		П	
<i>*e</i>	НП		П
<i>*ü</i>		П	

Дентальные, таким образом, палатализируются в энецком только перед **i* (а также перед совпавшим с ним в ненецком и энецком **e*) и перед **ü*. При этом в энецком максимально упрощена та зона, в которой нганасанский устроен сложнее, чем ненецкий: ненецкому *C'e* (< **ä*) могут соответствовать и нг. *C'e*, и нг. *Ca*, ненецкому *C'a* (< **a*) — и нг. *C'a*, и нг. *Ca*. В энецком — единственном — не только совпадают рефлексy **a* и **ä*, но и согласные перед ними не палатализируются (поскольку в нганасанском рефлексy этих гласных могут и палатализовать, и не палатализовать согласные).

3.4. Выводы

Итак, можно предложить единообразное объяснение причин возникновения в энецком языке основных инноваций в сфере фонетики, отличающих энецкий как от ненецкого,

¹⁶ В таблицy внесены только те гласные, перед которыми согласный палатализуется хотя бы в одном из языков.

¹⁷ Обозначения в таблицe: НП — непалатализованный, П — палатализованный.

так и от нганасанского. Фонетическая система энецкого изначально, как мы предполагаем, была в целом близка к ненецкой (в результате перехода с языка типа нганасанского на язык типа ненецкого). Затем в какой-то период энецкий вошел в ситуацию сильного вторичного контакта с нганасанским, в которой нганасанский выступал как «модельный» язык-донор, а энецкий — как язык-реципиент. Эта контактная ситуация спровоцировала возникновение ряда инноваций в тех случаях, в которых система нганасанского оказалась сложнее системы «ненецкого типа». При этом все эти инновации имеют один общий вектор: радикальное упрощение фрагментов энецкой системы, ведущее к их максимальному расподоблению с соответствующими более сложно устроенными фрагментами нганасанской системы¹⁸. Важно еще раз подчеркнуть, что относительная датировка этих инноваций, то есть отнесение их к периоду вторичного контакта энецкого и нганасанского, а не к более раннему периоду перехода с языка типа нганасанского на язык типа ненецкого, основано на следующем соображении. Причиной этих инноваций, на наш взгляд, является именно устранение ряда «проблемных» соответствий при столкновении более простой системы (ненецко-энецкого типа) с более сложной системой (нганасанского типа), в ситуации же перехода, напротив, более сложная система заменялась более простой, так что все перечисленные фонологические контексты потенциальной проблемы в этом случае не создавали бы. Эти вторичные контакты (как и первичные) относились к общеэнецкой эпохе, то есть затронули и будущий лесной, и будущий тундровый энецкий (сепаратные контакты тундрового энецкого с нганасанским в статье не рассматриваются).

К объяснению, апеллирующему к ситуации языковых контактов, нас подталкивает и более общее соображение: энецкий занимает центральное положение в континууме близкородственных языков¹⁹, где ожидалась бы скорее консервация языковых явлений, так что значительные изменения в сфере исторической фонетики вряд ли можно объяснить исключительно внутривидовыми соображениями.

¹⁸ Здесь в качестве аналогии можно привести анализ результата контактов родственных идиомов, предложенный Ю. К. Кузьменко: «Появление германо-кельтско-италийской просодической изоглоссы (фиксация ударения на корне. — *A.V.*) можно представить себе и как явление, связанное с контактами родственных языков германского, италийского и кельтского (до Первого перебоя и закона Вернера, и до многих других германских изменений), когда протогерманский, протоиталийский и протокельтский еще не слишком отличались друг от друга. Модель их контактов могла бы соответствовать модели, предложенной для объяснения редукции флексий в английском языке, которая по Есперсену была предопределена сходством корней и несходством окончаний в древнеанглийском и скандинавском в эпоху датского завоевания Англии. Причем эти контакты предопределили даже не столько редукцию окончаний, сколько выделение корневой морфемы. Типологически сходный процесс выделения корня при контакте родственных языков со сходными корнями был характерен и для южнодатско-северофризско-нижненемецкой контактной зоны на юге Ютландии [Кузьменко 1995]. Можно предположить, что подобные процессы происходили и при германо-итало-кельтских контактах в то время, когда многие корни сохраняли сходство, так что была возможна семикоммуникация» [Кузьменко 2011: 43]. В данном случае интересно обратить внимание на выделенную тенденцию: при семикоммуникации носителей родственных языков сохраняются сходные элементы системы и устраняются различные. Именно с проявлением этой тенденции, как кажется, можно связать возникновение собственно энецких историко-фонологических инноваций — максимального упрощения тех фрагментов системы, в которой нганасанский устроен сложнее, чем ненецкий (и наследующий ему энецкий). Это именно те случаи, когда различие в системах наиболее очевидно: с одной стороны, системы (до возникновения инноваций) достаточно схожи, чтобы между ними было возможно отождествление, с другой стороны, из-за принципиально более высокой сложности нганасанской системы в этих случаях полное отождествление (в частности, за счет «пересчета» между системами) невозможно.

¹⁹ Это положение центрально географически, но не лингвистически: центром зарождения и распространения инноваций энецкий не мог быть в силу малого числа говорящих.

Интересно также отметить, что М. А. Кастрен фиксировал для энецкого переход анлаутного $*p > f$, аналогично нганасанскому того периода (в современном нганасанском произошел дальнейший переход $*p > f > h$). В современном энецком f возможен как маргинальный вариант, но основным является p — как в ненецком. Это, в частности, еще раз подтверждает, что энецкий — вероятно, в силу географически центрального положения и одновременно малочисленности носителей — на протяжении всей своей истории испытывал влияние соседних идиомов: нганасанского и ненецкого. Причем в силу экстралингвистических причин он дрейфовал между этими двумя центрами притяжения, что могло менять вектор развития отдельных фрагментов энецкой языковой системы.

4. Морфонологические субстратные явления

В энецком и нганасанском представлено фонетически немотивированное развитие ряда непродуктивных суффиксов с перфективным значением, оканчивающихся согласным, ср. материал в таблице 10. Прибавление гласного после этих суффиксов (по крайней мере два из них являются общесамодийскими) не мотивировано законами исторической фонетики ни для нганасанского, ни для энецкого. Кроме того, не каждый суффикс такой фонетической структуры развивается так в нганасанском и энецком: ни имперфективный $*-r$ (эн. $d'adu-r-$, нг. $d\acute{e}žy\grave{p}-$ ‘ходить’), ни суффикс $*-m$, образующий отыменные глаголы с инцептивным значением (эн. $d'eri-m-$, нг. $d\acute{y}ala-m-$ ‘рассвести’), так себя не ведут.

Таблица 10

Исторически немотивированное развитие ряда непродуктивных суффиксов
 $*(CI)C- > (CI)C\acute{a}$ в энецком и нганасанском

Суффикс	Значение	Ненецкий	Энецкий	Нганасанский
$*-k\acute{a}r$ ‘семельфактив’	‘проникнуть; зайти (о солнце)’	$p\acute{a}k\acute{a}l-$	$pokoro-$	$hok\acute{a}l\acute{a}-$
	‘дернуть’	$nek\acute{a}l-$	$nekoro-$	$ny\acute{a}k\acute{a}l\acute{a}-$
$*-r$ ‘перфектив’	‘запрячь’	$podyer-$	$podaro-$	$xutyra-$
	‘обтесать’	$syab\acute{a}r-$	$seboro-$	—
	‘ободрать’	$x\acute{a}b\acute{a}r-$	—	$k\acute{a}x\acute{a}r\acute{a}-$
$*-l$ ‘инхоатив’	‘начать ходить’	$yad\acute{a}l-$	$d'adu\acute{r}o-$	$d\acute{e}žy\grave{l}\acute{a}-$
	‘начать жить’	$yilyel-$	$irero-$	$nylyle-$
$*-m$ ‘моментатив’	‘взмолиться’	$nyan\acute{a}t-$	$edoo-$	$yan\acute{t}\acute{a}m\acute{a}-$
	‘взглянуть; заметить’	$x\acute{a}ed\acute{a}-m-$	$ki\acute{d}o-o-$	$k\acute{a}chem\acute{a}-$

5. Именное словообразование: субстратные явления

В таблице 11 представлены диминутивные показатели северносамодийских языков²⁰.

²⁰ Их алломорфы обсуждались в разделе 2.3 выше.

Таблица 11

Диминутивные показатели северносамодийских языков

ПСС	Нганасанский	Тундровый ненецкий	Энецкий
*-jeŋ ^o ko	-AŋKU ₀ диндаŋку 'лучок' (оружие)	-yeko myadyeko 'чумик'	-iku buŋiku 'собачка', tiŋiku 'жилка'
*-jeŋko	-AŋKU ₀ мазанку 'чумик'	—	-igu meŋigu 'чумик' < meŋ 'чум' bexirigu 'сопочка' < bexiŋ (ŋ < r) 'сопка'
		-ko sirako 'снежок' < sira 'снег'	-ku siraku 'снежок' < sira 'снег'

Энецкий непродуктивный диминутивный показатель -igu (~ -id'u) является общесамодийским, ср. диминутив от слова *mât 'чум': нг. *мазанку*, т. эн. *meŋigu*, ск. *māduŋga* 'чумик'. Показательно, что в энецком (как в нганасанском) этот суффикс в качестве непродуктивного сохраняется, тогда как в ненецком нет ни следа данного показателя. Характеристика, объединяющая энецкий с нганасанским, может рассматриваться только как архаизм, но не как вторичное контактное влияние (в энецком этот суффикс, во-первых, непродуктивен, во-вторых, имеет сложное и архаичное распределение алломорфов, обсуждавшееся выше). Теоретически, конечно, можно было бы говорить и о независимом сохранении унаследованного от праязыка суффикса, но совокупность рассмотренных явлений, как кажется, делает более логичной трактовку этого совпадения между нганасанским и энецком в рамках модели смены языка, предложенной в данной статье.

6. Именное словообразование: вторичное влияние нганасанского

Слова, содержащие атрибутивный показатель *-jâ, демонстрируют в энецком двоякое фонетическое развитие последовательностей *-V-jâ. Эта последовательность развивается в одиночный гласный прежде всего в составных суффиксальных морфемах, ср. показатели пассивного перфективного причастия *-mâ-jâ > эн. -i, т. нен. -wi^o, нг. -мээ или суффикс порядковых числительных (родственный имперфективному причастию) *-ntâ-jâ, например, в эн. *пехо-de*, т. нен. *пуах^oрэм-tey^o*, нг. *нагэм-туо* 'третий'. Можно видеть, что эта монофтонгизация в целом повторяет ненецкую модель: во-первых, изменения гласных перед атрибутивным суффиксом (*â > i, *â > e) совпадают в этих двух языках, во-вторых, в ненецком эти последовательности также монофтонгизированы (последний гласный в абсолютном ауслауте редуцируется, а при словоизменении вообще отпадает).

Одиночный гласный выступает в энецком также при утрате атрибутивным суффиксом статуса отдельной морфемы, то есть в лексикализовавшихся формах. Это незначительное количество старых субстантивированных производных типа *baŋi* 'замена' от *baŋa* 'место' или *sire* 'двухгодовалая важенька' от *sira* 'зима'.

Напротив, *-V-jâ сохраняется в энецком в виде вокалической последовательности в том случае, если морфема *-jâ, прибавляясь непосредственно к основе, сохраняет атрибутивное значение, ср.:

т. эн. *baroo* 'крайний; край, берег', ср. т. нен. *wari^o* 'крайний' и нг. *бэрээ* 'берег', производные от т. эн. *baro*, т. нен. *war*, нг. *бэрэ* 'берег, край';

т. эн. *ud'aa*, л. эн. *osae* ‘много кушающий, ненасытный’; т. нен. *ḡatseyo* ‘отечный, нездорово полный (о человеке)’, нг. *ḡamsyo* ‘мясной’ — все три слова являются семантически разошедшимися производными от слова т. эн. *ud'a*, л. эн. *osa*, т. нен. *ḡatca*, нг. *ḡamsy* ‘еда’;

т. эн. *d'erio*, т. нен. *yalyeyo*, нг. *дялыэ* ‘дневной’ производны от т. эн. *d'ere*, т. нен. *yalya*, нг. *дялы* ‘день’.

Итак, в том случае, когда атрибутивный суффикс имеет отчетливую словообразовательную функцию, рефлексация в энецком следует нганасанскому типу, с сохранением вокалической последовательности. Определенно более старыми выглядят морфологические контексты, где атрибутивный суффикс уже утратил статус отдельного словообразовательного показателя и где рефлексом является одиночный гласный. Соответственно, в развитии модели с вокалической последовательностью мы можем видеть вторичное влияние нганасанского: резаймствование атрибутивного суффикса (который в нганасанском является совершенно продуктивным показателем) с сохранением его нганасанских морфонологических особенностей.

7. Именное словоизменение: субстратное влияние нганасанского

В именном склонении интересно рассмотреть различия в образовании форм локативных падежей во множественном числе (отражающее как субстратное, так и вторичное влияние).

Показатели локативных падежей в северносамодийских языках состоят из коаффикса **-kâ* (> нен. *-xə*, эн. *-xo*, нг. *-Kə*) либо **-ntâ* (> нен. *-nə*, эн. *-do*, нг. *-NTə*) и показателя собственно локативного падежа уральского происхождения (LAT **-ŋ*, LOC **-nâ*, EL **-tâ*):

LAT.SG **-ntâ-ŋ* > т. нен. *-nə-h*, т. эн. *do-ʔ*, нг. *-NTə-^C*;

LOC.SG **-kâ-nâ* > т. нен. *-xə-na*, т. эн. *-xo-ne*; **-ntâ-nâ* > нг. *-NTə-NU*.

Множественное число в формах локативных падежей маркируется в позиции после коаффикса: **коаффикс + *-jt + показатель локативного падежа*, ср. формы LOC.SG и LOC.PL:

Таблица 12

Образование показателей локативных падежей множественного числа
в северносамодийских языках: пример

	Ненецкий	Энецкий	Нганасанский
LOC.SG	<i>*-kâ-nâ</i> >		<i>*-ntâ-nâ</i> >
	<i>-xə-na</i>	<i>-xo-ne</i>	<i>-NTə-NU</i>
LOC.PL	<i>*-kâ-jt-nâ</i> >		<i>*-ntâ-jt-nâ</i> >
	<i>-xə-q-na</i>	<i>-xi-ne</i>	<i>-NT^C-NU</i>

В энецком представлен отклоняющийся от этой модели показатель латива множественного числа *-xiđo*:

Если бы в энецком показатель LAT.PL развивался из праформы **-kâ-jt-ŋ* (как в ненецком), этимологически корректной была бы форма *-xiʔ* (соответствующая т. нен. *-xəq*), и она, скорее всего, сохранилась бы в энецком. Однако, если форма латива была адаптирована энецким из адстратной системы ненецкого типа (при переходе с языка типа нганасанского, где в лативе используется другой коаффикс!), возможно, препятствием для ее освоения послужила морфологическая нечленимость этой формы в ненецком: в силу

Таблица 13

Латив множественного числа в северносамодийских языках

	Ненецкий	Энецкий	Нганасанский
LAT.SG	-nə-h	*-ntə-ŋ > -do-ʔ	-NTə- ^c
LAT.PL	*-kə-ʃt-ŋ > -xəq	??? -xiɖo	*-ntə-ʃt-ŋ > -NTIʔ

законов исторической фонетики в этой форме не представлен показатель латива т. нен. -h (< *-ŋ). Поэтому, вероятно, при освоении этой формы к ней был прибавлен плеонастический показатель собственно латива *-ŋ: протоэн. *-kit (соответствует т. нен. -xəq) + *ŋ > *-kit-ə-ŋ > протоэн. *-kido-ʔ > эн. -xiɖoʔ > эн. -xiɖo.

8. Именное словоизменение: вторичное влияние нганасанского

Ряд других изменений в субпарадигме локативных падежей множественного числа в энецком можно объяснить, напротив, вторичным влиянием нганасанского.

Как видно из таблиц 12 и 13 выше, на синхронном уровне падежные показатели мн. ч. отличаются в энецком и нганасанском от показателей соответствующих падежей единственного числа изменением гласного в коаффиксе. При этом в нганасанском наряду со стандартным показателем EL.PL -KI^cTə (соотв. т. нен. -xə-tə из *-xə-q-tə, т. эн. -xito) непритяжательного склонения представлены показатели нг. -KI^cTIʔ (в непритяжательном склонении) и нг. -KI^cTI^c (в притяжательном склонении, где это единственный вариант). В этих показателях множественное число маркируется дважды: после коаффикса и после падежного показателя, что на синхронном уровне легко определить по изменению качества гласного. В нганасанском основанием для развития таких форм аблатива могла послужить чисто формальная аналогия с показателем латива мн. ч., ср. формы притяжательного склонения с показателем 1SG от слова *kuhu* 'шкура' в таблице 14. Как можно видеть из сопоставления приведенных ниже словоформ, при различии глубинно-морфологического представления (и этимологии стоящих за ним показателей) на фонологическом уровне в этих формах перед притяжательным показателем в единственном числе выступает сегмент *tə*, во множественном — сегмент */ti/*:

Таблица 14

Формы латива и аблатива притяжательного склонения с показателем 1SG нг. *kuhu* 'шкура'

Форма	Фонологическое представление	Морфофонологическое представление
LAT.SG-1SG	<i>kubutənə</i>	< KUHU-NTə-Nə
LAT.PL-1SG	<i>kubutinə</i>	< KUHU-NTI ^c -Nə
EL.SG-1SG	<i>kuhugətənə</i>	< KUHU-Kə ^c Tə-Nə
EL.PL-1SG	<i>kuhugitinə</i>	< KUHU-KI ^c TI ^c -Nə-

В энецком также есть аналогические формы с двойным маркированием множественного числа в притяжательной парадигме (то есть с перегласовкой гласного непосредственно

перед притяжательным показателем). В энецком они представлены в парадигме более широко — это не только аблатив (эн. *-xiti*, в точности соответствующий нганасанскому), но также и локатив и пролатив. Так, в записях М. А. Кастрена в притяжательном склонении представлены показатели LOC.PL *-xini* и PROL.PL *-oni* / *-ini*. В лесном энецком такие показатели сохранились до настоящего времени — во всяком случае, этот вариант в притяжательном склонении является наиболее частотным в текстах на лесном энецком в [Сорокина, Болина 2005]. В энецком появление таких форм внутренними причинами (аналогическим выравниванием внутри парадигмы) объяснить невозможно. Вероятнее, что это вторичная адаптация двойного маркирования из нганасанских форм аблатива и латива с дальнейшим распространением на другие энецкие падежи.

Интересно обратить внимание на еще одну исторически аналогическую форму парадигмы именного склонения в энецком. Форма пролатива с показателем **-tānā* имеет позднее, послеложное происхождение. В записях М. А. Кастрена в энецком показатель пролатива мн. ч. имеет два варианта: для непритяжательного склонения *-mone* и *-mine*, для притяжательного — *-moni* и *-mini*. Нас в данном случае будет интересовать гласный первого слога в показателях *-mine* и *-mini*. При сравнении этих вариантов показателя пролатива мн. ч. с этимологически корректным *-mone* очевидно, что в энецком первая часть двусложного показателя пролатива была реинтерпретирована — по аналогии с первой частью двусложных показателей локатива и аблатива — как коаффикс, после которого требуется присоединение показателя мн. ч., ср.:

Таблица 15

**Реинтерпретация показателя пролатива мн. ч.
по аналогии с показателями локатива и аблатива мн. ч. в энецком
(непритяжательное склонение)**

	SG	PL
LOC	<i>-xone</i>	<i>-xine</i>
EL	<i>-xodo</i>	<i>-xito</i>
Аналогический пролатив	<i>-mone</i>	<i>-mine</i>

Итак, в парадигме именного склонения мн. ч. именно для энецкого мы находим наибольшее число исторически немотивированных аналогических изменений (двойное маркирование мн. ч. в притяжательном склонении, реинтерпретация первой части пролативного показателя как коаффикса). Двойное маркирование мн. ч. в притяжательном склонении можно достаточно уверенно объяснить вторичным влиянием нганасанского. В целом же аномально большое — в сравнении с ненецким и нганасанским — число исторически немотивированных аналогических изменений в парадигме, приводящих к усилению маркированности грамматических значений, можно, вероятно, рассматривать как результат языковых контактов.

Интересно, что устранение таких аналогических форм в тундровом энецком (где сейчас имеем LOC.PL в притяжательной парадигме *-xine*, а не *-xini*, PROL.PL *-mone*, а не *-moni* или *-mini*) совпадает по времени с устранением возникшего под нганасанским влиянием перехода **p > f* за счет восстановления *p*. То есть эти скоррелированные во времени признаки преодоления вторичного нганасанского влияния, вероятно, совпадают с более активным развитием энецко-ненецкого билингвизма. Иначе говоря, происходит очередная смена центра притяжения в истории развития энецкого (о реконструкции мульти- и билингвизма на Нижнем Енисее см. [Khanina 2021]).

9. Глагольное словообразование: субстратное влияние нганасанского

9.1. Деривационные суффиксы образных глаголов

Особенностью образных глаголов в самодийских языках является то, что они обладают собственным репертуаром деривационных морфем (для ненецкого эти показатели подробно описаны в [Урманчиева 2020], здесь приведем информацию о релевантных для нас показателях).

В энецком представлены имперфективные образные глаголы с суффиксом *-na*. Они имеют соотносимые каузативы (*-na* заменяется на *-nta*) и производные от этих каузативов стативно-декаузативные глаголы с суффиксом *-nti*:

kirana- ‘орать, визжать’, *kiranta*- ‘довести до плача’, *kiranti*- ‘плакать после того, как раздразнили’;

sesana- ‘шелестеть, шуршать’, *sesanti*- ‘бренчать, позвякивать’;

tôjana- ‘быть твердым, жестким’, *toĵanta*- ‘делать твердым’;

tujana- ‘стучать’, *tujanta*- ‘заставлять издавать стук’, *tujanti*- ‘издавать стук’;

d’irana- ‘пищать’, *d’iranta*- ‘скрипеть’, *d’iranti*- ‘издавать скрип’.

Часть из этих глаголов имеет корреляты в ненецком — но как с суффиксом *-ə-nə* (*ses-ə-nə* ‘шуршать’, *yir-ə-nə* ‘скрипеть, пищать’), так и с суффиксом *-na ~ -ana ~ -yena* (*toy-ena* ‘быть твердым’), некоторые соотносимых глаголов не имеют. При этом между ненецким и энецким нерегулярны соответствия гласных не только в суффиксе, но и в предшествующей суффиксу гласной второго слога, что говорит о вторичном характере таких соответствий.

Каузативы на *-nta* (и соотносимые с ними стативно-декаузативные производные глаголы на *-nti*) также являются в энецком поздними вторичными образованиями: для энецкого не характерно ни синкопирование гласного перед каузативным суффиксом, ни тем более кластер *nt*. Такое количество фонетических, морфологических и морфонологических нерегулярностей говорит о том, что в энецком эта группа образных глаголов имеет позднее происхождение.

В энецком и нганасанском у ряда образных основ находим словообразовательные гнезда с суффиксами нг. *-^cTU*, эн. *-da* (непереходные имперфективы), нг. *-MÛS*, эн. *-us* (инцептивы) и нг. *-MƏBTU*, эн. *-ota* (каузативы с инцептивным значением).

Таблица 16

Словообразовательные гнезда образных глаголов в нганасанском и энецком

Праформа суффикса	Нганасанский	Энецкий
*-ntā	<i>-^cTU</i> дяба?туса ‘краснеть; быть красным’	<i>-da</i> d’areda- ‘стоять наклонно’
*-mus	<i>-MÛS</i> дяба?мұса ‘покраснеть’	<i>-us</i> d’areus- ‘стать на бок’ naguus- ‘стать красным’
*-məbtā	<i>-MƏBTU</i> дяба?мəбтудя ‘сделать красным; покраснеть’	<i>-ota</i> d’areota- ‘поставить на бок’ naguota- ‘сделать красным’

Таким образом, в энецком сосуществуют две системы суффиксации образных глаголов: одна — по многим признакам достаточно инновационная — находит параллели в системе

суффиксации образных глаголов ненецкого языка (т. нен. имперфективы с суффиксами *-nə* и *-na* ~ *-ana* ~ *-yena*), вторая же полностью изоморфна подсистеме аспектуальных суффиксов образных глаголов нганасанского языка.

9.2. Деривационные суффиксы необразных глаголов

В глаголах, не относящихся к группе образных, большинство аспектуальных суффиксов в трех северносамодийских языках совпадает, однако некоторые различия есть как в наборе аспектуальных суффиксов, так и в употреблении и продуктивности совпадающих маркеров.

Разительнее всего различия в сфере показателей дуратива (имперфектива). У непереходных глаголов во всех трех северносамодийских языках используется общесамодийский показатель **-ntā*, ср. т. нен. *ηom-tə-*, эн. *ηoō-do-*, нг. *ηuom-tə-* ‘собираться, объединяться’. В ненецком, однако, более продуктивен показатель **-ηkă*, который отсутствует как в энецком, так и в нганасанском.

Система имперфективных показателей переходных глаголов в ненецком отличается от систем, представленных в нганасанском и энецком.

В ненецком продуктивен общесамодийский показатель дуратива-результатива т. нен. *-mpə* < **-mpaw* (соотв. эн. *-be*). Присоединяясь к глагольным основам с переходной семантикой, он сохраняет их модель управления. В нганасанском этого показателя нет, в энецком он не употребляется при переходных глаголах. У непереходных глаголов этот суффикс непродуктивен как в ненецком, так и в энецком, в нганасанском он отсутствует. В ненецком общесамодийский показатель имперфектива т. нен. *-nə* < **-ntā* (эн. *-do*, нг. *-NTƏ*) при присоединении к переходным глаголам блокирует возможность выражения прямого дополнения, хотя глагол и сохраняет семантику переходной ситуации.

Покажем на примерах употребление этих суффиксов в ненецком. Рассмотрим два примера с глаголом *təwa-* ‘принести, притащить’: (1) *t'ĩββan niññe n'allođā tūēββā-bi-n* [притащить-DUR.AOR-OP.L1] ‘я тащу их, ползя на животе’ [Lehtisalo 1956: 294b] и (2) *pākkoĩse dāββa-nā* [притащить-IPFV.AOR.S3], *lēkkoĩse maiōrā* ‘палочка несет, косточка режет’ (загадка про ложку и зубы) [Ibid.: 231a]. В первом примере глагол оформлен суффиксом *-mpə* (в данном случае *bi*) и имеет прямое дополнение, выраженное формой объектного спряжения, во втором примере глагол оформлен суффиксом т. нен. *-nə* (в данном случае *-nā*) < **-ntā* и прямого дополнения при себе не предполагает. В энецком и нганасанском суффиксы эн. *-do*, нг. *-NTƏ* не оказывают влияния на модель управления глагола.

Еще один имперфективный показатель т. нен. *-ηko* < **-ηkoj* (эн. *-gu*) со значением ‘заниматься чем-л.’ в ненецком ведет себя так же, как *-nə*: сохраняет семантику переходного действия, но блокирует выражение прямого дополнения: *syurteηko-* ‘обходить, объезжать кругом (неперех.)’ ← *syurte-* ‘обойти, объехать кругом что-л. (перех.)’. В энецком суффикс *-gu* не оказывает влияния на модель управления глагола: *биз моза-га* [принести-DUR.AOR.S3], *пязда моза-га* [принести-DUR.AOR.S3] ‘воду ему носит, дрова ему носит’ [Сорокина, Болина 2005: текст 20, предл. 7].

В нганасанском представлен показатель дуратива нг. *-KUJ* < **-koj*, который также не влияет на модель управления глагола, ср. *сеймы-м-ты дятэбту-гуй-ся* [глаз-АСС-3 направить-DUR-INF] ‘следить за чем-то’ (букв. ‘направлять свой глаз’) (из нганасанских текстов, любезно предоставленных в мое распоряжение В. Ю. Гусевым).

Данные, кратко описанные в тексте и суммированные в таблице 17, свидетельствуют о том, что развитие системы дуративных суффиксов в энецком демонстрирует существенно большую общность с нганасанским, чем с ненецким. Во-первых, в энецком, как и в нганасанском, — но не в ненецком — утрачена продуктивность суффикса **-mpə*. Во-вторых, в ненецком развитие дуративных суффиксов шло таким путем, что близкие по значению

Таблица 17

Дуративы от переходных глаголов в северносамодийских языках

Производные глаголы не меняют модель управления		Производные глаголы имеют значение ‘заниматься чем-л.’ и не присоединяют прямого дополнения	
Ненецкий	<i>*-mpaw</i>	<i>*-ntâ</i>	<i>*-ŋkoj</i>
Энецкий	<i>*-ntâ</i>	<i>*-ŋkoj</i>	
Нганасанский		<i>*-koj</i>	

общесамодийские суффиксы **-mpâ* и **-ntâ* оказались в итоге в значительной степени разведены в зависимости от модели управления как исходного, так и производного глаголов; так же, как **-ntâ*, ведет себя в ненецком **-ŋkoj*. В нганасанском и энецком дуративные суффиксы модель управления глаголов не затрагивают. При этом энецкий показатель дуратива *-gu*, хотя и соответствует материально т. нен. *-ŋko*, является функционально тождественным показателю нг. *-KUᵒJ*.

Обратим внимание на то, что в ненецком и энецком, как уже говорилось, в суффиксах в переднерядных основах произошли переходы **t³k > *t³s* и **ŋk > *ns*. В ненецком это чередование сохранилось в формах с диминутивным суффиксом (*vidy-ecyo* ‘водичка’) и с дуративным показателем **-ŋkoj* (*ŋeda-ncyo-* ‘защищаться’), а в энецком переход **ŋk > *ns > эн. d’*, **t³k > *t³s > эн. t’* представлен только в формах с диминутивными показателями (*pi-id’u* ‘горностай’, *bid-ii’u* ‘водичка’), архаичность которых мы обсуждали в разделе 5. Ни одной лексикализованной формы с дуративным показателем с таким переходом не найдено. Это также может подтверждать вторичность, инновационность данного показателя в энецком — вероятно, он адаптирован из адстратного языка типа ненецкого, но употребляется по модели субстратного нганасанского показателя (к тому же достаточно близкого фонетически).

Затронем также различия показателей аспектуальных категорий в северносамодийских языках. В показателе хабитуалиса нг. *-MUMHAᶜ* (*муну-мумба-тум* [сказать-нав-s1sg] ‘я всегда (так) говорю’), эн. *-mobi* (*d’adu-mobi* [ходить-нав.s3sg] ‘постоянно ходит’) выделяется компонент, который, вероятно, исторически является обсуждавшимся выше показателем дуратива **-mpaw*²¹ — в этом случае его исчезновение из репертуара продуктивных глагольных показателей в энецком и нганасанском как раз может объясняться его грамматикализацией в составном показателе хабитуалиса. В ненецком показатель хабитуалиса имеет вид *-syəqtâ*, и этот показатель никак не связан с энецким и нганасанским. Можно предположить, что показатель хабитуалиса в энецком унаследован от субстрата нганасанского типа: если бы он был заимствован из нганасанского в результате вторичных контактов, это вряд ли оказало бы влияние на судьбу дуративного показателя в энецком, однако и в энецком, и в нганасанском утрата дуративного показателя **-mpaw* и возникновение хабитуального показателя могут рассматриваться как взаимосвязанные процессы.

²¹ Показатель дуратива имеет в энецком вид *-be*, показатель хабитуалиса — вид *-mobi*. Разница в гласном (*e* или *i*) объясняется тем, что хабитуалис не входит в лексическую основу глагола, то есть является грамматическим, а не деривационным показателем. Соответственно, в отличие от деривационного дуративного показателя, мы видим его только в составе финитных форм. Как показывает нганасанская форма *муну-мумба-тум* [говорить-нав-s1sg] ‘я всегда (так) говорю’, после хабитуального показателя употребляется показатель какого-либо времени (или наклонения). В энецком наиболее частотная финитная форма — аорист. Эта форма образуется от дуративных глаголов перегибкой *e → i*: *edebe-* ‘радоваться’ → *edebi-ᵒᵒ* [радоваться.aor-s1] ‘радуясь’. Можно предположить, что хабитуальный показатель *-mobi* представляет собой результат грамматикализации такой аористой формы.

10. Глагольное словоизменение: нганасанское субстратное влияние

10.1. Образование аориста

Образование аориста непереходных глаголов в энецком демонстрирует противопоставление нганасанского субстрата и ненецкого адстрата. Аорист — одна из общесамодийских темпоральных форм (предельные глаголы в аористе имеют значение прошедшего времени, неопределённые — настоящего). В энецком форма аориста от непроизводных непереходных глаголов с основами на эн. *o* (< **ə*) и *-a* (< **ā*) образуется двумя способами. Первый способ — неагглютинативный: показатель аориста *-a* (в рефлексивном спряжении *-e*) вытесняет финальный гласный основы: *ado*- ‘сесть’ → AOR.R3 *adeđo?*, *kođa*- ‘лечь спать, уснуть’ → AOR.R3 *kođeđo?*. Второй способ — агглютинативный: вышеназванные показатели прибавляются после финального гласного основы: *koto*- ‘плавать’ → AOR.S1 *kotaado?*, AOR.S3 *kotaa*, *d’ogo*- ‘отделиться’ → AOR.R3 *d’ogoeđo?*.

В словарных материалах, собиравшихся в 1970-е годы Е. А. Хелимским по тундровому энецкому и в 1990-е — нами по лесному энецкому, найдено 28 непереходных глаголов, образующих аорист первым способом, из них 20 имеют соответствие в нганасанском и ненецком, четыре — в ненецком, но не в нганасанском (но три из них, вероятнее всего, общесамодийские), четыре имеют соответствие в нганасанском, но не в ненецком.

Найдено 35 непереходных глаголов, образующих аорист вторым способом, из них 28 имеют соответствие только в ненецком (четыре — общесамодийские), три имеют соответствия в обоих языках, два — только в нганасанском, два не имеют соответствий.

При этом все переходные глаголы на *-a* и *-o* в энецком и все суффиксальные глаголы образуют аорист неагглютинативно: *bata*- ‘вылить’ → S3 *bata*, то есть этот способ образования аориста сохраняет в энецком продуктивность. Также следует отметить, что ни в одном самодийском языке правила образования аориста не различаются для переходных и непереходных глаголов. Почему же в энецком у ряда непереходных глаголов на *-a* и *-o* аорист образуется путем прибавления гласного *-a*? Можно обнаружить вторичную связь этого алломорфа аористного показателя с непереходностью. Агглютинативный алломорф *-a* выступает только в глаголах, в которых этому показателю предшествует узкий гласный: *-i* либо *-u*. При этом в энецком все глаголы на *-u* и подавляющее большинство глаголов на *-i* являются производными глаголами со стативно-декаузативной семантикой:

(*kae*- ‘оставить’ →) *kai*- ‘остаться’ → S3 т. эн., л. эн. *kai-a*;

(*bano*- ‘лечь (об олене)’ →) *banu*- ‘лежать (об олене)’ → S3 т. эн., л. эн. *banu-a*.

Стативно-декаузативные глаголы все — непереходные, таким образом, формируется контекстуальная связь алломорфа аористного показателя *-a* с непереходностью, что, вероятно, и способствует его распространению как аористного показателя в группе непереходных *a*-основ и *o*-основ.

Существенно обратить внимание на то, что в группе непереходных *a*-основ и *o*-основ этот способ образования аориста распространяется прежде всего на глаголы, имеющие параллель в ненецком, но не в нганасанском. То есть аналогическое выравнивание аористой парадигмы непереходных глаголов по парадигме стативно-декаузативных непереходных глаголов сопровождается только «освоение» отсутствовавших в субстратном языке (нганасанского типа) глаголов ненецкого языка и не распространяется на основы, представленные в субстратном языке.

Важно также обратить внимание на случаи, когда глагол является общесамодийским, но отсутствует в нганасанском. Таких случаев пять, и в четырех из них энецкий глагол образует аорист не этимологическим, а аналогическим способом:

tuja- ‘разлиться (о водоеме)’ → AOR.S3 *tujaa*;

pora- ‘сгореть’ → AOR.S3 *porea*;

teka- ‘спрятаться’ → AOR.S1 *tekaado?*;

т. эн. *sôa-*, л. эн. *suja-* ‘родиться’ → AOR.S1 т. эн. *sôaado?*, л. эн. *sujaado?*;

rida- ‘быть похожим’ → AOR.S3 *rida*.

Это возможно объяснить, только если принять гипотезу о смене языка: если бы энецкий эволюционировал независимым образом как один из диалектов северносамодийского праязыка, эти глаголы, сохранившиеся от праязыкового состояния, не сменили бы тип образования аориста. Но, если в период перехода с «языка типа нганасанского» на «язык типа ненецкого» в «языке типа нганасанского» эти глаголы были утрачены, это означает, что в энецком они являются вторичным приобретением и потому образуют аорист аналогическим образом.

10.2. Система причастных форм

Основные причастные формы северносамодийских языков представлены в таблице 18; если показатели являются этимологически тождественными, разделитель в таблице не ставится. Система форм лесного энецкого совпадает с системой тундрового ненецкого, а система форм тундрового энецкого — с системой форм нганасанского.

Таблица 18

Причастия в северносамодийских языках

	Ненецкий	Лесной энецкий	Тундровый энецкий	Нганасанский
Пассивное причастие прошедшего времени	<i>-mi°</i>	<i>-mi ~ i</i>	<i>-mi ~ i</i>	<i>-Məθ</i>
Активное причастие прошедшего времени			<i>-si-i (< -se-i)</i>	<i>-SUə(-D'əθ)</i>
Причастие настоящего времени	<i>-na</i>	<i>-nta</i>	<i>-de</i>	<i>-NTUə</i>
Причастие будущего времени	<i>-manta</i>	<i>-mode</i>	<i>-i'udo</i>	<i>?SUTə</i>

Обсудим различия причастных форм прошедшего времени. В ненецком и лесном энецком общесеверносамодийское причастие с показателем **-mā-jā* (> т. нен. *-mi°*, эн. *-mi*, нг. *-Məθ*) является залогово нейтральным, в тундровом энецком и нганасанском это причастие имеет только пассивное значение, в активном употребляется причастие нг. *-SUə-D'əθ*, эн. *-se-i*. Лексикализовавшиеся формы второго причастия (с показателем т. нен. *-sey°*, соответствующим т. эн. *-se* в *-se-i*, нг. *-SUə* в *-SUə-D'əθ*) при этом обнаруживаются и в ненецком: *temta-sey°* ‘покупка, купленная вещь’, *pya-sey°* ‘пай, доля добычи’ (< *pya-* ‘начать’), *talye-sey°* ‘кража’, *xanye-sey°* ‘добыча’ (< *xanye-* ‘охотиться’), *pire-sey°* ‘варево’. Это означает, что причастие т. нен. *-sey°*, т. эн. *-se*, нг. *-SUə* было равным образом присуще всем трем северносамодийским языкам. Соответственно, система причастных форм прошедшего времени тундрового энецкого и нганасанского более архаична, и устранение показателя т. нен. *-sey°* из системы причастных форм — это собственно ненецкая инновация (вторично распространившаяся на лесной энецкий). Примечательно, что в лесном энецком

таких лексикализовавшихся форм нет, это подтверждает, что система претеритальных причастных форм ненецкого типа в этом языке — вторичное контактное явление: если в ненецком это причастие выходило из употребления постепенно, с сохранением некоторых форм в качестве лексикализовавшихся, в лесном энецком такие формы лексикализуются, по всей видимости, не успели.

Различия в причастных формах будущего времени можно прокомментировать схожим образом. Во-первых, показатели эн. *-t'udo*, нг. *-ʔSUTə* (< **-tʰ-psutə*) имеют общесамодийское происхождение, ср. селькупское дебитивное причастие *-psotəl'*, тогда как показатели т. нен. *-manta*, л. эн. *-mode* не имеют когнатов в других языках. Во-вторых, в тундровом ненецком есть лексикализовавшиеся формы *ɲædabcod*²⁰ ‘то, во что можно выстрелить, цель’, *pyiryebcod*²¹ ‘то, что предназначено для варевa’, *xayebcod*²² ‘то, что уезжающему следует оставить на память тем, кто остается’, *xetʰbcod*²³ ‘то, что следует сказать’ и т. п. В-третьих, подобных лексикализовавшихся форм в лесном энецком не обнаружено. Эти факты позволяют нам выстроить ту же цепочку рассуждений, что и для претеритального причастия: система, представленная в тундровом энецком и нганасанском, более архаична, появление причастия на *-manta* в тундровом ненецком — собственно ненецкая инновация, энецкий не затронувшая, появление этого причастия в лесном энецком имеет вторичную контактную природу.

10.3. Эвиденциальный показатель инферентива-репортатива

В ненецком представлена единая инферентивно-репортативная форма с показателем *-wi*²⁴, восходящая к перфективному причастию. В энецком также представлена форма с инферентивно-репортативным значением, но ее показатель *-bi* не соответствует ненецкому. В нганасанском семантические оппозиции внутри эвиденциальной зоны устроены иначе: в нганасанском есть специализированная форма инферентива *-HA-TU* и два репортативных показателя, из которых *-HA-MHU* используется в утвердительных предложениях, а *-HA* — в вопросительных (значение этого показателя — ‘меня послали спросить, верно ли Р’ либо ‘верно ли, что, как говорят, Р?’). Вопросительная форма репортатива оказывается морфологически более простой, чем утвердительная форма; в [Гусев 2006] было показано, что именно *-HA* и было исходным суффиксом эвиденциальных форм. Этот суффикс в указанной работе возводится к праформе **-pe* и сопоставляется с энецким *-bi*. Итак, энецкий показатель может быть унаследован из субстрата нганасанского типа. Трактовка его как вторичной контактной энецко-нганасанской инновации маловероятна²⁵. В этом случае следовало бы предполагать, что более ранняя энецкая система эвиденциальных форм совпадала с ненецкой и в качестве инферентивно-репортативной формы использовалось перфективное причастие с показателем **-mā-jā*. Однако выше при анализе системы причастных форм мы постарались показать, что более архаичная система причастных форм в энецком, вероятнее всего, совпадала с нганасанской, а в этой системе причастие с показателем **-mā-jā* было ограничено пассивными употреблениями (в отличие от ненецкого, где оно нейтрально в отношении залога). При таком залоговом ограничении маловероятно, что эта форма успела бы грамматикализироваться в инферентив-репортатив активного залога, но позже полностью исчезнуть из употребления.

²² Как и с другими архаизмами, из двух возможных объяснений (независимое сохранение архаизма энецким и нганасанским vs. наследование формы энецким из субстрата нганасанского типа) мы выбираем второе потому, что оно представляется нам более вероятным на фоне других фактов субстратного влияния, проанализированных в статье.

10.4. Формы с конъюнктивно-оптативной семантикой

В ненецком и энецком представлен общесамодийский показатель конъюнктива: т. нен. -*yí*, л. эн. -*ñi*, т. эн. -*i* / -*ʔi* / -*ñi*, в нганасанском этот показатель конъюнктива не сохранился, однако представлены две другие формы с интересующим нас значением: конъюнктив с показателем -*HA*₁-*A*₀ (4) и дубитатив с показателем -*LI*. Дубитатив употребляется в значении авертива, т. е. в значении ‘смотри не сделай что-л.’ (5)–(6), в значении эпистемической возможности (7), в риторических вопросах о будущем (8)–(9)²³:

- (4) *aiʔ kəta-baa-m tənə kəta-baa-m.*
EXCL поймать-OPT-1SG ты.ACC поймать-OPT-1SG
‘Эх, поймать бы мне тебя!’ [Wagner-Nagy 2018: 247]
- (5) *бызə дя чалəбə-лы-у!*
вода.GEN к упасть-DUB-S2
‘Смотри, в воду не кувыркнись!’
- (6) *динда-ʔку-мə нейкə-лы-ры-аи?*
лук-DIM-ACC.1 испортить-DUB-O2PL-EXCL
‘Смотрите, лучок-то мой не испортите!’
- (7) *əmə-ni-ə tə kəriʔə-nə дязыкү-мəну сиркə-уы-ры?*
ЭТОТ-LOC-ADJ ВОТ хорей-GEN.1SG К-PROL ВЫКОПАТЬ-IMP-2PLS/O
маа-балə или-Ø.
что-то быть-DUB.S3
‘Копайте по направлению того, как хорей мой воткнулся, что-то там будет, полагаю’.
- (8) *Ку-дү-ми ня-мту уəтə-лы-Ø?*
который-SPEC-1DU товарищ-ACC.3SG найти-DUB-S3
[Давай играть в прятки!] ‘Который же из нас первым товарища найдет?’
- (9) *Тə хуə-гүə-нə куə-ʔ ни-лы-ри, маа-м-сутə-Ø?*
а позади-то-LAT.1 умереть-CN NEG-DUB-S2DU что-VDN-FUT-S3
‘А после меня вы разве не умрете, что [иное] произойдет?’

В энецком, помимо общесамодийского конъюнктива (как в ненецком), есть две редкие формы: форма оптатива с показателем -*bi* (сохранилась только у отрицательного глагола в лесном энецком, см. (10)–(11)) и форма с показателем -*ri*. Показатель -*ri* зафиксирован в неопубликованных полевых грамматических материалах Е. А. Хелимского, в самостоятельном употреблении в его материалах он переведен как авертив (12), как оптатив (13), в сочетании с показателем претерита — как вопрос к будущему (14). В материалах по лесному энецкому встретилось однократное употребление в составе бытийного глагола (парадигма которого может, вероятно, сохранять некоторые архаизмы) в сочетании с показателем претерита в значении вопроса о будущем (15):

- (10) *дязуй-да меон и-би-з кань*
след-GEN.3 В.PROL NEG-OPT-S1 пойти-CN
‘по его следам пойдут-ка’ [Сорокина, Болина 2005: текст 45, предл. 30]
- (11) *Тазыбе-зи, мана, и-би-ра таза-’*
шаман-DESTPL.1 сказать.AOR.S3 NEG-OPT-S2PL дать-CN
‘Шаманов мне, говорит, привели бы’. [Сорокина, Болина 2005: текст 6, предл. 19]

²³ Примеры без ссылки на источник выбраны из нганасанских текстов, любезно предоставленных в мое распоряжение В. Ю. Гусевым.

- (12) *d'u-li-bo*
потерять-DUB-01
'не потерять бы мне'
- (13) *moga-xone a-ri-do?*
лес-LOC быть-DUB-S1
'хорошо бы мне быть в лесу'
- (14) *s'i-i? koma-ri-Ø-s'i*
гроз.ACC-1 любить-DUB-S3-PRAET
'будет ли он меня любить?'
- (15) *Онэй нээ тале моды-э-за, кунь э-ри-Ø-сь*
настоящий женщина воровать.INF смотреть-AOR-03 как быть-DUB-S3-PRAET
'Энка украдкой на нее (= на ведьму) смотрит, что дальше будет'. [Сорокина, Болина 2005: текст 12, предл. 60]

Энецкий показатель *-bi*, очевидно, можно отождествить по употреблению с нганасанским показателем опатива *-HA₁-A₀*, энецкий показатель *-ri* — с нганасанским показателем дубитатива *-li*. В результате имеем следующую систему форм.

Таблица 19

Система конъюнктивно-опативных форм северносамодийских языков

Конъюнктив	—	<i>-bi</i> (л. эн.)	<i>-HA₁-A₀</i>
	<i>-yi</i>	<i>-ni</i>	—
Дубитатив	—	<i>-ri</i>	<i>-li</i>

В энецком эти показатели естественнее трактовать как представляющие субстрат нганасанского типа. Показатель опатива *-bi* является морфологическим реликтом, сохранившимся только в парадигме отрицательного глагола в единственной идиоматизированной конструкции 'не сделать ли Р, хорошо бы сделать Р', то есть он во всяком случае не выглядит более молодым элементом системы в сравнении с общесамодийским конъюнктивом, продуктивным образом выражающим это значение в современном энецком. Показатель дубитатива, судя по всему, сохраняет в энецком также только реликтовые употребления. При этом возможность его сочетания с показателем претерита в обоих энецких диалектах (см. (14) и (15)) позволяет рассматривать этот показатель как субстратный элемент, получивший в энецком самостоятельное развитие в конструкции с адстратным постфиксальным показателем претерита, а не как позднее копирование нганасанской конструкции (в нганасанском постфиксального показателя претерита нет).

11. Периферия грамматической системы:
нганасанское субстратное влияние

11.1. Личные местоимения

В энецком языке личные местоимения 1 и 2 лица с высокой вероятностью близки к тем же праформам, которые имеют в нганасанском личные местоимения эмфатической серии (ср. в особенности формы неединственного числа типа эн. *mod'ini?* 'мы двое' < **mən-siə-nij*, нг. *mjñšianic* 'мы двое' < **men-siə-nij*), см. таблицу 20.

Таблица 20

Личные местоимения северносамодийских языков

	Ненецкий	Энецкий	Нганасанский	
			эмф.	прост.
1SG	<i>məny^o</i>	<i>mod'i</i>	<i>mjñśiənə</i>	<i>mənə</i>
1DU	<i>mənyih</i>	<i>mod'ihī?</i>	<i>mjñśiəni^c</i>	<i>mii</i>
1PL	<i>mənyaq</i>	<i>mod'ina?</i>	<i>mjñśiəni?</i>	<i>mij</i>
2SG	<i>pir^o ~ pədər^o ~ pəd^oyer^o24</i>	<i>tod'i</i>	<i>tjñśiətə</i>	<i>tənə</i>
2DU	<i>pəd^oryih piry^oh</i>	<i>tod'idi?</i>	<i>tjñśiəti^c</i>	<i>tii</i>
2PL	<i>pəd^oraq piraq</i>	<i>tod'ida?</i>	<i>tjñśiəti?</i>	<i>tij</i>

Вместе с тем энецкий демонстрирует некоторые отличия от нганасанского, которые говорят скорее об определенном периоде независимого развития форм в энецком и нганасанском, чем о недавнем копировании, что позволяет нам относить эти местоимения в энецком к субстратному слою. Ниже приведены праформы местоимений в нганасанском и энецком и прокомментировано их развитие.

1SG: нг. **men-siə-nə*, эн. **mən-siə* (в нганасанском используется специальная основа личного местоимения с особой огласовкой, в энецком отсутствует лично-притяжательный показатель²⁵, в суффиксе произошло сокращение дифтонга).

1DU: нг. **men-siə-nij*, эн. **mən-siə-nij*.

1PL: нг. **men-siə-năt*, эн. **mən-s' iə-năt*.

2SG: нг. **ten-siə-ntə*, эн. **tən-siə* (в нганасанском используется специальная основа личного местоимения с особой огласовкой, в энецком отсутствует лично-притяжательный показатель).

2DU: нг. **ten-siə-ntij*, эн. **tən-siə-ntij*.

2PL: нг. **ten-siə-ntăt*, эн. **tən-siə-ntăt*.

11.2. Вопросительные и указательные местоимения

В энецком представлено несколько основ вопросительных локативов и дейктических / анафорических местоимений:

²⁴ Фонетические варианты.

²⁵ Это можно объяснить как раз переходом на язык типа ненецкого: в ненецком местоимения 1DU и 1PL так же, как и в энецком, по-видимому, образованы от местоимения 1SG, что показывает форма с лимитативной интраклатикой *ryi* т. нен. *məny^o-ryi-naq* 'только мы', где *məny^o* совпадает с местоимением т. нен. *məny^o* 'я'.

а) чистая основа:

Праформа	т. нен.	т. эн.	л. эн.	нг.
вопр. * <i>ku-</i>	<i>xu-d</i> ° ‘откуда?’	<i>ku-do</i> ‘откуда?’		<i>ku-tə</i> ‘откуда?’
дальний дейксис * <i>tä-</i>	<i>təh</i> ‘тот’	<i>too(?)</i> ‘туда’ <i>too</i> ‘тот’		<i>təə</i> ‘тот’
ближний дейксис * <i>äm(ä)-</i>		<i>eo?</i> ‘этот’	← <i>əmə</i> ^{-C} ‘туда’	
		—	<i>otoone</i> ‘по этому месту’	<i>əmə-ənĭ</i> ‘по этому месту’

б) основа + *-*t³kä(jä)*

Праформа	т. нен.	эн.	нг.
вопр. * <i>ku-t³kV</i>		<i>ku-ke</i> ‘кто, который?’ <i>ku-ko-xorio-ne</i> ‘хоть где’	
дальний дейксис * <i>ti-t³kV</i>	<i>tyi-ki</i> ° ‘тот’ <i>tyi-ka-xəd</i> ° ‘оттуда’	<i>t'i-ke</i> ‘тот’ <i>t'i-ko-xodo</i> ‘оттуда’	
ближний дейксис * <i>äm(ä)-t³kV</i> * <i>tü-kV</i>		<i>e-ke</i> ‘этот’ <i>e-ko-xodo</i> ‘оттуда’	
	<i>tyu-ku</i> ° ‘этот’, <i>tyu-ko-</i> <i>xəd</i> ° ‘оттуда’		

в) основа + *-*ñi(ä)*

Праформа	т. нен.	т. эн.	л. эн.	нг.
вопр. * <i>ku-ñi-</i>		<i>ku-u=no-ne</i> ‘где?’	<i>ku-u=ne-ne</i> ‘где?’ <i>ku-ni-ne</i> ‘где?’	← <i>ku-ñi-ni</i> ‘где?’
дальний дейксис * <i>tä-ñi-</i>		<i>te-i=no-ne</i> ‘там’	<i>te-i=ne-ne</i> ‘там’ <i>to-ni-ne</i> ‘там’	← <i>tä-ñi-ni</i> ‘там’
ближний дейксис * <i>ämä-ñi-</i>		<i>i-i=no-ne</i> ‘там неподалеку’	<i>i-i=ne-ne</i> ‘там неподалеку’	<i>ämä-ñi</i> ^{-C} ‘туда в сторонку’

Серия а) является наиболее архаичной. Серия б), очевидно, возникла в энецком под вторичным влиянием ненецкого: во-первых, мы видим, что этот суффикс аналогически распространился на большее число форм, чем в ненецком, во-вторых, в энецком он фонетически более регулярен (регулярна огласовка суффикса как в номинативе, так и в формах косвенных падежей, за счет чего нет прямого исторического соответствия ненецких и энецких форм). Серия в), очевидно, отражает субстратное влияние нганасанского: в ненецком подобных форм нет. Прокомментируем несколько наиболее интересных моментов:

1) Следует обратить внимание на общий для нганасанского и энецкого корень местоимения близкого дейксиса **äm(ä)-*, отсутствующий в ненецком. На то, что этот элемент является субстратным, указывает, во-первых, особое фонетическое развитие форм с этим корнем в энецком (анлаутный *e-*, *i-* вместо регулярного *o-*), что было бы невозможно при вторичном заимствовании. Во-вторых, в местоимениях серии в) мы видим специфически энецкую морфологическую структуру местоимений (т. эн. *i-i=no-ne*, л. эн. *i-i=ne-ne* ‘там неподалеку’ — из стянувшейся послеложной конструкции), что также указывает на определенный возраст этих форм в энецком.

