



*А.В. Циммерлинг.
Фото М.Н. Лабовского*

Книга содержит работы по теоретической и исторической грамматике. Они разбиты на шесть разделов, носящие название «Теория. Метод. Доказательность», «Диахрония. Синхрония. Динамика», «Семантика в зеркале грамматики», «Делить на два. Неканонические подлежащие», «Порядок слов. Теория и типология» и «Текст и язык». Обсуждается соотношение частной и общей лингвистики, вклад первой во вторую. Обосновывается принцип множественности лингвистических описаний и приоритет аспективных описаний над интегральными, претендующими на полный охват некоторой области языковых явлений. Лингвистика является наукой в той мере, в которой она позволяет строить модели, предсказывающее распределение языковых черт на некотором эмпирическом материале. Если предиктивная модель содержит лишнюю информацию, предсказание будет неточным, что можно установить, проверяя модель на некотором множестве языков мира.

ISBN 978-5-4469-1792-1



9 785446 917921

От интегрального к аспективному

Антон Циммерлинг

Антон Циммерлинг

От интегрального
к аспективному



Антон Циммерлинг

**От интегрального
к аспектвному**



**Нестор-История
Москва — Санкт-Петербург
2021**

УДК 81-139

ББК 382-5

Ц613

Исследование выполнено при поддержке проекта Российского научного фонда, проект РНФ 18-18-00462, реализуемом в Государственном институте русского языка имени А. С.Пушкина.

Рекомендовано к печати Ученым советом Института языкознания РАН и Ученым советом Государственного института русского языка имени А. С.Пушкина.

Рецензенты:

доктор филологических наук, академик РАН *В. А. Плузган*;
доктор филологических наук, член-корреспондент РАН *А. В. Дыбо*

Циммерлинг А. В.

Ц613 От интегрального к аспективному / А. В. Циммерлинг. — М. ; СПб. : Нестор-История, 2021. — 652 с.

ISBN 978-5-4469-1792-1

DOI 10.31754/nestor4469-1792-1

Книга содержит работы по теоретической и исторической грамматике. Они разбиты на шесть разделов, носящие название «Теория. Метод. Доказательность», «Диахрония. Синхрония. Динамика», «Семантика в зеркале грамматики», «Делить на два. Неканонические подлежащие», «Порядок слов. Теория и типология» и «Текст и язык». Обсуждается соотношение частной и общей лингвистики, вклад первой во вторую. Обосновывается принцип множественности лингвистических описаний и приоритет аспективных описаний над интегральными, претендующими на полный охват некоторой области языковых явлений. Лингвистика является наукой в той мере, в которой она позволяет строить модели, предсказывающее распределение языковых черт на некотором эмпирическом материале. Если предиктивная модель содержит лишнюю информацию, предсказание будет неточным, что можно установить, проверяя модель на некотором множестве языков мира.

УДК 81-139

ББК 382-5

ISBN 978-5-4469-1792-1



9 785446 191792 1

© А. В. Циммерлинг, 2021

© Издательство «Нестор-История», 2021

СОДЕРЖАНИЕ

Вместо предисловия	4
<i>I. Теория. Метод. Доказательность</i>	
[1] От интегрального к аспектвному (2019)	7
[2] Что и зачем. Научная статья. Научные статьи (2019)	39
<i>II. Диахрония. Синхрония. Динамика</i>	
[3] Die unpersönlichen Satzmodelle in der altisländischen Sprache (1989)	67
[4] Между синхронией и диахронией. Просодические противопоставления в скандинавских языках (1999)	82
[5] Раннегерманское ударение: Фонетика и компаративистика (2006)	99
[6] Кому <i>проспати заутреня</i> и <i>ловити рыба</i> ? Именительный падеж дополнения инфинитива как предмет грамматического описания (2018–2019)	148
[7] Два диалекта русской грамматики: корпусные данные и модель (2018)	184
<i>III. Семантика в зеркале грамматики</i>	
[8] История одной полемики (1998)	203
[9] Субъект состояния и субъект оценки (1999)	236
[10] Русские предикативы в зеркале эксперимента и корпусной грамматики (2017)	243
[11] Имперсональные конструкции и дативно-предикативные структуры в русском языке (2017–2018)	261
[12] Эпистемические предикаты и ментальные состояния (2017)	288
[13] Обладать и быть рядом (1999)	308
[14] Логика прогноза (1997)	319
<i>IV. Делить на два. Неканонические подлежащие</i>	
[15] Dative Subjects and Semi-Expletive Pronouns in Russian (2007–2008)	331
[16] Неканонические подлежащие в русском языке (2011)	346
[17] Sentential Arguments and Event Structure (2014)	369
[18] Корпусная грамматика количественных групп в русском языке (2019)	385
[19] Топикальность и неканонические подлежащие в русском языке (2019)	403
<i>V. Порядок слов. Теория и типология</i>	
[20] Локальные и глобальные правила в синтаксисе (2008)	415
[21] Системы порядка слов в славянских языках (2012)	435
[22] A unified analysis of clitic clusters in world's languages (2012)	478
[23] Clitic templates and discourse marker <i>ti</i> in Old Czech (2013)	495
[24] Parametrizing Verb-Second Languages and Clitic-Second Languages (2015)	507
[25] 1P orders in 2P languages (2015)	519
<i>VI. Текст и язык</i>	
[26] Порядок, метр, прием (2003)	543
[27] The text-building function of names and nicknames in «Sverris saga» and «Böglunga sogur» (1994)	554
[28] Бремя чужих слов (1998)	566
Литература	569
Тематический указатель	604
English summaries	635

ВМЕСТО ПРЕДИСЛОВИЯ

Эта книга содержит работы, написанные в течение тридцати лет, с 1989 по 2019 г. Они разбиты на шесть разделов, носящих названия «Теория. Метод. Доказательность», «Диахрония. Синхрония. Динамика», «Семантика в зеркале грамматики», «Делить на два. Неканонические подлежащие», «Порядок слов. Теория и типология» и «Текст и язык». Первый раздел написан хронологически самым последним. Главная тема книги — соотношение частной и общей лингвистики, вклад первой во вторую. Обосновываются принцип множественности лингвистических описаний и приоритет аспективных описаний над интегральными, претендующими на полный охват некоторой области языковых явлений. Вторая тема книги — обоснование роли статьи как жанра лингвистического исследования столь же важного, как монография и квалификационная работа. Последние по условию должны стремиться к полному охвату явления. Задача статьи — точно выделить сам объект и поместить его в правильный контекст. Лингвистика является наукой в той мере, в которой она позволяет строить модели, предсказывающие распределение языковых черт на некотором эмпирическом материале. Если объект выделен неверно, ничего предсказать не удастся. Если предиктивная модель содержит лишнюю информацию, предсказание будет неточным, что можно установить, параметризуя языковое многообразие и проверяя модель на некотором множестве языков мира. Все интегральные концепции языка, претендующие на полный охват языковых явлений, на деле хорошо описывают лишь малую их часть. Преодолеть этот изъян можно, применяя принцип множественности описания, аналогичный принципу дополнительности Бора в естественных науках. Проверка того, действительно ли лингвистика является естественной наукой, зависит от того, в какой мере аспективные исследования языка поддаются гармонизации. Современный лингвист должен уметь строить предиктивные модели, которые хорошо работают на некотором материале и не зависят от выбора какой-то одной глобальной теории языка. Это налагает на него обязательство очертить контуры набора теоретических объяснений, с которыми совместимо предложенное им объяснение. Если угодно, он должен очертить контуры возможных теорий более общей области явлений сам. И это, при всех трудностях, более перспективный путь, чем подстраиваться под догматические допущения существующих интегральных теорий языка.

Антон Циммерлинг, январь 2020 г.

I.
ТЕОРИЯ. МЕТОД.
ДОКАЗАТЕЛЬНОСТЬ

[1] ОТ ИНТЕГРАЛЬНОГО К АСПЕКТИВНОМУ

Аннотация. Обосновывается приоритет принципа множественности лингвистических описаний над описанием с позиций интегральных теорий языка, опирающихся на строго монистические принципы. Сложность естественного языка такова, что интегральные теории хорошо описывают лишь часть явлений, на объяснение которых они претендуют, и задают стандарт огрубления прочих языковых явлений. Поэтому точные аспективные описания, совместимые с разными глобальными теориями, должны считаться приоритетными. Критерием точности является возможность построить предиктивную модель, предсказывающую распределение языковых черт на некотором эмпирическом материале. Эффективность предиктивных моделей повышается при параметризации языкового разнообразия и проверке модели на некотором множестве языков или идиомов. Основной вклад интегральных теорий в науку состоит не в доказательстве заявленных деклараций о природе языка, а в разработке аппарата описания отдельных явлений языка. Попытки его распространения на другие явления языка снижают точность предиктивных лингвистических моделей, поскольку в саму формулировку этих моделей заложена избыточная информация. Лингвисты используют два типа модельных понятий, первый из которых соответствует логико-математическому, а второй — естественно-научному пониманию моделей. Модельные₁ понятия — метаязыковые категории, являющиеся частью некоторого согласованного набора категорий описания. В своей основе модельные₁ понятия универсальны. Модельные₂ понятия — гипотезы, проверяемые на некотором материале. Вопреки делавшимся утверждениям, модельные₁ понятия в полном объеме неустранимы из описания, но между модельными₁ понятиями и модельными₂ понятиями не всегда имеется соответствие. Глава иллюстрируется аспективным анализом двух проблем общей и частной грамматики — соотношения нулевой и ненулевой форм связки «быть» в русском языке и принципов упорядочения энклитик в классе языков мира с законом Ваккернагеля. Эти проблемы могут, и в большинстве случаев должны, рассматриваться как автономные, однако диахронический анализ на материале истории русского языка показывает, что они при определенных условиях могут рассматриваться и как связанные друг с другом.

Ключевые слова: естественный язык, предиктивные модели, интегральные теории, параметризация, монизм, множественность описания, принцип дополнительности, общая грамматика, частная грамматика.

I

Один выдающийся ученый (А. Гардинер) в свое время начал книгу с подзаголовком «Полемическое эссе» словами о том, что «любая дискуссия должна начинаться с пунктов согласия» [Gardiner 1940: 5]. Его современник утверждал, что «прежде, чем объединяться, и для того, чтобы объединиться, мы должны сначала решительно и определенно размежеваться» [Ленин 1900]. В известном смысле это одно и то же. Правда, первая формулировка характеризует научную процедуру, а вторая — условие взаимодействия людей, которых объединяют и противопоставляют

разные идеи, цели и интересы. Хорошо это или плохо, но прогресс науки в большинстве случаев связан не с прямым взаимодействием научных идей, а с коммуникацией ученых. Формально это всегда так, ведь наукой занимаются люди. Но применительно к наиболее фундаментальным результатам, будь то открытие Америки или нового химического элемента, закон всемирного тяготения или доказательство теоремы Ферма, можно считать, что в обществе современного типа, основанном на открытом доступе к информации, они будут быстро востребованы и осмыслены, независимо от особенностей коммуникации между учеными. Воспринятый обществом, в котором мы живем, идеал позитивного научного знания, благодаря которому общество финансирует науку, состоит в том, что таковы вообще все научные результаты в разных областях знания, и что даже самые частные и специальные исследования вносят вклад в построение общенаучного здания, предлагая если не открытие Америки, то хотя бы правильно сформулированные наблюдения, которые при надлежащем осмыслении могут в дальнейшем стать основой более общих гипотез.

Увы! Большинство работ в науке о языке не таковы. Даже если отвлечься от сопутствующих факторов, таких как непонятность проблемы для широкой публики, включая лиц, осваивающих и даже преподающих некоторый язык — английский, русский, древнегреческий, корейский и т. д., наличия симуляций с применением наукообразной терминологии, валидность многих добросовестных работ под вопросом. Остаются сомнения в том, что проблема поставлена правильно, объект выделен верно, а методы описания адекватны. Именно это, наряду с обстоятельствами, которые характеризуют внутреннюю организацию сообщества лингвистов: зависимость отклика от языка и места публикации, харизмы автора, его принадлежности к той или иной школе и коллективу ученых, — и является причиной того, что большинство лингвистических работ игнорируется коллегами.

В качестве иллюстрации обсудим частный вопрос описания языка, имеющего литературную традицию и обеспеченного лингвистическими ресурсами в виде грамматических очерков, словарей и корпусов текста. Есть ли в русском языке два разных глагола со значением «быть» или только один? Этому вопросу повезло (или не повезло) больше других: разные ответы на него запечатлены в нормативных описаниях русского языка, в том числе — в версиях школьных и вузовских курсов, ср. [Лекант 1974: 108; 2001; 2010; Ильенко 2016]. По одной точке зрения, впервые высказанной в знаменитой работе [Щерба 1928], связка ‘быть’, выражаемая нулевой формой (∅) в настоящем времени и ненулевыми формами *был* и *будет* в прошедшем и будущем времени, где в некоторых конструкциях она требует или допускает форму творительного предикативного падежа существительного в позиции присвязочного члена, ср. *Лев — ∅ автор ряда романов / был автором ряда романов, Маше ∅ холодно / было холодно, Директором будет Вася*, и глагол существования ‘быть’, реализующийся в настоящем времени в виде формы ‘есть’, — разные слова, чуть ли не разные части речи, при этом связка ‘быть’ по семантическим основаниям выносится из класса глагольных слов в особую часть речи, так как она якобы не имеет значения существования и лексического значения

вообще [Виноградов 1972: 529]. Тезис о полном отсутствии у связки лексического значения восходит к работе [Пешковский 1928: 303]. Эта точка зрения опирается на неортодоксальную и, возможно, ошибочную концепцию частей речи, ср. критику в [Шапиро 1955; Апресян 1985], и в любом случае на не разделяемые значительной частью лингвистов соглашения том, что такое часть речи. Альтернативная точка зрения, последовательно выраженная в [Апресян 1996: 527–533], признает в русском языке только один глагол ‘быть’, имеющий связочные и несвязочные употребления, причем в некоторых контекстах (ср. *негде спать ~ было негде спать, у меня есть деньги ~ деньги у меня*), как утверждается, происходит «перетекание» несвязочных употреблений в связочные [там же: 530–532]. Эта точка зрения тоже опирается на предварительные соглашения о том, что такое «класс слов», «синтаксический признак» и каковы критерии тождества «лексемы», т. е. слова как единицы словаря.

Компромиссная точка зрения состоит в том, что связка ‘быть’ (Ø / был / будет) и полнозначный глагол ‘быть’ (*есть* / был / будет) — разные слова, парадигматически противопоставленные в настоящем времени. Ю. Д. Апресян [там же: 515] усматривает ее в схеме толкования в «Малом академическом словаре» под редакцией А. П. Евгеньевой [МАС]. То же представление отражено в статье Э. Бенвениста о глаголах ‘быть’ и ‘иметь’ в языках мира, где о противопоставлении «нулевая морфема vs *есть*», выражающем в русском языке контраст между связочным и полнозначным употреблением глагола ‘быть’, говорится как об установленном факте [Benveniste 1960; 1966; Бенвенист 1974: 209], а также в позднейших типологических работах, где русский язык со ссылкой на Э. Бенвениста помещается в класс языков мира с нулевой связкой в настоящем времени, куда также входят венгерский, арабский и многие австронезийские языки [Stassen 1994; Pustet 2003]. Утверждение о закреплённом в русской грамматике и словаре противопоставлении связки *быть* (реализующейся в настоящем времени как Ø) и глагола существования *быть* (реализующегося в настоящем времени в виде формы *есть*) делается также в статье [Тестелец 2008: 774]. Я. Г. Тестелец доказывает, что статус связки, в том числе нулевой, проясняется формальными грамматиками составляющих, так как связка является (**есть*) вершиной группы особого типа, так называемой «малой клаузы». Такой анализ требует дополнительных соглашений о «структуре составляющих» и о том, что такое «клауза» и «малая клауза». Впрочем, два первых понятия, а возможно, и третье, могут быть причислены к фундаментально значимым. На практике это значит, что они признаются необходимыми более чем в одной версии теории синтаксиса. Но есть и эмпирическая трудность: факты русского языка соответствуют гипотезе о парадигматическом противопоставлении двух глаголов Ø / был / будет vs *есть* / был / будет лишь частично.

Во-первых, как отмечено в [Апресян 1996] и в [Дымарский 2018], нулевая форма ‘быть’ может реализоваться в ряде экзистенциальных и посессивных конструкций, ср. *А у нас Ø в квартире газ*. Сторонник гипотезы о двух глаголах ‘быть’ будет вынужден либо постулировать ad hoc для соответствующих контекстов

«нейтрализацию», стирание противопоставления связка vs полнозначный глагол¹, либо, наоборот, постулировать ad hoc для предложений в парах типа *А у нас в квартире ∅ газ vs А у нас в квартире **есть** газ* разные значения во всех случаях употребления. Агностик скажет, что если в русском языке и есть два разных глагола ‘быть’, контексты типа *А у нас в квартире ∅ / **есть** газ* не годятся для их диагностики, поскольку в них допустимы оба варианта без явного различия в значении². Во-вторых, у формы ‘есть’ есть связочные употребления, ср. *Коммунизм **есть** электрификация плюс Советская власть, Вася **и есть** автор этой работы, Вы, батюшка, **и есть** наш лидер* и т. п. В контексте <X **есть** Y> связка *есть* имеет значение 3 л., а в контексте <X **и есть** Y> комбинируется с субъектными местоимениями в 1–3 л. в ед. ч. и мн. ч. (*Я, ты, Вы, он, она, оно, мы, вы, они и есть Y*).

Если изложенные данные верны, возобладавшая в русистике и за ее пределами точка зрения о том, что в русском языке современной эпохи (XX–XXI вв.) выделяются два разных глагола ‘быть’, которым якобы соответствуют разные парадигмы, искажает факты. Хотя эта точка зрения проникла в обзорные работы по типологии и теории грамматики, она — лингвистический миф, который должен быть отвергнут в претендующих на адекватность описаниях русской глагольной лексики. Отдельный вопрос состоит в том, в какой мере искажения фактов в исходных описаниях уменьшают валидность обобщений типолога. Для помещения русского языка в класс языков с нулевой связкой, видимо, достаточно констатации того, что в одних предложениях русского языка нулевую связку нельзя заменить на ненулевую, а в других предложениях она является одним из вариантов. Правда, типологу и его читателям придется допустить, что исходные наблюдения над дистрибуцией связок в русском и других языках верны, а понятие нулевой связки универсально, иначе все разбиение языков мира на классы может пойти прахом. Но *есть* (*∅) другой, не менее важный вопрос: почему, несмотря на то, что лексикографический подход [Апресян 1996], с каталогизацией всех употреблений глагола ‘быть’ в русском языке XX–XXI вв., выявляет ложность гипотезы о двух лексемах ‘быть’, эта гипотеза продолжает использоваться? Ответ (**есть*) на поверхности: лексикографический подход, в терминах Ю. Д. Апресяна — «портретирование» глагола ‘быть’, не очерпывает всех проблем, встающих в связи с дистрибуцией синтаксических элементов ‘∅^{БЫТЬ}’ и ‘*есть*’. Для теории грамматики важно, что на материале русского

¹ Без дополнительного обоснования, т. е. фактически — без разработки общей теории грамматики на основе представления о релевантных признаках грамматических форм и конструкций, устроенных аналогично дифференциальным признакам фонем, неочевидно, применимо ли фонологическое понятие «нейтрализации» [Trubetzkoy 1939] к грамматике или же речь идет о метафоре.

² Такой же подход может быть распространен на другие случаи конкуренции грамматических форм и конструкций, в том числе — на противопоставление форм им. п. и тв. п. в предикативной позиции. Ср. *он был дураком / умным человеком vs он был дурак / умный человек*. Разброс мнений о том, синонимичны ли предложения в подобных парах [Пешковский 1928: 288; Гиро-Вебер 2007; Nichols 1981], отражает не только различие вкусов лингвистов (которое вполне естественно в ситуации внутриязыкового варьирования), но и желание последних расставлять знаки приемлемости предложений и приписывать им значения ad hoc в соответствии с предсказаниями отстаиваемой теории, т. е. эффект самовнушения (priming effect).

языка прослеживается (= *есть*) устойчивая корреляция между предположительно универсальными логико-семантическими типами предложения (бытийные, possessивные, локативные предложения, предложения тождества, идентификации, характеристики, приписывания признака) и исторически сложившимся в русском языке к XX в. распределением нулевой и ненулевой форм глагола 'быть', что показано в работе [Арутюнова 1983]. Для изучения данной корреляции можно, а порой и нужно, отвлечься от полного каталога употреблений лексемы 'быть' и ограничиться 3–5 контекстами, ср. обзор некоторых контекстов в [Янко 2001: 322–325]. Словарный и грамматический подходы к описанию употреблений глагола 'быть' тем самым дополняют друг друга. Каждый из них «правилен», но важно понять, какие исходные данные теория получает на входе и какие ограничения использованный метод налагает на интерпретацию результатов на выходе³. Принцип дополнительности и множественности описания применяется в лингвистике редко, среди исключений см. [Гак 1998; Крылов 2004], что, возможно, отражает небольшой стаж современной лингвистики как автономной научной дисциплины.

Освоенное лингвистикой модельное понятие *нулевой связки*, выдвинутое А. М. Пешковским [Пешковский 1928: 303], опирается не только на специфический материал русского языка, но на общие соображения о связи между означаемым и означающим: эта модель, позже развитая Р. О. Якобсоном и И. А. Мельчуком [Якобсон 1985: 225–227; Mel'čuk 1995: 169–179], обосновывает существование нулевых лексем, аналогичных выделенным ранее нулевым словоизменительным морфемам. Противники гипотезы о двух глаголах 'быть' в русском языке, насколько нам известно, не отвергали представление о нулевой связке, в то время как для первых сторонников гипотезы о двух глаголах 'быть' — А. М. Пешковского и Л. В. Щербы — эта гипотеза была дополнительным подтверждением реальности нулевой связки: чем в большее количество противопоставлений вовлечена постулированная категория в конкретном языке, тем больше оснований верить в ее существование и в позитивную ценность описания, оперирующего понятием 'нулевого знака'. Тем самым дискуссия о двух глаголах 'быть' в русском языке действительно опирается на «пункты согласия» в смысле А. Гардинера — в том, что а) связочные употребления поддаются диагностике и б) применение гипотезы о нулевой связке 'быть' к материалу русского языка оправдано.

II

Вернемся к общей проблематике, которую вновь будем рассматривать на частном примере. Современная лингвистика претендует на роль *естественной* науки.

³ Некоторые соображения в данном конкретном случае указывают на то, что лексикографический подход к определению числа глаголов 'быть' заслуживает предпочтения, поскольку вопрос о числе лексем и критериях тождества должны решать лексикографы, а не синтаксисты. Однако понятие «лексемы» и способ толкования тоже может заключать в себе грамматическую информацию, поэтому сопоставление разных подходов к тому, чем является грамматика с точки зрения лексикографа и чем является словарь с точки зрения синтаксиста, необходимо.

Естественные науки предсказывают факты на основе предиктивных моделей. В контексте случившегося в XX в. окончательного размежевания лингвистики со смежной наукой, *филологией*, эта претензия правомерна. Предметом интереса филологии являются *тексты*, понимаемые как фиксированные вербальные произведения, а также классификация жанров текста. Предсказание уникального невозможно, а утверждение о месте конкретного текста в системе жанров нельзя считать предсказанием. Напротив, лингвистика притязает на выявление автоматически реализуемых механизмов, проявляющихся независимо от конкретного языка и конкретного говорящего / пишущего. Предмет лингвистики — все уровни и аспекты существования естественного языка, где такие механизмы могут быть выявлены — от фонетики до семантики, от внутри- и межъязыкового варьирования до факторов языковой эволюции и филиации языков.

Эффективность предсказания зависит от избранной модели и от сопротивления материала. Противопоставление *описания* фактов их *объяснению*, т. е. предсказанию на основе правил, корректно, но языковые факты не даны в откровении и не существуют в отрыве от теорий и модельных понятий. Они, кроме того, слабо рефлектируются. Распределение внешне невыраженной (\emptyset) и выраженной (*есть*) форм глагола 'быть' доступно любому носителю русского языка, но разные решения профессиональных лингвистов мотивированы тем, что они преследуют разные цели — обосновать важность противопоставления связочных vs несвязочных предложений, экзистенциальных vs неэкзистенциальных предложений, дать лексикографический портрет глагола 'быть' и т. п. — и по-разному трактуют такие модельные понятия, как «лексема», «парадигма форм слов», «грамматическая конструкция», «часть речи», «нулевой знак». Такого «факта», как объединение элементов \emptyset и *есть* в парадигме одного и того же русского слова или отнесение их к разным парадигмам, вне процедуры анализа нет.

Слово «модель» в обыденном языке многозначно. Слово «модель» в языке науки должно стремиться к однозначности. Но научные тексты пишутся людьми, использующими элементы обыденного языка. Остаются подводные камни. В тривиальном смысле модельной₁ можно признать любую категорию описания, являющуюся частью согласованной системы метаязыковых понятий⁴. Упомянутые выше понятия *нулевого знака* и *лексем* — модельные₁, поскольку они связаны с другими понятиями, такими, как «знак», «означающее», «значение», «парадигма». В специальном смысле естественных наук модель₂ — это тестируемая на конкретном материале гипотеза о соответствии наблюдаемого образцу. Между моделью₁

⁴ Уместно потребовать, чтобы в описаниях разных языков мира одна часть метаязыковых категорий была универсальными понятиями, т. е. понятиями, потенциально применимыми ко всем языкам мира, где есть соответствующие явления, а прочие понятия были производны от первых. Слово 'потенциальный' означает, что некоторые явления, например, морфологический падеж, нулевые морфемы, нулевые лексемы или энклитики второй позиции могут отсутствовать в конкретном языке L, что не лишает их универсальной значимости в открытом классе падежных языков $\{K_1 \dots K_n\}$, открытом классе языков с нулевыми морфемами $\{M_1 \dots M_n\}$, открытом классе языков с энклитиками второй позиции $\{W_1 \dots W_n\}$ и т. д. Выдвинутое не так давно требование устранить

и моделью₂ не всегда обнаруживается соответствие, так как не все модели₁ эмпирически проверяемы. В 1892 г. Я. Ваккернагель обосновал существование класса *энклитик второй позиции*, т. е. слабоударных элементов, располагающихся после начального элемента предложения [Wackernagel 1892]. Это понятие — часть модели₁, оно тесно связано с модельными₁ понятиями «ударного слова», «безударного слова», «энклитики», «порядка слов», «синтаксического расстояния». Одновременно Я. Ваккернагель показал, что энклитики второй позиции можно выделить в ряде древних индоевропейских языков, включая древнегреческий и древнеиндийский. Это открытие, названное «законом Ваккернагеля», можно интерпретировать по-разному. Во-первых, можно предположить, что такая метаязыковая категория, как «энклитики», нужна в описании лишь тогда, когда у элементов данного класса есть дистрибуция, предсказываемая законом Ваккернагеля, иными словами, «вторая позиция» — это дефиниционное свойство энклитик. Такая формулировка озвучивается редко, но соответствующее понимание *de facto* отражено во многих исторических и синхронных грамматиках языков мира, где раздела об энклитиках нет, несмотря на то, что элементы, подводимые под определение энклитик, т. е. безударных слов, стоящих после ударных элементов, в этих языках есть. Накопленные в XX в. наблюдения над языками мира указывают на разные типы расстановки энклитик и проклитик [Klavans 1985; Sadock 1995; Aikhenvald 2002; Lee, Billings 2005; Cysouw 2005]. Поэтому в перспективе описания всего многообразия энклитик закладывать в их определение критерий «второй позиции» неразумно. Тем не менее прощаться с этой точкой зрения рано. В последние десятилетия появились работы, где утверждается, что энклитики и другие предполагаемые элементы второй позиции, например, финитные глаголы (*finite verbs*), на некотором уровне репрезентации всегда стоят во второй позиции в пределах своей синтаксической области [Rizzi 2004; Benincà, Poletto 2004]. Есть также работа, где утверждается, что энклитики второй позиции, вопреки наблюдаемому на поверхности порядку слов, на самом деле стоят в *первой* позиции, а предыдущие элементы относятся к другой синтаксической области [Agbayani, Golston 2010]. Подобные априорные модели₁ не являются моделями₂, позиция энклитик определяется в них на основе допущений о соотношении разных уровней репрезентации структур, куда вставляются энклитики, а разметка синтаксических областей является следствием этих допущений.