2) На то, что серия б) отражает вторичное влияние ненецкого адстрата, указывает также местоимение *e-ke* ‘этот’, *e-ko-xodo* ‘оттуда’: исторически корректной была бы форма

e-ge**, *e-go-xodo** (< ***äm-kä-jä**, ***äm-kä-xätä**), однако в данном случае суффикс прибавляется ко вторичной, вычлененной по аналогии основе, представленной в других местоимениях с этим корнем в энецком.

3) На то, что серия в) отражает нганасанское субстратное влияние, указывает прежде всего сложное фонетически нерегулярное развитие этих форм в энецком. Местоимения являются квазиграмматическими элементами, за счет чего затрагивающие их фонетические процессы могут отличаться от тех, которые характерны для основного массива лексики. Заметим, что, если в случае вторичной серии б) энецкие формы максимально регулярны (что отличает их от фонетически нерегулярных ненецких), в случае субстратной серии мы видим фонетически нерегулярные формы уже в самом энецком:

ku-ni-** > т. эн. *ku-u=no-ne*, л. эн. *ku-u=ne-ne* вместо ожидаемого т. эн. *ku-i=no-ne**, л. эн. ****ku-i=ne-ne**;

tā-ni-** > т. эн. *te-i=no-ne*, л. эн. *te-i=ne-ne* вместо ожидаемого т. эн. *to-i=no-ne**, л. эн. ****to-i=ne-ne**;

äm-ni-** > т. эн. *i-i=no-ne*, л. эн. *i-i=ne-ne* вместо ожидаемого т. эн. *e-i=no-ne**, л. эн. ****e-i=ne-ne**.

4) В лесном энецком сохранились возникшие под вторичным нганасанским влиянием формы, которые не демонстрируют никаких нерегулярностей (такие формы обозначены в таблице стрелками). Наиболее показательна форма л. эн. *отоопе* ‘по этому месту’ (соотв. нг. *әтә-әңі*), в которой корень ***äm(ä)-** уже имеет регулярный анлаутный *o-*.

Местоимения ‘какой?’, ‘такой’ имеют следующий вид в северносамодийских языках:

	т. нен.	л. эн.	т. эн.	нг.
‘какой?’	<i>xurka</i>	<i>kurse</i> ↑	<i>kuroe</i>	<i>kuräd'i</i>
‘такой’	<i>tärcya</i> →	<i>torse</i> ↑	<i>toroe</i>	<i>täräd'i</i>

Местоимения ‘какой?’, ‘такой’ в лесном энецком, вероятнее всего, возникли под поздним сепаратным влиянием тундрового ненецкого. Во-первых, кластеры не должны были бы сохраниться, если бы эти формы представляли собой результат исторически регулярного развития общей с ненецким праформы. Во-вторых, очевидно, что из ненецкого было заимствовано местоимение т. нен. *tärcya* > л. эн. *torse* ‘такой’, затем в энецком из этого местоимения был вычленен суффикс л. эн. *-rse*, с помощью которого по аналогии было образовано местоимение *kurse* ‘какой?’: суффикс л. эн. *-rse* соответствует т. нен. *-rcya*, который, в свою очередь, является выступающим в переднерядных основах вариантом суффикса т. нен. *-rka*. Однако основа местоимения ‘какой’ является этимологически заднерядной (ср. этимологически корректное т. нен. *xurka*). Напротив, т. эн. *toroe* и *kuroe* точно соответствуют нг. *täräd'i* и *kuräd'i*, так что можно предполагать, что в общезнецком были представлены именно эти местоимения нганасанского субстратного типа, которые позже были вторично заменены в лесном энецком.

Вопросительное местоимение ‘что?’ имеет следующий вид в северносамодийских языках:

	т. нен.	л. эн.	т. эн.	нг.
	<i>әәм-ке</i>			
‘что?’	<i>әәмка-рыі</i> (ср. также <i>әәwoл/су°</i> ‘для какой цели служить’)	<i>obu</i>	<i>mii?</i> (? < r)	<i>maa</i>

Местоимение ‘что?’ лесного энецкого родственно ненецкому, однако выглядит как вторичная адаптация (в ходе смены языка) основы GEN.PL ненецкого местоимения: л. эн. *obu* ←

т. нен. *ḡəwoq* (форма GEN.PL от *ḡəm-ke*). Параллелизм основы т. нен. и л. эн. обнаруживается также в л. эн. *obu-ru* [что-LIM] ‘вещи’, ср. т. нен. *ḡəwoyiq* [что-LIM-GEN.PL] ‘вещи’. Местоимение т. эн. *miiʔ* ($\text{ʔ} < r$) ‘что?’ явно родственно нг. *maa* ‘что?’, нганасанский отражает праформу **meḡ*, т. эн. — праформу **meḡ-r* с дополнительной перегласовкой **ə > i* перед ауслатным суффиксальным *r*. Различие морфологической структуры местоимений нганасанского и тундрового энецкого заставляет рассматривать эту параллель между двумя идиомами не как случай вторичного контактного влияния нганасанского, а как сохранение общего архаизма — в противовес сближению лесного энецкого с ненецким, о вторичности которого свидетельствует адаптация энецким морфонологически и морфологически более удобной формы местоимения GEN.PL²⁶.

12. Заключение

Можно заметить, что практически никогда историческая фонетика не помогала нам отделить в энецком ни нганасанские субстратные явления от ненецких адстратных, ни результаты вторичного влияния нганасанского от хронологически более ранних явлений: для всех рассмотренных материальных единиц выполняются все исторические фонетические законы, которые мы знаем для энецкого. В связи с этим интересно следующее наблюдение: Е. А. Хелимский рассказывал в личной беседе, что один из лучших его информантов по тундровому энецкому, Холю Николаевич Каплин, также в совершенстве владевший ненецким, иногда мог выдавать «фантомные» энецкие слова, существование которых не подтверждали потом другие квалифицированные носители тундрового энецкого. Дело в том, что Е. А. Хелимский собирал от Х. Н. Каплина лексические материалы, причем Холю Николаевичу (который не владел свободно русским языком) было проще работать не с русскими стимулами, а с ненецкими (из словаря [Терещенко 1965]). Обсуждающиеся «фантомные» слова возникали как побочный результат такой работы: как оказалось, для Х. Н. Каплина было естественно пересчитать, как фонетически должна была бы выглядеть энецкая форма, соответствующая ненецкому стимулу (хотя в реальности такого энецкого слова не было).

При этом в завершение статьи нам хотелось бы обсудить один пример, показывающий, что в ситуации такого предполагаемого «билингвизма с пересчетом» изредка даже лексическая и грамматическая части одного и того же слова могут быть «обработаны» различным образом. Разберем случай с названием мамонта (см. также [Кахейнен 2022], где обсуждается история этого слова в самодийских языках). Т. эн. *kari-a* ‘живой мамонт’ соответствует нг. *kala-ḡu* ‘живой мамонт’, а т. эн. *kari-i* ‘мамонтовая кость’ соответствует нг. *kala-d’əḡ* ‘мамонтовая кость’. Суффиксы обоих слов, безусловно, проще объяснить в нганасанском, чем в энецком, ср. суффиксы в нг. *hulu-ḡu* ‘корпус нарты’, *horə-ḡu* ‘объект ругани, оспариваемый объект’, *mənə-ḡu* ‘шаманское пение, песенное камлание’ (при *mənə-mu* ‘совет, рекомендация’) и суффикс именного претерита в *bəbə-d’əḡ* ‘бывшее место (на котором прежде находился какой-то предмет, раньше сидел человек, лежал олень и т. п.)’: в нганасанском первый суффикс непродуктивен, второй продуктивен, в энецком первого суффикса нет вообще, а второй непродуктивен.

Парадокс этих слов состоит в том, что соответствия гласных в корневой части т. эн. *kari-* и нг. *kala-* выглядят как очень поздние (исторического соответствия т. эн. *a* = нг. *a* в первом слоге нет). Однако соответствия в суффиксальной части т. эн. *-a* = нг. *-ḡu* и т. эн. *-i* = нг. *-d’əḡ* выглядят как исторически корректные, подтверждаясь и остальным массивом

²⁶ Этимология обеих основ также говорит в пользу именно такой трактовки: местоимение **ḡm-* имеет ненадежную параллель в камасинском [Janhunen 1977: 17], местоимение **meḡ-* имеет надежные общесамодийские параллели [Ibid.: 17].

лексики. В [Кахейнен 2022] высказывается предположение, что энецкое слово заимствовано из нганасанского (в нганасанском, действительно, историческим продолжением самодийского слова со значением ‘мамонт’ является нг. *koliinj* ‘кит’). Однако этому противоречат два обстоятельства: во-первых, оснований предполагать производность этих слов на энецкой почве нет — в энецком нет оснований выделять подобные суффиксы, тогда как из нганасанского материала их производность объясняется вполне естественно. Во-вторых, надежных примеров заимствований из (тундрового) энецкого в нганасанский нет. Так что, каким бы ни был источник нганасанского слова, в энецком это слово выглядит так, что корень, образно говоря, представляет собой случай заимствования без пересчета вокализма, а суффиксы — случай заимствования с пересчетом вокализма. Такие парадоксы, вероятно, порождаются ситуацией билингвизма.

СПИСОК СОКРАЩЕНИЙ

1, 2, 3 — 1, 2, 3 лицо	IMP — императив
ACC — аккумулятив	INF — инфинитив
ADJ — атрибутивный показатель	IPFV — имперфектив
ADV — показатели падежей, употребляющиеся в наречиях и послелогах	LAT — латив
AOR — аорист	LOC — локатив
CARIT — каритив	NEG — отрицательный глагол
CN — коннегатив	O — объектное спряжение
COND — условное дееспричастие	OPT — оптатив
DEST — дестинатив	PL — множественное число
DIM — диминутив	PRAES — настоящее время
DU — двойственное число	PRAET — прошедшее время
DUB — дубитив	PROL — пролатив
DUR — дуратив	PRON — личное местоимение
EL — аблатив	R — рефлексивное спряжение
EXCL — восклицание	S — субъектное спряжение
FUT — будущее время	SG — единственное число
GEN — генитив	SPEC — выделительная форма ‘один из нескольких’
HAB — хабиуталис	VDN — суффикс, образующий отыменные глаголы

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Гусев 2006 — Гусев В. Ю. О сохранении архаичных форм в неассертивных контекстах: материал самодийских языков. *Проблемы типологии и общей лингвистики. Международная конференция, посвященная 100-летию со дня рождения А. А. Холодовича: материалы*. Храковский В. С., Дмитренко С. Ю., Заика Н. М. (ред.). СПб.: Нестор-История, 2006, 41–45. [Gusev V. Yu. On retainment of archaic forms in non-assertive contexts: Samoyedic data. *Problemy tipologii i obshchei lingvistiki. International conf. on the 100th birthday of A. A. Kholodovich: Proceedings*. Xrakovskij V. S., Dmitrenko S. Yu., Zaika N. M. (eds.). St. Petersburg: Nestor-Istoriya, 2006, 41–45.]
- Гусев 2022 — Гусев В. Ю. Рефлексы прасамодийских *a и *ä и нганасанский сингармонизм. [Gusev V. Yu. Reflexes of Protosamoyedic *a and *ä and Nganasan vowel harmony.] *Sibirica et Uralica: In Memoriam Eugene Helinski*. Anikin A., Gusev V., Urmancieva A. (eds.). Szeged: Univ. of Szeged, 2022, 39–62.
- Кахейнен 2022 — Кахейнен К. К вопросу о ранних контактах между самодийскими языками: заметки об истории нган. *колиин*. [Kaheinen K. On early contacts between Samoyed languages: Essays on the history of Nganasan *колиин*.] *Sibirica et Uralica: In Memoriam Eugene Helinski*. Anikin A., Gusev V., Urmancieva A. (eds.). Szeged: Univ. of Szeged, 2022, 77–94.
- Костеркина и др. 2001 — Костеркина Н. Т., Момде А. Ч., Жданова Т. Ю. *Словарь нганасанско-русский и русско-нганасанский*. СПб.: Просвещение, 2001. [Kosterkina N. T., Momde A. Ch.,

- Zhdanova T. Yu. *Slovar' nganasansko-russkii i russko-nganasanskii* [Nganasan-Russian and Russian-Nganasan dictionary]. St. Petersburg: Prosveshchenie, 2001].
- Кузьменко 1995 — Кузьменко Ю. К. Этнические контакты в Ютландии и языковые изменения в датских диалектах. *Этнолингвистические исследования. Этнические контакты и языковые изменения*. Бородин М. А., Кузьменко Ю. К. (ред.). СПб: Наука, 1995, 28–60. [Kuzmenko Yu. K. Ethnic contacts in Jutland and linguistic changes in Danish dialects. *Etnolingvisticheskie issledovaniya. Etnicheskie kontakty i yazykovye izmeneniya*. Borodina M. A., Kuzmenko Yu. K. (eds.). St. Petersburg: Nauka, 1995, 28–60.]
- Кузьменко 2011 — Кузьменко Ю. К. *Ранние германцы и их соседи. Лингвистика, археология, генетика*. СПб.: Нестор-История, 2011. [Kuzmenko Yu. K. *Rannie germantsy i ikh sosedi. Lingvistika, arkheologiya, genetika* [Early Germans and their neighbors. Linguistics, archeology, genetics] St. Petersburg: Nestor-Istoriya, 2011.]
- Лабанаускас 2002 — Лабанаускас К. И. *Родное слово. Энецкие песни, сказки, исторические предания, традиционные рассказы, мифы*. СПб.: Просвещение, 2002. [Labauskas K. I. *Rodnoe slovo. Enetskie pesni, skazki, istoricheskie predaniya, traditsionnye rasskazy, mify* [Native word. Enets songs, tales, historical legends, traditional stories, myths]. St. Petersburg: Prosveshchenie, 2002.]
- Напольских 2006/2018 — Напольских В. В. Булгарская эпоха в истории финно-угорских народов Поволжья и Предуралья. *Очерки по этнической истории*. Напольских В. В. Казань: Казанская недвижимость, 2018, 64–86. [Napolskikh V. V. The Bulgarian era in the history of the Finno-Ugric peoples of the Volga and Cis-Ural regions. *Ocherki po etnicheskoi istorii*. Napolskikh V. V. Kazan: Kazanskaya nedvizhimost', 2018, 64–86.]
- Сорокина, Болина 2005 — Сорокина И. П., Болина Д. С. *Энецкие тексты*. СПб.: Наука, 2005. [Sorokina I. P., Bolina D. S. *Enetskie teksty* [Enets texts]. St. Petersburg: Nauka, 2005.]
- Терещенко 1965 — Терещенко Н. М. (сост.) *Ненецко-русский словарь*. М.: Советская энциклопедия, 1965. [Tereshchenko N. M. *Nenetsko-russkii slovar'* [Nenets-Russian dictionary]. Moscow: Sovetskaya entsiklopediya, 1965.]
- Урманчиева 2020 — Урманчиева А. Ю. Деривационная морфология образных глаголов тундрового ненецкого языка. [Urmanchieva A. Yu. Derivational morphology of depictive (onomatopoeic) verbs in Tundra Nenets.] *Acta Linguistica Petropolitana*, 2020, 16(3): 570–612.
- Филиппова 2011 — Филиппова Т. М. Особенности верхнеобского говора южных селькупов (тайзаковско-старосондровская группа). *Вестник НГУ. Серия: история, филология*, 2011, 2 (10): 81–87. [Filipova T. M. Peculiarities of the Upper Ob subdialect of the Southern Selkups (Tayzakovo-Starosondorovo group). *Vestnik NGU. Series: History and Philology*, 2011, 2(10): 81–87.]
- Хелимский 1994 — Хелимский Е. А. Очерк морфонологии и словоизменительной морфологии нганасанского языка. *Таймырский этнолингвистический сборник*. Хелимский Е. А. (ред.). М.: РГГУ, 1994, 190–221. [Helimski E. A. Essay on morphonology and inflectional morphology of the Nganasan language. *Taimyrskii etnolingvisticheskii sbornik*. Helimski E. A. (ed.). Moscow: Russian State Univ. for the Humanities, 1994, 190–221.]
- Хелимский 1995/2000 — Хелимский Е. А. Proto-Uralic gradation: Continuation and traces. *Компаративистика. Уралистика: Лекции и статьи*. Хелимский Е. А. М.: Языки русской культуры, 2000, 167–190. [Helimski E. A. Proto-Uralic gradation: Continuation and traces. *Komparativistika. Uralistika: Lektsii i stat'i*. Helimski E. A. Moscow: Yazyki russkoi kul'tury, 2000, 167–190.]
- Alatalo 2004 — Alatalo J. *Sölkupisches Wörterbuch aus Aufzeichnungen von Kai Donner, U. T. Sirelius und Jarmo Alatalo*. Helsinki: Suomalais-Ugrilainen Seura, 2004.
- Castrén 1855 — Castrén M. A. *Wörterverzeichnisse aus den samojedischen Sprachen*. St. Petersburg: Be-druckerei der Kaiserlichen Akademie der Wissenschaften, 1855.
- Donner 1932 — Donner K. *Samojedische Wörterverzeichnisse*. Helsinki: Suomalais-Ugrilainen Seura, 1932.
- Donner, Joki 1944 — Joki A. J. *Kai Donners kamassisches Wörterbuch nebst Sprachproben und Hauptzügen der Grammatik*. Helsinki: Suomalais-ugrilainen seura, 1944.
- Helimski 1997 — Helimski E. *Die matorische Sprache*. Szeged: Szeged Univ., 1997.
- Helimski 1997/2022 — Helimski E. Zur Stellung des Matorischen innerhalb der samojedischen Sprachen. *Siberica et Uralica: In memoriam Eugen Helimski*. Gusev V., Urmanchieva A., Anikin A. (eds.). Szeged: Szeged Univ., 2022, 479–495.
- Helimski 2005 — Helimski E. The 13th Proto-Samoyedic vowel. *Mikola-konferencia 2004*. Wagner-Nagy B. (ed.). Szeged: SzTE Department of Finnougristics, 2005, 15–26.
- Janhunen 1977 — Janhunen J. *Samojedischer Wortschatz: gemeinsamojedische Etymologien*. Helsinki: Suomalais-Ugrilainen Seura, 1977.

- Janhunen 1998 — Janhunen J. Samoyedic. *The Uralic languages*. Abondolo D. (ed.). London; New York: Routledge, 1998, 457–479.
- Katz 1987 — Katz H. Zur Phonologie des Motorisch-Karagassisch-Taigischen. *Studien zur Phonologie und Morphologie der uralischen Sprachen. Akten der dritten Tagung für uralische Phonologie, Eisenstadt, 28. Juni — 1. Juli 1984*. Rédei K. (ed.). Wien: Verband der wissenschaftlichen Gesellschaften Österreichs, 1987, 336–348.
- Khanina 2021 — Khanina O. Language and ideologies at the Lower Yenisei (Siberia): Reconstructing past multilingualism. *International Journal of Bilingualism*, 2021, 25(4): 1059–1080.
- Lehtisalo 1956 — Lehtisalo T. *Juraksamojedisches Wörterbuch*. Helsinki: Suomalais-Ugrilainen Seura, 1956.
- Salminen 1998 — Salminen T. *A morphological dictionary of Tundra Nenets*. Helsinki: Suomalais-Ugrilainen Seura, 1998.
- Wagner-Nagy 2018 — Wagner-Nagy B. *A grammar of Nganasan*. Leiden: Brill, 2018.

Получено / received 14.05.2024

Принято / accepted 21.05.2024

Information structure and syntax of Uralic languages: A foreword

© 2024

Elena K. Skribnik

Ludwig-Maximilians-Universität in Munich, Munich, Germany; skribnik@lmu.de

Daria O. Zhornik

Institute of Linguistics, Moscow, Russia;

Ludwig-Maximilians-Universität in Munich, Munich, Germany;

daria.zhornik@yandex.ru

Acknowledgements: The authors would like to express deep gratitude to Gerson Klumpp and Jeremy Bradley for their valuable comments.

DOI: 10.31857/0373-658X.2024.5.60-64

Информационная структура и синтаксис уральских языков: предисловие

Елена Константиновна Скрибник

Мюнхенский Университет Людвиг-Максимилиана, Мюнхен, Германия; skribnik@lmu.de

Дарья Олеговна Жорник

Институт языкознания РАН, Москва, Россия;

Мюнхенский Университет Людвиг-Максимилиана, Мюнхен, Германия;

daria.zhornik@yandex.ru

DOI: 10.31857/0373-658X.2024.5.60-64

This issue includes papers by specialists working on Uralic languages spoken in Russia from both descriptive and typological perspectives. Uralic languages as the second largest language family in Europe have been attracting the interest of linguists for several centuries already [Stipa 1990], but the level of their description differs greatly. On the one hand, there are long-standing grammaticography traditions for national languages like Finnish, Hungarian, and Estonian, partly with their own slightly idiosyncratic terminology [Bakró-Nagy et al. 2020]. On the other hand, for minority languages modern typological grammars are scarce, although in recent years there has been a rise in theoretical research on them, see, e.g., the recent volumes on Moksha Mordvin [Toldova, Kholodilova (eds.) 2018] and Hill Mari [Kashkin (ed.) 2023] published at the Department of Theoretical and Applied Linguistics of the Lomonosov Moscow State University as a result of numerous field trips, continuing the tradition established by A. E. Kibrik, see [Fedorova et al. (eds.) 2014]. Extensive data on these languages is also analyzed in the recent handbooks on Uralic languages [Bakró-Nagy et al. (eds.) 2022; Abondolo, Valijärvi (eds.) 2023].

The majority of papers in this issue represent what has been called “an information structural turn in Uralic studies” [Blokland et al. 2023: 241]. A striking feature of some Uralic languages is the significant influence of information structure on the expression of grammatical categories, such as voice, object agreement, and non-possessive uses of possessive markers. Of course, many older grammars took note of these phenomena, introducing different descriptive terms for them (e.g. “logical stress”, “definite conjugation”, “possessive suffixes as definite articles”, etc.). Nevertheless, the systematic study of the way Uralic languages use morphosyntactic means to express information structure started only after the development of the general information structure theory (see, e.g., [Lambrecht 1994; Krifka 2007; Aissen 2023]). In fact, it follows the more general tendency: from the definition of a linguistic sign as a form-semantics pair through the triangle of form, semantics, and function to the more recent understanding of it as a combination of form, semantics, syntactics, and pragmatics. Otherwise, a definition of, e.g., voice cannot be seen as complete if it is defined only as a rearrangement of syntactic functions and semantic roles, as it was done in the 1970s, without an explanation of its function.

The aforementioned “information structural turn” in Uralic studies began with the Ob-Ugric languages, as they present a unique case of grammatical systems heavily governed by information structure. The idea of pragmatic sentence organization with topichood as a central notion first appears in [Nikolaeva et al. 1993], followed by [Nikolaeva 1999], though foundations for it were laid out already in [Rombandeeva 1979] and [Kulonen 1989]. It is claimed that in Khanty and Mansi varieties a whole range of phenomena is largely determined by the information-structural status of the core situation participants. Among them are voice [Kulonen 1989; Skribnik 2001; Filchenko 2006; Virtanen 2015; 2023; É. Kiss 2019; Urmanchieva, Plungian 2021; Muravyev, Zhornik 2024], object agreement [Nikolaeva 1999; Koshkareva 2002; Skribnik 2004; Virtanen 2015; É. Kiss 2019; Klumpp 2023], differential object marking [Virtanen 2015; Klumpp 2023], syntax of ditransitive constructions [Virtanen 2015; Bíró, Sipőcz 2017], use of possessive markers in non-possessive functions [Janda 2015; 2019], and constituent order [Virtanen 2023]. This systemic combination of morphosyntactic means was aptly defined as “communicative sentence paradigm” in [Koshkareva 2002], see also similar notions in [É. Kiss 2019; Klumpp, Skribnik 2022].

This approach was promptly applied to other branches of Uralic. For Samoyedic, research on Enets ditransitive constructions [Khanina, Shluinsky 2020], Nganasan passive [Leisiö 2006], Nenets object agreement [Dalrymple, Nikolaeva 2011; Nikolaeva 2014] and passive [Stenin 2023] was carried out. An interesting attempt to combine a generativist approach with the information structure theory is presented in [Däbritz 2021], which analyzes the data of five North-West Siberian languages, three of them Uralic (Nganasan, Enets, Khanty).

Regarding Mordvin languages, studies on differential object marking (DOM) [Toldova 2017] (see also a more general work on DOM in Finno-Ugric [Serdobolskaya, Toldova 2017] and another one on the information structure of direct object in general [Serdobolskaya, Toldova 2016]) and passive in Moksha Mordvin [Stenin 2018] were performed. For Mari languages, we find recent works on DOM [Serdobolskaya 2015], discourse particles [Kozlov, Zakirova 2023], and possessive marking in non-possessive function [Khomchenkova 2022]. For Komi and Udmurt, publications appeared on the interaction of DOM with information structure and other parameters, see [Klumpp 2012; 2014] for Komi and [Serdobolskaya 2020] for Beserman Udmurt.

As for the Western-most Finno-Ugric languages, pragmatic sentence organization has long been considered relevant for Hungarian and Finnish grammaticography, mainly in connection to constituent order, cf. the characterisation of Hungarian as a “discourse configurational language” [É. Kiss 1995] and the description of “discourse conditioned word order” in Finnish [Vilkuna 1989]. However, as far as we know, such systemic descriptions are still lacking for Fennic and Saami languages of Russia.

Three papers presented here continue this approach. Nikita Muravyev’s article discusses the influence of internal possessors on information structure in Northern Khanty. The study deals with a number of contexts in Northern Khanty, in which the presence of an internal possessor

in the noun phrase of one of the core arguments influences the choice between active and passive voice; three hypotheses explaining this behavior are considered. The paper by Daria Zhornik considers different grammatical categories that are influenced by information structure in Northern Mansi, giving an overview of the previous research on them and suggesting additional parameters to complement the existing information-structural approaches to these phenomena. Johannes Hirvonen in his paper provides a first descriptive overview of information structure in Meadow Mari, focusing on the types of morphosyntactic marking which interact with the information structural notions of focus, topic, and contrastive topic. The results of this research confirm previous claims about Meadow Mari information structure but at the same time show a more varied picture of the available information structure marking strategies.

This issue also contains articles on Uralic syntax. Irina Khomchenkova's article is dedicated to conditional and concessive constructions with Russian conjunctions *esli* 'if' and *xotja* 'although' in Hill Mari speech. The author examines the possibility of duplication of the Russian conjunctions with the corresponding Hill Mari conjunctions and shows that it is possible for the conditional conjunctions, which exhibit incomplete structural congruence, but not for the concessive one. Finally, the issue includes Natalia Serdobolskaya's review of "Clause linkage in the languages of the Ob-Yenisei area: Asyndetic constructions" [Behnke, Wagner-Nagy (eds.) 2024], a recent collection of articles dealing with clause linkage in Ob-Ugric and Samoyedic languages, among others, with focus on constructions with non-finite dependent clause. This special issue continues the recent series in *Voprosy Jazykoznanija*, following the issues № 5/2022 and № 6/2023.

REFERENCES

- Abondolo, Valijärvi (eds.) 2023 — Abondolo D., Valijärvi R. L. (eds.). *The Uralic languages*. 2nd edn. Routledge: London; New York, 2023.
- Aissen 2023 — Aissen J. Documenting topic and focus. Key topics in language documentation and description. *Language Documentation & Conservation*, 2023, 26: 11–57.
- Bakró-Nagy et al. 2020 — Bakró-Nagy M., Laakso J., Skribnik E. "As we say it in Finno-Ugric". Some thoughts on making Uralic language studies more accessible to outsiders. *Ёмас сымын нэ́кве вёртуп ётно́ст самы́н паму́м. Scripta miscellanea in honorem Ulla-Maija Forsberg*. Saarikivi J., Virtanen S., Holopainen S. (eds.). Helsinki: Suomalais-Ugrilainen Seura, 2020, 27–41.
- Bakró-Nagy et al. (eds.) 2022 — Bakró-Nagy M., Laakso J., Skribnik E. (eds.). *The Oxford guide to the Uralic languages*. Oxford: Oxford Univ. Press, 2022.
- Behnke, Wagner-Nagy (eds.) 2024 — Behnke A., Wagner-Nagy B. (eds.). *Clause linkage in the languages of the Ob-Yenisei Area: Asyndetic constructions*. Leiden; Boston: Brill, 2024.
- Bíró, Sipőcz 2017 — Bíró B., Sipőcz K. The Mansi ditransitive constructions. *Finno-Ugric Languages and Linguistics*, 2017, 6(1): 41–55.
- Blokland et al. 2023 — Blokland R., Bradley J., Klumpp G. Preface. *Linguistica Uralica*, 2023, 59(4): 241–242.
- Dalrymple, Nikolaeva 2011 — Dalrymple M., Nikolaeva I. *Objects and information structure*. Cambridge: Cambridge Univ. Press, 2011.
- Däbritz 2021 — Däbritz Ch. L. *Topik, Fokus und Informationsstatus: Modellierung am Material nordwest-sibirischer Sprachen*. Berlin; Boston: Walter de Gruyter, 2021.
- É. Kiss 1995 — É. Kiss K. NP movement, operator movement, and scrambling in Hungarian. *Discourse configurational languages*. É. Kiss K. (ed.). New York; Oxford: Oxford Univ. Press, 1995, 207–243.
- É. Kiss 2019 — É. Kiss K. Fused grammatical and discourse functions in Ob-Ugric: Case, agreement, passive. *Proc. of the workshop "Clause Typing and the Syntax-to Discourse Relation in Head-Final Languages"*, Allensbach-Freudental, May 15–17, 2019. Bayer J., Viesel Y. (eds.). Konstanz: Universität Konstanz, 2019, 163–174.
- Fedorova et al. (eds.) 2014 — Fedorova O. V., Lyutikova E. A., Daniel M. A., Plungian V. A., Tatevosov S. G. (eds.). *Язык. Константы. Перемennые. Памыати Aleksandra Evgen'evicha Kibrika* [Language. Constants. Variables. In memoriam of Aleksandr Evgen'evich Kibrik]. St. Petersburg: Aletheia, 2014.

- Filchenko 2006 — Filchenko A. The Eastern Khanty locative-agent constructions: A functional discourse-pragmatic perspective. *Demoting the agent: Passive, middle and other voice phenomena*. Lynghfelt B., Solstad T. (eds.). Amsterdam: John Benjamins, 2006, 47–82.
- Janda 2015 — Janda G. E. Northern Mansi possessive suffixes in non-possessive function. *Referential devices in Uralic languages*. Pajusalu R., Klumpp G., Laury R. (eds.). Tartu: Tartu Univ. Press, 2015, 243–258.
- Janda 2019 — Janda G. E. *Funktionen von Possessivsuffixen in den ugrischen Sprachen*. Cologne: Modern Academic Publ., 2019.
- Kashkin (ed.) 2023 — Kashkin E. V. (ed.). *Elementy gornomariiskogo yazyka v tipologicheskom osveshchenii* [Elements of Hill Mari in a typological light]. Moskva: Buki Vedi, 2023.
- Khanina, Shluinsky 2020 — Khanina O., Shluinsky A. Competing ditransitive constructions in Enets. *Functions of Language*, 2020, 27(3): 247–279.
- Khomchenkova 2022 — Khomchenkova I. A. Diskursivnye funktsii posessivnogo pokazatelya tret'ego litsa edinstvennogo chisla v gornomariiskom yazyke [Discourse functions of the 3rd singular possessive marker in Hill Mari]. *Ural-Altaic Studies*, 2022, 45(2): 148–168.
- Klumpp 2012 — Klumpp G. Differential object marking and information structure: On the function of two different pronominal accusatives in Komi and Khanty dialects. *Eesti ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 2012, 3(1): 343–372.
- Klumpp 2014 — Klumpp G. Identifiability, givenness and zero-marked referential objects in Komi. *Linguistics*, 2014, 52(2): 415–444.
- Klumpp 2023 — Klumpp G. On dative-lative encoded direct objects in West Mansi. *Linguistica Uralica*, 2023, 59(4): 307–338.
- Klumpp, Skribnik 2022 — Klumpp G., Skribnik E. Information structuring. *The Oxford guide to the Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 1018–1036.
- Koshkareva 2002 — Koshkareva N. B. Kommunikativnaya paradigma khantyiskogo predlozheniya (na materiale kazym'skogo dialekta) [Communicative paradigm of the Khanty sentence (based on the Kazym dialect data)]. *Yazyki korennykh narodov Sibiri*. Iss. 12. Koshkareva N. B., Oktyabr'skaya I. V. (eds.). Novosibirsk: Institute of Philology of the Siberian Branch of the Russian Academy of Sciences, 2002, 29–44.
- Kozlov, Zakirova 2023 — Kozlov A., Zakirova A. The exhaustive particle =ok in Hill Mari and beyond. *Linguistica Uralica*, 2023, 59(2): 154–188.
- Krifka 2007 — Krifka M. Basic notions of information structure. *The notions of information structure*. Féry C., Fanselow G., Krifka M. (eds.). Potsdam: Universitätsverlag, 2007, 13–55.
- Kulonen 1989 — Kulonen U.-M. *The passive in Ob-Ugric*. Helsinki: Suomalais-Ugrilainen Seura, 1989.
- Lambrecht 1994 — Lambrecht K. *Information structure and sentence form*. Cambridge: Cambridge Univ. Press, 1994.
- Leisiö 2006 — Leisiö L. Passive in Nganasan. *Passivization and typology: Form and function*. Abraham W., Leisiö L. (eds.). Amsterdam; Philadelphia: John Benjamins, 2006, 213–230.
- Muravyev, Zhornik 2024 — Muravyev N., Zhornik D. Passive in Ob-Ugric: Information structure and beyond. *Finnisch-Ugrische Mitteilungen*, 2024, 47(4): 67–95.
- Nikolaeva 1999 — Nikolaeva I. Object agreement, grammatical relations, and information structure. *Studies in Language*, 1999, 23: 331–376.
- Nikolaeva et al. 1993 — Nikolaeva I., Kovgan E., Koshkareva N. Communicative roles in Ostyak syntax. *Finnisch-Ugrische Forschungen*, 1993, 51: 125–167.
- Nikolaeva 2014 — Nikolaeva I. *A grammar of Tundra Nenets*. Berlin: Mouton de Gruyter, 2014.
- Rombandeeva 1979 — Rombandeeva E. I. *Sintaksis mansiiskogo (vogul'skogo) yazyka* [Syntax of Mansi (Vogul)]. Moscow: Nauka, 1979.
- Serdobolskaya 2015 — Serdobolskaya N. Pseudoincorporation analysis of unmarked direct objects in Mari. *The syntax and semantics of pseudo-incorporation*. Borik O., Gehrke B. (eds.). Leiden; Boston: Brill, 2015, 298–328.
- Serdobolskaya 2020 — Serdobolskaya N. A corpus analysis of differential object marking in Beserman Udmurt. *Linguistica Uralica*, 2020, 56(4): 275–308.
- Serdobolskaya, Toldova 2016 — Serdobolskaya N. V., Toldova S. Yu. Strukturnaya pozitsiya pryamogo dopolneniya i ego kommunikativnyi status [Structural position of the direct object and its communicative status]. *Arkhitektura klauzy v parametricheskikh modelyakh. Sintaksis, informatsionnaya struktura, poryadok slov*. Zimmerling A. V., Lyutikova E. A. (eds.). Moscow: Yazyki slavyanskikh kul'tur, 2016, 417–444.
- Serdobolskaya, Toldova 2017 — Serdobolskaya N. V., Toldova S. Yu. Direct Object marking in Finno-Ugric languages: Between sentence and discourse. *Ural-Altaic Studies*, 2017, 27(4): 92–112.

- Skribnik 2001 — Skribnik E. Pragmatic structuring in Northern Mansi. *Congressus Nonus Internationalis Fenno-Ugristarum* 2. Pars VI. *Dissertationes sectionum: Linguistica III*. Seilenthal T. (ed.). Tartu: Auctores, 2001, 222–239.
- Skribnik 2004 — Skribnik E. K. Kategorii mansiiskogo glagola i aktual'noe chlenenie predlozheniya [Mansi verb categories and communicative sentence structure]. *Tipologicheskie obosnovaniya v grammatike. K 70-letiyu prof. V. S. Khrakovskogo*. Volodin A. P. (ed.). Moscow: Znack, 2004, 445–458.
- Stenin 2018 — Stenin I. A. Passiv [Passive]. *Elementy mokshanskogo yazyka v tipologicheskom osveshchenii*. Toldova S. Yu, Kholodilova M. A. (eds.). Moscow: Buki Vedi, 2018, 490–545.
- Stenin 2023 — Stenin I. A. O passivnykh glagolakh s pokazatelem *-ra / -rye* v tundrovom nenetskom yazyke [On passive verbs with the marker *-ra / -rye* in Tundra Nenets]. *Voprosy Jazykoznanija*, 2023, 4: 21–46.
- Stipa 1990 — Stipa G. J. *Finnisch-ugrische Sprachforschung: von der Renaissance bis zum Neopositivismus*. Helsinki: Suomalais-Ugrilainen Seura, 1990.
- Toldova 2017 — Toldova S. Yu. Kodirovanie pryamogo dopolneniya v mokshanskom yazyke [Marking of the direct object in Moksha]. *Acta Linguistica Petropolitana*, 2017, 13(3): 123–157.
- Toldova, Kholodilova (eds.) 2018 — Toldova S. Yu, Kholodilova M. A. (eds.). *Elementy mokshanskogo yazyka v tipologicheskom osveshchenii* [Elements of Moksha Mordvin in a typological light]. Moscow: Buki Vedi, 2018.
- Urmanchieva, Plungian 2021 — Urmanchieva A. Yu., Plungian V. A. Passiv v zapadnom dialekte mansiiskogo [Passive in Western Mansi]. *Tomsk Journal of Linguistics and Anthropology*, 2021, 1: 94–131.
- Vilkuna 1989 — Vilkuna M. *Free word order in Finnish: Its syntax and discourse functions*. Helsinki: Suomen Kirjallisuuden Seura, 1989.
- Virtanen 2015 — Virtanen S. *Transitivity in Eastern Mansi: An Information Structural approach*. PhD thesis. Helsinki: Univ. of Helsinki, 2015.
- Virtanen 2023 — Virtanen S. Pragmatic reasons for non-clause-initial placement of subjects of active and passive clauses in Northern Mansi. *Linguistica Uralica*, 2023, 59(4): 286–306.

Получено / received 26.06.2024

Принято / accepted 24.09.2024

The influence of internal possessors on argument coding in Northern Khanty

© 2024

Nikita A. Muravyev

University of Hamburg, Hamburg, Germany; nikita.muraviev@gmail.com

Abstract: The study examines contexts in Northern Khanty where the presence of internal possessors in the noun phrases of the core arguments influences their morphosyntactic coding. Based on the corpus data from the Kazym dialect of Khanty, I investigate contexts in which the presence of a third-person topical possessor in the agent or patient NP leads to agent demotion and passivization. The article provides an analysis of these contexts in terms of the interaction between discourse prominence and processing efficiency. It is demonstrated that passive voice is used as an ambiguity-resolving strategy when several unmarked participants are present in the clause, or when referential conflict may arise due to unclear reference of an anaphoric possessive marker. This strategy applies whenever the agent is less prominent than either the patient or the possessor of the patient, or whenever both possessors of the agent and the patient are present, and both core arguments are low on animacy, agentivity, and topicality hierarchies, in which case passivization is due not to their relative prominence but rather to considerations of processing efficiency.

Keywords: argument structure, Northern Khanty, passive, possession, Uralic, voice

Acknowledgements: The research was conducted as part of the project “Verb Agreement and Voice in Northern Khanty Discourse”, supported by the Alexander von Humboldt Foundation. I am grateful to my colleagues and collaborators, Stiopa Mikhailov and Vsevolod Masliukov, for discussing my data from the perspective of hierarchical alignment and obviation. I also thank the volume editors and the anonymous reviewer who helped improve the quality of this paper. Any remaining mistakes are my own.

For citation: Muravyev N. A. The influence of internal possessors on argument coding in Northern Khanty. *Voprosy Jazykoznanija*, 2024, 5: 65–79.

DOI: 10.31857/0373-658X.2024.5.65-79

Влияние внутреннего посессора на кодирование аргументов в севернохантыйском языке

Никита Алексеевич Муравьев

Университет Гамбурга, Гамбург, Германия; nikita.muraviev@gmail.com

Аннотация: В исследовании рассматриваются контексты в севернохантыйском языке, в которых наличие внутреннего посессора в именных группах ядерных участников влияет на кодирование этих актантов. На основе корпусных данных казымского диалекта хантыйского рассматриваются контексты, в которых присутствие топикального посессора третьего лица в именной группе агенса или пациенса приводит к понижению агенса и пассивизации. В статье проводится анализ таких контекстов с точки зрения взаимодействия дискурсивной значимости и эффективности обработки речи. Установлено, что пассивный залог используется как стратегия разрешения неоднозначности, когда в предложении присутствуют несколько немаркированных участников, а также когда может возникнуть референциальный конфликт из-за неоднозначной референции анафорического посессивного маркера. Эта стратегия применяется, когда агенс менее дискурсивно значим, чем пациенс или посессор пациенса, или когда присутствуют посессоры и агенса, и пациенса, и оба основных аргумента занимают низкую позицию на иерархиях одушевленности,

агентивности и топиальности — в данном случае использование пассивного залога обусловлено не их иерархической позицией, а соображениями удобства восприятия.

Ключевые слова: аргументная структура, залог, пассив, поссесивность, севернохантыйский язык, уральские языки

Благодарности: Исследование проведено в рамках проекта «Глагольное согласование и залог в северохантыйском дискурсе», поддержанного Фондом Александра фон Гумбольдта. Я благодарен моим коллегам и соавторам Степану Михайлову и Всеволоду Маслокову за обсуждение моих данных с точки зрения иерархического кодирования и обвации. Я также благодарю редакторов спецвыпуска и анонимного рецензента, которые помогли улучшить качество этой статьи. Все оставшиеся ошибки являются моими собственными.

Для цитирования: Muravyev N. A. The influence of internal possessors on argument coding in Northern Khanty. *Voprosy Jazykoznanija*, 2024, 5: 65–79.

DOI: 10.31857/0373-658X.2024.5.65-79

Introduction

Any discussion of morphosyntactic coding will primarily focus on the properties of the core clausal arguments, namely S(ole) argument of an intransitive clause, A(gent), or transitive subject, and P(atient), or transitive object, following Dixon's [1972] SAO model. One such property that is rarely if at all considered is whether or not a core argument has an internal possessor. In Northern Khanty (< Ob-Ugric < Uralic), the presence of an internal possessor of A and/or P triggers the use of passive voice, as in example (1) below, which does not straightforwardly follow from the identity and status of either of the two core arguments or the event itself.

- (1) *śāłta mewəl nepək-əł pox-əł-ən śiwałə-s-i*
 then chest paper-POSS.3SG son-POSS.3SG-LOC see-PST-PASS[3SG]

‘Then his_i (the old man’s) son saw his_i chest paper.’ (NHC¹, The eagle: 18)

This study explores the idea that such use of the passive voice is due to morphological ambiguity between overt unmarked core arguments (here A ‘his son’ and P ‘his chest paper’) and their possessors, as well as referential ambiguity of possessive markers. It is suggested that the passive voice resolves this ambiguity by distinctly marking A with an oblique locative case, thus reducing by one the number of unmarked clausal arguments. Based on the corpus data of Kazyms Khanty, I determine the exact conditions under which such ambiguity-resolving passives appear.

1. Argument coding in Northern Khanty

Northern Khanty belongs to the Ob-Ugric branch of the Uralic language family. It is spoken in the western part of Northern Siberia, in the Khanty-Mansi and Yamalo-Nenets Districts, by approximately 13,900 speakers, according to the All-Russian Population Census of 2020². Case marking in Northern Khanty follows the neutral/accusative alignment [Nikolaeva 1999]. Full noun phrases as core transitive arguments are always unmarked (2), while personal pronouns distinguish between the unmarked nominative form for the S- and A-arguments and the marked accusative form for the P-argument (3).

¹ The example is taken from the Northern Khanty corpus.

² Available online at: <https://rosstat.gov.ru/vpn/2020>.

- (2) *pox an šukat-əs*
 boy cup break-PST[3SG]
 ‘A boy broke a cup.’ (elicited)
- (3) *luw mānti χātsə-s*
 he I.ACC hit-PST[3SG]
 ‘He hit me.’ (elicited)

Verbs in Northern Khanty obligatorily agree with the S/A-argument in person and number. Transitive verbs show non-obligatory agreement with P in number, determined primarily by information structure and/or givenness [Nikolaeva 2001; Koshkareva 2002; Dalrymple, Nikolaeva 2011; É. Kiss 2019; Däbritz 2020; Muravyev 2023]. In typological literature, this phenomenon is referred to as differential object indexing, see [Lemmolo 2011; Haig 2018]. Morphologically, verbs that index only the S/A-argument are inflected for subjective conjugation, whereas verbs indexing both A and P arguments take a separate objective conjugation paradigm. Compare a subjective form in a clause with a new and/or focused P-argument (4a) and an objective form with a given and/or topical P-argument (4b). The person-number affix follows the tense marker.

- (4) a. *juwan pe:tra re:sk-əs* / **re:sk-əs-li*
 John Peter hit-PST[3SG] hit-PST-3SG>SG
 [Whom did John hit?] ‘John hit PETER.’
- b. *luw pe:tra re:sk-əs-li* / **re:sk-əs*
 he peter hit-PST-3SG>SG hit-PST[3SG]
 [What did John do to Peter?] ‘He HIT Peter.’ [Nikolaeva 2001: 29–30]

In addition, Northern Khanty has a passive voice that is used whenever the A-argument is not given or topical [Kulonen 1989; Nikolaeva 2001; Koshkareva 2002; É. Kiss 2019]. The P-argument appears in the subject slot and is indexed on the verb by means of the subjective conjugation, while the A-argument is either unexpressed or, if overt, coded as a locative oblique, cf. (5). The passive marker *-a(j)* appears between tense and person-number affixes, with a non-word-final allomorph *-aj* and word-final *-a* [Kaksin 2010: 112–114]. In some Northern dialects, an additional pair of allomorphs, *-i(j)*, is used with vocalic verb stems.

- (5) *kalan xoj-na we:l-s-a?* — *Juwan-na we:l-s-a*
 reindeer who-LOC kill-PST-PASS[3SG] john-LOC kill-PST-PASS[3SG]
 ‘Who killed the/a reindeer? — JOHN did.’ [Nikolaeva 2001: 25]

Together, the passive voice and both conjugation types form a hierarchical alignment system based on animacy, givenness, and topicality (see [Muravyev 2023]). In languages with hierarchical alignment, the subject slot is accessed by the highest-ranked core argument [Nichols 1992; Thompson 1994; Zúñiga 2006]. Whenever A outranks P, neutral *direct* construction is used, as in example (6a) from Plains Cree with Speech-act participant (SAP) A > 3rd person P. Otherwise *inverse* construction is used, as in (6b) with 3 A > SAP P.

- (6) a. *ni-wāpam-ā-w* b. *ni-wāpam-ikw-w*
 1-see-3 1-see-INV-3
 ‘I see him/her.’ ‘He/she sees me.’ [Zúñiga 2006: 2]

Assuming active voice to correspond to direct and passive to inverse, we get a similar person contrast in Northern Khanty. In SAP A > 3 P contexts, active, as in (7a), is obligatory, because there are no locative forms for personal pronouns, which makes a passive construction impossible. On the other hand, in 3 A > SAP P contexts, passive is strongly preferred (7b), yet active, as in (3), may also occur if A is highly topical.

- (7) a. *ma waśaj-en χătśə-s-əm*
 I Vasya-POSS.2SG hit-PST-1SG
 ‘I hit Vasya.’
- b. *ma waśaj-en-ən χătśə-s-ij-əm*
 I Vasya-POSS.2SG-LOC hit-PST-PASS-1SG
 ‘Vasya hit me.’ (elicited)

There is often a further distinction between “proximate” (discourse-prominent, hierarchically high) and “obviative” (non-prominent) third-person arguments, see Section 2.3 for further details. For example, (8a) and (8b) from Plains Cree illustrate a proximate- and obviative-marked ‘dog’, respectively.

- (8) a. *niwapamaw atim* b. *pakamahwew napew atimwa*
 see(1–3) dog(3) hit(3–3’) man(3) dog(3’)
 ‘I see the dog.’ ‘The man hits the dog.’ [Aissen 1997: 707]

In most dialects of Northern Khanty, there are no formal distinctions in the domain of third person. However, Kazym Khanty employs a 2SG possessive affix (cf. (9a)) as a salient article, which specifically marks prominent or “proximate” arguments [Mikhailov (to appear)]. This salient article follows the proximate uniqueness principle proposed in [Goddard 1990] and [Aissen 1997; 2001] for languages with obviation marking. For instance, in (9b), only one noun phrase is compatible with the proximate reading of the 2SG possessive marker.

- (9) a. *amp-en λow šuwəλ-əs*
 dog-POSS.2SG bone see-PST[3SG]
 ‘The dog saw the bone.’
- b. *amp-en λow-en šuwəλ-əs*
 dog-POSS.2SG bone-POSS.2SG see-PST[3SG]
 ‘The dog saw your/*the bone.’ (elicited)

2. Passive and ambiguity resolution

Passive constructions in Northern Khanty are highly frequent in texts and have a variety of uses, some of which may seem counterintuitive at first. Consider the following pair of examples:

- (10) *məś pəxatur imij-a mə-s ar sərni wox, pox-le*
 Moshch hero woman-DAT give-PST[3SG] much gold money boy-DIM
wu-s-λe pa joxli pəla mən-ti pit-əs
 take-PST-3SG>SG ADD backwards to go-NFIN.NPST AUX-PST[3SG]
 ‘The Moshch hero gave the woman a lot of gold and money, took her son and went back.’ (NHC, The hero: 29)
- (11) *śi kut-ən ime-λ nox wərλ-əs pa jəś-əλ*
 DEM gap-LOC wife-POSS.3SG up wake-PST[3SG] ADD hand-POSS.3SG
jira katəλ-s-a
 to.side grab-PST-PASS[3SG]
 ‘Meanwhile his wife woke up and grabbed his hand (lit. “his hand was grabbed”).’ (NHC, Three wise words: 70)

In (10), the same actor is expressed throughout the passage as the subject. In contrast, in (11), the actor is first mentioned as the subject but then demoted in a passive construction. Typologically, A-arguments performing successive actions normally continue to be expressed

in the subject slot (see [Givón (ed.) 1983]), as in (10). Thus, the use of the passive in (11) seems out of place. The only obvious difference between the contexts is the presence of an anaphoric third-person possessor ‘husband’ in the nominal expressions ‘his wife’ and ‘his hand’ in the latter example. One may wonder what the link is between expressing a possessor and using the passive voice.

The idea I argue for in this paper is that the passive, as seen in (11), is used to resolve ambiguity, similar to the use of the inverse in Tlapanec to disambiguate between subjects and possessors [Wichmann 2007]. Northern Khanty tends to avoid overt expression of core arguments and instead uses zero-anaphora or pronominal expressions. Combined with the presence of a possessor in (11), this creates ambiguity: it is unclear, especially without contextual support, who grabbed whose hand. The passive efficiently resolves this ambiguity by a) demoting the agent ‘wife’ so that it can no longer be interpreted as the antecedent of the third-person possessive marker on ‘hand’, and b) keeping NPs with coindexed third-person possessors in the subject slot to ensure that their possessor is interpreted as the same referent, ‘husband’.

If this analysis is correct, we would expect the passive to resolve ambiguity in the presence of a possessor in most or all of the available corpus data. My research questions are as follows:

1. **Does passive resolve ambiguity between A, P, and possessors?**
2. **In what contexts does the ambiguity resolving passive apply?**

The first question can be answered by examining a sample of transitive clauses with possessors and determining the ratio of passive to active examples. If all or nearly all examples are passive, then the passive can be viewed as a dedicated means of ambiguity resolution. Otherwise, the passive is just one among several ambiguity-resolving means, and its use is restricted to certain contexts, which the second question aims to identify. Identifying these contexts requires an in-depth analysis of the event properties and all participants involved.

My general proposal can be formulated as follows:

Under conditions that require morphological ambiguity resolution in clauses simultaneously containing a) core arguments A and P, and b) possessor(s) of A and/or P, such ambiguity is fully or partially resolved by passivization, which demotes A to a locative oblique.

The main research objective was to identify the exact conditions for morphological ambiguity resolution. The following three hypotheses were considered:

- **Processing efficiency hypothesis:** Obligatory A-demotion and passivization occur whenever a clause contains more than two participants from the set {A, P, PsrA, PsrP}.
- **Possessor affectedness hypothesis:** Obligatory A-demotion and passivization occur whenever P’s possessor is affected by the event and needs to be morphosyntactically integrated into the proposition.
- **Proximate hypothesis:** Obligatory A-demotion and passivization occur whenever either P or its possessor is more prominent in the current discourse than A.

To test these hypotheses, all 1,035 transitive clauses, including 350 examples with possessors, were extracted from the Kazym dialect subcorpus of the unpublished Northern Khanty corpus (42,174 tokens), compiled by Egor Kashkin. For general purposes and to test the first hypothesis, the corpus examples were manually annotated for active and passive forms of the verb, the presence/absence of possessors on A and P, their formal expression (i.e., full NP or person-number index on the head), and whether they correspond to A, P, a speech-act participant, or another discourse referent. To test the second hypothesis, the examples were annotated according to whether the possessor of P is affected by the event. Annotation for the third hypothesis included parameters that are typologically relevant for hierarchical systems: person

(1, 2, 3), animacy (human, animal, inanimate), definiteness (definite, indefinite specific, indefinite non-specific), and topicality (topical > non-topical). Additionally, some field examples collected in the village of Kazym (Russia) in 2022 were used for illustration purposes.

2.1. Processing efficiency hypothesis

According to this hypothesis, the passive voice occurs in transitive clauses when A and P are present along with a possessor of either one or both, without any further conditions. The underlying assumption is that, since neither of the non-pronominal participants is case-marked, the simultaneous presence of these participants creates parsing difficulties. In his discussion of efficiency and complexity in English relative clauses, J. Hawkins [2004] demonstrates that similar parsing difficulties can be addressed by diversifying the coding in a clause, which he calls the “Maximize On-Line Processing” principle.

For example, in a “garden-path sentence” (12a), the hearer cannot unambiguously assign a grammatical function to the NP *the boy* before the sentence is finished. In contrast, (12b) resolves the issue by explicitly introducing the dependent clause with a complementizer *that*.

- (12) [Hawkins 2004: 51]
 a. $I_{VP}[\text{realize}_{S2}[\text{the boy knows the answer}]]$.
 b. $I_{VP}[\text{realize}_{S2}[\text{that the boy knows the answer}]]$.

In Northern Khanty, the passive voice appears to serve the same purpose by distinctly coding A as a locative oblique, as in (13) below. This allows the hearer to process the sentence as early as possible, avoiding erroneous role assignments and referential conflicts. First, such use of passive signals that neither of the unmarked participants is the agent. Second, it helps to avoid confusion of the antecedents of the two possessive markers: the former, on A $xqj\text{-}\acute{\alpha}\lambda\text{-}\acute{\alpha}n$ ‘man’, referring to the topical set of characters, and the latter, on P $\acute{\tau}ep\acute{\alpha}\lambda\text{-}\acute{\alpha}\lambda$ ‘belly’, to the overtly expressed possessor $s\acute{o}rt$ ‘pike’. This is similar to the concept of precluding referential conflict [Fedorova 2014] and the idea of passive as a “deconflicter” in A. Kibrik’s [2011] terms. In the majority of the annotated examples, the passive voice is indeed chosen as predicted.

- (13) $\acute{s}\acute{\alpha}\lambda\acute{t}a$ *in* $pup\acute{i}$ $sox\text{-}\acute{\alpha}p$ $xqj\text{-}\acute{\alpha}\lambda\text{-}\acute{\alpha}n$ *in* $w\acute{o}n$ $s\acute{o}rt$
 then DEM bear coat-PROPR man-POSS.3SG-LOC DEM big pike
 $\acute{\tau}ep\acute{\alpha}\lambda\text{-}\acute{\alpha}\lambda$ $p\acute{e}lki$ $\acute{s}om\acute{a}rm\acute{\alpha}\text{-}s\text{-}i$
 belly-POSS.3SG apart crush-PST-PASS[3SG]
 ‘Then the Man in the Bear Coat crushed the belly of the large pike.’ (NHC, Younger daughter of the sun: 80)

On the other hand, there are still plenty of examples in the active voice. For instance, A-arguments that are speech-act participants, as ‘we’ in (14), are never demoted from the subject slot. This can be explained by the fact that first and second persons are easily distinguished from third persons, so no ambiguity occurs.

- (14) $\acute{s}aj$ $ja\acute{n}i\acute{s}i$ $\acute{o}m\acute{\alpha}s\text{-}s\text{-}\acute{\alpha}m$, $\acute{s}\acute{o}\acute{s}\text{-}\acute{\alpha}t$ $j\acute{\alpha}n\acute{x}\text{-}m\acute{\alpha}t\text{-}s\text{-}\acute{\alpha}t$, $ku\acute{s}$ $mu\acute{n}$
 tea drink:NFIN.NPST sit-PST-1SG drake-PL walk-PNCT-PST-3SG though we
 $\lambda u w$ $\acute{n}oxe\text{-}\lambda$ $\acute{\alpha}nt$ $\lambda\acute{\epsilon}\text{-}\lambda\text{-}ew$
 they meat-POSS.3SG NEG eat-NPST-1SG>SG
 ‘I sat down to drink tea, the drakes flew in, even though we don’t eat their meat.’ (NHC, A gunless hunter: 10)

However, even third-person A-arguments can sometimes remain in the subject slot in the presence of a possessor, as shown in (15) and (16) below.

- (15) *λωw-αλ* *ξρηx-s-αλλε*
 horse-POSS.3SG kick-PST-3SG>SG
 ‘He_i kicked his_i horse.’ (NHC, God given wealth: 19)
- (16) *ike-λ* *λαλ'-a* *λεśat-s-a*, *ime-λ* *λωw*
 husband-POSS.3SG war-DAT prepare-PST-PASS[3SG] wife-POSS.3SG horse
pānən *tq-s-λε*
 with.oneself bring-PST-3SG>SG
 ‘They sent her husband to war; he took his wife’s horse with him.’ (NHC, Golden horse: 42)

Examples like (15) with zero A acting on P ‘his horse’ do not pose serious problems for interpretation, as A and the possessor of P are coreferent, meaning that there are only two participants involved, not three. In such contexts, the active voice is the only available option in the corpus. Note that in a mirror-image setup with A ‘his wife’ acting on zero P where P and the possessor of A are coreferent, as in (17), only the passive voice is used. Examples of this kind with non-distinct possessors will be briefly mentioned in Section 2.3 but otherwise remain outside the scope of this study.

- (17) *joxi* *λοη-ας*, *ime-λ-ən* *ińśas-λ-a*
 home enter-PST[3SG] woman-POSS.3SG-LOC ask-NPST-PASS[3SG]
 ‘He_i came home and his_i wife asks him_i.’ (NHC, The eagle: 55)

Examples like (16) with zero A acting on P ‘his wife’s horse’, on the other hand, do pose a problem for the first hypothesis. They suggest that certain properties of the arguments either block the use of the passive voice or make it unnecessary. In the following sections, I will discuss other factors potentially influencing the use of passive, namely whether P’s possessor is affected by the event (see Section 2.2 on the possessor affectedness hypothesis) and whether A is prominent/topical (see Section 2.3 on the proximate hypothesis). Table 1 below presents corpus counts for the processing efficiency hypothesis. All instances of overtly expressed or anaphorically marked possessors of A and P (including cases of multiple embedded possessors, as, e.g., in (16)) were counted as PsrA and PsrP respectively.

Table 1

Active and passive with different sets of arguments

	Active	Passive
A, P	722	200
A, P, PsrP	18	43
A, P, PsrA	22	19
A, P, PsrA, PsrP	1	10
Total with possessors	41	72
Total	763	272

In contrast to sentences without possessors of core arguments, in which the active voice clearly dominates in frequency, the presence of possessors gives the passive voice some advantage³. The passive voice appears twice as often as the active voice with A, P, and a possessor of P, and ten times more often than the active voice with A, P, and both possessors. However, with A, P, and a possessor of A, the active voice has a slight edge. Hence, the processing efficiency hypothesis needs further elaboration.

³ The Fisher exact test statistic value is < 0.00001. The result is significant at $p < 0.05$.

2.2. Possessor affectedness hypothesis

Since the simultaneous presence of A, P, and possessor(s) alone cannot account for the use of the passive voice, there must be additional factors at play. One potential factor is the affectedness of the possessor, which triggers passivization, as in example (18) below from Japanese. Such construction is usually referred to as adversative passive [Kortlandt 1992; Shibatani 1994], as typically the promoted participant is negatively affected by the event.

(18) JAPANESE [Kortlandt 1992: 99]

John wa dareka ni ie o yakareta
John TOP someone DAT house ACC burn.PASS.PST

‘John’s house was burnt by someone; John was negatively affected by it.’

In this example, John is the owner of the house and suffers the loss of his property. To emphasize the effect of the burning event on the possessor, the verb appears in the passive voice, and the possessor is topicalized. Northern Khanty exhibits a similar passive construction; however, the affected possessor, as ‘Masha’ in (19) and the speaker in (20) below, remains internal and is promoted to subject along with its host.

(19) *mašaj-en χət χojat-ən loŋ-s-a*
Masha-POSS.2SG house who.INDEF-LOC enter-PST-PASS[3SG]

‘Masha’s house was entered by someone.’ (elicited)

(20) *χət-εm xojat-ən nox wu-s-i*
house-POSS.1SG who.INDEF-LOC up take-PST-PASS[3SG]

‘My house has been taken by someone.’ (NHC, Little chipmunk: 13)

However, examples of the active voice with an affected P’s possessor, such as (21) below with zero A acting on P ‘the front door of the Vein-making woman’s house’, are also found in the corpus, which questions the possessor affectedness hypothesis.

(21) *meŋ kim εt-əs, šiw mǎn-əs pa*
daughter-in-law out appear-PST[3SG] DEM.LAT go-PST[3SG] ADD
λən wεrti ime-n χət ɔw-əl tut-ən
vein_making_woman-POSS.2SG house door-POSS.3SG fire-LOC
pon-s-əλλe
put-PST-3SG>SG

‘The daughter-in-law came out, went there and set the front door of the Vein-making woman’s house on fire.’ (NHC, Vein-making woman: 11)

Moreover, in both active (see (16) repeated below as (22)) and passive (23) examples with possessed Ps ‘his wife’s horse’ and ‘the son of those man and woman’ respectively, the possessor of P is not affected by the event. Examples with A, P, and A’s possessor but no P’s possessor are also left unaccounted for.

(22) *ike-λ λal’-a leśət-s-a, ime-λ λɔw*
husband-POSS.3SG war-DAT prepare-PST-PASS[3SG] wife-POSS.3SG horse
pǎnən tq-s-λe
with.oneself bring-PST-3SG>SG

‘They sent her husband to war; he took his wife’s horse with him.’ (NHC, Golden horse: 42)

(23) *šǎλta ime-ŋən-ike-ŋən pox šiw ši wɔx-s-a*
then woman-DU-man-DU boy DEM.LAT FOC call-PST-PASS[3SG]

‘Then they called the son of those man and woman.’ (NHC, Younger daughter of the sun: 46)

Table 2 below presents corpus counts for the possessor affectedness hypothesis.

Table 2

**Active and passive with different sets of arguments
with an affected and a non-affected possessor of P**

	PsrP affected		PsrP not affected	
	Active	Passive	Active	Passive
A, P, PsrP	8	23	11	11
A, P, PsrA	—	—	22	19
A, P, PsrA, PsrP	1	3	0	6
Total	9	26	33	36

The passive voice does correlate clearly with the affectedness of the P's possessor⁴. It appears three times more often with A, P, and the possessor of P, as well as with A, P, and both possessors. Yet, many examples without an affected possessor of P or without such a possessor are left unexplained by the possessor affectedness hypothesis. Hence, other conditions should also be taken into consideration.

2.3. Proximate hypothesis

Another potential factor is the relative ranking of A, P, and possessor(s) in terms of discourse prominence. Here I understand prominence in the sense of [von Heusinger, Schumacher 2019: 119] as “a relational property that singles out one element from a set of elements of equal type and structure”. Based on the conventions for describing languages with hierarchical alignment, I will distinguish two prominence statuses: proximate, or discourse-prominent, and obviative, or non-prominent. Proximate is the highest ranked argument among all the arguments in a clause with respect to the hierarchies of animacy, topicality, agentivity and with possessors preferred over possesseees [Aissen 2001: 6]. Proximate assignment obeys the Proximate Uniqueness Constraint stating that there can only be one proximate argument per obviation span which is minimally a clause [Ibid.: 12]. All other arguments are assigned obviative status.

In languages with hierarchical alignment, possessors are often treated on par with core arguments. Possessors can be either proximate or obviative, and proximate possessors have distinct coding from obviative ones [Aissen 1997]. In some languages, proximate possessors can even compete with A and P for the subject slot [Nikolaeva et al. 2019]. In a closely related Uralic language, Tundra Nenets, proximate possessors show indexing on the head, whereas obviative ones do not [Bárány, Nikolaeva 2019; Nikolaeva, Bárány 2019]. Compare (24), where the previously mentioned referent ‘Wera’ is assigned proximate status and is thus obligatorily indexed as a possessor, with (25), where the new referent ‘Petya’ is assigned obviative status with no indexing on the head.

TUNDRA NENETS

(24) [Why did Wera hit you?]

məŋ^o Wera-h mašina-m^o (-ta) taxabta^o-dəm-ć
I Wera-GEN car-ACC-3SG break-1SG-PST

‘I broke Wera’s car.’ [Nikolaeva, Bárány 2019: 232–233]

⁴ The Fisher exact test statistic value is 0.0356. The result is significant at $p < 0.05$.

- (25) [Where is your cup?]

*Pet'a-h weñako(-*da) taxabta°-da-s°*
 Petya-GEN dog-3SG break-3SG>SG-PST

'Petya's dog broke it.' [Ibid.]

Similar contrast is observed in Kazym Khanty, cf. *pet'aj-en xot-əl* [Petya-poss.2SG house-poss.3SG] 'Petya's house' where the indexed possessor 'Petya' is proximate and *pet'aj-en xot* [Petya-poss.2SG house] where it is non-indexed and obviative. Moreover, proximate possessors do participate in the hierarchical coding system alongside core arguments. Thus, settings with obviative A acting upon obviative P with a proximate possessor in almost all of the examples lead to passivization. In both examples below, passive is used, because the main characters appear as proximate possessors of obviative P-arguments, 'feet' (26) and 'wife' (27), while A-arguments 'those people' and 'Seven Canopies, Big Canopy' are obviative as well.

- (26) *kur-λ-əl xəšməλt-λ-aj-ət, isa potər*
 foot-PL-POSS.3SG warm.up-NPST-PASS-3PL UQ speech
oš-λ āntəm
 mind-POSS.3SG NEG.EX

'(Those people) are warming up her feet, she can't speak at all (from fear).' (NHC, Woman with children abandoned: 67)

- (27) [When he traveled, he used to leave his heart at home. While he was once driving through the forest, a man named Seven Canopies, Big Canopy, comes to his house.]

ime-λ śi lipət-s-a
 wife-POSS.3SG FOC pester-PST-PASS[3SG]

'He clung to his wife [— Where is his heart?].' (NHC, Younger daughter of the sun: 56)

In settings where an obviative A with an obviative possessor acts upon a proximate P, as shown in (28) below, passive also occurs without exceptions. Here P 'the small box' is assigned proximate status because it is more topical in the current stretch of discourse than the possessor of A, 'father'.

- (28) *pox-əl-ən aj λaraś nox punš-s-a*
 boy-POSS.3SG-LOC small box up open-PST-PASS[3SG]

'His son opened the small box.' (NHC, The eagle: 16)

On the other hand, contexts where a proximate A acts upon P with an obviative possessor, as illustrated in (29) below, consistently demonstrated the use of the active voice in all attested cases. The obviative status of the possessor 'husband' is supported by the fact that it is not indexed on the head noun 'heart'. Therefore, the higher prominence of A compared to the possessor naturally explains the preference for the active voice over the passive.

- (29) *in ne pakən-əl pa śi mā-λ-λe*
 DEM woman be.scared-NPST[3SG] ADD FOC give-NPST-3SG>SG
ike-λ sāṃ
 man-POSS.3SG heart

'The woman gets scared and gives her husband's heart.' (NHC, Younger daughter of the sun: 40)

Importantly, examples (15) and (17) with non-distinct possessors discussed earlier in Section 2.1 also fit naturally into this proposal. The obligatory use of the active voice when P is possessed by A and the obligatory use of the passive voice when A is possessed by P are due to the possessor always being more prominent than its possession, as described by the principle "possessor is proximate, possessee is obviative" [Aissen 1997].

One problematic context arises in sentences where both A and P have possessors. Under a hierarchical analysis, both A and P in such cases should be assigned obviative status, hence

no preference for either active or passive voice is expected. However, in reality, we observe a surprising dominance in the use of the passive voice. There are isolated examples in the corpus where there are two distinct possessors, with only one being proximate, as in (30) where the possessor of P ‘its sound’ is proximate and A ‘our adults’ has a first-person possessor, which has no obviation status, since speech-act participants do not participate in obviation [Goddard 1990; Aissen 1997; 2001].

- (30) *mǎn-əm* *sij-λ* *wɔn* *ot-λ-am-n* *ši* *xɔλ-s-a*,
 go-NFIN.PST sound-POSS.3SG big thing-PL-POSS.1SG-LOC FOC hear-PST-PASS[3SG]
ma isa xol pakən-s-əm
 I UQ away be.scared-PST-1SG
 ‘The noise with which she flew away was heard by adults; I was completely scared.’ (NHC, The tackle: 22)

However, in the majority of examples like sentence (1) repeated below with A ‘his son’ acting on P ‘his chest paper’, A and P share the same proximate anaphorically expressed possessor, and in this case, passive invariably occurs.

- (31) *šǎlta mewəl nepek-əl* *pox-əl-ən* *šiwalə-s-i*
 then chest paper-POSS.3SG son-POSS.3SG-LOC see-PST-PASS[3SG]
 ‘Then his_i (the old man’s) son saw his_i chest paper.’ (NHC, The eagle: 18)

This results in an asymmetrical situation where a possessed P is promoted to subject whenever A is obviative, whether possessed or not, while a possessed A can fill the subject slot only if P is obviative and has no proximate possessor. One possible explanation for this asymmetry, which again involves the processing efficiency hypothesis, is that when a clause contains a full set of A, P, and their possessors, there is no clearly proximate, zero, or pronominal expression of A. It becomes too difficult to keep track of all the roles and possessive relations of the participants; hence A is demoted to oblique via passivization. Table 3 presents corpus counts for the proximate hypothesis. Relative prominence of arguments and possessors was assessed based on their respective positions on the animacy, agentivity, and topicality hierarchies.

Table 3

Active and passive with different sets of arguments
and varying prominence of the arguments

	A/PsrA > P/PsrP		A/PsrA < P/PsrP		PsrA = PsrP	
	Active	Passive	Active	Passive	Active	Passive
A, P, PsrP	17	0	1	39	—	—
A, P, PsrA	20	0	0	18	—	—
A, P, PsrA, PsrP	2	0	0	1	0	9
Total	39	0	1	58	0	9

The proximate hypothesis successfully accounts for almost all of the available data⁵. Active voice consistently appears when A is hierarchically higher than P or its possessor, and the use of passive voice when A is lower than P or its possessor is also almost without exception. Hence, this hypothesis is accepted. The only category of examples not accounted for are sentences with both hierarchically low A and P where only passive voice is used.

⁵ The Fisher exact test statistic value is < 0.00001. The result is significant at $p < 0.05$.

3. Summary and discussion

The Northern Khanty data presented here demonstrates that, although possessors are not verbal arguments, they still belong to the morphosyntactic core of the system. Similar to A- and P-arguments, they are not marked for case and are indexed for person and number on the head, with the distinction that this indexing occurs inside the noun phrase rather than on the predicate. Functionally, treating possessors in this manner serves the purposes of discourse status tracking by signalling that the current core arguments are linked to a more prominent discourse referent.

On the formal side, however, this treatment of possessors creates ambiguity that is challenging to resolve without proper morphological support. As demonstrated above, this role is fulfilled by the passive voice. First, in the passive construction A is coded as a locative oblique, which clearly distinguishes it from other unmarked noun phrases. Second, even in the absence of an overt A, passive form of the verb signals that none of the unmarked noun phrases is the A-argument. Third, by demoting A from the subject slot, the anaphoric third-person possessor of P can no longer have A as its antecedent.

However, the decision of whether to passivize the clause or not is not solely based on formal considerations when ambiguity arises, but rather depends on the relative prominence of A, P, and their possessors. A prominent A-argument retains the subject position regardless of the presence or prominence of the possessor of the P-argument. If A is not prominent and the possessor of P is prominent, then A is demoted. The prominence of the possessor on A does not seem to play a role, except to signal that A itself is non-prominent and should be demoted when P is prominent.

The treatment of possessors in the system of Northern Khanty raises several questions. First, why is disambiguation achieved through passivization? The answer to this question lies in the central role of topicality in the morphosyntax of Northern Khanty, in contrast to the role-based morphosyntax found in many European languages. In languages where core arguments are distinguished by case marking, such as German or Russian, core arguments and possessors are coded with distinct cases or postpositions depending on their roles. Prominent possessors can also be integrated into the clause as participants in their own right, for example, receiving the dative case (see, e.g., [Shibatani 1994]).

In Northern Khanty, participants are primarily distinguished by whether they are topical or not. If a topical participant is not A, it is then promoted to the subject slot through passivization. Unlike languages with dedicated possessor case markings or postpositions, dative in Northern Khanty appears limited to marking ditransitive recipients and directional participants. Therefore, a possessor can only be indirectly promoted through the promotion of its host, the P-argument, via passivization.

Second, why do we need a broader notion of discourse prominence for Northern Khanty rather than simply adopt the idea of discourse topicality and topic maintenance⁶ proposed in [Givón (ed.) 1983] and applied in [Filchenko 2012; Sosa 2017] to Eastern Khanty and by [Koshkareva 2002] to Northern Khanty data? Corpus data show that, although this idea is in most cases applicable, there is evidence that discourse topicality is not always the primary factor in subject selection. Consider in (32) below the context of the earlier cited example (29) in more detail:

- (32) [When he traveled, he used to leave *his heart* at home. While he was riding through the forest, a man named Seven Canopies, Big Canopy, came to his house. He clung to **his wife**. “Where is his heart? Show me,” said the man named Seven Canopies, Big Canopy. “If you don’t show it, I’ll kill you right here.”]

<i>in</i>	<i>nε</i>	<i>pakən-əl</i>	<i>pa</i>	<i>ši</i>	<i>mă-λ-λe</i>	<i>ike-λ</i>	<i>sām</i>
DEM	woman	be.scared-NPST[3SG]	ADD	FOC	give-NPST-3SG>SG	man-POSS.3SG	heart

‘The woman gets scared and gives her husband’s heart.’ [The man took *the heart* and threw *it* into the waters of the Ob river.] (NHC, Younger daughter of the sun: 40)

⁶ I thank the anonymous reviewer for pointing out this competing explanation.

In this passage, four discourse participants are involved: the husband, his heart, the wife, and a sudden visitor named Seven Canopies, Big Canopy. The heart is globally topical, as it is not only repeatedly mentioned in this passage but is also important for the remaining part of the story. The woman and the visitor are both local discourse topics. The visitor, being the main actor throughout the fragment, is more topical than the woman. Since the woman is the least topical out of the three referents, we would expect the verb ‘give’ in the target sentence to be passivized and the A-argument demoted to oblique, much like in the example (11) discussed earlier. Yet we see an active clause which seems to be motivated by the fact that not only A ‘the woman’ is to certain extent topical but also that it outranks P ‘heart’ in animacy and also holds the subject pivot from the immediately preceding clause, hence being broadly discourse prominent rather than simply topical. Thus, the discourse prominence approach has a slightly better explanatory potential, though it remains yet to be systematically tested on the available data.

Third, what prevents prominent A from being demoted in the presence of possessor(s) and under threat of multiple ambiguity? In the majority of cases, prominent A is referred to with zero-anaphora or takes a pronominal form, i.e. personal pronoun or verbal indexing, whereas less prominent P and its possessor need more expressive material and thus are referred to with full NPs. Such a difference between A and other participants prevents the ambiguity, hence no need for the disambiguating passive.

P’s possessor can potentially be mistaken for an overt A, as ‘Vein-making woman’ in example (21) repeated below. However, were it for the topic shift, ‘Vein-making woman’ would have probably been stressed prosodically, segmentally and/or with word order.

- (33) *meñ kim et-əs, šiw mǎn-əs pa*
 daughter-in-law out appear-PST[3SG] DEM.LAT go-PST[3SG] ADD
λɔn werti ime-n xɔt ɔw-əl tut-ən pon-s-əλλe
 vein_making_woman-POSS.2SG house door-POSS.3SG fire-LOC put-PST-3SG>SG

‘The daughter-in-law came out, went there and set the front door of the Vein-making woman’s house on fire.’ (NHC, Vein-making woman: 11)

No clear cases of active examples with a prominent A expressed as full NP and with a possessed P are found in the corpus. There are still cases like sentence (11) repeated below that opened our discussion in Section 2, where both A ‘wife’ and P’s possessor ‘husband’ are prominent and zero or pronominally expressed.

- (34) *ši kut-ən ime-λ nox werλ-əs pa jɔš-əl*
 DEM gap-LOC wife-POSS.3SG up wake-PST[3SG] ADD hand-POSS.3SG
jira katəλ-s-a
 to.side grab-PST-PASS[3SG]

‘Meanwhile his wife woke up and grabbed his hand (lit. “his hand was grabbed”).’ (NHC, Three wise words: 70)

In such cases, regardless of A’s prominence, it is demoted via passivization. This appears to be the only effective strategy to avoid referential ambiguity between zero A and the possessive marking on P’s ‘hand’. Therefore, applying J. Hawkins’s “Maximize On-Line Processing” principle to the choice between active and passive in Northern Khanty is constrained to instances deviating from the canonical setting where A is zero/pronominal and P’s possessor is overt. These situations include cases where A is overt and P’s possessor is pronominal, or where both A and P’s possessor are equally overt or zero/pronominal, posing a high risk of role misassignment and referential conflict.

There are at least two directions for further research. First, it is necessary to expand the investigation into possessors and possessive marking in Northern Khanty and Uralic languages in general. These categories can be examined from a discourse perspective, focusing on anaphora, reference, and narrative structure. Specifically, it would be valuable to explore when a possessive noun phrase is chosen as a referential expression, how anaphoric links between possessors and their antecedents are established, and how topical chains involving possessors are constructed.

Second, gaining a deeper understanding of ambiguity resolution involving possessors requires a comprehensive study not only of morphological cues such as passive voice and objective conjugation, but also of other factors including word order, prosody, and semantics. For instance, it could be hypothesized that overtly expressed possessors are prosodically integrated with their head nouns, lacking their own intonation contour and featuring shorter pauses compared to other constituents, which facilitates their recognition as possessors.

All of the above topics would also be intriguing to explore within the realms of typology, historical linguistics, and language acquisition.

Conclusion

The study examines transitive clauses that include core arguments and their possessors. The presence of at least one possessor alongside A and P simultaneously creates ambiguity, which is resolved through various means such as passive voice, which is in focus of this investigation. The passive ambiguity-resolving strategy is employed when A is less prominent than either P or its possessor. Additionally, passive is used to resolve ambiguity when both possessors of A and P are present, and both core arguments are hierarchically low, a condition driven by considerations of processing efficiency rather than relative prominence. Affectedness of P's possessor also influences the choice of passive, although it is considered a secondary factor.

ABBREVIATIONS

1, 2, 3 — 1st, 2nd, 3rd person

ACC — accusative

AUX — auxilliary

ADD — additive

DAT — dative

DEM — demonstrative

DIM — diminutive

DIR — direct

DU — dual

EX — existential

FOC — focal

GEN — genitive

INDEF — indefinite

INV — inverse

LAT — lative

LOC — locative

NEG — negation

NFIN — non-finite

PASS — passive

PL — plural

PUNCT — punctive

POSS — possessive

PROPR — proprietive

PsrA — possessor of A

PsrP — possessor of P

PST — past

SG — singular

TOP — topic

UQ — universal quantifier

REFERENCES

- Aissen 1997 — Aissen J. On the syntax of obviation. *Language*, 1997, 73(4): 705–750.
- Aissen 2001 — Aissen J. The obviation hierarchy and morphosyntactic markedness. *Linguistica Atlantica*, 2001, 23: 1–34.
- Bárány, Nikolaeva 2019 — Bárány A., Nikolaeva I. Possessors in switch-reference. *Glossa*, 2019, 4(1): 1–36.
- Dalrymple, Nikolaeva 2011 — Dalrymple M., Nikolaeva I. *Objects and information structure*. Cambridge: Cambridge Univ. Press, 2011.
- Däbritz 2020 — Däbritz Ch. L. *Topik, Fokus und Informationsstatus: Modellierung am Material nordwestsibirischer Sprachen*. Berlin: Walter de Gruyter, 2020.
- Dixon 1972 — Dixon R. M. W. *The Dyirbal language of North Queensland*. Cambridge: Cambridge Univ. Press, 1972.

- É. Kiss 2019 — É. Kiss K. Fused grammatical and discourse functions in Ob-Ugric: Case, agreement, passive. *Proc. of the workshop "Clause Typing and the Syntax-to Discourse Relation in Head-Final Languages"*, Allensbach-Freudental, May 15–17, 2019. Bayer J., Viesel Y. (eds.). Konstanz: Universität Konstanz, 2019, 163–174.
- Fedorova 2014 — Fedorova O. The role of potential referential conflict in the choice of a referring expression. *The Russian Journal of Cognitive Science*, 2014, 1(1–2): 6–21.
- Filchenko 2012 — Filchenko A. Continuity of information structuring strategies in Eastern Khanty. *Argument Structure and Grammatical Relations. A crosslinguistic typology*. Suihkonen P., Comrie B., Solovyev V. (eds.). Amsterdam: John Benjamins, 2012, 115–131.
- Givón (ed.) 1983 — Givón T. (ed.). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam; Philadelphia: John Benjamins, 1983.
- Goddard 1990 — Goddard I. Aspects of the topic structure of Fox narratives: Proximate shifts and the use of overt and inflectional NPs. *International Journal of American Linguistics*, 1990, 56: 317–340.
- Haig 2018 — Haig G. The grammaticalization of object pronouns: Why differential object indexing is an attractor state. *Linguistics*, 2018, 56(4): 781–818.
- Hawkins 2004 — Hawkins J. A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford Univ. Press, 2004.
- Iemmolo 2011 — Iemmolo G. *Towards a typological study of differential object marking and differential object indexation*. PhD thesis. Pavia: Univ. of Pavia, 2011.
- Kaksin 2010 — Kaksin A. D. *Kazymskii dialekt hantyiskogo yazyka* [Kazym dialect of Khanty]. Khanty-Mansiysk: Ob-Ugric Institute of Applied Researches and Development, 2010.
- Kibrik 2011 — Kibrik A. A. *Reference in discourse*. Oxford: Oxford Univ. Press, 2011.
- Kortlandt 1992 — Kortlandt F. On the meaning of the Japanese passive. *Acta Linguistica Hafniensia*, 1992, 24(1): 97–108.
- Koshkareva 2002 — Koshkareva N. B. Kommunikativnaya paradigma khantyiskogo predlozheniya (na materiale kazymskogo dialekta) [Communicative paradigm of the Khanty sentence (based on the Kazym dialect data)]. *Yazyki korennykh narodov Sibiri*. Iss. 12. Koshkareva N. B., Oktyabr'skaya I. V. (eds.). Novosibirsk: Institute of Philology of the Siberian Branch of the Russian Academy of Sciences, 2002, 29–44.
- Kulonen 1989 — Kulonen U. M. *The passive in Ob-Ugrian*. Helsinki: Suomalais-Ugrilainen Seura, 1989.
- Mikhailov (to appear) — Mikhailov S. Diagnosing Northern Khanty unpossessives, or how to tell a synchronically independent marker from its diachronic source. *Linguistica Uralica*. To appear.
- Muravyev 2023 — Muravyev N. Passive in Kazym Khanty and the interaction of givenness, topicality and animacy. *Linguistica Uralica*, 2023, 59(1): 49–66.
- Nichols 1992 — Nichols J. *Linguistic diversity in space and time*. Chicago: Univ. of Chicago Press, 1992.
- Nikolaeva 1999 — Nikolaeva I. *Ostyak*. Munich: LINCOM Europa, 1999.
- Nikolaeva 2001 — Nikolaeva I. Secondary topic as a relation in information structure. *Linguistics*, 2001, 39(1): 1–50.
- Nikolaeva, Bárány 2019 — Nikolaeva I., Bárány A. Proximate possessors. *Prominent internal possessors*. Bárány A., Bond O., Nikolaeva I. (eds.). Oxford: Oxford Univ. Press, 2019, 228–258.
- Nikolaeva et al. 2019 — Nikolaeva I., Bárány A., Bond O. Towards a typology of prominent internal possessors. *Prominent internal possessors*. Bárány A., Bond O., Nikolaeva I. (eds.). Oxford: Oxford Univ. Press, 2019, 1–38.
- Shibatani 1994 — Shibatani M. An integrational approach to possessor raising, ethical datives, and adversative passives. *Proc. of the 12th Annual Meeting of the Berkeley Linguistics Society*, 1994: 461–486.
- Sosa 2017 — Sosa S. *Functions of morphosyntactic alternations, and information flow in Surgut Khanty discourse*. Doctoral diss. Helsinki: Univ. of Helsinki, 2017.
- Thompson 1994 — Thompson C. Passive and inverse constructions. *Voice and inversion*. Givón T. (ed.). Amsterdam; Philadelphia: John Benjamins, 1994, 47–63.
- von Heusinger, Schumacher 2019 — von Heusinger K., Schumacher P. B. Discourse prominence: Definition and application. *Journal of Pragmatics*, 2019, 154: 117–127.
- Wichmann 2007 — Wichmann S. The reference-tracking system of Tlapanec: Between obviation and switch-reference. *Studies in Language*, 2007, 31(4): 801–827.
- Zúñiga 2006 — Zúñiga F. *Deixis and alignment: Inverse systems in indigenous languages of the Americas*. Amsterdam; Philadelphia: John Benjamins, 2006.

Interaction between grammatical phenomena and information structure in Northern Mansi

© 2024

Daria O. Zhornik

Institute of Linguistics, Moscow, Russia;
Ludwig-Maximilians-Universität in Munich, Munich, Germany;
daria.zhornik@yandex.ru

Abstract: This article provides an overview of grammatical phenomena that are related to information structure in Northern Mansi, such as passive voice, object agreement, variation of participant coding in ditransitive constructions, non-possessive uses of the 3SG possessive marker, zero-reference, and discourse particles. I provide both a survey of previous research on the topic as well as present my suggestions on expanding the information-structural approach to these phenomena by considering other discourse-relevant parameters, such as the referential status and animacy of the participants, text genre, and narrative strategies.

Keywords: agreement, argument structure, discourse, information structure, Mansi, particles, passive, Uralic, voice

Acknowledgments: The author would like to thank Ksenia Shagal for detailed discussions about the article and extensive comments on its text, as well as the anonymous reviewer for valuable suggestions.

For citation: Zhornik D. O. Interaction between grammatical phenomena and information structure in Northern Mansi. *Voprosy Jazykoznanija*, 2024, 5: 80–96.

DOI: 10.31857/0373-658X.2024.5.80-96

Взаимодействие между грамматическими явлениями и информационной структурой в северномансийском языке

Дарья Олеговна Жорник

Институт языкознания РАН, Москва, Россия;
Мюнхенский Университет Людвига-Максимилиана, Мюнхен, Германия;
daria.zhornik@yandex.ru

Аннотация: В статье представлен обзор грамматических явлений, связанных с информационной структурой, в северномансийском языке, таких как пассивный залог, объектное согласование, вариативность кодирования участников в дитранзитивных конструкциях, дискурсивные употребления посессивного показателя 3SG, нулевая референция и дискурсивные частицы. В работе обобщаются результаты предыдущих исследований по теме, а также предлагается расширить информационно-структурный анализ указанных явлений за счет обращения к дополнительным параметрам, релевантным для структуры дискурса, включая референциальный статус и одушевленность участников, особенности жанра текста и нарративные стратегии.

Ключевые слова: аргументная структура, дискурс, залог, коммуникативная структура, мансийский язык, пассив, согласование, уральские языки, частицы

Для цитирования: Zhornik D. O. Interaction between grammatical phenomena and information structure in Northern Mansi. *Voprosy Jazykoznanija*, 2024, 5: 80–96.

DOI: 10.31857/0373-658X.2024.5.80-96

Introduction

The aim of this article is to provide an overview of grammatical phenomena influenced by information structure in Northern Mansi (< Ob-Ugric < Ugric < Finno-Ugric < Uralic¹). Ob-Ugric languages present a unique case of grammatical systems heavily governed by information structure² [Klumpp, Skribnik 2022]. It is reported that in Khanty and Mansi varieties, voice [Skribnik 2001; Filchenko 2012; Virtanen 2015; 2023; É. Kiss 2019; Urmanchieva, Plungian 2021; Muravyev, Zhornik 2024], object agreement [Nikolaeva 1999; Koshkareva 2002; Skribnik 2004; Virtanen 2015; É. Kiss 2019; Klumpp 2023], syntax of ditransitive constructions [Virtanen 2015; Bíró, Sipőcz 2017], and use of possessive markers in non-possessive functions [Janda 2015; 2019] are determined at least in part by the information-structural status of the core participants.

A recent overview of information structure in Uralic [Klumpp, Skribnik 2022] lists the following phenomena related to information structure in Northern Mansi: possessive marking (3SG in non-possessive function), object agreement, passive as a “heavy-duty topicalizer”, dative shift, pro-drop (3rd person). The main research question with regard to these categories is to determine how speakers make the choice between different categories (e.g. subject vs. object conjugation for describing the same situation) and what factors influence this choice.

In this paper I will, on the one hand, provide an overview of existing literature on information structure in Northern Mansi and Mansi in general. On the other hand, I will describe some of the phenomena connected to discourse structure in Northern Mansi that have not been previously studied as well as suggest a broader approach encompassing both information structure and a wider range of discourse-related parameters. The space allocated to each phenomenon in this article reflects the significance of their impact on the whole discourse system as well as the availability and thoroughness of previous research on the topic.

The structure of this article is as follows: In Section 1, I briefly introduce the Northern Mansi language. In Section 2, I discuss object agreement, passive voice, and variation of participant coding in ditransitive constructions. In Section 3, I describe phenomena that have been less researched, such as non-possessive uses of 3SG possessive marker and prosody, or not studied at all, such as zero-reference and particles *ta* and *ti*. In conclusion, I present a summary of the grammatical phenomena that interact with information structure in Northern Mansi and propose directions for future research in this domain.

1. The Northern Mansi language

The Northern Mansi language is the only living Mansi variety. Together, Khanty and Mansi languages form the Ob-Ugric subgroup of the Ugric branch of the Finno-Ugric language group within the Uralic language family. Previously, three other Mansi varieties existed: Eastern Mansi, Western Mansi, and Southern Mansi. Each of these languages had an individual history of language contact, and they had significant phonetic and grammatical differences (e.g. in the number

¹ This genealogical affiliation represents the traditional views on the existence of the Ob-Ugric and Ugric branches, as in [Honti 1998], see, however, [Skribnik, Laakso 2022] for a discussion of the validity of this grouping.

² In this article, I understand information structure in terms of [Krifka 2007] as local management of common ground, which includes the notions of topic, focus, and givenness, mostly analyzed at the level of a single utterance. According to [Lambrecht 1994: 131, 213], topic is defined as “whatever the proposition is about”, while focus is “the semantic component of a (...) proposition whereby the assertion differs from the presupposition”. In contrast, I use the term “discourse structure” to refer to a broader system of speech/text organization, including global common ground, narrative strategies, discourse topicality in terms of [Givón (ed.) 1983], etc.

of cases), leading to limited mutual intelligibility. There are different views on the exact number of Mansi languages. Some researchers distinguish only two or three based on lexicostatistical data, see [Koryakov 2017: 94; Koryakov et al. 2022: 74–75; Fedotova 2021].

Northern Mansi is spoken in Russia, in two neighboring regions of Western Siberia: Khanty-Mansi Autonomous Okrug (Yugra) and Sverdlovsk Oblast³. According to the latest 2020 All-Russian Population Census³, 1346 people claimed Mansi⁴ language proficiency and 1008 reported Mansi language use. Native speakers and local researchers⁵ estimate that the real number of speakers is unlikely to be higher than a thousand, possibly, it is even lower. Northern Mansi is considered a critically endangered language, see, e.g., [Skribnik, Koshkaryova 1996; Bíró, Sipőcz 2009].

The Northern Mansi language is divided into several dialectal varieties: Upper Lozva, Sosva, Sygva, and Ob'. In this article, I will provide examples from two Northern Mansi varieties: Upper Lozva and Sosva. Information structure and the grammatical phenomena in question are quite similar in different Northern Mansi varieties (although there is variation across Mansi languages, see, e.g., a study on Western Mansi passive in comparison with other Mansi varieties in [Urmanchieva, Plungian 2021]), so the exact choice of dialect is not crucial for this paper. The main source of examples for this paper is Valery Chernetsov's archive [Ryndina et. al 2022], containing Northern Mansi narratives from 1925–1938. Apart from that, examples from my personal field data⁶, studies on Mansi syntax and information structure, as well as published folklore collections [Rombandeeva 2005; Popova, Rombandeeva 2010; Kumaeva 2015] are used. The exact source is cited next to the translation of each example, except for the elicited ones, and the spelling of the original source is preserved.

2. Argument expression

2.1. General information

Ob-Ugric languages have an argument coding system which includes differential object agreement, passive voice, and variation of participant coding in ditransitive constructions. Existing works assume that these phenomena are governed to a significant degree by information structure and that different coding options represent various information-structural configurations. In Sections 2.2–2.3, I will explore previous research in this area and discuss factors responsible for the choice of argument expression strategies.

Object agreement is expressed morphologically in a dedicated subject-object agreement paradigm (2), distinct from the subject paradigm (1).

- (1) *am xūl puw-s-um*
1SG fish catch-PST-1SG
'I was fishing / I caught a fish.'
- (2) *am ti xūl ta puw-as-l-um*
1SG this fish PTCL catch-PST-SG.O-1SG.S
'I finally caught this fish.'

³ Available online at: <https://rosstat.gov.ru/vpn/2020> (Volume 5, Table 5).

⁴ In all Russian documents and sources Northern Mansi is referred to simply as Mansi.

⁵ The results of surveys carried out by the Ob-Ugric Institute of Applied Research and Development in Khanty-Mansiysk can be found at: <https://ouipiir.ru/content/мониторинг>.

⁶ Texts, as well as limited elicitation, were gathered during my fieldwork on the Upper Lozva dialect in the villages of Ushma and Treskol'je of the Idvel' District of the Sverdlovsk Oblast', Russia, in 2017–2021. Examples taken from these texts are marked with ULMC (Upper Lozva Mansi Corpus).

In a passive construction, the O-participant⁷ is promoted to subject and controls verb agreement, the verb receives the passive marker *-we*, and the A-participant is optionally expressed in the lative case, cf.:

- (3) *am ti-xōtal alpəl k̄wāl-s-um kwāl-tt-uwe-s-um ruē*
 1SG this-day morning get.up-PST-1SG get.up-CAUS-PASS-PST-1SG Russian
p̄y-na
 youngster-LAT
 ‘Today I woke up in the morning, the Russian guy woke me up.’ [ULMC: AEJ_DZh_SP_310118_myday]

Passive in Ob-Ugric is very productive, even intransitive verbs can be passivized and participants other than those expressed as the direct object in an active construction can be promoted to the subject position, as in (4), where a participant with the role of location is promoted to subject.

- (4) a. *xājtnut nēyl-ap-as m̄ān ēl-i-pāl-uw-t*
 wolf appear-MOM-PST[3SG] 1PL front-ATTR-side-POSS.1PL-LOC
 ‘A wolf appeared in front of us.’
 b. *m̄ān xājtnut-n nēyl-ap-awe-s-uw*
 1PL wolf-LAT appear-MOM-PASS-PST-1PL
 ‘A wolf appeared in front of us’ (lit. ‘We were appeared by a wolf’). [ULMC: AAA_DZh_SP_300118_wolfs]

The third coding alternation involved in this system is the so-called “dative shift” in ditransitive constructions, where there is variation in participant encoding, see [Bíró, Sipőcz 2017]. One option is to mark the recipient as a direct object, with the theme receiving instrumental case marking; the other is to mark the theme as a direct object, and the recipient as an indirect object in the lative case. This also interacts with the choice of subject vs. object agreement: in the first case both subject and object agreement can be used (in both cases S-IO-DO and S-DO-IO word orders are possible), see (5a)–(5b) from [Nikolaeva 2001], while in the latter case the object agreement is usually preferred, as in elicited (5c).

- (5) a. *am naḡən-n sūp jūnt-ēy-əm* [Nikolaeva 2001: 24]
 1SG 2SG.OBL⁸-LAT shirt sew-NPST-1SG
 b. *am sūp naḡən-n jūnt-i-l-um* [Nikolaeva 2001: 24]
 1SG shirt 2SG.OBL-LAT sew-NPST-SG.O-1SG.S
 c. *am naḡən sūp-əl jūnt-i-l-um*
 1SG 2SG.OBL shirt-INS sew-NPST-SG.O-1SG.S
 ‘I sew a shirt for you.’

Ditransitive verbs can also occur in a passive construction, and in this case the recipient is promoted to subject and the theme bears oblique case marking; the agent is usually omitted, see:

⁷ Here I introduce the notions of A- and O-participant, which refers to the syntactic roles of the constituents in a construction with two core arguments. The A-participant is either a subject of an active construction or a demoted participant (usually the agent) in a passive construction. The O-participant is either a direct object in an active construction or a subject in a passive construction.

⁸ In Northern Mansi, there is no accusative case markers, and the “accusative” forms of personal pronouns are formed by attaching a possessive marker of the same person and number to the stem of the personal pronoun, e.g. the “accusative” form of *taw* [3SG] would be *taw-e* [3SG-POSS.3SG]. All other marked cases are formed by attaching the case marker to the “accusative” form of the personal pronoun, which we gloss as an oblique form.

- (6) *n'āl-əl wār-we-s jōwt-əl wār-we-s*
 arrow-INS make-PASS-PST[3SG] bow-INS make-PASS-PST[3SG]
 'An arrow and a bow were made for him.' [Kálmán 1976: 68]

Thus, in the case of constructions with ditransitive verbs we have at least four syntactic options of participant encoding. As we will see in 2.2, previous research indicated that the choice of a particular construction in each case is determined at least in part by pragmatics.

2.2. Previous research

Recent research claims that the coding system outlined above is largely pragmatically driven. This view arose around the 1960–70s. The Ob-Ugric voice has attracted attention of researchers since the first half of the 20th century, see, e.g., [Bouda 1936]. The first accounts of Ob-Ugric and Mansi passive voice [Lavotha 1960; Honti 1982; Schiefer 1985] tried to approach this phenomenon from a semantic standpoint and determined different types of constructions depending on semantic roles of the core participants, most notably, see [Kulonen 1989]. In the studies of Evdokiya Rombandeeva, a native Mansi speaker and linguist, we find the first ideas that the choice of voice and agreement may be affected by definiteness of the constituents and the emphasis put on them, see [Rombandeeva 1979]. The early 2000s marked the beginning of the information-structural turn in Uralic studies, starting with the works of Irina Nikolaeva [1999; 2001], where she introduced the notion of secondary topic and tied passive and object agreement to the information structure in Khanty. This approach was developed by Elena Skribnik [2001; 2004] specifically for Mansi, as well as by Natal'ya Koshkareva [2002] for Khanty. This line of analysis ties agreement, voice, and variation of participant coding in ditransitive constructions into a single system, which allows for a more comprehensive description of argument expression in Ob-Ugric.

According to this line of research, argument coding in Ob-Ugric is closely tied to information structure, and different coding options represent various information-structural configurations. An active construction is used when the main topic of the utterance corresponds to the A-participant. Subject agreement is used when the O-participant is in focus, while the object agreement is triggered when the O-participant bears the information-structural role of secondary topic in Nikolaeva's terms. Passive voice is used when the O-participant, not the A-participant, is the most topical constituent in a clause; this category is viewed as the main topicalizing device in Ob-Ugric. When analyzed more broadly, beyond the scope of a single clause, passive can be viewed as a means of maintaining the current discourse topic by promoting the topical participant to subject when it is not agentive, as in (7).

- (7) {The heroes were hunting in the forest on their own, then they met Mis-hum⁹ and he decided to bring them to his village.}
ta=kōs poxt-as-iy misxum ta=min-i ta=tot-awe-y
 PTCL=CONC hold-PST-3DU Mishum PTCL=go-NPST[3SG] PTCL=carry-PASS-3DU
 'They held tight (onto Mishum), Mishum is going, and **they are being carried** (to his village).' [Chernetsov's archive 50, text 2]

Variation in ditransitive constructions is explained in the same terms. The coding can change depending on the alternations of information structure, as in the following example from [Kálmán 1976: 68] cited in [Bíró, Sipőcz 2017], which demonstrates three consecutive utterances in a narrative. In (8a), there is no topical object, thus subject agreement is used. In (8b), the recipient is topical and promoted to direct object, which triggers object agreement. In (8c), the main hero

⁹ Mis-hum is a Mansi mythological creature, a man living in the forest.

(the mythical *ēkwa pīryie*, who is asking for an arrow and a bow) is the most topical constituent; thus, he is promoted to subject via passive voice.

- (8) a. *akwēkw anəm-n n'al wār-en jōwt wār-en*
 aunt 1SG.OBL-LAT arrow make-IMP.2SG bow make-IMP.2SG
 'Auntie, make me an arrow, make me a bow.'
- b. *n'al-əl wār-i-l-əm jowt-əl wār-i-l-əm*
 arrow-INS make-NPST-SG.O-1SG.S bow-INS make-NPST-SG.O-1SG.S
 'I will make you an arrow, I will make you a bow.'
- c. *ja-t'i n'al-əl wār-we-s jowt-əl wār-we-s*
 well-PTCL2 arrow-INS make-PASS-PST[3SG] bow-INS make-PASS-PST[3SG]
 'Well, an arrow and a bow were made for him.' [Kálmán 1976: 68]

After these theoretical developments, numerous works on various Khanty and Mansi languages have appeared, where the authors analyze the phenomena of active vs. passive voice, object vs. subject agreement and variation in ditransitive constructions. Apart from Irina Nikolaeva's work on Obdorsk Khanty [Nikolaeva 2001; Dalrymple, Nikolaeva 2012], the main research in this domain for Khanty varieties was carried out by Andrey Filchenko [2006; 2007; 2012] based on Eastern Khanty data. Studies on Mansi languages are even more numerous: Susanna Virtanen [Virtanen 2015] tied all the aforementioned phenomena into a unified system for Eastern Mansi, Anna Urmanchieva and Vladimir Plungian analyzed passive in Western Mansi based on archival data gathered by Artturi Kannisto in 1901–1905 [Urmanchieva, Plungian 2021]. Bernadett Bíró and Katalin Sipőcz performed extensive research on ditransitive constructions in Ob-Ugric in general [Sipőcz 2015] as well as specifically in Mansi with main focus on Northern Mansi [Bíró, Sipőcz 2017]. Finally, Katalin É. Kiss analyzed voice and agreement phenomena in Ob-Ugric from a generative standpoint, see, e.g., [É. Kiss 2019] for an analysis of the passive as a means of demoting the [-topic] highest argument and promoting a [+topic] internal argument.

These works sought to apply the information-structural approach with topic-promotion as the main mechanism behind the phenomena in question to data from individual Ob-Ugric varieties or to describe the interaction between various parts of the argument coding system in more detail. In recent years it became apparent that, even though the information-structural approach has immensely deepened the understanding of argument coding in Ob-Ugric, there are numerous exceptions to the topic-promotion rule, at least in some Ob-Ugric varieties, specifically in Northern Mansi. For example, (9) is the first sentence of the text, where there is no topical constituent whatsoever, and still, passive is used. This might be explained by the need to demote an unknown or irrelevant A-participant in a generic context. However, such contexts are not described in the literature which considers topic-promotion to be the main function of the passive as existing exceptions.

- (9) *sāṅkwəltap jīw-əl wār-awe*
 sankwyltap wood-ABL make-PASS[NPST.3SG]
 'A sankwyltap¹⁰ is made of wood.' [ULMC: ASN_DZh_SP_310717_sankwiltap]

In (10) we see an example from a similar type of text, where generic actions during different holidays are discussed. The "wedding" is mentioned for the first time and is still promoted to subject via passivization.

- (10) *tuwəl os pūri pūri wār-awe*
 then ADD wedding wedding make-PASS[NPST.3SG]
 [Also a holiday is celebrated when a child is born.] 'Then also a wedding is celebrated.'
 [ULMC: BTP_DZh_140217_prazdniki]

¹⁰ A Mansi musical instrument.

Even though existing literature did not suggest any differences in the use of passive voice among various genres, we might suspect that texts with generic instructions function differently than narratives, and most of the examples in the existing literature come from narrative texts. However, non-topical subjects of passive constructions are found in narrative texts as well. Example (11) contains the first clauses of a narrative text, and in the third clause we find passive, although the subject “sable” is also mentioned for the first time and cannot be topical, whereas the omitted agent is topical, as the two forest men who chased down the sable are the main heroes of the text.

- (11) *kit wōr xum ōl-ē-y wōr-n min-as-iy akw mā-t*
 two forest man live-NPST-3DU forest-LAT go-PST-3DU one place-LOC
n'oxs xāŋxt-uwe-s
 sable chase.down-PASS-PST[3SG]

‘Two men live in the forest. They went to the forest. In one place they **chased down** a sable.’ [They killed the sable and went further. In one place they chased down another sable.] [Chernetsov’s archive 50, text 2]

Such examples are not that rare, and they suggest that the information-structural status of the core participants is a very significant but not the only parameter influencing the choice of argument coding in Ob-Ugric. In the most recent works, attempts to identify other parameters were undertaken, in [Muravyev, Zhornik 2023; 2024] for passive in Ob-Ugric, in [Muravyev 2023] for passive in Northern Khanty, in [Klump 2023] for differential object marking in West Mansi¹¹. We present this new line of analysis in the next section.

2.3. Parameters complementing the information-structural approach

The recent approach to studying argument coding in Ob-Ugric outlined in [Klump 2023], [Muravyev 2023], and [Muravyev, Zhornik 2024] considers information structure to be an important factor but also stresses the need to include other parameters for a more comprehensive analysis. This is in line with recent typological developments, such as in [Just 2024], where differential A- and O-indexing is analyzed cross-linguistically with regard to identifiability, animacy, and topicality of the core arguments. For Finno-Ugric direct objects, a similar set of parameters was outlined in [Serdobolskaya, Toldova 2012] and [Toldova, Serdobolskaya 2017].

In [Klump 2023], we find a broader set of parameters, including (but not limited to) possessor marking, givenness, animacy, the syntactic function of the antecedent, distance from it, and word order. Applying frequency analysis, the author concludes, for example, that dative-lative marking in West Mansi appears on given objects referred to with a full noun phrase and occurs significantly more often with referents of high animacy and referents with direct givenness (and not, e.g., givenness via a given possessor). Such conclusions are not possible within a purely information-structural analysis, and a broader set of parameters might account for the “exceptions” we encountered earlier. In [Muravyev 2023], a similar set of parameters is used for analyzing the choice of voice in Northern Khanty. The author concludes that at least givenness and animacy, apart from topicality, influence the use of passive voice, and that other parameters may be added in the future to enable more comprehensive studies of these phenomena in Ob-Ugric.

¹¹ West Mansi is unique among the Mansi languages in that it has a category of dative-lative objects, which are described in [Klump 2023]. Northern Mansi lacks differential object marking because the accusative case is absent in the language, and the dative-lative case is not used to mark direct objects. Thus, all objects in Northern Mansi bear the unmarked “nominative” form, except for personal pronouns which have an “accusative form”, and the differential object marking is carried out only by subject vs. object agreement. Eastern and Southern Mansi possess DOM via nominative vs. accusative marking.