Во-вторых, закон Ваккернагеля можно интерпретировать как статистически значимую тенденцию порядка слов, которая может проявиться или не проявиться в конкретном языке. Возможен сценарий, когда одна часть энклитик в некотором языке подчиняется закону Ваккернагеля, а другая — нет, а также сценарий, когда некоторый язык в один период времени подчиняется закону Ваккернагеля, а в

из описаний языков мира вообще все модельные понятия, претендующие на универсальную применимость [Haspelmath 2010; 2015a], едва ли реализуемо. Впрочем, и выдвигалось это требование в рамках специализированной области лингвистики — лингвистической типологии, и при этом в контексте специальной инженерной задачи — упрощения и стандартизации описаний языков мира при подготовке типологических баз данных и атласов языков мира.

другой — нет. В этих случаях выдвижение гипотезы о том, что закон Ваккернагеля действует в конкретном языке L, придает ей статус модели₂. Остановимся на двух прецедентах. В.И. Абаев в работах 1930–40-х гг. показал, что закон Ваккернагеля действует в осетинском языке для ряда местоименных и неместоименных слов [Абаев 1949]. Это открытие сделано *самим носителем осетинского языка*, почерпнувшим модельные₁ понятия энклитик и закона Ваккернагеля из описаний других языков и решившим, что они повышают точность описания родного языка. А.А. Зализняк в начале 1990-х гг. показал, что данный закон действовал в древнерусском языке и предсказывал позиции энклитических частиц, местоимений и связок [Зализняк 1993: 280–303]. Это плод наблюдений над *чужим языком*: современный русский не дает здесь ориентира, поскольку в нем закон Ваккернагеля, по крайней мере, по выдвинутым А.А. Зализняком критериям, не действует, а предыдущие описания древнерусского языка скорее запутывают картину, чем проясняют ее. Толчок дали открытие и расшифровка ранее недоступного массива текстов — берестяных грамот. Оказалось, что гипотеза о действии закона Ваккернагеля в древнерусском языке (модель₂) успешно предсказывает расположение энклитик не только в этом массиве⁵, но и в ряде ранее известных текстов, предположительно близких к устной речи, например, в прямой речи персонажей «Повести временных лет» и «Киевской летописи», а также в ряде памятников XIII–XV вв., в той или иной мере сохраняющих точки соприкосновения с устной речью и собственными последней механизмами расстановки энклитик. Последняя часть тезиса подробно доказывается в книге [Зализняк 2008]. Не вникая в историко-филологические тонкости, можно констатировать, что при отказе от гипотезы о действии закона Ваккернагеля не только расстановка энклитик, но весь порядок слов древнерусского языка оказывается зоной хаотического распределения, в то время как при ее выдвижении порядок слов можно описать на основе небольшого числа ограничительных условий⁶.

Ложное представление об отсутствии у древнерусских энклитик фиксированного места объясняется тем, что авторы предшествующих работ пытались установить место связочных, местоименной энклитики и энклитических частиц для каждого элемента отдельно, не различая два аспекта правил — *внутренний*, т. е. позицию клитик относительно друг друга (clitic-internal ordering), и *внешний*, т. е. позицию энклитики / цепочки энклитик в предложении (clitic-external ordering). Характер взаимодействия этих правил неспецифичен для древнерусского языка. То же соотношение было выявлено более 120 лет назад еще Я. Ваккернагелем на другом материале. Но если предположить, что человек, перед которым поставлена задача описать порядок энклитик, не имеет готовой модели₁, а должен вывести ее

⁵ Базис для верификации увеличивается с открытием каждой новой грамоты, содержащей энклитики.

⁶ Мы полагаем, что А.А. Зализняк имел в виду именно это, говоря о том, что «древнерусский традиционнo (и в значительной мере ошибочно) рассматривается как язык со свободным порядком слов. Этим можно объяснить тот факт, что рассматриваемые здесь правила <относительного расположения энклитик — А. Ц.> не были выявлены ранее» [Зализняк 1993: 281].

самостоятельно на основе дистрибуции элементов в текстах языка L, взаимодействие двух типов правил нетривиально. Правила внутреннего расположения энклитик опираются на математическое понятие *непосредственного предшествования*. Предполагается, что для любой пары энклитик a и b устанавливается отношение $a > b$ «a непосредственно предшествует b» и что последовательность энклитик, удовлетворяющая данному критерию, является объектом особого рода, так называемой *цепочкой* (clitic cluster), т. е. упорядоченным множеством $\{a, b, c \dots n\}$, занимающим ту же позицию, что любая одиночная энклитика $k \in \{a \dots n\}$. Специфика языков с законом Ваккернагеля такова, что правила внешнего расположения энклитик опираются не только на понятие непосредственного предшествования, но и на понятие относительного расположения элементов (*предшествования*). Для оценки *внешней* позиции цепочки энклитик / одиночной энклитики $k \in \{a \dots n\}$ в общем случае не важно, стоит ли она после глагола, существительного, наречия или чего-то еще. Важно *синтаксическое расстояние*: между цепочкой энклитик / одиночной энклитикой $k \in \{a \dots n\}$ и левой границей предложения должен быть ровно один элемент. В примере (1) четыре энклитики *же, ми, ся, еси* образуют цепочку, которая стоит после первого ударного слова *ако*, т. е. во второй позиции от начала предложения. То, что смежные элементы = *же* = *ми* = *ся* = *еси* не являются случайным объединением безударных слов, подтверждается тем, что пишущий расположил их именно в том порядке, который предсказывается шаблоном внутреннего упорядочения энклитик, выдерживаемым во всех древнерусских текстах.

(1) ДРЕВНЕРУССКИЙ

Ако = [*же* = *ми* = *ся* = *еси*] самъ върекль (Хлебн., л. 147, под 1150 г.).

Представленное в [Зализняк 1993; 2008] доказательство того, что в русском языке домонгольского периода часть энклитик, включая возвратную частицу *ся*, подчинялась закону Ваккернагеля, основано на комбинации базового правила, ставящего все энклитики после первого ударного слова, с контрправилом, смещающим энклитики в постпозицию ненаначальному элементу Y, ср. нотацию $[X] — Y = CL$, соответствующую колебаниям типа др. рус. *того* (X) = *ся* (CL) *отъступилъ* \Rightarrow [*того*] (X) *отъступил* (Y) = *ся* (CL). Эта процедура сближает описание Зализняка (предиктивную модель₂, предсказывающую порядок энклитик) с априорными моделями₁, где энклитики всегда стоят во второй позиции. Действительно, А. А. Зализняк пишет, что в предложениях вида $[X] — Y = CL$, начальный элемент [X] («барьер», в его терминах) «как бы можно не учитывать»⁷, и допускает, что «отчлененная часть» была отделена паузой [там же]. Однако полного сходства нет. Во-первых, А. А. Зализняк воздерживается от экстраполяции своего описания на другие языки. Во-вторых, эффективность описания А. А. Зализняка не зависит от допущения, что начальный «барьер» [X] отделен от последующей части фонетически. Это лишь эвристика, догадка автора модели₂ о том, как правило $X — CL \Rightarrow$

⁷ «...правила о расположении энклитик действуют после этого так, как если бы отчлененной части в предложении вообще не было» [Зализняк 1993: 286].

[X] — Y = CL реализовалось в устной речи носителей мертвого языка. Из двух равноценных объяснений — считать или не считать барьер [X] частью той синтаксической области, где стоит энклитика, А. А. Зализняк выбрал второе. Возможно потому, что оно показалось ему проще, а в его материале бесспорных примеров, где работало бы только одно из этих объяснений, нет. Зато в других языках с законом Ваккернагеля, в частности в региональных вариантах сербохорватского, словацкого языков и в древнечешском языке [Čavar, Wilder 1999; Циммерлинг 2002: 85–87; Циммерлинг 2013: 484], такие случаи есть. В предложениях с барьером возникает фиксированный порядок с глаголом (символ ‘V’ в нотации) на втором месте и энклитикой на третьем, $X — CL... V \Rightarrow [X] — V — CL$, т. е. закрепляется место еще одного элемента относительно начала предложения. Этот факт опровергает гипотезу о том, что категория [X] не играет роли для упорядочения последующей части предложения, содержащей энклитики.

Если допустить, что расстановка энклитик в языках с законом Ваккернагеля регулируется одними и теми же ограничительными условиями, то при описании древнерусского языка следует предпочесть синтаксическое объяснение, так как оно имеет большую предсказательную силу на множестве уже описанных языков с правилом барьера. Неясно, можно ли параметризовать этот класс, выделив предполагаемые языки с «несинтаксическим» и «синтаксическим» барьером: речь скорее идет о разных способах описания одного и того же материала, ср. [Progovac 1996] vs [Radanović-Kosić 1996]. Зато другая черта языков с законом Ваккернагеля — постановка энклитик после первого фонетического слова (X) vs после первой группы члена предложения (XP) — параметризуется хорошо. В особую группу выделяются X/XP языки типа сербохорватского, луисеньо и варльпири, где энклитики могут стоять как после первого фонетического слова, так и после первой группы [Anderson 1993]. Эти и прочие настройки параметров, определяющих возможный набор начальных элементов, к которым присоединяются энклитики, обсуждаются в [Циммерлинг 2013: 108–110; 355–384]. А. А. Зализняк, как и до него Я. Ваккернагель, принимает в расчет лишь языки подтипа X — CL, где энклитики / цепочки стоят после первого фонетического слова, трудно их за это упрекать, поскольку в их подборке XP — CL языков и X/XP — CL языков нет. Но любому лингвисту, какой бы пиетет перед предшественниками он не испытывал, все же придется параметризовать признак ‘± наличие полной группы перед клитиками’, если он поставит себе задачу описать известные языки с законом Ваккернагеля или пожелает установить сходства и различия в правилах упорядочения энклитик в таких языках, как древнерусский, древнегреческий, хеттский (X — CL), чешский, словацкий, осетинский, пашто (XP — CL), сербохорватский, варльпири (X/XP — CL).

Классы и подклассы языков мира в параметрической типологии — тоже теоретические конструкты, которые держатся на используемой сетке параметров и представлениях об их внутренней структуре [Longobardi 2001; 2005; Baker 2008; Roberts 2012]. Лингвисту остается утешать себя тем, что решения о помещении языков L_1 и L_2 в общий класс L верифицируемы в той мере, в которой значения признака, признаваемого диагностическим для L, в L_1 и L_2 совпадают и сам признак таков,

что принимаемые им значения поддаются проверке в массиве данных. Процедура описания варьирования, основанная на ранжировании параметров, имеет свои изъяны, но она операциональна. Надо лишь помнить, что такого дотеоретического «факта», как отнесенность древнерусского или сербохорватского к «классу языков с законом Ваккернагеля» или «классу языков с цепочками клитик» нет и быть не может.

Подведем итоги нашего экскурса и выдвинем три методологических положения:

- (i) При наличии альтернативных описаний следует выбрать то, которое имеет большую предсказательную силу.
- (ii) При прочих равных условиях имеют приоритет те модели₂, которые предсказывают варьирование в более обширном классе языков.
- (iii) Параметризация языкового многообразия повышает предсказательную силу моделей₂.

III

В последние десятилетия обрело популярность мнение, будто подлинное объяснение в лингвистике возможно лишь на уровне *глобальных концепций языка*, ср. характерные формулировки в учебнике [Radford 1995]. Такие концепции в англоязычной традиции называются термином «framework», а в русской именуются «интегральным описанием языка». Часто они связаны с выдвижением или отрицанием гипотезы о существовании врожденной универсальной грамматики, но в контексте нашего рассуждения это привходящее обстоятельство, показывающее, какого рода идеи в конце XX — начале XXI в. могут служить стимулом для очередной глобальной концепции языка. Framework — не дедуктивно замкнутая теория, а лингвистический идеал, инструкция, как описывать естественный язык на основе небольшого числа базовых принципов. У интегральных концепций языка два достоинства. Во-первых, они обогатили лингвистику большим числом модельных₁ понятий. Во-вторых, они внесли вклад в унификацию языка описания и облегчили коммуникацию лингвистов, работающих в рамках данного framework.

Недостатки глобальных концепций языка не менее очевидны. Frameworks претендуют на *интегральность*, т. е. на объединение всей релевантной информации, касающейся структуры языка: чем шире спектр исследований в пределах избранного framework, тем меньше его приверженцы нуждаются в выходах за его пределы. Проблема в том, что хорошо описывается лишь небольшая часть фактов, а в других случаях задается некоторый предел огрубления на фоне существующих точных *аспективных* описаний и предиктивных моделей₂⁸. Положение можно пытаться исправить — правда, на это не хватит всей жизни, — настойчиво привлекая

⁸ Если считать, что язык как предмет описания интегральной теории языка является цельным объектом, выдвинутое нами утверждение о неизбежности огрубления в описании с позиций

внимание законодателей мод в том или ином framework к явлениям, которые те понимают хуже всего⁹, но основной дефект, видимо, неустраним. Сложность естественного языка такова, что любая формализованная интегральная теория, основанная на строго монистических принципах, недостаточна: делая акцент на анализе явлений, которые неточно описываются конфликтующими с ней интегральными теориями, она проигрывает в точности описания других явлений. Кроме того, встает вопрос: если применение предиктивных моделей₂ не зависит от выбора между глобальными теориями языка А, В и С (см. выше в п. II комментарий к возможным трактовкам правила барьера в древнерусском языке), какую новую информацию мы получим от их адаптации к понятийному аппарату этих теорий?

Предложим свой вариант ответа: новая информация при адаптации предиктивной модели M_1 к аппарату глобальной теории А возникнет, если это позволяет строить модифицированную предиктивную модель M_{1a} , дающую лучшие результаты на материале одного или нескольких языков. Используем прежнюю иллюстрацию с правилом барьера $X — CL... V \Rightarrow [X] — V — CL$ в парах предложений типа др. рус. *того* (X) = *ся* (CL) *отъступилъ* \Rightarrow [*того*] (X) *отъступил* (Y) = *ся* (CL). Само наличие альтернатив данного типа в классе языков с законом Ваккернагеля — более важный факт, чем выбор одной из формальных моделей этой альтернативы, не потому что эти модели плохи, а потому, что они не нужны для решения ряда задач: (под)класс языков с правилом барьера можно выделить независимо от соображений относительно того, какими механизмами эти правила порождаются. Правда, без дополнительных данных неясно, пересекается ли этот класс с классом языков с законом Ваккернагеля или же один из этих классов является собственной частью другого.

Из двух возможных объяснений правил барьера — фонетического и синтаксического — первое не принимает в расчет информации о синтаксической структуре предложения с энклитиками, а второе учитывает ее. Рассуждая абстрактно, нельзя установить, что лучше: под рукой будут разные глобальные теории, с помощью которых можно обосновать любое решение, см. обзор [Halpern 1996: xi — xiii]. Эмпирика подсказывает, что есть языки с правилом барьера, где от синтаксического состава предложений с барьером и энклитиками нельзя отвлекаться. Поэтому, по постулату (ii), второе объяснение имеет приоритет. Синтаксис, однако, может присутствовать в формулировке правила барьера в разном объеме. Необходимый минимум на «входе» касается сведений о *наличном составе* элементов (глагол, именная группа и т. п.). Целесообразно также отразить представление о *направленности* преобразования $X — CL... V \Rightarrow [X] — V — CL$, чтобы выражение слева от стрелки \Rightarrow трактовалось как исходное, а выражение справа — как производное. Это уточнение помещает синтаксическое представление правила барьера в класс

интегральной теории при допущении, что некоторые характеристики объекта данная теория все же описывает точно, равнозначно утверждению, что лингвистические описания подчиняются принципу неопределенности Гейзенберга.

⁹ В разных глобальных теориях уязвимыми местами могут быть просодия, морфология, лексическая семантика, порядок слов и т. п.

трансформационных грамматик с заданным вектором преобразования. Такие объяснения используются в самых разных областях лингвистики, но трансформации в моделях синтаксиса имеют свою специфику. Имеет смысл уточнить, за счет какой операции возникает порядок слов в производном предложении $[X] — V — CL$, где начальная категория $[X]$ ведет себя как барьер, сдвигающий энклитику вправо на шаг дальше. Вариант ответа предлагает теория (= модель) формальных грамматик с оператором *перемещения* (movement), ср. [Stabler 1997]. Предложения с производным порядком слов, т. е. тем, что в традиционной лингвистике называется «инверсией», ср. рус. *Возвращается Петька из разведки*, порождаются на основе предполагаемых исходных предложений типа *Петька возвращается из разведки* не путем слепого отображения порядка XY в порядок YX, а за счет изменения положения какого-то одного активно перемещающегося элемента, в данном примере — глагола, ср. запись, где символ ‘ $_i$ ’ указывает на исходное место глагола: *Возвращается, Петька $_i$ из разведки*¹⁰. Тот же аппарат применим и к предложениям типа *Того отъступил, ся $_i$* , где гипотеза о перемещении глагола в позицию, предшествующую позиции энклитик, уместна. Если лингвист, объясняющий правило барьера в терминах перемещения, захочет отразить тот факт, что в языке с законом Ваккернагеля ограничения, накладываемые на место энклитик в предложении, более значимы, чем ограничения, накладываемые на место глагола в предложении — а это утверждение верно по отношению к большому числу языков данного класса, — и что порядок $X — CL \dots V$ с энклитиками во второй позиции от начала является источником порядка $[X] — V — CL$ с энклитиками в третьей позиции от начала, то ему, возможно, придется постулировать не только перемещение глагола *справа налево* (ближе к началу предложения, right-to-left movement), но и перемещение энклитик в обратном направлении, *слева направо* (ближе к концу предложения, left-to-right movement). Это понимание запечатлено в нотации (2), модифицированной вслед за [Zimmerling, Kosta 2013]: нижний индекс ‘i’ указывает на место энклитик, а нижний индекс ‘j’ — на место глагола.

$$(2) \quad X — CL \dots V \Rightarrow [X] — _i V_j — CL_i \dots _j.$$

Лингвист, объясняющий правило барьера в терминах перемещения, может отказаться от идеи о перемещении энклитик, ограничившись гипотезой о перемещении

¹⁰ Технически, «перемещенный» элемент α в теории формальных грамматик является участком дерева предложения (поддеревом) в структуре вида $X_\alpha Y Z_\alpha$, где допускается более одного вхождения α . Перемещения α не создают синтаксических позиций (поддеревьев) X, Z, U, W, а лишь меняют их заполнение. Операции преобразования уже построенного дерева, в результате которых возникают новые категориальные объекты вида X_α , т. е. «поддерево, соответствующее категории X, заполненное перемещенным элементом α », каузально мотивированы наличием некоторого абстрактного свойства (feature-driven), делающим их необходимыми или возможными. Такая модель имеет обширную область применения в естественном языке хотя бы потому, что порядок слов предсказуем и изменение линейного порядка всегда мотивировано теми или иными семантическими или коммуникативными причинами.

глагола, как поступили, например, Д. Чавар и К. Вильдер в описании хорватского варианта сербохорватского языка [Čavar, Wilder 1999]. Такие решения мотивированы не только формализмом (в том варианте генеративной грамматики, который отстаивают Д. Чавар и К. Вильдер, перемещения слева направо (left-to-right movement) запрещены), но и той задачей, которую решает лингвист.

Итак, по мере насыщения простого правила линейного порядка, каким является правило барьера, информацией о синтаксической структуре предложения, содержащего энклитики, это правило все теснее привязывается к классам формальных моделей₁, использующих понятие трансформации, перемещения (или эквивалентные ему представления об операциях, меняющих порядок слов) и постулатам о свойствах перемещения. Понятие перемещения было предложено в рамках глобальной теории (генеративной грамматики Н. Хомского), поэтому у читателя может возникнуть вопрос: не лучше ли принять ее исходные положения открыто, тем более что в ней эксплицитно оговариваются постулаты о перемещениях, см. [Pollock 1989], а также предлагаются готовые ответы на другие важные вопросы — как порождается предложение, как выглядит универсальная иерархия членов предложения, чем именно клитики отличаются от других элементов и т. д. и т. п. Не повысит ли это точность предсказаний? Увы. Объединение в одном «флаконе» того, что нужно для «лечения» (допущений, необходимых для построения предиктивной модели₂ и решений, повышающих ее точность на избранном материале), и готовых ответов на все вопросы — ловушка. Сличение априорной модели₁ с материалом — законная проблема. Но лингвист, который для построения предиктивной модели₂ предсказывающей распределение линейных порядков вида XYZ ~ XXУ в конкретном языке, нуждается в избыточной информации, например, в сведениях о последней версии гипотезы об универсальной иерархии фразовых категорий, или в постулатах о том, имеет ли язык модулярную или уровневую природу и присутствует ли универсальная грамматика в мозгу маленьких детей, подобен русскому писателю XVII в. Этот замечательный человек для обоснования своей жизненной и теологической правоты начал автобиографию с Символа Веры, а затем подробно обсудил (Псевдо-)Дионисия Ареопагита¹¹.

В используемой в минималистской программе модели₁ перемещения уязвимое место очевидно. Это то, что составляет предмет гордости ее разработчиков, — привязка линеаризации к алгоритму, выстраивающему дерево предложения снизу вверх (down-to-top). Постулат Фиенго [Fiengo 1977] разрешает перемещения элементов снизу вверх, т. е. при заданных условиях ветвления — *справа налево*, и запрещает перемещения в обратном направлении. Это догматическое требование, которое усилено в более поздней работе [Кауне 1994]. Как показали построенные почти одновременно с выходом книги Р. Кейна математически строгие минималистские грамматики с оператором перемещения [Stabler 1997], эффективность

¹¹ Преклоняюсь перед мощью и подвигом этого человека, оставившего замечательные образцы русского языка. Но параллели между объяснениями на основе глобальных теорий языка и теологическим спором в моих глазах не повышают престиж науки.

вычисления не меняется из-за запрета/разрешения на перемещение в двух направлениях. Для описания большинства альтернатив линейных порядков вида $XYZ \sim XZY$ информация о порождении дерева предложения не нужна, требуется понять отношение этих порядков по отношению друг к другу и установить, является ли один из них источником другого. В языках мира есть случаи, которые хорошо описываются правилами перемещения слева направо и хуже описываются иначе, ср. [Otsuka 2005; Manetta 2012; Overfelt 2015; Polinsky 2019]. Неудивительно, что такие случаи есть и в языках с правилами барьера¹². Выдвинем еще два методологических положения:

- (iv) При наличии альтернативных способов описания языковых явлений приоритет имеют те модели₂, которые не связаны с единственной глобальной моделью₁, но совместимы с классом моделей₁, опирающихся на разные исходные допущения.
- (v) Предиктивные модели₂ не должны содержать избыточной информации. Избыточная информация не повышает предсказательной силы модели и нужна для привязки к другим явлениям языка, характер связи которых с описываемым явлением не определен.

IV

Лингвистическая типология — область, которая должна соединять общую и частную лингвистику¹³. Но типологи тоже люди и тоже лингвисты. Лингвисты, как видно из нашего текста, склонны одновременно игнорировать друг друга и бороться с теориями друг друга, критиковать свою науку и апеллировать к чужим. Глобальные теории языка — развернутая апелляция к математическим и естественным наукам. Такие модельные₁ понятия, как «сегмент», «тождество», «тождественная единица», «структура», «дерево подчинения», «ветвление», «предшествование», «линейный порядок», «инвариант» и многие другие, с ними связанные,

¹² Эмпирические данные указывают на то, что правила барьера выделены не только в языках с законом Ваккернагеля, но и в других языках с цепочками клитик. Соотношение классов языков, по-видимому, таково:

(а) Языки с цепочками клитик \supset Языки с законом Ваккернагеля; Языки с законом Ваккернагеля \cap Языки с Правилами барьера.

В ряде неваккернагелевских языков с цепочками клитик, в частности, раннем новогреческом и европейском португальском, где клитики занимают смежные позиции с глаголом, правила барьера меняют просодический тип клитик, создавая альтернацию линейных порядков с энклизой и проклзой $V = CL \sim CL = V$ в одном и том же типе клауз. Запрет на перемещения слева направо вынудит лингвиста объявить конфигурацию с энклизой во всех случаях исходной. Для европейского португальского такое описание оправдано, ср. [Rouveret 1999; Циммерлинг 2002: 85], а для раннего новогреческого койнэ возможны разные решения, ср. [Rollo 1989; Mackridge 1993; 2000; Кисилиер 2011].

¹³ О соотношении типологии и сопоставительной (контрастивной) лингвистики см. проникающие соображения в книге [Гак 1983: 25–27].

разумеется, в основе своей математические. Трудно представить себе теоретическое описание языка, которое может без них обойтись (речь не о терминах, а о самих понятиях). В последнее время одна из школ в типологии поставила себе задачу полностью очистить описания языков мира от априорных модельных₁ понятий. Таковые в [Haspelmath 2010] названы термином «cross-linguistic categories», а дозволенные метаязыковые понятия, пригодные для сопоставления языков, — «comparative concepts». Один из доводов М. Хаспельмата — многие, если не все претендующие на универсальность cross-linguistic categories, такие как «падеж», «подлежащее», исторически возникли на основе наблюдения над небольшим числом новых и древних языков Европы (латынь, английский и т. п.) и экстраполированы на все языки. Это серьезная проблема. Идею о том, что каждый язык должен быть описан в терминах специфических для него категорий, М. Хаспельмат связывает с наследием Ф. Боаса (1858–1942), но она буквально носилась в воздухе. Тот же Л. В. Щерба, который не чурался в описании русского языка таких модельных₁ понятий, как «нулевая связка», «часть речи» и т. п., говорит в опубликованной посмертно статье о том, что проецирование категорий языка описания на описываемый язык приводит к недискретному представлению, которое он называет смешанным двуязычием [Щерба 1945; 2008: 40–42]. Щерба подчеркивает существование того, что сейчас принято называть «лингвоспецифическими концептами», и чуть ниже высказывает сомнение в том, что такое понятие, как «слово», универсально и имеет одинаковое содержание в языках разного типа [там же: 44]. Это понравилось бы нынешним борцам с модельными понятиями.

Понятию «клитики» (обобщению случаев проклитизы и энклитизы, см. [Zwicky 1977]) ныне тоже отказывают в валидности на том основании, что в описаниях разных языков этим термином обозначаются разные сущности: порой класс морфем, а порой — класс слабоударных слов [Haspelmath 2015b]¹⁴. В достаточности этого довода можно усомниться. «Клитика» — это понятие морфологической классификации: те лингвисты, которые пользуются представлением о тернарной классификации морфологических элементов, постулируют для клитик статус, промежуточный между статусом морфемы (аффикса или корня) и статусом словоформы [Aikhenvald 2002]. Это, бесспорно, модельное₁ понятие. Но отрицание тернарной классификации морфема vs клитика vs слово и замена ее бинарной классификацией морфема vs слово / комплекс морфем тоже есть модельное₁ понятие, причем совместимое не только со взглядом типолога-агностика, но с вполне формализованными глобальными теориями языками, такими, как грамматика HPSG, в рамках которой выдвигался тезис о том, что клитики — вид аффиксов (см. [Tseng 2003]). Кроме того, хотелось бы знать, почему, если клитики как особый класс элементов — фикция, предиктивные модели₂ предсказывают распределение этих якобы фиктивных сущностей, причем не только в пределах одного языка, но на материале

¹⁴ Призывы избавиться от понятия «клитик» и свести их к предположительно более ясным сущностям, которыми, по мнению ряда сторонников Н. Хомского, являются «проверяемые морфологические признаки», озвучивались и некоторыми генеративистами.