In [Muravyev, Zhornik 2023], we made a first attempt of creating a parametric voice model in Ob-Ugric, where we introduced the parameters of animacy, definiteness, and topicality (discourse topic as described in [Givón (ed.) 1983]) for both A- and O-participants, as well as volitionality for the A-participant and affectedness for the O-participant. However, it was lacking some important parameters accounting for the use of voice in Ob-Ugric. In [Muravyev, Zhornik 2024], we modified this approach by adding a more fine-grained set of parameters related to the referential status of the core participants, see parameters in [Gundel et al. 1993], as well as definiteness according to [Roberts 2003], and specificity according to [von Heusinger 2002]. We concluded that the use of passive is driven not by information structure alone but by a complex interaction of various properties of both core arguments. This approach can explain cases of A-demotion when A is specific indefinite, non-specific indefinite, generic, or definite new and P is not topical; hence, it has no apparent reason to be promoted. We also discovered that different parameters have various significance for the A- and O-participants, for example, animacy is a more significant parameter for the A-participant than for the O-participant. Based on the combination of all relevant factors, we proposed the following hierarchy of passivization for Ob-Ugric (the higher a participant is on the hierarchy, the more likely it is to be promoted in case of P or demoted in case of A):

A-demotion cline: Inanimate > Non-specific or generic > Specific indefinite > Definite non-topical > Topical > Speech act participant

P-promotion cline: Speech act participant > Topical > Episode-central > Definite > Weak definite > Specific indefinite > Non-specific or generic

This hierarchy is a preliminary attempt to unify within one scheme all the relevant parameters, namely person, animacy, topicality, definiteness, and specificity, resulting in a combination of the prominence hierarchy (see prominence hierarchies¹² in [Aissen 2003]) and the referential hierarchy¹³ (see [Gundel et al. 1993]), which also accounts for the fact that some parameters are more relevant for one participant than the other.

Moreover, we noticed that text genre plays a significant role in argument coding. In (9) repeated below as (12), we encountered an example from a procedural text, containing instructions on how to create a traditional Mansi instrument.

- (12) *sāṅkwəltap jīw-əl wār-awe*
 sankwyltap wood-ABL make-PASS[NPST.3SG]
 ‘A sankwyltap is made of wood.’ [ULMC: ASN_DZh_SP_310717_sankwiltap]

In this text, there is no definite A-participant, and the O-participant is generic, as well as the actions described in the text, and the discourse structure is significantly different from the one we encounter in narratives. Our recent research suggests that in narratives such factors as episode boundaries, narrative roles of the participants (protagonist, antagonist, etc.), and other plot characteristics may influence argument coding. However, this is a topic for future studies.

Thus, the current goal of further research on argument coding in Ob-Ugric is to determine the set of factors that influence the choice of voice and agreement and to provide a comprehensive description of the mechanism discussed in this section. This type of research requires a unified annotation for Ob-Ugric narrative texts, which would then facilitate a qualitative analysis. Since

¹² Aissen [2003] proposes two prominence hierarchies:

- a. Animacy scale: Human > Animate > Inanimate
- b. Definiteness scale: Personal pronoun > Proper name > Definite NP > Indefinite specific NP > Non-specific NP

¹³ In [Gundel et al. 1993], the following givenness hierarchy is proposed to represent the cognitive statuses relevant for the choice of referring expressions in discourse:

in focus > activated > familiar > uniquely identifiable > referential > type identifiable

text type influences argument coding, it would seem fruitful to zoom in on the use of these phenomena in narratives, as they are the most abundant type of material for Northern Mansi and Ob-Ugric in general, and then expand the analysis to other genres, such as dialogues and procedural texts.

3. Less-researched phenomena

Phenomena described in this and the next section have received much less attention than Ob-Ugric argument coding, and they probably exert less influence on the whole linguistic system. Nevertheless, they are important for understanding the general discourse structure of Northern Mansi and are mentioned in the overview [Klumpp, Skribnik 2022]. These categories include non-possessive uses of the 3SG possessive marker, zero-reference, discourse particles, and prosody. With regard to prosody, an analysis of audio data is required, hence I am not able to provide any insight into that domain as of yet. Some initial research results specifically on the prosodic marking of focal constituents in Northern Mansi were outlined in [Pokrovskaya 2022], where it was concluded that the focal constituent is marked by pitch rising on the stressed syllable of the corresponding word. Apart from that, no research on the interaction between Northern Mansi information structure and prosody has been carried out.

3.1. Non-possessive functions of POSS.3SG

Northern Mansi, as other Uralic languages, exhibits a set of possessive markers for each person and number combination. They can be used to express the usual range of possessive semantics (13). However, the POSS.3SG marker *-(t)e* is also used in so-called “non-possessive” contexts, see [Janda 2015], where the presence of the possessive marker is determined by the information-structural context and no possessive relation is present (14):

- (13) *ɬay-äyi-te kēr sāli-ūɬ χum-ɬy ʃh̄s-i-t-e*
 elder-girl-POSS.3SG iron reindeer-animal man-TRANS have-NPST-SG.O-3SG.S

‘His sister has the Iron Wolf as her husband.’ [Chernetsov’s archive 50, text 1]

- (14) {The Iron Wolf wants to eat the main hero. The main hero was in the bathhouse.}

kon ta=kwal-ap-vs, sar lāsāl,
 outside PTCL=GO.OUT-MOM-PST[3SG] HORT slow

uɬ-e nupəl lāw-i
 animal-POSS.3SG towards say-NPST[3SG]

‘He went outside, “Wait a moment”, he says to **the beast**.’ [Chernetsov’s archive 50, text 1]

This function of possessive suffixes is attested in most Uralic languages (see [Janda et al. 2022; Fraurud 2001; Nikolaeva 2003; Simonenko 2014]). However, the range of contexts where such possessives are used as well as which person and number markers can be used in a non-possessive function differ, see, e.g., [Khomchenkova 2022] for Hill Mari or [Serdobolskaya et al. 2017] for Beserman Udmurt. In Northern Khanty, the POSS.2SG marker can be used outside of possessive contexts [Mikhailov 2023; 2024], while in Northern Mansi such uses are only possible for the POSS.3SG marker, see [Janda 2015]. Previous research referred to these non-possessive uses as definiteness marking, see [Gerland 2014]. In [Janda 2015; 2019], however, it is concluded that the Northern Mansi POSS.3SG marker is a referential device used for information structure purposes. In some other Uralic languages (e.g. Hill Mari, see [Khomchenkova 2022]) possessive markers are used to denote selection from a set, and in Northern Mansi we also find such

contexts, see (15) adapted from [Janda 2015], where in (15a) and in the text before this sentence two men are considered as a set, and in (15b) only one member of this set is being referred to:

- (15) a. *kittiy ta=min-as-iy*
 apart PTCL=go-PST-3DU
 ‘They went their own ways’.
- b. *mān’ei jāniy ōjka-te noms-i*
 Mansi big man-POSS.3SG think-NPST[3SG]
 ‘The big Mansi man is thinking.’ [Janda 2015: 254]

It seems important to further research the connection between non-possessive uses of the POSS.3SG marker and the information structural status of the constituents, as well as determine other parameters relevant for the use of POSS.3SG.

3.2. Zero-reference to 3rd-person participants

In Northern Mansi narratives, full noun phrases and even pronouns are quite rarely used for referring to a participant, and zero-reference is common (see [Klumpp, Skribnik 2022: 1033] for a discussion of this phenomenon in Uralic under the notion of “pro-drop”). Full noun phrases are usually used to introduce a participant for the first time (16), to reactivate it when it has not been mentioned for a long time (17), or to distinguish between participants when several of them perform actions in the current episode (18).

- (16) {As the first sentence of the text:}
la an xum laγ-āyi-tēn-təl 5l-ēly
 so one man elder-girl-POSS.3DU-INS live-NPST.3DU
 ‘So one man lives with his sister.’ [Chernetsov’s archive 50, text 1]
- (17) {The Iron Wolf was mentioned once over 100 clauses ago.}
taχ kēr sāli āi-n ta=al-awe-m taχ
 FUT iron reindeer animal-LAT PTCL=kill-PASS[NPST]-1SG FUT
 ‘(It seems that) the Iron Wolf will kill me.’ [Chernetsov’s archive 50, text 1]
- (18) {The animal companions are trying to remove a fang which killed their master from his ear and change turns.¹⁴}
 a. *s5rup 5jka ta=sipy-i*
 moose man PTCL=suck-NPST[3SG]
 ‘The moose is sucking (the fang out).’
 b. *P’eu i5wer sipy-i*
 lion animal suck-NPST[3SG]
 ‘The lion is sucking (the fang out).’
 c. *tuvəl ras 5jka sipy-unjke pat-s*
 then lynx man suck-INF begin-PST[3SG]
 ‘Then the lynx started to suck (the fang out).’ [Chernetsov’s archive 50, text 1]

The 3rd person pronouns are used even rarer, usually to put emphasis on the participant, e.g. to contrast it with another participant, see (19).

¹⁴ The whole story is published with Russian glossing and translation in [Zhornik 2020].

- (19) {The hero is washing in the bathhouse to delay being eaten by the Iron Wolf. A raven comes to tell him that his animal companions are coming to save him, and he needs to buy a little more time.}
- χulaχ-e ta=min-as tau ta=lout-χat-i*
 raven-POSS.3SG PTCL=go-PST[3SG] 3SG PTCL=wash-REFL-NPST[3SG]
 ‘The raven left, and he (the main hero) continues to wash himself.’ [Chernetsov’s archive 50, text 1]

Thus, most of the topical and activated participants in a Northern Mansi narrative are referred to with zero-reference. In the literature, it is usually suggested that topical constituents enable such reduced referential means. However, in Northern Mansi both participants of a transitive clause are frequently referred to with zero-reference, and the participants are only encoded within the verb form via subject and object marking, as in (20).

- (20) *kolas nākw-əl ta=titt-i-t-e ti=titt-as-t-e ti=pusm-as*
 flour dust-INS PTCL=feed-NPST-SG.O-3SG.S PTCL2=feed-PST-SG.O-3SG.S PTCL2=recover-PST[3SG]
 ‘(The brother) is feeding (his sister) with flour dust, (he) fed (her), (she) recovered (from her illness).’ [Chernetsov’s archive 50, text 1]

In Northern Mansi, zero-reference is a means of tracking a highly activated or topical referent and it is intertwined with the system of voice and agreement (see also [Toldova, Volkova 2021] for Northern Khanty), whereby, for example, a topical constituent can be promoted to subject in passive construction and also be referred to with zero anaphora to signify its highly activated status, see (21).

- (21) {The main hero was killed by his sister.}
- lay-ayi-tè-n ta=tot-wè-s kusnès pält*
 elder-girl-POSS.3SG-LAT PTCL=bring-PASS-PST[3SG] blacksmith to
 ‘His sister brought him to the blacksmith (to put him inside a barrel).’ [Chernetsov’s archive 50, text 1]

Thus, it is important to research the connection between the choice of voice and agreement and referential means used to refer to the participants.

3.3. Particles *ta* and *ti*

Another phenomenon in Northern Mansi possibly pertaining to information structure includes some of the particles. Northern Mansi has a large number of various particles, and it has been suggested that some of them attach to focal constituents, see examples (22) and (23) (cited in [Klump, Skribnik 2022: 1021]), where the conditional particle *=ke* is attached to the elements under focus:

- (22) *am=ke särt täyint-i naŋ māj-t-ən am lēwat-ēy-əm*
 1SG-COND first fill-NPST[3SG] 2SG liver-POSS.2SG 1SG pull.out-NPST-1SG
 ‘If MINE (yarnwinder) will be pulled first, I will pull out your liver.’ [Kannisto, Liimola 1951: 205]
- (23) *juw jox-t-ēy-ən akʷ-ən mat-as=ke*
 to.home come-NPST-2SG aunt-POSS.2SG age-PST[3SG]=COND
manər wār-i-l-ən
 what do-NPST-SG.O-2SG.S
 ‘When you come home, if your aunt has become old, what will you do (with her)?’ [Chernetsov’s archive 44, text 5]

However, in this case only the position of the particle is determined by the information structure, not its presence or absence. In my research, I encountered other particles that may be dependent on certain information-structural, or at least broader discourse properties of a clause or a chunk of text where it appears. Two particles exhibit this behavior: *ta*, which is very frequent, and *ti*, which is significantly rarer. These particles are diachronically connected to the demonstrative pronouns *ta* ‘that’ and *ti* ‘this’. However, the pronouns and particles can be discerned by their position. The demonstratives appear as dependents in a noun phrase, while the particles usually appear as proclitics before verbs, as in (24), and they even can be inserted between a preverb and a verb, see (25).

- (24) *ti l'ōηχ χ̣ṣit ti=tot-ima*
 this way along PTCL2=bring-MIR.PASS[3SG]
 ‘He was brought along this way.’ [Chernetsov’s archive 50, text 1]

- (25) *tuwəl t̄ān xot=ta=tip-s-ət*
 then 3PL off=PTCL=get.lost-PST-3PL
 ‘Then they got lost.’

Particles are barely mentioned in existing literature on Northern Mansi: grammars [Rombandeeva 1973] and [Riese 2001] only provide a short list labeling them as “specifying”. In the latest overview [Bakró-Nagy et al. 2022], they are not mentioned at all. Several pages in [Saynakhova 1966: 179–185] are dedicated to the particles *ta* and *ti*. However, examples given in the dissertation include the demonstrative pronouns *ta* ‘that’ and *ti* ‘this’ and do not provide a clear picture of the functions of the particles.

My analysis showed that the particle *ta* is extremely frequent in narrative texts: in the narratives from Valery Chernetsov’s archive (containing mostly Mansi texts from 1925–1938), *ta* is present on average in 25 % of clauses¹⁵. One of the more frequent contexts where it appears includes two consecutive clauses, which are usually identical except for the presence of *ta* in the first one, see (26) and (27).

- (26) a. *aku lot luu ta=min-as-iy*
 one COM to.home PTCL=go-PST-3DU
 ‘(They) went home together.’
 b. *luu min-as-iy, lun χ̣l-s-iy*
 to.home go-PST-3DU home stay.overnight-PST-3DU
 ‘Went home, spend the night at home.’ [Chernetsov’s archive 50, text 1]
- (27) a. *luu=ta=salt-s-ət*
 home=PTCL=enter-PST-3PL
 ‘(They) went inside.’
 b. *luu salt-s-ət, uq-an-e kolkan-na lot=ta=χ̣u-l̄yt-as-ət*
 inside enter-PST-3PL animal-PL-POSS.3SG floor-LAT down=PTCL=lie.down-MOM-PST-3PL
 ‘(They) went inside, his animals laid down on the floor.’ [Chernetsov’s archive 50, text 1]

These instances can be classified as recapitulative bridging linkage constructions [Guérin, Aiton 2019], which include verbatim or very close repetition of the preceding clause before a new chunk of narrative. In this case, *ta* marks the beginning of a bridging construction, which might signal a turn in the narrative or a beginning of a new episode. However, this is

¹⁵ Ten narratives from Valery Chernetsov’s archive (ranging from 300 to 900 clauses) were split into clauses, for each text I counted the total amount of clauses and the amount of clauses containing *ta* and calculated the percentage of clauses with the particle for each text and the average percentage for the ten texts.

only one type of context where *ta* can appear and further research on the role of this particle is needed. Rarely, such repetition is possible without the particle *ta*, but in most cases it is present. It seems that in cases where the particle is absent, we find a more stative, non-perfective situation, see (28).

- (28) a. *ɬay-āyi-te* *ɬun* *ʃl-i*
 elder-daughter-POSS.3SG at.home be-NPST[3SG]
 ‘His older sister is at home.’
- b. *ɬun* *ʃl-i*, *akw* *ēr-t* *kēr* *sāli-ū*
 at.home be-NPST[3SG] one moment-LOC Iron reindeer-animal
sarit-ime *āras-n* *tu* *min-ās*
 burn-PTCP.RES fire-LAT there go-PST[3SG]
 ‘(She) is at home, at one point she went to the fire where the Iron Wolf was burnt.’ [Chernetsov’s archive 50, text 1]

The particle *ti* is much rarer, ranging approximately between 5–7 % of the clauses, e.g. in a text “The tale about the Iron Wolf” from Valery Chernetsov’s archive among 390 clauses there are 122 instances of *ta* and 24 instances of *ti*. The functions of *ti* (see (29)) are even less clear than those of *ta* and require further investigation.

- (29) {The hero’s sister got sick and sent her brother to find a special type of flour to heal her.}
añ *ɬay-āyi-te* *iŋ* *aym-əŋ*, *kolas* *ñakw-əl*
 now elder-girl-POSS.3SG still illness-PROP flour dust-INS
ta=titt-i-t-e, *ti=titt-əs-t-e*, *ti=pusm-əs*
 PTCL=feed-NPST-SG.O-3SG.S PTCL2=feed-PST-SG.O-3SG.S PTCL2=recover-PST[3SG]
 ‘Now his sister is still sick, he is feeding her with flour dust, he fed her, she recovered.’
 [Chernetsov’s archive 50, text 1]

Further analysis of both of these particles is needed to figure out their functions and types of contexts where they are used. The next step of analysis could be to figure out whether these particles specifically interact with other phenomena discussed in this article.

Conclusion

Northern Mansi, as other Ob-Ugric languages, has a multitude of grammatical phenomena that are at least partially conditioned by information structure. This peculiarity has attracted the attention of many researchers over the years and most of their efforts were aimed at investigating the argument coding system and its connection to information structure proper, that is, to the notions of topic, focus and, in some instances, givenness. Recent research shows two possible pathways of development in this area. First, categories related to argument expression, such as choice of agreement, voice and participant coding in ditransitive constructions might be to some degree influenced by other factors, such as the referential status or animacy of the participants. It seems beneficial to combine them with the information-structural status of the constituents into a single comprehensive analysis. Secondly, it is important to understand that the argument coding system does not function independently from other phenomena in the language. Some of these phenomena, e.g. referential means, non-possessive uses of possessive markers and discourse particles have been reported to be influenced by information structure as well, see [Klumpp, Skribnik 2022]. At first, more research on specific phenomena is needed to understand their functions and the parameters affecting their choice more clearly. The next step could be to unify all the phenomena mentioned above, including the argument coding system, into a single analysis, where the discourse structure of Northern Mansi would be analysed

as a whole comprehensive system. A significant challenge at this stage is dealing with differences among text genres in Northern Mansi. Narratives should be analyzed differently from texts containing generic descriptions and instruction. However, the general direction of such research looks promising and a step-by-step approach to gradually covering more ground within the domain of Northern Mansi information structure and discourse might be successful.

ABBREVIATIONS

1, 2, 3 — 1st, 2nd, 3rd person

ABL — ablative

ADD — additive

ATTR — attributive

CAUS — causative

COM — comitative

CONC — concessive

COND — conditional

DU — dual

FUT — future

HORT — hortative

IMP — imperative

INF — infinitive

INS — instrumental

LAT — lative

LOC — locative

MIR — mirative

MOM — momentative

NPST — non-past

O — object

OBL — oblique

PASS — passive

PL — plural

POSS — possessive

PROP — proprietive

PST — past

PTCL — particle *ta*

PTCL2 — particle *ti*

PTCP.RES — resultative participle

REFL — reflexive

S — subject

SG — singular

TRANS — translativ

REFERENCES

- Aissen 2003 — Aissen J. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 2003, 21(3): 435–483.
- Bakró-Nagy et al. (eds.) 2022 — Bakró-Nagy M., Laakso J., Skribnik E. (eds.). *The Oxford guide to the Uralic languages*. Oxford: Oxford Univ. Press, 2022.
- Bakró-Nagy et al. 2022 — Bakró-Nagy M., Sipőcz K., Skribnik E. North Mansi. *The Oxford guide to the Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 537–564.
- Bíró, Sipőcz 2009 — Bíró B., Sipőcz K. Language shift among the Mansi. *Variation in indigenous minority languages*. Stanford J., Preston D. R. (eds.). Amsterdam: John Benjamins, 2009, 321–346.
- Bíró, Sipőcz 2017 — Bíró B., Sipőcz K. The Mansi ditransitive constructions. *Finno-Ugric Languages and Linguistics*, 2017, 6(1): 41–55.
- Bouda 1936 — Bouda K. Das obugrische Passivum. *Ungarische Jahrbücher*, 1936, 16: 286–293.
- Dalrymple, Nikolaeva 2012 — Dalrymple M., Nikolaeva I. *Objects and information structure*. Cambridge: Cambridge Univ. Press, 2012.
- É. Kiss 2019 — É. Kiss K. Fused grammatical and discourse functions in Ob-Ugric: Case, agreement, passive. *Proc. of the workshop “Clause Typing and the Syntax-to Discourse Relation in Head-Final Languages”, Allensbach-Freudental, May 15–17, 2019*. Bayer J., Viesel Y. (eds.). Konstanz: Universität Konstanz, 2019, 163–174.
- Fedotova 2021 — Fedotova I. V. Dialektnoe chlenenie mansiiskogo yazyka v diakhronii po dannym bazisnoi leksiki [Dialect separation of the Mansi language in diachrony according to the basic vocabulary]. *Bulletin of Ugric Studies*, 2021, 11(2): 338–346.
- Filchenko 2006 — Filchenko A. The Eastern Khanty locative-agent constructions: A functional discourse-pragmatic perspective. *Demoting the agent: Passive, middle and other voice phenomena*. Lynghfelt B., Solstad T. (eds.). Amsterdam: John Benjamins, 2006, 47–82.
- Filchenko 2007 — Filchenko A. *A grammar of Eastern Khanty*. Doctoral thesis. Houston: Rice Univ., 2007.
- Filchenko 2012 — Filchenko A. Continuity of information structuring strategies in Eastern Khanty. *Argument structure and grammatical relations: A crosslinguistic typology*. Suihkonen P., Comrie B., Solovyev V. (eds.). Amsterdam; Philadelphia: John Benjamins, 2012, 115–131.

- Fraurud 2001 — Fraurud K. Possessives with extensive use: A source of definite articles? *Dimensions of possession*. Baron I., Herslund M., Sørensen F. (eds.). Amsterdam; Philadelphia: John Benjamins, 2001, 243–268.
- Gerland 2014 — Gerland D. Definitely not possessed? Possessive suffixes with definiteness marking function. *Frames and concept types: Applications in language and philosophy*. Gamerschlag Th., Gerland D., Osswald R., Petersen W. (eds.). Cham; Heidelberg: Springer, 2014, 269–292.
- Givón (ed.) 1983 — Givón T. (ed.). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam; Philadelphia: John Benjamins, 1983.
- Guérin, Aiton 2019 — Guérin V., Aiton G. Bridging constructions in typological perspective. *Bridging constructions*. Guérin V. (ed.). Berlin: Language Science Press, 2019, 1–44.
- Gundel et al. 1993 — Gundel J. K., Hedberg N., Zacharski R. Cognitive status and the form of referring expressions in discourse. *Language*, 1993, 64: 274–307.
- Honti 1982 — Honti L. Das Passiv in den obugrischen Sprachen. *Acta Linguistica Academiae Scientiarum Hungaricae*, 1982, 32(1–4): 39–51.
- Honti 1998 — Honti L. Die ob-ugrischen Sprachen. *The Uralic languages*. Sinor D. (ed.). Leiden: Brill, 1998, 147–196.
- Janda 2015 — Janda G. E. Northern Mansi possessive suffixes in non-possessive function. *Referential devices in Uralic languages*. Pajusalu R., Klumpp G., Laury R. (eds.). Tartu: Tartu Univ. Press, 2015, 243–258.
- Janda 2019 — Janda G. E. *Funktionen von Possessivsuffixen in den ugrischen Sprachen*. Cologne: Modern Academic Publ., 2019.
- Janda et al. 2022 — Janda G. E., Laakso J., Metslang H. Person marking. *The Oxford guide to the Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 894–903.
- Just 2024 — Just E. A structural and functional comparison of differential A and P indexing. *Linguistics*, 2024, 62(2): 295–321.
- Kálmán 1976 — Kálmán B. *Wogulische Texte mit einem Glossar*. Budapest: Akadémiai Kiadó, 1976.
- Kannisto, Liimola 1951 — Kannisto A. (comp.), Liimola M. (ed.). *Wogulische Volksdichtung 1: Texte mythischen Inhalts*. Helsinki: Suomalais-Ugrilainen Seura, 1951.
- Khomchenkova 2022 — Khomchenkova I. A. Diskursivnye funktsii posessivnogo pokazatelya tret'ego litsa edinstvennogo chisla v gornomariiskom yazyke [Discourse functions of the 3rd singular possessive marker in Hill Mari]. *Ural-Altaic Studies*, 2022, 45(2): 148–168.
- Klumpp 2023 — Klumpp G. On dative-lative encoded direct objects in West Mansi. *Linguistica Uralica*, 2023, 59(4): 307–338.
- Klumpp, Skribnik 2022 — Klumpp G., Skribnik E. Information structuring. *The Oxford guide to the Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 1018–1036.
- Koshkareva 2002 — Koshkareva N. B. Kommunikativnaya paradigma khantyiskogo predlozheniya (na materiale kazym'skogo dialekta) [Communicative paradigm of Khanty sentence (based on the Kazym dialect data)]. *Yazyki korennykh narodov Sibiri*. Iss. 12. Koshkareva N. B., Oktyabr'skaya I. V. (eds.). Novosibirsk: Institute of Philology of the Siberian Branch of the Russian Academy of Sciences, 2002, 29–44.
- Koryakov 2017 — Koryakov Yu. B. Problema «yazyk ili dialekt» i popytka leksikostatisticheskogo podkhoda [Language vs. dialect: A lexicostatistic approach]. *Voprosy Jazykoznanija*, 2017, 6: 79–101.
- Koryakov et al. 2022 — Koryakov Yu. B., Davidiyuk T. I., Kharitonov V. S., Evstigneeva A. P., Syuryun A. A. *Spisok yazykov Rossii i statusy ikh vital'nosti* [List of languages of Russia and their vitality statuses]. Moscow: Institute of Linguistics of the Russian Academy of Sciences, 2022.
- Kumaeva 2015 — Kumaeva M. V. *Skazki, pesni, zagadki naroda mansi* [Fairy tales, song, and riddles of the Mansi people]. Khanty-Mansiysk: Yugorskii format, 2015.
- Krifka 2007 — Krifka M. Basic notions on information structure. *Interdisciplinary studies on information structure*. Vol. 6: *The notions of information structure*. Féry C., Fanselow G., Krifka M. (eds.). Potsdam: Universitätsverlag Potsdam, 2007, 13–55.
- Kulonen 1989 — Kulonen U.-M. *The passive in Ob-Ugrian*. Helsinki: Suomalais-Ugrilainen Seura, 1989.
- Lambrecht 1994 — Lambrecht K. *Information structure and sentence form*. Cambridge: Cambridge Univ. Press, 1994.
- Lavotha 1960 — Lavotha Ö. Das Passiv in der wogulischen Sprache. *Journal de la Société Finno-Ougrienne*, 1960, 62: 3–34.
- Mikhailov 2023 — Mikhailov S. K. Northern Khanty possessives and determiner typology. *Rodnoy Yazyk*, 2023, 2: 6–51.
- Mikhailov 2024 — Mikhailov S. K. Semantics of the Northern Khanty Salient Article: Definiteness, saliency, and obviation. *Voprosy Jazykoznanija*, 2024, 1: 7–38.

- Muravyev 2023 — Muravyev N. Passive in Kazym Khanty and the interaction of givenness, topicality and animacy. *Linguistica Uralica*, 2023, 59(1): 49–66.
- Muravyev, Zhornik 2023 — Muravyev N., Zhornik D. Parametricheskaya model' zaloga v obsko-ugorskih yazykakh [A parametric model of voice in Ob-Ugric]. *Bulletin of Ugric Studies*, 2023, 13(1): 112–124.
- Muravyev, Zhornik 2024 — Muravyev N., Zhornik D. Passive in Ob-Ugric: Information structure and beyond. *Finnisch-Ugrische Mitteilungen*, 2024, 47(4): 67–95.
- Nikolaeva 1999 — Nikolaeva I. Object agreement, grammatical relations, and information structure. *Studies in Language*, 1999, 23: 331–376.
- Nikolaeva 2001 — Nikolaeva I. Secondary topic as a relation in information structure. *Linguistics*, 2001, 9(1): 1–50.
- Nikolaeva 2003 — Nikolaeva I. Possessive affixes in the pragmatic structuring of the utterance: Evidence from Uralic. *Proc. of the International Symposium on Deictic Systems and Quantification*. Suihkonen P. (ed.). Izhevsk: Udmurt State Univ.; Leipzig: Max Planck Institute for Evolutionary Anthropology, 2001, 130–145.
- Pokrovskaya 2022 — Pokrovskaya S. V. Fokus i ego prosodicheskoe markirovanie v verkhneloz'vinskom dialekte severnogo mansiiskogo yazyka [Focus and its prosodic marking in Upper Lozva Mansi]. *Tomsk Journal of Linguistics and Anthropology*, 2022, 1(35): 46–55.
- Popova, Rombandeeva 2010 — Popova S. A., Rombandeeva S. M. *Wōrajan māxum mōjīat. Skazki okhotnikov* [Hunters' fairy tales]. Khanty-Mansiysk: Pechatnoe delo, 2010.
- Roberts 2003 — Roberts C. Uniqueness in definite noun phrases. *Linguistics and Philosophy*, 2003, 26(3): 287–350.
- Riese 2001 — Riese T. *Vogul*. Munich: Lincom Europa, 2001.
- Rombandeeva 1973 — Rombandeeva E. I. *Mansiiskii (vogul'skii) yazyk* [The Mansi (Vogul) language]. Moscow: Nauka, 1973.
- Rombandeeva 1979 — Rombandeeva E. I. *Sintaksis mansiiskogo (vogul'skogo) yazyka* [Syntax of Mansi (Vogul)]. Moscow: Nauka, 1979.
- Rombandeeva 2005 — Rombandeeva E. I. *Mify, skazki, predaniya mansi (vogulov)* [Myths, fairy tales, and stories of the Mansi (Voguls)]. Moscow-Novosibirsk: Nauka, 2005.
- Ryndina et. al 2022 — Ryndina O. M., Zhornik D. O., Barsukov E. V., Pokrovskaya S. V. Arkhiv V. N. Chernetsova: mansiiskaya etnografiya, fol'klor i lingvisticheskaya perspektiva [Valery Chernetsov's archive: Mansi ethnography, folklore and linguistic perspective]. *Voprosy Jazykoznaniya*, 2022, 5: 35–50.
- Saynakhova 1966 — Saynakhova A. I. *Sluzhebnye slova v mansiiskom yazyke* [Functional words in Mansi]. Doctoral diss. Moscow: Institute of Linguistics of the Russian Academy of Sciences, 1966.
- Serdobolskaya, Toldova 2012 — Serdobolskaya N. V., Toldova S. Yu. Differentsirovanoe markirovanie pryamogo dopolneniya v finno-ugorskih yazykakh [Differential marking of the direct object in Finno-Ugric]. *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisaniya. Formal'nyi i funktsional'nyi podkhody*. Kuznetsova A. I. (ed.). Moscow: Rukopisnye pamyatniki Drevnei Rusi, 2012, 59–142.
- Serdobolskaya et al. 2017 — Serdobolskaya N., Usacheva M., Arkhangel'skiy T. Grammaticalization of possessive markers in the Beserman dialect of Udmurt. *Possession in languages of Europe and North and Central Asia*. Johansson L., Mazzitelli L. F., Nevskaya I. (eds.). Amsterdam; Philadelphia: John Benjamins, 2017, 291–312.
- Schiefer 1985 — Schiefer E. *Das Passiv und der mediale Gebrauch passiver Formen im Obugrischen*. München: L. & E., 1985.
- Simonenko 2014 — Simonenko A. Microvariation in Finno-Ugric possessive markers. *NELS 43: Proc. of the 43th annual meeting of the North East Linguistic Society*. Vol. 2. Huang H.-L., Poole E., Rysling A. (eds.). Amherst: GLSA, 2014, 127–140.
- Sipőcz 2015 — Sipőcz K. Ditransitivity in the Ob-Ugric languages. *Congressus Duodecimus Internationalis Fenno-Ugristarum. Plenary Papers*. Mantila H., Leinonen K., Brunni S., Sivonen J., Palviainen S. (eds.). Oulu: Univ. of Oulu, 2015, 133–159.
- Skribnik 2001 — Skribnik E. Pragmatic structuring in Northern Mansi. *Congressus Nonus Internationalis Fenno-Ugristarum 2. Pars VI. Dissertationes sectionum: Linguistica III*. Seilenthal T. (ed.). Tartu: Auctores, 2001, 222–239.
- Skribnik 2004 — Skribnik E. K. Kategorii mansiiskogo glagola i aktual'noe chlenenie predlozheniya [Mansi verb categories and communicative sentence structure]. *Tipologicheskie obosnovaniya v grammatike. K 70-letiyu prof. V. S. Khrakovskogo*. Volodin A. P. (ed.). Moscow: Znak, 2004, 445–458.
- Skribnik, Koshkaryova 1996 — Skribnik E., Koshkaryova N. Khanty and Mansi: the contemporary linguistic situation. *Shamanism and Northern ecology*. Pentikäinen J. (ed.). Berlin; New York: Mouton de Gruyter, 1996, 207–217.

- Skribnik, Laakso 2022 — Skribnik E., Laakso J. Ugric: General information. *The Oxford guide to the Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 523–536.
- Toldova, Serdobolskaya 2017 — Toldova S. Yu., Serdobolskaya N. V. Oformlenie pryamogo dopolneniya v finno-ugorskikh yazykakh: mezhdru predikatsiei i diskursom [Direct Object marking in Finno-Ugric languages: Between sentence and discourse]. *Ural-Altaic Studies*, 2017, 4(27): 91–112.
- Toldova, Volkova 2021 — Toldova S., Volkova A. Reflexivity in Kazym Khanty. *Reflexive constructions in the world's languages*. Janic K., Puddu N., Haspelmath M. (eds.). Berlin: Language Science Press, 2021, 260–292.
- Urmanchieva, Plungian 2021 — Urmanchieva A. Yu., Plungian V. A. Passiv v zapadnom dialekte mansiiskogo [Passive in Western Mansi]. *Tomsk Journal of Linguistics and Anthropology*, 2021, 1: 94–131.
- Virtanen 2015 — Virtanen S. *Transitivity in Eastern Mansi: An information structural approach*. PhD thesis. Helsinki: Univ. of Helsinki, 2015.
- Virtanen 2023 — Virtanen S. Pragmatic reasons for non-clause-initial placement of subjects of active and passive clauses in Northern Mansi. *Linguistica Uralica*, 2023, 59(4): 286–306.
- von Heusinger 2002 — von Heusinger K. Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 2002, 19(3): 245–274.
- Zhornik 2020 — Zhornik D. O. Mansiiskaya skazka o Zheleznom volke [A Mansi folktale about Iron Wolf]. *Rodnoy Yazyk*, 2020, 2: 151–201.

Получено / received 28.06.2024

Принято / accepted 24.09.2024

The information structure of Meadow Mari

© 2024

Johannes Hirvonen

Ludwig-Maximilians-Universität in Munich, Munich, Germany;
johannes.hirvonen@campus.lmu.de

Abstract: This paper provides a first descriptive overview of morphosyntactic strategies of marking information structure in contemporary Meadow Mari, based on both elicited and corpus data. I describe syntactic strategies of marking focus, topics, and contrastive topics in the language. Focus is variably realized in situ, in immediately preverbal or postverbal position, or via pseudocleft. The strategies differ in terms of their “markedness”, in the sense that more marked strategies encode subsets of the types of foci which the less marked strategies encode. Nevertheless, no type of focus requires a certain strategy. Topics are placed in the left periphery of the clause. In addition to these syntactic strategies of information structure marking, the suffix -ŽE is described as a morpheme denoting the selection of an individual from a superset and contrasting it with other individuals — essentially as a marker of contrastive topics. With its broad scope, the paper lays the groundwork for comparative research on information structure in Meadow Mari and other Uralic languages and for more detailed studies on information structure in Meadow Mari itself.

Keywords: contrast, focus, information structure, Mari, possession, topic, Uralic

Acknowledgements: This paper is an updated and condensed version of the author’s Master’s thesis (2023). The thesis and this paper have benefitted greatly from discussions with and comments by many people, including (in alphabetical order) Jeremy Bradley, Daniel Buring, Ekaterina Georgieva, Andreas Pregla, Ksenia Shagal, Elena Skribnik, Kriszta Szendrői, Elena Vedernikova, and an anonymous reviewer — thank you.

For citation: Hirvonen J. The information structure of Meadow Mari. *Voprosy Jazykoznanija*, 2024, 5: 97–115.

DOI: 10.31857/0373-658X.2024.5.97-115

Информационная структура в луговом марийском языке

Йоханнес Хирвонен

Мюнхенский университет Людвига-Максимилиана, Мюнхен, Германия;
johannes.hirvonen@campus.lmu.de

Аннотация: В статье представлен обзор морфосинтаксических стратегий маркирования информационной структуры в современном луговом марийском языке, основанный на элицированных и корпусных данных. Мы описываем синтаксические стратегии маркирования фокуса, топиков и контрастивных топиков в языке. Фокус может быть выражен *in situ*, в первой предглагольной или заглагольной позиции или с помощью псевдоклефта. Данные стратегии различаются «маркированностью», в том смысле, что более маркированные стратегии выражают такие подмножества типов фокуса, которые выражают и менее маркированные стратегии. Тем не менее ни один из типов фокуса не предполагает особого средства выражения. Топики находятся на левой периферии клаузы. Помимо указанных стратегий маркирования информационной структуры, суффикс -ŽE описывается как морфема, маркирующая выбор индивида из объемлющего множества и отношение контраста с другими индивидами — по сути маркером контрастивного топика. Учитывая широкую перспективу, статья закладывает основу для сравнительного изучения информационной

структуры в луговом марийском и других уральских языках, а также для более детального исследования информационной структуры в самом луговом марийском.

Ключевые слова: коммуникативная структура, контраст, марийские языки, поссесивность, топик, уральские языки, фокус

Для цитирования: Hirvonen J. The information structure of Meadow Mari. *Voprosy Jazykoznanija*, 2024, 5: 97–115.

DOI: 10.31857/0373-658X.2024.5.97-115

1. Introduction

Although it has previously been recognized that information structure is an important factor governing variation in word order in Meadow Mari, e.g., in [Chkhaidze 1941: 29–31; Pengitov 1961: 122; Vilkuna 1998; Saarinen 2022: 463], information structure marking in the language has not been systematically described so far. This paper provides a first descriptive overview of information structure marking in contemporary Meadow Mari, based on both elicited and corpus data. At the center of this investigation are types of morphosyntactic marking which interact with the information structural notions of focus, topic, and contrastive topic.

For focus marking, Meadow Mari employs four different strategies: (1) prosodic in-situ marking, (2) immediately preverbal constituent focus and clause-initial verbal focus, (3) postverbal focus, and (4) focussing via pseudocleft. The focus marking strategies differ in terms of their “markedness”, in the sense that more marked strategies encode subsets of the types of foci which the less marked strategies encode: constituent focus is mostly marked in situ; the immediately preverbal and postverbal positions have restrictions as to when non-object constituents may appear there (e.g. in correction contexts); and finally, pseudoclefts as the structurally most complex type of focus marking are employed only for identificational focus. Furthermore, focussed verbs can appear in a left-peripheral position. These facts show that Meadow Mari does not have a designated focus position, and that no type of focus *requires* a certain strategy. Topics are located in the left periphery. This is true for both aboutness topics and contrastive topics. The most important morphological means of information structure marking is the morpheme *-žE*.¹ This suffix has several discursive functions, which include the selection of an individual from a superset and triggering a contrastive interpretation. In this function *-žE* is used for marking contrastive topics and the shift to new aboutness topics.

The paper is structured as follows: After a short description of the data sources in Section 2, I discuss focus, (contrastive) topics, and the functions of the morpheme *-žE* in Sections 3, 4, and 5, respectively. In Section 3, I describe focus marking via the four available strategies (in-situ focus, immediately preverbal, immediately postverbal placement, pseudocleft) and investigate their functions. Section 4 contains a short overview of the placement of topics in the left periphery of the clause. In Section 5, I examine the discursive functions of *-žE* in more detail. Section 6 summarizes the most important findings.

2. Data

The variety investigated in this study is contemporary (written and spoken) Meadow Mari. Meadow Mari is a Uralic language with 318,495 speakers, according to the 2020 All-Russian Population Census², spoken mainly in the titular republic Mari El in the Russian Federation.

¹ *žE* in capital letters is used as a shorthand for various allomorphs of the suffix, e.g. *-že*, *-žö*, *-še*, etc. (see Section 5).

² Available online at: <https://rosstat.gov.ru/vpn/2020> (last accessed 04.07.2024).

The data for this study comes from different sources: First, I draw from already existing work on Meadow Mari, including data gathered for my Master's thesis on the same topic [Hirvonen 2023]. Another source of data were corpora of Meadow Mari. These were used mainly for the investigation of topic marking and the functions of the morpheme *-ŽE*. The largest corpus of Meadow Mari which allows to display context is the Corpus of Contemporary Written Literary Meadow Mari [Arkhangelskiy 2019]. This corpus has a size of 5.53 million tokens and mostly consists of contemporary newspaper texts. Context can be shown up to four sentences before and after the target sentence. A similar corpus is the Corpus of Social Media in Meadow Mari which consists of 3.59 million words of written text from the social network “VKontakte” [Ibid.].³ In addition to these corpora of written Meadow Mari, there are two more important sources of contemporary spoken Meadow Mari: a small printed corpus of data from the village Saryj Torjal collected in the year 2000 [Serdobolskaya, Toldova 2012], and the larger Spoken Corpus of Meadow Mari, consisting of data from 2000–2004 and 2018 gathered in the same village [Volkova et al. 2024].⁴ The section on topics in this paper relies mostly on the smaller printed corpus. Thirdly, for data on all types of information structure marking investigated in this paper, I worked with a native speaker consultant. My consultant is a female speaker of Meadow Mari in the age group 40–50 and grew up in the central Meadow Mari (Morki / Sernur) dialect area. Most data in this study is elicited.

3. Focus

This section deals with strategies of focus marking in Meadow Mari. Focus is understood here as indicating the presence of alternatives relevant for the interpretation of the marked linguistic expression, see, e.g., [Krifka 2008]. In the descriptively oriented approach taken here, I investigate the marking strategies in different contexts known to trigger focus marking. Relevant contexts include the expression answering a *wh*-question or alternative question (1), and correction contexts (2). The expression marked with focus is formatted with SMALL CAPITAL LETTERS:

- (1) — *What do you put in your pasta sauce? / Do you put tarragon or thyme in your pasta sauce?*
— I put *THYME* in my pasta sauce. [Büiring 2009: 178–179]
- (2) *I put THYME in my pasta sauce, not tarragon.* [Büiring 2009: 179]

The focus marked expressions can be of different sizes, e.g. whole clauses in the case of clausal focus, constituents, or single constituents, as in the previous examples. Another special type investigated here, *verum focus*, targets the truth value of a sentence. Presentational focus, as in (3), has been explained as indicating a covert question around which discourse is structured, such as *What was there?*, see, e.g., [Krifka 2008; Büiring 2009; van der Wal 2016; 2021].

- (3) *Once upon a time, there was a PRINCESS.* [Krifka 2008: 251]

Finally, the last distinction I make here is the distinction between information focus and identificational focus (sometimes also called exhaustive or contrastive focus) [É. Kiss 1998]. The latter identifies the subset of elements for which the predicate can hold, and thus excludes the other alternatives.

³ These two corpora are available online at: <https://meadow-mari.web-corpora.net/> (last accessed 04.07.2024).

⁴ Available online at: http://lingconlab.ru/spoken_meadow_mari/ (last accessed 04.07.2024). I want to thank one of the creators of the corpus, Aigul Zakirova, here for sending me the whole corpus data. The data in this corpus includes part of the data from Serdobolskaya and Toldova [2012], but I cited the source where I initially found the respective example.

It has long been recognized that the position of foci, and to some extent topics and backgrounded constituents, is associated with the word order type, viz. the headedness of a language. In head-final languages, constituent foci are typically found in the immediately preverbal position, see, e.g., [Dezső 1978; Kim 1988; Herring 1990], among others, and the postverbal position is often occupied by some type of backgrounded material, although there is much variation [Pregla 2024]. Meadow Mari exhibits a number of characteristics typical of a rigid head-final language: nouns follow their modifiers, compounds are head-final, the language employs postpositions, modal verbs follow lexical verbs, and clausal complements generally precede the main clause. The neutral word order is SOV with the dative object typically preceding the accusative object. Time and place adverbials are usually placed left in the clause; they may occur before or after subjects or topical constituents. Irrespective of their position relative to other parts of the sentence, time adverbials precede place adverbials [Chkhaidze 1941: 29–31; Pregla 2024: 23]. In the Uralic context, Meadow Mari word order has been characterized as “quite flexible” SOV as regards possible word order variation [Vilkuna 2022: 950], but as rigid in terms of argument reordering for focus purposes [Pregla 2024: 92]. A clause showing the word order in a neutral context is given in (4).⁵

- (4) *Tače kolhoz pasu-što vič rveze traktor dene mlandə-m udarno*
 today kolkhoz field-INE five lad tractor with land-ACC enthusiastically
kural-ət.
 plough-3PL
 ‘Today on the kolkhoz field, five lads enthusiastically plough the land with a tractor.’
 [Chkhaidze 1941: 30]

Previous research indicates that focussed constituents in Meadow Mari, including both subjects and objects, can either occupy the preverbal position, see, e.g., [Chkhaidze 1941: 29–31; Pengitov 1961: 122–123; Pregla 2024: 92–95], or the same position as in the unmarked clause, i.e. stay in situ [Nelson et al. (to appear); Saarinen 2022: 463; Vilkuna 1998: 195]. Additionally, two studies have found that “contrastively” focussed constituents may occupy a clause-final position as well [Vilkuna 1998: 195; Pregla 2024: 176–177]. In this section, I will discuss focus, paying attention to the interpretational effect of different morphosyntactic marking strategies. There are four different strategies employed to express focus: (1) in-situ focus, (2) immediately preverbal focus and clause-initial verbal focus, (3) postverbal focus, (4) usage of a special pseudocleft construction. Taking these four constructions as a starting point, this section discusses the focus marking strategies and makes an attempt to delimit the discourse contexts in which they occur.

3.1. In-situ focus marking

In Meadow Mari, focus can be marked prosodically. In the case of constituent focus, the marking involves a rise on the stressed vowel of the focussed expression, see the discussion in [Hirvonen 2023: 51–54]. In-situ focus via prosodic marking is a versatile strategy: it can be employed for both clausal focus and constituent focus, in answers to *wh*-questions, and in correction contexts. Furthermore, this marking strategy is underspecified for the type of focus and can mark both information and identificational focus. My consultant preferred in-situ focus in most contexts, but accepted preverbal answer and correction foci if presented with them.

Full clausal focus after all-new questions involves only the canonical word order without any special prosodic marking:

⁵ Abbreviations used in the glosses are given at the end of the paper.

- (5) *Mo tāšte?* — *ÜDƏRAMAŞ JOŠKAR ARAKA-M JÜ-EŠ.*
 what here woman red wine-ACC drink-3SG
 ‘What’s (going on) here? — THE WOMAN IS DRINKING RED WINE.’ (elicited)
- (6) *Mo tāšte kaj-a?* — *RVEZE MEČE DENE MOD-EŠ.*
 what here go-3SG boy ball with play-3SG
 ‘What’s going on here? — THE BOY IS PLAYING WITH A BALL.’ (elicited)

For constituent focus as answers to *wh*-questions, my consultant preferred prosodic in-situ marking (and answering the question only with the focussed constituent, or together with the verb).

- (7) *Kō tače una-lan kočkəš-əm kond-en?* — *MASKA tače (una-lan kočkəš-əm) kond-en.*
 who today guest-DAT food-ACC bring-PST2.3SG bear today guest-DAT food-ACC bring-PST2.3SG
 ‘Who brought the guests food today? — The BEAR brought the guests food today.’ (elicited)
- (8) “*Mo-m maska kond-en?*” *ur jod-o.* “*Maska KOČKƏŠ-ƏM kond-en,*” *krab vašešt-əš.*
 what-ACC bear bring-PST2.3SG squirrel ask-PST1.3SG bear food-ACC bring-PST2.3SG crab answer-PST1.3SG
 “‘What did the bear bring?’” asked the squirrel. “The bear brought the FOOD,” answered the crab.’ (elicited)

Focus in correction contexts is realized just like other types of information focus. Both subjects and objects are preferably placed in the same position as in the unmarked clause:

- (9) *Məj təj deč=at šukə-rak kogəl’ə-m əšt-en kert-am.* — *Uke,*
 1SG 2SG from=ADD much-CMPR pie-ACC do-CVB can-1SG no
MƏJ šukə-rak kogəl’ə-m əšt-en kert-am.
 1SG much-CMPR pie-ACC do-CVB can-1SG
 ‘I can bake more pies than you. — No, I can bake more pies.’ (elicited)
- (10) [‘In the garden, is the woman hitting a girl?’]
Uke, sad-əšte üdəramaš RVEZƏ-M šel-eš.
 no garden-INE girl boy-ACC hit-3SG
 ‘No, in the garden, the girl is hitting a BOY.’ (elicited)

Note, furthermore, that prosodic marking can express identificational focus as well, where focus identifies the exhaustive subset of which the predicate phrase holds [É. Kiss 1998] (for another strategy see Section 3.3):

- (11) [In a situation where only Mikhail is holding a rifle:]
MIXAIL pəčal-əm kuč-a.
 Mikhail rifle-ACC hold-3SG
 ‘MIKHAIL is holding the rifle (and not anyone else).’ (elicited)

3.2. Ex-situ focus marking

3.2.1. Immediately preverbal and clause-initial verbal focus

The second strategy of constituent focus marking is placing a focussed constituent in an immediately preverbal position, illustrated with elicited data from Pregla [2024]. The preverbal position is a marked position for non-object constituents as immediately preverbal subjects are

only acceptable when focussed, and not in neutral contexts (12). As mentioned above, my consultant consistently used in-situ focus for subjects but accepted preverbal subject foci if presented with them.⁶

- (12) *Ikeče jal-əšte joča-lan jomak-vlak-əm KOVA / *kova ojlə-š.*
 recently village-INE child-DAT tale-PL-ACC grandmother grandmother tell-PST1.3SG
 ‘GRANDMA / *Grandma recently told stories to the children in the village.’ [Pregla 2024: 94]

Importantly, Pregla [2024] has shown that what matters for the focal interpretation of a constituent is not the absolute position in a clause, but the relative position immediately preceding the verb. The focus-verb string can either appear in the clause-final position or clause medially, with the preceding constituent in focus:

- (13) *Poškudo TEDGEČE ERDENE puə-š joča-lan pört-əštö pərəs-əm.*
 neighbor yesterday morning give-PST1.3SG child-DAT house-INE cat-ACC
 ‘The neighbor gave a cat to the child in the house YESTERDAY MORNING.’ [Pregla 2024: 176]

In a related construction, clause-initial placement of a finite verb also allows focussing the verb itself [Hirvonen 2023: 50–51]. This placement marks types of focus associated with finite verbs, e.g. verum focus targeting the truth value of the sentence (14)–(15).⁷ This construction is commonly associated with correction contexts. Note, however, that for verbal focus a prosodic realization is possible as well, cf. [Hirvonen 2023: 50].

- (14) *Tenij māj Budapest-əške š-əm kaj. — KAJ-əš-əč taj tenij*
 this.year 1SG Budapest-ILL NEG.PST1-1SG go.CNG go-PST1-2SG 2SG this.year
Budapest-əške!
 Budapest-ILL

‘I did not go to Budapest this year. — You DID GO to Budapest this year!’ (elicited)

- (15) *Vrač joča-lan em-əm jüktə-š. — Əš JÜKTÖ*
 doctor child-DAT medicine-ACC give.to.drink-PST1.3SG NEG.PST1.3SG give.to.drink.CNG
vrač joča-lan em-əm!
 doctor child-DAT medicine-ACC

‘The doctor gave the child medicine to drink. — The doctor DID NOT GIVE the child medicine to drink!’ (elicited)

This phenomenon can best be illustrated with data from verb clusters involving displacement of one part of the cluster only, cf. [Georgieva et al. 2021: 463]. The following examples involve a verb cluster consisting of the modal verb *kert-* ‘can’, the negation auxiliary, and the lexical verb *əšte-* ‘do, make’. The order of these three verbs in a neutral context is lexical verb > negation > modal (16).⁸ As in example (15), the negation auxiliary is inflected and selects the invariant connegative form of the modal verb, which in turn selects the converbial form of the lexical

⁶ A reviewer raises the question of whether there are any restrictions on the usage of preverbal focus. My consultant did not categorically refuse a preverbal placement of constituents under focus in *wh*-question contexts, but consistently preferred and chose in-situ placement. I did not discover any strict constraints on the usage of preverbal constituent focus. The reasoning made here, that the preverbal position is a more marked position, thus comes from the preference for in-situ focus in *wh*-contexts and the general observation that the preverbal position is not available for non-focussed subjects, as shown in (12).

⁷ In example (15), the negation auxiliary occupies the clause-initial position together with the lexical verb in the connegative form. See the discussion and references in the remainder of this subsection for more detailed information on this phenomenon.

⁸ The order negation auxiliary > finite verb is unexpected in a head-final language, but found across Uralic languages [Yurayong et al. 2024: 736]. The relevant part for the argument made here is that the negation auxiliary and the verb in the connegative form cannot be separated from each other, but can be separated from a third verbal form. See Georgieva et al. [2021] for an in-depth analysis of this phenomenon.

verb. The negation auxiliary and the verb which it selects form a cluster in which nothing can intervene,⁹ but for focussing purposes this cluster can occupy the clause-initial position while leaving the lexical verb behind in a clause-final position (17):

- (16) *Məj tidə-m niguze əšt-en o-m kert.*
 1SG this-ACC no.way do-CVB NEG-1SG can.CNG

‘I cannot do this in any way.’ (elicited)

- (17) *O-M KERT məj tidə-m niguze əšt-en.*
 NEG-1SG can.CNG 1SG this-ACC no.way do-CVB

‘I CANNOT do this in any way.’ (elicited)

In sum, the position immediately preceding the verb is associated with constituent focus marking, irrespective of where in the clause the verb is placed. Clause-initial placement of the verb is associated with verbal or verum focus. What unites these constructions is that the immediately preverbal position can be employed for constituent focus, and that the verb receives a focussed interpretation if the verb itself is not preceded by any other constituent. In these constructions the number of postverbal elements is in principle unlimited and their ordering is the same as in the neutral context. Furthermore, the pre-focal elements do not receive any special information structural interpretation, cf. [Pregla 2024].