открытых классов языков мира. Влияние языка-эталона (древнегреческий и прочие мертвые языки) и первого прецедентного описания (Я. Ваккернагель, 1892) на лингвистов, усмотревших сходное распределение в других языках — сербохорватском (Р. О. Якобсон, 1935), осетинском (В. И. Абаев, 1949), древнерусском (А. А. Зализняк, 1993) — налицо. Но ведь эти модели₂ повышают точность описания соответствующих языков, а кроме того, позволяют изучать свойства некоторого класса языков мира, соответствующего заданным параметрам расстановки клитик. Доказывает ли закон Ваккернагеля и аналогичные предиктивные модели₂ существование «клитик» и «цепочек клитик»? Возможно, это вопрос веры. Но ведь и обратная установка, утверждающая что модели₂, повышающие предсказуемость описаний языков мира, для описаний этих языков и обобщений над этими описаниями излишни, тоже в лучшем случае — предмет веры.

Очищение типологических описаний от нечетких понятий и стандартизация метаязыка — важная задача, особенно при подготовке таких специализированных ресурсов, как базы данных и атласы языков мира. Пока что многое в декларациях и практике идеологов борьбы с глобальными теориями языка остается неясным. Адресуют ли они свой посыл только типологам, отказывая в валидности тем описаниям, которые содержат отсылки к *cross-linguistic categories*, или всем лингвистам, призывая их ввести единый стандарт описания языков для всех целей? Имеют ли они в виду преимущественно технические или теоретические аспекты? Верят ли они сами в то, что модельные₁ понятия устранимы из описания языков? Так, Д. Матич и Дж. Веджвуд в перспективе различения модельных понятий (*cross-linguistic categories*) vs категорий сравнительного описания (*comparative categories*) подвергают критике популярный термин «focus»¹⁵, утверждая, что в описаниях разных языков этот термин указывает на разные вещи [Matić, Wedgwood 2013: 158–159]. Они заключают, что *focus* не *cross-linguistic category*, а в лучшем случае — эвристика для различения контекстов, где нечто в языке L, судя по внешним признакам (использование особых конструкций, особой интонации, особого порядка слов и т. п.), является объектом прагматического выделения, и контекстов, где таких признаков нет¹⁶. С мнением о том, что ярлык «focus» привешивается к разнородным явлениям, предположительно связанным с коммуникативным выделением и при этом не отнесенным к сфере употребления «топика» (*topic*), другой загадочной категории, мы склонны согласиться¹⁷. Но само рассуждение Д. Матича и Д. Веджвуда ошибочно. Проблемы с применением термина «focus» в описаниях языков возникают не от того, что это не модельное понятие, а от того, что это неудачное модельное понятие, возникшее на стыке теории актуального членения

¹⁵ т. е. «фокус», иной перевод на русский язык затруднен.

¹⁶ «The notion of focus is then a means of identifying structural patterns which languages use to generate a certain number of related pragmatic effects» [Matić, Wedgwood 2013: 160].

¹⁷ Эта эвристика обычно используется так. Допустим, что частица F является *фокусной* (*focus particle*). Тогда предложения, где используется F, будут конструкциями с *фокусным выделением* (*focusing*), а позиции, где F возможно, — *позициями фокусных элементов* (*focus positions*). Тем же способом можно «объяснять» *фокусные* линейные порядки, акценты и т. п.

и логической семантики при попытке интерпретировать первую в терминах второй. Понятие «focus», см. [Vallduvi, Engdahl 1996; Buring 1997; Erteschik-Shir 1997] — попытка свести воедино три сущности, за каждой из которых стоит свое модельное, понятие: 1) *рему / собственно иллокутивный компонент*, т. е. цель сообщения; 2) *контраст*, т. е. логико-семантическое значение «выбора из множества альтернатив» [Янко 2001: 43–61]; 3) *сферу действия* операторных выражений (например, квантора, отрицания, вопросительного слова и т. п.) [Богуславский 1996: 43–48; McCawley 1995]. Детальное обсуждение явлений, которые модели₁, смешивающие статус ремы, контраста и сферы действия, описывают неверно, уведет в сторону. Укажем лишь, что нетождественность ремы и контраста, т. е. наличие контрастных сегментов высказывания, которые не являются ремами [Dryer 1996], а также автономность сферы действия элементов высказывания от бинарного членения высказывания на рему vs не-рему являются причиной того, что на основе модельного₁ понятия фокуса нельзя построить надежную предиктивную модель₂.

В недавней работе [Haspelmath 2019], где критикуются модельные₁ понятия «подлежащего» и «эргативности», высказано соображение, перекликающееся с нашим: преимущества от введения сложных модельных понятий там, где в описании достаточно более элементарных (в контексте статьи М. Хаспельмата это, прежде всего, *семантические роли*), иллюзорны, более простые описания для типологии приоритетны, а неverified допущений, например о том, что переходные предложения в некотором языке на «глубоком» (deep) уровне рассмотрения являются непереходными, а именное предложение в этом же или другом языке содержит узел скрытой категории, выраженной нулевой формой, следует избегать. В общей форме с этим можно согласиться. Но кто и с какой целью устанавливает глубину рассмотрения (depth) языковых явлений и предел сложности модельных₁ понятий? Мы полагаем, что ответ на этот вопрос должен зависеть не от стажа употребления модельных₁ понятий в той или иной глобальной теории языка, а от решаемой задачи и характера самого понятия.

Модельное₁ понятие *синтаксического подлежащего* таково, что допускает разбиение языков мира на подлежащные и бесподлежащные. По крайней мере, такая типология была предложена в [Li, Thompson 1976; Kibrik 1997], и статья М. Хаспельмата вносит вклад в ее разработку¹⁸. Модельные₁ понятия перемещения и фокуса таковы, что они имеют смысл только как универсальные. Абсурдно делить языки мира на **языки с перемещением* и **языки без перемещения*: такой шаг равнозначен делению языков мира на **языки с грамматикой составляющих* и **языки с грамматикой зависимостей*. «Грамматика составляющих» и «грамматика зависимостей» — разные (и эквивалентные по своей мощности) математические модели дерева подчинения, а не явления конкретных языков. То соображение, что специфические черты английской грамматики в свое время стимулировали

¹⁸ Требование убрать категорию подлежащего из описания подлежащих языков, например, английского или русского, на том основании, что в описании *некоторых других языков* можно без нее обойтись, выглядит чрезмерным.

ученых, владевших этим языком, к разработке формализма составляющих, а специфические черты русской грамматики в свое время стимулировали другую группу ученых на разработку формализма зависимости, не относится к делу¹⁹. Понятие «перемещение» математическое, а не лингвистическое, оно описывает операцию над деревом подчинения в классе формальных грамматик особого типа. Поэтому, строго говоря, нельзя утверждать, что описание в терминах перемещения таких явлений языка, как механизм приписывания падежа или связывания возвратного местоимения, менее «законно», чем описание альтернатив порядка слов. Сомнения вызывает то, что все эти явления глобальная теория с опорой на некоторые догматические допущения описывает одинаково. При применении понятия перемещения только в сфере порядка слов оно остается полезным инструментом. Модельное₁ понятие «фокус» — это лингвистический конструкт, а не чистое понятие логической семантики, хотя два из трех его компонентов являются таковыми. Сторонники этого конструкта уверены, что объединению трех разных сущностей — «ремы», «контраста» и «сферы действия» — обязательно соответствуют внешние языковые средства во всех языках мира. «Фокус» повторяет путь «подлежащего»: особый статус последнего пытались нащупать на основе гипотезы о том, что подлежащее — категория универсальная и как-нибудь в любом языке да проявится [Keenan 1976]. Накопленный с тех пор опыт описания не дает возможность исключить то, что подлежащее может и не быть универсалией. Зато в классе *подлежащих языков* на основе гипотезы о наличии подлежащего можно строить правдоподобные предсказания (модели₂). Фокус может быть только универсальной категорией, разбиение языков мира на *фокусные и *бесфокусные бессмысленно. Проблема лишь в том, что объект приложения этого понятия эфемерен.

V

«Глубина» рассмотрения — понятие, которое относится не только к *сложности* представления, но и к дистанции во времени. Если звучат призывы ограничить глубину рассмотрения синхронных описаний, может быть, лингвистам можно рекомендовать ограничить или увеличить глубину диахронического анализа ради лучшего понимания самой синхронии? В качестве иллюстрации рассмотрим сюжет, который объединяет два ранее рассмотренных нами, — эпизоды функционирования связки 'быть' в разные периоды существования русского языка. Дело в том, что в древнерусском языке связочные формы глагола 'быть' были *энклитиками* и подчинялись закону Ваккернагеля, который упоминался в п. III–IV. Более

¹⁹ С равным успехом можно отвергнуть предложенные в качестве универсальных коммуникативных категорий модельные₁ понятия темы и ремы на том основании, что они исторически были впервые предложены на материале славянских языков со свободным порядком слов [Mathesius 1947] и лингвоспецифическим соотношением между порядком слов и фразовыми акцентами [Ковтунова 1976]. Понятия «темы» и «ремы» операциональны: тезис о том, что в любом языке есть высказывания, которые делятся на два коммуникативных компонента, один из которых является темой, пока что не опровергнут.

того, в древнерусском языке уже был аналог *нулевой связки*, так как связка ‘быть’ в 3 л. настоящего времени была выражена нулем. Правда, диагностический для связки контекст не совпадает с современным: это аналитическая конструкция прошедшего времени с причастием –л, которое еще не стало глагольной формой, так называемый древнерусский перфект²⁰. В 1–2 л. ед. ч., мн. ч. и дв. ч. уже к XI–XII вв. говорили *пришель есмь* «я пришел», *пришель еси* «ты пришел», *пришла еси* «ты пришла», *пришли есмь* «мы пришли», *пришли есте* «вы пришли», *пришла есѣ* «мы двое пришли», *пришла еста* «вы двое пришли», но в 3 л. *пришель* ∅ «он пришел», *пришла* ∅ «она пришла», *пришло* ∅ «пришло», *пришли* ∅ «они пришли» [Зализняк 1993: 285].

Таблица 1.1

Энклитические формы глагола ‘быть’ в древнерусском языке, по [Зализняк 1993; 2008]

	SG	PL	DU
1	есмь	есмь	есѣ
2	еси	есте	еста
3	∅	∅	-

Указанная дистрибуция свойственна именно энклитикам настоящего времени, а не полнозначным формам ‘быть’, которые были в древнерусском языке ударными словами и в 3 л. реализовались в виде *есть* и *суть*.

Таблица 1.2

Полноударные формы настоящего времени глагола ‘быть’ в древнерусском языке, по [Зализняк 1993; 2008]

	SG	PL	DU
1	есмь	есмь	есѣ
2	еси	есте	еста
3	есть	суть	–

Наличие в 3 л. настоящего времени парадигматического противопоставления энклитик и полноударных форм глагола ‘быть’, вообще говоря, является основанием для признания этих серий форм разными словами. Примеры внешне выраженной связки в 3 л., типа *пришель есть* вместо *пришель* ∅, по А. А. Зализняку [там же], есть лишь в книжных древнерусских текстах, но эти тексты в любом случае не относятся к тому же самому идиому, который описывается соотношением таб. 1.1 и таб. 1.2. Описание А. А. Зализняка, которое мы воспроизвели, добавив в него категорию ‘∅’, соответствует постулату (iv) и теоретически нейтрально в том плане, что оно совместимо с разными подходами. Но это не значит, что

²⁰ По своей функции л-форма в древнерусском языке скорее является нарративом, чем перфектом.

оно не опирается на модельные₁ понятия. Не только ‘Ø’ (предполагаемая нулевая словоформа), но и «связка» и даже «древнерусский язык» — модельные₁ понятия, нетривиальным образом взаимодействующие с другими понятиями. Начнем с символа ‘Ø’. За ним в таб. 1.1 может стоять три разных явления: А) дефектная морфологическая парадигма (morphological gap), ср. отсутствие по разным причинам формы 1 л. ед. ч. настоящего времени у таких глаголов, как *тошнить* и *жужжать*, в современном литературном русском языке; В) нулевая словоформа; С) эллипсис внешне выраженных форм 3 л. *есть* и *суть*. Первые две гипотезы апеллируют к понятию «парадигма слова», и их можно выдвигать только в том случае, если клитика *быти* и полноударный глагол *быти* — разные слова, у которых есть разные парадигмы. При этом модельное₁ представление о дефектной парадигме, обходящееся без гипотезы В, не менее сложно, чем модели₁ синтаксиса, где использовано представление о нулевом синтаксическом элементе. Что касается гипотезы С, она как будто отсекается решением А. А. Зализняка отрицать существование в описываемом идиоме древнерусского языка, где закон Ваккернагеля действовал в полном объеме и связочных клитик 3 л. не было (далее будем обозначать данный идиом сокращением «ДР-1»), варьирования *пришел* Ø «он пришел» ~ **пришел есть* «то же». Тексты, где такое варьирование разрешено (далее будем обозначать соответствующий идиом сокращением «ДР-2»), конечно, есть. При этом ДР-1 и ДР-2 иногда соответствуют разным регистрам одного и того же древнерусского памятника: так, в «Повести Временных Лет» и в «Киевской Летописи» летописцы от себя пишут на ДР-2, но персонажи их летописей говорят на ДР-1 [Зализняк 1993: 280; 2008: 23–24]. Процедуру выделения идиомов ДР-1 (содержательно — «разговорная форма древнерусского языка, близкая к устной речи») и ДР-2 (содержательно — «книжная форма древнерусского языка, испытавшая влияние церковнославянской традиции») по А. А. Зализняку можно критиковать. Но в пользу постулирующей различие ДР-1 ~ ДР-2 модели₂ А. А. Зализняка, объясняющей распределение энклитик и полноударных слов, а также нулевых vs ненулевых форм связки, говорят два соображения. Во-первых, эта модель₂ хорошо предсказывает употребление энклитик и связочных форм в памятниках русского языка XI–XV вв. Во-вторых, есть тексты, хотя их и относительно немного (берестяные грамоты), где регистра ДР-2 нет. Поэтому произведенное А. А. Зализняком моделирование древнерусского языка как идиома, близкого к ДР-1, операционально и может быть использовано для решения ряда задач. Оно не нейтрально теоретически. Не нейтральными при ближайшем рассмотрении оказываются и все прочие модели ‘древнерусского языка’ как набора идиомов, существовавшего в обществе, где параллельно использовался другой язык — церковнославянский²¹.

²¹ Допущение, что эти языки были распределены строго дополнительно — на древнерусском языке книжные памятники вообще не писались, а на церковнославянском языке в Древней Руси вообще не говорили — известно, как гипотеза *диглоссии* [Успенский 1994: 4–10]. Разумеется, это тоже модельное₁ понятие и модельное₁ допущение. Большинство русистов по ряду причин — от культурно-исторических до идеологических — не готово признать все письменные памятники

Если гипотеза В для идиома ДР-1 принимается, резонно уточнить, нулевой формой чего именно является ‘Ø’ в 3 л. настоящего времени в конструкции перфекта. Противопоставление употреблений, отраженных в таб. 1.1 и таб. 1.2, идет по линии клитика (Ø) vs неклитика (*есть, суть*). Подвох состоит в том, что модельное₁ понятие клитики *исключает наличие нулевой формы*: этот термин указывает на элементы, имеющие фонетическую форму и свойство «безударность» / «просодическую дефектность» [Zwicky 1977]²². Тем самым свойство «иметь нулевую форму» характеризует не статус клитики, а статус связки, точнее, вспомогательного глагола, в конструкции древнерусского л-перфекта. А. А. Зализняка упрекали и за то, что он смешал терминологически «связку» (компонент структуры именного сказуемого) и «вспомогательный глагол» (компонент аналитической глагольной формы или конструкции), различение которых свойственно исторической славистике. Заменить ярлык нетрудно. Из описания древнерусского материала, однако, следует, что дистрибуция нулевой и ненулевой форм глагола ‘быть’ в именном сказуемом напоминает распределение соответствующих форм в конструкции перфекта [Зализняк 2008: 256–261]. Тем самым использованное понятие «связки» тоже является модельным₁, указывая на употребления, сближающие клитики и компоненты именного сказуемого в ДР-1.

Таблица 1.3.

Связки и связочные клитики в древнерусском языке (на основе идиома ДР-1)

Связки		Не-связки
Клитики	Не-клитики	
пришелъ Ø, *пришелъ есть	Гюрги есть силенъ ~ Гюрги Ø силен	Въверице твое въ городе оу Жирослава соуть

Нужны ли эти исторические сведения для описания современного русского языка, в частности, для решения вопроса о том, сколько глаголов ‘быть’ в современном русском языке и стала ли связка ‘быть’ в нем самостоятельным словом? С одной стороны, памятники XI–XV вв., и конкретно идиом ДР-1, не находятся в отношении непосредственной преемственности с литературной нормой современного языка, истоки которой обычно прослеживают не далее XVIII в., когда предположительно наметилось объединение книжных и некнижных идиомов русского языка. С другой стороны, дистрибуция нулевой и ненулевой связки в ДР-1 сходна с ситуацией в современном языке. Более того, в книжной форме современного языка употребляется не только этимологически связанная с 3 л. ед. ч. форма *есть*, но и исконно связанная с 3 л. мн. ч. реликтовая форма *суть*. [МАС] признает для современного языка форму *суть* стилистическим вариантом формы *есть*, на что Ю. Д. Апресян справедливо возражает, что у русского *суть*, в отличие

Древней Руси («церковнославянскими» или «гибридными»), при этом в разных учебных пособиях и специальных трудах границу между «древнерусским» и «церковнославянским» в области грамматики проводят по-разному, ср. [Горшков 1969: 32–38; Филин 1981: 7–33].

²² Можно ad hoc задать такую дефиницию клитик, которая допускает наличие нулевой формы. Но это уже будет совсем другое понятие.

от русского *есть*, нет несвязочных употреблений [Апресян 1996: 528]. Предложения типа (3а–б) неграмматичны, при том что *есть* может использоваться в экзистенциально-посессивном контексте типа (3с).

(3) РУССКИЙ

- а. *Твои товары **суть** в городе у Жирослава.
- б. *У Жирослава **суть** основания не отдавать тебе товары.
- с. У Жирослава **есть** основания не отдавать тебе товары.

Неодинаковый статус форм *есть* и *суть* в действующей норме русского языка не является чем-то совершенно новым по сравнению с предшествующими состояниями языка. Ср. характерный пример из текста образованного автора XVIII в., А. Н. Радищева (1773 г.), где *суть* при именном подлежащем во мн. ч. (*возвышения*) используется как связка, а *есть* при местоименном подлежащем во мн. ч. (*они*), отсылающем к той же именной группе во мн. ч. (*возвышения_i... они_i*), — как бытийный глагол:

(4) РАННЕНОВОРУССКИЙ

Первая представляет везде землю ровную, плоскую, и **возвышения_i**, где **они_i есть, суть** не иное что, как берега рек к укреплению вод [А. Н. Радищев. Сокращенное повествование о приобретении Сибири [отрывок] (1773)].

Мн. ч. подлежащего в примере (4) не мешает употребить автору историческую форму ед. ч. Это не только индивидуальная черта. Как показали произведенные нами подсчеты на материале Национального корпуса русского языка (НКРЯ), на период 1700–1800 гг. приходится не менее 340 примеров употребления полнозначной формы *есть* в значении 3 л. мн. ч. (более 30% от общего числа употреблений полнозначной формы *есть* при именном или местоименном подлежащем). Есть и обратные случаи, когда форма *суть* в функции связки спорадически используется в значении 3 л. ед. ч. и даже 2 л. ед. ч. Такие примеры встречаются не только у автодидакта В. Н. Татищева (1658–1750), но и у наиболее образованных писателей — А. Д. Кантемира (1708–1744), Д. И. Фонвизина (1745–1792), И. А. Крылова (1769–1844) и того же А. Н. Радищева (1749–1802), которые были обучены церковнославянскому языку, но при этом позволяли себе употребления вроде (5) — (8).

(5) РАННЕНОВОРУССКИЙ

Вы **суть** заблудшие овцы, коих против воли желаем мы возвратить на пажить господню [И. А. Крылов. Почта Духов, или Ученая, нравственная и критическая переписка арабского философа Маликульмулька с водяными, воздушными и подземными духами (1789)].

(6) РАННЕНОВОРУССКИЙ

Что же склада и силы стихов касается, в том полную власть всяк имеет осудить, охулить и исправить, понеже **суть** юношества произведение и плод весьма малой науки [А. Д. Кантемир. Предисловие на пятую сатиру (1731–1743)].

(7) РАННЕВОРУССКИЙ

Словом, обнаружил он сердце свое и тем хотел сделать услугу человечеству, показав ему в самой слабости, каково **суть** человеческое сердце [Д. И. Фонвизин. Письма П. И. Панину (1777–1778)].

(8) РАННЕВОРУССКИЙ

А местами берег Оки, хотя низкий, являет гранит, что побуждает думать, что и вся подстилка сего возвышения **суть** гранит [А. Н. Радищев. Записки путешествия из Сибири (1797)].

Примеры (5) — (8) показывают, что в ряде идиомов русского языка XVIII в. форма *суть* употреблялась просто как связка настоящего времени, а не как согласовательный показатель. Те же авторы в других случаях используют форму *суть* в ее исконном этимологическом значении 3 л. мн. ч. Но возможно, употребления (5) — (8) и употребления того же автора, где формы *есть* и *суть* правильно распределены по 3 л. ед. ч. и 3 л. мн. ч., — это разные регистры или «штили» русского языка, аналогичные идиомам ДР-1 и ДР-2 в текстах древнерусских летописцев. В любом случае, примеры (5) — (8) ничем не отличаются от разбираемого Ю. Д. Апресяном [Апресян 1996, 528] примера из М. Горького *Сие не **суть** угроза*. Утверждение Ю. Д. Апресяна, будто этот пример М. Горького и аналогичные ему примеры, на основании которых [МАС] признал реликтовую форму *суть* в значении 3 л. ед. ч. стилистическим вариантом связки *есть*, стоят «за гранью литературной нормы», вызывает вопросы.

Нормы какого именно языка — церковнославянского, древнерусского, русского языка XXI, XVIII, XX вв. — нарушены? Как оценить правильность употребления реликтовой формы? Как далеко надо смотреть назад? Стоит ли оценивать «правильность» реликта *суть* с позиций церковнославянской грамматики (она влияла на русскую по крайней мере до середины XIX в.) и таб. 1.2 (она характеризует первые письменные тексты древнерусского языка)? В перспективе множественности описаний это одно из возможных решений. Но тогда нужно оценить пример из М. Горького и в контексте непрерывной традиции употребления формы *суть* в 3 л. ед. ч., которая по корпусным данным прослеживается, по крайней мере, до XVI в. В памятнике светского характера, «Дневнике» русского путешественника конца XVII в. П. А. Толстого (1645–1729), Исторический корпус НКРЯ указывает 10 примеров употребления *суть*, из них в 6 случаях из 10 оно употреблено в значении 3 л. ед. ч.

(9) СТАРОРУССКИЙ

Варшава **суть** место великое, на левом берегу реки Вислы положенное. [Путешествие стольника П. А. Толстого по Европе. 1697–1699 (1699)].

П. А. Толстой не подлаживался под книжную речь, он просто употребляет связки так, как было принято в его идиоме. Макароническое употребление *суть* рядом с *есть* в памятнике веком старше (*яже **суть** и ныне **есть***) может быть сбоем. Но этот сбой отражает неуверенность писца в том, что церковнославянской формы

суть достаточно в экзистенциально-локативном контексте, где русский язык его времени предпочитал форму *есть*, независимо от числа.

(10) СТАРОРУССКИЙ

И оттоле бысть страх велии на поганых Нѣмецъ, и дасть Магнушъ король свеиской на собя писмо и клятву, отнюдь никако не приходити на Русь воиною и до скончания вѣка, яже **суть** и доныне **есть** в Великомъ Новѣгородѣ во святеи Софеи» [Окончание списка Оболенского (1562–1612)].

Для периода 1500–1700 гг. Исторический корпус НКРЯ, по нашим подсчетам, дает 148 примеров (большинство относится к XVII в.) формы *есть* при существительном во мн. ч. или при сочиненной группе типа *писмо* и *клятва*, что соответствует 25,5% от общего числа случаев употребления формы *есть* в несвязочной позиции. В период 1300–1499 гг. выгрузка меньше, но меньше и самих текстов. Знаменательно, что 24 из 124 несвязочных примеров *есть* использовано в значении мн. ч. (19%): 23 примера в значении 3 л. мн. ч., см. (11) — (12) и 1 пример употребления в значении 2 л. мн. ч., см. (13):

(11) СТАРОРУССКИЙ

Два бо **пути есть** разума: ли глаголати, или молчати правѣ [Послание мудрого Феофана монаху Прохору (1375–1400)].

(12) СТАРОРУССКИЙ

Духовные, господине, и купчие, и даные грамоты у нас нет, а **есть**, господине, у нас на ту землю **знахори Прокофей Шипулин да Легчан Кондратов сын**, а се, господине, тот Прокофей и Легчан пред тобою [Правая грамота суда Ивана Сухого, по докладу великому князю Ивану Васильевичу, митрополиту Геронтию на землю у реки Конши, в Костромском уезде, бывшую в споре у игумена митрополичьего Пречистенского монастыря Василия с детьми боярскими Д. Рожновым с братьею (1473–1489)].

(13) СТАРОРУССКИЙ

«Господиа моя, аще въспросят мѣнящи из мѣшець от иконъ, что сътворю? А **вы** кто **есть**?» [Повесть о Петре, царевиче ордынском (1470–1490)].

Итак, распространение формы *есть* на мн. ч. и превращение ее в обобщенную презентную форму в позиции полнозначного глагола — явление давнее, его отражают уже тексты XIV–XV вв., как светские, так и церковные. Для *суть* НКРЯ с 1300 г. дает крайне мало употреблений в экзистенциально-посессивных и локативных контекстах типа <*В X-е суть Y-и> или <*У X-а суть Y-и>²³. Конкуренция *есть* и *суть* статистически значима лишь в подлежащих связочных конструкциях: в бесподлежащих связочных предложениях с несогласуемым предикативом

²³ Из исключений см. в переводном тексте, где в первом предложении *суть* употреблено без подлежащего в им. п. Далее при выраженном подлежащем *птици*, стоящем в им. п., переводчик сбивается на форму *есть*: «И *мышев* бо множицею в домовех царских *суть*, но непотребни

(тип *прискорбно естъ*) форма *суть* почти не встречается, за вычетом единичных сбоев. Тем самым, при попытке найти в диахронии истоки дистрибуции *есть* vs *суть* в современном языке, где *суть* сохраняется как факультативный вариант связки *есть*, обнаруживается примерно то же распределение в аналогичных конструкциях, но на 500–600 лет раньше.

В древнерусском языке энклитические формы глагола *быти* были показателями лично-числового согласования, порой единственными, поскольку в предложениях типа *приишел есмь* «я пришел» внешне выраженное подлежащее при наличии клитик было необязательно, а в идиоме ДР-1 обычно опускалось [Зализняк 1993: 291]. Напомним, что идиомы ДР-1 и ДР-2 отличаются набором связочных энклитик: в ДР-1 энклитик 3 л. в конструкции перфекта не было, а в ДР-2 энклитические формы *есть* и *суть* спорадически употреблялись [Зализняк 2008: 257–259]. В конструкции именного сказуемого оба древнерусских идиома использовали энклитики 1–3 л., но в ДР-2 их употребление было обобщено. ДР-1 допускал вариацию типа *Гюрги естъ силенъ* ~ *Гюрги ∅ силенъ*, в то время как ДР-2 требовал внешне выраженной связки [там же].

Синхронные лингвисты порой ссылаются на «исторический фактор», когда не могут полностью объяснить сложившееся распределение. Со времен Н. С. Трубецкого и Р. О. Якобсона принято подчеркивать, что языковые изменения а) системны и б) носят чуть ли не телеологический характер [Jakobson 1931; Якобсон 1985: 132; Trubetzkoy 1931]. За вторым утверждением стоит оптимистическое модельное₁ допущение, что естественный язык устроен так, что эволюция любого языка предопределена внутренней логикой развития его фонологической и/или грамматической системы, стремящейся принять равновесное (оптимальное) состояние. Это допущение тесно связано с другой программной идеей ранних структуралистов — о том, что выделение фонологической и/или грамматической *системы*, т. е. сетки отношений между элементами одного и того же уровня, является самодостаточным средством описания языка в синхронии: тем самым структуралистский рецепт, как надо описывать изменения языка в диахронии, строится на заданных заранее представлениях о том, как язык устроен в каждом синхронном срезе. Такая связь модельных₁ допущений, каждое из которых само по себе неочевидно, побуждает подозревать, что структурализм — это не только инвентарь методов, но и интегральная теория языка, пусть и вышедшая из моды.

Дает ли история форм *есть* и *суть* в русском языке XV–XVIII вв. ключ к современной русской литературной норме и новые данные о древнерусском состоянии? Произведенный экскурс в историю употреблений *суть* и *есть*, видимо, опровергает два предубеждения. Во-первых, использование форм *есть* и *суть* не в функции показателей лично-числового согласования, а в функции презентных маркеров экзистенциальных и связочных конструкций не есть инновация современного русского периода (XIX в. по настоящее время). Сходные идиомы были и в XVIII в.

пребывают, аще и близ суть; птици же, нарицаемии фалкон, аще дивии естъ, но за потребы своя призываются и приемлются и на царских руках съдят [Стефанит и Ихнилат (1300–1500)].

(обозначим соответствующий идиом сокращением «РНР-1» = «ранненоворусский идиом 1») и в XV–XVII вв. (обозначим соответствующий идиом сокращением «СР-1» = «старорусский идиом 1»). Во-вторых, хотя форма *суть* и является в современном языке реликтом, напрямую оценивать правильность употреблений этой формы в паре *суть* ~ *есть* с парадигмами глагола *быти* в церковнославянском языке и древнерусских идиомах ДР-1 и ДР-2 неверно: в этих языках формы глагола *быть* выступают в роли показателей лично-числового спряжения, а в современном русском языке сопоставимой функции у них нет. Поэтому с точки зрения описания ровно двух форм — *есть* и *суть* — действительно можно рекомендовать лингвистам не заглядывать дальше конца XIV–XV вв.

На этом можно было бы поставить точку, если бы принцип множественности описания действовал только в синхронной лингвистике. Но он столь же важен в диахронии. Если смотреть не на две, а на все три презентные формы глагола *быти* — \emptyset , *есть* и *суть*, то все-таки придется вернуться к соотношению древнерусских идиомов ДР-1 и ДР-2. Идиом ДР-1 не исчез вместе с берестяными грамотами: его диахроническим «следом», как можно предполагать, и является нулевая связка ‘ \emptyset ’, распространившаяся на все настоящее время. Сходный путь проделали ненулевые формы *есть* и *суть*, в силу чего менее частотная из них стала избыточной. Но это уже диахронический след идиома ДР-2, связанного с книжной традицией. Слегка формализуем эту гипотезу в виде положений (vi) — (viii).

- (vi) Не различающая лиц и чисел нулевая форма настоящего времени глагола *быть* в современном русском языке (СРЯ) является непосредственным продолжением не различавшей числа нулевой формы 3 л. ‘ \emptyset ’, использовавшейся в значении 3 л. ед. ч. и 3 л. мн. ч. в древнерусском идиоме ДР-1 и распространившейся на 1–2 л. в идиомах СР-1 и РНР-1, покрывающих период от 1350 до 1800 гг. Существует непрерывная диахроническая связь между идиомами ДР-1 \Rightarrow СР-1 \Rightarrow РНР-1 \Rightarrow СРЯ.
- (vii) Формы настоящего времени *есть* и *суть*, первая из которых в СРЯ не различает числа и лица, а вторая различает их пережиточно, являются продолжением показателей лично-числового согласования в 3 л. ед. ч. *есть* и 3 л. мн. ч. *суть* в древнерусском идиоме ДР-2, связанном с книжной традицией. Существует диахроническая связь между идиомами ДР-2 \Rightarrow СР-2 \Rightarrow РНР-2 \Rightarrow СРЯ.
- (viii) Полная совокупность употреблений презентных форм глагола *быть* в СРЯ диахронически соответствует не одному, а двум разным идиомам. Та же гипотеза может быть применена к усредненному представлению о других временных срезах русского языка с 1350 до 1800 гг.

Модельное₁ представление о том, что в текстах одного и того же говорящего / пишущего могут выделяться разные регистры и идиомы, является для смежной науки, филологии, обычным. Заслугой А. А. Зализняка является то, что он связал выделение идиомов ДР-1 и ДР-2 с параметризацией синтаксиса древнерусского языка и использовал ее для построения предиктивных моделей₂, предсказывающих употребление

энклитик и неэнклитических форм глагола *быти*. Та же модель может применяться к идиомам старорусского и ранненоворусского периодов. Книжные идиомы «второй линии» — ДР-2, СР-2 и РНР-2 — генерализовали употребление ненулевых форм глагола ‘быть’ во всех трех лицах в конструкции перфекта с причастием на *-л* и в составе именного сказуемого, независимо от того, были ли эти формы энклитиками или нет. В модельных¹ терминах это означает, что во всей серии книжных идиомов второй линии формы 1–3 л. настоящего времени глагола ‘быть’ имели статус показателей лично-числового согласования. В некнижных идиомах «первой» линии — ДР-1, СР-1 и РНР-1 — энклитик 3 л. в конструкции перфекта не было изначально, а в составе именного сказуемого в ДР-1 они употреблялись редко. Тем самым, для носителя идиома «первой» линии, осваивавшего идиом «второй линии» с целью порождения книжных текстов, ненулевые формы 3 л. *есть* и *суть* были элементами чуждого ему языка. К тому же, к концу старорусского периода, причастия на *-л* типа *прииель <есь, еси>* вследствие отмирания энклитик 1–2 л. изменили свой статус в идиоме СР-1 и из именного компонента аналитического перфекта в XVI–XVII вв. превратились в собственно глагольные формы [Зализняк 2008: 255]. Поэтому сохранение в идиомах СР-2 и РНР-2 конструкции аналитического перфекта с показателем лично-числового согласования во всех трех лицах стало двойной нагрузкой для носителей идиомов «первой» линии, которые допускали грубые ошибки: они sporadически употребляли связку не только при церковнославянских формах аориста, см. ниже пример (14), и имперфекта, путали 3 л. *суть* с причастием *сущи*, но и в форме настоящего времени, ср. аномальный пример (14).

(14) СТАРОРУССКИЙ

Они же начаша дивитися, како се увѣдѣ, о немже они **придоша суть**, но понеже познаша челоуѣка того Божиа быти, и посланиа давшѣ от князя святому [Житие Кирилла Белозерского (1450–1455)].

(15) СТАРОРУССКИЙ

И погребень бысть честно близ келии святаго у дуба на Покров Святыя Богородица, еже и донынѣ **суть ту пребывает** под церковию Благовѣщения Пресвятыя Богородица во славу Христову, ему же слава [Житие Никандра Псковского (1595–1605)].

Встречаются и ошибки, связанные с паразитическим употреблением формы 3л. ед. ч. *есть*, ср. пересказ эпизода «Повести временных лет» в летописи начала XV в.:

(16) СТАРОРУССКИЙ

И видѣ ту люди сущая, какъ **есть** обычаи **имуть**, и како ся мыют и хвошутся; и удивися имѣ [Новгородская Карамзинская летопись. Первая выборка (1400–1450)].

Макаронические употребления (14) — (16) — факты иного рода, чем употребления (5) — (13), где *есть* и *суть* использованы как презентные формы глагола ‘быть’ вне связи со значениями конкретного лица и числа. Первые из упомянутых — это

ошибки носителей идиомов «первой» линии при освоении книжных идиомов «второй линии». Вторые из упомянутых — вероятный результат *ассимиляции* ненулевых форм *есть* и *суть* в идиомах «первой» линии. Распространение формы *есть* на мн. ч. в несвязочной позиции носит массовый характер. В подлежащих связочных структурах с вторичным согласованием числовая форма именного компонента сказуемого подсказывает, которую из форм надо избрать в идиоме «второй линии», но ошибки бывают и здесь. Так, архимандрит Дионисий, носитель СР-1, неудачно перейдя на СР-2 в 1658 г., породил фразу *нынѣ есть сдѣланы*, см. (17). Эта ошибка предсказуема на основе гипотезы о том, что Дионисий был носителем СР-1, поскольку в СР-1 в данной позиции употреблялась нулевая связка (*что нынѣ ∅ сдѣланы*), а пересчете на СР-2 интуиция не подсказала Дионисию, которую из книжных форм — *есть* или *суть* — надо избрать.

(17) СТАРОРУССКИЙ

А жернокововъ бы тебѣ кормить и поить добрѣ, чтобъ имъ ничимъ скудно не было; а камневь убавливать, что нынѣ **есть сдѣланы**, не велѣтъ; да вамъ же бѣ велѣтъ тотчасъ камень дѣланое къ мельницы привозить, чтобъ имъ тутъ камень справить [Память архимандрита Иверского монастыря Дионисия посельским старцам Аврамию и Алимпию о даче жерноковам в помощь работника, вырубке в Вышнем Волочке валов и о привозе их в Выдропуск на мельницу (1658)].

Итак, апелляция к диахронии не снимает неясностей синхронного описания. Зато синхронное и диахроническое описание одного и того же материала дополняют друг друга. Взаимодополняющими могут быть и разные версии диахронического анализа. Здесь тоже есть неустранимые модельные₁ понятия и сохраняется возможность выдвигать предиктивные модели₂, которые до известной степени проверяемы. При ретроспективном анализе, в том числе при реконструкции эволюции фрагментов грамматических систем, глубина рассмотрения зависит от избранного материала. В избранной иллюстрации (соотношении презентных форм *есть*, *суть* и нулевой связки '∅' в литературной норме современного русского языка) частичная синонимия ненулевых форм *есть* и *суть* — относительно поздний факт, который отражает распределение, сложившееся в XV–XVII вв., но не всегда последовательно отраженное в текстах. Противопоставление нулевой и ненулевых форм ∅ vs *есть* ~ *суть* — более древнее явление, оно прослеживается вплоть до раннедревнерусского состояния (XI–XII вв.). Представление о единстве литературной нормы / усредненном употреблении языка в некотором синхронном срезе, как и альтернативная гипотеза о сосуществовании разных идиомов, которым соответствуют разные грамматические системы, являются модельными₁ допущениями.

В заключение, на «десерт», еще об одном эпизоде полемики о двух разных vs едином глаголе 'быть', где выдвигался скрытый диахронический аргумент. Полемизируя с точкой зрения о том, что связка 'быть' в современном русском языке является автономной лексемой, отличной от полнозначного глагола 'быть', Ю. Д. Апресян в 1992 г. обратил внимание на то, что основы БЫТЬ и ЕСТЬ распределены

«строго дополнительно» относительно категории времени: основа БЫТЬ представлена в прошедшем и будущем времени, а ЕСТЬ в настоящем, что соответствует определению морфологического супплетивизма и, по мнению Ю. Д. Апресяна, делает гипотезу об автономности связки лишней [Апресян 1996: 528]. Это уже экскурс в реалии примерно 5000-летней давности, когда в праиндоевропейском языке в парадигме глагола ‘быть’ сформировался супплетивизм основ **bhē/*bhū* (ср. рус. *был, будет*, англ. *be* и т. д.) и **es* (ср. рус. *есть, суть*, др. рус. *есмь, еси, есть, есмь, есте, суть, есвѣ, еста*, англ. *am, is* и т. д.). Могут возразить, что тезис Ю. Д. Апресяна непосредственно относится к современному русскому, а не к праиндоевропейскому или праславянскому языкам. Но все дело в том, что на основе строго синхронного анализа принципиально невозможно установить, что нулевая связка ‘Ø’ соотнесена именно с основой **es*, а не с каким-то иным рядом форм. Это — *этимологическое соображение*, не так важно, возникло ли оно в результате изучения исторической грамматики или стало результатом созерцания единственной оставшейся в русском языке активно используемой формы от основы **es* — рус. *есть*. Правда, догадаться о том, что и реликтовая форма *суть*, наличие которой в книжном варианте нормы признает сам Ю. Д. Апресян, тоже восходит к презентной основе **es*, и что соотношение *есть* ~ *суть* не является случаем супплетивизма, без специальных знаний трудно, как и в случае фонетически далеких форм типа англ. *is* ~ *am*, нем. *ist* ~ *sind* и т. п. Можно, конечно ввести соответствующее решение декретом, объявив, что связочные и несвязочные структуры записываются примерно так:

(18) РУССКИЙ

- a. Бодуэн — Ø^{ЕСТЬ. НАСТ. ВР} гений.
- b. У Бодуэна Ø^{ЕСТЬ. НАСТ. ВР} много работ.
- c. Бодуэн был^{БЫТЬ. ПРШ. ВР} гением.
- d. У Бодуэна было^{БЫТЬ. ПРШ. ВР} много работ.

Но это доказывает лишь то, что интегральное описание русского языка по Ю. Д. Апресяну — глобальная теория, которая использует модельные₁ допущения о соотношении словаря и грамматики, супплетивизме, нулевом морфе и т. д. Сторонник гипотезы о двух глаголах ‘быть’ спросит, почему нельзя пользоваться нотацией (19) или даже (20): последний вариант связан с отказом признавать нулевую форму полнозначного глагола:

(19) РУССКИЙ

- a. Бодуэн — Ø^{СВЯЗКА. НАСТ. ВР} гений.
- b. У Бодуэна Ø^{БЫТЬ. НАСТ. ВР} много работ.
- c. Бодуэн был^{СВЯЗКА. ПРШ. ВР} гением.
- d. У Бодуэна было^{БЫТЬ. ПРШ. ВР} много работ.

(20) РУССКИЙ

- a. Бодуэн — ØСВЯЗКА. НАСТ. ВР гений.
- b. У Бодуэна ØСВЯЗКА. НАСТ. ВР много работ.
- c. Бодуэн был СВЯЗКА. ПРШ. ВР гением.
- d. У Бодуэна было БЫТЬ. ПРШ. ВР много работ.

Итак, историческое соображение о супплетивизме ничего не добавляет к тщательно продуманной и убедительной аргументации Ю. Д. Апресяна в пользу тезиса о том, что при лексикографическом подходе гипотеза о наличии русского глагола 'быть' заслуживает предпочтения. Возможно, это просто инженерное решение, обоснование разметки, принятой в некоторой прикладной системе (синтаксическом словаре, парсере и т. д.). Но при перенесении его в теоретическую плоскость оно становится лишней информацией.

VI

Чем отличается провозглашенный нами приоритет множественности описания при отказе от последовательного применения интегральных теорий языка от того, что в гуманитарных науках иронически называется «постмодернизмом»? Скатиться к постмодернизму и эклектике — реальная опасность. Так, возможно, и произойдет, если лингвист не будет строить предсказательные модели и проверять их или если он, доверившись борцам с глобальными теориями, заранее, до попытки анализа, сбросит с «парохода современности» все модельные понятия, которые накопила лингвистика. Ученый, так поступающий, подобен человеку, внявшему совету М. Кузмина, которому, впрочем, он сам не последовал: *рассуждая логически, всем нам надо повеситься, а там видно будет*. Трагедии, возможно, не произойдет, но применять модели заставит сам материал языка: отказавшись от того, чтобы делать модельные допущения самостоятельно, лингвист останется с теми допущениями, которые за него сделали другие люди, которых не устраивают те или иные модельные понятия. Если этим последним не понравилось понятие клитики, которое является частью тернарной классификации морфологических элементов морфема vs клитика vs словоформа, он останется с бинарной классификацией морфема vs все-что-не-является-морфемой, которая тоже опирается на модельные допущения. Не в дверь, так в окно. Программа войны на два фронта — и с глобальными теориями, и с универсальными понятиями в описании — еще окончательно не оформилась. Но при усилении догматических черт она имеет шансы стать очередной глобальной теорией языка (неясно, притязает ли она на эту роль сейчас).

Было бы правильно закончить обсуждение, применив постулат о приоритете множественности описаний к предмету этой статьи. Не отказываясь от тезиса о недостатках глобальных теорий языка, усредняющих и уплощающих описание языковых явлений, выдвинем дополнительный тезис. Возможно, дело не в самих теориях, а в том, как люди их применяют, в *человеческом факторе*. Разработчики

и пользователи лингвистических моделей, как показывает наш экскурс, склонны делать две ошибки. Во-первых, многие лингвисты склонны относиться к модельным понятиям, а иногда и к целым теориями, как к волшебной палочке или паролю, который откроет все языковые тайны сам, избавив от скрупулезного анализа. Если при произнесении слов типа: «фокус», «рема», «клитика», «перемещения», и т. п. — такого эффекта не происходит, многие разочаровываются. Во-вторых, разработчики глобальных теорий, в той или иной мере использующие формализацию, математическое моделирование и/или занимающиеся подготовкой лингвистических ресурсов — корпусов текстов, баз данных и/или разработкой лингвистического процессора, склонны смешивать инженерные вопросы с концептуальными. И это наша общая беда, страдают этим все эксперты в области грамматики, лексикографии, типологии и корпусные лингвисты. При составлении словаря или базы данных, разметке корпуса, применении алгоритма построения дерева предложения, определении эллипсиса, разрешении неоднозначности или выделения анафорической связи какие-то допущения конструктивно необходимы или целесообразны. Но лингвист-теоретик утверждает, что так «на самом деле» устроен сам язык.

Одна из наиболее влиятельных книг в лингвистике XX в. [Chomsky 1965] в своем названии содержит слово «аспект» и в качестве опыта аспекттивного описания грамматики интересна. Судьба положенной в основу этого описания модели, а также моделей, разработанных автором упомянутой книги и его соратниками, равно как и противниками, позже и обросших догматическими ограничениями, показывает, что интегральные теории сами по себе не решают кардинальных проблем лингвистики, связанных с точным предсказанием языковых черт. Лучшее, на что можно надеяться, — это то, что сторонник одной из интегральных теорий, отказавшись от части навязанных догм, предложит новое аспекттивное описание, которое может быть использовано всеми лингвистами.

[2] ЧТО И ЗАЧЕМ. НАУЧНАЯ СТАТЬЯ. НАУЧНЫЕ СТАТЬИ

Аннотация: Обсуждается жанр лингвистической статьи как формы исследования, ориентированной на аспекттивные описания явлений, поддающихся изучению с позиции тех или иных моделей языка. Продуктивность статьи обусловлена тем, насколько точно определен объект анализа. Обсуждается пять проблемных направлений — гармонизация методов исторической и синхронной лингвистики, грамматическая семантика, субъектно-предикатные отношения, типология линейных порядков и анализ текста с применением лингвистических методов.

Ключевые слова: научное исследование, лингвистика, доказательность, частная лингвистика, общая лингвистика, типология.

Лингвист <математику>:

— Так почему Вы занимаетесь ...-ми? Что специально в них интересного?

Математик:

— У них есть свойства, которые позволяют их изучать.

I. Факты и модели

Вынесенный в эпиграф разговор — был. У математика и у лингвиста были имена. Хотя это неважно. В естественном языке тоже есть объекты, обладающие свойствами, о которых говорил математик: *группы явлений*, которым соответствуют *модели*. Умение выделять такие объекты — условие успешного лингвистического исследования. Сменяющие друг друга версии «общей теории всего» (пародийный ярлык Станислава Лема), сводящиеся к декларациям о том, чем *на самом деле* является естественный язык, и к переходящим в жанр рекламных объявлений рекомендациям о том, к какой группировке ученых надо примкнуть, чтобы описывать язык *правильно*, таких исследований заменить не могут. Не должны заменять их и прозрения авторитетных ученых либо неофитов о месте лингвистики в кругу наук будущего, и написанные в жанре «взгляд и нечто» тексты без языковой конкретики. Увеличение числа публикаций последнего типа в журналах и сборниках связано не только с конъюнктурой, но и с глубинными потребностями академического сообщества адаптироваться к междисциплинарным задачам при недовольстве устоявшимся аппаратом своей науки. Но при знакомстве с некоторыми из подобных публикаций возникают сомнения в том, что их авторы в состоянии выделить конкретный объект исследования и анализировать,

например, прилагательные, глаголы, тональные акценты, значение лексических единиц и т. п.

Жанр текста, где решаются задачи описания языковых объектов с поддающимися изучению свойствами, известен — это научная статья. За тридцать лет работы у меня накопилось несколько опытов подобных исследований, которыми я решил поделиться. Для этой книги я отобрал 26 текстов, которые уместились в пять рубрик: принцип размещения — тематический, не хронологический. Было понятно, что квалификационные работы, фрагменты монографий и изданий литературных памятников, обзоры, рецензии и статьи из энциклопедий не подходят. Не стоит включать и работы, написанные в соавторстве, чтобы не перекладывать на других ответственность за композицию книги. Названия лингвистических статей отражают предпочтения эпохи и зависят от твердости редактора, который терпит либо не терпит предложения автора статьи¹. В эпоху позитивизма поощрялись демонстративно скромные формулировки типа «Заметка об акцентовке несклоняемого причастия на -ль» и «К истории звука Ъ в московском говоре в XIV–XVII веках» [Васильев 1905a; 1905b], а попытки обобщения в названии типа «Система мер и весов в Полоцке во второй половине XV в.» не приветствовались. Сейчас упомянутые статьи Л. Л. Васильева, оставившие след в науке, при том же охвате материала могли бы называться «Принципы фразовой акцентуации в древнерусском языке», ср. [Дыбо 1975], или даже «Универсальные принципы сегментации (на материале древнерусского языка)». Я живу не в начале XX в., среди статей, вошедших в эту книгу, одна названа «A unified analysis of clitic clusters in world's languages». Стыдно. Но не очень. Предмет названной статьи — диагностика языковых объектов некоторого типа («цепочки клитик» = «clitic clusters») в открытом классе языков мира, где выделяются объекты другого типа, предположительно связанного с первым («клитики» = «clitics»). Если кому-то предложенная процедура не нравится или он на своем материале пришел к иным выводам, применив ту же процедуру, пусть опровергнет. И пишет свою статью.

Установление, точнее — попытки установления, *общих* закономерностей какой-то группы явлений, будь то звуковые переходы, цепочки слабоударных слов или падежные системы на основе индуктивных наблюдений над некоторым числом конкретных языков, — тоже в известном смысле *частная* лингвистика. Область проверки таких попыток (= гипотез лингвистической типологии) — открытый класс языков мира, число которых неизвестно, и большинство из которых,

¹ Если отвлечься от цензуры редактора, названия научных статей и докладов всегда отражают один главный фактор выбора — набор названий других статей, докладов или книг, которые автор научного текста считает достойными подражания или полемической реакции. Так, например, написанная по-немецки статья Р. О. Якобсона «Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kasus» [Jakobson 1936], в русском переводе А. А. Холодовича — «К общему учению о падеже. Общее значение русского падежа» [Якобсон 1985: 133–175], своим названием намекает на не указанную в списке литературы книгу Б. Дельбрюка «Synkretismus. Ein Beitrag zur germanischen Kasuslehre», т. е. «Синкретизм. К теории германского падежа» [Delbrück 1907].

включая праязыки ныне существующих языковых семей, исчезло, не будучи документировано какими-либо текстами. Говорить «исчезло бесследно» все же неверно, так как методы современной сравнительно-исторической лингвистики позволяют в благоприятном случае установить отдельные соотношения между исчезнувшими языками. Параметризация многообразия в доступных сегодня для наблюдения языках *частично* восполняет этот пробел, позволяя делать прогнозы о более вероятных и менее вероятных комбинациях значений параметров. Это и делает лингвистику *естественной* наукой.

Статьи автора, включенные в данную книгу, как указано, датируются 1989–2019 гг. Каждая статья — продукт своего времени. Если они в каких-то отношениях не устарели, это значит, что попытка анализа и постановка задачи сохраняет смысл и сейчас. Частично изменились представления автора об оформлении научного текста. При включении в настоящую книгу редактирование текста статей в основном ограничивалось следующими операциями: 1) была составлена сводная библиография и упорядочено цитирование; 2) в ряде ранних статей цитируемые примеры для облегчения восприятия оформлены в виде нумерованного списка; 3) текст ряда статей разбит на параграфы, которым даны названия; 4) часть развернутых примечаний перенесена в основной текст статьи; 5) к некоторым статьям добавлены аннотации и списки ключевых слов. При подготовке главы [10] был восстановлен полный текст статьи, сокращенный в публикации 2017 г. из-за редакционных ограничений на объем. Восстановленные разделы статьи [10] содержат технические сведения о процедуре социолингвистического эксперимента и корпусной разметки и приводят количественные данные в том объеме, в котором они были обработаны к моменту публикаций, с учетом тогдашнего состояния НКРЯ. Новые подсчеты не проводились.

II. Диахрония. Синхрония. Динамика

Одна из тем, которая занимала меня на протяжении многих лет, — оптимальное сочетание методов синхронного и диахронического описания. Базовое для нашей науки различие диахронии и синхронии как разных аспектов существования языка — во времени и вне временных изменений, в одной метафизической «точке», которая в зависимости от принятой шкалы измерения может соответствовать и периоду в сто лет или более, — как известно, введено Ф. де Соссюром, автором, чье научное наследие в последние десятилетия подвергается критике. В нашей стране и до этого к фигуре Соссюра относились с подозрением. Считалось хорошим тоном, а вместе с тем подтверждением преданности отечественной научной мысли, что весьма кстати для ученых, привыкших публично демонстрировать свой конформизм, — пересказывая Соссюра, разоблачать его и напоминать, что классик «оторвал» синхронию от диахронии, систему «языка» (*langue*) от «речевой деятельности» (*langage*) и т.п., см. [Материалы 1960]. Вне зависимости от того, плох ли Соссюр (это отдельная тема), следует поставить вопрос, устарели ли проведенные им различия и можно ли их безнаказанно отменить. Я отвечаю

на этот вопрос отрицательно, так как знаю фундаментально важные задачи, которые не только допускают, но и требуют абстрагирования либо от синхронного, либо от диахронического описания. Это, соответственно, — доказательство языкового родства и системный анализ лексики и грамматики одного отдельно взятого языка, направленный на выявление действующих в любом языке в любой момент времени ограничительных условий, накладываемых на деятельность говорящих и пишущих. Встречающиеся утверждения о том, что новые технологии конца XX в., в частности, подготовленные лингвистические ресурсы — корпусы текстов языков мира, якобы отменяют жесткое разделение синхронии и диахронии, переводя основной акцент в плоскость изучения статистически значимого языкового варьирования, мне трудно воспринимать иначе, чем рекламу вспомогательной лингвистической дисциплины — *корпусной лингвистики* и ее продуктов, в частности — исторических *корпусов текстов*. Действительно, языки мира могут быть представлены не только в виде реально написанных или виртуальных словарей и грамматических описаний, но и в виде корпусов текстов. Вот только корпус, в отличие от словаря — предмета изучения теоретической лексикологии, ср. [Апресян 1974; 1996], и грамматики — предмета изучения теории грамматики, ср. [Jespersen 1924; Diderichsen 1946; Chomsky 1965; Шведова 1970], виртуальным не бывает. У бесписьменных праязыков нет корпусов текстов. Исторические корпусы большинства древних письменных языков содержат невосполнимые лакуны, а их объем слишком мал для того, чтобы давать статистически значимую выгрузку по ключевым запросам, проверяющим ограничительные условия и действующие в языке тенденции.