3.2.2. Postverbal focus

Immediately postverbal constituent focus is a more specialized strategy and is available mainly in two contexts — presentational focus and correction focus. My consultant dispreferred postverbal constituent focus in answers to *wh*-questions. A subtype of information focus, presentational focus indicates a covert question in the discourse. This strategy is typically employed in Meadow Mari to introduce a referent which is then continued as an aboutness topic (see Section 4 on topics) in the following clause, cf. *klad* ‘shed, barn’ in example (18), and a person called *Zina* in example (19). The new topics are indicated by boldfacing:

- (18) *Memnan pört vokane əl'e təgaj KLAT. Vot klat... pusta klat,*
 2SG.GEN house besides be.PST1.3SG such shed see shed empty shed
nigō tušto il-en og-əl. Tide klat ümbake küz-en
 nobody there live-CNG NEG.3SG-BE.CNG this shed onto climb-CVB
šinč-ən-am no.
 sit.down-PST1-1SG PTCL

‘Next to our house there was this SHED. Well, a shed... an empty shed, nobody lived there. **This shed**, I climbed up and sat down on it.’ [Serdobolskaya, Toldova 2012: 712]

- (19) [“I had a hard time learning to knit.”]
Məj pər'l'a tunem-ən-na ZINA dene, Morkinskij rajon-et, tudo peš
 1SG together learn-PST2-1PL Z. with M. district-PX2SG 3SG very
čot pid-eš əl'e. Tudə-n den tunem-əm məj tunem-aš... pid-ən-am,
 much knit-3SG AUX 3SG-GEN with learn-PST1.1SG 1SG learn-INF knit-PST2-1SG
tudo məj-əm tunəkt-en.
 3SG 1SG-ACC teach-PST2.3SG

‘I learned together with **ZINA** from Morkinskiy district, **she** used to knit a lot. **With her** I learned, to learn... I knit, she taught me.’ [Volkova et al. 2024: Biografiya Z. I. E. (Part 2)]

⁹ Jeremy Bradley (p.c.) remarks that this is not entirely as clitics like the additive clitic *=at* can indeed intervene, e.g. *o-m=at kert* [NEG-1SG=ADD can.CNG]. Georgieva et al. [2021] discuss the same phenomenon for Udmurt. In any case, the argument made here is not affected by this additional complication.

Additionally, constituents in an immediately postverbal position can receive a focussed interpretation in correction contexts [Pregla 2024]:¹⁰

- (20) (*Uke*,) *erdene joča-lan pört-əštö pəras-əm pu-əš AŠNƏŠE PÖRJED!*
 no in.morning child-DAT house-INE cat-ACC give-PST1.3SG caring person
 ‘(No,) A CARING PERSON gave a cat to our child in the house in the morning [and not the neighbor]!’ [Pregla 2024: 186]
- (21) *Ašnəşe pörjeŋ joča-lan pört-əštö pəras-əm pu-əš ERDENE!*
 caring person child-DAT house-INE cat-ACC give-PST1.3SG in.morning
 ‘A caring person gave a cat to our child in the house IN THE MORNING [and not in the evening]!’ [Pregla 2024: 187]

Postverbal focussing appears to be more marked than in-situ focus, as it is available mainly for presentational and correction focus. Furthermore, the strategy is positionally more restricted as only the focussed constituent is allowed to appear after the verb. The more constituents appear after the verb, the more the sentences are degraded. In the following clause, *pərasəm* ‘the cat’ occurs in the immediately postverbal position and receives a focussed interpretation. If the verb is further left (and the number of constituents between the verb and the focussed constituent increases), the sentence is degraded (22). This is a clear difference to preverbal focus, where the number of postverbal elements is in principle unlimited (see Pregla [2024] for further discussion).

- (22) *Poškudo (ʔʔpu-əš) erdene (ʔʔpu-əš) joča-lan (ʔpu-əš) pört-əštö*
 neighbor give-PST1.3SG yesterday give-PST1.3SG child-DAT give-PST1.3SG house-INE
 (*pu-əš*) *PƏRƏS-ƏM.*
 give-PST1.3SG cat-ACC
 ‘The neighbor gave A CAT to the child in the morning.’ [Pregla 2024: 187–188]

3.3. Pseudocleft

In addition to the strategies of focus marking underspecified for the type of focus, there is a more marginal strategy to mark exhaustive focus, namely a pseudocleft. This construction involves a clause with a *wh*-phrase in subject position, followed by the determiner *tide* ‘this’ and the focussed constituent. Pseudoclefts are only felicitous if a sentence provides the exhaustive extension of the predicate. Consider the following examples. When the context indicates that there is only one individual for which the predicate holds, usage of a pseudocleft is felicitous (23). However, when the context indicates that a predicate holds of more than one individual in a given situation, usage of a pseudocleft is not felicitous. This is shown in example (24) where the predicate “holding a bottle” is true of all people:¹¹

- (23) [In a situation where only Mikhail is holding a rifle:]
Kö pəçal-əm kuč-a, tide MIXAIL.
 who rifle-ACC hold-3SG this Mikhail
 ‘It is MIKHAIL who is holding the rifle.’ (elicited)

¹⁰ In my previous work, I reported judgments about postverbal focus being dispreferred in correction contexts [Hirvonen 2023: 49]. The factors influencing the different judgments are not clear to me, but in light of the available evidence for postverbal focus in correction contexts in Vilks [1998] and Pregla [2024], I reject my earlier claim for it being dispreferred in this context.

¹¹ Note that these examples were not produced spontaneously, but elicited. My consultant did not consider the construction very natural, but their judgments about its felicity were very clear. For the parallel Russian construction see, for example, [Burukina et al. 2024: 31–32].

- (24) [In a situation where three people are holding one bottle each:]

#*Kö butəlka-m kuč-a, tide JəVAN.*

who bottle-ACC hold-3SG this Jəvan

‘It is JəVAN who is holding the bottle.’ (elicited)

3.4. Summary of focus marking strategies

In sum, there are four different strategies of constituent focus marking: prosodic in-situ marking, immediately preverbal constituent focus and clause-initial verbal focus, immediately postverbal focus, and focussing via pseudocleft. Prosodic in-situ marking is a versatile strategy for focus marking on any constituent, and, furthermore, underspecified for the type of focus as it can encode both information and identificational focus. Preverbal constituent focus is more marked for subject focus, for example, but does not categorically mark a different type of focus. Postverbal focus is commonly employed for presentational focus and in correction contexts. Finally, the pseudocleft is felicitous only for identificational focus. There are no strategies specific to clausal focus.

Importantly, no focus marking strategy discussed in this section *is required* by any type of focus. For example, correctional focus can be marked with the constituent in a postverbal position, but both the preverbal position and in-situ marking are available marking strategies as well. The focus marking strategies are characterized by increasing “markedness”, in the sense that more marked strategies encode subsets of the types of foci which the less marked strategies encode. Constituent focus is mostly marked in situ; the immediately preverbal and postverbal positions have pragmatic restrictions as to when non-object constituents may appear there; and finally, pseudoclefts as the most complex type of focus marking are employed only for identificational focus. The distribution conforms to the markedness scale of structural complexity of focus marking by Skopeteas and Fanselow [2010], which predicts that more marked types of focus require more marked strategies. In the broader typological picture, the data shows that Meadow Mari behaves like a rather typical SOV language, since it allows preverbal focus. But at the same time, it is not a strict-position focus marking language, as an alternative in-situ marking strategy exists, cf. [Büring 2009].

4. Topics

Information is not added to the discourse in an unstructured way. Rather, speakers attempt to organize this information and give hints about how it is to be classified. Reinhart [1981] argues that the discourse context is not an unordered set of propositions, rather it is structured into subsets of propositions, which are stored under defining entries, for example, propositions about individuals — this is the notion of topic. In example (25a), the new information provided by the sentence is stored as information about Aristotle Onassis, and in example (25b) it is stored as information about Jacqueline Kennedy, the respective topic. Topical constituents are marked by boldfacing:

- (25) [Krifka 2008: 265]

a. ***Aristotle Onassis** married Jacqueline Kennedy.*

b. ***Jacqueline Kennedy** married Aristotle Onassis.*

The most influential approach to topics, Reinhart [1981], defines topics as the part that a given sentence is “about”, e.g. an individual, as illustrated above. Such topics are usually called aboutness topics. In relation to the discourse context, topics can serve different purposes. For example, some topics continue reference to previously introduced, given referents, while others are newly

introduced into the discourse or mark the shift to another topic. Givón [1983: 9] characterizes topics like the latter as “newly-introduced, newly-changed or newly-returned” topics, and like the former ones as “continuous topics”. Similar classifications have been made by many scholars, see, e.g., [Erteshik-Shir 2007; Frascarelli 2018; Frascarelli, Hinterhölzl 2007; van der Wal 2021], and while these proposals differ in their details, all these scholars usually make a distinction between some form of continued (or familiar, backgrounded) topics and shifted (or switch) topics. After the discussion about topics in this section, I will argue in Section 5 that the Meadow Mari suffix *-ŽE* can mark a topic shift, but not a topic continuation.

Another notion which is commonly distinguished is that of contrastive topic. I follow the definition by Büring [2003; 2016], which states that contrastive topics trigger a conventional implicature that there is an identifiable alternative question including the referent of the contrastive topic expression.¹² Typical contexts that evoke sequences with contrastive topic and focus are multiple *wh*-questions, single *wh*-questions containing plurals, or implicational topics (26). This third type, like presentational foci, indicates additional questions such as *Where was the cook?* Contrastive topics are boldfaced in the examples.

(26) [Büring 2016: 68, 70]

- a. (*Which guest brought what?* —) **Fred** brought the BEANS.
- b. (*Where do your siblings live?* —) My **sister** lives in STOCKHOLM.
- c. (*Where was the gardener at the time of the murder?* —) The **gardener** was in the HOUSE.

In Meadow Mari, aboutness topics and contrastive topics are located in the left periphery of the clause, while expressions referring to given individuals can appear postverbally. There does not seem to be a designated structural position for topics, however, as they can be preceded or followed by various adverbs, as in example (29) below. In addition to topics which are part of the clause, Meadow Mari has left-dislocated aboutness topics as well. In the former case, the topics necessarily have a corresponding gap in the clause (e.g. the position for objects in the unmarked order), other than left dislocated topics. For aboutness topics, two main diagnostics can be employed: Most importantly, non-subjects as topics in a topic progression are preposed and located in the left periphery, e.g. objects in an OSV order (if preverbal subject focus is not the reason for reordering). In this case, they can still be preceded by adverbs. The second diagnostic, the classic aboutness-test, see, e.g., [Erteshik-Shir 2007: 19–20], reveals that aboutness topics are preferred in a position left of the verb and dispreferred in a postverbal position.¹³

In a clause uttered in a neutral context, subjects occupy a left position in the clause, and the subject functions as sentence topic as well. But if a non-subject becomes the topic of a clause in the sense that the clause is “about” the non-subject, the constituent can be placed in the sentence-initial position to express the progression, that is, the shift to this new topic. These topics may receive a contrastive interpretation, but not obligatorily. In example (27), the first sentence tells about what happened after a charity concert and introduces ‘200 000 rubles’. The following sentence describes what happened to that money, and the accusative-marked

¹² Note that Büring [2003; 2016] considers contrastive topics as independent of topics, in contrast to, e.g., Krifka [2008], who considers them a combination of topic and focus. Regardless of that fact, I discuss contrastive topics in the section about topics.

¹³ Based on data from Bradley and Hirvonen [2022], I previously suggested another diagnostic for the topichood of left-peripheral constituents—that placing a constituent in the clause-initial position increases its acceptability as antecedent for a null subject in the following clause [Hirvonen 2023: 64–66], based on the hypothesis that in consistent null-subject languages like Meadow Mari, null subjects are resolved with respect to the aboutness topic, see, e.g., [Holmberg 2010; Frascarelli 2018]. While there is indeed a small effect in the data, and topichood thus could play a role for null subject resolution, null subjects and preposed aboutness topics regularly occur together in the same clause. Therefore, null subject resolution and topichood are likely two different phenomena. I will not further discuss the availability of this diagnostic in this article.

phrase ‘this money’ occupies the sentence-initial position before the subject of the clause, in an OSXV order. Similarly, in example (28), belletristic books are continued as topic ‘such books’ in an OSXV order:

- (27) *Tunam, koncert ertarə-me dene kok šüdü tüžem tenge*
 then concert carry.out-PTCP.PASS with two hundred thousand ruble
pogən-en. Tide šijvundə-m murəzo onkologij čer dene
 be.gathered-PST2.3SG this money-ACC singer oncology illness with
čerlanə-še joča-vlak-əm eml-aš kusar-en.
 be.sick-PTCP.ACT child-PL-ACC cure-INF transfer-PST2.3SG

‘In the course of the concert 200 000 rubles were gathered. **This money**, the singer donated to the treatment of children with cancer.’ [Arkhangelskiy 2019, Journal *Kidšer* (26.03.2016)]

- (28) *Səlnəmut-an kniga-m avtor-vlak ške küšeš aļe sponsor polšəmo dene*
 literature-ADJ book-ACC author-PL own at.expense or sponsor help with
vele lukt-ət. A tagaj kniga-m pisatel'-vlak en ondak užal-aš
 only publish-3PL but such.kind.of book-ACC writer-PL SUP before sell-INF
tərš-at <...>.
 try-3PL

‘Belletristic books, the authors only publish on their own expense or with the support of sponsors. But **such books**, the writers first of all try to sell <...>.’ [Arkhangelskiy 2019, Newspaper *Marij El* (27.05.2005)]

The classic aboutness-test can also be applied. Example (29) shows that subjects as aboutness topics are preferably placed left in the clause, and the acceptability of a sentence degrades if they appear further right. Such subjects were not accepted in a postverbal position by my consultant. Only expressions referring to given individuals can be located there. This example also shows that the aboutness topic can be preceded by adverbs, cf. [Hirvonen 2023: 67–68].

- (29) [“Have you heard anything about Serge?”]
*(Serge), pialeš, (Serge) u paša-m (ʔSerge) mu-ən, (*Serge).*
 Serge fortunately Serge new work-ACC Serge find-PST2.3SG Serge
 ‘Fortunately, **Serge** has found a new job.’ (elicited)

Mere givenness does not explain the left-peripheral placement of a given constituent alone, as given constituents can also be located elsewhere in the clause. This can be observed in the first clause of example (28) above, where the topic precedes the given subject (*pisatel'-vlak* ‘the writers’), and in the following newspaper example (30) about the capital of Mari El, Yoshkar-Ola. Here, parts of the city are introduced in the first sentence and referred to by a phrase headed by the proximal demonstrative *tə* ‘this’ in the next sentence. The phrase is placed after the subject in the clause:

- (30) *Pətartəš žap-əšte Joškar-Ola-šte u mikrorajon-vlak pišə-n kušk-ət:*
 last time-INE Yoshkar-Ola-INE new microdistrict-PL quick-INS grow-3PL
Festival'naj, Mirnəj, Molodežnəj, Sportivnəj. Jeŋ-vlak tə kundem-əšte
 F. M. M. S. person-PL this area-INE
pačer-əm nal-ət, posnak samərak-vlak.
 apartment-ACC buy-3PL especially young-PL

‘In recent times, new microdistricts in Yoshkar-Ola are growing fast: Festival'naj, Mirnəj, Molodežnəj, Sportivnəj. People are buying apartments in these areas, especially the young ones.’ [Arkhangelskiy 2019, Newspaper *Marij El* (20.03.2017)]

Contrastive topics are located in the left periphery of the clause as well. This is true for subjects (31) and objects as contrastive topics (32):

- (31) [“Where do your siblings live?”]

Iza-m *MOSKVA-ŠTE*, *a* *šūžar-em* *JOŠKAR-OLA-ŠTE*.
 older.brother-PX1SG MOSCOW-INE but little.sister-PX1SG Joškar-Ola-INE

‘My older brother (lives) in MOSCOW, and my little sister (lives) in YOSHKAR-OLA.’ (elicited)

- (32) *A* *Irina* *kok* *tarelka-m* *vele* *mušk-ən* *šuktə-š*. *Kumšo* *tarelka-žə-m*
 but Irina two plate-ACC only wash-CVB manage-PST1.3SG third plate-ŽE-ACC
tudo *mušk-ən* *əš* *šukto*.
 3SG wash-CVB NEG.PST.3SG manage.CNG

‘But Irina only managed to wash two plates. The third plate she did not manage to wash.’ (elicited)

One indication for a topical interpretation of left-peripheral non-subjects is the restricted availability of non-subject negative indefinite pronouns in this position. They cannot receive a topical interpretation because they do not refer to any individual. This can be observed in example (33) — an object negative indefinite pronoun cannot appear in a topical position even if the structure favors the clause-initial placement of the object in contrast to the topic in the preceding clause:¹⁴

- (33) *Araka* *manmašte*, *tudə-m* *iktaž-kö* *jü-ən* *mo?* — *Araka-m*
 vodka about 3SG-ACC INDEF-who drink-PST2.3SG Q vodka-ACC
iza-m *jü-ən*, *a* *{*ñimo-m}* *šoləm*
 older.brother-PX1SG drink-PST2.3SG and nothing-ACC younger.brother.PX1SG
{ñimo-m} *jü-ən* *og-əl*.
 nothing-ACC drink-CVB NEG.3SG-be.CNG

‘What about the vodka, did anybody drink it? — The vodka my older brother drank, but my younger brother drank nothing.’ (elicited)

Apart from topics in the left periphery of the clause, Meadow Mari allows for left-dislocated topics outside of the core clause as well. Example (34) shows such a topic (which is additionally marked with -ŽE, see Section 5). It is not part of the clause as the topic is in the nominative case like a subject, but the finite verb does not agree with it. Instead, the verb shows 3PL morphology agreeing with a 3PL null subject, which is typical for generic statements, cf. [Bradley, Hirvonen 2022]. These topics can optionally be introduced by the phrase *manmašte* ‘as regards, concerning’ (35). Optionally, a resumptive pronoun within the core clause (e.g. *tudo* ‘it, this’) can refer back to this topic.

- (34) *Marij* *televidenij-žə* *tidə-m=ak* *vele* *ončəkt-at* *kuze* *ubiratl-at*.
 Mari television-ŽE this-ACC=EMPH only show-3PL how harvest-3PL

‘Mari television, they only show such stuff, how they harvest crops.’ [Serdobolskaya, Toldova 2012: 731]

- (35) [“Before Miľa returns from work, she’s really busying herself ⟨...⟩.”]

Paša *manmašte*, *Sergej* *ñimənar-at* *jara* *ok* *šinče* ⟨...⟩.
 work about Sergej no.amount.of=ADD idle NEG.3SG sit.CNG

‘Speaking about work, Sergej does not sit around idly either ⟨...⟩.’ [Arkhangelskiy 2019, Newspaper *Marij El* (01.09.2007)]

After the discussion about positions for topics in this section, and foci in Section 3, a short remark about the other positions in the Meadow Mari clause is in order. The aboutness topics and contrastive topics discussed so far are located left in the clause. Section 3 on focussed

¹⁴ Note, however, that non-subject negative indefinite pronouns can be found in the initial position in corpus data, but the exact contexts have yet to be investigated. As contrastive topic in example (33), the negative indefinite was not accepted in any case.

constituents showed that if constituent focus is realized by positioning, then the constituent is placed in the immediately preverbal or postverbal position. Besides the left periphery as a place for topics, and the immediately preverbal and immediately postverbal position for focussed constituents, the other parts of the clause do not appear to receive any special interpretation. Pregla [2024: 92–95] noted this for constituents preceding focus. Regarding the postverbal field, Hirvonen [2023: 73–84] showed that it is a heterogeneous place in the Meadow Mari clause, and that constituents can end up there for various reasons, including: (1) as afterthoughts and completions of a clause, containing expressions referring to given individuals;¹⁵ (2) constituents can epiphenomenally appear to the right of the verb when the verb itself occupies a left-peripheral position for verbal focussing purposes; (3) prosodically heavy constituents can be extraposed to a right-peripheral position; (4) genre-specific reasons. Furthermore, presentational and corrective constituent focus can appear postverbally (see Section 3.2.2). Therefore, constituents can appear in the postverbal field for various unrelated reasons, and do not necessarily receive a certain interpretation. Reason (2) was described already in Section 3.2.1, and for (1), (3), and (4) I refer the reader to the discussion in Hirvonen [2023: 73–84].

Summarizing, the place for topics in the Meadow Mari clause is the left periphery. Aboutness topics and contrastive topics are located in that part of the clause, and this is most clearly visible when non-subjects appear as topics. Furthermore, Mari employs left-dislocated topics for topic shift. The postverbal field can be filled for various unrelated reasons, and the remaining positions in the clause do not appear to serve any specific functions related to information structure marking.

5. Functions of the morpheme -ŽE

Besides syntactic means of marking information structure, Meadow Mari also has at its disposal some morphemes interacting with information structure, including focus particles (*vele*, *gəna* ‘only’), the additive clitic (=at ‘also, even’), and the morpheme -ŽE, which will be described in this section in more detail. In contemporary Meadow Mari, this morpheme is used in two main distinct discursive functions: it can mark the selection of an individual from a superset, including contrastive topics (36) and a topic shift, and it marks uncontroversial information (37), cf. the category of “enimitive”, introduced by Panov [2020]. This section is mainly concerned with the former usage. The relevant parts of the sentence are highlighted by boldfacing in this section.

- (36) *No pūrtūs-še peš saj, a uslovij-že saj ogəl il-aš.*
 well nature-ŽE very good but condition-ŽE good NEG live-INF
 ‘Well, **nature** is very pretty (there), but **conditions** aren’t good to live.’ [Serdobolskaya, Toldova 2012: 728]

- (37) *Tače təlāt nimogaj munə-m kond-en o-m kert.*
 today 2SG.DAT no.kind.of egg-ACC bring-CVB NEG-1SG can.CNG
Rəvəž-še čəvə-na-m kočk-ən.
 fox-ŽE chicken-PX1PL-ACC eat-PST2.3SG
 ‘I cannot bring you any eggs today. The **fox** has eaten our chicken, **you see**.’ (elicited)

-ŽE is used as a shorthand for the suffix’s various allomorphs that appear as the result of processes like voicing assimilation (e.g. -že vs. -še), avoidance of consonant clusters (e.g. -əže vs. -že) or vowel

¹⁵ In Hirvonen [2023] I treated familiar topics separately from afterthoughts and completions. However, it has not been systematically investigated yet whether there are any prosodic or syntactic differences that justify treating familiar topics in Meadow Mari as a separate descriptive category, instead of treating them simply as given elements (thanks to Andreas Pregla for discussion).

harmony (e.g. *-že, -žo, -žō*) [Alhoniemi 1993: 72; Riese et al. 2022: 46–48]. A homophonous suffix also serves as the possessive suffix of the third person singular, but due to differences not only in meaning but also in distributional properties (and for ease of exposition), it is reasonable to treat these functions separately. This is also reflected in the glossing. *-ŽE* as discussed here has the following distributional properties differing from the 3SG possessive suffix usage, cf. also [Georgieva 2022; 2024]:

- 1) Possessive suffixes agree with the possessor in person and number, but *-ŽE* is invariant. This is visible most clearly in cases of ellipsis where the antecedent involves a non-3SG possessor, but the remnant bears the suffix *-ŽE* and thus does not agree with the possessor. The remnant can still receive a non-3SG interpretation:
- (38) [“You have one heavy and one light bag.”]
Nele sumka-t ūstembalne, a kuštalgo / kuštalgo-žo polka-što.
 heavy bag-PX2SG on.table but light light-ŽE shelf-INE
 ‘Your heavy bag is on the table, and **your light one** is on the shelf.’ [Georgieva 2022: 4]
- 2) The suffix can be stacked onto other possessive suffixes, in which case *-ŽE* follows the possessive suffix and does not agree with any constituent, cf. [Georgieva 2022; Tuzharov 1987: 67; Simonenko 2014]:
- (39) [“What food do your pets like to eat?”]
Pij-em(-že) pəzə-m kočk-aš jōrat-a, pəras-em(-že) kol-əm kočk-aš jōrat-a.
 dog-PX1SG-ŽE meat-ACC eat-INF love-3SG cat-PX1SG-ŽE fish-ACC eat-INF love-3SG
 ‘My dog likes to eat meat, (and) my cat likes to eat fish.’ (elicited)
- 3) In the enimitive function it can mark other constituents as well. In example (40), *-ŽE* attaches to a whole PP:
- (40) [*Araka deč posna*]-žē veselitl-aš ok lij mo?
 vodka from separate-ŽE have.fun-INF NEG.3SG be.CNG Q
 ‘Can’t you have a good time **without alcohol**?’ [Arkhangelskiy 2019, VKontake]

The suffix has been described before as a marker of definiteness, e.g., in [Alhoniemi 1993: 75–76], identifiability [Nikolaeva 2003], specificity [Georgieva 2022; Simonenko 2017], contrast [Riese et al. 2022: §9.1.5; Tuzharov 1987: 67–68; Nikolaeva 2003; Simonenko 2014], contrastive topics [Georgieva 2024; Hirvonen 2023; Yurayong 2020: 115], topic shift [Hirvonen 2023], as a marker establishing reference to previously mentioned individuals in the discourse [Riese et al. 2022: §9.1.5; Saarinen 2022: 445], or some other kind of emphasis [Riese et al. 2022: §9.1.5; Tuzharov 1987: 67]. Georgieva [2022: 6] notes that some Meadow Mari speakers also allow usage of the suffix in anaphoric contexts. For closely related Hill Mari, Khomchenkova [2022; 2023] has argued that the suffix marks topics and has a contrastive use with topics, but less with focus.

Here I show that the main syntactic and interpretational properties of discursive, non-enimitive *-ŽE* are as follows: (1) it attaches to the very right of any nominal constituent; (2) the marked constituents are mostly left-peripheral; (3a) it expresses the selection from a superset and a contrast, but (3b) its usage is optional if the context already provides a superset and a contrast; (4a) it can mark a topic shift; and (4b) marking a shift with *-ŽE* appears to have an adversative function, or expresses unexpectedness. In contemporary usage, the suffix, furthermore, seems to mirror usage of the Russian discourse particles *že* (же) (including the enimitive) and *-to* (-то), cf. [Hirvonen 2023: 109–115], but these functions will not be discussed here.

Property (1), the attachment to the very right of a nominal constituent, is visible most clearly in cases where the suffix follows other possessive suffixes, e.g. in (39). In most cases where the suffix appears on another phrase, it receives an enimitive interpretation, as in, e.g., (40). Property (2), the left-peripheral placement of the marked constituent, is indicative of the fact that *-ŽE* often marks constituents which receive a topical interpretation, since these constituents are

placed in the left periphery (see Section 4). This can be seen in most examples in this section as well. Of special interest are sentences with topical objects in sentence-initial position preceding subjects, cf. example (32) repeated below, as they show this very clearly. This example also illustrates that both of the contrasted expressions need not be marked at the same time, cf. [Hirvonen 2023: 112–113; Georgieva 2024: 13]:

- (41) *A Irina kok tarelka-m vele mušk-an šuktā-š. Kumšo tarelka-žə-m*
 but Irina two plate-ACC only wash-CVB manage-PST1.3SG third plate-ŽE-ACC
tudo mušk-an əš šukto.
 3SG wash-CVB NEG.PST.3SG manage.CNG
 ‘But Irina only managed to wash two plates. **The third plate** she did not manage to wash.’
 (= (32))

Most commonly, the suffix marks a contrast between referents of a superset which is contextually available (property 3a), cf. [Simonenko 2014; 2017; Georgieva 2022]. But importantly, neither superset nor contrasted referent have to be explicitly introduced in the context. Property (3a) can be shown in contexts which typically elicit sequences of contrastive topic and focus, namely if a superset is introduced in the *wh*-question setting the context, and the answer provides a pair-list reading, e.g. (39). When the context provides the superset, its usage is optional, however (property 3b) — in example (39) my consultant did not express a preference for using or omitting the suffix.

Summarizing so far, the suffix marks that, from an available superset, one alternative is chosen and contrasted with another. This happens mostly with topics, and marking is optional if the context already provides the superset. The pragmatic effect of actually using the morpheme can be illustrated best with cases in which the context neither provides nor excludes a superset. In such a context, the suffix implies the presence of a contrast, although the contrast does not have to be known to the speaker (42). This is a typical property of contrastive topic markers [van der Wal 2021: 36]:

- (42) *Maša kočkəš-əm jamdəl-en mo? — Šür-žə-m tudo šolt-en.*
 Maša food-ACC prepare-PST2.3SG Q soup-ŽE-ACC 3SG cook-PST2.3SG
Ves kočkəš-əm tudo jamdəl-en aʔe uke — o-m pale.
 other food-ACC 3SG prepare-PST2.3SG or NEGEXIST NEG-1SG know.CNG
 ‘Did Maša prepare food? — **The soup** she cooked. Whether she cooked other food, I don’t know. (But some other food was cooked [by some one else] as well.)’ (elicited)

The suffix also marks the shift to a topic which is newly introduced, newly changed, or newly returned to (property 4a). Here, a superset could be understood as the set of discourse participants or possible topics, of which then one is chosen as an aboutness topic for the clause (cf. [Khomchenkova 2022; 2023] on Hill Mari). In the following dialogue example, speaker A introduces *pamaš* ‘springs’ into the discourse in their turn. Speaker B shifts the aboutness topic in the sentence to this previously introduced constituent. But then, speaker A again rapidly shifts the topic to *snege* ‘wild strawberries’. In both cases, the topic is marked with -ŽE:

- (43) A. *Pamaš šuko ulo tušto, srazu pamaš gəč vüd-əm numal-ət.*
 spring many exist there at.once spring from water-ACC carry-3PL
 ‘There’s lots of springs, people just go right away to get some water from the springs.’
 B. *Memnan pamaš-əže nimat kod-ən og-əl <...>.*
 1PL.GEN spring-ŽE nothing stay-CVB NEG.3SG-be.CNG
 ‘At our place, **springs**, there is nothing left <...>.’
 A. *Snegə-že vet čot šuko ulo, šoldəra möre gaj.*
 wild.strawberry-ŽE see very much EXIST large garden.strawberry like
 ‘(And) **wild strawberries**, you know, there’s lots of, like large strawberries from the garden.’ [Serdobolskaya, Toldova 2012: 729]

Again, its usage appears to be optional in this function, as most topic shifts are not marked with -ŽE. But if the suffix is used, it may mark adversativity or unexpectedness in topic shift contexts (property 4b). In the beginning of the narration which the following example is taken from, a shed (*saraj* ~ *levaš* ‘shed, barn’) is introduced. The shed is mentioned a few times, but only in the last clause it becomes the sentence topic and is marked with -ŽE, which, in turn, stresses the unexpectedness of the events:

- (44) [“Once, there was a brick factory besides this spring. In addition to the brick factory, they built a shed near the spring.”]

No ikana šorgəkt-en jogə-mo šošo vūd saraj-əm mušk-ən
but once beat.down-CVB flow-PTCP.PASS spring water shed-ACC wash-CVB

nanğaj-en. Tiddeč vara levaš-əm ves vere čon-en-ət.
take.away-PST2.3SG from.this then shed-ACC other to.place build-PST2-3PL

Revolucij žap-əšte kerməč jamdələmaš čarn-en-ət. Saraj-že šuko žap
revolution time-INE brick production stop-PST2.3PL shed-ŽE much time

eše šinč-en.
still sit-PST2.3SG

‘But one time the wildly flowing water in spring washed away the shed. After that they built the shed in a different place. After the Revolution they stopped making bricks. **The shed, however**, was still standing for a long time.’ [Serdobolskaya, Tol-dova 2012: 724]

Usage of the suffix is not felicitous if the context does not allow the construal of a superset, and thus violates property (3a), cf. [Georgieva 2022: 5; Hirvonen 2023: 101–106]. Consider examples (45)–(46). Above I argued that -ŽE can mark a topic shift, and Section 4 discussed that these topics are typically placed in the left periphery. Conversely, expressions referring to known individuals, including topics, can be encountered in the postverbal position. In example (45), -ŽE marks a constituent in the postverbal position and the context does not provide or even allow the construal of a superset. My consultant did not accept marking this kind of expression with -ŽE. However, if the context provides a superset and/or a contrast or allows its construal superset and/or a contrast (e.g. a cat and a dog), such marking is felicitous (46):

- (45) [“What about your lost cat — did you look for it?”]

*Da, i mu-ən lukt-ən-am, pəras-em-əm / *pəras-em-žə-m.*
yes and find-CVB lead.out-PST2-1SG cat-PX1SG-ACC cat-PX1SG-ŽE-ACC

‘Yes, and I did find (him), **my cat**.’ (elicited)

- (46) [“What about your lost cat and dog — did you look for them?”]

Da, i mu-ən lukt-ən-am, pəras-em-əm / pəras-em-žə-m.
yes and find-CVB lead.out-PST2-1SG cat-PX1SG-ACC cat-PX1SG-ŽE-ACC

A pij-em-əm eše mu-ən om-əl.
but dog-PX1SG-ACC yet find-CVB NEG.1SG-be

‘Yes, and I did find (him), **my cat**. But **my dog** I have not found yet.’ (elicited)

Summarizing, the morpheme -ŽE has two main functions: It marks uncontroversial information, serving as an “enimitive” marker, cf. [Panov 2020], or selects an individual from a superset while implying a contrast to the other members of the set. In the latter function, the suffix is commonly used with topical constituents, including contrastive topics and a topic shift, or in contexts of unexpectedness and adversativity. Importantly, the suffix is optional in this function in the sense that the semantic well-formedness of the sentence does not depend on the morpheme.

6. Summary

This paper provided an overview of the morphosyntactic strategies of marking topic, focus, and contrastive topic in Meadow Mari. In this section the most important findings are summarized. Constituent focus can variably be marked by prosodic in-situ marking, immediately preverbal or postverbal positioning, and marginally, a pseudocleft construction. Importantly, no type of focus *requires* a certain strategy, and any type of focus can be marked prosodically, but the marking strategies increase in “markedness” — postverbal focus, for example, is available only in a subset of in-situ focus contexts. Aboutness topics and contrastive topics are placed in the left periphery of the clause. Furthermore, left-dislocated topics can be employed for a topic shift. Finally, the morpheme -ŽE marks the selection of an individual from a superset while implying a contrast to the other members. The morpheme is commonly used with contrastive topics and for topic shift.

As regards focus, Meadow Mari behaves like a rather typical Eurasian SOV language, since it allows preverbal focus. But at the same time, it is not a strict-position focus marking language, as an alternative in-situ marking strategy exists, cf. [Büring 2009]. Of further special interest in a comparative Uralic context is the suffix -ŽE, since a number of other Uralic (and Altaic) languages also employ possessive suffixes in discursive functions, see, e.g., [Nikolaeva 2003; Simonenko 2014; Gerland 2014; Yurayong 2020: 115–119], among many others. This overview of information structure in Meadow Mari contributes to the growing comparative body of work on information structure marking in the Uralic languages, see [Surányi et al. (to appear)], and lays the foundation for more detailed studies on topics in Meadow Mari information structure.

ABBREVIATIONS

1, 2, 3 — the 1st, 2nd, 3rd person

ACC — accusative case

ADD — additive clitic

ADJ — adjectivizing suffix

CMPR — comparative

CNG — connegative verb form

CVB — converb

EMPH — emphatic clitic

EXIST — existential copula

ILL — illative case

INDEF — indefinite prefix

INE — inessive case

INF — infinitive

INS — instrumental

NEG — negation auxiliary or constituent negation

NEGEXIST — negative existential predicate

PL — plural

PST1 — first past tense

PST2 — second past tense

PTCP.ACT — agent relativizing participle

PTCP.PASS — passive participle

PTCL — particle

PX — possessive suffix

Q — polar question marker

SG — singular

SUP — superlative

ŽE — Mari discourse marker -ŽE

REFERENCES

- Alhoniemi 1993 — Alhoniemi A. *Grammatik des Tscheremisschen (Mari)*. Hamburg: Buske, 1993.
- Arkhangelskiy 2019 — Arkhangelskiy T. Corpora of social media in minority Uralic languages. *Proc. of the 5th Workshop on Computational Linguistics for Uralic Languages, Tartu, Estonia, January 7–8, 2019*.
- Pirinen T. A., Kaalep H.-J., Tyers F. M. (eds.). Association for Computational Linguistics, 2019, 125–140.
- Bradley, Hirvonen 2022 — Bradley J., Hirvonen J. Null subjects in Mari. *Null subjects in Slavic and Finno-Ugric: Licensing, structure and typology*. Dalmi G., Tsedryk E., Ceglowski P. (eds.). Boston; Berlin: De Gruyter, 2022, 281–306.
- Büring 2003 — Büring D. On D-trees, beans, and B-accent. *Linguistics and Philosophy*, 2003, 26(5): 511–545.

- Büring 2009 — Büring D. Towards a typology of focus realization. *Information structure*. Zimmermann M., Féry C. (eds.). Oxford: Oxford Univ. Press, 2009, 177–205.
- Büring 2016 — Büring D. (Contrastive) topic. *Oxford handbook of information structure*. Féry C., Ishihara Sh. (eds.). Oxford: Oxford Univ. Press, 2016, 64–85.
- Burukina et al. 2024 — Burukina I., Borise L., den Dikken M. A ‘big DP’ analysis of Russian copular constructions with *ëto*. *Glossa*, 2024, 9(1): 1–37.
- Chkhaidze 1941 — Chkhaidze M. P. *Sintaksis lugovo-vostochnogo mariiskogo yazyka* [Syntax of Meadow Mari]. Moscow: Uchpedgiz, 1941.
- Dezső 1978 — Dezső L. Towards a typology of theme and rheme: SOV languages. *Wortstellung und Bedeutung*. Vol. 1. Conte M.-E., Giacalone Ramat A., Ramat P. (eds.). Tübingen: Niemeyer, 1978, 3–11.
- É. Kiss 1998 — É. Kiss K. Identificational focus versus information focus. *Language*, 1998, 74: 245–273.
- Erteschik-Shir 2007 — Erteschik-Shir N. *Information structure: The syntax-discourse interface*. Oxford: Oxford Univ. Press, 2007.
- Frascarelli 2018 — Frascarelli M. The interpretation of *pro* in consistent and partial null-subject languages: A comparative interface analysis. *Null subjects in Generative Grammar*. Cognola F., Casalicchio J. (eds.). Oxford: Oxford Univ. Press, 2018, 211–239.
- Frascarelli, Hinterhölzl 2007 — Frascarelli M., Hinterhölzl R. Types of topics in German and Italian. *On information structure, meaning and form: Generalizations across languages*. Schwabe K., Winkler S. (eds.). Amsterdam; Philadelphia: John Benjamins, 2007, 87–116.
- Georgieva 2022 — Georgieva E. Fake possessives, partitives and definites sharing the same D. *Talk at the 45th GLOW Colloquium, Queen Mary Univ. of London, April 28, 2022*.
- Georgieva 2024 — Georgieva E. On sluicing in Meadow Mari and Udmurt. *Talk at Jezik & Linguistics Colloquium, Univ. of Nova Gorica, May 30, 2024*.
- Georgieva et al. 2021 — Georgieva E., Salzmann M., Weisser Ph. Negative verb clusters in Mari and Udmurt and why they require postsyntactic top-down word-formation. *Natural Language and Linguistic Theory*, 2021, 39: 457–503.
- Gerland 2014 — Gerland D. Definitely not possessed? Possessives with non-possessive function. *Frames and concept types: Applications in language and philosophy*. Gamerschlag Th., Gerland D., Osswald R., Petersen W. (eds.). Dordrecht: Springer, 2014, 269–292.
- Givón 1983 — Givón T. Topic continuity in discourse: An introduction. *Topic continuity in discourse: A quantitative cross-language study*. Givón T. (ed.). Amsterdam: John Benjamins, 1983, 1–41.
- Herring 1990 — Herring S. C. Information structure as a consequence of word order type. *Annual Meeting of the Berkeley Linguistics Society*, 1990, 16(1): 163–174.
- Hirvonen 2023 — Hirvonen J. *Focus and topic in Meadow Mari*. Master’s thesis. Vienna: Univ. of Vienna, 2023. <https://theses.univie.ac.at/detail/69244/>.
- Holmberg 2010 — Holmberg A. Null subject parameters. *Parametric variation: Null subjects in Minimalist theory*. Biberauer Th., Holmberg A., Roberts I., Sheehan M. (eds.). Cambridge: Cambridge Univ. Press, 2010, 88–124.
- Khomchenkova 2022 — Khomchenkova I. A. Diskursivnye funktsii posessivnogo pokazatelya tret’ego litsa edinstvennogo chisla v gornomariiskom yazyke [Discourse functions of the 3rd person singular possessive marker in Hill Mari]. *Ural-Altaic Studies*, 2022, 45(2): 148–168.
- Khomchenkova 2023 — Khomchenkova I. A. Diskursivnye funktsii posessivnogo pokazatelya 3-go litsa edinstvennogo chisla [Discourse functions of the 3rd person singular possessive marker]. *Elementy gornomariiskogo yazyka v tipologicheskoy osveshchenii*. Kashkin E. V. (ed.). Moscow: Buki Vedi, 2023, 614–630.
- Kim 1988 — Kim A. H.-O. Preverbal focusing and type XXIII languages. *Studies in syntactic typology*. Hammond M., Moravcsik E. A., Wirth J. (eds.). Amsterdam: John Benjamins, 1988, 149–171.
- Krifka 2008 — Krifka M. Basic notions of information structure. *Acta Linguistica Hungarica*, 2008, 55(3–4): 243–276.
- Nelson et al. (to appear) — Nelson D., Vedernikova E., Bradley J. The syntax of Meadow Mari: Some issues around TAM and phrase structure. *Uralic syntax*. Tamm A., Vainikka A. (eds.). Cambridge: Cambridge Univ. Press. To appear.
- Nikolaeva 2003 — Nikolaeva I. Possessive affixes in the pragmatic structuring of the utterance: Evidence from Uralic. *International Symposium on Deictic Systems and Quantification in Languages Spoken in Europe and North and Central Asia: Coll. of papers*. Suihkonen P., Comrie B. (eds.). Izhevsk; Leipzig: Udmurt State Univ.; Max Planck Institute for Evolutionary Anthology, 2003, 130–145.

- Panov 2020 — Panov V. The marking of uncontroversial information in Europe: Presenting the enimitive. *Acta Linguistica Hafniensia*, 2020, 52(1): 1–44.
- Pengitov 1961 — Pengitov N. T. Poryadok slov v prostom rasprostranennom predlozhenii [Word order in a simple expanded sentence in Mari]. *Trudy MarNIIYaLI*, 1961, 15: 113–124.
- Pregla 2024 — Pregla A. *Word order variability in OV languages: A study on scrambling, verb movement, and postverbal elements with a focus on Uralic languages*. PhD thesis. Potsdam: Univ. of Potsdam, 2024. <https://doi.org/10.25932/publishup-64363>
- Reinhart 1981 — Reinhart T. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 1981, 27(1): 53–94.
- Riese et al. 2022 — Riese T., Bradley J., Yefremova T. *Mari (марий йыме): An essential grammar for international learners*. Vienna: University of Vienna, 2022. grammar.mari-language.com.
- Saarinén 2022 — Saarinen S. Mari. *The Oxford guide to the Uralic Languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 432–470.
- Serdobolskaya, Toldova 2012 — Serdobolskaya N. V., Toldova S. Yu. Teksty na lugovom mariiskom yazyke (starotor’yal’skii govor) [Texts in Meadow Mari (Starotoryal dialect)]. *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisanija. Formal’nyi i funktsional’nyi podkhody*. Kuznetsova A. I. (ed.). Moscow: Rukopisnye pamyatniki Drevnei Rusi, 2012, 709–745.
- Simonenko 2014 — Simonenko A. Microvariation in Finno-Ugric possessive markers. *NELS 43: Proc. of the 43th annual meeting of the North East Linguistic Society*. Vol. 2. Huang H.-L., Poole E., Rysling A. (eds.). Amherst: GLSA, 2014, 127–140.
- Simonenko 2017 — Simonenko A. Towards a semantic typology of specific determiners. *Proc. of the 21st Amsterdam Colloquium*. Cremers A., van Gessel Th., Roelofsen F. Amsterdam: Institute for Logic, Language and Computation, 2017, 425–434.
- Skopeteas, Fanselow 2010 — Skopeteas S., Fanselow G. Focus types and argument asymmetries. *Comparative and contrastive studies of information structure*. Breul C., Göbbel E. (eds.). Amsterdam; Philadelphia: John Benjamins, 2010, 169–197.
- Surányi et al. (to appear) — Surányi B., Brattico P., Asztalos E., Sahkai H. Information structure in Finno-Ugric languages. *Uralic syntax*. Tamm A., Vainikka A. (eds.). Cambridge: Cambridge Univ. Press. To appear.
- Tuzharov 1987 — Tuzharov G. M. *Grammaticheskie kategorii imeni sushchestvitel’nogo v mariiskom yazyke* [Grammatical categories of the noun in Mari]. Yoshkar-Ola: Mari Book Publ., 1987.
- van der Wal 2016 — van der Wal J. Diagnosing focus. *Studies in Language*, 2016, 40(2): 259–301.
- van der Wal 2021 — van der Wal J. *The BaSIS basics of information structure*. Aug. 2021. Manuscript.
- Vilkuna 1998 — Vilkuna M. Constituent order in European Uralic. *Constituent order in the Languages of Europe*. Siewierska A. (ed.). Berlin: de Gruyter, 1998, 173–233.
- Vilkuna 2022 — Vilkuna M. Word order. *The Oxford guide to the Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 950–960.
- Volkova et al. 2024 — Volkova A., Zakirova A., Voronov M., Dolgodvorova M., Klyucheva Z., Kokoreva S., Makarchuk I., Khomchenkova I., Arkhangel’skiy T., Sokur E. *Spoken corpus of Meadow Mari (as spoken in the village of Staryj Torjal, Novyj Torjal district, Mari El Republic, Russia)*. Moscow: HSE Univ., 2024. Electronic resource: http://lingconlab.ru/spoken_meadow_mari/ (accessed on 01.07.2024).
- Yurayong 2020 — Yurayong Ch. *Postposed demonstratives in Finnic and North Russian Dialects*. Doctoral diss. Helsinki: Univ. of Helsinki, 2020.
- Yurayong et al. 2024 — Yurayong Ch., Niva H., Miestamo M., Kittilä S., Sandman E. Miscellanea typologica uralica ex oriente et occidente. *Itämeren kieliapajilta Volgan verkoille: Pühendusteos Riho Grünthalile 22. mail 2024*. Björklöf S., Jantunen S., Junttila S., Kuokkala J., Laitinen S., Pasanen A. (eds.). Helsinki: Suomalais-Ugrilainen Seura, 2024, 729–760.

Условные и уступительные конструкции с русскими союзами *если* и *хотя* в горномарийской речи

© 2024

Ирина Андреевна Хомченкова

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

Институт русского языка им. В. В. Виноградова РАН, Москва, Россия;

Институт языкознания РАН, Москва, Россия; irina.khomchenkova@yandex.ru

Аннотация: В горномарийской речи встречаются как исконные союзы (например, *gән' (ә)* ‘если’, *gән'üt* ‘хотя’), так и заимствованные из русского (например, *если*, *хотя*). Мы рассмотрели устройство условных и уступительных клауз в горномарийском языке и сравнили их с русскими конструкциями. Структурно эти конструкции в русском и горномарийском похожи (финитная форма глагола, наличие союза), отличается лишь позиция союза. На основе корпусных данных мы показали, что в горномарийской речи союз *если* употребляется препозитивно, как и в русском, возможно также его дублирование союзом *gән' (если... gән')*. Союз *хотя* в наших данных продублирован не был; он также в большинстве случаев употребляется препозитивно. Мы продемонстрировали, что выбор русского vs. горномарийского союза зависит от конкретного носителя (корреляций с какими-либо социолингвистическими параметрами выявлено не было). Также отмечена тенденция к более частому употреблению русских союзов по сравнению с более ранними данными других исследователей. Мы предполагаем, что использование русских союзов можно объяснить структурной конгруэнтностью между стратегиями оформления условных клауз в языке-реципиенте (горномарийском) и языке-доноре (русском). При этом конгруэнтность должна быть неполной — дублирование союзов возможно, поскольку для русского союза остается незанятый слот в начале протатиса.

Ключевые слова: заимствования, марийские языки, русский язык, синтаксис, союзы, уральские языки, условные конструкции, уступительность, языковые контакты

Благодарности: Исследование выполнено при поддержке Российского научного фонда (проект № 24-18-00199 «Структура клаузы и позиционные феномены в языках SOV»), выполняемый в Институте языкознания РАН).

Для цитирования: Хомченкова И. А. Условные и уступительные конструкции с русскими союзами *если* и *хотя* в горномарийской речи. *Вопросы языкознания*, 2024, 5: 116–134.

DOI: 10.31857/0373-658X.2024.5.116-134

Conditional and concessive constructions with Russian conjunctions *если* ‘if’ and *хотя* ‘although’ in Hill Mari speech

Irina A. Khomchenkova

Lomonosov Moscow State University, Moscow, Russia;

Vinogradov Russian Language Institute, Russian Academy of Sciences, Moscow, Russia;

Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia;

irina.khomchenkova@yandex.ru

Abstract: In Hill Mari speech, there are both native conjunctions (for example, *gән' (ә)* ‘if’, *gән'üt* ‘although’) and borrowed Russian conjunctions (for example, *если* ‘if’, *хотя* ‘although’). I examined the structure of conditional and concessive clauses in Hill Mari and compared them with the Russian ones.

Structurally, these constructions are similar in Russian and Hill Mari (finite verbs, presence of a conjunction), only the position of the conjunction differs. Based on corpus data, I showed that in the Hill Mari speech, the conjunction *esli* 'if' is used prepositively, as in Russian, and can also be duplicated by the conjunction *gән* (*esli... gән*). The conjunction *xotja* 'although' was not duplicated in my data; it also tends to be used prepositively. I have demonstrated that the choice between the Russian and Hill Mari conjunction depends on the speaker (no correlations with sociolinguistic parameters were found). A tendency towards more frequent use of Russian conjunctions was also noted in comparison to earlier data of other researchers. I assume that the use of Russian conjunctions can be explained by structural congruence between the strategies of conditional clause formation in the recipient language (Hill Mari) and in the donor language (Russian). At the same time, the congruence should be incomplete — duplication of conjunctions is possible, since there is an unoccupied slot at the beginning of the protasis for the Russian conjunction.

Keywords: concessive, conditional constructions, conjunctions, language contacts, loanwords, Mari, Russian, syntax, Uralic

Acknowledgements: This research was supported by the Russian Scientific Foundation (project No. 24-18-00199 "Clause structure and positional phenomena in SOV languages" realized at the Institute of Linguistics, Russian Academy of Sciences).

For citation: Khomchenkova I. A. Conditional and concessive constructions with Russian conjunctions *esli* 'if' and *xotja* 'although' in Hill Mari speech. *Voprosy Jazykoznanija*, 2024, 5: 116–134.

DOI: 10.31857/0373-658X.2024.5.116-134

1. Введение

Горномарийский язык — один из марийских языков (наряду с луговым марийским), относящийся к финно-угорской группе уральской семьи. Он имеет высокую степень сохранности: передается детям, используется в качестве основного языка общения в бытовой коммуникации (в первую очередь старшим и средним поколениями), см. [Шабыков и др. 2020; Кашкин 2023а: 3–4].

Марийские языки находятся в тесном контакте с русским языком уже долгое время. Одним из результатов этих контактов стали многочисленные русские заимствования, проникшие в марийскую лексику, см. подробное исследование в [Саваткова 1969]. Большая часть марийцев является билингвами — согласно переписи населения 2020 г.¹, русским языком владеют 98 % марийцев (416087 человек из 423803). Помимо русских заимствований, в марийской речи многочисленны случаи переключения кодов, которые рассматриваются, например, в [Гаврилова 2013; Дьячков 2020].

В частности, в марийской речи встречаются русские союзы. В. Г. Гаврилова приводит несколько иллюстраций из лугового марийского языка, в том числе изъяснительный союз *што*, условный союз *если*, целевой союз *штобы* и причинный союз *потому што*, и отмечает, что «некоторые виды сложных предложений строятся по моделям русского языка, либо с использованием только союзов русского языка, либо дополнительно его дублированием на марийском языке» [Гаврилова 2013: 18].

А. А. Саваткова среди заимствованных подчинительных союзов в горномарийском языке упоминает *ыштовы*, *ышитвы* 'чтобы', *потому шты* 'потому что', *хоть*, *хотя* [Саваткова 1969: 70], см. также [Саваткова 2002: 258]. И. С. Галкин [1964: 179] приводит союзы *хоть* и *потому что*. Перечисляются русские союзы, зафиксированные в различных источниках, и в работе [Сибатрова 2016: 36–38].

¹ Итоги Всероссийской переписи населения 2020 г. доступны по ссылке https://rosstat.gov.ru/vpn_roru/ (Том 5 «Национальный состав и владение языками», Табл. 3. Население по национальной принадлежности, владению русским языком и его использованию).

Объектом нашего исследования являются условные и уступительные конструкции с русскими союзами *если*² и *хотя* (*хоть*). Такие конструкции состоят из двух частей: протасиса (или антецедента) и аподосиса (или консеквента). Например, в предложении *Если будет солнце, мы пойдем в парк* первая клауза является протасисом, а вторая — аподосисом.

В качестве материалов для анализа мы использовали корпус горномарийского языка (далее ГК; на текущий момент он включает 63522 словоупотребления, см. hillmari-exr.tilda.ws/corpus), собранный коллективом экспедиции ОТиПЛ МГУ под руководством Е. В. Кашкина в с. Кузнецово и близлежащих деревнях Горномарийского района республики Марий Эл в 2016–2019 гг. Корпус также включает относящиеся к тому же говору тексты, опубликованные в [Акцорин 1991] (год записи — 1965³) и [Саватеева 2005] (года записи — 2003 и 2005).

Условные конструкции были выбраны для исследования, поскольку по абсолютной частотности зафиксированных русских союзов в корпусе *если* является наиболее распространенным (так, условный союз *если* встретился в два раза чаще, чем целевой союз *чтобы*). При этом, несмотря на то что союз *хотя* не является частотным (88 вхождений *если* при 7 вхождениях *хотя*), мы рассматриваем и уступительные конструкции, так как горномарийский уступительный союз *gән 'ät* 'хотя' образован от условного союза *gән* 'если' с помощью аддитивной частицы; в текстах встречается и употребление *gән 'ät* с композиционным значением 'даже если'. Таким образом, рассматривать условные конструкции в отрыве от уступительных не является целесообразным.

Вопрос о заимствовании (подчинительных) союзов обсуждается в ряде работ, см., например, [Майтинская 1982: 102; Thompson et al. 2007: 268–269; Grant 2012: 350]. С учетом наблюдений, представленных в них, на материале горномарийского языка мы попробуем ответить на вопрос, почему и как именно язык может включать в свою систему неисконный союз.