Вместе с тем области исследований, где приходится комбинировать методы диахронического и синхронного анализа, безусловно, есть. Это и процедура описания древних языков, и изучение предыстории современных языков. Важно понять, в каком порядке эти методы применять, какова степень доказательности полученных с их помощью выводов, в чем состоит динамика развития языка и как ее измерять. Про это — пять статей второго раздела книги, озаглавленного «Диахрония. Синхрония. Динамика».

Л. В. Щерба в конце 1920-х гг., обсуждая сосюровское противопоставление «языка как системы» (*langue*) «речевой деятельности» (*langage*) в контексте обоснования метода *лингвистического эксперимента*, заметил, что в его время было принято описывать живые языки как мертвые, в то время как правильно описывать мертвые языки как живые, применяя к ним метод эксперимента [Щерба 1931; 2008: 31–32]. Л. В. Щерба, по крайней мере в письменных работах, не настаивал на буквальном прочтении своих рабочих терминов, даже таких, как «категория состояния». То, что он называет «лингвистическим экспериментом», позже называлось «методом отрицательных языковых данных» или «введением в описание звездочки» (*). Имеется в виду тестирование действующих в любом языке ограничительных условий, можно или нельзя сказать на рус. **Лингвисты написало статью*, нем. **Lew den Aufsatz hat geschrieben*, рус. **Я увидел наук*, англ. **The boat capsized by purpose*, исл. **Bátnum hvolfði viljandi*, букв. *«лодке перевернуло умышленно»

и т. п. В современных языках отрицательные данные могут быть собраны путем интроспекции, социолингвистических методов или изучения корпусной статистики. Разумеется, в случае с плохо документированными или вымирающими языками, где непосредственный контакт лингвиста с носителями языка отсутствует или затруднен, все эти методы могут не действовать. Применительно к древним языкам интроспекция и опрос информантов обычно не считаются валидными методами², но корпусная статистика при большом количестве текстов в некоторой степени эмулирует результаты эксперимента и дает представление о том, какие комбинации элементов возможны, а каких ждать не следует. Примером такого языка служит древнеисландский, материал которого обсуждается в статье [3], озаглавленной «Die unpersönlichen Satzmodelle in der altisländischen Sprache». Об объеме словаря этого языка дает представление [Cleasby, Vígfússon 1957], а о разнообразии форм текста — [Heusler 1913; 1969].

Одна из задач диахронической лингвистики, помимо доказательства языкового родства, — выявлять архаизмы и инновации. Это задача решается не только и не столько путем анализа статистики в массивах текстов на древних языках, сколько путем внутренней реконструкции, опирающейся на системные отношения единиц языка и языковых структур. Формулируя иначе: внутренняя реконструкция может опираться на статистику, но в общем случае это необязательно. Моя работа [3] написана в 1989 г. и опубликована 3 года спустя. Я принимаю гипотезу своего учителя О. А. Смирницкой [Смирницкая 1969; 1970; Smirnickskaja 1972] о том, что необычное на фоне других древних германских языков обилие безличных моделей предложения в древнеисландских памятниках является результатом двух, причем противоположенных, процессов. С одной стороны, древнеисландский язык, в силу своей консервативности и изолированного географического положения, сохранил многие исконные черты прагерманского и поздне-индоевропейского синтаксиса, отраженные в инвентаре безличных конструкций. С другой стороны, продуктивные модели безличного предложения, связанные с оформлением семантического субъекта связочного предиката дат. п., являются поздней инновацией самого древнеисландского языка, которую без дополнительных доказательств нельзя проецировать не только в общегерманскую, но и в праскандинавскую эпоху. Мой вклад в верификацию этой гипотезы состоит в оправданном прогнозе. Я предположил, что раз древнеисландский язык обобщил трехчастную схему дативно-предикативных структур дат. п. лица — связка — предикатив, то должен существовать класс слов, обслуживающих эту схему и реализующихся преимущественно или исключительно в позиции неизменяемого именного слова при связке. Такой класс слов в русской грамматической традиции принято называть «категорией состояния». Прогноз оказался верным. Когда я в конце 1990-х гг. ознакомился

² Бывали случаи, когда знатоки древних языков отождествляли себя с их носителями и позволяли себе категорические суждения о древних языках, фактически основанные на интроспекции. К сожалению, обычно у таких ученых находились оппоненты, чья интроспекция была устроена иначе.

с картотекой Арнемагнеанского института в Копенгагене, которая регистрирует все употребления древнеисландских слов, оказалось, что в этом языке существует несколько сот неизменяемых присвязочных слов (!), которые реализуются только в позиции сказуемого и не употребляются в функции наречий и атрибутивных прилагательных. Это на порядок больше, чем в современном русском языке, где соответствующий класс слов был впервые открыт Л. В. Щербой, ср. рус. *стыдно, совестно, жаль, пора, надо*. Счастливым для древнего языка стечение обстоятельств состоит в том, что при таком количестве текстов и таком количестве предикативов отсутствие сразу у 200–300 слов наречных и атрибутивных употреблений не может быть случайным. Лакуна — не в языковых данных, а в умах германистов, которые не могли распознать класс предикативов потому, что не владели понятийным аппаратом для его выделения. Пораженный этим казусом, я в 1998 г. написал статью [8], где сопоставляются данные русского и древнеисландского языков, но уже вне связи с внутренней реконструкцией.

Статья [4] «Между синхронией и диахронией. Просодические противопоставления в скандинавских языках» (1999 г.) посвящена традиционной теме исторической акцентологии — происхождению тональных словесных акцентов в скандинавских языках. В шведском и норвежском языках и диалектах, близких к их произносительной норме, противопоставление тональных акцентов носит бинарный характер, в связи с чем используются ярлыки «акцент 1» и «акцент 2». Часть ученых проецирует это бинарное противопоставление в индоевропейские глубины, в то время как их оппоненты связывают его с поздней инновацией скандинавских языков — развитием суффигового артикля. В других германских языках суффигового артикля нет, равно как нет и словесных акцентов шведско-норвежского типа [Кузьменко 1990; Kusmenko 2008]. Дополнительное методологическое расхождение между двумя лагерями ученых, наслаивающееся на первое, состоит в том, что часть акцентологов (мейнстрим) верит в то, что звуковые признаки актуализируются, используя профессиональный жаргон — «становятся релевантными», «фонологизируются», в то время как их оппоненты утверждают, что фонологизация аллофонов и просодий — миф, и верят в то, что языку в ходе его эволюции сопутствует некоторый устойчивый набор фонетических и просодических признаков. По такой логике получается, что если в праязыке была гортанная смычка, глоттализация или восходящий тон, никуда они исчезнуть не могут. И наоборот, если, допустим, в датском, латышском и тагальском языках есть глоттализация или восходящая тональная просодия, которая не выделяется в родственных языках, мы обязаны возводить ее к праязыку. В статье [4] я предлагаю оценить дистрибуцию скандинавских просодий в эпоху распространения бесспорной инновационной черты — суффигового артикля. Выясняется, что морфологические свойства артикля в разных группах древнескандинавских языков различны: в наиболее консервативных диалектах, сохраняющих четырехпадежную систему и многосложные именные словоформы, суффиговый артикль фактически выступает в качестве второй флексии, в то время как в более радикальных диалектах, где падежная система свернута и все старые двусложные слова стали односложными,

суффиксированного артикля нет. Закрепление бинарного противопоставления акцент 1 vs акцент 2 связано со средними темпами развития диалекта, когда падежная система уже упростилась, но в парадигме существительных еще есть формы с разным числом слогов. С моей точки зрения, эти синхронные данные, если и не опровергают гипотезу об индоевропейском характере шведско-норвежских акцентов, то серьезно ее компрометируют.

В статье [5], озаглавленной «Раннегерманское ударение: Фонетика и компаративистика» (2006 г., опубликовано в 2007 г.), я посягаю на понятие звукового закона и разбираю два исторических соответствия, считающиеся таковыми — *закон Вернера* и *закон Хольцманна*. Здравый смысл подсказывает, что звуковой закон — это некоторый переход, звуковой процесс. Но опыт компаративистики учит, что здравого смысла и языковых данных не всегда достаточно, чтобы реконструировать механизм перехода в реконструированном языке. Парадокс состоит в том, что звуковое соответствие между языками а и б (в частном случае — между разными стадиями одного и того же языка) иногда устанавливается точно, при том что полные описания фонетики и просодии сопоставляемых языков недоступны. Разумеется, это относится к соответствиям в любой области, но особенно досадно, что успешная сравнительно-историческая реконструкция соответствий в области фонетики может не сопровождаться удовлетворительным объяснением фонетических процессов, эти соответствия породивших. Именно так обстоит дело со знаменитым законом, носящим имя Карла Вернера (1846–1896). Вернер в 1876 г. установил, что звонкие германские спиранты вместо ожидаемых глухих *f*, *þ*, *x*, *x^w*, *s* встречаются в медиальной и финальной позиции в индоевропейских словах с неначальным ударением: ударный слог восстанавливается с опорой на те индоевропейские языки, которые сохранили разноместное ударение. Это замечательное открытие было обставлено целым рядом произвольных соображений Вернера и его последователей о фонетической природе предполагаемого перехода. Не будучи частью компаративистской процедуры, эти соображения воспроизводятся в справочниках в силу того, что индоевропеистика за полтора столетия не предложила им равноценной замены. Вернеру нужно было решить, какой именно процесс отражают древние германские языки — *озвончение* спирантов перед ударным гласным или *ослабление* конца слова. Он выбрал второй ответ, потому что заранее принял два сделанных до него априорных допущения о природе ударения в языке а (= «индоевропейский язык») и продолжающем его языке б (= «прагерманский язык»): ударение в языке а якобы было музыкальным и разноместным, а в языке б якобы стало динамическим и сдвинулось на первый слог в словах с неначальным ударением. Чтобы объяснить, как музыкальное ударение языка а превратилось в динамическое ударение языка б, Вернер принял третье допущение: это якобы было следствием сдвига (ретракции), аналогичного вокальному приему *portamento*, ударения на первый слог. В учебниках параллели между раннегерманским ударением и певческой техникой признаются нестрогими. Зато другие, куда более обязывающие, допущения о том, что звонкость германских спирантов якобы была следствием якобы имевшей место в общегерманском языке ретракции ударения, выдаются

за установленные факты. Между тем имеется столько же оснований считать, что закон Вернера отражает не потерю разноместного ударения, а наоборот, состояние, когда оно еще не было утрачено. Я сознаю, что выдвинутое мной в [5] объяснение является одной из многих гипотез в области, где нет консенсуса, — реконструкции фонетического механизма просодического сдвига в германском праязыке. Но полагаю, что оно проясняет больше фактов, чем гипотезы, опирающиеся на догмы индоевропеистики.

Статья [6] «Кому *проспати заутреня* и *ловити рыба*? Именительный падеж дополнения инфинитива как предмет грамматического описания» (2019) — о синтаксической особенности древнерусского языка, указанной в заглавии. Здесь, как и в случае со скандинавскими акцентами в [4], конкурируют два объяснения, каждое из которых поддержано именами видных ученых. Либо перед нами унаследованная черта индоевропейского синтаксиса, архаические свойства формы инфинитива, всплывшие ок. 1000–1100 гг. в группе диалектов русского языка. Либо грамматическое заимствование из смежных или субстратных финно-угорских языков, распространенных на западе и северо-западе Руси. Синхронные данные об употреблении конструкции им. п. + инфинитив в текстах — а их за 700 с лишним лет накопилось немало — свидетельствуют в пользу второй гипотезы, хотя окончательно ее не доказывают. Но один камень в огород исторической русистики проведенный анализ позволяет бросить. Историки языка на протяжении более ста лет были озабочены реконструкцией «идеального» состояния грамматической конструкции, по неясным причинам считая распределение оборотов им. п. + инфинитив vs вин. п. + инфинитив в реальных текстах древнерусского и старорусского периода хаотическим. Между тем, гипотетическое «идеальное» состояние конструкции, когда любое дополнение инфинитива переходного глагола оформляется им. п., по-видимому, не реконструируется вообще: если что и было заимствовано древнерусским языком в результате ареальных контактов, то это — закономерное распределение разных падежей (на русской почве — им.п. vs вин. п.) в зависимости от коммуникативного статуса дополнения, его одушевленности и типа инфинитивной клаузы. Такое распределение в синхронной типологии описывается моделью *дифференцированного маркирования аргументов* [Aissen 2003], а представление лингвиста о хаосе в распределении падежных форм и стадийальных сдвигах в синтаксисе инфинитива возникает из-за того, что данная модель русистикой пока что освоена слабо.

Статья [7] «Два диалекта русской грамматики: корпусные данные и модель» (2018 г.) завершает второй раздел книги. Тут диахроническое измерение связано не с внутренней реконструкцией, а с ретроспективой и параметризацией грамматики. Слово «диалект» в данной работе используется не в значении «исторически сложившаяся территориальная форма языка», а в значении «идиом, идиолект, соответствующий заданному набору ограничительных условий». Анализ текстов основного подкорпуса Национального корпуса русского языка (НКРЯ) подвел к нетривиальному выводу о том, что разные группы носителей русского языка употребляют такую падежную форму, как «тв. предикативный падеж» в соответствии с разными типами правил. При этом с XVIII по XXI в. в НКРЯ устойчиво

прослеживается «диалект» русской грамматики, где допустимы предложения типа рус. *“Нам было известным, что Вы придете*, в которых тв. предикативный падеж приписывается прилагательному при отсутствии стандартного подлежащего в им. п. Это явный случай внутриязыкового варьирования. Все же от тезиса о единстве русской грамматики отказываться рано: условия приписывания тв. предикативного падежа в разных «диалектах» русского языка можно описать в терминах разных настроек одного и того же параметра, связывающего предикативный падеж с синтаксическим контролем со стороны подлежащих разного типа.

III. Семантика в зеркале грамматики

Описание межъязыкового и внутриязыкового варьирования эффективно, когда есть *tertium comparationis* — семантика. В семи работах раздела III «Семантика в зеркале грамматики» обсуждается *грамматическая семантика*, отвлеченная от употребления конкретных лексических единиц, но связанная с употреблением конструкций, классов слов и тематических групп лексики.

Правомерно потребовать, чтобы определения семантических ролей и базовых типов предикатов — действий, свойств, состояний, результатов, отношений и т. п. были универсальны, а не подтаскивались под явления конкретного языка. Правда, на представления лингвиста всегда влияет тот языковой материал, который он лучше знает. В статье [8] «История одной полемики» (1998 г.) сравнивается класс слов «категории состояния» в двух языках — русском и древнеисландском. Я преклоняюсь перед смелостью Л. В. Щербы, который в далеком 1928 г. предположил, что предикатный тип «состояний» может выражаться специальным классом неглагольных слов, близких к особой части речи, и горжусь тем, что моя статья [8] вышла в первом номере журнала со знаковым названием «Язык и речевая деятельность». Но мне кажется, что доказательность рассуждений Л. В. Щербы и его соратников В. В. Виноградова и А. В. Исаченко существенно снижается из-за того, что они недискретно объединили три разных группы проблем: 1) диагностику выделения предикативов «категории состояния» и их частеречный статус в языках типа русского; 2) семантическое описание слов «категории состояния» и уточнение состава предикативов; 3) наличие типологических параллелей в плане становления класса слов «категории состояния». На последний вопрос В. В. Виноградов и А. В. Исаченко, поддавшись гипнозу схемы, по которой «категория состояния» формируется в русском языке лишь в последние века, с ходу ответили отрицательно, объявив ситуацию в современном русском языке уникальной. Это не позволило им найти параллели в родственных славянских языках и других языках мира. В паре древнеисландский vs русский язык семантика предикативов оказалась сходной, а различие между классами предикативов в этих двух языках лучше всего описывается с помощью классификации производящих основ: в древнеисландском языке намного больше основ адъективной и причастной морфологии, от которых образуются только несогласуемые предикативные формы, а не атрибутивные прилагательные и причастия.

В небольшой статье [9] «Субъект состояния и субъект оценки» (1999 г.) уточняется семантический аппарат описания слов состояния по отношению к другим типам предикатов. Я утверждаю, что специфика современного русского языка в плане реализации стативных значений состоит не только и не столько в противопоставлении неглагольных предикатов состояния глагольным предикатам действия³, сколько в выделенном положении предикатов *внутреннего состояния*, связанных с одушевленным субъектом⁴. Противопоставление значений типа рус. *мне весело, душно* значениям типа рус. *здесь пыльно, ветрено* столь же релевантно, как и щербовское противопоставление *мне весело* vs *я веселюсь* vs *я веселый* [Щерба 2008: 90]. Эта ситуация интерпретируется с опорой на параметры описания, которые я называю *субъектом состояния* и *субъектом оценки*. Предикатный признак <здесь> *пыльно, солнечно* принимает одинаковое значение для любого наблюдателя / оценивающего субъекта, поэтому в стандартном русском языке нельзя сказать **мне пыльно, солнечно*. При предикатах с валентностью на дат. п. лица, ср. *мне весело, скучно, холодно, душно* предикатный признак принимает соответствующее значение для конкретного одушевленного субъекта, и только для него. Это и есть внутреннее состояние: одному человеку в данной ситуации может быть *холодно* или *скучно*, а другой этих состояний не испытывает. При предикатах типа <это> *аморально, алогично* тоже есть встроенная оценка, претендующая на общезначимость: предполагается, что любой наблюдатель оценит данное явление как *аморальное* или *алогичное*. Но нет информации о переживаемом состоянии.

Парадоксальным образом рассуждения о классе слов состояния как о «визитной карточке» русского языка на протяжении 90 лет не подкреплялись лексикографическим описанием входящих в него единиц. Многие охотно рассуждали о том, что количество слов «категории состояния» возрастает, но никто не указал ни примерного объема этого класса, ни перечня его элементов. Возможно, первая попытка была сделана в 2017 г. в моей статье [10], озаглавленной «Русские предикативы в зеркале эксперимента и корпусной грамматики». Проблему представляет не подсчет элементов типа *жаль, надо*, которые употребляются только в функции сказуемого (таких слов в русском языке, в отличие, например, от древнеисландского, относительно мало), а исчисление производных форм с финалью *-о*, которые способны формировать дативно-предикативную структуру (*мне холодно / душно / грустно* и т. п.), сохраняя при этом согласуемые корреляты (*холодный, душный, грустный X*). Я принимаю семантический критерий и предполагаю, что валентность на дат. п. лица диагностирует класс внутренних состояний в русском языке. Число предикативов, имеющих такую валентность, определяется методом

³ Этот тезис в начале 1980-х гг. убедительно развенчала Т. В. Булыгина, с чьей работой [Булыгина 1982] должен считаться каждый, кто пишет на данную тему. Суть в том, что далеко не каждый глагол является предикатом действия, что ошибочно утверждают Л. В. Щерба и те, кто буквально воспроизводят его тезис: разброс таксономических предикатных значений для глагольных предикатов ничуть не меньше, чем для неглагольных.

⁴ Термин «внутренние состояния» ранее использовался в книге Анны А. Зализняк [Зализняк 1992], но на несколько ином материале.

двойного — социолингвистического и корпусного — эксперимента. Результаты показывают, что средний объем класса предикативов с валентностью на дат. п. лица в активном словаре опрошенных информантов превышает две сотни элементов. Однако лишь часть этих слов воспроизводится носителями русского языка как готовые лексемы со встроенной валентностью на дат. п. лица. Другая часть достраивается носителями по грамматической аналогии, при этом разные носители русского языка используют в дативно-предикативной структуре разные предикативы. По методике, апробированной в [10], позже были описаны и предикативы в других языках — болгарском, македонском, сербском, словенском и украинском [Митковска 2018; Марич, Керкез 2018; Ухлик 2018; Кулинич 2018; Градинарова 2018; Ivanova, Zimmerling 2019].

В статье [11] «Имперсональные конструкции и дативно-предикативные структуры в русском языке» (2018 г.) делается обобщение о природе семантико-грамматического интерфейса в русском языке. Предлагается считать, что признак ‘± одушевленность семантического субъекта’ играет решающую роль в распределении трех конструкций — дативно-предикативных структур и двух глагольных конструкций, обычно в русской грамматической традиции называемых неопределенно-личными (рус. *улицу засыпали песком*) и переходными безличными (рус. *улицу засыпало песком*). На типологической шкале русский язык занимает промежуточное место между языками, где употребление сопоставимых единиц генерализовано и мало зависит от тематической группы предикатной лексики, и языками, где употребление имперсональных глаголов и предикативов с дат. п. лица полностью лексикализировано. В течение своей письменной истории русский язык претерпел значительную эволюцию, результатом которой является установившееся сложное соотношение между предикатной лексикой и моделями синтаксиса.

В написанной годом ранее статье [12] «Эпистемические предикаты и ментальные состояния» (опубликована в 2019 г.) тема более специальная. Прослеживаются контексты, где предикаты БЫТЬ и КАЗАТЬСЯ противопоставлены, и контексты, где они сближаются. Контексты последнего типа прослеживаются на материале русского, болгарского и древнеисландского языков. В этих языках реально засвидетельствованы как ситуации, где семантический оператор КАЗАТЬСЯ может сигнализироваться дативно-предикативной структурой с глаголом «быть», так и ситуации, где употребление предикатива при глаголе «казаться» не связано с контрастом БЫТЬ vs КАЗАТЬСЯ. Название статьи не является загадкой. В названных языках предложения с составом элементов типа болг. *Кораво ми е*, букв. «жестко мне» вне контекста неоднозначны и могут выражать как внутреннее состояние «Х-у жестко», так и эпистемическую оценку «Х-у кажется, что здесь жестко».

Статья [13] «Обладать и быть рядом» (1999 г.) — о неполном соответствии семантики и синтаксиса. Некоторые ярлыки, например «посессивность», вроде бы указывают на бинарное семантическое отношение вида «Х имеет Y / у X-а есть Y», но они одновременно применяются и для обозначения синтаксических конструкций, которые либо используются для выражения более широкого круга предикатных значений, либо для выражений отношений, которые вообще не связаны

с выделением роли посессора. Проблема не только в том, что лингвистические термины неоднозначны, но и в том, что подход, признающий грамматику изоморфным отражением семантики, серьезно искажает описание языковых фактов.

Последняя (хронологически — первая) статья раздела — [14], «Логика прогноза» (1997 г.), — является попыткой сопоставить разным логическим моделям будущего разные типы футуральных конструкций и разные нарративные стратегии. Логическая семантика разного типа, возможно, мотивирована разными моделями грамматикализации будущего времени. И наоборот: разные пути моделирования будущего времени как звена видовременной системы отражают фундаментально различные представления о ситуации прогноза.

IV. Подлежащие и неканонические подлежащие

Деление на два — одна из самых простых математических операций. С лингвистикой она связана двойко. Во-первых, естественный язык в некотором смысле является исчислением. Во-вторых, многие базовые понятия лингвистического описания, например, представление о «субъектном аргументе» или «подлежащем», опираются на идею бинарного членения. Бинарное представление фонологического или грамматического признака, например '± звонкость' или '± одушевленность' — это, прежде всего, способ описания. Бывают ситуации, когда бинарное vs небинарное представление явлений, противопоставленных по глухости / звонкости, не одинаково удачны. Но обычно никто не настаивает на том, что альтернативное бинарному описанию в терминах значений признака невозможно. К тому же небинарное представление, например, стремление учесть не 2, а 4 степени фонологического признака 'подъем языка' или 4 значения (граммемы) некоторой грамматической категории, легко приводится к бинарному за несколько шагов: для начала, например, можно ввести противопоставление 'верхний vs неверхний подъем' и далее выделить все нужные значения признаков. Для одних лингвистов описание в терминах бинарных признаков — добродетель, для других — красная тряпка и часть отвергаемой ими традиции, но большой проблемы здесь не видно.

Описание в терминах «подлежащего» и бинарного членения синтаксической структуры на «подлежащее» vs «сказуемое», «вершину» (head) vs «зависимый элемент» (complement) претендует на большее: предполагается, что такова *действительная природа* языковых объектов и что выделенный статус подлежащих, вершинных и зависимых элементов подтверждается набором языковых средств. Выдающиеся умы (Б. Дельбрюк) еще в конце XIX в. усомнились в том, что структура предложения всегда бинарна, и допустили, что идея об обязательности подлежащего как элемента, противопоставленного второй части предложения (сказуемому), является не только априорным постулатом, но и экстраполяцией ситуации в таких языках, как современные немецкий и английский, где описание в терминах подлежащего удобно и просто, на языки, где природа синтаксического отношения точно неизвестна [Delbrück 1893]. Эта точка зрения скептиков конца XIX в. напоминает критику метаязыковых понятий в недавней работе [Haspelmath 2010],

где модельное понятие подлежащего (syntactic subject) исключено из списка так называемых *сравнительных категорий* (comparative concepts), на основе которых разрешается проводить сопоставление языков. Но к доводам скептиков тоже надо относиться со скепсисом. Шансы полностью исключить из типологии любые априорные модельные понятия — отдельная тема, которая прямо не связана с опытом применения понятия «подлежащее» в описаниях языков мира. Некоторый диалог между сторонниками и противниками идеи (не)универсальности синтаксических отношений ведется с 1970-х гг., им трудно полностью игнорировать друг друга. Распространена точка зрения о том, что наряду с классом «подлежащих» языков, с бинарным членением структуры предложения на подлежащее и сказуемое, где каждый из компонентов может быть представлен в виде синтаксической группы ($S \rightarrow NP, VP$), есть какие-то другие языки, называемые «языками с выдвиганием топики» [Li, Thompson 1976], «неконфигурационными языками» [Hale 1983], где бинарное членение предложения либо отсутствует, либо связано с выделением не подлежащих, а каких-то гипотетических иных сущностей, например «топики». Обычно считается, что принадлежность языка мира к классу неподлежащих должна специально доказываться, а принадлежность языка мира к классу подлежащих — нет, ср., однако, статью [Kibrik 1997], где утверждается прямо обратное. Господствующая установка подтверждается тем, что в описательные грамматики иногда добавляют раздел о том, что данный австралийский или австронезийский язык является «неконфигурационным», см. [Austin 2001], в то время как разделов, специально доказывающих, что некоторый язык является подлежащим, в описательных грамматиках обычно нет: авторы описательных грамматик чаще ограничиваются указанием на то, что подлежащее в языке L стоит в таких-то позициях и имеет такие-то признаки.