В [Майтинская 1982] рассматривается наличие русских союзов в речи носителей финно-угорских языков. К. Е. Майтинская отмечает, что в финской и венгерской речи заимствованные союзы практически не встречаются, поскольку в языках, «литературные варианты, а также письменность которых развивались относительно равномерно и постепенно в течение многих веков, средства связи между предложениями-компонентами формировались постепенно по мере необходимости оформления сложных синтаксических конструкций, особенно характерных для письменного литературного языка. В “готовых” средствах связи срочной необходимости не возникало, слова собственного языка в современных союзы преобразовывались постепенно» [Там же: 102]. Заимствованные союзы часто встречаются в языках, «которые развивались под особенно сильным влиянием окружающих языков и литературные варианты, а также письменность которых были слабо развиты» [Там же].

Аналогичное рассуждение проводится в [Thompson et al. 2007: 268–269], где обсуждаются адвербиальные клаузы различных типов. Авторы предполагают, что языки, лишенные специализированных средств для соединения клауз, будут их заимствовать при контактах с языками, в которых такие средства есть.

Однако этим можно объяснить случаи заимствования таких союзов, для которых нет аналогов в языке-реципиенте, например *потому что* (1). Условные и уступительные союзы же в горномарийском присутствуют.

- (1) *a krol'ik juk-äm=at a-k lâk potomu što*
 а кролик звук-ACC=ADD NEG.NPST-3 извлекать потому что

² В [Саваткова 2002: 258] также упоминается условный союз *ras* (русс. *раз*), однако примеры не приводятся; в наших данных он не был засвидетельствован.

³ Тексты из этого источника не включали условных и уступительных клауз; таким образом, мы рассматриваем только современные тексты.

tädə vosp'itannâj âl-ân.
 тот воспитанный быть-ПРЕТ

‘А Кролик не издает ни звука, потому что он воспитанный’. [ГК: «Пересказ мультфильма о Винни-Пухе», 83]

В работе [Grant 2012] приводится следующее обобщение: «...менее частотные показатели зависимых клауз, по-видимому, заимствуются чаще, чем такие формы, как ЕСЛИ или ПОТОМУ ЧТО, которые можно считать более базовыми, поскольку они являются “прототипическими” маркерами своих типов зависимых клауз (условных и причинных соответственно)»⁴. Однако горномарийский заимствовал «базовый» условный союз *если*, несмотря на наличие исконного союза.

В [Thompson et al. 2007: 269] приводится еще одно обоснование заимствования союзов: оно происходит, потому что носители-билингвы стремятся «создавать паттерны в одном из их языков, которые структурно параллельны тем, которые есть в другом языке»⁵. Мы обсудим условные и уступительные конструкции с русскими союзами с этой точки зрения, используя идею (ин)конгруэнтности между стратегиями оформления этих клауз в языке-реципиенте (горномарийском) и в языке-доноре (русском).

Понятие конгруэнтности упоминалось еще в [Weinreich 1953]: «...совершенно ясно, что наличие большой конгруэнтности в структуре облегчает перенесение морфем, ибо сильно связанная морфема настолько зависит от своей грамматической функции, что, если в чужой системе для нее не окажется готовой функции, эта морфема окажется там бесполезной»⁶ [Вайнрайх 1979: 65]. Об этом понятии как о механизме регулирования структурных ограничений на контактно-обусловленные феномены см. также [Harris, Campbell 1995: 123–125; Aikhenvald 2007; Sebba 2009; Besters-Dilger et al. (eds.) 2014].

Мы опишем способы выражения условной и уступительной семантики в горномарийском языке и обсудим, как русские союзы встраиваются в существующую систему, сравнив наши данные с приведенными в работах [Riese 1984] и [Гаврилова 2013].

2. Показатели с условной / уступительной семантикой

2.1. Союзы *gän'ä*, *gän'üt*

В горномарийском языке в условных клаузах употребляется союз *gän'* (имеющий также форму *gän'ä*, которая с ним взаимозаменяема; далее приводится только вариант *gän'*) [Саваткова 2002: 258], ср.:

- (2) *a-t čuč gän'ä madäš gäc läkt-ät.*
 NEG.NPST-2SG попадать если игра EL выходить-NPST.2SG

‘Если не попадешь — из игры выйдешь’. [ГК: «Игры-5», 175]

⁴ «Less frequently used dependent clause markers seem to be borrowed more readily than forms such as IF and BECAUSE, which can be considered to be more basic inasmuch as these are ‘prototypical’ markers of their kinds of dependent clauses (conditional and causal respectively)» [Grant 2012: 350]. (Здесь и далее перевод наш. — И.Х.)

⁵ «...to create patterns in one of their languages which are structurally parallel to those found in the other» [Thompson et al. 2007: 269].

⁶ «It stands very much to reason that the transfer of morphemes is facilitated between highly congruent structures; for a highly bound morpheme is so dependent on its grammatical function (as opposed to its designative value) that it is useless in an alien system unless there is a ready function for it» [Weinreich 1953: 35].

В горномарийском корпусе встретилось 85 условных конструкций с союзом *gän'*. В большинстве случаев этот союз располагается на правой периферии клаузы (протасиса), как в (2). Лишь в двух случаях он оказался в предпоследней позиции — в (3) за ним следует частица *molo* 'еще', в (4) — отрицательное местоимение с элативным послелогом.

- (3) *i vot vara ves t'el'ežkă-m čäl'-en ke-ä,*
и вот потом другой тележка-ACC привязывать-CVB идти-NPST.3SG

kâm täžem cävă gän' molo.
три тысяча курица если еще

‘И вот потом привязывает другую тележку, если там 3000 куриц’. [ГК: «Птицефабрика», 41]

- (4) *a a-t lüd gän' n'i-ma gäc=ät,*
а NEG.NPST-2SG бояться если NEG-что EL=ADD
cilä nelä-m=ok säng-äš li-eš.
весь тяжелый-ACC=EMPH побеждать-INF становиться-NPST.3SG

‘А если не боишься ничего, то можно осилить все трудности’. [ГК: «Описание фотографий», 42]

Подробное описание конструкций с *gän'* доступно в [Riese 1984]. Т. Ризе также отмечает, что *gän'* располагается строго на последнем месте протасиса, и исключения немногочисленны [Ibid.: 174–175].

От условного союза *gän'* образован уступительный союз *gän'ät* ‘хотя’ (5). Согласно [Галкин 1964: 182], он включает аддитивную частицу *=at* (=ät после передних гласных в соответствии с правилами сингармонизма); такая модель образования уступительных союзов типологически распространена [König 1991: 82]. 10 примеров с союзом *gän'ät* в корпусе имеют уступительную семантику.

- (5) *män'-än äškal-em ulâ kăzät=ät šket-em=ok*
я-GEN корова-POSS.1SG EXST сейчас=ADD в_одиночку-POSS.1SG=EMPH
äl-am gän'=ät.
быть-NPST.1SG если=ADD

‘Я до сих пор держу корову, хоть и одна’. [ГК: «В старину», 88]

В нескольких случаях сочетание союза *gän'* и частицы *=at* имеет композиционное значение ‘даже если’, которое типологически часто является исходным для союзов со значением ‘хотя’, имеющих такую структуру [König 1991: 82]. Конструкции такого типа (‘even if’) называются условными уступительными (concessive conditionals), см. [Haspelmath, König 1998; Храковский 2004]. В этих контекстах в горномарийской речи ни одиночный, ни продублированный русский союз *если* употреблен не был. В корпусе встретилось пять примеров с сочетанием условного союза *gän'* и аддитивной частицы *=at* в условно-уступительных контекстах, см., например, (6). При этом в двух случаях значение частицы *=at* усиливалось русской лексемой *даже*, как в (7).

- (6) *ves-ä-vlä-m=ät pâr-t-aš li-eš vara*
другой-FULL-PL-ACC=ADD входить-CAUS-INF становиться-NPST.3SG потом
cilä-n-äštä-m, kü-vlä-m vyšybaj-en lâkt-än-ät
весь-FULL-POSS.3PL-ACC кто-PL-ACC вышибать-CVB извлекать-PRET-3PL
gän'=ät tädä-m pâr-t-aš li-eš.
если=ADD тот-ACC входить-CAUS-INF становиться-NPST.3SG

‘Других пускать можно, всех, кого пусть и вышибли, их пускать можно’. [ГК: «Игры-2», 34]

- (7) *dažä m'insk motocikâl r'eg'istr'irâvannâj agâl gän'=ät.*
даже Минск мотоцикл зарегистрированный NEG если=ADD

[Там никакого ГАИ нет, прав не надо, можно кататься просто так.] ‘Даже если мотоцикл “Минск” не зарегистрирован’. [ГК: «Мотоцикл “Минск”», 29]

2.2. Маркер условной конструкции *âlgescă*

В корпусе встретился один пример с показателем *âlgescă* ‘COND’ (8); он упоминается в [Саваткова 2002: 258], см. также [Riese 1984: 191]. Эта единица представляет собой форму кондиционала (условно-сослагательного наклонения) от глагола *âlaş* ‘быть’. Это наклонение описывается в [Alhoniemi 1993: 116–117; Саваткова 2002: 195–197], однако оно не используется в рассматриваемом говоре [Мордашова 2023б: 350]. Застывшая форма *âlgescă* стандартно используется в конструкциях со значением ирреального условия. При употреблении союза *gân* в таких контекстах нереальность маркируется формой прошедшего времени с показателем ретроспективного сдвига (9) [Мордашова 2023а: 343].

- (8) *mašină man-ân-at âlgescă, pâl-em âlân srază.*
 машина говорить-PRÉT-2SG COND знать-NPST.1SG RETR2 сразу

‘Если бы сказала, на машине, я бы узнал сразу’. [ГК: «Игра в шляпу-2», 142]

- (9) *irgod-âm âşkal-âm a-m orolâ âl’â gân’,*
 завтрашний_день-ACC корова-ACC NEG.NPST-1SG пасти RETR1 если
măn’ tă don-da păşă-l-em âl’â.
 я вы с-POSS.2PL работа-DENOM-NPST.1SG RETR1

‘Если бы я завтра не пасла коров, я бы с вами поработала’. [Мордашова 2023а: 343]

2.3. Союз *dâk*

Также в корпусе встретился один пример с показателем *dâk* (10). В форме *âk / âk⁷* он упоминается в [Саваткова 2002: 258] при перечислении условных союзов. Согласно [Riese 1984: 177], союз *tak / dak* употребляется в яранском и ветлужском диалектах (которые объединяются в рамках «северо-западного наречия марийского языка»), см. также [Иванов, Тужаров 1970: 200]. Т. Ризе отмечает, что в этих диалектах не используется союз *kân* (соответствующий горномарийскому *gân*). Союз *dâk* также располагается на правой периферии протасиса [Riese 1984: 177]. В [Ibid.: 178–179] предполагается, что этот союз развился из русской частицы *так*. Возможно также сопоставление этой единицы с частицей *дак*, распространенной в северных русских диалектах, см., например, обсуждение условных и временных конструкций с *dak* в камасинской речи [Гусев 2020].

- (10) *şudă koşk-en dâk⁸ tädă-m âdâr-aş kel-eş.*
 трава сохнуть-PRÉT так он-ACC собирать-INF быть_нужным-NPST.3SG

‘Если сено высохло, надо его сгребать (надо метать стога)’. [ГК: «В старину», 78]

Согласно [Галкин 1964: 182], *dâk* может употребляться в уступительном значении в сочетании с частицей *=at*; в северо-западном наречии зафиксирована также форма *dâkănăt*,

⁷ Фонема, обозначаемая как /d/ в ГК и в текущей работе, может реализовываться как зубной [d] (в большинстве случаев) и межзубный [ð] звуки. Она соответствует фонеме /ð/ по [Alhoniemi 1993; Саваткова 2002]. См. о системе записи [Кашкин 2023б].

⁸ Отметим, что в этом предложении мы не можем доказать, что *dâk* находится в конце протасиса, ср. *Сено высохло, так его собирать надо*. Цитируемый текст опубликован в [Саватеева 2005: 293–295] и не содержит аудиозаписи.

представляющая собой сочетание единиц *dāk* и *gān 'ūt*. В [Саваткова 2002: 259] также упоминается уступительный союз *ḍākat / ḍākāt*, однако в нашем корпусе такие единицы не были зафиксированы.

2.4. Прочие способы выражения условной / уступительной семантики

Условное значение может иметь деепричастие на *-mākā* 'CVB.ANT', которое употребляется в контекстах предшествования (11). Таких примеров (с семантикой условия) в корпусе 15, см., например, (12)–(13). Наличие условного значения у *-mākā* в горномарийском языке также фиксируется в [Riese 1984: 185–186] («an adverbial infinitive being used in a conditional sense»). В [Thompson et al. 2007: 257] отмечается, что в некоторых языках условные («'if' clauses») и временные («'when' clauses») клаузы не различаются.

- (11) *dā jāt li-mākā tō-škā pār-en vaz-āt.*
и ночь становится-CVB.ANT тот-ILL входит-CVB ложиться-NPST.3PL
[Себе яму копают.] 'И, когда наступила ночь, туда ложатся'. [ГК: «Два старика», 9]

- (12) *tol-mākā-štā kogo-n susu li-ām*
приходить-CVB.ANT-POSS.3PL большой-ADV веселый становится-NPST.1SG
āl-ne-žā.
быть-DES-POSS.3SG
'Если придут, был бы очень рад'. [ГК: «Разговоры о жизни», 32]

- (13) *kāzāt-šā sravn'ivajā-mākā kāzāt-šā ālā-māš-ām*
сейчас-ATTR сравнить-CVB.ANT сейчас-ATTR жить-NMLZ-ACC
sravn'ivajā-mākā voopš'e kāzāt c жиру бесятся.
сравнить-CVB.ANT вообще сейчас
'Если сравнивать с тем, что сейчас, если сравнивать с жизнью сейчас, сейчас вообще с жиру бесятся'. [ГК: «Биография», 33]

Однако в [Муравьев 2023: 778–779] наличие у *-mākā* условного значения оспаривается: «...такое употребление *-mākā* является следствием отнесения контекста к плану будущего и все равно подразумевает отношение предшествования между событиями, поэтому не может быть признано самостоятельным значением», см. неграмматичный пример (14), где событие аподосиса предшествует событию протасиса.

- (14) **vas'a to-kā-žā tol-mākā tādā-m anzā-c kolt-en-āt.*
Вася дом-ILL2-POSS.3SG приходить-CVB.ANT он-ACC перед-EL посылать-RET-3PL
Ожидаемое значение: 'Если Вася пришел домой, значит его раньше отпустили'. [Муравьев 2023: 779]

В восьми случаях в русских переводах предложений, не содержащих никаких условных показателей, были использованы условные конструкции, см., например, (15). Такие примеры рассматриваются и в [Riese 1984: 182]. Во всех этих случаях представлена последовательность событий, поэтому данная стратегия также не является собственно условной.

- (15) *kol'ca-š cama-m popaz-ākt-at kāč-at vara.*
кольцо-ILL жеребенок-ACC попадать-CAUS.DIST-NPST.3PL держать-NPST.3PL потом
'Если попадали жеребенком в кольцо, то их тогда ловили'. [ГК: «Акрям», 8]

Условную семантику можно усмотреть и в некоторых других примерах, где в русском переводе не используется союз *если*. Например, в (16) «в переводе на русский» можно перефразировать как «если переводить на русский», а в (17) в переводе можно добавить *если* — «если сказала бы».

- (16) *ruš-la-škâ sārä-mäkâ «Долина мельниц» man-alt-eš.*
 русский-SIM-ILL вращать-CVB.ANT говорить-MED-NPST.3SG
 ‘В переводе на русский будет «Долина мельниц»’. [ГК: «О Вакшлапе», 8]
- (17) *vākšlap-âštâ rovojat-en man-ân — srazu päl-em âl’â.*
 Вакшлап-IN работать-PRET говорить-CVB сразу знать-NPST.1SG RETR1
 ‘Сказала бы, что «в Вакшлапе работала» — сразу понял бы’. [ГК: «Игра в шляпу — 2», 152]

В некоторых случаях русские переводы похожих предложений могут включать как условный союз *если* (18), так и временной союз *когда* (19).

- (18) *ärz-ält-eš, arava-žâ ärz-ält-mäkâ kâce*
 трясти-MED-NPST.3SG телега-POSS.3SG трясти-MED-CVB.ANT как
man-alt-eš?
 говорить-MED-NPST.3SG
 ‘Трясется, если колесо трясется, как называется?’ [ГК: «Игра в шляпу-3», 132]
- (19) *t’el’ega, arava-žâ ärz-ält-mäkâ kâce man-alt-eš?*
 телега колесо-POSS.3SG трясти-MED-CVB.ANT как говорить-MED-NPST.3SG
 ‘Телега, когда колесо трясется, как называется?’ [ГК: «Игра в шляпу-3», 133]

Таким образом, мы исключаем из рассмотрения способы выражения условной семантики, описанные в этом разделе, поскольку названные стратегии не являются собственно условными, а также единицу *dâk*, ввиду неоднозначности единственного примера с ней. В качестве аналогов условных клауз с *если* мы будем рассматривать конструкции с *gân* ‘(ä) ‘если’ (85 примеров) и *âlgecä* ‘COND’ (1 пример), а при описании уступительных клауз с *хотя* — конструкции с *gân’ât* ‘хотя’ (10 примеров).

3. Конструкции с русскими союзами *если* и *хотя*

Структура условных конструкций в русском и горномарийском языках различается позицией союза: *если* занимает начальную позицию в клаузе, а *gân* — конечную.

В горномарийском корпусе зафиксированы 88 условных клауз с русским союзом *если*, см., например, (20).

- (20) *a jesl’i teve d’iab’et’ik-vlä saxâr-âm kačk-ât nânâ jasâlan-at.*
 а если вот диабетик-PL сахар-ACC есть-NPST.3PL они болеть-NPST.3PL
 ‘А если вот диабетики сахар едят, они болеют’. [ГК: «Зимовье пчел», 40]

Также возможно сочетание русского и горномарийского союза в одном предложении (21); такие примеры приводятся и в [Гаврилова 2013: 18] для лугового марийского языка. «Смешанных» примеров в корпусе встречается 36 — примерно половина от всех примеров с русск. *если*. («Смешанные» примеры включались в категорию предложений с русскими союзами и не учитывались при подсчете общего количества примеров с горномарийскими союзами.)

- (21) *jesl’i čädä-n âl-at gân’â ik igra vele prostâ.*
 если мало-FULL быть-NPST.2SG если один игра только просто
 ‘Если мало народу, то только одна игра просто’. [ГК: «Игры-7», 20]

Союз *если* может располагаться не только в абсолютном начале протасиса, но и во второй позиции:

- (22) *tädä-m jesl'i jār-žä káč-aš gän'ä,*
 тот-ACC если вокруг-POSS.3SG держать-INF если
văc m'etär dä pelä kel-eš.
 пять метр и половина быть_нужным-NPST.3SG
 [Но в нашей деревне есть такое дерево, дуб.] 'Если его обхватить вокруг, нужно пять
 с половиной метров'. [ГК: «Дуб», 6]

В одном примере встретилось дублирование *gän'* в середине протасиса (23). В другом же примере союз *gän'* был также продублирован, однако располагался в начале протасиса после союза *если* (24).

- (23) *jesl'i äl-äš tumaj-edä gän' ti vār-äšt=ok gän'ä,*
 если жить-INF думать-NPST.2PL если этот место-IN=EMPH если
to karem tor-eš sola-m šänd-ädä.
 то овраг поперек-LAT деревня-ACC сажать-IMP.2PL
 'Если думаете жить на этом же месте, то ставьте деревню поперек оврага'. [ГК: «Хозяйство», 112]
- (24) *jesl'i gän' mäch-äm šu-en koltâ-mâ, tã, kom... prot'iv... ma,*
 если если мяч-ACC бросить-CVB посылать-NMLZ тот что
ploš'adkâ-štâ edem xvat'-a gän', sv'ečkâ-m xvat'-a
 площадка-IN человек хватать-NPST.3SG если свечка-ACC хватать-NPST.3SG
gän', značit (...).⁹
 если значит
 'Если, когда бросили мяч, человек, который на площадке, поймает, если поймает свечку, значит, [это, считается, эта команда проигрывает, местами меняются]'. [ГК: «Игры», 26]

В корпусе есть лишь один пример с конструкцией со значением ирреального условия, содержащий русский союз *если* (25)¹⁰. В связи с ограниченным количеством примеров, по корпусным данным нельзя сделать вывод о возможности сочетания союза *если* и единицы *älgecä*, которая типична в таких контекстах (см. раздел 2.2).

- (25) *projekt semän' teve ti štr'ixujâ-m vār-vlâ-m cišti*
 проект по вот этот заштриховать-PTCP.PASS место-PL-ACC совсем
potopâ-šâš-lâk âlân, jesl'i čeboksarsk'ij ges-än
 затопить-ORT-DEST RETR2 если Чебоксарский ГЭС-GEN
urov'en'-žž-m kud lu kändäkš-šä otm'etkâ jakte
 уровень-POSS.3SG-ACC шесть десять восемь-ORD отметка до
lukt-äl-ät âl'â.
 поднимать-ATT-NPST.3PL RETR1
 'По проекту вот эти заштрихованные места все должны были быть затоплены, если бы уровень Чебоксарской ГЭС подняли до 68-й отметки'. [ГК: «Карта», 55]

В русском языке условные конструкции с *если* могут включать частицу *то* (о различиях между условными конструкциями с этой частицей и без нее см., например, [Пекелис 2015]). В горномарийском корпусе встретились примеры с частицей *то* (см., в частности, (23)). В 12 случаях был также использован горномарийский союз, как в (26), в 13 случаях он отсутствовал (27). Эти данные представлены в табл. 1.

⁹ В этом примере мы учитывали конструкцию *jesl'i gän'... gän'* как продублированную русскую, а последнее вхождение *gän'* (*sv'ečkâ m xvat'a gän'* 'если поймает') как часть горномарийской конструкции.

¹⁰ Ирреальность в данном контексте передается финитными глагольными формами в сочетании с показателями ретроспективного сдвига *âlân* и *âl'â*, см. [Мордашова 2023а: 343–344].

- (26) *jesl'i samân' keles-ä gän', to vidä-šä vaštalt-eš.*
 если неправильно говорить-NPST.3SG если то вести-PTCP.ACT меняться-NPST.3SG
 'Если неправильно скажет, то ведущий меняется'. [ГК: «Игры-6», 35]
- (27) *jesl'i šokšâ li-eš to cilä pospej-en*
 если теплый становиться-NPST.3SG то весь поспевать-CV
šo-kt-a každâj i-n.
 достигать-CAUS-NPST.3SG каждый год-GEN
 'Если будет тепло, то всё успеет поспеть каждый год'. [ГК: «Интервью о хозяйстве», 36]

Таблица 1

Горномарийские примеры с русским союзом *если*

	Есть частица <i>то</i>	Нет частицы <i>то</i>	Всего
Нет союза <i>gän'</i>	13	39	52
Есть союз <i>gän'</i>	12	24	36
Всего	25	63	88

Можно было бы предположить, что частица *то* будет меньше использоваться в конструкции с дублированием, однако в наших данных такая закономерность не наблюдается. Более того, в корпусе встретилось два примера с горномарийским союзом *gän'* и частицей *то* без русского союза *если*, см. (28) и (29). Отметим, впрочем, что в (29) употребление *gän'* кажется избыточным, ср. наличие союза *kânam* и перевод этого примера (который не включает фальстарты и паузы).

- (28) *jarat-âkt-a, a kid-et poka lüd-mä-žž...*
 любить-CAUS.DIST-NPST.3SG а рука-POSS.2SG пока бояться-NMLZ-POSS.3SG
lüd-ân kolt-en kâdal-ân gän', to bol'sâ
 бояться-CV посылать-CV бежать-PRET если то больше
kâč-aš a-k li tädä-m.
 держать-INF NEG.NPST-3 становиться тот-ACC
 [Поласкаешь, если успеешь рукой дотронуться.] 'Дает себя погладить, а если испугался руки и убежал (букв. "пока боязнь... испугавшись, убежал если"), то больше не поймает его'. [ГК: «Кошки», 24]
- (29) *kânam âl-ân p'ečnoj otopl'en'ij gän', to t'et'ä-vlâ, každâj t'et'ä*
 когда быть-PRET печной отопление если то ребенок-PL каждый ребенок
kok kubâm'etr rädä pu-m pil'-en, šel-ân pu-en.
 два кубометр по дерево-ACC пилить-CV раскалывать-CV давать-PRET
 'Когда было печное отопление, то дети, каждый ребенок по два кубометра дров пилил и колот'. [ГК: «Пятая четверть», 37]

Согласно [Саваткова 2008: 48], лексема *gän'* может употребляться не только как условный союз, но и как усилительная частица (*ёшке гъинь* 'сам же'). В корпусе встретился один пример, где *gän'* не имеет явного условного значения (30), который мы не учитывали в подсчетах. Можно предположить, что в (29) *gän'* также может использоваться не как условный союз.

- (30) *ände ik-tä lu=at kâm i ta-ma rovojaj-del-am*
 теперь один-INDEF2 десять=ADD три год INDEF1-что работать-NEG.PRET-1SG
gän'â kolhoz-âštâ.
 если колхоз-IN
 'Теперь, наверное, лет тринадцать прошло, как не работала, в колхозе'. [ГК: «В старину», 11]

Интересно, что в материалах, проанализированных в [Riese 1984: 181], русский союз *если* был засвидетельствован только в козьмодемьянском говоре (т. е. в горномарийском языке). Т. Ризе приводит несколько примеров, содержащих конструкцию с дублированием (*если... gän*’); союз *если* без союза *gän*’ встретился лишь один раз.

В [Гаврилова 2013: 19] приводятся количественные данные об использовании союзов *гын* и *если* в речи носителей лугового марийского языка: «...союз *гын* в спонтанной речи используют 51,6 % студентов, союз *если* — 24,2 %, *если... гын* — 24,2 %», а также о предпочтении опрошенных студентов: *гын* — 28,5 %, *если* — 14,3 %, *если... гын* — 57,5 %. В. Г. Гаврилова делает вывод о том, что на текущий момент в речи преобладает союз *гын*, но союз *если* также распространен и может вытеснить *гын*.

Обобщение представлено в табл. 2. Сравнение горномарийских данных из текстов разных корпусов (более старого, проанализированного Т. Ризе¹¹, и современного, ГК) показывает, что стратегия с русским союзом *если* стала более распространенной; она встречается в речи чаще, чем «смешанная» конструкция с дублированием союза. Также можно отметить, что русский союз *если* проник и в другие говоры: в [Riese 1984] отмечалось, что *если* употребляется лишь в горномарийском языке, однако, согласно [Гаврилова 2013], он используется и в луговом марийском. Также в [Riese 1984] не отмечается возможности использования частицы *то* в условных конструкциях, которая встречается как в наших данных, так и в данных В. Г. Гавриловой.

Таблица 2

Условные конструкции в марийских языках

	[Riese 1984] (горномарийский язык)	ГК (горномарийский язык)	[Гаврилова 2013] (луговой марийский язык)
Марийский союз	много	49,7 % (87 примеров)	51,6 %
Смешанная стратегия	встречается	20,6 % (36 примеров)	24,2 %
Русский союз	мало	29,7 % (52 примера)	24,2 %

Также можно предположить, что использование русского / горномарийского условного союза будет варьировать в зависимости от носителя. К сожалению, надежные статистические выводы сделать невозможно ввиду ограниченного объема корпуса. Мы рассмотрим данные по семи носителям, в текстах от которых больше всего вхождений условных конструкций, см. нормированную гистограмму с накоплением (рис. 1, с. 127).

Из рис. 1 видно, что носители предпочитают разные стратегии. Например, носители с кодами Н1 и Н2 чаще других используют горномарийские союзы. В речи носителя Н3 часто встречаются горномарийские союзы и конструкции с дублированием, в отличие от русских союзов, а в речи носителя Н4 — горномарийские и русские союзы, в отличие от конструкций с дублированием. Носитель Н5 использует все три стратегии примерно с одинаковой частотой. Носители Н6 и Н7 в основном используют русские союзы.

Рассмотрим некоторые социолингвистические параметры, представленные в табл. 3 (носители упорядочены по году рождения).

Корреляции с годом рождения в наших данных не наблюдается. Самый молодой носитель в этой выборке (1980 г. р.) Н6 действительно часто употребляет русские союзы; однако это свойственно и носителю Н7 1948 г. р. При этом носитель Н1 1974 г. р. практически не использует русские союзы в своей речи.

Уровень образования также не влияет на выбор союза: все носители в выборке имеют высшее образование (за исключением носителей Н2 и Н5 со средним специальным

¹¹ В [Riese 1984] цитируются примеры из таких работ, как [Ramstedt 1902; Wichmann 1916; Beke 1938].

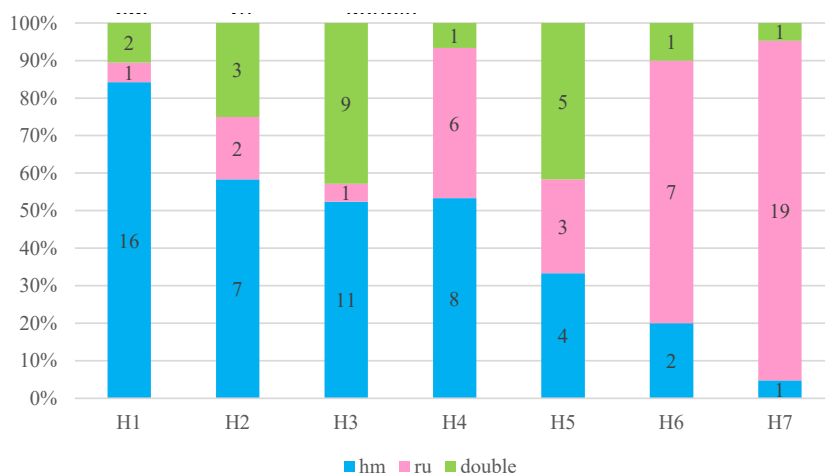


Рис. 1. Количество условных клауз в речи конкретных носителей: горномарийский союз (hm), русский союз (ru), горномарийский и русский союз в одной клаузе (double)

образованием). Носитель H4 с филологическим образованием часто использует смешанную стратегию, а не только горномарийские союзы, как могло бы ожидатьс.

Места, в которых носители проживали ранее, также однозначно не определяют использование рассмотренных стратегий. С одной стороны, превалирование русских союзов в речи носителя H7 (1948 г. р.) может объясняться тем, что он провел много времени за пределами Республики Марий Эл и в повседневной речи часто использует русский язык. С другой стороны, носитель H1 (1974 г. р.), который также выезжал за пределы республики, предпочитает горномарийские союзы.

Вместе с тем (двух) носителей, предпочитающих русские союзы (стратегия ru), объединяет более частое использование русского языка в повседневном общении. Относительно прочих носителей с другими стратегиями однозначных выводов сделать не представляется возможным.

Таблица 3

Социолингвистическая информация о носителях и используемые ими стратегии в выборе союзов

Код	Год рождения	Длительные выезды	Язык повседневного общения	Стратегия
H6	1980	нет	русский	ru
H1	1974	да	горномарийский	hm
H4	1968	нет	горномарийский	hm / double
H2	1963	нет	горномарийский / русский	hm
H5	1958	нет	горномарийский / русский	ru / hm / double
H3	1953	нет	горномарийский	hm / ru
H7	1948	да	русский / горномарийский	ru

Русский союз *хотя* был употреблен 6 раз. Он может располагаться на абсолютной левой периферии (4 примера в корпусе), во второй позиции (1 случай в корпусе), см. пример (31), который иллюстрирует обе эти стратегии, а также на правой периферии (2 примера в корпусе), см. (32).

- (31) *xot'a škál-et ál-ân. — škál-em xot'a ál-ân.*
 хотя корова-POSS.2SG быть-PRÉT корова-POSS.1SG хотя быть-PRÉT
 [— Детей молоком не поили, поили кислым молоком и всё, больше ничего нет...]
 '— Хотя у тебя корова была. — Хотя у меня корова была'. [ГК: «Диалог о жизни», 33–34]
- (32) *kəzāt jažo xot'a: karem-əštə šudə-m kačk-āt.*
 сейчас хороший хотя овраг-IN трава-ACC есть-NPST.3PL
 [У кого какая скотина есть, за всеми вот ухаживаешь: запираешь, кормишь.] 'Сейчас хорошо хотя: в овраге едят траву (что, зеленую траву)'. [ГК: «Работа», 58]

Несмотря на аналогичное структурное устройство условных и уступительных конструкций, в корпусе не встретились «смешанные» случаи, которые содержали бы оба союза. Впрочем, это может объясняться меньшим количеством уступительных конструкций (17 примеров) по сравнению с условными (174 примера). Однако при этом именно в уступительных конструкциях встретился порядок с русским союзом на последнем месте (32). В условных конструкциях союз *если* во всех примерах располагался на левой периферии. Вместе с тем, согласно [Кашкин 2023а: 4], в русской речи носителей горномарийского языка в условных конструкциях был отмечен порядок с постпозитивным *если*: *Он немножко начитанный ребенок если, он правильно составит и предложение и всё*.

4. Обсуждение: стратегия с дублированием

Как мы показали в разделе 3, почти 50 % условных конструкций в ГК включают русский союз (29,7 % — конструкции с русским союзом вместо горномарийского, 20,6 % — с дублированием горномарийского союза). «Смешанная» стратегия оказывается доступной, поскольку между стратегиями оформления условных клауз в горномарийском и русском языке наблюдается конгруэнтность, и частотной, поскольку эта конгруэнтность является частичной.

Говоря о конгруэнтности, А. Айхенвальд отмечает, что «заимствование происходит гораздо чаще между структурно схожими системами, чем в других случаях. Если контактирующие языки имеют схожие конструкции, они усиливают друг друга» [Aikhenvald 2007: 32]¹². В качестве примера она приводит заимствование префиксов из языка карамоджонг в язык лабвор (нилотская семья) [Storch 2007], что возможно благодаря существующему слоту для префиксов.

Структура условных клауз в русском и горномарийском совпадает по двум параметрам: наличие союза и использование (в общем случае) финитных глагольных форм. Различается только позиция союза — препозитивная в русском («если» ____») и постпозитивная в горномарийском («____ 'если'»). В нашем случае русский союз *если* вставляется именно в «незанятый» слот, что отличается от примера А. Айхенвальд. Такие случаи можно также анализировать как случаи одновременного заимствования материала (MAT(ter)-borrowing) и модели (PAT(tern)-borrowing) в терминологии [Matras, Sakel (eds.) 2007], т. е. MAT&PAT заимствование по [Gardani 2020: 272].

Мы считаем, что полная конгруэнтность может, наоборот, блокировать употребление русских слов. Одним из аргументов может служить следующее. В горномарийском есть временной союз *kānam* 'когда'. Он является препозитивным и употребляется с финитными глагольными формами, см. (33). При этом в корпусе ни разу не встретился русский союз *когда*, предположительно из-за слишком большой конгруэнтности — конструкции совпадают с точностью до позиции союза.

¹² «Cross-linguistically, borrowing is much more frequent between structurally similar systems than otherwise. If languages in contact have similar constructions and patterns, they reinforce each other» [Aikhenvald 2007: 32].

- (33) *ti ajo-m m'edv'ed'ev-vlā-n semn'ä tumaj-en lākt-ān,*
 этот праздник-ACC Медведев-PL-GEN семья думать-СВВ извлекать-PRET
kānam m'edv'ed'ev pr'ez'ident āl-ān.
 когда Медведев президент быть-PRET
 'Этот праздник придумала семья Медведевых, когда Медведев был президентом'.
 [ГК: «День семьи», 2]

Механизм совмещения структур двух языков обсуждается в [Hicks 2012]. Подход, представленный в [Hicks 2012], основывается на теории автолексического синтаксиса [Sadock 1985; 1991], где различаются морфологический и синтаксический модули, см. рис. 2. Наличие двух отдельных модулей, изображенных на рис. 2 сверху и снизу, и является основной идеей, которая развивается в [Hicks 2012].

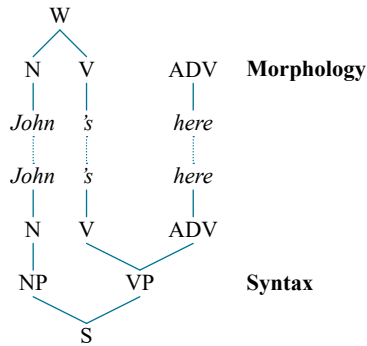


Рис. 2. Морфосинтаксический интерфейс в автолексическом синтаксисе [Sadock 1985: 385]

Рассмотрим анализ «смешанных» структур по [Hicks 2012] на примере. В (34) представлено английское предложение из речи японско-английского билингва, в котором предлог *for* 'для' продублирован японским послелогом *ni* с аналогичным значением. Для анализа таких «дублированных» элементов предлагается использовать подход с отдельным модулем для каждого языка. На рис. 3 (с. 130) изображена синтаксическая структура примера (34): верхняя часть представляет собой структуру английского, а нижняя — японского. В адложной группе есть один общий элемент — именная группа *Sean* 'Шон', благодаря этому общему элементу структуры языков могут совмещаться. Вокруг этого элемента есть два слота — слева для английского предлога, справа для японского послелога.

- (34) ...look at the things she buys **for** *Sean ni*.
 'Посмотри на вещи, которые она покупает для Шона'. [Nishimura 1986: 140]

В горномарийском языке также встречаются похожие на (34) примеры, см. (35). Однако они единичны (см. [Дьячков 2020: 279]); в (35) также представлены речевые сбои. Малочисленность таких примеров объясняется инконгруэнтностью структур: горномарийские падежные показатели присоединяются к основе (см. *kryl'cā-štā* [крыльцо-IN] 'на крыльце'), их присоединение к русской предложной группе затруднено, поскольку существительное выступает в различных падежных формах, в (35) — в предложном падеже (*на крыльчке*). Таким образом, русские конструкции с предлогами в подавляющем большинстве не присоединяют ни горномарийских падежных показателей, ни послелогов (существительное в послеложной конструкции в горномарийском языке стоит в номинативе), ср. (36a) и (36б) — «смешанной» конструкции *до войны jakte* мы в данном случае не ожидаем.

- (35) *tāxen' ma postol-āštā, [na kryl'ečkā]-štā, kryl'cā-štā, kryl'oc...*
 такой что подобно-IN на крыльчке-IN крыльцо-IN крыльцо-IN

nu, pört anzâl-nâ.

ну дом перед-IN2

‘Лежит на чем-то похожем, на крылечке, на крыльце, ну, перед домом’. [ГК: «Игры-7», 50–51]

- (36) а. *p'edučil'iš'e âl-ân, do войны tidâ cilâ*
педучилище быть-PRET этот весь

‘Было педучилище, до войны это все’. [ГК: «Диалог о проблемах», 68]

- б. *tâmansola značit âl-ân vojna jakte-ok eče*
Тюманово значит быть-PRET война до=EMPH еще

‘Деревня Тюманово, значит, была еще до войны’. [ГК: «Хозяйство», 97]

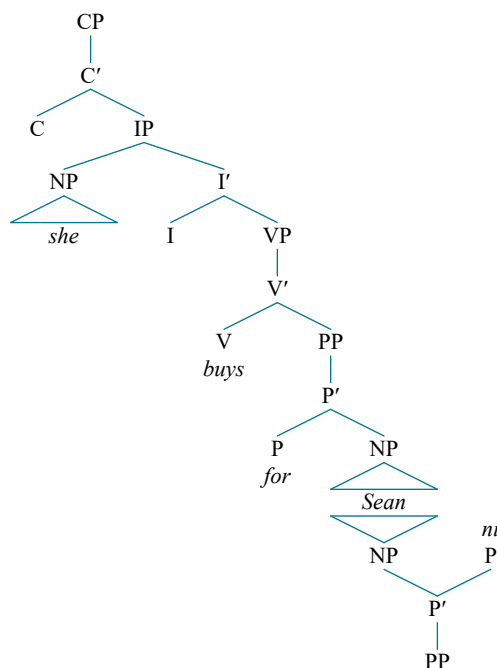


Рис. 3. Двойная структура предложения из речи англо-японского билингва с дублированным адлогом [Hicks 2012: 52]

Что касается союзов, то в случае инконгруэнтности ожидается, что их частотность будет не так высока. Например, в нанайском языке условные конструкции включают финитную глагольную форму и постпозитивный союз *oseni*, в то время как для временных конструкций используется специализированная нефинитная форма [Khomchenkova, Stoynova 2023: 11]. Русский союз более частотен в условных конструкциях (с частичной конгруэнтностью), чем во временных (с инконгруэнтностью), см. табл. 4 (с. 131) и примеры (37). Различия между двумя конструкциями статистически значимы (точный тест Фишера, $p = 0.0001$).

НАНАЙСКИЙ [Khomchenkova, Stoynova 2023: 11]

- (37) а. *jesli bəum-bə wā-ri*
если лось-ACC убивать-PRS

‘если они убьют лося...’

- б. *kogda* *ʒaŋg'ar-i-do-i*
 когда судить-PTCP.PRS-DAT-REFL.SG
 'когда [он] судит...'

В лесном энецком языке в условных конструкциях употребляется специализированное деепричастие (501 пример в корпусе). В 22 случаях (5,3 %) это деепричастие употребляется вместе с русским союзом *если* (38а), несмотря на инконгруэнтность. В еще меньшем количестве случаев — в 5 (1 %) — происходит изменение конструкции в соответствии с конструкцией русского языка («*если* + финитная форма глагола») (38б), см. [Khomchenkova, Stoynova 2023].

ЛЕСНОЙ ЭНЕЦКИЙ [Khomchenkova, Stoynova 2023: 21]

- (38) а. *jesli* *ŋa-za* *bɔa* *ɛ-bu-ta*
 если небо-NOM.SG.3SG плохой быть(IPFV)-CVB.COND-OBL.SG.3SG
 'если погода плохая'
- б. *jesli* *ɔsa-d* *kɔma-d*
 если мясо-DAT.SG хотеть(IPFV)-2SG.S
 'если хочешь мяса'

Это может также служить аргументом в пользу того, что русский союз *если* распространен в горномарийской речи (как в дублированном, так и в единичном употреблении), поскольку условные конструкции в этих двух языках изначально похожи и не требуют серьезной перестройки.

5. Заключение

Итак, на основе корпусных данных мы показали, что в горномарийской речи встречаются русские союзы *если* и *хотя* в условных и уступительных клаузах, наряду с горномарийскими союзами *gəŋ* 'а', *gəŋ'at* 'хотя', а также *əlgecə* 'COND'. Союзы *если* и *хотя* употребляются в большинстве случаев препозитивно; *если* также может дублироваться горномарийским союзом *gəŋ*.

Займствование этих союзов нельзя объяснить «заполнением лакун» ([Thompson et al. 2007: 268–269]) ввиду наличия исконных союзов. По-видимому, использование *если* и *хотя* обусловлено усилением влияния русского языка, ср. тенденцию к более частому употреблению русских союзов по сравнению с более ранними данными других исследователей.

Возможность сочетания двух союзов в одной клаузе и ее распространенность можно объяснить структурной конгруэнтностью между стратегиями оформления условных клауз в языке-реципиенте (горномарийском) и языке-доноре (русском). Русская условная конструкция «'если' ____» и горномарийская «____ 'если'» совмещаются за счет сходного устройства, заключающегося в наличии «общей» части — финитной клаузы. В результате получается конструкция с двумя слотами для условного союза: «'если' ____ 'если'». Соответственно, горномарийское предложение может включать русский союз (препозитивный), горномарийский союз (постпозитивный) либо оба союза в двух позициях. Мы предполагаем, что именно такая «неполная» конгруэнтность способствует использованию русских союзов.

По имеющимся данным можно также сделать предварительный вывод о том, что структурная конгруэнтность важнее, чем семантическая (*если* и *когда* соответствуют по значению постпозитивному союзу *gəŋ* 'если' и препозитивному *kənat* 'когда', но *если* в корпусе частотен, а *когда* не употреблен ни разу). Для верификации этого предположения в дальнейшем планируется подробный анализ прочих русских союзов, встречающихся в горномарийской речи.

Интересен также вопрос сравнения свойств конструкций с заимствованными союзами в разных языках (см. некоторые данные в [Khomchenkova, Stoynova 2023]). Первая

масштабная попытка была сделана в [Stolz, Levkovych 2022], однако в этой работе приводятся данные из грамматических описаний языков; корпусные данные не привлекаются. Проведение подробного анализа корпусных данных, подобного представленному в настоящей статье, для других союзов и других языков поможет лучше понять роль семантической и структурной (ин)конгруэнтности при заимствовании союзов с учетом социолингвистического статуса языка.

СПИСОК СОКРАЩЕНИЙ

1, 2, 3 — 1, 2, 3 лицо	INDEF1, INDEF2 — неопределенное местоимение
ACC — аккузатив	INF — инфинитив
ADD — аддитивная частица	IPFV — имперфектив
ADV — адвербиализатор	LAT — латив
ATT — аттенуатив	MED — медиопассив
ATTR — атрибутивизатор	NEG — отрицание
CAUS — каузатив	NMLZ — номинализация
CAUS.DIST — дистантный каузатив	NOM — номинатив
COND — маркер условной конструкции	NPST — непрошедшее время
CVB — дееспричастие	OBL — косвенный падеж
CVB.ANT — дееспричастие с семантикой предшествования	OPT — оптатив
DAT — датив	ORD — порядковое числительное
DENOM — отыменный глагол	PL — множественное число
DES — деизидератив	POSS — посессивность
DEST — показатель предназначения (дестинатив)	PRET — претерит
EL — элатив	PRS — настоящее время
EMPH — эмфатическая частица	PTCP — причастие
EXST — показатель с экзистенциальной семантикой	PTCP.ACT — активное причастие
FULL — полная форма	PTCP.PASS — пассивное причастие
GEN — генитив	REFL — рефлексив
ILL, ILL2 — иллатив	RETR1, RETR2 — показатель ретроспективного сдвига
IMP — императив	S — субъектное спряжение
IN, IN2 — инессив	SG — единственное число
	SIM — симилиатив

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Акцорин 1991 — Акцорин В. А. *Марийский фольклор. Мифы, легенды, предания*. Йошкар-Ола: Марийское книжное изд-во, 1991. [Aktorsin V. A. *Mariiskii fol'klor. Mify, legendy, predaniya* [Mari folklore. Myths, legends, traditions]. Yoshkar-Ola: Mari Book Publ., 1991.]
- Вайнрайх 1979 — Вайнрайх У. *Языковые контакты. Состояние и проблемы исследования*. Жлуктенко Ю. А. (пер. с англ.). Киев: Вища школа, 1979. [Weinreich U. *Yazykovye kontakty. Sostoyaniye i problemy issledovaniya* [Languages in contact: Findings and problems]. Zhluktenko Yu. A. (trans. from English). Kyiv: Vishcha shkola, 1979.]
- Гаврилова 2013 — Гаврилова В. Г. Марийско-русское переключение и смешение кодов. *Вестник Удмуртского университета. Серия «История и филология»*, 2013, 2: 16–22. [Gavrilova V. G. Mari-Russian code-switching and code-mixing. *Bulletin of Udmurt University. History and Philology Series*, 2013, 2: 16–22.]
- Галкин 1964 — Галкин И. С. *Историческая грамматика марийского языка. Морфология. Ч. I*. Йошкар-Ола: Марийское книжное изд-во, 1964. [Galkin I. S. *Istoricheskaya grammatika mariiskogo yazyka. Morfologiya* [Mari historical grammar. Morphology]. Part I. Yoshkar-Ola: Mari Book Publ., 1964.]
- Гусев 2020 — Гусев В. Ю. О камасинском постпозитивном условно-временном союзе *dak*. *Труды Института русского языка им. В. В. Виноградова*, 2020, 4: 76–89. [Gusev V. Yu. On the postpositive conditional-temporal conjunction *dak* in Kamas. *Proc. of the V. V. Vinogradov Russian Language Institute*, 2020, 4: 76–89.]

- Дьячков 2020 — Дьячков В. В. Структурные и социолингвистические характеристики горномарийско-русского переключения кодов: пилотное исследование. *Труды института русского языка им. В. В. Виноградова*, 2020, 4: 273–288. [Dyachkov V. V. Structural and sociolinguistic features of Hill Mari — Russian code switching: A pilot study. *Proc. of the V. V. Vinogradov Russian Language Institute*, 2020, 4: 273–288.]
- Иванов, Тужаров 1970 — Иванов И. Г., Тужаров Г. М. *Северо-западное наречие марийского языка*. Йошкар-Ола: Марийское книжное изд-во, 1970. [Ivanov I. G., Tuzharov G. M. *Severo-zapadnoe narechie mariiskogo yazyka* [Northwestern dialect of Mari]. Yoshkar-Ola: Mari Book Publ., 1970.]
- Кашкин 2023а — Кашкин Е. В. Общие сведения о горномарийском языке. Социолингвистическая ситуация. *Элементы горномарийского языка в типологическом освещении*. Кашкин Е. В. (отв. ред.). М.: Буки Веди, 2023, 1–5. [Kashkin E. V. General information about Hill Mari. Sociolinguistic situation. *Elementy gornomariiskogo yazyka v tipologicheskom osveshchenii*. Kashkin E. V. (ed.). Moscow: Buki Vedi, 2023, 1–5.]
- Кашкин 2023б — Кашкин Е. В. Фонология. *Элементы горномарийского языка в типологическом освещении*. Кашкин Е. В. (отв. ред.). М.: Буки Веди, 2023, 6–11. [Kashkin E. V. General information about Hill Mari. Sociolinguistic situation. *Elementy gornomariiskogo yazyka v tipologicheskom osveshchenii*. Kashkin E. V. (ed.). Moscow: Buki Vedi, 2023, 6–11.]
- Майтинская 1982 — Майтинская К. Е. *Служебные слова в финно-угорских языках*. М.: Наука, 1982. [Maytinskaya K. E. *Sluzhebnye slova v finno-ugorskikh yazykakh* [Function words in Finno-Ugric languages]. Moscow: Nauka, 1982.]
- Мордашова 2023а — Мордашова Д. Д. Ретроспективный сдвиг. *Элементы горномарийского языка в типологическом освещении*. Кашкин Е. В. (отв. ред.). М.: Буки Веди, 2023, 336–349. [Mordashova D. D. Retrospective shift. *Elementy gornomariiskogo yazyka v tipologicheskom osveshchenii*. Kashkin E. V. (eds.). Moscow: Buki Vedi, 2023, 336–349.]
- Мордашова 2023б — Мордашова Д. Д. Средства выражения модальности. *Элементы горномарийского языка в типологическом освещении*. Кашкин Е. В. (отв. ред.). М.: Буки Веди, 2023, 350–398. [Mordashova D. D. Means of expressing modality. *Elementy gornomariiskogo yazyka v tipologicheskom osveshchenii*. Kashkin E. V. (ed.). Moscow: Buki Vedi, 2023, 350–398.]
- Муравьев 2023 — Муравьев Н. А. Обстоятельственные конструкции с дееспричастиями на *-mâkâ* и *-meškâ*. *Элементы горномарийского языка в типологическом освещении*. Кашкин Е. В. (отв. ред.). М.: Буки Веди, 2023, 775–794. [Muravyev N. A. Adverbial constructions with participles ending in *-mâkâ* and *-meškâ*. *Elementy gornomariiskogo yazyka v tipologicheskom osveshchenii*. Kashkin E. V. (ed.). Moscow: Buki Vedi, 2023, 775–794.]
- Пекелис 2015 — Пекелис О. Е. Показатель *то* как средство акцентуации имплицативного отношения (на примере союза *если... то*). *Вопросы языкознания*, 2015, 2: 55–96. [Pekelis O. E. The marker *to* as a means for highlighting implicative relations (case study of *esli... to* 'if... then' conjunction)]. *Voprosy Jazykoznanija*, 2015, 2: 55–96.]
- Саватеева 2005 — Саватеева Г. А. *Лексические особенности правобережных говоров горномарийского языка*. Дис. ... канд. филол. наук. Йошкар-Ола: Марийский государственный ун-т, 2005. [Savateeva G. A. *Leksicheskie osobennosti pravoberezhnykh govorov gornomariiskogo yazyka* [Lexical features of the right-bank dialects of Hill Mari]. Candidate diss. Yoshkar-Ola: Mari State Univ., 2005.]
- Саваткова 1969 — Саваткова А. А. *Русские заимствования в марийском языке*. Йошкар-Ола: Марийское книжное изд-во, 1969. [Savatkova A. A. *Russkie zaïmstvovaniya v mariiskom yazyke* [Russian borrowings in Mari]. Yoshkar-Ola: Mari Book Publ., 1969.]
- Саваткова 2002 — Саваткова А. А. *Горное наречие марийского языка*. Szombathely: Berzsényi Dániel Főiskola, 2002. [Savatkova A. A. *Gornoe narechie mariiskogo yazyka* [Hill dialect of Mari]. Szombathely: Berzsényi Dániel Főiskola, 2002.]
- Саваткова 2008 — Саваткова А. А. *Словарь горномарийского языка*. Йошкар-Ола: Марийское книжное изд-во, 2008. [Savatkova A. A. *Slovar' gornomariiskogo yazyka* [Hill Mari dictionary]. Yoshkar-Ola: Mari Book Publ., 2008.]
- Сибатрова 2016 — Сибатрова С. С. Русские заимствования в системе союзов марийского языка. *Ежегодник финно-угорских исследований*, 2016, 10(3): 30–41. [Sibatrova S. S. Russian borrowings in the system of conjunctions of Mari. *Yearbook of Finno-Ugric Studies*, 2016, 10(3): 30–41.]
- Храковский 2004 — Храковский В. С. Уступительные конструкции: семантика, синтаксис, типология. *Типология уступительных конструкций*. Храковский В. С. (ред.). СПб.: Наука, 2004, 10–91. [Xrakovskij V. S. Concessive constructions: Semantics, syntax, typology. *Tipologiya ustupitel'nykh konstruksii*. Xrakovskij V. S. (ed.). St. Petersburg: Nauka, 2004, 10–91.]
- Шабыков и др. 2020 — Шабыков В. И., Кудрявцева Р. А., Зорина З. Г. Статус горномарийского языка в современном обществе. *Социодинамика*, 2020, 4: 74–84. [Shabykov V. I., Kudryavtseva R. A., Zorina Z. G. Status of Hill Mari in modern society. *Sociodynamics*, 2020, 4: 74–84.]