Я встану на позицию агностика, еще раз воспользуюсь идеей членения на два и поделю языки мира на два класса: класс I, куда входят языки, для которых гипотеза о наличии подлежащего по крайней мере столь же удобна, как и альтернативная гипотеза об его отсутствии, и класс II, о котором нет достаточной информации. Далее я предположу, что русский язык относится к классу I, а обратный вывод, о том, что категории подлежащего в русском языке нет, сделанный в замечательной работе [Козинский 1983], основан на неверной процедуре универсального определения подлежащего по множеству разнородных и разноуровневых свойств, ср. критику данного подхода в [Тестелец 2001]. Для выделения стандартного подлежащего в русском языке в общем случае достаточно двух критериев — предикатного согласования и наличия им. п. Стоит уточнить, что «стандартное» или «каноническое» подлежащее — это именная группа (которую можно *прономинализовать*, заменить на местоимение) или местоимение, вместо которого можно подставить именную группу. Именные группы в русском языке, как считалось до недавнего времени, имеют падеж, а так называемые беспадежные подлежащие группы, типа рус. *пять человек, два человека* [Лютикова 2017], все-таки являются не каноническими именными группами, а выражениями иного типа — *количественными группами* [Franks 1995]. В русской грамматической традиции

принято считать, что сентенциальные актанты — инфинитив, точнее — инфинитивная группа, ср. рус. [*Кататься на санках*] — *весело* [Золотова 1982; Шведова 1982б], а также придаточные, ср. рус. [*Что волки жадны*], *всем известно*, [*Что он постоянно опаздывает*], *всех раздражало*, тоже являются грамматическими подлежащими [Бондарко 1992]. Ясно, однако, что падежа у инфинитива в современном русском языке (в отличие от праиндоевропейского языка, где инфинитив был отглагольным именем в дат. п. [Попов 1881]) нет, равно как и у придаточных предложений. Сентенциальные актанты не контролируют, по крайней мере в русском языке, согласования сказуемого: оно стоит в 3 л. ед. ч., а в высказываниях в прошедшем времени — в 3 л. ед. ч. ср. р., точно так же, как в безличных предложениях типа рус. *света-л-о*, *его стошни-л-о*, обычно признаваемых бесподлежащими [Babby 2002]. У количественных групп типа *два человека*, *пять человек* другая странная особенность в русском языке — сказуемое при них может согласоваться как во мн. ч., так и в 3 л. ед. ч. ср. р., т. е. принимать форму, относительно которой а priori неясно, является она согласуемой или нет⁵. Таким образом, если сентенциальные актанты и количественные группы действительно являются подлежащими в русском языке, есть основания признать их подлежащими нестандартными, «неканоническими».

Есть и другие кандидаты на роль неканонических подлежащих. На материале разных языков обсуждалась возможность подлежащих в косвенных падежах (*oblique subjects*) [Moore, Perlmutter 2000], так называемых формальных или экзплетивных подлежащих типа англ. *it*, *there* [Partee 1979] и даже предполагаемых нулевых подлежащих [Meřčuk 1995; 2014]. Все эти постулированные сущности являются модельными понятиями, а не плодом дотеоретических наблюдений над языками. Их объединяет то, что, не удовлетворяя диагностическим для языка L критериям подлежащности (например, контролю предикатного согласования и наличию им. п. в русском языке), они не являются частью сказуемого, но противопоставлены ему примерно так же, как канонические подлежащие⁶.

Доводилось заниматься неканоническими подлежащими на материале нескольких языков и мне. Для этой книги я отобрал пять статей, в которых эта проблематика рассмотрена на материале современного русского языка. Как мне представляется, главная теоретическая и методологическая проблема в этой области — объяснить, как одна и та же система грамматических отношений может регулярно порождать и канонические, и неканонические предикатные актанты⁷. Соответственно,

⁵ В других славянских языках, в частности в польском, правила иные.

⁶ В другой терминологии, неканонические и канонические подлежащие подводятся под общую категорию «приоритетных актантов»: подлежащее является приоритетным актантом, но не всякий приоритетный актант является подлежащим [Тестелец 2001: 319–320].

⁷ Альтернатива только одна — объявить неканонические подлежащие реликтами какой-то иной системы грамматических отношений, и она непривлекательна. Даже если в некотором языке разные типы предикатных актантов и разные модели предложения прослеживаются до разной глубины, это не избавляет от необходимости разбирать соотношение этих актантов и моделей предложения в синхронии.

ценным будет то описание, которое даст новые данные об этой системе грамматических отношений. К примеру, вывод Б. Парти [Partee 1979] о том, что в предложениях типа англ. *There are three cats on the roofs (aren't there?)* и *there is so much time left* грамматическим подлежащим является формальное слово *there*, а не группа, стоящая после глагола (*three cats, much time*), важен не столько для повышения значимости формальных слов в сознании лиц, изучающих английский язык, и даже не столько для подтверждения общей идеи о том, что подлежащее в языках мира может быть выражено формальным словом, сколько для уточнения критериев подлежащности в самом английском языке. Как показала Б. Парти, критерий сохранения позиции в производной структуре, ср. [*there being so much time left*] bothers me, в ряде случаев — именно в английском языке! — важнее критерия контроля согласования сказуемого [ibid.]. В этой иллюстрации теоретическая грамматика одного отдельно взятого языка (английского) вносит больший вклад в систему научных знаний, чем практическая грамматика данного языка и чем претензия на обобщение уровня общей лингвистики (какие подлежащие вообще бывают и каких не бывает), не подкрепленная скрупулезным анализом материала. Разумеется, кому поп, а кому попадья, кому водка, а кому молодка: трудно настаивать на том, что практическое освоение английского или русского языков менее важно для человечества, чем анализ нюансов их теоретической грамматики.

Мой вклад в описание неканонических подлежащих в русском языке состоит в анализе системных отношений между ними и в разработке концепции *сильных* и *слабых* подлежащих, опирающейся на ряд тестов. В статье [15] «Dative Subjects and Semi-Expletive Pronouns in Russian» (2007–2008 гг., опубликована в 2009 г.) я доказываю, что эксплетивное безударное местоимение *это* имеет свойства формального подлежащего в предложениях с предикативами типа *приятно, известно (Это приятно / известно)*. Но те же предикативы допускают и два других типа неканонических актантов — сентенциальные группы ([*Кататься на санках*] приятно, Известно, [*что Вася уехал из города*] и т. п.) и группы со зн. дат. п. лица, указывающие на семантический субъект предикатива (*Мне было приятно*). При этом предикативы, имеющие валентность на дат. п. лица, допускают одновременную реализацию двух неканонических актантов, но не трех из них, ср. *это приятно, что Вася принес извинения / мне это было приятно*, при неграмматичном **мне это было приятно, что Р*. Стандартная для выдвижения гипотезы о подлежащем презумпция состоит в том, что в одной синтаксической области может быть только одно подлежащее. Если приведенные выше конструкции моноклаузальны, данная дистрибуция свидетельствует о том, что при предикативах, сочетающихся с дат. п. лица, в роли неканонического подлежащего могут выступать разные элементы, один из которых выбирается в соответствии с некоторой иерархией. Актанты такого рода, конкурирующие между собой при предикатах, не сочетающихся в русском языке с каноническими подлежащими в им. п., в статьях [15], [16] и [19] названы *слабыми*. *Сильные* подлежащие являются единственно возможными кандидатами на роль приоритетного актанта при данной форме и семантике предиката. Таковы, в частности, именные аргументы независимых дативно-инфинитивных

предложений, ср. рус. *Грузовикам здесь не проехать. Так тому и быть. Саду — цвесть. Всем — молчать!* В современном русском языке именной аргумент таких предложений всегда маркируется дат. п.

В статье [16] «Неканонические подлежащие в русском языке» (написана в 2011 г., опубликована в 2012 г.) концепция сильных и слабых подлежащих развернута подробнее. Я утверждаю, что внешне невыраженные (нулевые) подлежащие с ролевой семантикой Агенса и неопределенным референтным статусом, постулированные И. А. Мельчуком для предложения типа рус. \emptyset^{3PL} *улицу засыпа-л-и песком* и рус. \emptyset^{3SG} *улицу засыпа-л-о песком* [Meľčuk 1995], можно трактовать как нулевые неопределенные подлежащие местоимения, соответственно, 3 л. мн. ч. и 3 л. ед. ч., дополнительно распределенные по признаку ‘± одушевленность’. Между собой эти виды нулевых подлежащих в современном русском языке (в отличие от древнерусского, что показано в статье [10] из предыдущего раздела) конкурировать не могут. Если принять вывод И. А. Мельчука о том, что предложения с нулевым подлежащим в 3 л. мн. ч., типа \emptyset^{3PL} *улицу засыпа-л-и песком*, всегда несинонимичны предложениям с внешне выраженным показателем неопределенности, ср. *Люди засыпали / Кто-то засыпал улицу песком*, и \emptyset^{3PL} и \emptyset^{3SG} следует признать сильными подлежащими.

В статье [17] «Sentential Arguments and Event Structure» (2014 г.) утверждается, что придаточные и инфинитивные группы в русском языке могут быть только слабыми неканоническими подлежащими, как при предикативах, так и при глагольных предикатах. Продвижение сентенциальных групп в позицию подлежащего обставлено рядом дополнительных ограничений, касающихся наличного состава элементов, модели управления предиката и порядка слов. Поэтому следует считать, что сентенциальные актаны занимают на шкале слабых подлежащих одно из низших мест, вопреки укорененным в русской грамматической традиции представлениям.

Статья [18] «Корпусная грамматика количественных групп в русском языке» (2019 г.), в отличие от других статей данного раздела книги, опирается на репрезентативную выборку корпусных данных: одной из ее целей является анализ соответствия прогнозов формальной грамматики наблюдаемому в корпусе (НКРЯ) языковому варьированию. Однако не все положения формальных моделей языка проверяемы. Эмпирическая картина такова. Описательные грамматики утверждают, что группы типа *два человека, пять человек* в позиции подлежащего допускают двойное согласование сказуемого — и во мн. ч. и в 3 л. ед. ч. ср. р. (*два человека пришли / пришло*), в то время как группы, расширенные согласуемым препозитивным элементом (*эти два человека*) или согласуемым оборотом (*два человека, которые пришли первыми*) допускают только согласование во мн. ч. [Шведова 1982б]. Этот вывод подтверждается корпусными данными. В формальной модели [Лютикова 2017; 2018] считается, что подлинное (т. е. контролируемое подлежащим) согласование количественных групп, а также групп, их содержащих, ср. рус. *эти три женщины, три женщины и пять мужчин*, всегда реализуется в русском языке во мн. ч., в то время как форма сказуемого в 3 л. ед. ч. ср. р. реализует так

называемое дефолтное согласование, т. е. является несогласуемой в перспективе традиционной грамматики. Эту гипотезу без дополнительных соглашений об особенностях синтаксической структуры русского языка и способах ее описания методами корпусной грамматики проверить нельзя. Зато методов корпусной грамматики в силу некоторых особенностей порядка слов в русском языке достаточно для того, чтобы подтвердить статус количественных групп как неканонических актантов. Большинство типов подлежащих ориентировано на употребления с коммуникативным статусом темы, но количественные подлежащие группы в большинстве примеров имеют коммуникативный статус ремы или части рематической составляющей. Коммуникативные статусы в русском языке, в первую очередь, маркируются акцентно, и лишь во вторую очередь — средствами порядка слов [Ковтунова 1976; Янко 2001; 2008]. Но при порядке VS русский язык исключает возможность сделать подлежащее акцентно маркированной темой, поэтому выгрузки письменных текстов для решения поставленной задачи достаточно.

Статья [19] «Топикальность и неканонические подлежащие в русском языке» (2019 г.) дает общий обзор инвентаря доказанных и вероятных неканонических подлежащих русского языка, с опорой на такие параметры, как топикальность / нетопикальность, семантическая роль, различие сильных / слабых подлежащих. Поборники методологической строгости, которые считают, что *мухи отдельно, а котлеты — отдельно*, возразят, что топикальность, «наличие коммуникативного статуса темы» — не формально-грамматическое свойство, в отличие от «подлежащности». В синтаксическом минимализме Н. Хомского, где мерой всех вещей является линейная позиция (точнее — соответствие между порядком слов и деревом подчинения), будет и другое возражение, которое мне доводилось слышать: «подлежащее» и «топик» — не только категории разного типа, но и разные *позиции* в модели так называемой генерализованной фразовой структуры. Если некое выражение вынесено за пределы клаузы и стоит «в позиции топика», оно не может одновременно быть подлежащим. И наоборот, если некое выражение попадает в «позицию подлежащего», оно не может одновременно быть топиком. Это логично, но логикой не следует злоупотреблять (Э. Ренан). Критерии, которые хороши для канонических подлежащих, могут не работать для неканонических. Конкретно в ситуации современного русского языка есть основания считать, что сентенциальные актанты могут попадать в позицию подлежащего и выделяться из состава сказуемого лишь при получении коммуникативного статуса темы и выноса в линейную позицию, предшествующую сказуемому, целиком. В иных случаях, например, при разрыве инфинитивного оборота, могут нарушаться формально-синтаксические условия для подлежащих структур (запрет на извлечения синтаксических элементов из состава подлежащего).

V. Порядок слов. Теория и типология

Речь и ее письменная фиксация линейны. Линеаризация языковых структур — фундаментально важный аспект их существования, как и то, что порядок слов и разрешенные комбинации элементов предложения *предсказуемы на основе*

правил: нет языков, где синтаксическая структура во всех предложениях разворачивается единственным способом, равно как нет и языков, где распределение линейных порядков полностью хаотично и абсолютно все комбинации элементов разрешены во всех предложениях. 'Предшествование' и 'непосредственное предшествование' — математические, а не собственно лингвистические понятия. Вероятно, поэтому отдельные лингвисты считают изучение линейного порядка элементов предложения, или попросту «порядка слов», областью, где достаточно здравого смысла, а специальная подготовка не нужна. При этом недооценивается многообразие правил, определяющих разрешенные порядки слов в языках мира.

Выражения «языки с фиксированным порядком слов» и «языки со свободным порядком слов» в свете постулата о предсказуемости порядка слов — условные ярлыки. Их лучше понимать не как разные полюса континуума «свободы», а как отсылки к двум классам языков. В одном из них есть полностью формализованные *запреты* на некоторые порядки слов хотя бы в части предложений. Эти запреты действуют независимо от коммуникативной структуры и фонетической реализации элементов. В другом классе таких запретов нет. Это не единственная перспектива, в которой стоит рассматривать явления порядка слов. Принципиально важно также различие двух типов ограничительных условий, накладываемых на разворачивание предложения. Правила первого типа, названные в статье [20] «Локальные и глобальные правила в синтаксисе» (2008 г.) *локальными*, задают относительное расположение элементов в любых парах типа (x, y) : предполагается, что в любой паре произвольного типа можно найти приоритетный (*базовый*) порядок, который может быть либо не быть единственным: *рестриктивность* системы порядка слов, т. е. наличие прямых запретов на порядок (y, x) при базовом порядке (x, y) , при данном подходе не имеет особого значения, вопреки возражениям, высказанным в [Mithun 1992]. Такова типология порядков слов Дж. Гринберга с ее знаменитыми «формулами предложения» типа SVO (подлежащее, глагол, дополнение), SOV (подлежащее, дополнение, глагол), VSO (глагол, подлежащее, дополнение) и т. д. [Greenberg 1963; Dryer 2013], которая целиком построена на локальных правилах. Формула типа SVO является обобщением *порядка ветвления* во всех бинарных группах вида XY, где один из элементов (X^0) признается вершиной, а второй — зависимым, параметр ветвления задает место зависимого по отношению к вершине. В группе V^0O вершиной признается глагол, а зависимым — дополнение, в группе S^0V вершиной признается существительное, а глагол — зависимым. Формула SVO является обобщением этих случаев и при заданных допущениях отражает последовательное ветвление категорий предложения слева направо. Этот подход продуктивен, но не может быть единственным. В нотацию Гринберга можно добавлять новые «буквы» (символы категорий предложения), но ограничения, накладываемые моделью, не исчезнут. Столь же важны для естественного языка правила второго типа, которые могут быть названы *глобальными*, они обеспечивают упорядочение синтаксической области без непосредственного обращения к локальным правилам относительного расположения категорий предложения. Глобальные правила работают так. Выдвигается гипотеза о том, что в синтаксической области

$D [D \dots x \dots y \dots]^F$ расположение элементов x и y зависит от реализации некоторого признака F , который может принимать разные значения и при этом ассоциируется со всей областью D , а не с каким-то ее отдельным сегментом. Мне известны два основных случая, где применение глобальных правил не просто возможно, но и необходимо для моделирования порядка слов. В статье [20] на материале современного русского языка разбирается один из них, связанный с реализацией коммуникативной структуры и ее отображением на синтаксическую структуру. Приписывание элементам предложения коммуникативного статуса является следствием бинарного членения большинства высказываний, соответствующих речевому акту сообщения на два компонента, называемые «темой» и «ремой». В языке могут существовать и высказывания, нечленимые на тему и рему, так называемые «нерасчлененные» предложения (*thetic sentences*), но они определенным образом, в том числе деривационно, связаны с расчлененными [Sasse 1987]. Членение на тему и рему имеет смысл признать коммуникативно-семантической универсалией и с ее помощью интерпретировать формальную структуру предложения. В русском языке коммуникативный статус элементов маркируется двояко: при помощи фразовых акцентов и средствами порядка слов. В сохраняющей свое значение и поныне работе [Ковтунова 1976] было показано, что все многообразие комбинаций линейных порядков с наборами фразовых акцентов можно задать при помощи правил, *одновременно* меняющих порядок слов и интонацию, — «линейно-интонационных», или «линейно-акцентных», преобразований (ЛА-преобразований). Это и есть глобальные правила в указанном выше смысле: меняя расположение подлежащего и дополнения, глагола и подлежащего и т. п. в соответствии с заданным коммуникативным значением, они обычно меняют и расположение других элементов и/или их акцентную маркировку. Поэтому удовлетворительно объяснить механизмы коммуникативного выделения и порядок слов в языках типа русского на основе локальных правил, предположив, что при изменении коммуникативного статуса выделенный элемент как бы «подсвечивается», а прочая часть высказывания остается незатронутой, обычно не удается. ЛА-преобразования соединяют между собой пары или n -ки предложений с одинаковой лексико-синтаксической структурой, т. е. деревом предложения, узлы которого уже заполнены конкретным лексическим материалом. Эта структура служит инвариантом множества коммуникативных преобразований, тем самым теория ЛА-преобразований — это трансформационная грамматика, являющаяся дополнением к алгоритму, строящему дерево предложения (в разных теориях синтаксиса он может быть разным). Версии теории ЛА-преобразований развивались в [Падучева 1985; Янко 2001]. В моей статье [20] указаны пути формализации ЛА-грамматики, которая позволяет построить исчисление русских предложений с уже расставленными фразовыми акцентами. Такая грамматика относится к классу формальных грамматик с мягко-контекстно-зависимыми правилами (*mildly context-sensitive rules*) и оператором перемещения (*movement*). Я показываю также, как глобальные правила ЛА-преобразований взаимодействуют в русском языке с локальными правилами выбора места фразового акцента в синтаксической составляющей. Этот результат был получен до меня

в [Янко 2001: 69–84], я лишь демонстрирую, как его можно встроить в предложенный мной формализм.

В работе [21], озаглавленной «Системы порядка слов в славянских языках» (2012 г.), и последующих работах раздела подробнее обсуждается другой тип глобальных правил. Лингвисты давно обратили на него внимание в связи с тем, что некоторые языки мира грамматикализовали топологические шаблоны развертывания предложения. В этих языках *синтаксическое расстояние* от левой или правой границы предложения (предикатной клаузы) — более важный фактор, чем локальные правила относительного расположения большинства элементов предложения. Диагностическим в таких системах порядка слов обычно является позиция *второго от начала* элемента, либо позиция *предконечного* элемента, которая закреплена за категорией заданного типа (например, финитным глаголом или операторным словом), в то время как смежная с ней начальная resp. конечная позиция, наоборот, не закреплена за какой-то одной категорией и может быть интерпретирована как позиция произвольной категории (символическая запись ‘XP’). «Формула предложения» в языке такого типа, определяемая на основе глобального правила, кардинально отлична от формулы предложения в типологии Дж. Гринберга. Так, немецкий язык, где финитная и нефинитная формы глагола в главных повествовательных клаузах соответствует топологическому шаблону $\#XP — V_{\text{fin}} [_{\text{MIDDLE FIELD}} \dots] V_{\text{inf}}$, где финитный глагол является диагностической категорией, стоящей на втором месте от начала клаузы, а позиция XP может быть заполнена почти любой другой категорией. Между финитным и нефинитным глаголом в клаузах такого типа, в сегменте, условно называемом «средним полем» (нем. *Mittelfeld*, англ. *middle field*), могут стоять любые категории, которые не попадают в XP, например, подлежащее, дополнения, обстоятельства: фактически это зона свободного порядка слов, внутреннее упорядочение элементов «среднего поля» определяется глобальными правилами, сходными с теми, которые действуют в русском языке. В зависимых клаузах, в том числе в придаточных с подчинительным союзом (узел *Comp*), топологический шаблон имеет уже другой вид — весь глагольный комплекс размещается контактно на правом краю клаузы, причем финитный глагол стоит последним: $\#Comp \dots [V_{\text{inf}} — V_{\text{fin}}]\#$. В терминах типологии Дж. Гринберга немецкий язык интерпретируется как система порядка слов с непоследовательным ветвлением $V^0O \sim OV^0$. Такое описание вполне возможно, а для решения определенных задач и нужно. Но оно не ухватывает особой роли, которую в немецкой системе порядке слов играет позиция финитного глагола по отношению к началу клаузы ($\#XP — V_{\text{fin}}$). Тем не менее, если ограничиться бинарной группой (V, O), такое описание немецкого языка полезно. Зато если добавить еще одну «букву» в попытке обозначить в формуле предложения место подлежащего, хорошего описания в терминах Дж. Гринберга не получится: отождествление предфинитной позиции XP с позицией подлежащего серьезно искажает факты, а в «среднем поле», т. е. в сегменте между финитным и нефинитным глаголом в главной повествовательной клаузе, отдельной линейной позиции у подлежащего в немецком языке нет. Русский язык, в принципе, можно интерпретировать как язык с базовым порядком

SVO, но это требует сложного формализма, включающего глобальные правила. Лингвист может задать их либо в виде гипотезы о наличии так называемой генерализованной фразовой структуры, т. е. универсального топологического шаблона категорий предложения, который соблюдается всеми языками, включая русский, либо (если первое допущение отвергается) в виде модельных постулатов о соотношении коммуникативной и синтаксической структур: последний путь обосновывает работа [20].

В славянских языках, материал которых обсуждается в статье [21] на фоне типологически близких систем порядка слов, ровно таких правил линеаризации предложения с опорой на место глагола, как в немецком языке, нет. Хотя некоторые малые славянские языки, в частности, верхнелужицкий и нижнелужицкий, испытали долговременное влияние немецкого. Зато славянские языки отражают весьма распространенный в других языках мира случай, где диагностической категорией, стоящей на втором месте от начала клаузы, являются *сентенциальные клитики*, т. е. слабоударные предикатные элементы, относящиеся ко всему предложению: #...CL. Статус сентенциальных клитик в славянских языках могут иметь частицы, слабоударные местоимения и связки. Примечательным свойством славянских сентенциальных клитик является то, что они могут образовывать *цепочки*, т. е. синтаксические объекты, элементы которых при контактном расположении выстраиваются в единственно возможном порядке. Кроме сентенциальных, славянские языки используют *локальные клитики*, т. е. клитики уровня непредикатных групп, например, предложной или именной. Но у локальных клитик свойства образовывать цепочки в славянских языках нет. Все славянские языки, кроме современных русского, украинского и белорусского, входят в класс *языков мира с цепочками клитик*, в этот же класс входит древнерусский язык и некоторые украинские карпатские диалекты. Параметризация славянских систем порядка слов с цепочками клитик в целом соответствует представлениям о разнообразии этого класса за пределами славянского ареала. Конвергентное сходство славянских систем порядка слов с системами порядка слов далеких ареалов значительно. Так, набор ограничительных условий, накладываемых на позицию сентенциальных клитик и позиции глагольных форм в болгарском языке — клитики твердо стоят во второй позиции от начала, глагол при наличии клитик занимает смежные с ними позиции (#XP — CL — V ~ #V — CL; *#CL; *XP — [Y] — CL), поразительно напоминает наблюдаемую в тагальском и других филиппинских языках картину.

Языки с цепочками клитик используют два типа правил, связанных с синтаксисом цепочек. Первая группа правил касается *внешней позиции* цепочек и *кластеризуемых клитик*, т. е. клитик, способных быть частью цепочки. Болгарско-филиппинский сценарий, когда клитики жестко привязаны ко второй позиции от начала клаузы, а глагол занимает смежные позиции либо слева (т. е. в начальной позиции), либо справа (т. е. в третьей позиции от начала), не является ни единственным, ни самым распространенным. В большинстве случаев постановка кластеризуемых клитик / цепочек на второе место не скоррелирована с условием смежности глагола и клитик: глагол в предложении с клитиками сохраняет подвижность. Системы

порядка слов подобного рода известны как стандартные языки с законом Ваккер-нагеля. В романских и балканских языках (новогреческом и албанском), напротив, представлены системы порядка с *приглагольными* клитиками, где требование контактной позиции в группе $V + CL$ жестко выдерживается, но сам комплекс $[V + CL]$ не имеет фиксированной позиции в предложении. В Европе системы подобного типа составляют большинство от общего числа языков с цепочками клитик, но в других ареалах они встречаются значительно реже. Бесспорных примеров систем подобного типа в славянском ареале почти нет. Македонский язык, где комплекс $CL = V$ может начинать предложение и клитики могут стоять как во второй, так и в первой позиции, по ряду критериев не является системой порядка романского типа: точная параллель к нему неожиданно обнаруживается в совсем ином ареале — в языке рикавунг пуума на о. Тайвань, см. детали в [25]. Другие условия внешнего упорядочения клитик касаются инвентаря синтаксических категорий, к которым могут присоединяться клитики, и возможностей вставки клитик внутри первой синтаксической группы, состоящей из нескольких фонетических слов. Наконец, в части языков с законом Ваккер-нагеля, в том числе в древнерусском и нескольких современных славянских языках, действуют особые коммуникативно-мотивированные правила, смещающие кластеризуемые элементы вправо.

Второй тип правил в языках с цепочками клитик касается правил *внутреннего упорядочения* клитик друг по отношению к другу. Правила этого типа обсуждаются в работах [22] и [23]. В статье [22] «A unified analysis of clitic clusters in world's languages» (2012 г.) предложены универсальные критерии, позволяющие отделить цепочки как синтаксические объекты особого типа от последовательностей фонетически смежных клитик. Шаблоны внутреннего упорядочения клитик, так называемые *Правила Рангов*, в западной традиции называются термином «template». Правила Рангов в естественном языке, задающие внутренний порядок клитик в виде таблицы, — точный аналог простых формальных грамматик, называемых «финитными автоматами с магазинной памятью» или «грамматикой рангов»: за один шаг такие грамматики просматривают ровно один участок, затем переходят к следующему за ним и т. д. Если упорядочение клитик в цепочке вида $[CLP \ abc\dots n]$ твердо, имеется однозначное соответствие между множеством клитик вида $\{a_1, a_2\dots a_n\}$ и *рангом* клитик, т. е. номером того столбца таблицы, куда попадают любая клитика $a_k \in \{a\}$. В зависимости от поставленной задачи, Правила Рангов клитик можно понимать либо просто как дескриптивные обобщения над множеством текстов некоторого языка, имеющего клитики, либо как *алгоритмы упорядочения* клитик. Последний аспект особенно важен для анализа эволюции языка или описания варьирования. Если в языке появляются новые клитики, важно установить, имеются ли доступные клетки Правила Рангов, куда эти клитики могут быть добавлены, или же требуется расширить Правило Рангов и добавить новые клетки — с левого, правого края или в ином месте. Главная ценность Правил Рангов как инструмента анализа состоит в том, что они часто позволяют понять синтаксис самих клитик, не имея на входе полной информации о синтаксисе всего предложения. Это — важный ресурс, позволяющий снимать омонимию форм и разрешать неоднозначность

предложений. Кроме того, из Правил Рангов можно пытаться извлечь информацию диахронического плана о последовательности клитизации элементов.

В статье [23] «Clitic templates and discourse marker *ti* in Old Czech» (2014 г.) Правило Рангов использовано как инструмент, верифицирующий наличие дискурсивной частицы *ti* (*t'*) «действительно» в древнечешском языке. Доказательство основано на гипотезе, аналогичной той, которая ранее была доказана А. А. Зализняком на материале древнерусского языка [Зализняк 1993: 298–308]. Если энклитика ti_1 — не тот же самый элемент, что энклитическое местоимение ti_2 ‘тебе’, а его омоним, и оба этих элемента — кластеризуемые клитики, следует ожидать, что ti_1 и ti_2 будут занимать разные клетки в Правиле Рангов. Для этого надо параллельно решить на материале древнечешского языка задачу, аналогичную той, которую А. А. Зализняк в 1993 г. решил на материале древнерусского: доказать, что древнечешский язык подчинялся закону Ваккернагеля, и выделить полный состав Правила Рангов. Это и является второй (или первой) задачей, которая решается в статье [23] на материале избранного древнечешского памятника — «Далимиловой Хроники».