- Aikhenvald 2007 — Aikhenvald A. Y. *Grammars in contact: A cross-linguistic perspective. Grammars in contact: A cross-linguistic typology*. Aikhenvald A. Y., Dixon R. M. W. (eds.). Oxford: Oxford Univ. Press, 2007, 1–66.
- Alhoniemi 1993 — Alhoniemi A. *Grammatik des Tscheremissischen (Mari)*. Hamburg: Helmut Buske Verlag, 1993.
- Beke 1938 — Beke Ö. *Tscheremissische Märchen, Sagen und Erzählungen*. Helsinki: Suomalais-Ugrilainen Seura, 1938.
- Besters-Dilger et al. (eds.) 2014 — Besters-Dilger J., Dermarkar C., Pfänder St., Rabus A. *Congruence in contact-induced language change: Language families, typological resemblance, and perceived similarity*. Berlin: Mouton de Gruyter, 2014.
- Gardani 2020 — Gardani F. Borrowing matter and pattern in morphology. An overview. *Morphology*, 2020, 30: 263–282.
- Grant 2012 — Grant A. Contact, convergence, and conjunctions: a cross-linguistic study of borrowing correlations among certain kinds of discourse, phasal adverbial, and dependent clause markers. *Dynamics of contact-induced language change*. Chamoreau C., Léglise I. (eds.). New York: Mouton de Gruyter, 2012, 311–358.
- Harris, Campbell 1995 — Harris A., Campbell L. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge Univ. Press, 1995.
- Haspelmath, König 1998 — Haspelmath M., König E. Concessive conditionals in the languages of Europe. *Adverbial constructions in the languages of Europe*. van der Auwera J. (ed.). Berlin: Mouton de Gruyter, 1998, 563–640.
- Hicks 2012 — Hicks C. A dual-structure analysis of morphosyntactic doubling in code switching. *Studies in the Linguistic Sciences: Illinois Working Papers*, 2012, 37: 44–57.
- Khomchenkova, Stoynova 2023 — Khomchenkova I., Stoynova N. Adverbial clauses with Russian conjunctions in three languages with different subordination strategies. *Talk at the 56th Annual Meeting of the Societas Linguistica Europaea (SLE, 29.08–01.09.2023, Athens, Greece)*.
- König 1991 — König E. *The meaning of focus particles: A comparative perspective*. London: Routledge, 1991.
- Matras, Sakel (eds.) 2007 — Matras Y., Sakel J. (eds.). *Grammatical borrowing in cross-linguistic perspective*. Berlin, New York: Mouton de Gruyter, 2007.
- Nishimura 1986 — Nishimura M. Intrasentential code-switching: The case of language assignment. *Language Processing in Bilinguals: Psycholinguistic and neuropsychological perspectives*. Vaid J. (ed.). Hillsdale: Lawrence Erlbaum Associates, 1986, 123–144.
- Ramstedt 1902 — Ramstedt G. J. *Bergtscheremissische Sprachstudien*. Helsinki: Suomalais-Ugrilainen Seura, 1902.
- Riese 1984 — Riese T. *The conditional sentence in the Ugrian, Permian and Volgaic languages*. Wien: Studia Uralica, 1984.
- Sadock 1985 — Sadock J. M. Autolexical syntax: A proposal for the treatment of noun incorporation and similar phenomena. *Natural Language and Linguistic Theory*, 1985, 3: 379–439.
- Sadock 1991 — Sadock J. M. *Autolexical Syntax: A theory of parallel grammatical representations*. Chicago: Univ. of Chicago Press, 1991.
- Sebba 2009 — Sebba M. On the notions of congruence and convergence in code-switching. *The Cambridge handbook of linguistic code-switching*. Bullock B. E., Toribio A. (eds.). Cambridge: Cambridge Univ. Press, 2009, 40–57.
- Stolz, Levkovych 2022 — On loan conjunctions: A comparative study with special focus on the languages of the former Soviet Union. *Susceptibility vs. resistance: Case studies on different structural categories in language-contact situations*. Levkovych N. (ed.). Berlin; Boston: Walter de Gruyter, 2022, 259–392.
- Storch 2007 — Storch A. How long do linguistic areas last? Western Nilotic grammars in contact. *Grammars in contact: A cross-linguistic typology*. Aikhenvald A. Y., Dixon R. M. W. (eds.). Oxford: Oxford Univ. Press, 2007, 94–113.
- Thompson et al. 2007 — Thompson S. A., Longacre R. E., Hwang Shin Ja J. Adverbial clauses. *Language typology and syntactic description. Vol. 2*. Shopen T. (ed.). Cambridge: Cambridge Univ. Press, 2007, 237–299.
- Weinreich 1953 — Weinreich U. *Languages in contact: Findings and problems*. New York: Linguistic Circle of New York, 1953.
- Wichmann 1916 — Wichmann Y. *Syrjänische Volksdichtung*. Helsinki: Suomalais-Ugrilainen Seura, 1916.

William Croft's approach to constructions in comparison with the system of constructions in the Russian Constructicon

© 2024

Valentina A. Zhukova

UiT The Arctic University of Norway, Tromsø, Norway; valentina.a.zhukova@uit.no

Alexander B. Letuchiy

HSE University, Moscow, Russia; Vinogradov Russian Language Institute,
Russian Academy of Sciences, Moscow, Russia; alexander.letuchiy@gmail.com

Abstract: In this article, we discuss the approach to description of constructions adopted in William Croft's monograph "Morphosyntax. Constructions of the World's languages" (2022) and compare it to the approach to description and classification that is used in the Russian Constructicon. We conclude that Croft's system and the Russian Constructicon show several substantial differences. First, Croft's approach is based on the notion of Comparative Concepts and on an a priori classification of grammatical domains. By contrast, in the Russian Constructicon, a bottom-up approach is taken: it presupposes collecting the most representative inventory of constructions and, then, creating a system for classifying them. Second, the Russian Constructicon represents properties of constructions as a system of tags, where a construction can bear multiple tags, while Croft does not discuss cases of multiple tagging or grammatical class intersection. Finally, Croft's system focuses on the core of grammar and includes mainly those values that are grammaticalized, while in the Russian Constructicon, attention is given not only to grammatical constructions, but also to constructions that can be termed 'quasi-grammatical' or 'lexicalized' — they have narrow semantics and combinational properties (here belong, for instance, iterative/frequentative constructions, such as *to i delo* 'frequently', and constructions with the terminative/resultative meaning, such as *svoë otguljal* '[he] is done with having fun').

Keywords: survey, Construction Grammar, constructions, Russian Constructicon

Acknowledgements: We are grateful to Ekaterina V. Rakhilina, Laura A. Janda and other colleagues from the Russian Constructicon group for their valuable help and discussion of the text. We also thank the anonymous reviewers for valuable remarks and questions.

For citation: Zhukova V. A., Letuchiy A. B. William Croft's approach to constructions in comparison with the system of constructions in the Russian Constructicon. *Voprosy Jazykoznanija*, 2024, 5: 135–153.

DOI: 10.31857/0373-658X.2024.5.135-153

Подход Уильяма Крофта к описанию конструкций в сравнении с системой представления конструкций в Русском конструктиконе

Валентина Александровна Жукова

Университет Тромсё — Арктический университет Норвегии, Тромсё, Норвегия
valentina.a.zhukova@uit.no

Александр Борисович Летучий

Национальный исследовательский университет «Высшая школа экономики», Москва, Россия;
Институт русского языка им. В. В. Виноградова РАН, Москва, Россия; alexander.letuchiy@gmail.com

Аннотация: В статье мы обсуждаем подход к описанию конструкций, принятый в монографии Уильяма Крофта “Morphosyntax. Constructions of the World’s languages” (2022) и сопоставляем его с подходом к описанию и классификации конструкций, применяемым в “Русском конструктиконе”. Сравнение позволяет нам прийти к выводу, что между двумя методами описания есть существенные различия. Во-первых, подход Крофта опирается на «сравнительные категории» как основу типологических исследований и априорную классификацию значений и конструкций. Напротив, в «Русском конструктиконе» описание ведется «снизу вверх»: сначала составляется как можно более представительный инвентарь конструкций, а уже потом создается система для их классификации. Во-вторых, в «Русском конструктиконе» свойства конструкций представлены в виде системы помет, которых у каждой конструкции может быть несколько, тогда как в описании Крофта не обсуждается возможность принадлежности конструкции сразу к нескольким классам. Наконец, система Крофта ориентирована в первую очередь на ядерные зоны грамматики и включает в основном значения, грамматикализованные в тех или иных языках, тогда как в «Русском конструктиконе» внимание уделяется не только грамматическим, но и «квазиграмматическим» и «лексикализованным» конструкциям — данные конструкции обладают узкой семантикой и сочетаемостью (к ним, в частности, относится класс конструкций повторяемости, таких как *то и дело*, или конструкция со словом *своё* и глаголом с приставкой *от-*, например, *своё отгулял*, обозначающая окончание некоторого занятия и невозможность его продолжения).

Ключевые слова: обзор, грамматика конструкций, конструкции, русский конструктикон

Благодарности: Мы благодарим Екатерину Владимировну Рахилину, Лору Янду и других коллег из коллектива «Русского конструктикона» за ценные советы и обсуждение текста. Кроме этого, мы благодарны анонимным рецензентам за важные замечания и вопросы.

Для цитирования: Zhukova V. A., Letuchiy A. B. William Croft’s approach to constructions in comparison with the system of constructions in the Russian Constructicon. *Voprosy Jazykoznanija*, 2024, 5: 135–153.

DOI: 10.31857/0373-658X.2024.5.135-153

1. Introduction

The monograph by William Croft “Morphosyntax. Constructions of the World’s languages” [2022] describes various language phenomena from a constructivist perspective. The book takes on an ambitious task to discuss very different phenomena, their function, form, and semantics in a typologically-oriented manner.

The monograph has a twofold aim that combines language theory and linguistic typology. The author seeks to review the main domains of grammar, to describe questions, ambiguities, and trouble spots they pose for a researcher and, at the same time, to show the variance of formal organization in the given domain across languages.

A comprehensive understanding of a language is impossible without discovering the language-specific patterns of its syntax, morphology, and semantics. The idea of understanding a language on its own terms dates back to the structuralist tradition (cf. [Boas 1911]). A reasonable question, however, is whether it is possible to create a cross-linguistic classification that would be universally applicable and meaningful. William Croft [2022] has made an attempt to provide such a universal inventory of Comparative Concepts (CCs) that could be employed to describe the morphosyntax of any language and, therefore, would provide grounds for cross-linguistic comparison of constructions in the world’s languages. He takes a Comparative Concept to mean “a concept that can be used to compare the morphosyntactic structure of different languages” [Ibid.: 676].

In this paper, we describe the general principles adopted in the book and its specific features and compare its approach with the practically-oriented framework used in the Russian Constructicon.¹

¹ The resource is available online at <https://constructicon.github.io/russian/about/>.

2. Organization and characteristic features of the book

The book covers both well-described topics that have become the subject of descriptive, theoretical, and typological studies and problems and areas of grammar that are rarely considered to be relevant for this domain. The former include complex clauses and voice alternations, while among the latter there are speech act constructions (their relation to modality is often dealt with in discussions of grammatical semantics but not of morphosyntax), as well as nonpredicational clauses, as in (1) and stative complex predicates, as in (2).

- (1) Tiwi (isolate, Australia)
Purukupali ma.intina
 Purukuparli boss
 ‘Purukuparli is boss.’ [Croft 2022: 295]
- (2) *We painted the door red.* [Ibid.: 432].

The language phenomena and semantic domains analyzed in Croft’s book are divided into several groups that correspond to the parts of the monograph.

Part II “Argument phrase structure: reference and modification” describes mechanisms used to refer to an entity and to its specific properties: referential accessibility; specificity; types of modification; classifiers; anaphora. In Part III “Clause structure: predication and arguments” the author focuses on phenomena that are usually regarded as simple clause phenomena: voice and transitivity; argument coding and valency; nonverbal predication; complex predicates. Among topics that are rarely included in the discussion of simple clauses are speech act typology (exclamatives, hortatives, and so on), salience of the participants of the situation, and information packaging. Finally, Part IV “Complex sentences” deals with not only canonical problems related to complex sentences and clause combining (syntactic typology of clauses: relative, complement, and adjunct clauses; the opposition of balanced and deranked clauses), but also topics regarding semantic and discourse relations in complex sentences (reference tracking, zero expression of arguments).

Apart from these phenomena-oriented descriptions, the monograph contains an introductory Part I consisting of two chapters: the first one outlines the structure of the book and the author’s approach to cross-linguistic comparison, and the second one discusses the constructivist orientation of the book.

A special feature of the book is the absence of a clear boundary between main levels/areas of language, such as semantics, syntax, morphology, and discourse. As a consequence, topics generally believed to be mainly syntactic (transitivity, clause combining) are analyzed side by side with semantic (motion events, stative verbs) and discourse phenomena (reference, salience of participants). This approach has its strengths and weaknesses. On the one hand, it allows the author to demonstrate links between the various levels of language, usually identified only for individual language categories (cf., e.g., studies by Hopper and Thompson [1980] and Tsunoda [1985], which present semantic grounds of syntactic transitivity opposition). However, this approach makes the monograph heterogeneous: for instance, the rather narrow issue of stative predicates is given a separate chapter, while another chapter focuses on the complement clause typology, which is a much broader topic.

The reader finds some lacunas in the set of topics. For example, the monograph does not contain a detailed discussion of tense and aspect typology: the grammaticalization of tense, aspect, modality, and polarity is only touched upon in sections about complex sentences and complex predicates. Similarly, categories like nominal case and number are not analyzed in detail. The reason seems to be that the author, due to the constructivist approach, adopts (implicitly and, to some extent, also explicitly) the following hierarchy of linguistic description:

- 1) the semantic / functional domain that the phenomenon represents;
- 2) its impact on the syntactic relations / the whole construction;
- 3) properties of elements of the constructions.

To illustrate, the domain of voice is given much attention because voice meanings affect the whole sentence, including information structure types, discourse relevance of participants, and syntax of the clause / sentence (2) above). By contrast, case is a category marking a dependent element, but it does not affect the features of the head. This is why case is, so to say, “not very morphosyntactic” (that is, not directly related to the syntax of the whole constituent) and is peripheral to the approach adopted in the monograph.

This choice, resulting from the perspective taken by the author, is not free from disadvantages: some phenomena that could be relevant for the syntax of languages are either not analyzed or scattered across different parts of the book. For instance, case has been shown (e.g., by Wierzbicka [1980] and Kittilä [2002]) to be relevant for categories of the whole sentence, such as transitivity — which, in turn, have to do with constructions and information structure — however, in Croft’s monograph, case, argument structure, and information structure are considered in different chapters, and their relations are not explicitly mentioned. This is why, in some cases, it would have been useful (especially for readers specialized in typology of grammatical categories) to emphasize in each chapter relations between close phenomena or phenomena with similar functions, including head- and dependent-related categories.

Some descriptions of specific phenomena also call for questions and comments, for example, Chapter 6 on transitivity and argument structure. The point of view that transitivity should be described as a prototype and a set of peripheral phenomena that do not fully correspond to the prototypical description is adopted by many linguists. However, many current studies also seek to answer another question: why is it the transitive construction (and not, for instance, the one with a dative object) that is so important for our description of the syntactic structure? It seems that the answer to this question requires a plurifunctional approach that takes into account the (i) semantics of events (e.g., breaking / change of state) and of participants (see the treatment of animate / inanimate participants by Hopper and Thompson [1980] and Tsunoda [1985]); (ii) formal properties of the transitive construction; (iii) behavior of the transitive construction, including available voice alternations. In other words, Croft’s chapter on transitivity and argument structure lacks some theoretical depth that is needed to describe the class of transitive verbs and their role in the grammatical theory.

The last point, the properties of constructions and / or their elements, is not always given sufficient attention throughout the book. A case in point is Chapter 10 “Nonprototypical predication and nonpredicational clauses”. Since the type of clauses under analysis lacks an explicit predicate form, an *a priori* classification of nonpredicational clauses is problematic. At the same time, Roy [2013] and Testelefs [2008] show that some aspects of syntactic behavior differentiate predicational, presentational,thetic, and other types of nonpredicational clauses. Croft, however, mainly takes into account semantic and contextual features and suggests a classification mainly based on (im)possibility of a particular interpretation in the given context. Of course, it is impossible to consider the whole range of syntactic tests used for delimitation of types in one study. However, if some examples of syntactic contrasts were given, this would have made the classification of nonpredicational clauses more empirically founded.

A notable feature of the book is the wide use of quantitative data, see the statistics on word order in Chapter 6 “Event structure and argument coding: semantics, transitivity, and alignment” and Table 8.4 on argument expression in Huallaga Quechua in Chapter 8 “Argument coding and voice: discourse factors”. On the one hand, the prominence of statistics gives the language data a new perspective and distinguishes the book from many previous typologically-oriented descriptions. On the other hand, it seems that the statistical data are not always accompanied by an explanation.

Another important feature of Croft's monograph is that the domains of grammar are defined on different grounds. While some of them are defined semantically (for instance, Chapter 15 "Temporal and causal relations between events" illustrates various means for expression of the same semantic and pragmatic relations, see also Chapter 12 "Speech act constructions" and Chapter 14 "Stative complex predicates, including manner"), others are based on formal grounds (in Chapter 10 "Nonprototypical predication" the absence of a verb or another type of predicate serves as the sole defining factor). This asymmetry is, to a large extent, an integral part of today's syntactic and semantic typology. It results from the core-periphery structure that current theory assigns to grammar. For example, the core (canonical) sentence has a verb / predicate, and this is why the presence of the predicate goes without a comment; verbal sentences are semantically and syntactically classified into stative and dynamic, transitive and intransitive, and so on. By contrast, the absence of a predicate makes a clause / sentence noncanonical: this formal feature serves as the basis for further description of semantic or pragmatic distinctions, e.g., between 'possession', 'qualification', 'identification'. Roughly speaking, when we classify grammatical phenomena in various frameworks (including Croft's), we tend to classify them by the most visible and most unusual properties, even if these properties belong to different aspects of grammar and levels of language (semantics, syntax, illocutionary force, and so on).

We turn next to the concepts and principles proposed by the author and assess their strengths and weaknesses.

3. Constructivist orientation

A key feature of the book, as mentioned above, is its constructivist orientation. This perspective makes the task carried out by the author even more ambitious because constructions are more complicated to compare across languages and to distinguish from each other than, for example, grammatical markers.

It should be noted that Croft regards the construction as a basic concept of language organization and a basic instrument for language description. One of the key assumptions he puts forward in the introductory part is that "the proper unit for grammatical analysis is a (morphosyntactic) construction, such as the numeral modification construction" [Croft 2022: 4–5]. Word classes are also defined on constructivist grounds: the author states that "word classes are defined not by their semantics but by their occurrence in constructions — more precisely, a word's occurrence in a particular role in a construction" [Ibid.: 8]. This strength of the constructivist approach is that Croft emphasizes the importance of empirical analysis of basic elements of grammar (such as words and word classes) on the basis of contexts they are compatible with, instead of postulating an a priori classification or using vague semantic formulations, such as 'adjectives denote properties'. At the same time, this broad notion of a construction as a basic element of grammar does not provide results significantly different from notions of 'verb class' or 'syntactic relation' of standard grammatical typology. For instance, instead of saying that an adjective is a word that occurs as a modifier in a noun-modifier construction and in the predicate position in a copula construction, we can formulate it in a standard syntactic way: "An adjective can be a modifier of a noun or pronoun or the second argument of a copula verb". In this sense, Croft's construction is an instrument of description which does not provide results crucially different from those provided by formal or functional syntactic analysis.

Another possible problem of Croft's type of constructional analysis, is the lack of clear boundaries between different constructions. To illustrate, in Modern Standard Arabic, the numeral *wāḥidun* '1' agrees in case and gender with the noun, the numeral *ṭalāṭun* '3' subcategorizes for genitive case and takes the gender opposite to the gender of the noun, and *kḥamsūna* '50' does not have gender and requires the accusative case form (see [Grande 2001; Aoun et al. 2010] for a detailed discussion). Does this mean that '1' belongs to one word class and '3' and

‘50’ to another one (because the latter, but not the former, assign case to the noun) or that all three numerals are of different classes?

As mentioned above, Croft [2022] elaborates on the idea of comparative concepts and provides an arguably exhaustive list of them based on the typological analysis of constructions on the level from phrases to clauses and complex sentences. Let us repeat Croft’s definition: a comparative concept is “a concept that can be used to compare the morphosyntactic structure of different languages” [Ibid.: 676]. The taxonomy of Croft’s CCs includes several hundred CCs that are divided into four types: **construction** (*cxn*), **semantic content** (*sem*), **information packaging** (*inf*), and **strategy** (*str*). This four-dimensional distinction corresponds to the basic structure of a construction proposed by Croft. The notion of comparative concepts, i.e. theoretical constructs for typological comparisons, was introduced by Haspelmath [2010], who suggested differentiating COMPARATIVE CONCEPTS from language-specific DESCRIPTIVE CATEGORIES. The reason behind this differentiation is that “crosslinguistic comparison of morphosyntactic patterns cannot be based on formal patterns (because these are too diverse), but has to be based on universal conceptual-semantic concepts” [Ibid.: 665]. This idea is not new and follows the Greenbergian tradition of typological research, cf. [Greenberg 1963; Comrie 1989; Croft 2003].

Information packaging in Croft’s terms is contrasted to strategy. Strategy refers to the set of specific means used to convey a given meaning. By contrast, packaging refers to the role of information in discourse. As Croft [2022: 13] puts it, “any semantic content can be packaged in any information packaging function”. The main packaging functions are reference, modification, and predication. For instance, the property ‘tall’ can be predicated in *That tree is tall*, used to modify a referent in *a tall tree* and used as a referent in *the tallness of the tree*.

According to Croft [2022: 5], each construction is a combination of form and function. The form refers to the morphosyntax of the construction (i.e., strategy), while the function is further subdivided into semantic content (i.e., meaning) and information packaging, suggesting that any meaning can be packaged differently, see Figure below. Thus, four types of CCs correspond to the three end nodes (*inf*, *sem*, and *str*) and the root node (*cxn*) in the structure.

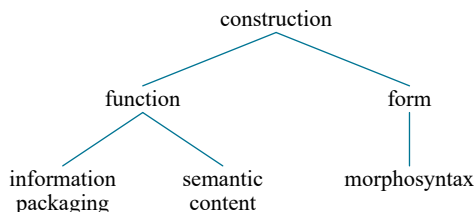


Fig. The basic structure of a construction [Croft 2022: 5]

Croft’s broad notion of a construction calls for a comparison with other, more practically oriented approaches to constructions. Recently, linguists’ interest in constructions has acquired a new, applied dimension, namely, the creation of constructicons (databases of constructions that contain their semantic and structural description, as well as annotation of relations between them, both formal and semantic). This is why the monograph under analysis is also important as a potential starting point for creating an inventory of grammatical constructions (as well as grammatical markers) and compiling new constructicon resources.

Constructicons are created for a limited number of languages, among others, for English, German, Japanese, Swedish, Brazilian Portuguese, and Russian. While the majority of constructicons are based on corresponding FrameNet resources (cf. [Fillmore 2008])², the Russian Constructicon is built independently from FrameNet and uses its own set of syntactic and semantic

² Available online at <https://framenet.icsi.berkeley.edu/fndrupal/>.

types to describe the constructions. In what follows, we compare Croft's approach to that adopted in the Russian Constructicon.³

4. Constructions in Croft's account and in the Russian Constructicon: a comparative account

The construction is a notion that is, on the one hand, very useful in current linguistic studies, and, on the other hand, rather challenging. Almost everyone recognizes the importance of constructions, at least in some of the senses used in linguistics and to some degree. At the same time, the great diversity of approaches and definitions points to the fact that different authors may employ the same notion differently — and, thus, several construction-oriented descriptions may in fact be technically and theoretically different.

A comparison across various approaches to constructions is of particular relevance because Lyngfelt et al. [2022] propose the application of Croft's CCs to the emerging field of multilingual constructicography, which aims to compare constructions across various languages. However, cross-linguistic comparison of constructions presents a challenge, particularly for typologically distinct languages (cf. an exploration of the alignment of constructions in English and Japanese via interactional frames in [Ohara (forthcoming)]). This difficulty extends even to typologically close languages, as differences in the choice of linguistic material across constructicon resources and variations in the granularity of grammatical constructions pose obstacles to cross-linguistic comparison. CCs are supposed to map constructions in different languages to conceptual domains, but that does not mean finding constructional equivalents [Lyngfelt et al. 2022].

But can we apply the set of CCs developed by Croft to the constructions in the Russian Constructicon? Are the two systems compatible? What is the conceptual overlap of the two classifications and how are they different? We address these questions below.

It is necessary to note that the notion of constructions used by Croft is not identical to that used in the Russian Constructicon. The main difference is that for the Constructicon, a **construction is an empirical object**, while for Croft's description of morphosyntax, a **construction is a theoretical object**. The following formulations illustrate each approach:

Croft: the main instrument for language description are constructions, and many aspects are easier to analyze in the construction-based description. The constructional mechanism is equally used in the description of the core grammatical mechanisms (e.g., complement clauses) and narrow quasi-grammatical phenomena (e.g., copula constructions).

Russian Constructicon: the class of constructions is heterogeneous, some of them are lexicalized (semantically and syntactically close to lexemes and/or peripheral grammatical markers with narrow combinational properties), and others belong to a core grammatical domain.

Of course, many language elements can be described as constructions in both frameworks. For instance, the Russian *èto*-construction that marks the real participant of the situation as contrasted to other, hypothetical, participants is a construction in both senses:

- (3) *Èt-o Vasj-a razbi-l okn-o.*
 this-NOM.N.SG Vasja-NOM.SG break-PST.M.SG window-ACC.SG
 'It was Vasja who broke the window.'

³ The Russian Constructicon is the result of many joint efforts; among the major contributors are Anna Endresen, Daria Mordashova, Ekaterina Rakhilina, Laura Janda, Olga Lyashevskaya, as well as over 40 students from both UiT The Arctic University of Norway and HSE University in Moscow.

In this case, not only the presence of *èto*, but also its linear position (it typically appears in the leftmost position), and the intonational emphasis on the argument following *èto* mark the contrast between *Vasja* and other possible participants. The construction functions as a complex multi-word unit and thus calls for a Construction Grammar analysis. The same seems to be the case with stative complex predicates, discussed in Chapter 14 of [Croft 2022], as well as by [Goldberg 1995; Himmelmann, Schultze-Berndt (eds.) 2005], among many others.

However, many constructions, such as transitive clause constructions; constructions with adjunct embedded clauses; complement clause constructions, and so on, are covered in the monograph but definitely feel to be less lexicalized and closer to the grammatical level in the proper sense. The reason is that Croft regards constructions as a convenient theoretical instrument for describing various types of constituents, mainly because elements of almost any constituent type are interconnected: for instance, though complementation is very often marked by a dedicated syntactic element (complementizer) or morphological form of the embedded verb (infinitive or another embedded form), its description also requires appealing to the main clause properties: Which classes of main predicates take complement clauses? Why do some quasi-synonymous predicates differ in ability to take complement clauses or in the type of complement clauses they take? How are the main and the embedded event in the complement clause construction situated in time?

By contrast, in the Russian Constructicon constructions are regarded as an empirical phenomenon. In this sense, Croft's broad notion of a construction is too theoretical to be directly applied to a constructicon. The Russian Constructicon team considers constructions a heterogeneous class of elements that include 'core grammar' as well as lexicalized constructions ('quasi-grammar') and lexicon.

For the 'quasi-grammatical' and 'lexical' subclass of constructions, the two key features that are often described are the **absence of markers and / or non-compositional use of markers** and **lexicalization**. A construction may, for example, (i) not contain an element whose main or only function is to connect other (meaningful) elements and (ii) the link between elements may be stronger than in most language structures of similar type (as is the case with the depictive construction: the link between the depictive and the main clause is stronger than between the clauses in a complex sentence, which leads to non-predictable selective restrictions). To illustrate, the construction *to i delo* 'often' with the meaning of frequency is semantically very close to a grammatical iterative marker. Furthermore, it contains several word forms and is not interpreted compositionally, based on the meaning of the words *to* 'that' and *delo* 'business'.

By definition, such a class of lexicalized / quasi-grammatical constructions cannot contain subtypes of "Transitivity expression" or "Voice expression". In general, voice mechanisms and transitivity features are systemic phenomena that characterize the language system as a whole, are marked in a standard way (e.g., voice markers, direct object case markers), and are far from being lexicalized. At the same time, some transitivization patterns, such as those with the word *svoë* 'one's' (*on svoë otbegal* 'he cannot run anymore', lit. 'he his own out-ran') can be included in construction databases due to the peculiar and non-compositional use of the transitive verb in combination with the reflexive possessive pronoun.

Some of the phenomena discussed by Croft should be assigned a peripheral position in a constructicon, because they show no sign of constructionalization in the narrow sense (they are interpreted compositionally, and their elements are productive grammatical markers or syntactic elements). This means that the conditional construction with *if* in English or the purpose construction with *čtoby* in Russian is not a core element of the constructicon (they are elements of 'big' syntax, namely, principal types of conditional and purpose subordinate clauses). On the contrary, constructions where grammatical elements are used in a noncanonical way, such as the "order" construction with *čtoby* (*Čtoby ja tebjja zdes' bol'se ne videl* 'I order you to never come here again!' (lit. 'that I would never see you here!')) are core elements of a constructicon: *čtoby* is here neither a subordinator, nor a marker of the purpose meaning, and the properties of this construction

type (such as the absence of the main clause) characterize the construction as a whole, rather than the word *čtoby*, which serves here as a quasi-grammatical marker of imperative.

Let us now describe the general makeup of constructicons (based mainly on the example of the Russian Constructicon) in more detail and compare it to Croft's proposal. The Russian Constructicon is an electronic linguistic database that consists of more than 2200 grammatical constructions of Russian (<https://constructicon.github.io/russian/>). The constructions in the resource are supplied with elaborate semantic, syntactic, and stylistic annotation, definitions translated into English and Norwegian, corpus-based example sentences, common fillers for construction slots, and constructional equivalents in English and Norwegian. Although constructicons are often described as “dictionaries of grammatical constructions” [Lyngfelt et al. 2022: 101], a constructicon is emphatically not a list but a structured network of constructions that are linked to each other [Fillmore 1988: 37; Goldberg, Herbst 2021: 286] or even a “set of networks” [Langacker 2008: 237]. Identifying the links between constructions makes it possible to model horizontal and vertical relations between them and to group constructions into families, clusters, and networks. The elaborate description of the semantic classification of constructions in the Russian Constructicon as well as the family-based expansion of the constructicon is available in [Janda et al. 2023] and [Endresen et al. (forthcoming)].

4.1. Bottom-up approach to classification

The semantic, syntactic, and stylistic classification of constructions in the Russian Constructicon has resulted from a bottom-up approach and emerged together with the expansion of the database. In practice, this means that constructions were added and annotated at the same time. Oftentimes the decisions about the annotation had to be reconsidered several times during the process when more data were added and constructions were analyzed in groups.

The bottom-up approach means collecting a maximally inclusive inventory of constructions and analyzing them on their own terms, thus allowing patterns to emerge from the data. The bottom-up approach is opposed to the top-down approach that presupposes a predefined system in which more granular distinctions are formed later.

This bottom-up approach means that the starting point for a Constructicon is always the data of an individual language. Not only may the items and meanings of a given semantic zone be very different across languages, but also the set of domains and subdomains in different languages can also be incomparable, meaning that its analogues cannot be found in other languages. This makes the “cross-constructicon” comparison problematic, which means that the items collected for each language and each domain have to be further checked for their membership in one or another small domain. In this sense, typological research based on the data of constructicons is “post-hoc”, built after and upon the individual-language descriptions. On the contrary, Croft's typology and typological studies, e.g., by Dahl [1985], Comrie [1989], Shopen [2007], from which Croft inherited many systemic features is mostly pre-hoc: they are built on broad (but not so finely distinguished and described) semantic zones and syntactic phenomena types that are often assumed rather than emergent from language data.

4.2. Semantic classification

Constructions are defined in Construction Grammar as conventionalized form-meaning (or function) pairings [Goldberg 2006: 5] (see also [Goldberg 1995; 2013]). Semantic types identified in the Russian Constructicon characterize the latter aspect of constructions, i.e. their meanings. As mentioned above, Croft's types are not always defined on semantic grounds: very often, their

classification is based on salient syntactic and/or morphological features. The semantic classification in the Russian Constructicon comprises 55 major semantic types and 182 more specific subtypes. Semantic types of constructions form larger units: subclasses and classes. All major semantic types are illustrated in Table 1 as bullet points. If a semantic type has subtypes, it is illustrated with a plus symbol next to the type. For example, the semantic type **Addressee** has four subtypes (**Core addressee**, **Audience**, **Beneficiary**, and **Maleficiary**), while the semantic type **Possession** has no subtypes. The description of all semantic types and subtypes is available on the Instructions page of the Russian Constructicon website⁴ and in [Janda et al. 2023]. The larger units (classes and subclasses) are illustrated with numbered black and blue categories in the table respectively.

Table 1

Semantic classification of constructions in the Russian Constructicon:
55 main semantic types included

1. Qualia			
1.1. Situation structure	1.2. Major roles	1.4. Logical relations	1.6. Sets and elements
• Timeline + • Taxis + • Actionality + • Pluractionality + • Phase of action + • Result + • Actuality	• Addressee + • Instrument • Possession • Comitative • Caritive • Non-standard subject +	• Cause • Purpose • Consequence • Condition • Concession +	• Additive • Inclusive • Exceptive • Exclusive • Subset • Options • Quantification +
1.3. Situation modifiers		1.5. Properties	1.7. Magnitude
• Spatial expressions + • Temporal expressions + • Manner		• Salient property + • Temporary characteristics + • Comparison +	• Non-existence + • Measure + • Calculation +
2. Modality and its neighborhood	3. Subjectivity	4. Discourse	5. Parameters
2.1. Core modal meanings	• Assessment + • Attitude + • Polarity value + • Source of opinion + • Mirative	4.1. Discourse organization	• Degree of intensity + • Degree of accuracy +
• Root modality + • Epistemic modality +		• Discourse structure + • Intersubjectivity • Objectivity • Source of information	
2.2. Neighborhood		4.2. Discourse clauses	
• Volition • Causation + • Prohibition + • Threat • Request • Apprehension + • Curse		• Reaction to the previous discourse + • Routine +	

Importantly, the principles of classification adopted in the Constructicon and in the monograph under analysis are different. Croft begins the classification from sentence types: for him,

⁴ Available online at <https://constructicon.github.io/russian/semantic-types/>.

a construction is a ‘type of utterance / sentence’. This does not mean that only whole-sentence categories are discussed; however, for example, voice is more likely to figure in Croft’s study than, for instance, case-marking or prepositional expression of the argument. The semantic classification of the Russian Constructicon works as a system of tags, and a construction can be assigned more than one semantic type. This is different from the system adopted in constructicons based on FrameNet classification (i.e., of semantic frames), where semantic multiple motivation of constructions is not allowed. The system of tags adopted in the Russian Constructicon is useful to capture the complex semantics of a given construction and is fruitful for modeling the horizontal relations across constructions. The combination of semantic and syntactic classification reveals a structured network of constructions.

In the Russian Constructicon, the semantic types are identified based on shared semantic properties of constructions. The structure of the semantic type, therefore, resembles a radial category with more prototypical members and a semantic periphery. For instance, both constructions (4) and (5) express the cause of some action. Therefore, they bear the semantic tag **Cause**. However, while (4) *introduces* a cause, (5) is a *question to the addressee* about the cause of his / her actions. Thus, (5) can be considered less prototypical than (4). The construction in (5) additionally conveys the speaker’s dissatisfaction with some action by the interlocutor and is also tagged as **Attitude** with the subtype **Dissatisfaction**. This additional semantical component places the construction in (5) even further away from the prototype.

- (4) ID 53 **vvidu togó, čto Cl, Cl** ‘due to the fact that...’

Izdatel'stv-o *objazan-o* *vyplati-t'* *avtor-u* *vs-e*
publishing_house-NOM.SG obliged-N.SG pay-INF author-DAT.SG all-ACC

100 % *gonorar-a* *vvidu* *t-ogo,* *čto* *rukopis'*
royalties-GEN.SG due_to that-GEN.SG that manuscript.NOM.SG

otklonen-a *uže* *posle* *odobreni-ja.*
rejected-F.SG already after approval-GEN.SG

‘The publishing house is obliged to pay the author 100 % of the royalties **due to the fact that** the manuscript was rejected after approval.’

- (5) ID 353 **čego (ěto) NP-Nom VP?** ‘why...’

Č-ego *ět-o* *ty* *lež-iš'?*
what-GEN.SG this-NOM.N.SG you.NOM lie-PRS.2SG

‘Why are you lying here?’

In Croft’s classification, (4) and (5) would definitely belong to different classes. While (4) and similar constructions are classified as a type of complex clauses (more precisely, adjunct clause constructions), (5) would be described as a question. Croft’s approach is in fact convenient for cross-linguistic comparison. The most visible and most accessible for formalization are aspects of language related to **marking**. While the utterance like ‘He left because he felt tired’ very often contains only a marker of cause, a question of the type ‘Why are you lying here?’ is often marked as a question (questions being in most languages more derived and more marked than affirmative utterances).

The radial structure of semantic zones, adopted in the Russian Constructicon, is not convenient for a typological study and is not used in his book. Croft’s semantic labels aim to be more clear-cut and well differentiated from each other for the sake of typological comparison. Furthermore, Croft’s book relies on the convenience of syntactic radial categories, for instance, for description of voice and transitivity.

Contrary to the Russian Constructicon, Croft’s study does not feature a detailed semantic classification. The wider the typological sample, the more the semantic description gets reduced to generalized domains and semantic types.

The semantic classification obtained using a bottom-up approach is, nevertheless, verified by typological studies, reflecting the types of meanings that are encoded grammatically in other

languages. In the Russian Constructicon, many semantic types correspond conceptually as well as terminologically with typologically attested grammatical categories. For instance, the construction that indicates the absence of a secondary agent or the absence of an object possessed by the main participant of the situation bears the semantic tag **Caritive** borrowed from the typological literature, see, e.g., [Plungian 2011: 125]. The semantic approach to constructions in the Russian Constructicon, however, often distinguishes more subtypes and includes more values than the typological approach as it has to deal with complex multiword constructions and not just grammatical affixes. The semantic classification also includes semantic types that are not attested in the typological literature, such as **Calculation** or **Curse**.

4.3. Syntactic classification in the Russian Constructicon

Unlike the semantic classification that works as a system of tags assigned to a construction, the syntactic classification of constructions adopted in the Russian Constructicon resembles a system of boxes, into which the constructions are distributed during the classification process. This means that, as a rule, one construction can belong to only one syntactic type. However, the syntactic annotation is four-dimensional to capture the complex structure of multiword constructions.

First, there is a **Syntactic type of construction**. This annotation category describes construction as a whole. Second, **Syntactic function of anchor** describes the syntactic function only of the fixed elements of the construction. Third, the anchor itself can be a complex entity, and thus we need to have **Syntactic structure of anchor** category for annotation. Finally, all anchor elements of construction are annotated according to **Part of speech of anchor**.

To illustrate, the example (4) can be analysed as shown in Table 2. The construction as a whole is a **Connection construction** because it contains a combination of words that perform as a conjunction. This combination of words is an anchor part of the construction and it functions as a **Subordinator**. Since this is not a single word conjunction, *vvidu togó, čto* is annotated as a **Multiword Conjunction**. If we decompose this multiword conjunction, we will see that it consists of a **preposition** (*vvidu*), a **pronoun** (*togó*), and a **conjunction** (*čto*).

Table 2

Syntactic annotation of constructions ‘ID 53 *vvidu togó, čto* Cl, Cl’
and ‘ID 353 *čego* (čto) NP-Nom VP?’

Cxn	Syntactic type of construction	Syntactic function of anchor	Syntactic structure of anchor	Part of speech of anchor
ID 53	Connection construction	Subordinator	Multiword Conjunction	Preposition Pronoun Conjunction
ID 353	Clause	Modifier	Not applicable	Pronoun

Construction (5) is less structurally complex because its anchor part does not comprise a complex entity. Thus, as a whole this construction is a **Clause**, where the anchor *čego* functions as a **Modifier**. It has an optional element *ěto* which performs an emphatic function. Structurally, *čego* and *ěto* do not comprise one complex entity; therefore, the syntactic structure of an anchor is not characterized for this construction (tag ‘Not applicable’). Both *čego* and *ěto* are **Pronouns**, which is reflected in the part of speech annotation.

While the semantic system is organized as a system of tags to capture the complex semantics of multiword constructions, the syntactic system is organized as a four-dimensional system of types in order to capture the complex syntactic structure of constructions.

The most complicated type of formal classification is the classification of constructional patterns. While within one language (and one constructicon) it is possible to describe the techniques a language uses to make constructions from what used to be free combinations, this cannot be achieved in the typological perspective and, thus, is not so crucial for Croft's project. However, in some examples it is useful to describe constructional patterns even in typological studies. For instance, what Croft describes as constructions with stative predicates is in fact a result of the widespread process of constructionalization. Here belongs the resultative construction like *We painted the door red* [Croft 2022: 432], which, according to [Goldberg 1995], made its way to a construction through constructionalization of the combination of a verb like *paint* and an adjective denoting the final state, such as *red*.

4.4. Communicative types

Unlike the large semantic and syntactic annotation systems, the system of communicative types in the Russian Constructicon is rather small. This is due to the fact that this annotation system is only applicable to those constructions that form a clause or a combination of clauses. The system of communicative types has only four categories: **Declarative**, **Interrogative**, **Exclamatory**, and **Interrogative/Exclamatory**. The names of the types are mostly self-explanatory, with the exception of the **Interrogative/Exclamatory** type, which means that the construction combines two types at the same time. This generally refers to constructions that function as rhetorical questions and in fact express negative attitude of the speaker towards the actions of the addressee, as the construction represented in (5).

5. Two classifications combined

In this section, we describe differences between Croft's model and the Russian Constructicon, applied to individual examples and generalized semantic domains. Croft's four-part model might be expected to fit neatly with the annotation model of the Russian Constructicon since the latter is also divided according to form-function distinctions. Presumably, one would search for the associative links between

- the semantic types in the Russian Constructicon and the semantic content (*sem*) type of the CCs;
- the syntactic types in the Russian Constructicon and the strategy (*str*) type of the CCs;
- the communicative types in the Russian Constructicon and the information packaging category (*inf*) type of the CCs.

Construction (*cxn*), which is the fourth type of CCs, is a more complex comparative concept since it consists of both form and function.

Let us first apply Croft's CCs to an already discussed example from the Russian Constructicon. If we look at the construction (5) from Section 4.2, we can characterize its semantic content as **causal** (*sem*), which Croft [2022: Appendix] defines as "the semantic relation between two events where one event causes the other". An additional semantic CC that is applicable here is **cause** (*sem*), defined as "a semantic role including participant roles for a participant, usually an event, that causes the event expressed by the predicate" [Ibid.: Appendix]. As we can see, this CC is related to a semantic role of the subordinate element (the pronoun *čego*) in the construction. The annotation of the Russian Constructicon also includes a set of semantic roles. They are assigned to the constructional slots and are marked in the definition of construction and in example sentences. Example (4), therefore, is annotated in the resource in the following way (6), with a Situation and a Cause filling the slots:

(6) ID 53 **vvidu togó, čto Cl, Cl** ‘due to the fact that...’

<i>[Izdatel'stv-o</i>	<i>objazan-o</i>	<i>vyplati-t'</i>	<i>avtor-u</i>	<i>vs-e</i>
<i>publishing_house-NOM.SG</i>	<i>obliged-N.SG</i>	<i>pay-INF</i>	<i>author-DAT.SG</i>	<i>all-ACC</i>
<i>100 %</i>	<i>gonorar-a</i> _{Situation}	<i>vvidu t-ogo,</i>	<i>čto</i>	<i>[rukopis']</i>
	<i>royalties-GEN.SG</i>	<i>due_to</i>	<i>that-GEN.SG</i>	<i>that manuscript.NOM.SG</i>

<i>otklonen-a</i>	<i>uže</i>	<i>posle</i>	<i>odobreni-ja.</i> _{Cause}
<i>rejected-F.SG</i>	<i>already</i>	<i>after</i>	<i>approval-GEN.SG</i>

‘The publishing house is obliged to pay the author 100 % of the royalties **due to the fact that** the manuscript was rejected after approval.’

Syntactically, the most relevant CC from the list for construction (4) is **adverbializer** (*str*) defined by Croft [2022: 664] as “a morpheme that overtly expresses the semantic relation in an adverbial clause construction”. Croft has an extended understanding of a morpheme and lists, for example, English *because* as a morpheme. Therefore, it seems legitimate to extend this terminology to the Russian *vvidu togó, čto* complex conjunction. Adverbializer is a subtype of **conjunction** (*str*) CC, “a free morpheme or clitic that encodes the relation between the events denoted by the two clauses in a complex sentence construction” [Ibid.: 680]. In Croft’s terms, construction (4) would also be considered an instantiation of an **adverbial clause construction** (*cxn*), which is a CC defined as “a complex sentence construction with a figure–ground construal / information packaging of the relation between the events denoted by the two clauses” [Ibid.: 664]. As we can see, this CC is still predominantly form-oriented rather than function-oriented.

Zooming out from the level of individual constructions to the level of semantic types within the Russian Constructicon, a broader perspective reveals systematic correspondences between these semantic types and Croft’s Comparative Concepts. A comprehensive comparison at this level seems rather promising, as the two systems demonstrate a substantial level of compatibility because they refer to similar conceptual domains. Thus, 22 out of 55 semantic types in the Russian Constructicon fully overlap with the CCs identified by Croft and 10 semantic types exhibit partial overlap with the set of CCs. We deliberately compare the semantic types and the CCs on a conceptual level because they can be distinct terminologically. For instance, we consider **Apodosis** (*cxn/sem*) [Croft 2022: 668] to fully correspond to the semantic type **Consequence** in the Russian Constructicon as they both denote the consequences of an action. Overall, full and partial overlap with CCs is attested for approximately 58 % of the semantic types in the Russian Constructicon.

Interestingly, for some semantic subclasses, Croft’s approach is more compatible with the account adopted in the Russian Constructicon than for others. For instance, the subclass **Logical relations** of the semantic class **Qualia** shows the highest degree of overlap between the system of the Russian Constructicon and Croft’s system (100 %), while the subclass **Discourse clauses** of the semantic class **Discourse** shows no overlap at all (0 %). The degree of overlap between the two systems is exemplified in Table 3. The table shows the semantic annotation in the Russian Constructicon (with types, subclasses, and classes). The dark green highlighting is used for semantic types that have a corresponding Comparative Concept in Croft’s classification. The light green highlighting shows that there is a partial correspondence between the classifications. For example, the semantic type **Actuality** in the Russian Constructicon has six subtypes: **Attenuative**, **Cumulative**, **Distributive**, **Gradual development**, **Punctual**, and **Saturative**, but in Croft’s classification only **punctual** (*sem*) is listed as a Comparative Concept.

The clearest difference between Croft’s classification and the Russian Constructicon is observed in semantic domains that contain many values not coded by grammatical markers. Compare, for instance, the aspectual domain: while the opposition of resultative and non-resultative is often coded by verbal inflectional markers, this is rare for saturative. This means that in Croft’s approach resultative or punctual has a greater chance to be classified as a typologically relevant constructional pattern, relevant for the whole-sentence level (punctuality is relevant for the semantic transitivity degree, cf. [Hopper, Thompson 1980], and resultative is closely

related to passive). By contrast, saturative only has a chance to be represented in constructions due to the existence of individual constructions like *On napilsja v drova / v nol'* / ... 'He is very drunk' (lit. 'He got drunk to the firewood / to zero', and so on).

Table 3

**Overlap of the semantic classification of constructions
in the Russian Constructicon with Croft's Comparative Concepts**

1. Qualia			
1.1. Situation structure	1.2. Major roles	1.4. Logical relations	1.6. Sets and elements
• Timeline + • Taxis + • Actionality + • Pluractionality + • Phase of action + • Result + • Actuality	• Addressee + • Instrument • Possession • Comitative • Caritive • Non-standard subject +	• Cause • Purpose • Consequence • Condition • Concession +	• Additive • Inclusive • Exceptive • Exclusive • Subset • Options • Quantification +
1.3. Situation modifiers		1.5. Properties	1.7. Magnitude
• Spatial expressions + • Temporal expressions + • Manner		• Salient property + • Temporary characteristics + • Comparison +	• Non-existence + • Measure + • Calculation +
2. Modality and its neighborhood	3. Subjectivity	4. Discourse	5. Parameters
2.1. Core modal meanings	• Assessment + • Attitude + • Polarity value + • Source of opinion + • Mirative	4.1. Discourse organization	• Degree of intensity + • Degree of accuracy +
• Root modality + • Epistemic modality +		• Discourse structure + • Intersubjectivity • Objectivity • Source of information	
2.2. Neighborhood		4.2. Discourse clauses	
• Volition • Causation + • Prohibition + • Threat • Request • Apprehension + • Curse		• Reaction to the previous discourse + • Routine +	

The second major finding from a comparison on the level of semantic types is that the semantic classification in the Russian Constructicon is often more granular than Croft's CCs. For example, the **Epistemic modality** type in the Russian Constructicon has two subtypes: **High degree of certainty** and **Low degree of certainty**, while the corresponding CC called **epistemic modality** (*sem*) has no subtypes. One would expect this from a large-scale typological classification which aims to cover a fully inclusive inventory of meanings rather than elaborate on fine-grained nuances of meanings. A more fine-grained analysis is, however, beneficial for describing constructions of a particular language.

The difference in the nature of semantic classification allows us to describe Croft's system as a **grammar of constructions** (not in the conventionalized sense 'construction grammar', but

denoting ‘the system representing the most general semantic domains, relevant for grammatical marking and features of the whole sentence’) and a constructicon as a **grammatical dictionary of constructions** (that contains much information about the behavior of constructions but includes very fine classification, characteristic of lexical descriptions).

The system of communicative types in the Russian Constructicon exhibits almost full correspondence with Croft’s CCs, as shown in Table 4.

Table 4

Communicative types compared to information packaging CCs

Communicative type in the Russian Constructicon	Comparative Concept [Croft 2022]
Declarative	declarative (<i>infl/cxn</i>)
Interrogative	interrogative (<i>or</i> question) (<i>infl/cxn</i>)
Exclamatory	exclamative (<i>infl/cxn</i>)
Interrogative / Exclamatory	—

The description of parts of speech poses a problem in both approaches. The Russian Constructicon includes a special feature **Part of speech of anchor** — however, problematic cases where the part of speech is not obvious are not addressed in a systematic way. In Croft’s approach, parts of speech are partially covered by the description of strategies (*str*) for a given semantic domain. For example, in Chapter 4, numerals and quantifiers (Section 4.1.3) are considered separately from nominal modification constructions (Section 4.1.4), and the description of comparatives and equatives in Chapter 17 distinguishes strategies employing a particle (English *than*, Latin *quam*) and those employing a preposition, a postposition, or a case marker (Mundari *ete* ‘from’, Breton *wid* ‘for’). However, Croft’s monograph does not answer the question whether the Comparative Concept approach is sufficient to describe basic syntactic notions, such as parts of speech.

6. Possible relationships between the two approaches

The comparison between Croft’s monograph and studies used for building constructicons reveals a difference of approaches — both may be adequate, but their domains of use and the range of tasks they are appropriate for are different. In particular, Croft’s approach, which builds on the a priori classification of grammatical and discourse phenomena, is not very convenient for building the constructicon of a specific language. Moreover, using an a priori classification may result in missing some peripheral (for grammatical typology) language phenomena, which, however, are richly represented by lexicalized constructions.

At the same time, it seems that adopting Croft’s approach can be useful on some stages of building a constructicon — not as the only source of language phenomena classification, but rather as a starting point. The most adequate approach seems to be to switch between the typological approach and individual language studies and, then, to analyze their results in a cross-linguistic perspective. A possible strategy is demonstrated below.

- 1) Croft’s approach can be used as a source point for constructicons, as it identifies a possible range of phenomena that could presumably be used for a constructicon. The list of notional categories and language elements proposed in the monograph is necessary, but not sufficient: domains of grammar mentioned by Croft should be included in any constructicon (or at least checked for being really ‘constructional’ in the language under analysis), though there are many domains and elements missing from Croft’s system.

- 2) Each of the domains proposed by Croft can be tested against the data of a particular language. This testing makes it possible to add new elements to any domain. Due to the fact that many of Croft's domains are labeled formally (e.g., "Nonpredicational clauses"), many resulting groups and entries of a construction may get a two-part description: for instance, the construction *u* + NP-GEN + NP-NOM (*U Pet-i gripp-Ø* [by Petja-GEN.SG flu-NOM.SG] 'Petja has the flu') can be semantically classified as a construction that marks a person's physical state. On the other hand, on the basis of syntax, it can be described as a construction with a zero copula (nonverbal predication). Similarly, the construction PRAEDIC + NP-ACC + V-INF (*Neobxodimo èt-o sdela-t'* [necessary.PRAEDIC this-ACC.N.SG do-INF] 'It is necessary to do it') is semantically a modal construction, and syntactically, a nonverbal predication with a predicative *neobxodimo*.

7. Conclusions

Let us sum up our conclusions based on the comparison of the two approaches. Analysis of the semantic and syntactic classification in the Russian Constructicon suggests that it exhibits considerable overlap with the system of Comparative Concepts introduced by Croft. Conceptually, the structure of annotation in the Russian Constructicon is compatible with Croft's system, suggesting that Semantic types predominantly correspond to Semantic content (*sem*) type of CCs, Syntactic types represent the Strategy (*str*) type of CCs, while Communicative types correspond to Information packaging (*inf*) type.

However, the semantic classification in the Russian Constructicon is much more detailed than that in Croft's system, while the communicative types are less developed compared to Croft's information structure types. Although this comparison provides a preliminary overview, further investigations are required for a more comprehensive analysis of the compatibility of two systems, especially on the level of individual constructions.

A bottom-up classification based on data from a single language and a full-scale typological classification based on multiple genetically distinct languages arrive at similar results, thus cross-validating each other. The complementary categories that are present in one system and absent from the other are a fruitful source for the expansion of the constructicon resource as well as for updating the list of Comparative Concepts. This also suggests that a closed universal list of categories may not be sufficient to capture the semantic types of many language-specific constructions. However, a list of Comparative Concepts can indeed be a useful first step for a multilingual mapping of constructions.

What is definitely missed by Croft's analysis or a similar approach built on Comparative Concepts is the difference between constructions organized as core grammatical phenomena (for instance, adjunct and complement clauses, voice, and so on) and constructions that include a closed set of elements and are closer to lexical (or quasi-grammatical) phenomena, such as saturative (*on napilsja* 'he is drunk, he has drunk enough') or special types of terminative (*on svoë otbegal* 'he cannot run anymore', lit. 'he his own out-ran'). Comparative Concepts are not suitable to distinguish between these levels of constructional description. The bottom-up approach to classification makes it possible to differentiate between 'grammar' and 'quasi-grammar'/'lexicon' in the description of constructions.

Although some technical and theoretical solutions proposed in Croft's monograph are not uncontroversial, the book will be of great use both for descriptive studies and for linguistic typology and theory. The monograph covers a great deal of linguistic phenomena that either have not received much attention or have rarely been systematically described, with phenomena of different types and different languages put together. In this sense, Croft's monograph is not only a significant contribution to Construction Grammar and linguistic typology, but also a potential

starting point for many studies and projects, lying at the crossroads of grammatical description, typology, and construction grammar.

The combination of Croft's Comparative Concepts and the construction technique developed by the Russian Construction provides a way for further development of constructional approaches. One of the questions that arise is whether a two-faceted approach is possible: on the one hand, constructions can be classified on the basis of argument marking and semantic roles of arguments (as well as a more fine-grained semantic classification), on the other hand, labels of the general constructional typology can be used.

A more ambitious task is to elaborate an opposition of assertive / non-assertive or head / non-head semantic components that form the meaning and the formal make-up of constructions. For instance, how formally and semantically different are constructions that are marked as questions and those that are questions only in the discourse level, but not formally marked as questions? What can help us to decide whether most features of a given construction result from it being 'a question construction' or 'a causal construction'? This research problem is waiting to be addressed.

ABBREVIATIONS

2 — 2nd person
ACC — accusative case
DAT — dative case
F — feminine
GEN — genitive case
INF — infinitive
M — masculine

N — neuter
NOM — nominative case
PRAEDIC — predicative (nonverbal predicate)
PRS — present tense
PST — past tense
SG — singular

REFERENCES

- Aoun et al. 2010 — Aoun J. E., Benmamoun E., Choueiri L. *The syntax of Arabic*. Cambridge: Cambridge Univ. Press, 2010.
- Boas 1911 — Boas F. Introduction. *Handbook of American Indian languages*. Vol. 1. Boas F. Washington, DC: Government Printing Office, 2011, 1–83.
- Comrie 1989 — Comrie B. *Language universals and linguistic typology: Syntax and morphology*. 2nd edn. Chicago: The Univ. of Chicago Press, 1989.
- Croft 2003 — Croft W. *Typology and universals*. 2nd edn. Cambridge: Cambridge Univ. Press, 2003.
- Croft 2022 — Croft W. *Morphosyntax: Constructions of the World's languages*. Cambridge: Cambridge Univ. Press, 2022.
- Dahl 1985 — Dahl Ö. *Tense and aspect systems*. Oxford: Blackwell, 1985.
- Endresen et al. (forthcoming) — Endresen A., Janda L. A., Zhukova V., Mordashova D., Rakhilina E. Turning a list into a network via family-based expansion of the Russian Construction. *Constructing constructions*. Ziem A., Willich A., Michel S. (eds.). Amsterdam: John Benjamins. Forthcoming.
- Fillmore 1988 — Fillmore Ch. J. The mechanisms of "Construction Grammar". *Proc. of the 14th Annual Meeting of the Berkeley Linguistics Society*, 1988: 35–55.
- Fillmore 2008 — Fillmore Ch. J. Border conflicts: FrameNet meets Construction Grammar. *Proc. of the 13th EURALEX International Congress*. Bernal E., DeCesaris J. (eds.). Barcelona: Universitat Pompeu Fabra, 2008, 49–68.
- Goldberg 1995 — Goldberg A. E. *Constructions: A construction grammar approach to argument structure*. Chicago: Chicago Univ. Press, 1995.
- Goldberg 2006 — Goldberg A. E. *Constructions at work: The nature of generalizations in language*. Oxford: Oxford Univ. Press, 2006.
- Goldberg 2013 — Goldberg A. E. Constructionist approaches. *The Oxford handbook of Construction Grammar*. Hoffmann Th., Trousdale G. (eds.). Oxford: Oxford Univ. Press, 2013, 15–31.
- Goldberg, Herbst 2021 — Goldberg A. E., Herbst T. The *nice-of-you* construction and its fragments. *Linguistics*, 2021, 59(1): 285–318.

- Grande 2001 — Grande B. M. *Kurs arabskoi grammatiki v sravnitel'no-istoricheskom osveshchenii* [Arabic grammar in a comparative and historical aspect]. Moscow: Vostochnaya literatura, 2001.
- Greenberg 1963 — Greenberg J. H. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*. Greenberg J. H. (ed.). Cambridge (MA): MIT Press, 1963, 40–70.
- Haspelmath 2010 — Haspelmath M. Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, 2010, 86(3): 663–687.
- Himmelman, Schulze-Berndt (eds.) 2005 — Himmelman N. P., Schulze-Berndt E. (eds.). *Secondary predication and adverbial modification: The typology of depictives*. Oxford: Oxford Univ. Press, 2005.
- Hopper, Thompson 1980 — Hopper P. J., Thompson S. A. Transitivity in grammar and discourse. *Language*, 1980, 56: 251–299.
- Janda et al. 2023 — Janda L. A., Endresen A., Zhukova V., Mordashova D., Rakhilina E. From data to theory: An emergent semantic classification based on a large-scale constructicon. *Constructions and Frames*, 2023, 15(1): 1–58.
- Kittilä 2002 — Kittilä S. *Transitivity: Towards a comprehensive typology*. Turku: Univ. of Turku, 2002.
- Langacker 2008 — Langacker R. W. *Cognitive Grammar: A basic introduction*. Oxford: Oxford Univ. Press, 2008.
- Lyngfelt et al. 2022 — Lyngfelt B., Bäckström L., Torrent T. Comparative Concepts as a resource for Multilingual Constructicography. *Valency and constructions: Perspectives on combining words*. Blenselius K. (ed.). Göteborg: Meijerbergs institut för svensk etymologisk forskning, 2022, 101–129.
- Ohara (forthcoming) — Ohara K. Frames, functions, and constructional meanings: Aligning constructions across typologically distinct languages. *Constructing constructicons*. Ziem A., Willich A., Michel S. (eds.). Amsterdam: John Benjamins. Forthcoming.
- Plungian 2011 — Plungian V. A. *Vvedenie v grammaticheskuyu semantiku: grammaticheskie znacheniya i grammaticheskie sistemy yazykov mira* [Introduction to grammatical semantics: Grammatical meanings and grammatical systems of the world's languages]. Moscow: Russian State Univ. for the Humanities, 2011.
- Roy 2013 — Roy I. *Nonverbal predication: Copular sentences and the syntax–semantics interface*. Oxford: Oxford Univ. Press, 2013.
- Shopen 2007 — Shopen T. (ed.). *Language typology and syntactic description*. Cambridge: Cambridge Univ. Press, 2007.
- Testelests 2008 — Testelests Ya. G. Struktura predlozhenii s nevyrazhennoi svyazkoi v russkom yazyke [The structure of sentences with an unexpressed copula in Russian]. *Dinamicheskie modeli. Slovo. Predlozhenie. Tekst*. Rozina R. I., Kustova G. I. (eds.). Moscow: Yazyki slavyanskikh kul'tur, 2008, 773–789.
- Tsunoda 1985 — Tsunoda T. Remarks on transitivity. *Journal of Linguistics*, 1985, 21: 385–396.
- Wierzbicka 1980 — Wierzbicka A. *The case for surface case*. Ann Arbor: Karoma, 1980.

[Рец. на: / Review of:] **A. Behnke, B. Wagner-Nagy (eds.). *Clause linkage in the languages of the Ob-Yenisei area: Asyndetic constructions***. Leiden; Boston: Brill, 2024. xvi + 455 pp. (Empirical Approaches to Linguistic Theory, 23.) ISBN 978-90-04-68476-8 (print), ISBN 978-90-04-68477-5 (e-book).

Наталья Вадимовна Сердобольская

Институт языкознания РАН, Москва, Россия;
serdobolskaya@gmail.com

Благодарности: Работа поддержана грантом
РНФ № 22-18-00528.

DOI: 10.31857/0373-658X.2024.5.154-160

Natalia V. Serdobolskaya

Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia; serdobolskaya@gmail.com

Acknowledgements: The work is supported by the
RSF grant No. 22-18-00528.

Clause linkage is a domain of syntax that does not often receive much attention in reference grammars. It often happens that in a short 200-pages grammar more than half of the book is dedicated to morphology and grammatical semantics and less than 30 pages to syntax, where subordination is represented by singular examples. As far as Uralic and Altaic languages of Russia are concerned, this problem is reinforced by the fact that traditional grammars written in the USSR times and earlier were oriented towards the description of lexical subordinators and conjunctions, while non-finite subordination was scarcely mentioned in chapters devoted to verbal non-finite forms. It is only in the 1980s when the Novosibirsk linguistic school elaborated a model of non-finite subordination [Cheremisina 1980; Ubryatova, Litvin 1986] that linguists started to systematically consider non-finite clauses as a subtype of subordination. Given that in the Uralic and Altaic languages non-finite subordination is the only or at least the prevailing model (though in some of them it has been recently replaced by finite subordination with lexical subordinators), such an approach resulted in inadequate description of subordination as a whole.

In light of that, the volume on clause linkage in the languages of the Ob-Yenisei area is highly welcome, since it provides a large amount of new interesting data, a great deal of which has been previously ignored by the typological and theoretical literature. A noteworthy feature of the book is the use of the unified format of language description throughout the volume for every examined language. Each chapter is organized along a specific scheme and provides, first, information about the language, second, details about non-finite verbal forms, and third, a sketch on clause linkage, including coordination, adverbial subordination, relative clauses, and complementation. The book comprises sketches of 9 languages: Enets, Tundra Nenets, Evenki, Ket, Eastern Khanty, Mansi, Selkup, Chulym Turkic, and Kamas (some of them represented by two or more varieties).

The languages under discussion show a number of common properties in the domain of clause linkage, that is:

- extensive use of juxtaposition of clauses as a means of coordination and for some typically subordinate semantic relations, such as reason (Ob-Ugric, Selkup), purpose (Ket, Evenki), time (Khanty, Selkup), condition (Mansi, Selkup, Chulym Turkic), and complementation with utterance and mental verbs;
- clause chaining (Chulym Turkic, Evenki, and Southern Samoyedic);
- converbs (all the languages except for Ket);
- specific forms encoding purpose and dependent clauses as complements of manipulation and speech causation verbs, i.e. infinitives and/or supines (all the languages except for Ket, see the arguments against the term “infinitive” for the Ket action nominal in Andrey Nefedov’s chapter; in Evenki, Enets, and Nenets the general converb is used in this function, which may be termed as “infinitive”, see the remarks in the corresponding chapters);

- use of possessive markers to signal agreement with the subject of the dependent clause, at least in some subordination types (all the languages studied in the volume, with a reservation that in Chulym Turkic the possessive occurs on the head noun of a relative clause; Ket is an exception, since it lacks the nominal possessive paradigm, which is found in the neighboring languages; some subordinators, though, require for the possessive augment, whose function is unclear, see footnote 14 on p. 221);
- use of participles in relative clauses (all the languages, except for Ket).

An interesting feature of the languages under discussion is the fact that, when differentiating between same-subject and different-subject constructions, possessed subjects, especially bodyparts and some other objects of possession, may be treated in the same way as the same subjects. For example, Enets manner clauses with the general converb are mostly same-subject; however, the bodypart of the subject of the matrix clause is also possible in the subject position in a converb clause:

- (1) ENETS
 seta oðie teʔinaad'uʔ, teinoʔ baðotia.
 [se-ta ɔzi-e teʔi-naad'u-ʔ],
 face-OBL.POSS3SG be.seen(IPFV)-CVB up-ADV-LAT.ADV
 teino-ʔ bazotia
 there-LAT.ADV lie(IPFV).3SG.S
 'She is lying there face up.' [Shluinsky, Wagner-Nagy 2024: 69]

Another remarkable feature of the languages under examination is the possibility to combine *wh*-morphemes with non-finite verbal forms in one dependent clause. This is largely observed in indirect questions in Enets (219), (220), Evenki (30c), Eastern Khanty (20), (42a) (see [Serdobolskaya et al. 2012] on similar constructions in Finno-Ugric languages).

Some languages of the Ob-Yenisey area demonstrate headless relative clauses (Enets, Eastern Khanty, Ket, Chulym Turkic), correlative constructions (Enets, Eastern Khanty, Evenki (30b), Chulym Turkic (28), Selkup). The authors explain the use of the correlatives by Russian influence; however, it must be noted that Standard Russian lacks definite correlative constructions of the *who... that* type (only constructions with universal reading are possible, *Kto ne rabotaet, tot ne est* 'The one who does not work, does not eat'; see [Belyaev, Haug 2020]). Hence, such generalizations should be tested based on the data of neighboring Russian dialects.

Many languages under consideration have a rich and elaborated domain of temporal adverbial clauses and purpose clauses, while adversative, contrastive, and reason clauses may not receive a specific encoding, and the corresponding relations are often inferred from the context and a specific linear ordering of the clauses. With reason clauses, this pattern is observed in Enets, Tundra Nenets, Evenki, and Eastern Khanty. Adversative relations may also be encoded by conjunctions borrowed from Russian, while Tundra Nenets, Selkup, Khanty, and Mansi have native adversative conjunctions. The contrastive relation is not encoded by native conjunctions in either of the languages in question.

The terminology that is generally used for non-finite forms by most linguists presumes that, in a given language, they can be clearly differentiated on the basis of subordinate clauses they introduce. That is, converbs are expected to introduce adverbial clauses, participles are used in relative clauses, and infinitives and nominalizations head complement clauses. This idealistic picture, however, finds its exceptions in most cases. Especially Uralic and Altaic languages have a large number of non-finite verbal forms that may be assigned two or even three different labels, even if we only consider their functions. In this respect, the descriptions offered in the book under discussion offer valuable comparison of terminological labels used in different grammars and provide important discussions of the terminology. See, for example, overviews of participles, action nouns, and converbs in Nikolett Mus's chapter on Tundra Nenets. However, we cannot fully adopt the motivation behind the choice of some terms in this and some other chapters. For example, when discussing a form that shows the properties of both infinitives and converbs, the author uses a functional motivation only: "Additionally, certain verbs and adjectival

predicates can take converbs as their complements (either as a subject or as an object) in Tundra Nenets (...). This might lead us to categorize them as infinitives, i.e., they appear as clausal complements just like infinitives do" (p. 142). However, infinitives are termed as such on the basis of both their function and morphosyntactic properties: lack of person/number agreement, tense and mood distinction, sometimes also voice and aspectual distinctions, see [Iljevskaya, Kalinina 2002]. Therefore, a functional criterion is not enough to call a verbal form an infinitive.

Another weak point of many existing grammars, in what concerns clause linkage, is the fact that many authors hardly ever give examples of different-subject constructions. However, it is crucial for description, because many languages make use of specific markers for same-subject structures (e.g. Russian converbs), while other markers can be used irregardless of the same-subject vs. different-subject differentiation. The chapters in the book under consideration analyze all verbal forms in detail and consistently comment on the possibility of expressing the subject different from that of the matrix clause overtly, as well as the possibility to express other arguments and adverbials. This is another problematic point in many reference grammars: examples of "full" non-finite clauses, that is, clauses containing all the arguments and adjuncts of the non-finite predicate are often absent from the grammars. Sometimes it is accidental; however, in some cases it is a result of the lexicalization of considered forms, which do not introduce full clauses, but tend to be used to refer to a situation without expressing its participants. An example is found in Shoksha Mordvin, where participles are hardly ever used in relative clauses and may be termed "deverbal adjectives" (Aralova, Brykina, p. c.).

An important point is the differentiation of purpose constructions with motion verbs and purpose constructions with other types of matrix verbs, which is made in the volume, see the chapter on Enets (p. 70), in Ket (p. 226), Mansi (pp. 322 and 338), Kamas (p. 436). These constructions often demonstrate different morphosyntactic properties, whereas the contexts of function, e.g. "a rag for cleaning the floor", and that of "give" often behave in the same way as that of motion verbs.

All these issues are covered in the book under review, thus offering valuable information for typological generalizations.

A similar problem occurs in the description of coordination: when illustrating the use of coordinating conjunctions, many grammars only provide examples of phrases (*John and Mary*, *here and there*), and the reader remains uninformed about the ability of certain conjunctions to coordinate clauses. It must be said that in some chapters of the discussed book, coordination, especially conjunction, is illustrated mostly by examples of coordinated verbs (or other types of phrases), while examples of whole clauses are scarce. Meanwhile, in some languages it is crucial, since conjunctions coordinating singular verbs or verb phrases ("He arrived, opened the door **and** started working") are different from those that coordinate clauses ("The sun is up **and** the birds are singing"). In a book dedicated to clause linkage more examples of coordinated clauses would be expected.

I should make a special mention of the chapter on Enets by Andrey Shluinsky and Beáta Wagner-Nagy, which offers a meticulous and fine-grained description and contains many valuable remarks revealing the morphosyntactic properties of subordinate clauses and their functions. All the types of subordinate clauses are characterized at length; each type is covered in tables providing information on relevant morphosyntactic parameters. Enets is rich in purpose constructions, which receive a minute description, and the authors give a meticulous examination of complement-taking verbs and the constructions they govern. It is important that the authors clearly show the difference between mere juxtaposition and the direct speech construction (based on interpretation of indexicals) with utterance verbs, which is not examined in full in other chapters. All the generalizations are carefully formulated separately for Forest Enets and Tundra Enets.

The book opens with an introduction written by Anja Behnke and Beáta Wagner-Nagy, which gives background information on clause linkage and an overview on morphosyntax, the functions and use of non-finite forms in various types of complex sentences in the discussed languages. The introduction contains interesting generalizations, in particular on different types of converbs. The authors propose to distinguish between converbs with verbal and nominal properties. The "verbal" converbs do not take possessive and case markers; these are attested

in Khanty, Selkup, Kamas, and Chulym Turkic. Converbs with nominal properties are found in Northern Samoyedic languages (Enets, Nenets), Evenki, and Mansi. It could be further suggested to differentiate between converbs that regularly introduce full clauses with all the verbal arguments and adjuncts vs. converbs that tend to be used in same-subject constructions, in periphrastic modal and aspectual constructions and as a part of a complex verb. Such forms are found in Kalmyk (converbs in *-ad* and *-žə*, according to [Mischenko 2009; Say 2009]) and in Eastern Mari (the converb in *-ən*, see [Arkhipov 2000]). The Mari converb may occasionally introduce full clauses; however, it is most often used to denote a background situation modifying the situation described by the matrix verb, which is foregrounded (e.g. “He entered the room **laughing**”, “He left **grudging**”), while full adverbial clauses are most often headed by other converbs.

However, it looks surprising that some important generalizations made in the introductory paper are scarcely followed by the authors of language sketches. For example, the differentiation of various types of converbs is hardly ever referenced to.

Although the book’s merits outweigh its drawbacks, there are some points that deserve further discussion. In a volume dedicated to a linguistic area it would be appropriate to provide a summary of contact-induced vs. native constructions in each chapter. At least, a comparison with neighboring languages suggests itself, given that all the necessary information is provided in the same book. However, such a summary is only given for Ket, Eastern Khanty, and Selkup, while other chapters merely mention Russian influence on the syntax of complex sentences and notice that particular conjunctions and subordinators are borrowed from Russian.

In discussions of complementation, it seems that some authors do not distinguish between indirect *wh*-questions, complementation, and sometimes also actor nominalization, cf. examples (9), (81)–(83) in the paper on Tundra Nenets, (31), (34), (38) on Evenki, (35a)–(35c) on Chulym Turkic, and (28a) on Selkup. It has been previously shown in typological studies that indirect question constructions may be formed differently from other types of complement clauses [Kahrel 1985; Morozova, Serdobolskaya 2020]; note that typological works on complementation, such as [Noonan 1985] and [Dixon, Aikhenvald 2006], do not consider indirect questions at all. In some languages indirect *wh*-questions may have morphosyntactic properties similar to relative clauses (Basque, Adyghe), unlike other complementation types. In languages with objective conjugation the *wh*-word in an indirect question is expected to trigger object agreement on the matrix verb, unlike clausal complements. Consider the following example from Tundra Nenets:

(2) TUNDRA NENETS

niša-nd *xăʔmaʔ-ma-mʔ* *nămda-ra-šʔ*
 father-GEN.POSS2SG say.what-ACTN.CTMP-ACC hear-2SG.SG.O-PS

‘You heard what your father said.’ [Mus 2024: 166]

The object agreement marker in this example could be attributed either to the theme of the verb ‘say’ (the words of the father), or to the complement clause as a whole. The first interpretation is more expected on general grounds, as entities are more likely to trigger agreement than situations; in any case, this question deserves more attention.

Interestingly, the languages under discussion make extensive use of direct quotation also in context of complement-taking verbs that do not involve speech, e.g. knowledge and perception verbs. However, in the book this claim is often not supported by examples showing that these verbs take the direct quotation construction, that is, examples with indexicals, exclamationatives, or vocatives. For example, the paper on Eastern Khanty has persuasive examples with utterance verbs, but no examples with indexicals are given to show that cognition and perception verbs also take the same construction as a complement. For Mansi, an example intended to illustrate direct quotation (64) has the 3rd person plural form referring to the subject of the matrix clause and, hence, showing indirect quotation.

The use of the term “action nouns” throughout the book seems infelicitous. This term has been suggested in Koptjevskaja-Tamm’s thesis [1988] and covers only (de)verbal nominals denoting action, while (de)verbal nouns denoting uncontrolled situations and states are excluded.

Thus, the use of this term imposes a restriction on the deriving verbal stem, which is conceivably absent from (at least some of) the languages under discussion.

When considering nominalizations, most authors give a very detailed account of the use of possessive and case marking with the nominalized verb. Another interesting piece of information would be the use of nominal plural marking with nominalizations. As all the chapters demonstrate, case marking and the use of adpositions with nominalizations is usually a way to express various adverbial relations (in adverbial subordination) or may be governed by the matrix predicate (in complementation). Possessive marking is most often used to encode agreement with the subject of the embedded clause. Therefore, both these properties, although showing the presence of the nominal head, are functionally reanalyzed as clausal properties. As opposed to these two parameters, the use of the nominal plural usually signals to what extent the nominalized form is semantically close to non-derived nominals.

In the sections dedicated to complementation, the authors scarcely ever mention the use of objective conjugation with clausal arguments:

(3) MANSI

kāsal-as-anəl, mātāpriš juw
notice-PST-3PL.O mouse come.PRS.3SG.S

‘They noticed that the mouse was coming.’ [Bíró 2024: 343]

According to the sources cited in [Ibid.], the objective conjugation marker in this example refers to the situation introduced by the complement clause. Given that many languages under consideration have objective conjugation (Tundra Nenets, Enets, Eastern Khanty, Mansi, Selkup, Ket), it is important to specify whether certain complementation strategies require/allow cross-reference by the 3rd person objective markers on the matrix verb, see the discussion of the Mansi example (3) above and the remarks in the discussion of morphosyntax of Ket action nominals in Section 3.2 of Chapter 4 by Andrey Nefedov. In Enets, this agreement pattern is found with verbs ‘understand’ and ‘know’ shown in (192) and (194), respectively; it is absent in case of ‘think’ in (193) and ‘say’ in (190)–(191). Compare also (88) and (90) from Tundra Nenets, (76) and (74) from Eastern Khanty, (57) and (58) from Mansi. Comments on the semantic factors behind this agreement pattern would be highly welcome. It could also be interesting to see whether there are examples with plural agreement (with multiple clausal arguments), given that the 3SG marker might be used as a default agreement marker or even reanalyzed as a complementizer.

In what concerns Selkup, the distribution of the subjective and objective conjugation should definitely be commented on, since the glosses are not transparent here. It could be made more clear that the gloss “o” only signals the presence of the direct object in a clause, without indicating its person and number (for example, “1SG.o” (19c) or “3SG.o” (16b) does not refer to the 1st or 3rd person singular object, but rather denotes agreement with the 1st or 3rd person singular subject in the presence of the direct object). It is remarkable that the variation between subjective and objective conjugation in Selkup is sometimes attested with one and the same complement-taking verb, see examples (36a)–(36b) with ‘know’, and sometimes with the verbs having similar semantics, e.g. ‘think’ in (28c) and (40a)–(40b). The choice of conjugation type with various complement-taking verbs may puzzle the reader.

A remarkable feature of Uralic and Altaic languages is the presence of participles that can relativize several discontinuous positions on the Accessibility Hierarchy (S > DO > IndO > Obl > Gen > OComp [Keenan, Comrie 1977]), see [Brykina, Aralova 2012]. This is also observed in the languages discussed in the book, see the examples from Chulym Turkic, where the participle in *-GAn* may be used to relativize direct objects and obliques such as a time adjunct below:

(4) CHULYM TURKIC

a. *sur-ya-ŋ epči-ŋ*
ask-GAN.PTCP-POSS2SG woman-POSS2SG

‘the woman you have asked for’

- b. *tu-yan* *kün*
be.born-GAN.PTCP day
‘birthday’ [Däbritz, Karakoç 2024: 392]

It must be noted that the question of the availability of various relativization positions for each of the participles is not discussed in detail in many of the chapters. As far as Mansi, Kamas, and Selkup are concerned, the examples of relativization are scarce, the reason apparently being the lack of sources. The Mansi and Selkup examples only feature subject and object relativization, while for Kamas an example of adjunct relativization is given as well (“the land which I walked”).

All the comments above do not affect the overall high assessment of the book and contribute to the future studies of various topics it addresses. The volume brings up new data relevant for typology and theory of clause linkage, and the results of the nine studies carried out by the authors open new and deeper perspectives to studying coordination and subordination constructions in the languages of Siberia. The book is therefore highly welcome to the domain of language-specific studies as well as the typologically-oriented research on clause linkage.

ABBREVIATIONS

2, 3 — 2 nd , 3 rd person	GAN.PTCP — participle in <i>-Gan</i>	POSS — possessive
ABL — ablative case	GEN — genitive case	PRS — present
ACC — accusative	IPFV — imperfective verb	PST — past
ACTN — action noun	LAT — lative case	s — subject
ADV — adverb	O — object	SG — singular
CTMP — contemporaneous	OBL — oblique	
CVB — converb	PL — plural	

REFERENCES

- Aralova, Brykina 2012 — Aralova N. B., Brykina M. M. Finitnye odnositel’nye predlozheniya v mariiskom i erzya-mordovskom yazykakh [Finite relative clauses in Mari and Erzya-Mordvin]. *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisaniya. Formal’nyi i funktsional’nyi podkhody*. Kuznetsova A. I. (ed.). Moscow: Rukopisnye pamyatniki Drevnei Rusi, 2012, 521–543.
- Arkhipov 2000 — Arkhipov A. *Sochinitel’nye i podchinitel’nye svoistva konstruktii s deeprichastiyami* [Coordinate and subordinate properties of converb constructions]. A fieldtrip report. Moscow. Ms., 2000.
- Belyaev, Haug 2020 — Belyaev O., Haug D. The genesis and typology of correlatives. *Language*, 2020, 96(4): 874–907.
- Bíró 2024 — Bíró B. Mansi. *Clause linkage in the languages of the Ob-Yenisei area: Asyndetic constructions*. Behnke A., Wagner-Nagy B. (eds.). Leiden; Boston: Brill, 2024, 319–348.
- Brykina, Aralova 2012 — Brykina M. M., Aralova N. B. Sistemy prichastii v mariiskom i permskikh yazykakh [Participle systems in Mari and Permian]. *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisaniya. Formal’nyi i funktsional’nyi podkhody*. Kuznetsova A. I. (ed.). Moscow: Rukopisnye pamyatniki Drevnei Rusi, 2012, 476–521.
- Cheremisina 1980 — Cheremisina M. I. O soderzhanii ponyatiya «predikativnost’» v sintaksise slozhnogo predlozheniya [On the notion of “predicativity” in the syntax of complex sentences]. *Polipredikativnye konstruktii i ikh morfologicheskaya baza (na materiale sibirskikh i evropeiskikh yazykov)*. Cheremisina M. I. (ed.). Novosibirsk: Nauka, 1980, 154–180.
- Däbritz, Karakoç 2024 — Däbritz C. L., Karakoç B. Chulym Turkic. *Clause linkage in the languages of the Ob-Yenisei area: Asyndetic constructions*. Behnke A., Wagner-Nagy B. (eds.). Leiden; Boston: Brill, 2024, 386–415.
- Dixon, Aikhenvald 2006 — Dixon R. M. W., Aikhenvald A. Y. *Complementation: A cross-linguistic typology*. Oxford: Oxford Univ. Press, 2006.

- Iljevskaya, Kalinina 2002 — Iljevskaya A. A., Kalinina E. Yu. Tipologicheskie kriterii infinitiva: sushchestvuyut li oni? [Typological criteria of the infinitive: Do they exist?] *Lingvisticheskii bespredel. Sbornik statei k 70-letiyu A. I. Kuznetsovoi*. Kibrik A. E. (ed.). Moscow: Moscow State Univ., 2002, 241–262.
- Kahrel 1985 — Kahrel P. Indirect questions. *Predicates and terms in Functional Grammar*. Bolkestein A. M., de Groot C., MacKenzie J. L. (eds.). Dordrecht: Mouton de Gruyter, 1985, 165–182.
- Keenan, Comrie 1977 — Keenan E. L., Comrie B. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry*, 1977, 8: 63–99.
- Koptjevskaja-Tamm 1988 — Koptjevskaja-Tamm M. *A typology of action nominal constructions*. Doctoral diss. Stockholm: Univ. of Stockholm, 1988.
- Mischenko 2009 — Mischenko D. F. Deeprichastiya s pokazatelyami -žə i -ad v kalmytskom yazyke [Converbs with suffixes -žə and -ad in Kalmyk]. *Issledovaniya po grammatike kalmytskogo yazyka*. Say S. S., Baranova V. V., Serdobolskaya N. V. (eds.). St. Petersburg: Nauka, 2009, 225–255.
- Morozova, Serdobolskaya 2020 — Morozova V. A., Serdobolskaya N. V. Towards the typology of indirect question markers. *Typology of Morphosyntactic Parameters*, 2020, 3(2): 125–144.
- Mus 2024 — Mus N. Tundra Nenets. *Clause Linkage in the Languages of the Ob-Yenisei Area: Asyndetic Constructions*. Behnke A., Wagner-Nagy B. (eds.). Leiden; Boston: Brill, 2024, 133–175.
- Noonan 1985 — Noonan M. Complementation. *Language typology and syntactic description*. Vol. 2: *Complex constructions*. Shopen T. (ed.). Cambridge: Cambridge Univ. Press, 1985, 42–140.
- Say 2009 — Say S. S. Grammaticheskii ocherk kalmytskogo yazyka [A sketch of Kalmyk grammar]. *Issledovaniya po grammatike kalmytskogo yazyka*. Say S. S., Baranova V. V., Serdobolskaya N. V. (eds.). St. Petersburg: Nauka, 2009, 622–710.
- Serdobolskaya et al. 2012 — Serdobolskaya N. V., Il'jevskaya A. A., Minor S. A., Miteva P. S., Fajnvejc A. V., Matvejeva N. S. Konstrukcii s sentencial'nymi aktantami v finno-ugorskikh yazykakh [Complementation in Finno-Ugric languages]. *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisanija. Formal'nyi i funktsional'nyi podkhody*. Kuznetsova A. I. (ed.). Moscow: Rukopisnye pamyatniki Drevnei Rusi, 2012, 382–476.
- Shluinsky, Wagner-Nagy 2024 — Shluinsky A., Wagner-Nagy B. Enets. *Clause Linkage in the Languages of the Ob-Yenisei Area: Asyndetic Constructions*. Behnke A., Wagner-Nagy B. (eds.). Leiden; Boston: Brill, 2024, 26–133.
- Ubryatova, Litvin 1986 — Ubryatova E. I., Litvin F. A. (eds.). *Strukturnye tipy sinteticheskikh polipredikativnykh konstruksii v yazykakh razlichnykh sistem* [Structural types of synthetic complex sentences in languages of diverse types]. Novosibirsk: Nauka, 1986.

[Рец. на: / Review of:] **D. G. Krivochen. *Syntax on the edge: A graph-theoretic analysis of sentence structure***. Leiden; Boston: Brill, 2023. xxi + 494 pp. (Empirical Approaches to Linguistic Theory, 21.) ISBN 978-90-04-54142-9 (print), ISBN: 978-90-04-54231-0 (e-book).

Marina B. Ermolaeva

Lomonosov Moscow State University, Moscow, Russia; mail@ermolaeva.com

Марина Борисовна Ермолаева

Московский государственный университет имени М. В. Ломоносова, Москва, Россия; mail@ermolaeva.com

Ekaterina A. Lyutikova

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

Екатерина Анатольевна Лютикова

Московский государственный университет имени М. В. Ломоносова, Москва, Россия; lyutikova2008@gmail.com

Acknowledgements: This work was done with the support of MSU Program of Development, Project No. 23-SCH02-10 “Linguistic competence of natural language speakers and neural network models”.

Благодарности: Работа выполнена при поддержке Программы развития МГУ, проект № 23-Ш02-10 «Языковая компетенция носителей естественного языка и нейросетевых моделей».

DOI: 10.31857/0373-658X.2024.5.161-167

The title of the monograph under review, *Syntax on the Edge*, may sound misleading to generative syntacticians. Within the Minimalist program, there is the concept of the left periphery, or edge, of a phase, as opposed to the complement of a phase with respect to the accessibility of its material for the higher heads [Chomsky 2000; 2001; 2008]. Any constituents that must move out of the phase head’s complement carry relevant edge features and are attracted to the edge of the phase. In this sense, the term *edge* is actively used in minimalist syntax; cf., e.g., the title of Svenonius’ [2004] *On the edge*. In the monograph under review, however, *edge* is a term from graph theory — a directed or undirected arc that connects vertices (or nodes) in a graph. The term has been used in linguistic models representing sentence structure as a system of dependencies, such as Relational Grammar [Perlmutter (ed.) 1983] and its successor Arc Pair Grammar [Johnson, Postal 1980], as well as, perhaps, the better-known in Russia Meaning ⇔ Text Model [Mel’čuk 1974; 1988]. This understanding of *edge* is represented, for example, in *Edge-based clausal syntax* [Postal 2010], as well as — not taking into account the idiomatic reading — in the monograph under review. And while the specific term *edge* is not universal to the entire class of dependency-based grammars, the idea of syntactic structure as a set of nodes with the dependency relation is at its core — as in Word Grammar, a unified approach treating the lexicon, morphology, and syntax as networks of nodes connected by links [Hudson 1984; 2021].

Here we have another syntactic theory based on the dependency relation. What is the motivation for a new syntactic model provided by the author, who primarily works with data from English and Spanish — languages whose syntax has been analyzed very thoroughly within the framework of existing grammatical theories? As Diego Krivochen himself writes, this is an attempt to answer a “what if...?” question, that is, a theoretically motivated study of available options in choosing a model for natural language. The main idea is to try to simplify the syntactic description by reducing the number of entities (projections, null categories, traces) and increasing the number of connections between terminal nodes. This idea is opposite to the direction of development observed in generative grammar, where the relations are minimized (strictly binary branching) at the cost of increasing the number of syntactic objects. In principle, both approaches are equally viable in a situation where the object being modeled does not impose the

presence of either multiple entities or multiple relations on the model. The former does not seem to be the case when it comes to modeling natural language syntax; in particular, maximal projections of syntactic categories are not a concession to the formalism that minimizes the number of local relations, but rather entities objectively represented in the language system and acting, along with heads, as subjects of grammatical rules.

Let us now summarize the main provisions of the proposed approach. The theory of grammar envisioned by the author includes a definition of basic expressions; a definition of derived expressions; and a set of constraints over graphs. Using this characterization, analyzing sentences of a natural language boils down to “the identification of basic and derived expressions and relations between these, and the formulation of admissibility constraints” (p. 61).

A graph is defined, as commonly accepted, as a pair (E, V) , where E is a set of vertices (nodes) and V is a set of edges (arcs). The structures used throughout the monograph are directed graphs (digraphs), whose edges are ordered pairs of vertices. This yields a binary asymmetric relation “immediately dominates”, denoted as p ; $p(v_1, v_2)$ means “ v_1 immediately dominates v_2 ” and is true iff there is an arc from v_1 to v_2 .

The dominance set, or p -set, of a graph G is the set of all immediate dominance relations in G . Predicates always directly dominate their arguments, and p -sets are ordered with respect to a hierarchy of grammatical functions (p. 165), which allows predication and grammatical functions of arguments to be read off the edges of the graph. Furthermore, p -sets allow repetitions to accommodate parallel edges, which technically makes these data structures ordered multi-sets. These p -sets serve as structural descriptions of sentences. Linearization of graphs — mapping structural descriptions to strings — is explicitly (p. 58) taken to fall outside of the scope of the monograph.

Nodes in a graph correspond to basic expressions of the language, drawn from some set Exp . Each node is assigned a unique address that points to its semantic value and allows it to be referred to unambiguously. The building blocks of the grammar are elementary graphs, which are a special case of an arbor, a single-rooted directed graph. An elementary graph includes the smallest set of connected nodes that contains a predicative lexical basic expression (referred to as the elementary graph’s anchor), its functional modifiers (such as tense and aspect), and its arguments. Single-rooted graphs, and elementary graphs in particular, can be composed into derived graphs by means of one or more common nodes. This is formally defined as graph union; given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, their union is given by $(V_1 \cup V_2, E_1 \cup E_2)$. Any nodes with the same address are collapsed into a single node (called a linking node), providing a built-in implementation of structure sharing.

Further characteristics of the model being developed partly follow from the main idea of minimizing syntactic entities, and partly are independent of it and specified as axioms. Minimization of entities is achieved due to the fact that the same unit can be a dependent in more than one relation (that is, one node can have more than one incoming arc). Thus, the author discards the single mother condition (SMC); moreover, for non-elementary syntactic structures, multiple root nodes are also allowed. The theory is constructed as model-theoretic, as opposed to proof-theoretic or derivational (which is characteristic, for example, of minimalism). As a model-theoretic approach, the theory must offer a way of mapping a linguistic expression to some representation and formulate constraints over such representations that characterize the grammar of a given language or natural language in general. Since the proposed representations are digraphs (i.e., directed graphs), it is necessary to determine which elements of the linguistic expression correspond to nodes of the graph and under which circumstances two nodes are connected by a directed edge. Grammatical phenomena are modeled in terms of constraints on graph structure: the availability of a trail between the two nodes representing variables in a syntactic rule.

Many axiomatic requirements for syntactic trees appear in a weakened form (and it is not entirely clear whether this is a consequence of the inevitable sketchiness of the first version of a large model or a fundamental feature of the model). Only categorematic expressions appear as nodes, which have a corresponding semantic interpretation that varies (or may vary) across

models. Syncategorematic expressions like articles, complementizers, grammatical prepositions, etc., are not represented. Moreover, only lexical predicates can serve as an anchor of an elementary tree, whereas nominal arguments are represented as single nodes without internal structure (with the exception of dependent clauses embedded in a noun phrase and *wh*-determiners and quantifiers they are headed by). Finally, there are multiword expressions, which constitute a single node, e.g. *act like*, Spanish *ser asesinado* ‘be killed’, etc.

Edges are directed from a predicate to its arguments. Functional elements, like lexical auxiliaries, and adjuncts take predicates as their arguments. That is, the basic structure of an elementary tree closely resembles the semantic representation of the Meaning $\Leftrightarrow$ Text Model. And, as in the Meaning $\Leftrightarrow$ Text Model, different linguistic expressions can correspond to the same semantic structure. Thus, Krivochen rejects the idea of a one-to-one correspondence between a linguistic expression and its structural description.

To summarize, instead of a model that permits nodes that are not represented in a linguistic expression, a model is proposed where the number of nodes does not exceed the number of words, but which departs from a whole slew of restrictions observed by traditional models: the principle of complete representation (all linguistic units are represented in the structural description), the further indivisibility of terminal nodes, the principle of distinctiveness (a linguistic expression can be unambiguously restored from its structural description, and vice versa).

One question that is likely of much interest to the more computationally-minded among the linguistic community is the expressive power of this theory. In Chapter 4, the author claims that the proposed formalism is capable of expressing context-free patterns, such as nested dependencies in Turkish, as well as mildly context-sensitive patterns, such as the crossing dependencies in Swiss German [Shieber 1985] and Dutch [Joshi 1985]. This is illustrated with the (unordered) ρ -set representation (1b) of a famous Dutch example (1a).

(1) DUTCH

- a. *Jan Piet Marie zag helpen zwemmen.*
 Jan Piet Marie saw help swim

‘Jan saw Piet help Mary swim.’ (p. 149; from [Joshi 1985: 245])

- b. $\rho = \{(zag, Jan), (zag, Piet), (zag, helpen), (helpen, Piet), (helpen, Marie), (helpen, zwemmen), (zwemmen, Marie)\}$ (p. 153)

As in the rest of the monograph, the author steers clear of the problem of linearization. However, without a defined algorithm of linearizing the graphs, the formalism is unable to distinguish between strings containing the same terminal symbols (basic expressions) in different order. The author does point this out, stating (on p. 154) that, if the ρ -sets encoded word order (which is what distinguishes cross-serial dependencies from center embedding and tail-recursion), they would differ between the Dutch example and its Turkish and English counterparts; however, this is the path taken in the present work (which we take to imply that in the current version of the theory they are indeed the same *modulo* the content of the basic expressions).

Without further clarification, this discussion seems to fall prey to the subset fallacy. While the formalism will certainly produce representations of strings with crossing dependencies, it cannot describe *just these* strings to the exclusion of others. Citing [Mohri, Sproat 2006: 434], ‘if a language L contains a subset L' which is at position P in the Chomsky hierarchy, in general this tells us nothing about the position P of L in the hierarchy.’ For example, $\{a, b, c\}^*$, the language of all strings over the alphabet $\{a, b, c\}$, does contain the mildly context-sensitive language of strings $\{a^n b^n c^n \mid n \geq 0\}$ as a subset. However, $\{a, b, c\}^*$ itself is not a mildly context-sensitive language, since it contains all the other strings that do not belong to this mildly-context-sensitive set.

At the same time, this formalism seems capable of expressing some complex patterns. For instance, it should be able to generate the context-free language $\{w \mid w \in \{a, b\}^* \ \& \ |w|_a = |w|_b \geq 0\}$ (all strings consisting of an equal number of a ’s and b ’s), but not the context-free language

$\{a^n b^n \mid n \geq 0\}$ (all strings consisting of a string of a 's immediately followed by a string of b 's of equal length). With this in mind, another path the author could have taken — and one that could align well with his desire to remain agnostic on the issue of linearization — would be to state that languages generated by this formalism are not sets of strings but sets of *multi-sets* of terminals. There exists work in formal grammar theory introducing a Chomsky-like hierarchy for sets of multisets [Kudlek et al. 2001]; examining it with respect to natural language syntax could potentially lead to an interesting direction of research.

One can certainly see the advantages of a graph-based approach that allows for multidominance (including multiple roots) and parallel edges; and certain traditionally challenging phenomena can be described in these terms quite naturally. To give a few examples, graphs are a good fit for representing constructions such as anchored parentheticals and amalgams, which involve otherwise apparently independent clauses sharing a common element. The node addressing system treats coindexation simply as node identity, allowing the author to recast binding in graph-theoretic terms. This solution deals with the basic examples rather elegantly — for instance, reflexives boil down to multiple arcs of the same predicate terminating at the same node. While it is not immediately clear how other cases (e.g., reciprocals, split antecedents) would work, this problem can probably be overcome with a more nuanced definition of what the nodes in a graph correspond to; the possibility of altering “the level of granularity at which nodes are defined (basic expressions, features or feature bundles, roots, etc.)” is mentioned by the author on p. 83.

Moving on to a critical assessment of the model, it should be emphasized that we have before us a partial model of the syntactic structure of a sentence, reflecting only those of its properties that follow from the definition of Relational Grammar: “A clause consists of a network of grammatical relations” [Perlmutter, Postal 1983: 9]. Accordingly, we should not expect to learn anything new about the structure of nominal constituents, prepositional phrases, or adjectival phrases. Incomplete representation of sentence structure leads to the fact that restrictions on representation in such an impoverished structure can be formulated only for a certain range of phenomena and, in general, are unable to distinguish grammatical sentences from ungrammatical ones. For example, the representation of (2a) looks like (2b). At the same time, this formalism is apparently unable to express the difference between grammatical (2a) and ungrammatical (2c).

- (2) a. What did Sue say?
 b. $\rho = \langle (\text{what, say}), (\text{say, Sue}), (\text{say, what}) \rangle$
 c. *What said Sue?

The issues discussed in the monograph relate to two important areas: the establishment of hierarchical grammatical relations within a clause (that is, an elementary tree) and ways of combining elementary trees into a derived structure corresponding to a complex sentence.

In both cases the model is vulnerable to criticism. In the structure of an elementary tree, we observe the consistent use of the direction of the semantic relationship between predicate and argument to determine the direction of syntactic dependence. However, since the time of Otto Jespersen [1924] a distinction has been made between syntactic dependencies that coincide in direction with semantic ones (nexus) and those opposite in direction to semantic ones (junction). Models based on semantic relationships between concepts use special rules that “reverse” the direction of dependency in syntagmas with junction at the syntactic level. The lack of distinction between nexus and junction leads to expressions with different meanings receiving the same representation (and, apparently, the same interpretation):

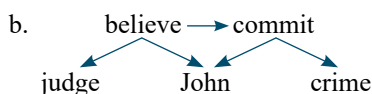
- (3) a. I like my tea strong.
 b. I like my strong tea.
 (4) $\rho = \langle (\text{like, I}), (\text{like, tea}), (\text{strong, tea}) \rangle$ (for (3a) (p. 195) and (3b))

The author explicitly rejects labeling the arrows (p. 162), suggesting that their interpretation is uniquely determined by the order of the relations listed in the ρ -set. However, it should be emphasized that the presence of an edge (v_i, v_j) between the vertices of the graph is used to encode relations of different levels: semantic relations (that is, hierarchies at the level of the thematic structure), grammatical relations (that is, hierarchies at the level of syntactic functions), logical relations (that is, the relationship between the antecedent and the anaphoric element, for example, in the identification of a *wh*-pronoun and a resumptive pronoun, or between an operator and the head of its nuclear scope, such as a *wh*-pronoun and a predicate). The need to distinguish between these hierarchies and the possibility of discrepancies in the ranks of arguments at different levels has been widely discussed in the syntactic literature. Some models, such as Lexical functional grammar (LFG), map multiple representations to linguistic expressions and formulate rules connecting argument structure, functional structure, and constituent structure. In other models, such as generative grammar, hierarchical relationships can be established within extended verb projections of different levels (vP, TP, CP). In the monograph under review, there is no articulated distinction between the levels to which specific relations belong, which, in our opinion, will cause serious problems when modeling phenomena oriented towards different hierarchies (cf. [Manning 1996] on the need to distinguish between the subject of argument structure and the subject of grammatical relations).

Furthermore, even if all edges in an elementary graph represent grammatical relations, their ordered sets do not solve the problem of identifying a particular relation. In the fundamental absence of null and syncategorematic units in the representation, the first relation in the tuple is inevitably interpreted as the subject, and the second as the direct object. This makes it difficult to analyze impersonal constructions (such as lenient passives), intransitive verbs with two arguments, and many other cases where a gap occurs in the hierarchy of grammatical relations (SU > DO > IO > Oblique). All of the above leads to the idea that one of the basic dichotomies of the monograph — that between transformations that change and those that do not change the grammatical relations — turns out to be an epiphenomenon of the adopted formalism.

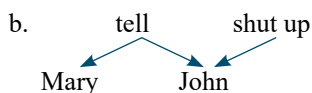
Let us now turn to derived structures. Biclausal constructions are formed by combining elementary graphs; possible methods of unification are through a root node (nesting one tree into another) and through a non-root node (in which case a multirooted graph is formed). According to the author, the ways of connecting graphs distinguish between English infinitive constructions and control and raising (p. 203). Thus, the ECM-construction in (5a) receives the representation (5b) which includes a relation between the matrix and nested predicate (see also the ordered ρ -set in (5c)), while the control configuration (6a) is represented as (6b)–(6c), where the dependent and main clauses are connected only through a common argument.

- (5) a. The judge believed John to have committed the crime.



- c. $\rho_{\text{derived}} = \langle (\text{believe, judge}), (\text{believe, John}), (\text{believe, commit}), (\text{commit, John}), (\text{commit, crime}) \rangle$

- (6) a. Mary told John to shut up.



- c. $\rho_{\text{derived}} = \langle (\text{tell, Mary}), (\text{tell, John}), (\text{shut up, John}) \rangle$

Without disputing the validity of such a solution (it seems to us that the dependent clause is an argument of the matrix predicate both in the case of control and in the ECM-configuration), we note that the proposed model does not in any way limit the configurations in which predicates can have a common argument, and does not place any limitations on which arguments can be shared. It appears that this model can represent any combination of control/raising and shared argument, non-local control, hyper-hyperraising, and other ungrammatical configurations that are consistently constrained, for example, in generative models of control and raising. Note also that when connecting graphs by means of a non-root node (as in (6b)), it is impossible to read the status of the main vs. dependent clause off the representation, so (5b) could also be the graph of the completely ungrammatical *John_i shut up Mary to tell e_i*. From a broader perspective, the model seems capable of representing both natural language expressions and those that are not characteristic of natural languages. Constraints on graphs are intended to distinguish between them, but there are surprisingly few such constraints in the monograph, and for the complementation constructions under consideration they are absent.

The scope of the review does not allow us to dwell in detail on all the grammatical phenomena considered in the monograph, which are quite diverse and include, for example, clitic climbing, *wh*-movement, restrictive and appositive relative constructions, and anaphora. However, despite a rather sophisticated analysis and the involvement of a huge amount of relevant literature from a variety of theoretical frameworks, one cannot avoid the impression that the model being developed does not allow us to learn anything fundamentally new about language. The monograph outlines a number of interesting ideas but stops just short of developing them, highlighting a (legitimately hard) challenge inevitably faced by any newly introduced linguistic formalism: the need for a balance between compatibility with different competing analyses vs. making specific decisions to enable the theory to make specific predictions. Traditional questions of the research program of formal linguistics about the human language ability, the acquisition and origin of language, the universal and specific in language, the parameters and limits of linguistic variation can hardly be adequately posed within the framework of the proposed model. The author explicitly rejects this view of the theory of language (pp. 437–438) and proposes to model the relational structure of linguistic expressions *an und für sich*. One can only hope that this research program will find its followers.

REFERENCES

- Chomsky 2000 — Chomsky N. Minimalist inquiries: The framework. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Martin R., Michaels D., Uriagereka J. (eds.). Cambridge (MA): MIT Press, 2000, 89–156.
- Chomsky 2001 — Chomsky N. Derivation by phase. *Ken Hale: A life in language*. Kenstowicz M. (ed.). Cambridge (MA): MIT Press, 2001, 1–52.
- Chomsky 2008 — Chomsky N. On phases. *Foundational issues in linguistic theory*. Freidin R., Otero C. P., Zubizarreta M.-L. (eds.). Cambridge (MA): MIT Press, 2008, 133–166.
- Hudson 1984 — Hudson R. *Word grammar*. Oxford: Basil Blackwell, 1984.
- Hudson 2021 — Hudson R. Word grammar. *The Routledge handbook of cognitive linguistics*. Xu W., Taylor J. R. (eds.). New York; London: Routledge, 2021, 111–126.
- Jespersen 1924 — Jespersen O. *The philosophy of grammar*. New York: Allen & Unwin, 1924.
- Johnson, Postal 1980 — Johnson D., Postal P. *Arc pair grammar*. Princeton: Princeton Univ. Press, 1980.
- Joshi 1985 — Joshi A. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? *Natural language parsing*. Dowty D., Karttunen L., Zwicky A. (eds.). Cambridge: Cambridge Univ. Press, 1985, 206–250.
- Kudlek et al. 2001 — Kudlek M., Martin-Vide C., Păun G. Toward a formal macroset theory. *Multiset processing: Mathematical, computer science, and molecular computing point of view*. Calude C. S., Păun Gh., Rozenberg G., Salomaa A. (eds.). Berlin; Heidelberg: Springer, 2001, 123–133.
- Manning 1996 — Manning Ch. *Ergativity: Argument structure and grammatical relations*. Stanford: CSLI Publications, 1996.

- Mel'čuk 1974 — Mel'čuk I. A. *Opyt teorii lingvisticheskikh modelei «Smysl ⇔ Tekst»* [Toward a theory of “Meaning ⇔ Text” linguistic models]. Moscow: Nauka, 1974.
- Mel'čuk 1988 — Mel'čuk I. *Dependency syntax: Theory and practice*. Albany: State Univ. of New York Press, 1988.
- Mohri, Sproat 2006 — Mohri M., Sproat R. On a common fallacy in computational linguistics. *SKY Journal of Linguistics*, 2006, 19: 432–439.
- Perlmutter (ed.) 1983 — Perlmutter D. (ed.). *Studies in relational grammar I*. Chicago: Univ. of Chicago Press, 1983.
- Perlmutter, Postal 1983 — Perlmutter D., Postal P. M. Some proposed laws of basic clause structure. *Studies in Relational Grammar I*. Perlmutter D. (ed.) Chicago: Univ. of Chicago Press, 1983, 81–128.
- Postal 2010 — Postal P. *Edge-based clausal syntax*. Cambridge (MA): MIT Press, 2010.
- Shieber 1985 — Shieber S. Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 1985, 8(3): 333–343.
- Svenonius 2004 — Svenonius P. On the edge. *Peripheries*. Adger D., de Cat C., Tsoulas G. (eds.). Dordrecht: Kluwer, 2004, 259–287.

Получено / received 14.04.2024

Принято / accepted 21.05.2024

ВОПРОСЫ ЯЗЫКОЗНАНИЯ
научный журнал Российской академии наук
(свидетельство о СМИ ПИ № ФС77-77284 от 10.12.2019 г.)

Оригинал-макет подготовлен С. С. Белоусовым

Адрес редакции: 119019, Москва, ул. Волхонка, 18/2,
Институт русского языка им. В. В. Виноградова РАН, редакция журнала «Вопросы языкознания»,
тел.: +7 495 637-25-16, e-mail: voprosy@mail.ru

Подписано к печати 99.99.2024. Дата выхода в свет 99.99.2024. Формат 70×100 1/16.
Усл. печ. л.: 99,9. Уч.-изд. л. 99,9. Тираж 999 экз. Заказ 999. Цена свободная

Учредители: Российская академия наук, Институт русского языка им. В. В. Виноградова РАН

Издатель: Российская академия наук, 119991 Москва, Ленинский просп., 14

Исполнитель по контракту № 4У-ЕП-03 -24 ФГБУ «Издательство «Наука»
121099, г. Москва, Шубинский пер., д. 6, стр. 1

Отпечатано в ФГБУ «Издательство «Наука»
121099, г. Москва, Шубинский пер., д. 6, стр. 1