Доказательство существования некоторого класса языков мира — столь же специальная и частная задача, как и анализ одной отдельно взятой энклитической частицы в конкретном языке. В статье [24] «Parametrizing Verb-Second Languages and Clitic-Second Languages» (2015 г.) проверяется гипотеза о том, что два класса языков с элементами «второй позиции» — языки со второй позицией финитного глагола (V2) и языки со второй позицией энклитик (CL2) — образуют общий класс. Параметризация каждого из этих классов дает ключ к доказательству основной гипотезы. Если удастся выделить сходные правила, регулирующие дистрибуцию финитных глаголов и клитик, гипотеза о существовании общего класса «языков второй позиции» операциональна, если нет — их объединение может быть иллюзией. Доступные мне данные скорее говорят в пользу гипотезы о том, что разные типы элементы второй позиции имеют ряд общих комбинаторных свойств, но полного параллелизма нет.

В последней статье раздела — [25] «1P orders in 2P languages» (2015 г.) специально исследуется одно из явлений, сближающих языки с ограничениями на позиции финитного глагола и языки с ограничениями на место клитик, — возможность опережающего появления диагностической категории не во второй, а в *первой* позиции в качестве запасной стратегии линеаризации предложения. Есть два сценария. При одном из них порядки 1P ~ 2P альтернируют в пределах одного и того же типа клауз, причем порядок 1P коммуникативно маркирован. Большинство языков, причисляемых к типу V2, на самом деле являются языками, разрешающими вынос глагола в начало предложения, т. е. языками V1/V2. Вынос глагола в начало предложения (verb fronting), по-видимому, всегда мотивирован потребностями реализации особой коммуникативной структуры. Разумеется, нужна особая настройка грамматики, которая разрешает такую операцию. Языков подтипа CL1/C2, разрешающих коммуникативно-мотивированный вынос клитик, обычно стоящих во второй позиции, в начало предложения (clitic fronting), напротив, мало. Мне удалось обнаружить два случая: так ведут себя

клитики в словенском языке и в языке кашибо-какатайбо (семья пано, Перу). Clitic fronting — это операция, которая может реализоваться лишь при условии, что кластеризуемые элементы не являются строгими энклитиками. Второй сценарий состоит в том, что одни и те же клитики в разных типах клаузов ведут себя по-разному: в клаузах одного типа они выступают в роли начальных проклитик, а в клаузах другого типа ведут себя как проклитики. Так устроены македонский язык и австронезийский язык рикавунг пуюма: в них смена просодического типа клитики служит маркером типа клаузы.

VI. Текст и язык

В заключительный, шестой, раздел вынесено три работы, которые обсуждают возможные пути лингвистической интерпретации нескольких филологических проблем, связанных с поэтическим языком. Мое внимание к этим проблемам возникло на почве анализа средневековых древнеисландских текстов, которые я параллельно изучал как литературные памятники, комментированные издания которых на русском языке я готовил.

Одна из главных тем стиховедения — соотношение ритмического и синтаксического членения, становление и эволюция метрики, которая может серьезно отклоняться от свойственных самому языку стратегий ритмизации и существенно их деформировать. Еще одна проблема — появление новых форм и приемов, например строфики, аллитерации, рифмы и т. п. Иногда происхождение таких инноваций неясно, так как поэтическая традиция может быть герметична, носить эзотерический характер и по своему характеру противиться внешнему влиянию. В подобных случаях филологи разбиваются на два лагеря: сторонников культурного заимствования, точный адрес которого неизвестен, и сторонников автохтонного развития, механизм которого непонятен. Полемика на эту тему сильно напоминает разбиение лингвистического сообщества на сторонников гипотезы о том, что инновации в лексике и грамматике возникают в силу внешнего влияния (субстратного, адстратного и т. п.), и на сторонников гипотезы о том, что внутренних причин вполне достаточно для объяснения эволюции языка. В моей статье [26] «Порядок, метр, прием» (2003 г.) вся эта проблематика обсуждается на примере скальдической строфической метрики, которая не имеет прямого аналога в германском стихосложении. Я прихожу к выводу, что ни одна поэтическая традиция не существует в полном вакууме и что признание внешнего культурного влияния на отдельные аспекты скальдической строфики не принижает оригинальности и культурной ценности эксперимента скальдов над своим языком. Сейчас мне кажется, что в спорных случаях при описании инноваций в языке, находящемся в ситуации контактного развития, обоснованная гипотеза о заимствовании или внешнем влиянии по умолчанию более правдоподобна, чем обоснованная с той же степенью тщательности гипотеза о внутреннем развитии. Хотя этимологи и историки языка обычно считают иначе, см. [Trubetzkoy 1931]. Возможно, они правы, поскольку консенсуса о том, что такое «спорный случай», в ситуации возможного

контактного влияния нет, а бездумное применение гипотезы о заимствовании дискредитирует этот подход⁸.

В статье [27], носящей название «The text-building function of names and nicknames in “Sverris saga” and “Böglunga sogur”» (1994 г.), обсуждается уже не поэзия, а *правда*, т. е. проза, претендующая на фактическую достоверность, но жанровая специфика присутствует и здесь. Исследуя так называемые «Саги о Посошниках»⁹, которые повествуют о недавних событиях, я обратил внимание на то, что огромное количество личных имен в этом тексте фактически структурировано в виде базы данных. Форма именования и частота употреблений персонажа указывают на его ранг в этой базе, а сама она отражает структурирование информации в сознании рассказчика. Это неудивительно, так как сага по своему происхождению — жанр устного текста, универсальная форма устного рассказа о любых событиях, претендующих на общезначимость. Одновременно она есть *форма хранения информации* при устной передаче: если сообщение неструктурировано или содержит лишнюю информацию, слушатель не запомнит его правильно и не сможет рассказать сам. В начале 1990-х гг. мне доводилось рассказывать сюжет статьи [27] публично. Один из критиков возразил, что так работает передача информации в любом устном рассказе. Для меня это комплимент: я не считаю правильным объяснение специфического материала, которое игнорирует универсальные основания. Специфика все же есть. Далеко не каждый устный рассказ выливается в сагу, которую можно исследовать как текст. К тому же память современного человека гораздо слабее и менее надежна, чем у человека младописьменного общества. В этом я убедился сам, пересматривая статьи этой книги.

Последняя включенная в книгу заметка [28], озаглавленная «Бремя чужих слов» (1994 г.), — это рассуждение о том, что такое «чужое», «свое» и «ничье» в языке. Я утверждаю, что есть связь между принципами цитации и порождения гномических текстов и такими языковыми механизмами, как аналогия. Принцип цитаты — характеристика культур, где есть сакральные тексты, приписываемые боже-ству или иному авторитету: авторитет признается *автором* текста, который нельзя воспроизводить, искажая его вербальную форму. Но у рядовых пользователей

⁸ Стоит прислушаться к Н. С. Трубецкому, который в 1931 г. утверждал, что выдвижение гипотезы о субстрате не имеет смысла применительно к единичным явлениям, пока не будет собран весь материал [Trubetzkoy 1931]. Это весьма разумное соображение, из которого вытекает, что на разной стадии исследования контактных явлений дефолтными, выбираемыми по умолчанию, могут быть разные объяснения. Вполне в духе принципа дополнительности и множественности описания.

⁹ Ударение на первом слоге. Переводя др. исл. *vaǰlar*, «сторонники партии епископа», букв. «посохи» (к *böggull* «епископский посох»), на русский язык словом «Посошники», я следую традиции, которую создал М. И. Стеблин-Каменский. В русском языке слово *посошный*, конечно, есть, но оно имеет иное значение. *Посошное налогообложение* — это обложение налогом *по сохе*. В русском языке выражение «посошник» неоднозначно — это либо лицо, платящее посошный налог (посошник), либо церковный служака, следящий за посохом (посошник). В реалиях Норвегии начала XIII в., которые описывает «Сага о Посошниках», *посошник* — это вооруженный ополченец и, в современных терминах, член политической партии, созданной епископом.

языка, по крайней мере в рамках традиционной культуры, авторских текстов может и не быть: буквальная запись того, что сказано, конечно, может трактоваться как фиксированная вербальная форма и в таком качестве включаться в корпуса текстов. Но это специальная задача лингвистов и специалистов по искусственному интеллекту, занимающихся извлечением информации и атрибуцией запечатленного сообщения конкретному человеку. Большинство людей не относится к сообщением друг друга как к авторским текстам и воспринимают их либо как информацию об окружающем мире, либо как выражение мыслей и эмоций, которые тоже, в сущности, являются типовыми, а не индивидуальными. В дописьюменной и младодписьюменной культуре иного пути вообще нет, хотя какая-то важная информация, например, уложения законов, описание ритуалов и божественное откровение, длительное время может передаваться изустно, в буквальной форме. Напротив, общие истины, трюизмы, приметы, поговорки в любой культуре не предполагают поисков автора и обращены к прототипической, родовой ситуации. Большинство текстов имеют жанровые клише — их несоблюдение приведет к отказу признать соответствие текста данному жанру. Но и экспансия этих клише приведет к нежелательному результату. Грамматические правила, как и трюизмы, — общее достояние и одновременно «ничье», каждый примеряет их к своему опыту. Многие речевые ошибки, которые лингвист объяснит грамматической аналогией, вроде субнормативных употреблений рус. *^{??}Я считаю о том, что, или ^{??}мне это органично, возможно возникают не в силу того, что в мозгу говорящего активизировалось иное правило (как задним числом объясняют лингвисты), а в силу того, что говорящий ориентируется на некое запавшее в ему в душу прецедентное клишированное употребление. Если это так, чужой речевой акт действительно может обладать магической силой. Сейчас мне, правда, кажется, что нужно сделать оговорку. Если такие случайные или целенаправленные эксперименты с языком на основе неточно освоенных чужих прецедентных употреблений (рус. *мне понятно, мне кажется понятным, это органично, это кажется органичным* → ! **мне органично, мне кажется органичным*) станут массовыми, все равно придется прибегать к объяснениям в терминах изменения правил или изменениям нормы (печально для тех, кто является носителем старой) или даже к гипотезе о наличии внутри некоторого языка идиомов, где можно и где нельзя сказать ^{??}мне органично. Но это уже тема для другой статьи, которой нет в этой книге. Возможно, такая статья еще не написана.

ANTON ZIMMERLING.
FROM INTEGRAL FRAMEWORKS TO ASPECTIVE ANALYSIS

I. Theory. Method. Conclusiveness

- [1] From Integral Frameworks to Aspective Analysis (2019)
- [2] The Whats and the Whys. Exploring Scientific Articles (2019)

II. Diachrony. Synchrony. Dynamics

- [3] Die unpersönlichen Satzmodelle in der altisländischen Sprache (1989)
- [4] Between Synchrony and Diachrony. Prosodic Oppositions in the Scandinavian Languages (1999)
- [5] Early Germanic Accent: Phonetics and Comparative Grammar (2006)
- [6] Who Will *Oversleep the Matins* and *Take Fish*? Nominative Object of the Infinitive as a Grammatical Phenomenon (2018–2019)
- [7] Two Dialects of Russian Grammar: Corpus Data and Formal Models (2018)

III. Semantics in the Grammar Mirror

- [8] The History of One Polemic (1998)
- [9] The Subject of State and the Subject of Evaluation (1999)
- [10] Russian Predicatives in the Perspective of Sociolinguistic Experiment and Corpus Grammar (2017)
- [11] Impersonal Constructions and Dative-Predicative Structures in Russian (2017–2018)
- [12] Epistemic Predicates and Mental States (2017)
- [13] To Own and to Be Near (1999)
- [14] The Logic of Prognosis (1997)

IV. A Two-Piece Solution. Non-canonical Subjects

- [15] Dative Subjects and Semi-expletive Pronouns in Russian (2007–2008)
- [16] Non-canonical Subjects in Russian (2011)
- [17] Sentential Arguments and Event Structure (2014)
- [18] The Corpus Grammar of Russian QPs (2019)
- [19] Topicality and Non-Canonical Subjects in Russian (2019)

V. Word Order. Theory and Typology

- [20] Local and Global Rules in Syntax (2008)
- [21] Word Order Systems in Slavic Languages (2012)
- [22] A Unified Analysis of Clitic Clusters in World's Languages (2012)
- [23] Clitic Templates and Discourse Marker *TI* in Old Czech (2013)
- [24] Parametrizing Verb-Second Languages and Clitic-Second Languages (2015)
- [25] 1P orders in 2P languages (2015)

VI. Text and Language

- [26] Order. Meter. Poetic Technique (2003)
- [27] The Text-building Function of Names and Nicknames in «Sverris saga» and «Böglunga sogur» (1994)
- [28] The Burden of Others' Words (1994)

PART 1
THEORY. METHOD. CONCLUSIVENESS

CHAPTER 1

“From Integral Frameworks to Aspective Analysis” (2019)

In this chapter, I argue that linguistic analysis based on the principle of complementarity should be preferred over framework-internal analysis based on purely monistic principles. Natural language has high complexity. Integral frameworks aim at covering a wide range of natural language phenomena but fit only to their minor part, e.g. selected modules or interface mechanisms, while the rest gets an inaccurate or distorted description. Therefore accurate linguistic explanations compatible with more than one framework have high priority. The accuracy depends on the linguist’s ability to propose a model that predicts the distribution of language features on some data collection. The parametrization of cross-linguistic and language-internal variation makes predictive models more efficient since it gives a chance to test them on a class of languages or dialects. I argue that the main contribution of the framework-based research in linguistics is the predictive models developed for the analysis of some specific mechanisms, e.g. word order or predicate-argument analysis or case or agreement, etc. The extrapolation of these models to other language phenomena decreases their accuracy since framework-oriented versions of such models contain redundant information and framework-internal assumptions. The linguists apply to two kinds of models: the first one corresponds to the mathematical notion of model, while the second one to the models of natural sciences. I introduce the distinction of model₁ and model₂ notions in linguistics. Models₁ are metalanguage notions belonging to some harmonized set of cross-linguistic categories: they apply to all world’s languages and can be treated as universal. Models₂ are hypotheses tested on some collections of language data. Contrary to the claims made elsewhere, model₁ notions cannot be eliminated from linguistic description, but model₂ notions do not always correlate with model₁ notions. In this chapter, I show the interplay of models₁ and models₂ on two selected issues in Russian and typological linguistics, notably — the syntax of the zero form of the verb “BE” in the history of Russian and the clitic ordering in the class of the world’s languages with Wackernagel’s law. These problems can and in the majority of cases must be studied on an independent basis yet the diachronic analysis proves that in some cases it is reasonable to treat them as parts of the same research complex.

Keywords: natural language, predictive models, frameworks, parametrization, cross-linguistic categories, monism, the principle of complementarity, general grammar, partial grammars, clitics, Wackernagel’s law, zero verb forms.

CHAPTER 2

“The Whats and the Whys. Exploring Scientific Articles” (2019)

This chapter discusses the perspectives and limitations of a research article as one of the main forms in linguistic research. A linguistic article usually includes an aspective description of language phenomena. The real impact of a linguistic article crucially depends on external factors such as the communication between scientists and scientific schools, place, and language of the publication, etc. Nevertheless, the productivity of an article, i.e. its potential to initiate or resume a polemic is determined by the fact, whether the research object is defined correctly. In this perspective, I introduce my collection of articles written in 1989–2019 and included in this book. They cover five research fields — 1) harmonization of the methods of historical vs synchronic linguistics, 2) grammatical semantics, 3) subject-predicate analysis, 4) word order typology, 5) text analysis.

Keywords: natural sciences, mathematic sciences, linguistic research, conclusiveness, general linguistics, partial linguistics, typology.

PART II DIACHRONY. SYNCHRONY. DYNAMICS

CHAPTER 3

“Die unpersönlichen Satzmodelle in der altisländischen Sprache” (1989)

This chapter is a study devoted to the interplay of innovative processes and archaisms. Innovations in grammar can contribute to the preservation of presumably archaic sentence patterns. I argue that this was the case with Old Icelandic. Old Icelandic is a language with rich inflexional morphology, its syntax is usually considered archaic. Meanwhile, many morphosyntactic features including the stability of the four-case system in Old Icelandic and the productivity of the dative case-marking on the grammatical subject and grammatical object are probably innovative. Basing on the productivity of Dative — Predicative Structures with slots for a copular element and a non-agreeing predicative complement, I argue that Old Icelandic developed a class of nominal predicatives similar to Slavic Category-of-State forms. The productivity of impersonal sentence patterns in Old Icelandic is a result of a relatively late process. In XI — XIII centuries, Old Icelandic generalized a three-part copular pattern $N_{\text{DAT}} — V_{\text{LINK}} — \text{PRED}$, with dative case-marking on the subject of a non-agreeing nominal predicative. This process involved a reanalysis of several verbal constructions as copular and was a prerequisite for the formation of a class of special predicative forms lacking adverbial, adjectival, or nominal side uses.

Keywords: Old Icelandic, internal reconstruction, archaism, innovation, impersonal sentences, semantic subject, copulas, predicatives.

CHAPTER 4

“Between Synchrony and Diachrony. Prosodic Oppositions in the Scandinavian Languages” (1999)

This chapter discusses a traditional problem of Scandinavian accentology — the origin of the tonal accents and their functional load. I argue that the geographical distribution of Scandinavian accents should be explained from the perspective of morphological change and different scenarios of language evolution. The binary opposition of word prosodies in Swedish and Norwegian results from an accentual restructuring, which took place in Swedish and Norwegian dialects in the period, when the suffixed definite article has established and the case morphology was becoming redundant. The binary opposition ‘Accent 1’ vs ‘Accent 2’ was generalized in the group of dialects with average tempi of language evolution: the case system was already simplified, but the nominal declension paradigms still contained word forms with a different number of syllables. The reconstruction of the binary opposition of word prosodies for other groups of Scandinavian dialects is unlikely since the dialects from these groups lacked morphological conditions for generalizing such an opposition. This holds both for the more radical dialects in the Danish area, which lost the case morphology earlier, and for the conservative dialects of the island type (Icelandic, Faeroese, Dalarna, etc), which preserved the 3–4 case system and reanalyzed the case-marked suffixed articles as a special series of secondary inflection forms. There are no grounds to reconstruct a binary opposition of word accents for Proto-Scandinavian, let alone for Proto-Germanic.

Keywords: Scandinavian languages, accentology, segmental basis, suffixed article, monosyllables, polysyllables, word accent, stød, tones.

CHAPTER 5

“Early Germanic Accent: Phonetics and Comparative Grammar” (2006)

I offer a new explanation of Verner’s Law (1876), which predicts the correlation of Germanic and Indo-European consonants and discuss the optimal procedure of combining the methods of diachronic

and synchronic linguistics. I show that Carl Verner proposed a phonetically untenable interpretation of the exact and accurate comparative correlations between Indo-European and Early Germanic and argue that the voiced Germanic fricatives in the word-internal position, contrary to Verner and the following mainstream research prove not the loss of the free Indo-European accent, but the preservation of some of its effects in Early Germanic. The traces of the post-root accent are also found in a group of Germanic verbs with Holzmann's law. The reinforcement (gemination) of the glides *-j-*, *-w-* >> *-jj-*, *-ww-* is attested in Germanic verbs historically belonging to the mobile accentual paradigm. The gemination *-j-*, *-w-* >> *-jj-*, *-ww-* in this group of verbs, just as the voiced Inlaut fricatives directly reflect not the accent retraction towards the root/preceding syllable, but the privileged status of the non-initial syllable in the words from the mobile accentual paradigm in Early Germanic. Holzmann's law, unlike Verner's law, has a narrow segmental basis, but their underlying phonetic mechanisms are similar. The failure of Verner and his followers to find a correct solution is due to their attempt to provide a uniform explanation for Germanic voiced fricatives in the Inlaut and Auslaut positions. The Auslaut position is a clear example of phonetic weakening, while the exact position of the word-internal fricatives and glides is a matter of debate: Holzmann's law, Verner's law as well as the reflexes of voiced fricatives in Gothic suggest that the consonant could belong to the subsequent accented syllable or be ambisyllabic.

Keywords: sound laws, Verner's Law, Holzmann's Law, phonetic change, voiced consonants, accent, timing, Germanic languages, Indo-European languages.

CHAPTER 6

“Who Will Oversleep the Matins and Take Fish? Nominative Object of the Infinitive as a Grammatical Phenomenon” (2018–2019)

This paper is addressed the East Slavic/ Old Russian construction with nominative case-marking on the infinitival object, cf. ЗЕМЛЯ ПАХАТИ, РЫБА ЛОВИТИ, which has been a matter of a long debate in Slavic studies. I discuss the probability and verifiability of diverse hypotheses on the origin, evolution, and distributive features of this construction in the history of Russian. I argue that the nominative case-marking on infinitival non-subject arguments instantiates the mechanism of differential object marking since the speakers were able to use both the accusative and the nominative objects in the same syntactic position: the choice of the nominative vs accusative case was triggered by such pragmatic factors as communicative status and referentiality. The data provided by the historical text corpora testify that the modal semantics of the ЗЕМЛЯ ПАХАТИ construction did not favor its use in most types of embedded clauses, especially — in infinitival complements of causative and volitive predicates. In these syntactic contexts, the accusative remained the strongly preferred option. With impersonal modal verbs and predicatives, the nominative and the accusative have comparable frequency figures.

Keywords: nominative case, nouns phrases, infinitive clauses, differential argument marking, Russian, Old Russian, historical text corpora.

CHAPTER 7

“Two Dialects of Russian Grammar: Corpus Data and Formal Models” (2018)

This paper is addressed the problem of parametric variation in Russian grammar, with the focus on copular constructions with agreeing and non-agreeing adjectival predicates. Basing on Russian National Corpus, I reconstruct two dialects in Russian morphosyntax. They differ regarding the assignment of the predicative instrumental case, raising conditions and the distribution of agreeing vs non-agreeing predicates after *быть* ‘be’, *стать* ‘become’ and *казаться* ‘seem’. The standard idiom tagged ‘Russian-A’ only licenses predicative instrumental on adjectives after SEEM (*казалось странным, что P*) and non-agreeing predicatives after non-zero forms of BE or BECOME (*было странно, что P*), while the idiom tagged ‘Russian-B’ allows non-agreeing forms after SEEM (*казалось странно, что P*) and forms of the

predicative instrumental case after non-zero forms of BE and BECOME (*было странным, что P*). I argue that the differences between Russian-A and Russian-B must be explained in terms of parametric settings and claim that Russian predicatives lack any case forms incl. the forms of the predicative instrumental. The assignment of the predicative instrumental to adjectival heads can be explained as subject control in all dialects, but only Russian-B allows raising of sentential arguments to the position of the matrix subject.

Keywords: Russian, corpus linguistics, parametric grammar, adjectives, predicatives, predicative instrumental case, agreement, control, raising, dialect variation.

PART III SEMANTICS IN THE GRAMMAR MIRROR

CHAPTER 8

“The History of One Polemic” (1998)

The paper is devoted to the so-called Category-of-State word class attested in Modern Russian, Old Slavic, Old Germanic languages, and elsewhere. The famous Russian linguist Lev Ščerba put forward a hypothesis that Russian tends to develop a special part of speech he dubbed ‘Category of State’. Category-of-State forms are nominal predicatives realized in copular sentences without overt agreement controllers. Many of them are used in the Dative-Predicative structures. I argue that while Ščerba’s claim that the Category of State makes a separate part of speech is wrong, Ščerba’s idea that Russian nominal predicatives always convey a special taxonomic meaning is correct. In more modern terms, Ščerba’s Category-of-State forms can be analyzed as Stage-Level predicates denoting a spatiotemporal stage of their object. I falsify the claims advanced by Ščerba’s followers Alexander Isačenko and V́ctor Vinogradov and demonstrate an exact parallel from Old Icelandic — a language, where the class of nominal Category-of-State forms conveying the same taxonomic meaning was even larger. The puzzling failure of the Germanicists to recognize this class in Old Icelandic is due to the fact that they were unaware of Ščerba’s hypothesis and the external typological parallels.

Keywords: category-of-state, Russian, Old Icelandic, word classes, predicate taxonomy, syntactic derivation, stative predicates, adjectives, predicatives, copula.

CHAPTER 9

“The Subject of State and the Subject of Evaluation” (1999)

This paper presents a semantic apparatus for the analysis of situations with an inactive animate participant. It is illustrated by the data of Modern Russian, which has different kinds of predicative constructions with an inactive animate participant. I discuss the possibility of recognizing an extra argument with the role like “outer world” or “locus”: this argument often remains covert or is expressed by a sentential complement. The situations described by verbal predicates like Rus. *menja zlit, trevožit, čto P* can be analyzed as including an Agent argument, while the overt animate participant is patient-like. True stative predicates lack any Agent argument. A special characteristic of Russian grammar is not the opposition of non-verbal statives to the agentive verbal predicates but the dedicated status of internal states requiring an animate subject and the consequent distinction of internal vs external states. The categorial opposition of the meanings like Rus. *mne veselo* ‘I am having fun’, *mne dušno* ‘I feel stuffy’ (internal state) vs the meanings like Rus. *zdes’ pyl’no, vetreno* ‘It is dusty, windy here’ (external state) is just as relevant as the opposition Rus. *mne veselo* ‘I am having fun’ (stage-level predicate) vs *ja vesel’us* ‘I am doing something that makes me have fun’ (action) vs *ja veselyj* ‘I am funny’ (individual-level predicate). The predicates *zdes’ pyl’no* ‘It is dusty here’, *sejčas solnečno* ‘It is sunny now’ have the same truth value for every observer: standard Russian does not license combinations with an overt dative subject like **mne pyl’no, *mne solnečno*. This kind of predicates is called ‘constatives’ in this paper. Russian predicatives selecting dative subjects, cf. *mne veselo, skučno, xolodno, dušno* denote

spatiotemporal situations, where the truth value (whether X is having fun, is bored, is cold, feels stuffy) can only be established for the specific individual, i.e. the subject of the internal state. The predicates like *èto amoral'no* 'This is immoral', *èto alogično* 'This is illogical' bring about a built-in assessment that claims to be universally valid: it is meant that every observer will value *p* as 'immoral' or 'illogical'. This second type of the animate participant is called 'subject of evaluation (assessment)' in this paper. I argue that internal states are always linked both to the subject of state and to the subject of evaluation, while the predicates of assessment generally lack a subject-of-state argument.

Keywords: predicate taxonomy, stative predicates, internal states, constatives, evaluations, subject of state, subject of evaluation.

CHAPTER 10

“Russian Predicatives in the Perspective of Sociolinguistic Experiment and Corpus Grammar” (2017)

This paper aims at measuring the active vocabulary of Russian predicatives licensing dative-predicative-structures (DPS) and introduces the results of the sociolinguistic experiment paired with the corpus study for the same set of stimuli. The sociolinguistic experiment aimed at measuring the active DPS vocabulary in the idiolects of 18 native speakers of Russian and produced a sample of 422 stimuli ranked according to their acceptance rate. The same set of stimuli was tested on data from the Russian National Corpus, which produced the second sample based on the frequency ranks of DPS clauses in the diagnostic context *мне Z-во*, with the priority dative argument placed contiguously to the predicative at the distance <-1:1> in the same clause. The correlation of two ranked samples predicts the volume of the DPS vocabulary in an average idiolect of Russian and helps to establish the proportion between the lexical nucleus of the DPS construction and its extension. The analysis verifies the hypothesis that native speakers of Russian apply to the same underlying principles of semantic selection for DPS elements while using non-identical sets of lexical elements. The paper introduces a formal model of lexical extension with the entire DPS being grouped into 15 thematic classes. The results prove that each thematic class includes high-, middle- and low-frequency DPS elements. Low-frequency DPS predicatives are modelled after high-frequency elements from the same thematic class.

Keywords: construction grammar, lexicon, predicatives, dative-predicative-structures, idiolect, lexical extension, lexical reduction, thematic class.

CHAPTER 11

“Impersonal Constructions and Dative-Predicative Structures in Russian” (2017–2018)

This paper analyzes two Russian constructions, verbal impersonals in 3Sg ~ 3Pl and dative-predicative structures with a non-agreeing predicative, and discusses the problem of lexical-syntactic interface and the borderline between lexicon and grammar. Productive constructions are defined as bilateral units and analyzed in the perspective of selective conditions imposed on lexical extension, not in the perspective of idiomaticity or non-compositionality. Languages with similar morphosyntax can develop similar constructions with the same underlying semantics but these constructions often show different scenarios of the lexicon-vs-grammar interaction. Modern Russian takes an intermediate position between languages where verbal impersonals and predicative structures are grammaticalized and languages where these constructions belong to the lexicon. The distribution of the impersonal patterns in 3Sg and 3Pl in Modern Russian encodes a non-trivial interaction of lexicon and grammar. The key factor in the stability of Russian dative-predicative structures is the distinction between three classes of adjectival stems and the preservation of stems blocking the derivation of predicates that denote internal states.

Keywords: constructions, Germanic languages, grammar, impersonals, lexicon, predicatives, Russian, Slavic languages, s-selection, zero subjects.

CHAPTER 12

“Epistemic Predicates and Mental States” (2017)

This paper discusses the behavior of value predicates and predicates denoting external states in the context of an epistemic operator as well as regular homonymy of constructions expressing internal states linked with a referential animate subject vs constructions of epistemic value. I focus on the data from Modern Russian, Modern Bulgarian, and Old Icelandic and prove that the epistemic operator SEEM is in some contexts expressed by auxiliary BE-constructions. At the same time, in contexts of a different type, the verb SEEM does not bring about the contrast with non-epistemic predicates. Two dialects of Modern Russian display similar though non-identical raising conditions for sentential arguments taking the subject position of the matrix verb *казаться* ‘seem’. The standard dialect only licenses adjectival forms in the predicative instrumental case, whereas another dialect also licenses non-agreeing predicatives in this context.

Keywords: internal states, external states, relational predicates, epistemic values, sentential arguments, raising, adjectives, predicatives, infinitive, finite clauses, Russian, Bulgarian, Old Icelandic.

CHAPTER 13

“To Own and to Be Near” (1999)

This paper discusses the interactions of the syntactic and semantic approaches to possessivity. Possessivity is a binary relation that can be interpreted in purely semantic terms as well as in mixed semantic-syntactic terms. I focus on the transpositions from the intransitive BE-type (X is at Y/ near to Y) to the transitive HAVE-type (X has Y). In many cases, the position of a possessee in these constructions is filled by a non-material object, cf, HOPE, SUSPICION, which gives rise to semi-possessive sentences, i.e. sentences where the possessive syntax is extended to non-possessive semantic fields. I also discuss the iteration of the possessive operator, notably, in the *U* + Gen type, cf. substandard Russian *U menja u teti pyatnadcatj detej* ‘My aunt has 15 kids’ lit. ‘At me at aunt 15 children’.

Keywords: predicate-argument analysis, syntactic constructions, possessivity, possessor, object of possession, localization, Slavic languages, Germanic languages.

CHAPTER 14

“The Logic of Prognosis” (1997)

This paper introduces and discusses three models of the linguistic representation of Future: a determinist model, a stochastic model, and the model of True Prophetic knowledge. All three models coexist in natural languages and are represented both in their grammatical systems as well as in the text-building / discourse strategies the speakers and authors apply to.

Keywords: logic of tense, future, prognosis, verifications, narrative strategies, speech acts.

PART IV

A TWO-PIECE SOLUTION. NON-CANONICAL SUBJECTS

CHAPTER 15

“Dative Subjects and Semi-expletive pronouns in Russian” (2007–2008)

The paper discusses Russian sentence patterns without overt nominative subjects and interprets them in a constraint-based generative framework with syntactic zero categories and oblique subjects, i.e. oblique NPs/DPs taking the subject position. These patterns are classified into the subclass with strong non-nominative subjects, including the patterns with the so-called Mel’čuk’s zero subjects (\emptyset^{3SG} *улицу засыпало песком*, \emptyset^{3PL} *в своем доме не гадят*), Dative-Infinitive Structures (*грузовикам здесь не проехать*, *им*

еще учиться и учиться, Всем встать!) and the subclass with weak non-nominative subjects including patterns with Dative-Predicative Structures (*ей было стыдно, ему стало холодно*). A prominent feature of Dative-Predicative Structures is that the predicates capable of taking dative subject NPs specified as [+animate] can also take clausal subjects (*Интересно, что P*) and the semi-expletive stressless pronoun *это* (*Это было интересно*) but cannot take all three types of weak subjects simultaneously.

Keywords: grammatical subject, strong subjects, weak subjects, expletives, sentential arguments, Russian.

CHAPTER 16

“Non-canonical Subjects in Russian” (2011)

The paper is devoted to non-canonical subject marking in Russian. I argue that the criterion of role opacity does not fit the non-canonical subject marking in languages like Modern Russian. Minor sentence patterns always exhibit special role-and-reference features triggered by the choice of a non-canonical subject. Russian has default nominative-accusative alignment with subject-verb agreement. In finite clauses, the nominative subject is the agreement controller, which to a certain extent is true for at least one type of non-canonical subjects — the so-called ‘Mel’čuk’s zeroes’, i.e. zero subject pronouns in 3 p., which have inflexional features and characteristic role-and-reference features. In infinitive clauses, where the infinitive is the predicate head there is no overt agreement and the subject is case-marked with dative (DIS structures). There is no one-to-one correspondence of role features and case-marking since dative subjects in finite Russian clauses with a nominal non-agreeing predicative (DPS structures) and in DIS structures convey different semantics. I argue that all kinds of grammatical subjects in Russian pattern with two groups — strong and weak subjects. The class of strong grammatical subjects includes standard nominative NP/DPs, the dative subjects of DIS and Mel’čuk’s zeros: they do not combine and do not compete with any other subject-like expressions. The class of weak subjects includes the dative subjects of DPS, expletive non-referential pronoun *это*, and sentential arguments: they exhibit distributional restrictions and compete for the unique subject position.

Keywords: subjects, non-canonical arguments, strong subjects, weak subjects, zero subjects, Russian.

CHAPTER 17

“Sentential Arguments and Event Structure” (2014)

This paper is addressed the interaction of subject marking and event structure in languages, which allow sentential arguments in the subject position. In Russian and other Slavic languages sentential subjects share several formal and semantic properties with zero subjects with role-and-reference features and with so-called oblique subjects, i.e. subject-like arguments marked with an oblique case. I argue that sentential subjects represented by bare that-clauses (Rus. *čto*-clauses) cannot have the roles of Agent/Causer, while zero subjects can. I also argue that the capacity of taking *to*, *čto P*-clauses, i.e. that-clauses headed by the correlative pronoun *to* serves as diagnostics for several verbal classes. Causative predicates like *vynudit’*, *zastavit’*, *sklonit’ k čemu-libo* only take *to*, *čto P*-clauses, but not bare *čto P*-clauses as surface subjects. Factive predicates like *znat’*, *razdražat’* etc. take *to*, *čto P*-clauses, but not bare *čto P*-clauses as surface subjects, while non-factive predicates like *dumat’*, *mereščit’ sa* only take bare *čto P*-clauses. Nominal predicatives licensing Dative-Predicative-Structures (DPS) with an oblique subject marked with dative case and specified as {+animate; +referential} split into two groups. Russian DPS predicatives from the *stydno*, *dosadno*, *protivno*, *vse ravno* group only take bare *čto P*-clauses and behave as non-factive verbs in all contexts with an overt oblique subject. Russian DPS predicatives from the *izvestno*, *neizvestno*, *stranno*, *bezrazlično* group both take bare *čto P*-clauses and headed *to*, *čto P*-clauses, i.e. can be used in factive contexts as well. That means that their sentential argument can both get the status of a fact, i.e. verified proposition *P*, logical truth, and an intentional situation, e.g. subjective evaluation of *P*, inner vision of *P*, etc. Russian has two expletives — *это* and *to*, but their syntax is different. *Èto* behaves as the surface subject of the matrix clause and alternates

with oblique subjects and sentential arguments in the subject position, while *to* cannot be separated from the complement clause and takes the subject position only in the combination with the CP.

Keywords: grammatical subject, sentential arguments, expletive subjects, zero categories, case marking, semantic roles, factivity, facts, situations, event structure.

CHAPTER 18

“*The Corpus grammar of Russian QPs*” (2019)

The paper is addressed the corpus grammar of Russian quantifier phrases (QPs), with the focus on two issues: (i) subject-predicate agreement patterns in sentences with a QP in the position of a grammatical subject, (b) the choice of the agreeing/non-agreeing form of the adjective in QPs with an embedded NP with the head noun in the feminine gender. QPs license both the plural and the singular form of the predicate. I argue that the singular form optionally shown on the predicate instantiates non-canonic agreement controlled by the QP and does not pattern with the default agreement in 3Sg.N. The analysis is based on the complete statistics of all Russian cardinal numerals used in the RNC in QPs of the type *два человека/пять человек* in the Russian National Corpora. I show the correlations between plural/singular agreement forms, word order (QP—V ~ V—QP), and communicative status of QP. The choice of the agreeing preposed NP-level adjective as in *dve interesnye knigi* does not constrain the form of the predicate agreement while agreeing DP-level elements as in *èti dve knigi* blocks the singular form on the predicate. Russian subject QPs are non-canonic arguments with a decreased topicality: in two-thirds of the corpus data, they lack the status of a theme.

Keywords: numerals, agreement, case, number, phrases, QP, NP, DP, information structure, communicative-syntax interface, Russian language, corpus grammar.

CHAPTER 19

“*Topicality and Non-Canonical Subjects in Russian*” (2019)

This paper discusses the relevance of the topicality feature for the syntax of non-canonic subject-like arguments in Russian. Russian non-canonic arguments include two types of dative noun phrases, quantifier phrases, sentential arguments, the expletive element *èto*, and zero subjects with the role-and-reference features of non-referential Agents. I argue that non-canonic arguments take the subject position according to some hierarchy of non-canonic subjects. Standard subjects case-marked with the nominative case can be used both in Russian active and passive clauses and get the communicative status of both theme and rheme. Non-canonical Russian subjects, except for the quantifier phrases are almost exclusively used in the active clauses, where they usually have a fixed communicative status. Most types of Russian non-canonical subjects are associated with the communicative status of theme in their prototypical uses, while Russian quantifier phrases typically get the status of rheme.

Keywords: syntactic relations, information structure, subject, topic, theme, rheme, non-canonical arguments, zero subjects.

PART V.

WORD ORDER. THEORY AND TYPOLOGY

CHAPTER 20

“*Local and Global Rules in Syntax*” (2008)

The paper discusses word order and phrasal prosody in Russian. I claim that both phenomena can be described in terms of two successive sets of rules — local rules vs. global rules. Combinations of these two sets of rules are typical of multilayer language models and for the algorithmic generation of complex

structural objects in formal grammars. Modern Russian applies to a highly formalized rule of selecting the locus of the main phrasal accent: the hierarchy of potential accent bearers is a mirror image of the grammatical hierarchy of arguments and adjuncts. The order of communicative constituents in Russian is shaped by 7–8 Linear-Accent Transformations (LA-transformations). LA-transformations are Movement rules, which both operate on constituent order and change accent markings of communicative constituents. In the preceding Russian linguistic tradition, cf. Padučeva and Yanko LA-transformations are defined as Context-Sensitive rules, which makes word order calculus impossible. I discuss the possibility to reformulate LA-transformations as pairs of the type <Active & Remnant Movement> and offer an analysis compatible with Mildly Context-Sensitive Grammars, e.g. with Stablerian Minimalist Grammars.

Keywords: locality, local rules, global rules, predictivity, formal grammars, mildly context-sensitive grammars, linear-accent grammars, phrasal accentuation, choice of the accent bearer, communicative structure.

CHAPTER 21

“Word Order Systems in Slavic Languages” (2012)

Slavic word order systems fall into five types, tagged W-systems, W⁺-systems, W^{*}-systems, V-systems, and C-systems in the paper. The first four types have clitic clusters. A cluster is defined as a contact sequence of elements from {a, b, c... n}, where the linear order for each pair a, b is fixed. In Slavic languages only clause-level clitics clusterize. This generalization holds for pronominal dative and accusative clitics that can both be clause-level and phrase-level elements. Clusters are phrases and must be distinguished from occasional sequences of adjacent clitics. Clustering elements represent one domain, adjacent clitics can represent two or more domains. The internal order in a cluster is triggered by Template Rules. There are three areal types of Slavic Template Rules. Slavic languages use up to six different types of clitic hosts, including CPs. All Slavic languages with clusters have Barrier rules that generate derived word orders. The Main Clause Barrier rule typically gives rise to V2 orders as XP — V — CL. Czech and Bulgarian do not apply the Main Clause Barrier rule in declaratives but allow the so-called ensembles, i. e. combinations of initial constituents preceding clusters with the order XP — CL. There is a mismatch between phonetic and syntactic aspects of enclisis. Some Slavic languages allow for post-syntactic parenthetical insertions that do not change well-formedness conditions for enclisis.

Keywords: word order, clitics, clusters, syntactic typology, templates, barriers, Slavic languages.

CHAPTER 22

“A Unified Analysis of Clitic Clusters in World’s Languages” (2012)

This paper proposes a unified analysis of complex syntactic objects defined as clitic clusters. A cluster is a string of elements [_{CLP} abc...n] which can function independently without combining but are arranged in a rigid order when they assume a contact position, so that for each pair (a, b) the linear order a > b, i. e. ‘a precedes b’ is fixed. Elements conforming to this definition are called clustering. Rules ordering language elements in clusters are called *Template Rules*. If clitic-internal ordering is rigid, there is a one-to-one correspondence between the set of clitics {a₁, a₂... a_n} and the clitic’s rank, i.e. the table column taking all clitics a_k ∈ {a}. In the first section, I analyze Template Rules as empiric generalizations made on text corpora representing the normative usage of the languages from the class of the world’s languages with clustering clause-level elements. In the final section, I analyze Template Rules as linearization algorithms and argue that these algorithms pattern with finite-state machines with store memory. The general conclusion is that clusters ordered by Template Rules are normally non-homogenous regarding their morphosyntactic and prosodic value. I furthermore argue that a unified theory of clusters can be built with little or no recourse to the prosody of the clitic elements.

Keywords: clusterization, clitics, syntax, linearization algorithms, template rules.

CHAPTER 23

“Clitic Templates and Discourse Marker TI in Old Czech” (2013)

This paper aims at solving two problems of Slavic syntax — a) to prove that Old Czech had a discourse clitic particle *ti_i* ‘indeed’ attested in other Old Slavic languages including the Old Novgorod dialect of Old Russian, b) to describe Old Czech clitic template for the first time. Both problems are related: the proof that Old Czech *ti_i* was a separate syntactic element is based on the claim that *ti_i* was part of Old Czech clitic template and had a unique slot in it. If the enclitic particles *ti_i* is a different element than the pronominal enclitic *ti₂* ‘you: 2SG.DAT’ and both elements clusterize they take different slots in the clitic template: this prediction is true. This special research task can be solved, if one verifies a more general hypothesis that Old Czech was a language with Wackernagel’s law and identifies all slots in the Old Czech clitic template. This hypothesis is proved for the language of the “Dalimil chronicle”.

Keywords: clitics, clitic clusters, template rules, Old Czech, discourse particles, Wackernagel’s law.

CHAPTER 24

“Parametrizing Verb-Second Languages and Clitic-Second Languages” (2015)

In this chapter, I discuss verb-second (V2) and clitic-second languages (CL2) as an object of the parametric typology. V2 systems are a small group of languages sharing several parameters constraining the clausal architecture and the finite verb placement. CL2 systems are a large group of languages sharing several parameters constraining the clausal architecture and clitic ordering. The definitional property of both V2 and CL2 syntax is the bottleneck condition licensing only one constituent in XP. The ordering of finite verbs in V2 languages and the ordering of clitics/clitic clusters in CL2 languages is explained as head movement of the diagnostic category (verb or clitic) to 2P accompanied by the XP-movement of phrasal categories. An extension of the class of rigid V2 languages is the class of V1/V2 languages which license IS-motivated V1 orders in declarative clauses, while CL1/CL2 languages are rare.

Keywords: parametric typology, movement, clausal architecture, verb-second languages, clitic second languages

CHAPTER 25

“P1 orders in 2P languages” (2015)

I discuss the phenomenon of Information-Structure (IS) motivated clausal initial placement of diagnostic categories conventionally labelled clausal second (2P) elements, notably, finite verbs in Verb Second languages (V2) and clustering clause-level clitics in Clitic Second Languages (CL2). The main research issue is whether V1 orders in V2 languages and CL1 orders in CL2 languages have identical IS-triggers or not. If the answer is positive, one gets additional support for the hypothesis that V2 languages and CL2 languages form together one class in terms of shared constraints on clause structure. The analysis of 1P orders in two classes of 2P languages — V2 languages and CL2 languages supports the hypothesis that V2 languages and CL2 languages form a class definable in terms of shared constraints on clause structure. 1P orders in 2P languages are best explained as overt syntactic movement of an operator category (finite verb or clustering clitic) to the clause-initial position. I demonstrate that 1P orders in 2P languages arise due to two different mechanisms — 1) IS-triggered optional alternation of 1P ~ 2P orders in the same clause type or 2) clause type shifting. V1/V2 languages have optional IS-triggered verb fronting in declarative clauses, while V1 orders in declaratives may encode several different IS types including rheme dislocation andthetic structures. CL1/CL2 languages and rigid CL2 languages may have optional IS-triggered alternation of CL1 ~ CL2 orders in some types of clauses. Clause type shifting may be manifested by the declarative vs non-declarative split (Germanic languages), nominal-type clause vs verbal-type clause split (Macedonian), or transitive vs intransitive split (Rikavung). The data

from CL1/CL2 languages with clause type shifting undermines the analysis in terms of edge features [+EF] and brings negative evidence against the practice of ascribing the characteristics of clause structure to syntactic heads in languages where different clause types have different positioning of the same fixed positioned operator elements. Kashibo-Kakataibo person markers and Macedonian clitic pronouns do not have any edge feature requiring proclisis or enclisis in imperative clauses but a different configuration in declarative clauses. What matters is that some of these clauses require IS-motivated fronting of non-clitic categories, while some other clauses block it. Rikavung clitic pronouns in transitive and intransitive clauses have different distributions not because they get an elusive [+EF] feature in intransitive non-negative clauses and lose it in transitive clauses but because Rikavung intransitive verbs must be fronted in non-negative clauses. CL1/CL2 ordering is an interface phenomenon since clitics must have non-syntactic characteristics allowing them to change prosodic orientation and to be fronted.

Keywords: 2P languages, verb-second languages, clitic-second languages, V2/V1 languages, CL1/CL2 languages, information structure, clause shifting, Kashibo-Kakataibo, Rikavung, Macedonian, Slovenian, German.

PART VI TEXT AND LANGUAGE

CHAPTER 26 “*Order. Meter. Poetic Technique*” (2003)

The paper discusses the phenomenon of Skaldic poetry, specifically its relation to presumably common Germanic poetic forms. Skaldic stanzas do not have direct counterparts in Old Germanic metrics. Most Skaldic meters can be classified as alliterative verse with the obligatory rhyme in the verse end. I discuss the relation of poetic and syntactic segmentation in the Skaldic meters and argue that certain features like obligatory rhyming in a long verse line are due to external European influence. The disintegration of the metric vs syntactic segmentation is realized on two different paths. The principal Skaldic meter (*dróttkvætt*) is characterized by severe violations of the word order, but the prosody of the poetic line does not impose strict restrictions on the selection of language material. In other Skaldic meters, the prosody of the line constrains the syntax, which leads to the parcelling of the sentence into groups of words united, first of all by the metric and only then by the syntax.

Keywords: meter, syntactic ordering, Skaldic poetry, stanzas.

CHAPTER 27 “*The Text-building Function of Names and Nicknames in «Sverris saga» and «Böglunga sögur»*” (1994)

This paper explores the hypothesis that proper names serve as anchors identifying the individuals in a possible or real world. This hypothesis is tested on Old Icelandic narratives. A prominent feature of Old Icelandic Sagas is that the narrative matter is not quite new. Moreover, a Saga is reliable, if and only if it refers to events relevant for the audience and accepted as true by the whole community. The present paper provides evidence that proper names must be regarded as references to the background knowledge of the Saga audience. I follow two contemporary sagas — “*Sverris saga*” and “*Böglunga sögur*” — written down shortly after their events took place and demonstrate that proper names in the Saga mind were arranged as a database, each type of personal designations having their characteristic text-building profile. The form of the personal designation and its frequency in the Saga are linked with the rank of the character. The abnormally high number of the people mentioned in the Sagas is explained by the fact that the Sagas aimed at describing the real world, where the names of the individuals and the events involving these individuals cross-identified each other. Every character had his identifiers linking him to real life, while a narrative without proper names

was not considered reliable. The total amount of proper names mentioned in any Saga is always higher than the number of active figures (characters). I classify Saga characters into five ranks — a) rulers (kings, earls, archbishops), b) chieftains and field commanders, c) episodic figures, d) names in the enumeration lists, e) background names, — and discuss their information-structural and communicative values (given vs new information, theme vs rheme).

Keywords: proper names, personal designations, defined descriptions, cross-identification of individuals, databases, contemporary Sagas, narrative strategies.

CHAPTER 28

“The Burden of Others’ Words” (1994)

This paper discusses the status of text quotations and proverbial wordings from the perspective of logical analysis and text linguistics. I claim that direct quotations are especially characteristic of cultures where sacred texts (to be reproduced word for word) have a wide circulation. Quotations have their authors responsible for both the form and content. Contrariwise, gnomic texts (proverbs, general statements, common wisdom) do not involve a reconstruction of authorship since they have generic reference and are linked to generic situations. There are at least two different scenarios for inserting such gnomic mini-texts into a larger text. They can be linked to implications $p \rightarrow q$ in two different ways: 1) Generalization: ‘The rule $p \rightarrow q$ can be extended to the individual spatiotemporal situation A_i ’; 2) Characterization: ‘The rule $p \rightarrow q$ is characteristic of spatiotemporal situations like A_i ’. I claim that the Generalization scenario is typical for modern European written languages, which include lexicons of general statements, while oral and young written cultures of the Medieval European type lacked such lexicons and used the Characterization scenario instead.

Keywords: speech acts, citation, generic situations, analogy, speech errors.

APPENDIX: ONE’S OWN AND CITED TERMINOLOGY <A PREFACE TO THE THEMATIC INDEX>

Linguistic terminology includes direct borrowings from mathematics (*function, argument, tree*) and natural sciences (*valency*), but its major part comes from the common language. Some terms originated as metaphors or illustrations, cf. *actant, strong verbs, weak adjectives, movement, scrambling, middle field*, etc., but quickly became part of the mainstream tradition. Russian linguistic terminology incorporates many calques and borrowings of foreign terms, whereby some of them get more than one Russian equivalent, e.g. the term <syntactic> *movement* is translated both *передвижение* or *перемещение*. Linguistic terminology is redundant since, in many contexts terms like *actant* vs *argument*, *phrase* vs *constituent* are synonymic. At the same time, there are homonymic terms conveying different semantics. This problem is discussed in **chapter 7** (regarding two divergent uses of the term *dialect*) and **chapter 20** (regarding four different uses of the term *local*). The index made for this book has eighteen homonymic entries (pairs, triplets and quadruplets of terms) distinguished by subscripts: *адресат*₁₋₂ (‘addressee’), *время*₁₋₂ (‘time’ vs ‘tense’), *диалект*₁₋₂ (‘dialect’), *инверсия*₁₋₂ (‘inversion’), *интенсивность*₁₋₂ (‘intensity’), *лицо*₁₋₂ (‘person’), *локальность*₁₋₄ (‘locality’), *модель*₁₋₃ (‘model’), *неоднозначность*₁₋₂ (‘ambiguity’), *нулевая категория*₁₋₂ (‘zero category’), *одушевленность*₁₋₂ (‘animacy’), *омонимы*₁₋₂ (‘homonyms’), *подъем*₁₋₃ (‘raising’ or ‘height’ or ‘rise’), *посессивность*₁₋₂ (‘possessivity’), *словарь*₁₋₂ (‘lexicon’), *составляющие*₁₋₂ (‘constituents’ or ‘components’), *стимул*₁₋₂ (‘stimulus’), *эксперимент*₁₋₂ (‘experiment’). All these terms are traditional, neither of them is coined by the book’s author. Introducing homonymic entries in the index is a last resort operation. It has been avoided, where possible: e.g. the terms *семантический контроль* ‘semantic control’ vs *синтаксический контроль* ‘syntactic control’ were introduced as autonomous entries instead of making a cover entry **control*₁₋₂. Eighteen pairs/triplets/quadruplets of homonymic terms differentiated by subscripts in the index for a book, which has a single author but consists of twenty-eight papers written during thirty years (1989–2019) for various editions in three languages (Russian, English,

German) reflect the status quo in current linguistic terminology: different research tasks and different schools make use of non-identical sets of terms.

In ca. 60 cases, the entries have been supplied with an arrow, cf. '<ENTRY> '→ X' or '<ENTRY> → 'X, Y' showing the source of the term, e.g. *однопадежный ряд* 'one-case set' (→ Andrej A. Zaliznjak), *синтаксический признак* 'syntactic feature' (→ Jurij D. Apresjan'), *эпические формулы* 'epic formulae' (→ M. Parry, A. B. Lord). Such tags do not necessarily mean that the compiler of the index distances himself from these terms. However, they show that the correct identification of the corresponding terms is facilitated by the knowledge of the tradition behind them. The presence of the referred terms which do not belong to the generalized set of the shared linguistic apparatus is a salient feature of modern linguistics. The minimization of the notational categories is desirable both in framework-internal linguistics and in task-oriented research, where the author selects a narrow set of the basic notations but can also hint at the ways of its extension.

In four cases the index entries are marked with capital letters referring to generalized predicative meanings *БЫТЬ* ('BE'), *ИМЕТЬ* ('HAVE'), *КАЗАТЬСЯ* ('SEEM'), *ПРАВИТЬСЯ* ('LIKE'). The same notation is used in the book itself, but less consistently. All entries except for fifteen are given in Russian: that implies the Russian variant of the linguistic term is equivalent to its English (German, French) counterpart. The exceptions represent the cases, where exact Russian counterparts of the English terms do not exist or do not help to identify the term, cf. *I-nominals*, *superiority*.

Научное издание

Антон Владимирович Циммерлинг

ОТ ИНТЕГРАЛЬНОГО К АСПЕКТИВНОМУ

Корректоры *Н.А. Крылова, Л.А. Мосионжик*
Оригинал-макет *Л.Е. Голод*
Дизайн обложки *С.И. Касьян*

Подписано в печать 12.10.2021.
Формат 70×100^{1/16}. Бумага офсетная. Печать офсетная
Усл.-печ. л. 51. Тираж 500 экз. Заказ № 2243

Издательство «Нестор-История»
197110 СПб., Петрозаводская ул., д. 7
Тел. (812)235-15-86
e-mail: nestor_history@list.ru
www.nestorbook.ru

Отпечатано в типографии издательства «Нестор-История»
Тел. (812)235-15-86